Target-Disease-Drug Association Network-Guided Classification and Drug Repurposing of Neurological Disorders

By

Muhammad Yasir Ateeque

Master of Science in Bioinformatics

Fall 21-MSBI-NUST00000362438

Supervisor

Prof. Dr. Ishrat Jabeen

Department of Sciences

A thesis submitted in partial fulfilment of the Master of Science degree requirements.

In

School of Interdisciplinary Engineering & Sciences (SINES)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

September, 2023

Annex A to NUST Letter No. 0972/102/Exams/Thesis-Cert dated 23 Dec 16.

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by Mr. **Muhammad Yasir Ateeque** Registration No. **00000362438** of **SINES** has been vetted by undersigned, found complete in all aspects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature with stamp: _____

Name of Supervisor: **Dr. Ishrat Jabeen**

Date: 13-09-2023

Signature of HoD with stamp: _____

Dr. Fouzia Malik
HoD Sciences
Associate Professor
SINES - NUST, Sector H-12
Islamabad

Date: 19-9-2023

Countersign by

Dr. Hammad M. Cheema
Principal & Dean
SINES - NUST, Sector H-12
Islamabad

Signature (Dean/Principal): _____

Date: 19 SEP 2023

Dedication

Foremost to Almighty Allah for giving me the willpower and strength to complete my dissertation and to my family especially my loving Mother and Father for their endless love, support, and encouragement throughout my pursuit of education. I hope this achievement will fulfil the dream they envisioned for me.

Certificate of Originality

I hereby declare that this submission is my work and that, to the best of my knowledge, it contains no materials previously published or written by another person nor material that has been accepted for the award of any degree or diploma at the Department of Sciences at or any other educational institute, except where due acknowledgement has been made in the thesis. Furthermore, any contribution to the research made by others, including those with whom I have worked or elsewhere, is explicitly acknowledged in the thesis. I also declare that, except for assistance from others in the project's deliverable or in style, presentation, and linguistics, the content of this research work is the product of my own work.

Author Name: **Muhammad Yasir Ateeque**

Acknowledgements

First, I want to thank and praise Almighty Allah, who has enabled me to understand, think and complete this case report work in His great mercy and benevolence. My heartiest praise goes to the holiest man in the whole galaxy, Almighty's beloved, Hazrat Muhammad (Peace Be upon Him), humanity's reformer. A special note of appreciation goes out to my dedicated teachers who have been instrumental in shaping my academic path.

I want to express my sincere gratitude and thanks to my supervisor Dr. Ishrat Jabeen for showing me support and guidance for my Post Grad study. Her mentorship has not only enriched my academic pursuits but also nurtured my growth as a researcher. I extend my sincere gratitude to the esteemed members of my guidance committee: Dr. Zamir Hussain and Dr. Uzma Habib from SINES, NUST. Their thoughtful insights and encouraging words have played a crucial role in refining my research and shaping my academic perspective.

It is with deep gratitude that I turn to my family, who have been my unwavering pillars of support for me. My mother stands as the epitome of selflessness and determination. To my father, whose love and guidance inspire me every moment. All credit goes to my beloved parents. Their sacrifices and boundless encouragement have paved the way for all my accomplishments throughout my career.

I want to extend a special express gratitude to the Computational Drug Design lab members for their support, discussions, compassion, and respect. Last but not least, I highly acknowledge Rizwan Ahmad for being with me throughout all the thick and thins of my research. The time I spent with them will always be cherished.

Muhammad Yasir Ateeque

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

ADMET	Chemical Absorption, Distribution, Metabolism, Excretion, & Toxicity
AGE	Advanced Glycation End-products
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
ANN	Artificial neural Network
APP	Amyloid Precursor Protein
ASD	Autism Spectrum Disorders
AUC	Area under the ROC Curve
Aβ protein	amyloid-beta protein
BNDF	Brain-Derived Neurotrophic Factor
C/T/D	Conversion/Transition/Distribution
C9orf72	chromosome 9 open reading frame 72
Ca^{2+}	Calcium
CADD	Computer-Aided Drug Design
ChEMBL	chemical, bioactivity and genomic database
CHMP2B	Charged Multivesicular Body Protein 2B
Cl^-	Chloride
CNS	Central Nervous System
CUMS	Chronic Ultra Mild Stress
CYP46A1	Cytochrome P450 46A1
DL	Deep Learning
DT	Decision Tree
DTI	Drug-Target Interactions
EAAT2	Excitatory Amino Acid Transporter 2
FDA	Food and Drug Administration
FGF	Fibroblast Growth Factor
FMR1	Fragile X Messenger Ribonucleoprotein 1

FN	False Negatives
FP	False Positives
FTD	Frontotemporal Dementia
FTLD	Frontotemporal Lobar Degeneration
FUS	Fused in Sarcoma
GBA	Glucocerebrosidase
GBD	Global Burden of Diseases
GBM	Gradient Boosting Machine
GridSearchCV	GridSearch Cross-Validation
GRN	Progranulin
IDE	Integrated Development Environment
K^+	Potassium
KEGG	Kyoto Encyclopedia of Genes and Genomes
KNN	K-Nearest Neighbors
LBDD	Ligand-Based Drug Design
LRRK2	Leucine-rich repeat kinase 2
MAPT	microtubule-associated protein Tau
ML	Machine Learning
MLP-NN	Multi-Layer Perceptron Neural Network
MSE	Mean Squared Error
Na^+	Sodium
NDDs	Neurodegenerative Disorders
NIH	National Institutes of Health
NLDs	Neurological Disorders
NumPy	Numerical Python
OCD	Obsessive-Compulsive Disorder
OS	Operating System
Pandas	Python Data Analysis Library
PINK1	PTEN-induced kinase 1
PP2A	Protein phosphatase 2A
PRKN	Parkin

PSA	Progressive Supranuclear Palsy
PSEN	Presenilin
PWS	Prader-Willi Syndrome
PyPI	Python Package Index
QSAR	Quantitative structure-activity relationship
Relu	Rectified Linear Unit
RF	Random Forest
ROS	Reactive Oxygen Species
S100B	S100 calcium-binding protein B
SBDD	Structure-Based Drug Design
Sklearn	scikit-learn
SMOTE	Synthetic Minority Oversampling TEchnique
SNCA	Synuclein Alpha
SSRIs	Selective serotonin reuptake inhibitors
STITCH	Search Tool for Interactions of Chemical
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
SVM	Support Vector Machine
TARDBP	TAR DNA Binding Protein
TBI	Traumatic Brain Injury
TDP-43	Transactive Response DNA-binding Protein with Molecular Weight 43 kDa
TN	True Negatives
TNF-α	Tumor Necrosis Factor alpha
TP	True Positives
TTD	Therapeutic Target Database
UniProt	Universal Protein Resource
VCP	Valosin-containing Protein
WHO	World Health Organization

Abstract

Accurate therapeutic intervention against many Neurological disorders is still not known. Only symptomatic treatments are being used for the cure of such devasting disorders. Therefore, it is crucial to probe the target associations with the drug followed by the subsequent disease associations which could aid in more accurate and effective treatment against neurological disorders. Additionally, there is a major overlap of targets in neurological and neurodegenerative disorders. In this study, a database for known protein-targets and FDA-approved drugs for 10 neurodegenerative disorders and 9 neurological disorders is developed from publicly available resources. The database contains 236 unique protein-targets with Protein-Protein Interactions (PPIs) ranging from 3 to 71, and 964 FDA-approved drugs against selected target-proteins for the 19 neuronal disorders. Network pharmacology approach was used to investigate the targets association and overlap in neurological and neurodegenerative disorders. Three networks i.e., Target-Disease, Disease-Drug and Target-Disease-Drug Networks, were built between protein-targets, FDA-approved drugs, and neuronal disorders, with datasets categorized into neurological and neurodegenerative disorders. Furthermore, five machine learning models were trained on the networks, with Decision Tree, Random Forest, and Gradient Boosting Classifiers emerging as optimal models for predicting disease association of protein-targets and drugs. The results provide a comprehensive view of drugs and protein-targets' association with specific neurological and neurodegenerative disorders, as well as target overlap among multiple neuronal disorders. Finally, a multi-variate Artificial Neural Network (ANN) to predict drug-target interactions linked to specific diseases has been developed. The model was trained using a multi-variate output configuration, enabling predictions for both target protein descriptors with 53% accuracy and disease class with 82% accuracy, for a given drug. This study contributes to database development and Network classification for FDA-approved drugs and protein targets associated with neurological and neurodegenerative disorders including the multi-variate model development, offering potential avenues for developing new therapeutics and personalized treatment strategies.

Chapter 1

Introduction

1 Introduction

1.1 Background:

Neurological disorders, a diverse group of conditions that includes both neurodevelopmental and neuropsychiatric disorders, present significant challenges to global healthcare systems [1]. These disorders inflict debilitating effects on patients, impairing their cognitive, motor, and emotional functions, leading to a diminished quality of life. Neurodegenerative diseases [2], such as Alzheimer's, Parkinson's [3, 4], Huntington's diseases [5], etc. are characterized by progressive degeneration of nerve cells and neural networks, resulting in the deterioration of cognitive and motor functions over time [6]. Neurodegenerative disorders are caused due to the inability of nerve cells to perform their function well [7]. The loss of nerve cells is caused due to Increased ROS, Imbalance of ions in nerve impulse transmission, Chronic neuroinflammation and intracellular and extracellular accumulation of misfolded proteins, causing abnormalities in normal cellular functions, and ultimately causing apoptosis of nerve cells as shown in Figure 1.1. As functional neuronal death is almost certainly the key factor that mediates functional impairment, preventing neuronal death and dysfunction will have a huge clinical benefit. Neurodegenerative diseases are incurable and debilitating, resulting in the progressive degeneration and/or death of neurons [8]. On the other hand, psychiatric disorders encompass a wide range of mental health conditions [9], including depression [10], anxiety [11], bipolar disorder [12], and schizophrenia [13], which significantly impact a person's thoughts, emotions, and behavior. In neuropsychiatric disorders,

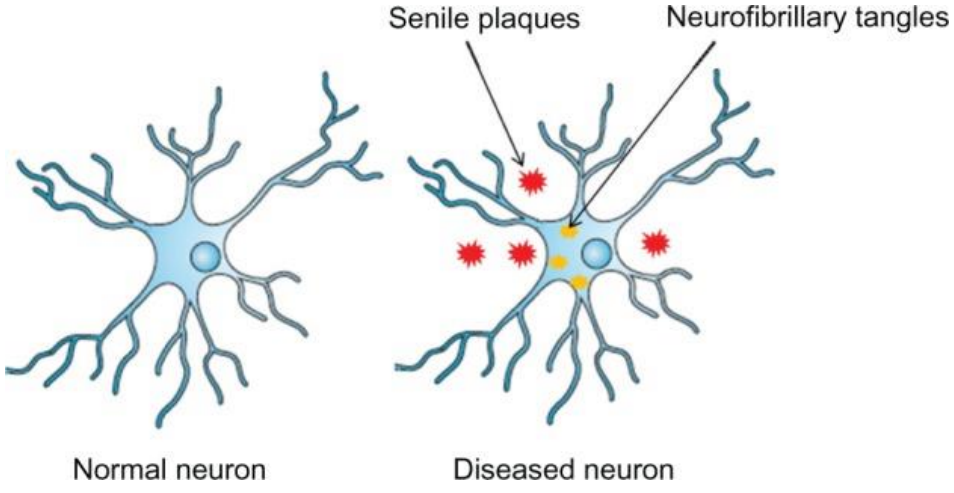

Figure 1.1 *Normal Vs Diseased Neuron*

there is an imbalance of ions and neurotransmitters between nerve cells, which occurs either due to abnormal secretions of excitatory, inhibitory, or modulatory neurotransmitters, affecting neurotransmission and ultimately causing many psychological problems. Neurons are the building blocks of the nervous system [14] and are different from other cells of the body as they do not reproduce or replace themselves, therefore, the body cannot replace them with other neurons when they are damaged.

1.2 Prevalence and Mortality Rate:

According to National Health Survey conducted in 2000 [15], the prevalence of hypertension in Pakistan is 33%, 91.1% for stroke, 22.9% for Migraine, 34% for depression and anxiety 1.5% for Epilepsy, 39% for depression, 51% for Parkinson's disease, and 60% for dementia due to Alzheimer's disease (AD).

According to Global Burden of Diseases, Injuries, and Risk Factors (GBD), Neurological disorders are an important cause of disability and the second-leading cause group of deaths worldwide [16]. Globally, the burden of neurological disorders has increased substantially over the past 25 years because of expanding population numbers and aging. The most prevalent neurological disorders are tension-type headache, migraine, Alzheimer's disease and other dementias. Between 1990 and 2015, the number of deaths from neurological disorders increased by 36.7%, and the number of DALYs by 7.4%. The number of patients who will need care by clinicians with expertise in neurological conditions will continue to grow in the coming decades.

World Health Organization (WHO) data suggests that neurological and psychiatric disorders are an important and growing cause of morbidity, mortality, and disability worldwide [17]. Neurological disorders are the third most common cause of disability and premature death in the Europe and their prevalence and burden will likely increase with the progressive ageing of the European population.

A recent study of GBD in 2019 showed that the prevalence of mental disorders (depression, anxiety, post-traumatic stress disorder, bipolar disorder, autism spectrum disorders and schizophrenia) is 22.1% and remained among the top ten leading causes of burden worldwide,

with no evidence of global reduction in the burden since 1990 [18]. The burden of mental disorders is high in conflict-affected populations. Given the large numbers of people in need and the humanitarian imperative to reduce suffering, there is an urgent need to implement scalable mental health interventions to address this burden.

Neurological disorders stand as the second leading cause of death in the elderly population worldwide. The demographic factors, particularly aging, might be related to an increase in the mortality of neurological disorders. As the prevalence of these conditions continues to rise with an aging population and changing lifestyles, the burden on healthcare systems and the economic impact on societies are mounting [19].

1.3 Common Neurodegeneration Mechanisms:

Chronic neuroinflammation, increased ROS, imbalance of ions in nerve impulse transmission, accumulation of misfolded proteins in nerve cells, has emerged as some common neurodegenerative mechanisms across many neurological disorders, contributing to disease progression and pathophysiology [20, 21]. The details of all the mechanisms are given below:

1.3.1 Chronic neuroinflammation:

Chronic neuroinflammation, characterized by persistent and prolonged activation of immune responses within the central nervous system, has garnered significant attention as a common underlying feature across a wide spectrum of neurological disorders. This chronic inflammatory state has been implicated in the pathogenesis of diverse conditions, ranging from neurodegenerative diseases like Alzheimer's and Parkinson's to psychiatric disorders such as depression and schizophrenia. The recognition of chronic neuroinflammation as a shared phenomenon in these seemingly disparate disorders has shed light on the intricate relationship between molecular targets, drugs, and disease manifestations, presenting a promising avenue for the development of more accurate and effective treatments [22]. Understanding the complex interplay between molecular targets and disease outcomes is essential in the quest to decipher the

underlying mechanisms driving neurological disorders. By elucidating the specific interactions between biological targets and therapeutic agents, we can unlock novel strategies for intervention and potentially halt disease progression.

1.3.2 Dysregulation of Reactive Oxygen Species (ROS):

Reactive Oxygen Species (ROS) are highly reactive molecules that contain oxygen and are produced as natural byproducts of cellular metabolism. In normal physiological conditions, cells maintain a delicate balance between ROS production and their elimination through antioxidant defense systems. However, several factors, including environmental stress, exposure to toxins, chronic inflammation, and dysfunction of mitochondria, can disrupt this balance and lead to an excessive generation of ROS, resulting in a state of oxidative stress. The dysregulation of ROS production stands out as a critical factor contributing to neuronal damage and inflammation. Within the context of neurological disorders, the impact of increased ROS levels on disease progression is profound [23]. Oxidative stress becomes a critical player in the pathophysiology of these disorders. The deleterious effects of oxidative stress are primarily attributed to the damaging effects on cellular components, including lipids, proteins, and DNA, within nerve cells. Such damage can lead to cellular dysfunction and, in severe cases, cell death. The detrimental consequences of oxidative stress extend to neuronal function and communication. Oxidative damage can impair neuronal function, disturb synaptic communication, and facilitate the process of neuroinflammation, which further exacerbates the disease's severity. A particularly significant consequence of ROS-induced damage is the accumulation of misfolded or abnormal proteins. This accumulation is particularly relevant to neurodegenerative diseases like Alzheimer's [24], Parkinson's [25], and Amyotrophic Lateral Sclerosis (ALS) [26], where protein misfolding plays a central role in disease pathogenesis. To address the adverse effects of increased ROS and oxidative stress in neurological disorders, researchers are actively exploring antioxidant therapies and interventions to restore redox homeostasis and protect neurons from oxidative damage. By mitigating oxidative stress, these treatments have the potential to slow down disease progression and offer neuroprotection. The intricate relationship between ROS, oxidative stress, and

neurological disorders underscores the significance of redox homeostasis in maintaining neuronal health.

1.3.3 Imbalance of Ions & Neurotransmitters:

The proper functioning of the nervous system relies on the precise transmission of nerve impulses, a process governed by the delicate balance of ions across neuronal membranes. Among these ions, sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), and chloride (Cl^-) play pivotal roles in neuronal signaling. Neurons employ specialized protein structures called ion channels and pumps to regulate the movement of these ions in and out of cells, leading to the generation of crucial electrical signals that facilitate communication between neurons. Maintaining ion homeostasis is of paramount importance for the overall health and function of neurons. Disruptions in this delicate balance can have profound consequences, leading to abnormal nerve impulse transmission and impairing neuronal communication and function [27]. An imbalance in ion concentrations in nerve impulse transmission has been implicated in the hyperexcitability of neurons and the manifestation of neurological symptoms as shown in Figure 1.2. In the context of neurological disorders, such imbalances can become particularly significant. Neurological disorders, including epilepsy and multiple sclerosis, are often characterized by disturbances in ion concentrations. In epilepsy, excessive neuronal excitability arises due to imbalances in ion concentrations, leading to the

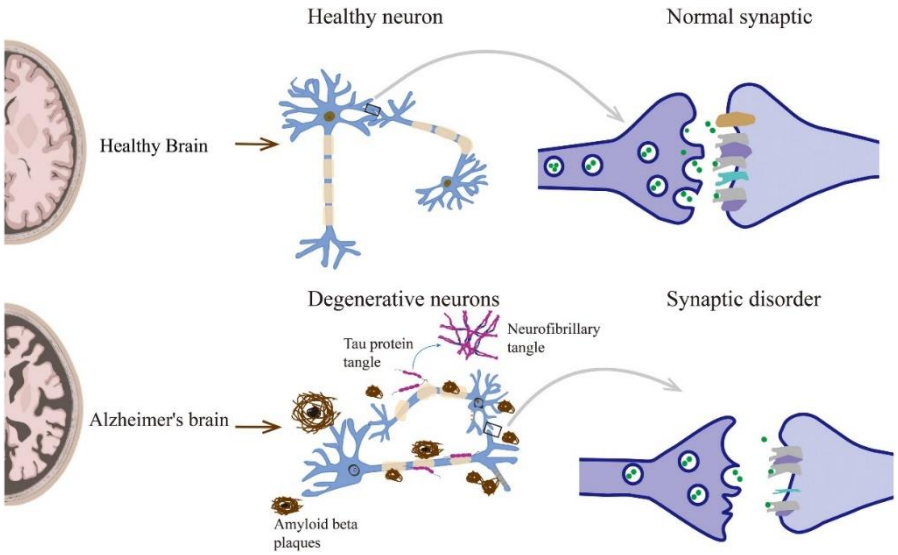

Figure 1.2 *Neurotransmitters Imbalance in Neurological disorders*

uncontrolled firing of neurons and the occurrence of seizures. On the other hand, multiple sclerosis is marked by altered ion homeostasis, which can hinder the smooth transmission of signals between neurons, resulting in motor dysfunction and cognitive impairments. These disruptions in ion homeostasis contribute to the diverse array of symptoms observed in neurological disorders, ranging from seizures to motor and cognitive impairments. As our knowledge of ion channel dysfunctions continues to expand, we can anticipate the emergence of novel and innovative treatment options for neurological disorders, providing hope and improved quality of life for those affected by these conditions [28].

1.3.4 Accumulation of misfolded proteins:

Proteins, being the workhorses of cellular functions, assume critical roles in maintaining the normal activities of living organisms. For proteins to function properly, their three-dimensional structure must be correctly folded. However, under certain circumstances, proteins may misfold due to genetic mutations, environmental influences, or cellular stress. Such misfolded proteins have a tendency to aggregate and form toxic clumps within nerve cells, thereby disrupting vital cellular processes and leading to impaired neuronal function [29]. The accumulation of misfolded proteins represents a hallmark feature observed in numerous neurodegenerative diseases, including Alzheimer's [3, 4], Parkinson's, and Huntington's diseases. In Alzheimer's disease, the formation of amyloid-beta protein [30] aggregates gives rise to plaques, while in Parkinson's disease, alpha-synuclein [31] aggregates manifest as Lewy bodies. These protein aggregates exert toxic effects on neurons, instigating dysfunction and eventual degeneration of these essential cells. A comprehensive understanding of these mechanisms opens the door to potential therapies that can protect neurons from the toxic effects of misfolded proteins, offering hope for improved treatment strategies and ultimately enhancing the quality of life for individuals impacted by neurodegenerative diseases.

In conclusion, neurological disorders represent a multifaceted and intricate web of interconnected mechanisms, comprising chronic neuroinflammation, increased levels of Reactive Oxygen Species (ROS), imbalances in ion homeostasis, and the accumulation of misfolded proteins as shown in the Figure 1.3. The intricate understanding of these underlying mechanisms

and their role in disease pathogenesis offers valuable insights that can pave the way for the development of targeted and effective therapies and interventions, ultimately enhancing the management and treatment of neurological disorders. Addressing these diverse facets of neurological disorders is of paramount importance in unlocking novel treatment approaches and mitigating the significant burden these disorders impose on affected individuals and global healthcare systems alike. By delving into the complexities of chronic neuroinflammation, oxidative stress, ion imbalances, and protein misfolding, researchers can identify potential targets for intervention, thus providing new avenues for therapeutic development. The quest for innovative treatments is a crucial endeavor, as it holds the promise of offering hope to patients and healthcare systems worldwide. Through persistent dedication and interdisciplinary collaboration among scientists, clinicians, and researchers, advancements in our understanding of neurological disorders will continue to unfold, ushering in a new era of improved management and care.

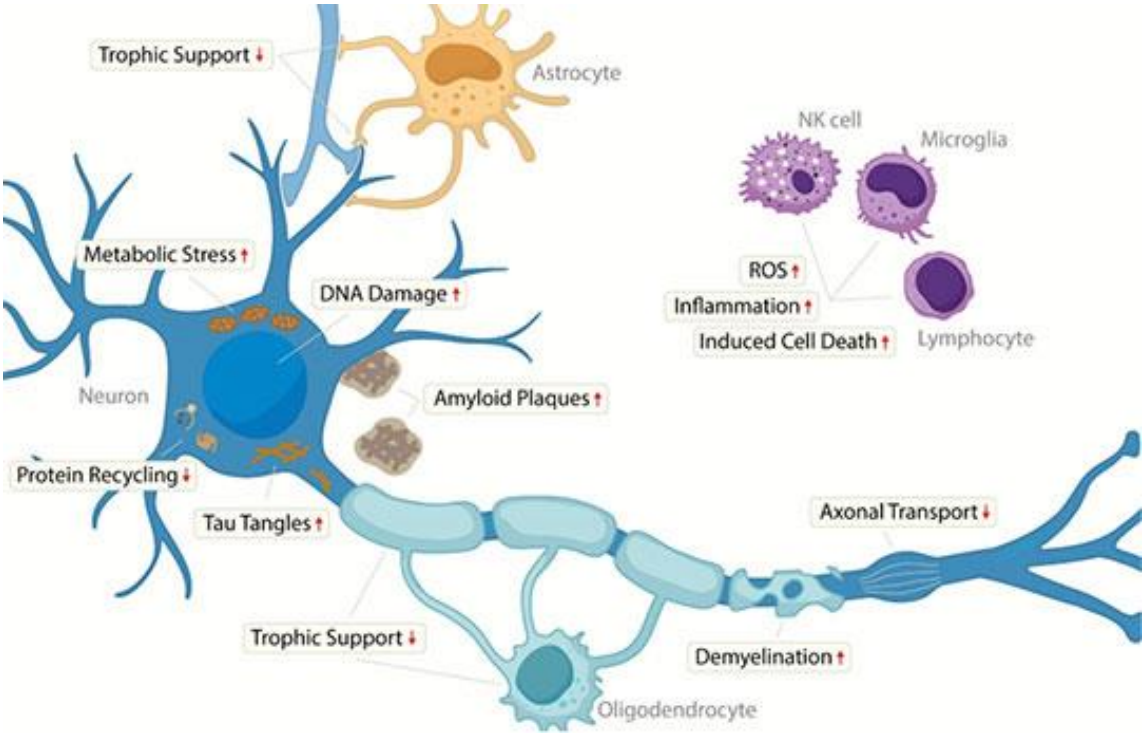

Figure 1.3 *Mechanisms of Neurodegeneration*

1.4 Disease Pathogenesis & Symptoms:

Neurological disorders encompass a broad spectrum of conditions that pose significant public health challenges and impact millions of individuals worldwide. Despite dedicated research efforts, developing precise and effective therapeutic interventions for many of these complex disorders remains a formidable task. Among the diverse array of neurological disorders, several stand out, each presenting its unique set of challenges. The selected disorders specifying their disorder category in the context of common neurodegenerative mechanisms, including increased Reactive Oxygen Species (ROS), imbalance of ions in nerve impulse transmission, accumulation of misfolded proteins in nerve cells, and chronic neuroinflammation:

1.4.1 Neurodegenerative Disorders:

- **Alzheimer's Disease** is characterized by the accumulation of misfolded proteins such amyloid-beta (*Aβ*) protein plaques, presenilin (*PSEN*), amyloid precursor protein (*APP*) and tau protein tangles in the brain [32]. These protein aggregates trigger chronic neuroinflammation, oxidative stress, and disruption of ion homeostasis, contributing to the degeneration of nerve cells and cognitive decline [33]. Symptoms include memory loss, confusion, and difficulty performing everyday tasks [32].
- **Huntington's Disease** is caused by a genetic mutation leading to the aggregation of mutant huntingtin protein. The accumulation of misfolded proteins, along with increased ROS and chronic neuroinflammation, contribute to the progressive degeneration of neurons. Symptoms include motor problems, cognitive decline, and psychiatric symptoms [5].
- **Amyotrophic Lateral Sclerosis (ALS)** is a progressive motor neuron disease involving the degeneration of motor neurons in the brain and spinal cord [34]. Imbalance of ions and increased ROS production contribute to the neurodegenerative process. ALS leads to muscle weakness and paralysis.
- **Frontotemporal Lobar Degeneration (FTLD)** encompasses a group of disorders characterized by the degeneration of the frontal and temporal lobes of the brain. Several genes have been associated with FTLD, and mutations in these genes play a significant role

in the pathogenesis of the disease, such as *C9ORF72*, *MAPT* (Microtubule-Associated Protein Tau) [35], Progranulin (*GRN*), *TDP-43* (Transactive Response DNA-binding Protein with Molecular Weight 43 kDa), *FUS* (Fused in Sarcoma), *CHMP2B* (Charged Multivesicular Body Protein 2B), and *VCP* (Valosin-containing Protein) [36]. Mutations in these genes disrupt various cellular processes, including protein homeostasis, RNA processing, inflammation, and cellular transport, leading to the characteristic neuronal degeneration. FTLD results in changes in personality, behavior, and language difficulties.

- **Multiple Sclerosis** is an autoimmune disorder in which chronic inflammation damages the myelin sheath around nerves. This chronic neuroinflammation and demyelination contribute to nerve dysfunction and causing symptoms like fatigue, vision problems, and difficulty walking [37].
- **Parkinson's Disease** is characterized by the degeneration of dopamine-producing neurons in the brain. Mutations in the several genes have been implicated in the development and progression of Parkinson's disease such as *FMR1, α-synuclein, Parkin, PINK1, DJ-1, PARK8* and *GBA* [38]. Increased ROS production, misfolded protein accumulation (alpha-synuclein), and chronic neuroinflammation contribute to the neurodegenerative process. Symptoms include tremors, rigidity, and difficulty with balance and coordination.
- **Dementia** is a syndrome associated with a decline in memory, cognitive function, and behavior. Several types of dementia, including Alzheimer's disease, involve common neurodegenerative mechanisms, such as misfolded protein accumulation, chronic neuroinflammation, and oxidative stress [39]. Dementia results in memory loss, impaired reasoning, and personality changes.
- **Prion Disease**, like Creutzfeldt-Jakob disease, involves the misfolding of prion proteins, leading to the formation of infectious protein aggregates [40]. This accumulation of misfolded proteins triggers neuroinflammation and neurodegeneration. Symptoms include Rapidly Progressive Dementia, Movement Abnormalities, Behavioral Changes, Visual Disturbances, Muscle Weakness, Difficulty Swallowing and Speaking.
- **Progressive Supranuclear Palsy** is a rare neurodegenerative disorder characterized by movement and balance problems. Misfolded tau proteins and chronic neuroinflammation contribute to the degeneration of brain cells [41]. Symptoms include Balance and Gait

Problems, Stiffness and Rigidity, Slow Movements, Cognitive Changes, Speech Difficulties, and Swallowing difficulties [42].

1.4.2 Neurological Disorders:

- **Epilepsy** is characterized by recurrent seizures resulting from abnormal electrical activity in the brain [43]. Imbalance of ions, particularly sodium and potassium, disrupts nerve impulse transmission and leads to epileptic seizures which can vary in intensity and manifestation [44].
- **Obsessive-Compulsive Disorder (OCD)** is characterized by repetitive, intrusive thoughts and behaviors. Abnormalities in brain neurotransmitters and neuroinflammatory processes have been implicated in the disorder [45]. OCD involves intrusive thoughts and repetitive behaviors, causing distress and impairment.
- **Migraine** is a complex neurological disorder with genetic and environmental factors contributing to its pathogenesis. The exact mechanisms behind migraine are still under investigation, but studies suggest that oxidative stress and neuroinflammation may play a role in migraine attacks [46]. Migraine causes severe headaches, often accompanied by nausea, sensitivity to light, and sound.
- **Psychotic Disorder** such as schizophrenia [47], involve altered brain connectivity and neurotransmitter imbalances. Dysregulated ROS production, oxidative stress, and chronic neuroinflammation may contribute to the pathophysiology of psychotic disorders. Such disorders lead to delusions, hallucinations, and disorganized thinking [48].
- **Autism Spectrum Disorder** is a complex neurodevelopmental disorder with diverse genetic and environmental influences. Imbalances in neurotransmitters and neuroinflammatory processes may contribute to the neurological manifestations of autism. ASD results in difficulties in social communication and repetitive behaviors [49].
- **Anxiety Disorder** involves dysregulation of brain circuits and neurotransmitters. Neuroinflammation and oxidative stress have been implicated in the pathophysiology of anxiety disorders [50]. Anxiety Disorders cause excessive worry, fear, and avoidance of certain situations.

- **Major Depressive Disorder** involves alterations in brain regions responsible for mood regulation and neurotransmitter imbalances [11]. Chronic inflammation and oxidative stress may contribute to the development and persistence of depressive symptoms [10]. Symptoms include persistent feelings of sadness and loss of interest in activities.
- **Prader-Willi Syndrome** results from genetic abnormalities and affects brain development and function. While the exact mechanisms linking neurodegeneration to this disorder are not fully understood, chronic neuroinflammation and oxidative stress could play a role [51]. Prader-Willi Syndrome affects brain development, leading to intellectual disability and behavioral problems.
- **Down Syndrome** is caused by the presence of an extra chromosome 21 [52]. Individuals with Down syndrome have an increased risk of developing Alzheimer's disease later in life, likely due to shared neurodegenerative mechanisms involving misfolded proteins and oxidative stress. Symptoms involve cognitive impairment, distinctive facial features, and an increased risk of Alzheimer's disease.
- **Williams-Beuren Syndrome** is a genetic disorder affecting multiple systems, including the nervous system. The mechanisms underlying neurodegeneration in this disorder are not fully understood but could involve oxidative stress and neuroinflammation [53]. It affects multiple systems, including the nervous system, leading to developmental delays and cardiovascular problems.

1.5 Ongoing therapies and their Limitations:

Below are some typical and well-known current treatment approaches for neurological, neurodegenerative, neurodevelopmental, and neuropsychiatric along with their general success rates and limitations:

1.5.1 Neurological Disorders:

- **Epilepsy:** Antiepileptic medications are the mainstay of treatment and can control seizures in about 60-70% of patients [54]. However, some individuals may not respond to medications and may require other treatments, such as surgery or a ketogenic diet.
- **Multiple Sclerosis:** Disease-modifying therapies can slow disease progression and reduce relapses in many patients, but they may not be effective for all individuals, and side effects can occur [55].
- **Migraine:** Triptans and other migraine-specific medications can provide relief for many patients [56], but they may not work for everyone and can have side effects.
- **Parkinson's Disease:** Levodopa is the primary medication for managing motor symptoms, and it is effective in improving mobility [57]. However, long-term use can lead to motor complications.
- **Depression:** Antidepressant medications and psychotherapy can be effective for many patients [58], but the success rate varies, and some individuals may require multiple trials of different medications before finding one that works for them.
- **Anxiety Disorders:** Medications such as SSRIs and benzodiazepines can provide relief for some patients [58], but they may not work for everyone and can have side effects.
- **Bipolar Disorder:** Mood stabilizers and antipsychotic medications are used to manage bipolar disorder [59] but finding the right combination of medications can be challenging, and some individuals may experience treatment-resistant symptoms.
- **Psychotic Disorder or Schizophrenia:** Antipsychotic medications can help manage symptoms, but they may not be effective for all patients [48], and some individuals may experience side effects.
- **Obsessive-Compulsive Disorder (OCD):** Treatment may involve cognitive-behavioral therapy and/or medications such as SSRIs [60]. Success rates vary among patients.
- **Autism Spectrum Disorder (ASD):** Treatment may involve behavioral therapies, speech therapy, and medications to manage associated symptoms [61].
- **Prader-Willi Syndrome:** There is no cure for Prader-Willi Syndrome, and treatment focuses on managing symptoms and providing supportive care [62].

- **Down Syndrome:** Treatment focuses on addressing associated medical conditions and providing support for developmental and cognitive challenges [63].
- **Williams-Beuren Syndrome:** Treatment focuses on addressing associated medical and developmental challenges and providing supportive care [64].

1.5.2 Neurodegenerative Disorders:

- **Alzheimer's Disease:** Current medications, such as cholinesterase inhibitors and memantine, can help manage symptoms, but they do not stop disease progression [65]. The success of these medications varies among patients, and they may not work for all individuals.
- **Huntington's Disease:** There is no cure for Huntington's disease, and current treatments focus on managing symptoms [66] and providing supportive care to improve the patient's quality of life.
- **Amyotrophic Lateral Sclerosis (ALS):** Riluzole is the only FDA-approved drug for ALS and may extend survival by several months [67]. However, its effects are modest, and there is a need for more effective treatments.
- **Prion Disease:** There is no cure for prion diseases, and treatment focuses on managing symptoms and providing supportive care [68].
- **Progressive Supranuclear Palsy (PSP):** Treatment is mainly supportive, focusing on managing symptoms and providing physical therapy [69].
- **Frontotemporal Lobar Degeneration (FTLD):** Treatment is mainly supportive, focusing on managing symptoms and providing cognitive and behavioral support [70].
- **Dementia:** Treatment aims to manage symptoms and improve quality of life, but there is no cure for most forms of dementia [71].

1.6 Network-based Drug Repurposing:

Network-based drug repurposing is a powerful approach in drug discovery and development that leverages the principles of network pharmacology and computational methods to identify new therapeutic uses for existing drugs [72]. This strategy involves constructing and analyzing complex biological networks that integrate information about molecular targets, diseases, and drugs. The process of network-based drug repurposing begins with the construction of Target-Disease-Drug Association Networks. These networks elucidate the relationships between specific molecular targets, neurological diseases, and available drugs or compounds as shown in Figure 1.4. Data for these networks is curated from reputable databases and sources, ensuring the reliability and accuracy of the information [73].

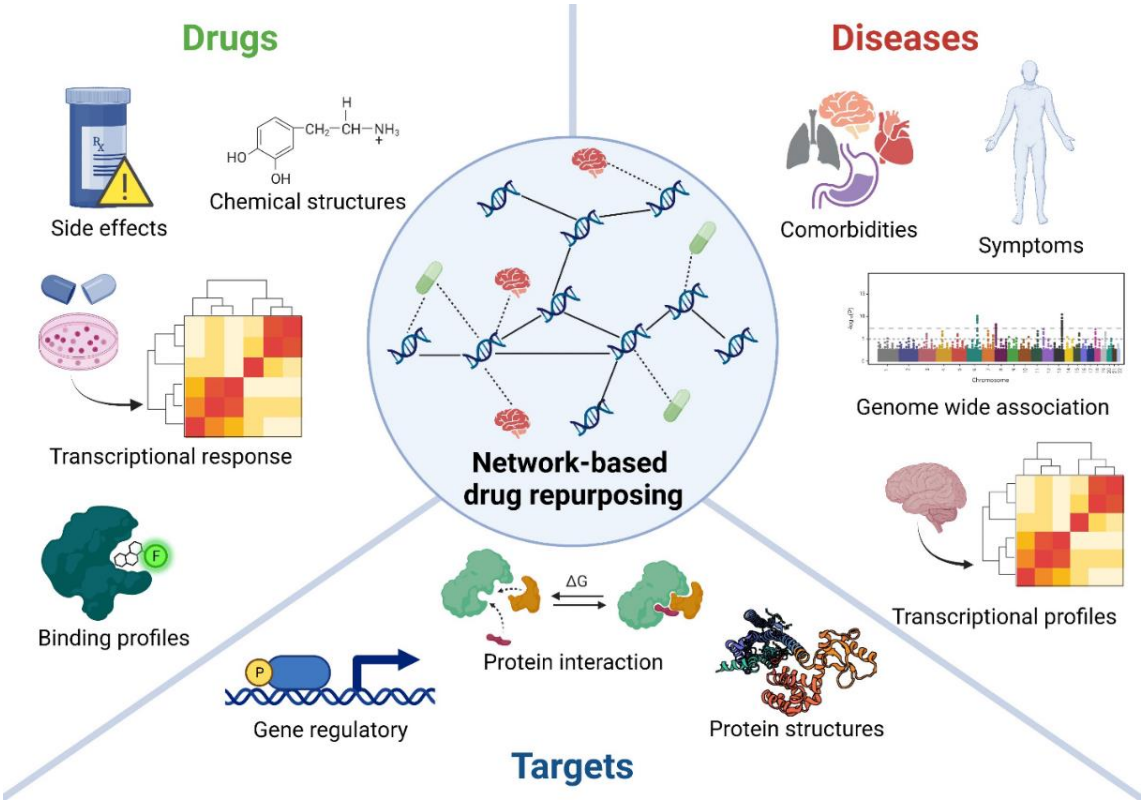

Figure 1.4 *Network-based Drug Repurposing*

In neurodegenerative and neurological disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, and autism, the underlying molecular mechanisms are often multifactorial

and involve intricate networks of molecular pathways [21]. Traditional drug discovery approaches may be time-consuming and costly, with a high failure rate. Network pharmacology offers a holistic perspective on the complex interactions between drugs, targets, and diseases at a molecular level, making it a valuable tool in tackling these challenging disorders. Once the Target-Disease-Drug Association Networks are in place, different Machine learning and deep learning methods can be applied to guide classification and drug repurposing efforts. These algorithms can effectively capture complex relationships within the dataset and make robust predictions, ultimately leading to the identification of potential therapeutic interventions for neurological disorders.

The integration of network pharmacology [74] and machine learning enables the identification of existing drugs that show potential efficacy against specific molecular targets associated with neurological disorders [75]. Drug repurposing offers an advantageous strategy as it capitalizes on the safety profiles and known mechanisms of approved drugs, leading to faster and more cost-effective drug development processes. By combining network-based drug repurposing with traditional drug development approaches, researchers can accelerate the discovery of effective treatments for neurological disorders. This approach holds great promise for revolutionizing neurological disorder management and significantly improving patient outcomes, paving the way for precision medicine in neurology. Overall, network-based drug repurposing is a data-driven, systematic, and innovative approach that has the potential to transform the landscape of drug discovery and significantly impact the field of neuroscience and neurological disorder therapeutics.

1.6.1 Network pharmacology:

Network pharmacology leverages network-based analysis to identify potential drug candidates for repurposing. It integrates various data sources, including protein-protein interactions, drug-target interactions, and disease-associated pathways, to construct comprehensive interaction networks [76]. These networks help reveal the underlying mechanisms of neurological disorders and identify key drug targets. In the context of drug repurposing, network pharmacology enables the identification of existing drugs that may have activity against multiple targets involved in

different neurological disorders. This approach allows researchers to explore new therapeutic options for known drugs, potentially accelerating the drug development process. Furthermore, network pharmacology can facilitate the prediction of drug-disease associations and identify drugs with potential off-target effects, leading to the discovery of new therapeutic uses for existing drugs. By leveraging the wealth of available molecular data and employing advanced computational techniques, network pharmacology provides a data-driven and efficient approach to drug repurposing for neurological disorders. The integration of network pharmacology with machine learning algorithms further enhances predictive power, allowing for the identification of drug-target-disease associations with higher accuracy. By considering the complex interplay of molecular interactions, network pharmacology contributes to a more targeted and personalized approach to drug repurposing in neurological disorders. Network pharmacology is implemented by the construction of Target-Disease-Drug Association Networks. Overall, network pharmacology [77] holds the promise of uncovering novel therapeutic opportunities and improving the treatment landscape for neurological disorders by repurposing existing drugs and expediting the translation of potential treatments from bench to bedside.

1.6.2 Target-Disease-Drug Association Networks:

The critical role of understanding the intricate relationship between molecular targets, drugs, and subsequent disease manifestations in the pursuit of more accurate and effective treatments for neurological disorders involves following aspects:

Neurological disorders arise from complex interactions between various molecular targets in the body. These targets can include proteins, enzymes, receptors, and other biomolecules that play key roles in cellular signaling and regulation. Dysregulation or dysfunction of these targets can lead to abnormal cellular processes, ultimately culminating in the manifestation of neurological diseases. Understanding these molecular targets is essential as they serve as potential points of intervention for therapeutic treatments. By identifying and targeting specific molecules involved in the disease pathogenesis, researchers can develop drugs that aim to correct or modulate the malfunctioning processes, leading to improved disease management and symptom relief. The development of effective drugs is a critical aspect of treating neurological disorders. Drugs interact with specific

molecular targets in the body, either promoting their activity or inhibiting their function, to bring about the desired therapeutic effect. However, the process of drug development is highly intricate and challenging, as researchers need to consider factors such as drug specificity, bioavailability, and potential side effects.

The successful identification and optimization of drugs for neurological disorders rely heavily on a deep understanding of the molecular targets and disease pathways involved [78]. Rigorous research and preclinical studies are essential to evaluate the efficacy and safety of potential drug candidates before they advance to clinical trials. Neurological disorders encompass a diverse range of conditions, each with unique clinical presentations and underlying causes. However, there is growing evidence to suggest that many neurological and neurodegenerative disorders share common pathophysiological mechanisms. These shared mechanisms can include oxidative stress, inflammation, protein misfolding, mitochondrial dysfunction, and synaptic abnormalities. The recognition of overlapping targets and mechanisms across different neurological disorders presents an opportunity for the development of novel therapeutic strategies. By understanding the common pathways that contribute to the progression of various neurological conditions, researchers can explore treatments that may have broader applications across multiple disorders. The intricate relationship between molecular targets, drugs, and disease manifestations in neurological disorders forms the foundation for developing more accurate and effective treatments. By comprehending the molecular underpinnings of neurological diseases and identifying drug targets, researchers can create therapies tailored to the specific mechanisms driving each disorder. Additionally, the discovery of shared pathophysiological mechanisms among neurological disorders opens avenues for innovative therapeutic approaches that could potentially have widespread benefits for patients with diverse conditions. This multidimensional approach to understanding and treating neurological disorders has the potential to revolutionize patient care and improve the quality of life for millions of individuals worldwide.

1.6.3 Machine Learning Algorithms:

Machine learning algorithms are computational methods that enable machines to learn patterns and make predictions from data without explicit programming. These algorithms are a core

component of artificial intelligence and data science applications. There are various types of machine learning algorithms, including:

- **Supervised Learning:** In this type of algorithm, the model is trained on labeled data, where each input is associated with a corresponding output. The goal is to learn a mapping between inputs and outputs so that the model can make accurate predictions on new, unseen data [79]. Examples of supervised learning algorithms [80] include Linear Regression, Support Vector Machines, Decision Trees, Random Forests, and Neural Networks.
- **Unsupervised Learning:** In contrast to supervised learning, unsupervised learning algorithms work with unlabeled data, and the model tries to find patterns and structure within the data [81]. Clustering and dimensionality reduction are common tasks in unsupervised learning. K-Means, Hierarchical Clustering, Principal Component Analysis (PCA), and Autoencoders are examples of unsupervised learning algorithms [82].
- **Semi-Supervised Learning:** This is a hybrid approach that combines elements of supervised and unsupervised learning [83]. The model is trained on a small amount of labeled data and a larger amount of unlabeled data. It uses the labeled data to learn patterns and then generalizes this knowledge to the unlabeled data.
- **Reinforcement Learning:** This type of learning involves an agent that interacts with an environment and receives feedback in the form of rewards or penalties based on its actions [84]. The agent's goal is to learn a policy that maximizes the cumulative rewards over time. Reinforcement learning is commonly used in applications like gaming, robotics, and autonomous vehicles.
- **Transfer Learning:** Transfer learning involves using knowledge gained from solving one problem to help solve a related but different problem [85]. This approach allows models to leverage pre-trained representations and fine-tune them for specific tasks, saving time and computational resources.
- **Ensemble Methods:** Ensemble methods combine multiple base models to make more accurate predictions [86]. Examples include Bagging (e.g., Random Forests) and Boosting (e.g., Gradient Boosting Machines).
- **Deep Learning:** Deep learning is a subfield of machine learning that uses artificial neural networks with multiple layers (deep architectures) to learn complex patterns from data [87].

Deep learning has shown remarkable success in tasks like image and speech recognition, natural language processing, and autonomous systems.

These algorithms are used in a wide range of applications, including image and speech recognition, natural language processing, recommendation systems, fraud detection, drug discovery, healthcare, finance, and more. As computational power and data availability continue to grow, machine learning algorithms are becoming increasingly powerful and capable of addressing complex real-world problems.

Machine learning algorithms play a crucial role in drug repurposing by identifying potential therapeutic candidates among existing drugs. These algorithms leverage the vast amount of data available on molecular targets, diseases, and drugs to make accurate predictions about drug efficacy and potential interactions [75]. The process of drug repurposing is more cost-effective and time-efficient compared to traditional drug discovery, as it capitalizes on the known safety profiles and mechanisms of approved drugs. Machine learning algorithms can be used in various mechanisms of drug repurposing such as Data Integration and Analysis, Network Pharmacology, Predictive Modeling, Drug-Target Prediction, Side Effect Prediction, Drug Combination Prediction, Drug-Drug Interaction Prediction, and Virtual Screening. Machine learning models can accelerate the process of drug discovery and facilitate precision medicine approaches for various diseases [88]. However, it is important to note that the success of drug repurposing using machine learning relies heavily on the quality and diversity of the data used for training and validation. Additionally, experimental validation is essential to confirm the predictions made by these algorithms before advancing to clinical trials.

1.7 Our Strategy:

The primary objective of the present thesis is to tackle the critical requirement for precise and personalized therapeutic solutions for neurological disorders. We aim to achieve this goal by concentrating on the creation and examination of Target-Disease-Drug association networks. These networks are informed by network pharmacology principles and bolstered by advanced machine learning algorithms. Through this approach, we seek to unravel the intricate relationships

between molecular targets, neurological diseases, and potential therapeutic agents for precise predictions of unknown drug efficacy against specific neurological disorder-associated targets.

In this thesis, we embark on a journey to address the crucial need for accurate and tailored therapeutic interventions for neurological disorders. To accomplish this goal, we begin with meticulous construction of Target-Disease-Drug Association Networks. These networks are designed to unravel the intricate relationships between biological targets, neurological diseases, and potential drugs. Through a rigorous curation process, we have compiled a comprehensive dataset comprising 374 known targets and 2,452 drugs associated with the nineteen neuronal disorders mentioned earlier. To ensure the reliability and accuracy of the data, we source information from reputable and established databases such as STRING, UniProt, DrugBank, Therapeutic Target Database, ChEMBL, and GeneCards. Drawing inspiration from network pharmacology, our research efforts are dedicated to building two fundamental networks: the Targets-Diseases Network and the Diseases-Drugs Network. The former seeks to shed light on the specific molecular targets' associations with the diverse range of neurological disorders under investigation. By identifying and understanding key targets crucial in the pathogenesis and progression of different neurological disorders, this network serves as a valuable resource for guiding future research and therapeutic development. In parallel, the Diseases-Drugs Network offers an extensive map of connections between neurological diseases and available drugs or compounds. This network is instrumental in exploring potential therapeutic options and opportunities for drug repurposing. By analyzing the associations between diseases and drugs, we can identify existing drugs that hold the potential to be repurposed for the treatment of specific neurological disorders. Drug repurposing presents a promising avenue for accelerating the development of effective treatments, capitalizing on the safety profiles and known mechanisms of established drugs. The construction of these association networks is a pivotal step towards gaining deeper insights into the complex interactions between targets, diseases, and drugs concerning neurological disorders. These networks lay the foundation for further analysis and exploration in subsequent objectives of the thesis. Leveraging the wealth of information encapsulated within these networks, we aim to identify promising therapeutic candidates and novel treatment strategies, ultimately advancing the field of precision medicine in neurology. Then we performed Machine Learning-Guided Classification and Drug Repurposing, by harnessing the constructed Target-Disease-Drug association network. With this network in place, we harness the power of state-of-

the-art machine learning algorithms to rank and link the role of different molecular targets in the context of subsequent diseases. The objective is to identify potential therapeutic interventions for neurological disorders that can significantly improve treatment outcomes. To achieve this objective, we employ a diverse set of non-linear Machine Learning algorithms, including Support Vector Machines (SVM), Decision Trees, Random Forests, Multi-layer Perceptron (MLP) Neural Networks, and Gradient Boosting Machines (GBM). These algorithms effectively capture complex relationships within the dataset and enable robust predictions. By training the machine learning models on the combined network dataset, which integrates information from the Targets-Diseases Network and the Diseases-Drugs Network, we gain a holistic perspective on the therapeutic landscape for neurological disorders. To ensure the accuracy and reliability of predictions, we undertake comprehensive hyperparameter tuning and rigorous validation procedures. Fine-tuning the algorithm parameters to achieve optimal performance and validating the models on independent datasets to assess their generalization capabilities are essential steps in this process. The ultimate goal is to obtain precise predictions of drug efficacy against specific neurological disorder-associated targets. Identifying drugs that exhibit potential efficacy against specific targets opens novel drug repurposing opportunities, expediting drug development processes in a cost-effective manner. The integration of network pharmacology and advanced machine learning offers a systematic and data-driven approach to identify promising drug candidates, revolutionizing neurological disorder management, and significantly impacting patient outcomes.

1.8 Thesis Objectives:

1. To build a comprehensive database of all the target-proteins and FDA-approved drugs for all the neurological and neurodegenerative disorders.
2. To develop Target-Disease, Disease-Drug and Target-Disease-Drug association networks to probe the target specificity for neuronal disorders.
3. To predict the target/drug attributes of neuronal disorders using linear and non–linear classification models.
4. To develop a model to repurpose an unknown drug against specific target and disorder type.

1.9 Thesis Outline:

The outline of this thesis is as follows: Chapter 1 introduces the research background, objectives, and significance, as presented in this introduction. Chapter 2 provides a comprehensive review of the literature, complexity and role of targets in neurological disorders, current state-of-the-art advancements in neurological disorders research, focusing on the molecular and cellular mechanisms of neurodegeneration progression, and the applications of network pharmacology, drug repurposing, and machine learning for target-disease-drug association prediction and treatment of neurological disorders. Chapter 3 describes the methodology, including data collection procedures, network construction, and machine learning model implementation. Chapter 4 presents the results and evaluation of the machine learning models using Target-Disease-Drug association networks, showcasing their performance in predicting drug efficacy against specific disorder types and their associated targets. The Discussion in Chapter 5 interprets the findings and discusses their implications for drug repurposing and precision medicine in neurological disorders. Moreover, we address the strengths and limitations of our approach and highlight potential avenues for future research. Finally, Chapter 6 concludes the thesis, summarizing the main contributions and emphasizing the potential impact of this research on advancing treatments for neurological disorders.

In summary, this thesis seeks to contribute to the growing field of drug repurposing and precision medicine for neurological disorders. By exploring the intricate associations between targets, diseases, and drugs, we aspire to pave the way for more accurate, effective, and personalized therapeutic interventions, ultimately improving the lives of patients and providing hope for a brighter future in the fight against neurological disorders.

Chapter 2

Literature Review

2 Literature Review

2.1 Neuronal disorders:

The progressive and continual death of brain's neurons, or nerve cells, is referred to as neurodegeneration. Numerous neurodegenerative illnesses are at the root of this phenomenon because nerve cells are unable to perform their tasks properly [89]. Nerve cell death is caused by the intracellular or extracellular accumulation of misfolded proteins. As a result of these disruptions, Apoptosis, or programmed cell death, is finally triggered in the regular functioning of these cells [90]. One of the main factors causing the reduction in overall brain function is the functional death of neurons. Because of this, the clinical importance of preventing neuronal death and dysfunction is enormous [91]. Neurological disorders are caused by the imbalance of ions and neurotransmitters between nerve cells. This imbalance results from aberrant neurotransmitter releases that can either stimulate, inhibit, or alter the activity of nearby neurons. Such imbalances influence neurotransmission, the mechanism through which nerve cells converse with one another, and they can cause a variety of psychological issues [92]. Neurons, which are the fundamental building blocks of the nervous system, differ from other cells in the body as they lack the ability to reproduce or replace themselves. The complexity of treating neurological illnesses is increased due to the limited ability for regeneration of the neurons. Additionally, ageing is a prominent example of a demographic characteristic that is associated with an increasing prevalence of neurological illnesses and may be a factor in the higher death rate for these disorders [93]. Neurodegenerative disorders are characterized by their persistent and disabling nature. They include the gradual deterioration and frequently the demise of neurons. The blood-brain barrier is a biological barrier that makes treating certain illnesses particularly difficult [94]. The purpose of this barrier is to keep dangerous chemicals out of the brain. Treatment strategies are further complicated in this situation by the complex regulatory networks in action in the body.

Every gene is linked with many other genes working in the cascade, forming a complex network in the nerve cell, and are involved in many cell regulation processes. Many proteins or enzymes play a supporting role in this regard. For the normal functioning of a cell, upregulation and downregulation of different genes play a vital role. Sometimes, specific genes show under and

overexpression in the nerve cell due to some genetic factors like mutations, and internal, or external factors, causing accumulation of unwanted proteins and abnormalities in the neuron. If we focus on one target with respect to one disease for drug binding, it may be possible that the target is involved in other normal functions, thus, affecting the normal functioning of other organs as well, which leads to abnormal cellular functions in the body. Therefore, predicting the target gene and understanding the complex network is crucial for the cure of neurodegenerative disorders [95]. There are trillions of neural pathways involved in carrying the information between the CNS and other organs of the body. These involvements of many neural pathways make the network more complex [91]. A huge number of genes/proteins perform different functions in the various organs of the body. More than one neurological disorder occurs by a mutation in a common gene but follows the different neural pathways and show different negative outcomes. That is why this network has become very complex and very challenging to handle.

2.2 Healthy versus degenerative neuron:

Each neurodegenerative disorder exhibit mutation in specific genes of genetic network. But some genes are common in more than one neurodegenerative disorder, that's why understanding the complete pathway regarding abnormalities in those genes would be beneficial. *Tau* proteins [96], *TDP-43* [97], *C9ORF72* [98], *PSEN1 & PSEN2, APP* [99], *LRRK2, APOE* [100], etc. are the genes involved in more than one neurodegenerative disorder. Let us talk about how tau protein is responsible for neurodegeneration. In normal neurons, tau phosphorylation is done by the activities of kinases such as *GSK, CDK, or ERK* and phosphatase 2A (*PP2A*), which is a heterotrimeric enzyme having 2 subunits i.e., structural subunit A and catalytic subunit C. Each subunit has two isoforms A and β. *PP2A* takes on regulatory subunits that compete in binding. Their binding is regulated by methylation of a catalytic subunit at C-terminal Leucine (L309) recruiting the cytoplasmic enzyme *LCMT1. PME-1* is a demethylating enzyme present in the nucleus. The C subunit is phosphorylated by *GSK3* at TYR(Y)307. There are two endogenous inhibitors, *I1PP2A* and *I2PP2A* (also known as *SET* or *TAF1*). Generally, *PP2A* activity is decreased by phosphorylation or via endogenous inhibitors and increased by methylation. Therefore, *PP2A* is active in the nucleus and inactive in the cytoplasm, preventing phosphorylation of tau in the

cytoplasm. Sometimes, the nuclear pores start to leak and *I2PP2A* again localizes to the cytoplasm. *I2PP2A* cleaved into two smaller fragments that can freely diffuse between the cytoplasm and nucleus. Phosphorylation of *I2PP2A* at Ser(S)9 by *CKII* (casein kinase II) causes its retention in the cytoplasm. Together with increased activity of *GSK3* in the cytoplasm, the net effect is increased tau phosphorylation.

Hyperphosphorylated tau disassembles microtubules that were assembled from normal tau and tubulin causing degeneration of neuron as sown in Figure 2.1. Fibrillar tau is toxic and underscored the importance of phosphorylation of tau in exerting this toxic effect [101]. Hyperphosphorylated

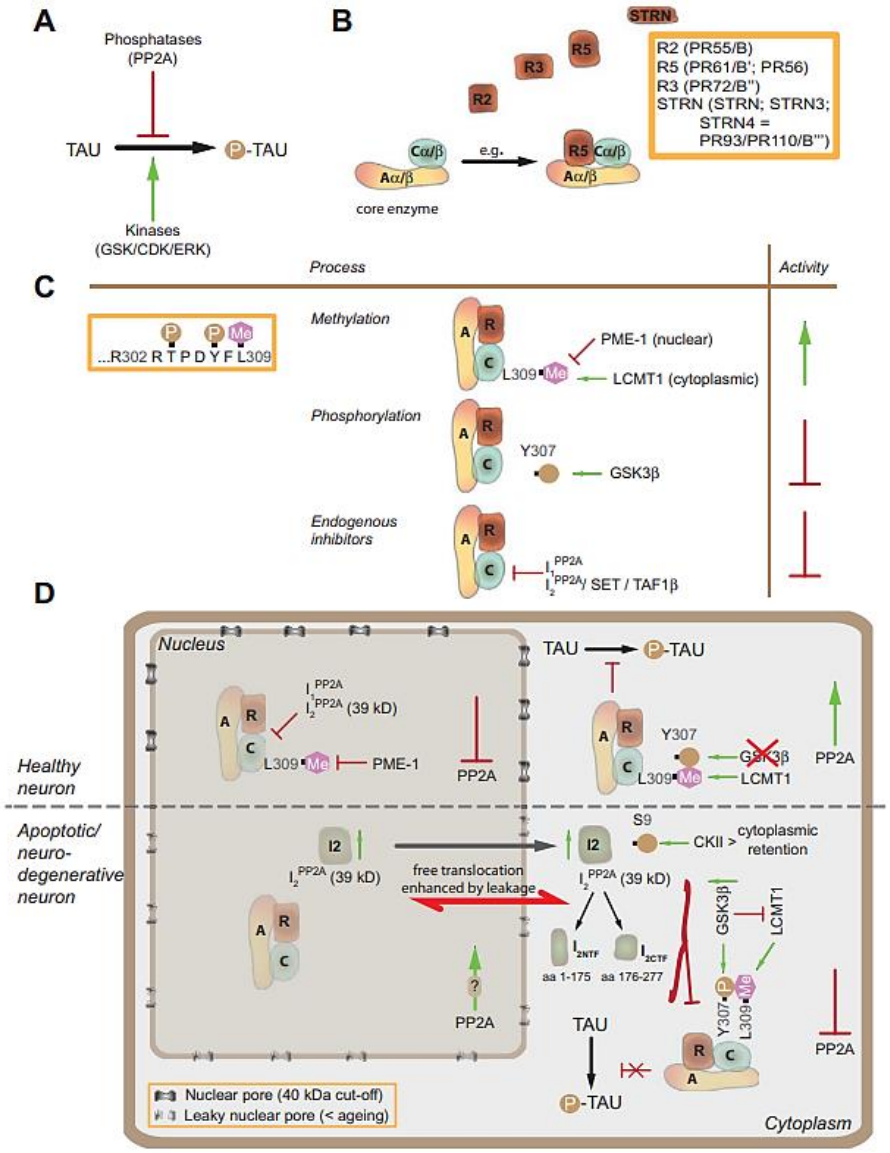

Figure 2.1 *Hyper phosphorylation of tau protein in Apoptotic Neuron*

tau is glycosylated, and abnormal phosphorylation might promote aggregation of tau, formed Neurofibrillary tangles (NFTs), and inhibition of the assembly of microtubules, glycosylation appeared to be responsible for the maintenance of the paired helical filaments (PHF) or tau structure. Of the tau sites required for microtubule binding, Ser262 and the AT180 epitope Thr231 are critical. Combined phosphorylation of tau at Thr212, Thr231 and Ser262 has been shown to cause neurodegeneration and has also been observed in the apoptosis of neurons.

2.3 Role of biological targets in neurological disorders:

Targets refer to specific molecular entities, such as proteins or genes, that are implicated in the pathogenesis or progression of these disorders. Knowing the function of targets can shed light on the disorder's underlying mechanisms. These targets play a crucial role in various cellular processes that contribute to the development and manifestation of neurological conditions. Different neurological disorders may share certain molecular targets, but the ways in which these targets are dysregulated can vary. The basis for logical medication design and precision medicine is the identification and understanding of targets in neurological illnesses. Finding targets enables the development of personalized medicine tactics, in which treatment plans are adapted to a person's genetic profile and the molecular reasons causing their condition [102]. Because they affect various pathways, not only the ones relevant to the condition, but many current therapies for neurological disorders also include side effects. Targeted medicines may be able to reduce side effects and increase patient quality of life by reducing off-target effects. Identifying and understanding targets in neurological disorders is essential for the development of effective therapeutic interventions for several important reasons [103].

A huge number of genes have been identified for different neurological disorders. The importance of certain protein targets such as *Amyloid-β (Aβ)* [104], *tau*, presenilin (*PSEN*), amyloid precursor protein (*APP*), in Alzheimer's disease (AD) are highlighted in literature. The creation of distinctive brain lesions, including plaques and tangles, which are hallmarks of AD, is primarily mediated by a buildup and alterations in tau protein. Cognitive decline and memory loss are brought on by these protein clumps, which interfere with neuronal activity.

Substantial research has been done on the function of many genes and proteins as targets in Parkinson's disease (PD). The importance of genes like *Parkin, PINK1*, and *LRRK2 FMR1, α(alpha)-synuclein, DJ – 1, PARK8,* and *GBA* in PD development is highlighted in literature [105]. Important physiological functions like protein breakdown, mitochondrial function, and neurotransmitter control are disrupted by mutations in these genes, which contributes to the selective death of dopaminergic neurons and the recognizable motor symptoms of Parkinson's disease (PD).

A few genes have also been identified as major targets in Frontotemporal Lobar Degeneration (FTLD) such as *C9ORF72, MAPT* (Microtubule-Associated Protein Tau), progranulin (*GRN*), *TDP-43* (transactive response DNA-binding protein with molecular weight 43 kDa) [12] [13], *FUS* (fused in sarcoma), *TARDBP* (transactive response DNA-binding protein), *CHMP2B* (Charged multivesicular body protein 2B), and *VCP* (Valosin-containing protein) mutations [106] are significant in FTLD and associated diseases. Changes in these genes impair cellular processes such as protein synthesis and breakdown, which causes an accumulation of aberrant proteins and neurodegeneration in particular brain regions. These examples from the literature show that it is crucial for understanding the underlying mechanisms and creating new therapeutic approaches to focus on specific genes or proteins linked to neurological illnesses. By concentrating on specific targets, scientists can try to bring back healthy cellular functions, stop the buildup of abnormal proteins, and eventually lessen the symptoms or stop the advancement of neurological illnesses. Finding new targets can also result in the repurposing of already approved medications i.e., existing drugs. It may be possible to assess the effectiveness of medications authorized for other ailments that also happen to interact with the identified target in the treatment of neurological disorders [107]. When compared to creating completely novel medications, this method is more cost- and time-effective.

In one of the studies, researchers found that higher expression levels of *A2A* and *P2X7* receptors in neurological disorders such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, multiple sclerosis, and epilepsy, further complicate the disease condition [108].

In another study, it was investigated that *S100B* protein plays a crucial role in Alzheimer's disease, Parkinson's disease, multiple sclerosis, Schizophrenia and epilepsy because the high expression of this protein directly targets astrocytes and promotes neuroinflammation. Under

stressful conditions, *S100B* produces toxic effects mediated through receptor for advanced glycation end products (AGE) binding. *S100B* also mediates neuroprotection, minimizes microgliosis and reduces the expression of tumor necrosis factor (TNF-alpha) but that are concentration-dependent mechanisms [109]. It was also proposed that *S100B* can be used as a potential therapeutic target to reduce the prevalence of neurological disorders.

In many cases, important roles of a particular gene in embryonic development have precluded the in vivo study of its function in the adult brain, which is usually the most relevant experimental context for the study of neurological disorders. Conditional knockout technology has been used successfully to generate viable mouse models with gene inactivation patterns in certain regions or cell types of the postnatal brain [110].

In amyotrophic lateral sclerosis, *AMPA* receptors allow cytotoxic levels of calcium into neurons, leading to motor neuron death. Likewise, in some epilepsies, overactivation of *AMPA* receptors leads to neuron damage. The same is true for ischemia, where oxygen deprivation leads to excitotoxicity. Conversely, Alzheimer's disease is characterized by decreased *AMPA* activation and synapse loss. Unfortunately, many clinical studies have had limited success by directly targeting *AMPA* receptors in these diseases. Indirectly affecting *AMPA* receptors or by regulating glutamatergic transmission, may provide new therapeutic potential for neurological disorders [111].

Gene expression changes in neuropsychiatric and neurodegenerative disorders, as well as gene responses to therapeutic drugs, offer new ways to identify central nervous system (CNS) targets for drug discovery. Targets for Alzheimer's disease and cognitive decline associated with normal aging and mild cognitive impairment include *τ, amyloid-β* precursor protein, *Aβ*, all three high-affinity neurotrophins receptors, fibroblast growth factor (FGF) system, synapse markers, glutamate receptors (GluRs and transporters), and dopamine (*DA*) receptors, particularly the *D2* subtype. Gene-based candidates for Parkinson's disease include the ubiquitin–proteosome system, scavengers of reactive oxygen species, brain-derived neurotrophic factor (*BDNF*), its receptor, *TrkB*, and downstream target early growth response 1, *Nurr-1*, and signaling through protein *kinase C* and *RAS* pathways. Studies in schizophrenia reveal robust decreases in genes for *GABA* function, including glutamic acid decarboxylase, *HINT1*, glutamate transport and *GluRs, BDNF* and *TrkB*,

numerous 14-3-3 protein family members, and decreases in genes for CNS synaptic and metabolic functions, particularly glycolysis and ATP generation [112].

In a study, it was proposed that the carboxy-terminus of Hsc70-interacting protein (*CHIP*) is a crucial molecular co-chaperone and ubiquitin E3 ligase that regulates various biological functions, including misfolded-protein refolding, autophagy, immunity, and necroptosis. Its ubiquitous expression in the central nervous system suggests its involvement in various neurological diseases. Recent studies have highlighted *CHIP*'s beneficial role in the pathogenesis of stroke, intracerebral hemorrhage, Alzheimer's disease, Parkinson's disease, and polyglutamine diseases. *CHIP* mutations could also cause neurodegenerative diseases. Overexpression of *CHIP* could be a promising therapeutic target for several neurological diseases, based on available literature [113].

In a finding, it was discovered that Acid-sensing ion channels (ASICs) are voltage-independent, proton-gated cation channels found in the central and peripheral nervous system. They detect pH changes during various activities, including pain perception, synaptic plasticity, learning, memory, fear, and neuronal degeneration. ASICs are potential therapeutic targets for manipulating pain and neurological diseases [114].

Authors, in another finding, investigated that Excitatory amino acid transporter 2 (*EAAT2*) is a crucial neurotransmitter in the central nervous system, responsible for clearing extracellular glutamate to prevent neuronal excitotoxicity and hyperexcitability. It regulates synaptic activity and plasticity and has been linked to various central nervous system disorders. *EAAT2*'s structure, pharmacology, physiology, and functions are essential in understanding its role in various diseases like stroke, Parkinson's disease, epilepsy, Alzheimer's disease, major depressive disorder, and addiction. Up-regulation of *EAAT2* protein has shown significant benefits in various disease models, suggesting its activation as a promising therapeutic approach [115].

Cholesterol is a crucial component of the cell membrane, affecting membrane-bound protein permeability and function. It plays a role in synaptogenesis, axonal growth, dendrite outgrowth, and microtubule stability. Cholesterol metabolism in the brain is primarily mediated by *CYP46A1*, or cholesterol 24-hydroxylase, which eliminates about 80% of cholesterol excess. Studies show that cholesterol and 24HC levels change during neurological diseases, suggesting that inhibition or activation of *CYP46A1* could be an effective therapeutic strategy. Preclinical studies have assessed its role in neurodegenerative disorders like Parkinson's, Huntington's, Alzheimer's,

multiple sclerosis, spinocerebellar ataxias, and amyotrophic lateral sclerosis. Recent development of soticlestat, a selective and potent *CYP46A1* inhibitor, has significant anti-seizure effects, indicating its importance for future drug developments. Both activation and inhibition of *CYP46A1* are of therapeutic value [116].

In one of the studies, Protein tyrosine phosphatase 1B (*PTP1B*) has been discovered as a key enzyme in the *PTP* family, responsible for regulating receptors and kinases. It has been linked to various diseases, including schizophrenia, anxiety, neurodegeneration, neuroinflammation, and depression. Inhibition of *PTP1B* can prevent microglial activation, promoting anti-inflammatory effects and potentially increasing cognitive function through stimulation of hippocampal insulin, leptin, and *BDNF/TrkB* receptors. However, most research on *PTP1B*'s clinical efficacy has focused on obesity and type 2 diabetes mellitus. Despite the link between metabolic alterations and neurodegeneration, no clinical trials have assessed the neurological benefits of *PTP1B* inhibition. Preclinical studies suggest that targeting *PTP1B* could reach various pathophysiological mechanisms simultaneously [117].

The *NLRP3* (NLRP3: NOD-, LRR-, and pyrin domain-containing protein 3) inflammasome is the best-described inflammasome that plays a crucial role in the immune system and various diseases, including neurological disorders. Its association with neurodegenerative diseases and strokes highlights its importance as a clinical target for pharmacological intervention. However, the mechanism of *NLRP3* activation remains indefinite. Emerging pharmacological approaches targeting *NLRP3* inflammasome in neurological diseases have clinical translational potential. Chinese herbal medicine and botanical ingredients have been specifically focused on as potential therapeutics for central nervous system disorders, potentially contributing new perspectives to neurological disease treatment [118].

It is proposed in the literature that mutations(repeats) in the *C90RF72* gene are involved in many neurological disorders such as Amyotrophic Lateral Sclerosis (ALS)/ Motor Neuron Disease (MND) and Frontotemporal Dementia (FTD). Recent studies show the disease-target association of Psychotic patients with the *C90RF72* gene. The study demonstrates that the genetic counseling of patients having psychotic disorders reveals the mutations(repeats) in the *C90RF72* gene [119].

Neurotransmitters are the chemicals released between presynaptic and postsynaptic neurons for the transmission of nerve impulses either by their excitatory or inhibitory role [120]. They play an

important role in the brain by influencing mood, which is why they are sometimes described as "feel-good" chemicals. Five important neurotransmitters include dopamine, serotonin, oxytocin, norepinephrine, and endorphins. Malfunctions of these neurotransmission processes can result in clinical disease. The Loss of memory in Alzheimer's disease is postulated to involve insufficiency of the neurotransmitter acetylcholine in synapses, which mediates the laying down of new memories. Certain drugs block the enzyme acetylcholinesterase (which breaks down acetylcholine) and thus increase the amount of acetylcholine in the synapse. As a result, memory function may improve [121].

Serotonin is a neurotransmitter affecting multiple physiological processes and cognitive brain functions, among them mood and emotions, which is why it has been linked to mood disorders such as depression [122]. Serotonin (5-hydroxytryptamine, or 5-HT) is generated by the raphe nucleus and midline neurons of the pons and upper brain stem. Serotonin levels are controlled by the uptake of tryptophan and intraneuronal monoamine oxidase (MAO), which breaks down serotonin. Ultimately, serotonin is excreted in the urine as 5-hydroxyindoacetic acid or 5-HIAA. Serotoninergic (5-HT) receptors are classified as 5-HT1, 5-HT2, and 5-HT3. Selective serotonin receptor agonists (e.g., sumatriptan) can abort migraines. Selective serotonin reuptake inhibitors (SSRIs) can also be used to treat several mental health disorders (e.g., depression, anxiety, obsessive-compulsive disorder, and post-traumatic stress disorder) [50]. SSRIs perform symptomatic treatment.

2.4 Network Pharmacology in Neurological disorders:

Network pharmacology is a bioinformatics-based research strategy used to explore the anti-epileptic mechanism of Rhizoma Coptidis. One of the studies predicted protein targets and validated the interaction between active components and predicted targets using molecular docking technology. Nine active compounds were selected, with 68 targets associated with Rhizoma Coptidis treating epilepsy. KEGG pathway enrichment analysis identified 89 signaling pathways related to epilepsy. Quercetin and NAÏVE-canadine exhibited good docking with key targets, suggesting Rhizoma Coptidis can regulate various signaling pathways and have therapeutic effects on epilepsy [123].

Epilepsy is the fourth most common neurological disease, with one-third of cases refractory to existing anticonvulsants. The study aims to discover new pharmacological targets for the treatment of Refractory Epilepsy (RE) using network pharmacology methods. The researchers selected 83 potential targets linked to 83 genes associated with RE development, and then selected 10 most promising targets based on published data. All selected target proteins play a key role in biological processes involved in RE development. Nine of the 10 targets have potential associations with different types of epilepsy, highlighting the potential of network pharmacology in finding new molecular targets for RE treatment [124].

Ayurvedic medications, such as Saraswatarishta (SWRT), are prescribed to control neurological disorders like slurred speech, anxiety, Parkinson's disease, and Alzheimer's disease. However, there is limited scientific research on SWRT's mode of action. This study uses network pharmacology to understand its neuroprotective role in neurological disorders. Out of the 18 ingredients in SWRT, five were considered in this study due to their elevated therapeutic action in neurological disorders. Further, nine active phytoconstituents were chosen from the five selected ingredients. Gene targets were screened and selected using STITCH, SwissTargetPrediction, and ChEMBL. Protein-Protein interaction and Gene Ontology enrichment analysis were performed using STRING and g:Profiler. Cytoscape 3.7.2 was used to create three networks, and bioactivity scores and blood-brain barrier probability scores were obtained. The phytoconstituents were found to be linked to gene targets involved in 10 major neurological disorders, with bioactivity scores in the active range and high BBB probability scores [125].

One of the studies explores a mechanism-based disease definition for network pharmacology, focusing on ischemic stroke and reactive oxygen species (ROS) forming NADPH oxidase type 4 (Nox4) as primary causal targets. The study used classical protein-protein interactions and metabolite-dependent interactions to identify suitable synergistic cotargets for network pharmacology. The nitric oxide synthase gene family is identified as the closest target to Nox4. Combining a NOS and a NOX inhibitor at subthreshold concentrations results in pharmacological synergy, reducing cell death, infarct size, stabilized blood-brain barrier, reduced reoxygenation-induced leakage, and preserved neuromotor function. This approach potentially reduces the risk of failure in single-target and symptom-based drug discovery and therapy [126].

C. pluricaulis Choisy, a perennial herb used in traditional folk medicine, has been extensively researched and analyzed for its phytochemistry, neuropharmacological, and toxicological properties. The herb and its metabolites have been found to have various in vitro and in vivo neuropharmacological effects, including memory enhancement, anxiolytic, tranquilizing, anti-depressant, anti-stress, neurodegenerative, anti-inflammatory, antioxidant, analgesic, sedative, anti-convulsant, and Alzheimer's disease-reversing effects. Network pharmacology results suggest that C. pluricaulis compounds interact with proteins, neuro synapses, signaling pathways, and serotonergic synapse, which are crucial in neurotransmission, Alzheimer's disease, long-term depression, alcohol addiction, cognitive disorders, psychological conditions, and increasing serotonin concentration in synapses [127].

Icariin is a biologically active substance in Epimedii herba that is used for the treatment of neurologic disorders. A comprehensive analysis of the molecular mechanisms of icariin is lacking. In a study, a brief overview of the history of icariin's used as a medication, the active chemical elements of Epimedii herba are discussed, and looked at the data from experimental investigations that have shown the molecular targets of icariin in various disorders. To predict the therapeutic effects of icariin in nervous system diseases like Alzheimer's disease, Parkinson's disease, ischemic stroke, depressive disorder, multiple sclerosis, glioblastoma, and hereditary spastic paraplegias, the researchers built a protein-protein interaction network and performed Gene Ontology and Kyoto Encyclopaedia of Genes and Genomes functional enrichment analyses using Network Pharmacology approach. The conclusions from analyses can guide future studies on the application of icariin to the treatment of neurologic disorders [128].

Major Depressive disorder is a common mental disorder characterized by depressed mood and loss of interest or pleasure. As the Herbal medicines are mainly used as complementary and alternative therapy for depression. A study investigates the antidepressant activity of Huang-lian Jie-du Decoction (HLJDD) and its potential depression-associated targets. HLJDD was administered to chronic unpredictable mild stress-induced (CUMS) depressive mice, and its effects were evaluated through force swimming test, novelty-suppressed feeding test, and open field test. Active components of HLJDD, potential targets, and metabolic pathways involved in depression were explored through systemic biology-based network pharmacology assay, molecular docking and metabolomics. The study identified 28 active compounds and ten biochemical pathways

involved in HLJDD. The findings of the study show that HLJDD exhibited antidepressant effects. SLC6A4 and MAOA may be the main antidepressant targets for HLJDD [129].

Inflammatory responses play an extraordinary role in the pathogenesis of cerebrovascular and neurological disorders. One significant flavonoid, baicalin, is obtained from Scutellaria baicalensis Georgi. Baicalin has recently been proven in multiple in vivo and in vitro investigations to have positive effects on anti-inflammatory and immunomodulatory processes, as well as to exert positive therapeutic benefits in cerebrovascular and neurological illnesses. In this review, anti-inflammatory effects of baicalin are studied via multiple pathways and targets, that affect the production of a variety of inflammatory cytokines and the neuroprotective process of neurological diseases. The related targets of the baicalin's anti-inflammatory effects were analysed using the tools of network pharmacology, providing theoretical support and novel ideas for the potential clinical use of baicalin in the future [130].

2.5 Drug Repurposing in Neurological disorders:

Drug repurposing or repositioning refers to the study of clinically approved drugs in one disease to see if they have therapeutic value and do not trigger side effects in other diseases. Today, it is crucial to examine potential therapeutic benefits of already available medications or drug candidates in a range of human diseases, including neurological disorders. The lack of funding and time constraints seen during conventional drug development are overcome by this method. It offers hope for some refractory illnesses, such as neurological conditions. Drug repurposing is especially crucial since neuropathological problems generally make it more difficult to produce new medications than diseases in other organs due to the nervous system's complicated structural makeup and the blood-brain barrier's influence. Drug repurposing can be used to treat neurological diseases, summarize the repurposing candidates that are presently being tested in clinical trials for neurological diseases, and present some early findings [131].

Traumatic brain injury (TBI) is a major global cause of death and disability, with no FDA-approved drugs to substantially attenuate its effects. This has led to the emergence of drug repurposing, which involves repurposing existing drugs with well-characterized mechanisms of

action and human safety profiles. Compared to the conventional discovery pathways, drug repurposing is less costly, relatively rapid, and poses minimal risk of adverse outcomes. Drug repurposing has been applied to various neurodegenerative diseases and neurological disorders, including brain injury. Edaravone, glyburide, ceftriaxone, levetiracetam, and progesterone are selected as potential TBI neurotherapeutic agents. Although FDA-approved for other purposes, they have shown efficacy in ameliorating the detrimental outcomes of TBI in preclinical and clinical studies [132].

Computational drug repurposing has the potential to significantly reduce drug development time and cost, particularly in neurodegenerative diseases like Alzheimer's disease. This approach involves in silico screening of FDA-approved compounds for new indications and has the potential to expedite the development of effective therapies for these diseases. Traditional drug development, which can take 15 years and over one billion dollars, involves discovery, pre-clinical research, safety review, clinical studies, FDA review, and post-market safety monitoring. However, many repurposed drugs have already been FDA approved, making them a cheaper and quicker route to the clinic. High throughput screening technologies and the growing repository of 'omics-based data across disease indications have catapulted computational drug repurposing methods to the forefront of attractive drug discovery techniques for neurodegenerative diseases. The integration of artificial intelligence and machine learning algorithms will enable the creation of large-scale transcriptomic and electronic medical record databases. However, this process presents unique challenges due to the lack of effective validation methods and the heterogeneous nature of the disease. Successful repurposed drugs exist in fields like oncology, diabetes, leprosy, and inflammatory bowel disease. This study examines existing approaches to computational drug repurposing, including molecular, clinical, and biophysical methods, and proposes data sources and methods to advance computational drug repurposing in neurodegenerative diseases [133].

Previously, fortunate discoveries in the laboratory and clinic resulted in the success of repurposed medications. One pertinent illustration of this is the way zonisamide is used to treat Parkinson's disease. Murata found that when zonisamide was used to treat a Japanese epileptic patient who also had Parkinson's disease (PD), the patient's PD symptoms also improved. In 2009, Japan approved zonisamide as an anti-PD medication based on this coincidental discovery. High throughput molecular, clinical, and structural biology technologies, along with the development of

large-scale computational capacity that is economically feasible, have recently given rise to a novel opportunity: the use of computational frameworks rather than random discoveries to rationally repurpose existing drugs [133].

The development of new treatments for acute stroke has been fraught with costly and spectacularly disappointing failures. Repurposing of drugs that are previously proven to be secure offers a less dangerous option. Drug repurposing involves taking use of commercial medications' secondary activities, and pursuing compounds with several modes of action, including vascular protection. Protecting the ischemic vasculature is expected to offer long-term advantages and support neural rehabilitation for stroke patients. Currently, acute aspirin therapy and drug-assisted reperfusion are employed clinically to lessen ischemic stroke-related impairment. The use of growth factors like erythropoietin and medications like statins, angiotensin II receptor blockers, and minocycline is possible in the future. A clinical experiment on acute ischemic stroke has already shown that the angiotensin II receptor blocker candesartan can protect blood vessels [134].

Drug repurposing refers to a reinvestigation of existing drugs for new therapeutic interventions. It is a promising, fast, and cost-effective method that can overcome traditional de novo drug discovery and development challenges in targeting neuropsychiatric and other disorders. Traditional methods are complicated due to limitations in understanding pathophysiological phenomena and are risky, expensive, and time-consuming. Various drug classes such as selective serotonin reuptake inhibitors (SSRIs), antipsychotic, cholinesterase inhibitors, and thrombolytic agents show polypharmacological features. In addition, amantadine was initially developed for influenza; however, after redirection, it is useful for Parkinson's disease. Zidovudine was intended for cancer treatment, and now it is redirected to targeting HIV/AIDS. An additional, but well-known example is Viagra (Sildenafil) that was intended to antianginal medication but redirected to penile erections. Drug repurposing takes advantage of off-target effects of existing drugs, identifying new opportunities by understanding their biological and pharmacological mechanisms. This approach is more effective in developing drugs against neuropsychiatric and other disorders [135].

2.6 Existing research on ML Algorithms for target-disease-drug association prediction:

Artificial intelligence (AI) and machine learning (ML) methods are increasingly being used to combine various types of data, including transcriptomic, structural, and clinical data. Companies like IBM have developed AI- and ML-based frameworks for drug discovery, using AI-based text-mining strategies to create a semantic model of ALS-associated RNA-binding proteins. ML is particularly attractive in computational drug repurposing, where molecular and biophysical data are integrated. A recent method used drug-induced gene expression signatures, molecular target information, and structural information as features to train a multiclass support vector machine to predict the therapeutic class of a given drug. This approach has shown a classification accuracy of 78%, demonstrating its potential usefulness in neurodegenerative disease. While AI and ML models have shown promise in disease prediction for Parkinson's disease, MS, and Alzheimer's disease, their full utility in computational drug repurposing for neurodegenerative diseases would be realized as molecular, structural, and clinical data resources for neurodegenerative diseases increase [133].

A study showed that supervised machine learning models have been proven feasible for Drug Target Interaction (DTI) prediction, but they often generate inaccurate predictive results due to their disregard for unlabeled drug-target pairs [136]. Similarity-based methods have limitations when extending to large data sets due to high complexity of similarity matrices computation. Nearest neighbor methods, bipartite local models, matrix factorization methods, semi-supervised learning, ensemble methods, and ensemble methods all have their strengths and weaknesses. However, machine learning has achieved favorable performance in DTI prediction. Factors such as problem formulation, evaluation data set, evaluation procedure, and experimental setting significantly impact prediction results. The imbalanced dataset problem is another challenge, as current models like decision trees and SVMs have a bias for recognizing the majority class, resulting in poor performance. Most machine learning models have poor interpretability properties, making it difficult to understand the underlying drug mechanism of action from a biological perspective. There are no uniform evaluation metrics special for DTI prediction, but AUPR and AUC are generally adequate metrics for evaluating the performance of machine learning-based

methods. In the currently accessible datasets, the number of unknown samples is much greater than the known ones, so false positives should be weighed more. Overall, machine learning methods have the potential to improve DTI prediction accuracy.

A study presents a network-based method called NEDNBI, which predicts disease-drug associations using a gene-disease-drug tripartite network. The architecture of this network is shown in Figure 2.2. This method is useful for drug repurposing, especially in finding old drugs for new diseases like COVID-19. The method requires no negative data and allows new diseases to be added to the network. The evaluation results show good performance, with 8 out of 20 predicted old drugs clinically tested for COVID-19 treatment, demonstrating the method's usefulness in drug repurposing [137].

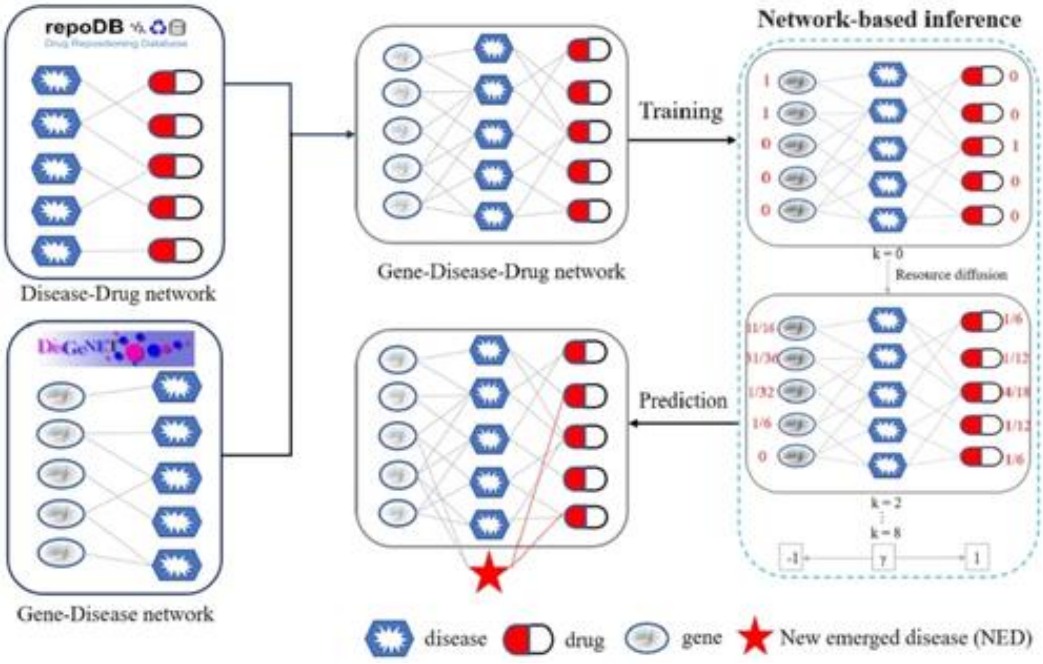

Figure 2.2 *Drug Repurposing by Gene-Disease-Drug Network*

A study presents a prediction method called multi-scale topology learning for drug-disease (MTRD) that integrates and learns multi-scale neighboring topologies and attributes of a pair of drug and disease nodes. It constructs multiple drug-disease heterogenous networks to integrate drug similarities and associations. The method uses a Bi-directional long short-term memory-based module to encode these embeddings and their relationships. Attention mechanisms at feature and scale levels are designed to obtain more informative pairwise features and topology

embeddings. MTRD achieves superior performance than other methods and retrieves more actual drug-disease associations in top-ranked candidates [138].

Predicting binding affinity between compounds and proteins is crucial in drug discovery, as it reduces the need for wet-lab experiments. Machine learning and deep-learning techniques, including ligand-based and target-based approaches, are used to improve drug-target interaction prediction. Popular machine-learning models include SVM, Random Forest, Naïve Bayes, KNN, GBT, GP, and XGBoost. Deep learning, using artificial neural networks, is used in medicinal chemistry for compound classification, QSAR studies, and drug identification. PCM models identify probable targets, predict binding affinity, and discover interactions between compounds and targets. Convolutional neural networks predict binding affinity using 1D representations of proteins and compounds [139].

Drug-Target interaction (DTI) is crucial for drug discovery, repositioning, and understanding drug side effects. However, the exponential growth of genomic and drug data makes it difficult to identify new associations between drugs and targets. A study addresses these challenges by developing a predictive model for DTI prediction using computational methods. The study is conducted on four protein classes: Enzyme, Ion Channel, G Protein-Coupled Receptor (GPCR), and Nuclear Receptor. The target protein sequence is encoded using the dipeptide composition and drug with a molecular descriptor. A machine learning approach is employed to predict DTI using wrapper feature selection and synthetic minority oversampling technique (SMOTE). To deal with the problem of DTI, various classifiers are used in this study. This method could identify one target protein interacting with many drugs and many drugs interacting with one protein, which are experimentally verified. It can be used for understanding and identifying new drug-target interactions. This method relies only on the dipeptide composition of the target descriptors [140].

A study utilizing machine learning methods to predict druggability of proteins used 443 sequence-derived features revealed the Neural Network as the most accurate classifier with 89.98% accuracy. The Support Vector Machine-Feature Selection (SVM-FS) algorithm had the most relevant features at 130. This led to the discovery of new drug targets for cell signaling pathways, gene expression, and signal transduction. Sequence properties determine a protein's targetability, and increasing the number of features is crucial for better prediction performance. Among the algorithms used, NN showed superior performance compared to Naïve Bayes, SVM,

kNN, RF, and DT models. Protein-drug interactions, including amino acid hydrophobicity, acidity, alpha helix, and sulfur atoms, are pivotal features in protein-drug interactions. This study demonstrates that combining different protein attributes and efficient machine-learning algorithms can significantly improve the predictability of target proteins [141].

Network medicine is a promising tool for understanding disease molecular complexities and identifying new drug targets. Computational approaches for drug repositioning integrate information from multiple sources and levels, providing valuable insights into complex relationships among drugs, targets, disease genes, and diseases at a system level. This article proposes a computational framework based on a heterogeneous network model for drug repositioning using existing omics data about diseases, drugs, and drug targets. The framework significantly outperforms several recent approaches, with case studies demonstrating its practical usefulness. The three-layer heterogeneous graph model captures inter- and intra-relationships among diseases, drugs, and targets for novel drug usage prediction. An iterative algorithm is developed to obtain final proximity scores between diseases and drugs, which can be used to rank candidate drugs for each disease [142].

Based on the essential findings discussed in this literature review, network pharmacology approach is beneficial for a better understanding of interactions between multiple targets associated in neurological as well as neurodegenerative disorders. It has been inferred that drug repurposing has provided a powerful technique for personalized treatment of neurological disorders. Through network pharmacology, researchers have been able to identify complex interaction patterns between multiple targets in a neurological disease and gain insight into the molecular mechanisms driving neurodegeneration and retarded nerve impulse transmission. Additionally, machine learning approaches used in drug repurposing has identified interaction between therapeutic targets and already approved drugs which helped to pave the way for personalized medicine approaches. However, there are still knowledge gaps in this field that need to be addressed, such as the better understanding of the target overlap exist between multiple neurological and neurodegenerative disorders and identification of accurate machine learning model and deep neural network that can be predictive for better treatment response to neuronal disorders. Another potential knowledge gap could be the development of database for all the protein druggable targets and all the FDA-approved drugs for all the important and lethal neurological and neurodegenerative disorders by

utilizing various online available databases for targets and drugs data. Machine learning models require huge amounts of data to be trained. A lot of work has been done on the Target-Drug predictions but with limited data and on specific individual disorders. Addressing these gaps can help to further advance our understanding of neurological and neurodegenerative disorders and improve outcomes for patients. Overall, the findings discussed in this literature review highlight the power of network pharmacology, machine learning, and drug repurposing for neuronal disorders research and the potential for this technique to revolutionize neuronal disorders treatment. The uniqueness of our work is the development of Target-Disease-Drugs Association network for all the lethal neurological and neurodegenerative disorders and using machine learning and deep neural network approaches for performing targets and drugs classification with respect to specific disorder class and ultimately performing drug repurposing for Neurological disorders.

Chapter 3

Methodology

3 Methodology

The aim of the current study is to probe the association of the numerous targets with the drugs and with the diseases i.e., neurological disorders (NLDs) and neurodegenerative disorders (NDDs). In this chapter, the materials and methods used to collect all the data of disease-associated targets and already approved drugs for the 19 neuronal disorders and to build the target-disease, drug-disease and target-disease-drug association networks, classification using state-of-the-art machine learning and deep-learning methods and ultimately drug repurposing for neurological disorders using multi-variate artificial neural network architecture are described in detail. A lot of publicly available databases were used to collect the datasets for targets and drugs for all the nine neurological and 10 neurodegenerative disorders, which were further analyzed in this study. The dataset was pre-processed by various data preprocessing techniques including handling missing values, deduction of less important feature variables, normalization of data etc. and to construct and classify networks using the Network pharmacology and Machine learning based approaches. A standardized pipeline for data processing was applied in this study, which included filtering out empty cells and normalizing feature variable values.

The methods described in this chapter were used to generate the results and insights presented in the following chapter. Details of the methodology employed in this study for Target-Disease-Drug Association Network-guided Classification are described in Figure 3.1 and multi-variate ANN model architecture for prediction of Targets descriptors and disease class is shown in Figure 3.1.

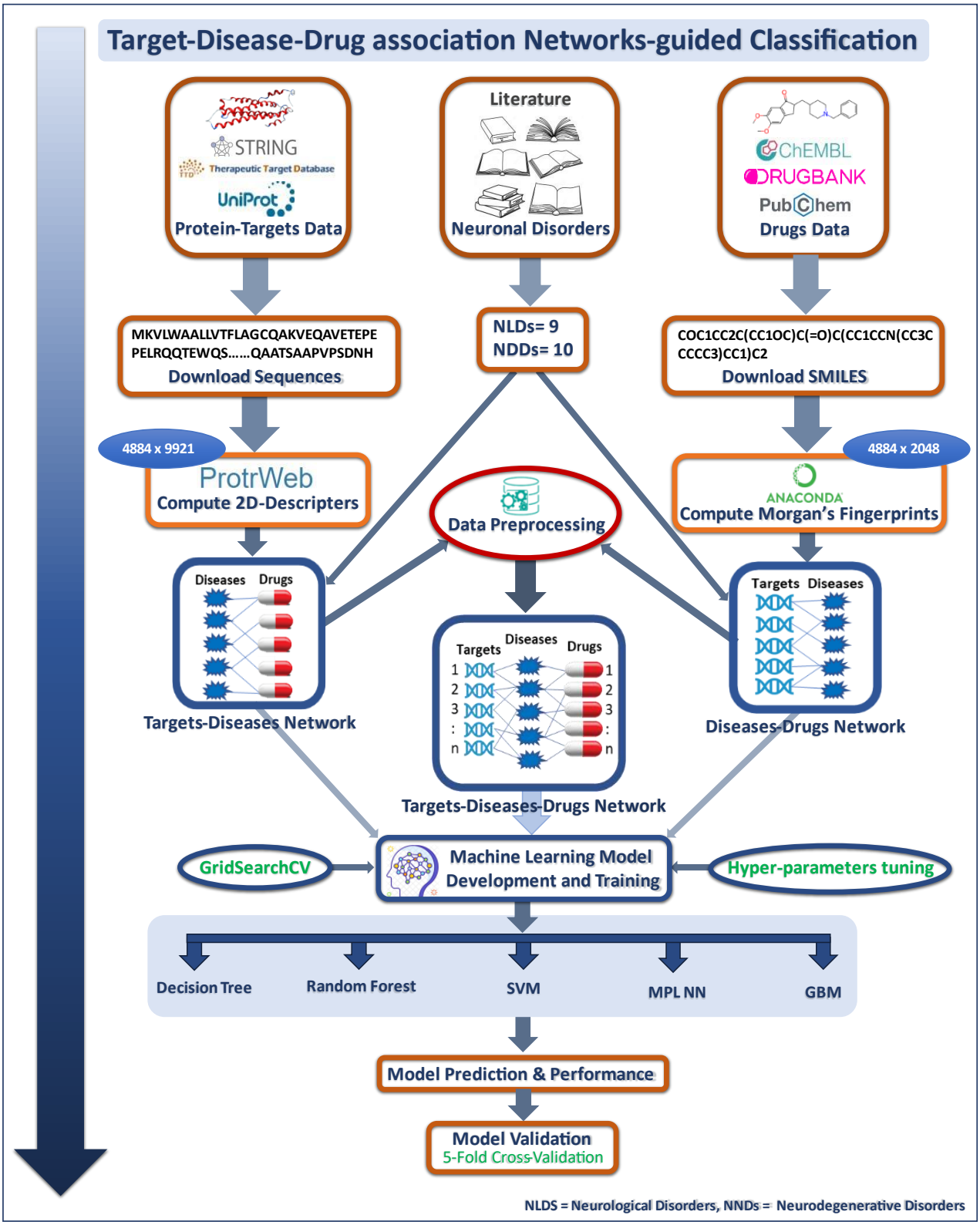

Figure 3.1 *Overall Methodology of Target-Disease-Drug Association Network-guided Classification*

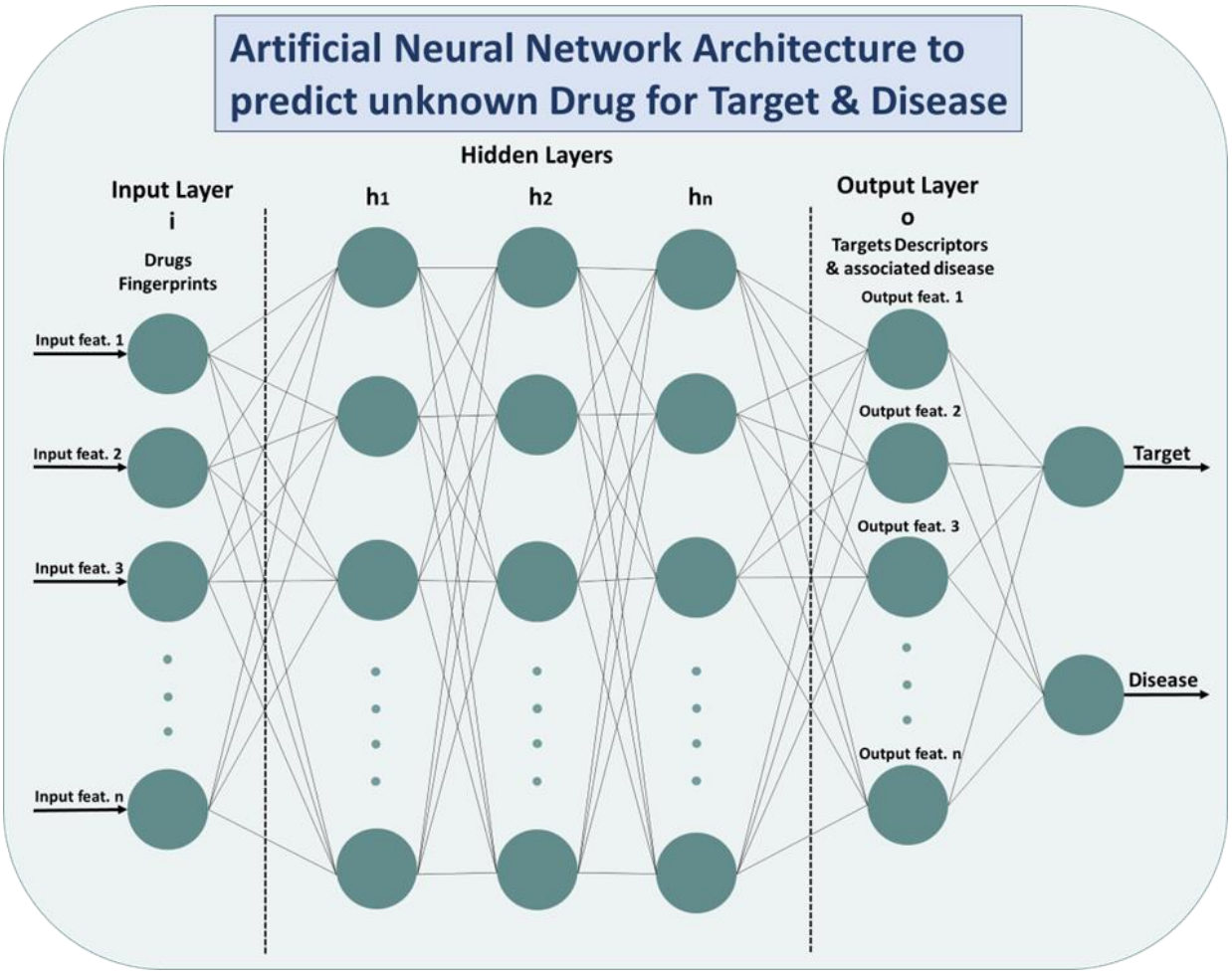

Figure 3.2 *multi-variate ANN Architecture for Target & Disease Prediction*

3.1 Data Collection:

3.1.1 Target Data Retrieval:

The data of all the protein-targets was retrieved from STRING database [145] and Therapeutic Target Database (TTD) [146] by giving the query of each Neurological and Neurodegenerative disorder individually. Subsequently, we merged the dataset of both databases to remove target duplicates. After removing duplicates, the data contained 821 unique targets protein between the protein-protein interaction edges zero to seventy-one. We shortlisted the targets datasets having edges of at least three to seventy-one. So, the total unique data of target proteins remained 236.

But the total dataset is composed of 4884 protein targets because of the target-overlap exists between multiple neurological as well as neurodegenerative disorders. The dataset contains the protein name, gene name, family, no. of edges and the UniProt IDs of all the target proteins. The sequences of all the selected target proteins were obtained from UniProt [147] using the given UniProt IDs and downloaded the data in an Excel file.

3.1.2 Drugs Data Retrieval:

The data of all the FDA-approved, experimental phase, investigational phase, nutraceutical, Illicit and withdrawn drugs was retrieved from multiple databases, such as, DrugCentral [148], e-Drug3D [149], ChEMBL [150] and DrugBank [151] by giving the query of all the selected protein-targets one by one and all for all the Neurological and Neurodegenerative disorder individually. Then we merged the dataset of all the databases to remove duplicates of all drugs. Then we keep only the data of all FDA-approved drugs. After shortlisting, the data contained 964 unique drugs. But the total dataset is composed of 4884 drugs because of the complementarity of same drug for more than one protein-target associated with multiple neurological as well as neurodegenerative disorders. The dataset contains the drug name, synonym, type, phase, physicochemical properties along with the ChEMBL IDs and DrugBank IDs of all the drugs. The SMILES of all the selected drugs were obtained from PubChem database using the given drug name and downloaded the data in an Excel file.

3.1.3 Selection of Diseases:

Neurological and Neurodegenerative disorders are selected due to their prevalence and mortality rate worldwide, with the help of literature and different organizations such as World Health organization (WHO) [152], Global Burden of Diseases (GBD) [153], etc. The data of all the protein-targets associated with the selected disorders, and all the FDA-approved drugs for those specific targets was collected. To identify and validate the protein targets that have a role in the progression of neurological and neurodegenerative disorders, STRING app plugin of Cytoscape tool version 3.7.1 [154] was used.

3.2 Descriptors Formation:

Once the data of the protein-targets, drugs and diseases was collected the descriptors were computed for proteins and drugs to train the various machine learning and deep learning models.

3.2.1 Protein Descriptors Computation:

The sequence-based descriptors of all the selected protein-targets were computed using web interphase of protr, a package of R. The ProtrWeb [155] is a freely available and used for the computation of the structural and physiochemical descriptors of the proteins. 9921 descriptors were calculated for each of the 4884 proteins including amino acid composition, dipeptide and tripeptide composition, C/T/D (Composition/Transition/Distribution), conjoint triad, sequence-order coupling number, Quasi-sequence-order descriptors, pseudo-amino acid, and amphiphilic pseudo-amino acid composition. After computation of all these descriptors, it was saved in an excel file having binary class label of neurodegenerative and neurological disorders. The order of the data matrix of the Morgan fingerprints of inhibitors was 4884×9921.

3.2.2 Drugs Descriptors Computation:

For the drugs dataset, Morgan fingerprints were generated which is the best molecular fingerprint used for drug discovery purposes [156]. For the calculation of the Morgan fingerprints, SMILES of the inhibitors were used as input. The SMILES of all the drugs were extracted from PubChem [157]. The ALLChem package of RDKit library [158] was imported in Python for the generation of Morgan fingerprints. Morgan fingerprints were in the form of bits string (0,1) with a length of 2048. The order of the data matrix of the Morgan fingerprints of inhibitors was 4884×2048. After computation of Morgan fingerprints for all the drug's SMILES, it was saved in an excel file having binary class label of neurodegenerative and neurological disorders.

3.3 Tools:

To generate Morgan fingerprints of all the drugs' SMILES in python language and all the machine learning/deep learning models development, training and validation for classification of networks datasets and for drug repurposing purposes, Jupyter notebook of Anaconda distribution was used. To Compute the descriptors of targets proteins in R, R-studio tool was used. Following is a brief description of Python language, Anaconda distribution, Jupyter notebook, R language, and R-Studio (an R language IDE).

3.3.1 Python-Language:

Python is a high-level, dynamically typed programming language renowned for its simplicity and readability. Created by Guido van Rossum in the late 1980s, Python emphasizes code clarity and uses indentation to define code blocks, which enhances human readability. Its versatile nature enables it to serve as a general-purpose language, suitable for web development, scientific computing, data analysis, artificial intelligence, machine learning, automation, and more. Python's extensive standard library provides a wealth of modules and functions for common tasks, contributing to faster development [159]. The language's object-oriented, imperative, and functional programming paradigms accommodate a range of coding styles. Python's popularity is fueled by an active community, frequent updates, and a plethora of third-party packages accessible via the Python Package Index (PyPI). Its cross-platform compatibility and ease of integration with other languages make it a preferred choice for both beginners and experienced developers.

3.3.2 Anaconda distribution:

Anaconda is an open-source Python distribution for data science and machine learning, offering a variety of pre-installed libraries and tools for data analysis, visualization, and scientific computing workflows. Developed by Anaconda, Inc., it encompasses a curated collection of tools, libraries, and packages that facilitate the development and deployment of data-intensive

applications [160]. Anaconda Distribution includes the Python programming language along with a multitude of libraries and tools for tasks like data manipulation, analysis, visualization, and machine learning. One of its key features is the Anaconda Navigator, a graphical user interface that aids in package management, environment creation, and launching applications. Anaconda also offers the conda package manager, which allows users to create isolated environments to avoid version conflicts among different packages. This distribution is widely used by data scientists, researchers, and developers to streamline the process of setting up development environments and to ensure consistent and reproducible results across various projects.

3.3.3 Jupyter notebook:

Jupyter Notebook is an interactive web-based environment that allows users to create and share documents containing live code, equations, visualizations, and narrative text. Jupyter Notebook is a powerful tool for creating interactive documents that combine code execution, visualizations, and explanations [161]. It supports various programming languages but is primarily used with Python. Key features of Jupyter Notebook include interactive execution, rich text support, data visualization, and easy sharing. When working with Anaconda, users can launch Jupyter Notebook from the Anaconda Navigator or directly from the command line using the jupyter notebook command. This allows users to create, open, and edit notebooks in a user-friendly environment. Overall, Jupyter Notebook offer a powerful platform for data scientists and analysts to work with Python, manage environments, and create interactive documents that showcase their analysis and findings.

3.3.4 R-Language:

R is a programming language and software environment that is widely used for statistical computing and graphics. It provides a wide variety of statistical and graphical techniques, including linear and nonlinear modeling, time-series analysis, classification, clustering, and more. R is open-source, free software and is available for Windows, Mac OS X, and Linux operating

systems. It has a large and active community of users and developers, and there are many resources available online for learning and using the language. R is used by statisticians, data scientists, and researchers in many different fields, including social sciences, finance, healthcare, and more. It is also popular in the field of data analysis and visualization, as it offers many tools for working with data sets and creating visualizations [162]. In summary, R is a powerful and versatile language with many applications in statistical computing, data analysis, and visualization, and it is widely used by professionals and researchers in many different fields.

3.3.5 R-Studio:

R Studio is a powerful and user-friendly integrated development environment (IDE) for the R programming language. It provides a comprehensive set of tools and features that make it easy for users to manage and analyze data, write code, and create visualizations. One of the key benefits of R Studio is its ability to streamline the development process for R code. It includes an intuitive code editor with features like syntax highlighting, code completion, and error highlighting to help users write code more efficiently and with fewer errors. It also includes tools for managing data, including importing and exporting data from a variety of file formats, and cleaning and transforming data using R's built-in functions. In addition to its data management and code development features, R Studio also offers a range of visualization tools, including plots, charts, and graphs. These tools allow users to create high-quality visualizations of their data and communicate their findings effectively. Another advantage of R Studio is its support for the development of R packages. R packages are collections of R code and functions that can be easily shared and reused by other users. R Studio provides tools for building, testing, and publishing R packages, which makes it easier for users to contribute to the R community and collaborate with others [163]. Overall, R Studio is a powerful and versatile IDE that is widely used by data scientists, statisticians, and researchers in many different fields. Its intuitive interface and comprehensive set of tools make it an essential tool for managing and analyzing data, writing code, and creating visualizations in R.

3.4 Data-preprocessing:

Data preprocessing is a crucial step in preparing data for training and developing machine learning models. It involves cleaning, transforming, and organizing raw data into a suitable format that can be effectively used by machine learning algorithms. The goal of data preprocessing is to enhance the quality and reliability of the data, thereby improving the performance of the trained models. One of the steps involved in data preprocessing is Data Cleaning. Identify and handle missing data, either by inputting with statistical measures or removing the instances with missing values. Data Transformation is another step of preprocessing. It includes feature scaling to ensure that all the features are on a similar scale, preventing one feature from dominating others. Common scaling methods include Min-Max Scaling and Standardization [164]. Converting categorical variables into numerical format, such as one-hot encoding for nominal variables and label encoding for ordinal variables, is important for model training. Splitting the preprocessed data into training, validation, and test sets is also important. The training set is used for model training, the validation set for hyperparameter tuning, and the test set for evaluating the final model's performance. Normalization of data is necessary if needed, such as when working with neural networks. Normalization ensures that the data falls within a specific range, often [0, 1].

3.5 Targets-Diseases Network Construction:

In the “STRING disease query” tab different keywords of each neuronal disorder were searched to find out the disease associated proteins. The names of 10 most frequent Neurodegenerative disorders were used as keywords to find out the disease associated proteins. The most common and lethal neurodegenerative disorders include Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, Amyotrophic lateral sclerosis, Dementia, Prion disease, Frontotemporal Lobar Degeneration, Multiple Sclerosis, Progressive Supranuclear Palsy, and Down Syndrome. The most prevailing neurological disorders include Migraine, Psychotic disorder, Obsessive-compulsive disorder, Epilepsy, Autism Spectrum Disorder, Williams-Beuren syndrome, Anxiety disorder, Amyotrophic lateral sclerosis, and Major Depressive Disorder. Each keyword resulted in

a list of related diseases, the network of all these diseases were selected and merged to form huge networks against each keyword. The networks for all the neurological disorders were merged to form a final network of proteins-associated with neurological disorders containing 171 nodes while the merged network of neurodegenerative disorders has 65 proteins that have role in progression of the disorders. After construction of the Targets-Diseases Network, the data is saved in a csv file.

3.6 Diseases-Drugs Network Construction:

By using the collected data of all the FDA-approved drugs associated with specific protein-targets and for all the selected disorders using different online available databases, we constructed the Diseases-Drugs network. For the ten most common and lethal neurodegenerative disorders and nine neurological disorders, the data of drugs collected for each disorder separately to form nineteen Diseases-Drugs Networks. The networks for all the neurological disorders were merged to form a final network of drugs associated with neurological disorders containing 3440 entries while the merged network of neurodegenerative disorders has total 1443 entries that have role for the specific target-proteins in these disorders. The data of the final Diseases-Drugs Network saved in an csv file.

3.7 Targets-Diseases-Drugs Network Construction:

We constructed Targets-Diseases-Drugs Network, which is a comprehensive network that combines diverse data types to understand the intricate relationships between neurological and neurodegenerative disorders and their potential therapeutic interventions, by combining the two networks: the Targets-Diseases Network and the Diseases-Drugs Network. The Targets-Diseases Network identifies protein targets and specific neuronal disorders, represented by 9921 descriptors that quantitatively represent their attributes and properties. The Diseases-Drugs Network highlights the interconnections between neurological disorders and potential drugs that hold promise for their treatment, represented by 2048 Morgan fingerprints encoded in binary format. The network provides an encompassing perspective, offering insights into the triadic relationship

among protein targets, disorders, and the drugs. Each facet of the network is represented using distinct techniques, with protein targets characterized by descriptors and drugs represented by binary-encoded Morgan fingerprints. Neurological disorders comprise nineteen distinct entities, each encoded using one-hot encoding. The class label of the dataset of Targets-Diseases-Drugs Network divided the instances into neurodegenerative disorders (0) and neurological disorders (1) for binary classification.

3.8 Compiling Data for Training Targets-Diseases Network Model:

Different classification models were trained to predict the disease association of the protein with respect to neurological and neurodegenerative disorders. Out of 4884 target proteins, 3440 came out to be associated with neurological disorders. While 1443 came out to be associated with neurodegenerative disorders. The models were built on the protein descriptors dataset used as the X-matrix while the class label was assigned based on the specific disorder type. All the protein targets associated with neurodegenerative diseases were labeled as 0 and all the protein targets associated with neurological diseases were labeled as 1. There is a difference of 1997 instances between both classes which clearly demonstrates the class-imbalance between two target classes. To prevent the model performance from deterioration, we perform hyperparameters tunning and its special feature i.e., GridSearchCV [165], short for Grid Search Cross-Validation, which is a hyperparameter tuning technique widely used in machine learning. It helps automate the process of finding the best combination of hyperparameters for a given machine learning model by exhaustively searching through a predefined set of hyperparameter values. The data matrix was of the order 4884×9921. The class labels, total no. of descriptors for each target protein, and total number of instances belonging to each class in the dataset of Targets-Diseases Network model are presented in Table 1.

Table 1 *Summary of Targets-Diseases network classification dataset*

Class Labels	Total no. of Instances	Total no. of descriptors
1	3440	9921

(Protein targets associated with neurological disorders)		
0 (Protein targets associated with neurodegenerative disorders)	1443	9921

3.9 Compiling Data for Training Diseases-Drugs Network Model:

There are different classification models which were built for the prediction of the drugs against targets- protein associated with neurodegenerative disorders and with neurological disorders. The labels were given to the drugs based on the activation/inhibition data collected from different databases. In the Diseases-drugs network dataset, 3440 of the 4884 drugs were active against targets associated with neurological disorders while the other 1443 were active against targets associated with neurodegenerative disorders. Classification of the drugs for activity against neurodegenerative disorders and with neurological disorders was done by taking morgan fingerprints of the inhibitors as X-matrix. The class labels were given as 1 for neurological disorders and zero for neurodegenerative disorders. To assign the class labels to the drugs, we took help from different databases. To compile a complete network dataset, total 4884 drugs as instances were taken in each class active against neurological and neurodegenerative disorders having 2048 features in X-matrix (4884×2048). Table 2 shows the class labels, no. of Morgan's Fingerprints for each drug, and total number of instances belonging to each class in the dataset of Diseases-Drugs Network model.

Table 2 *Summary of Diseases-Drugs network classification dataset*

Class Labels	Total no. of Instances	No. of Morgan's Fingerprints
1	3440	2048

(Drugs associated with neurological disorders)		
0 (Drugs associated with neurodegenerative disorders)	1443	2048

3.10 Compiling Data for Training Targets-Diseases-Drugs Network Model:

Different classification models were trained on the combined Targets-Diseases-Drugs Network to predict the disease association of the protein targets as well as their drugs with respect to neurological and neurodegenerative disorders. Out of total 4884 number of instances in the network, 3440 target proteins and drugs came out to be associated with neurological disorders. While 1443 target proteins and drugs came out to be associated with neurodegenerative disorders. The models were built on the complete network dataset having 9921 protein descriptors, 2048 morgan's fingerprints of all the drugs, one hot encoding of the nineteen neuronal disorders, drugs activity and no. of protein edges were used as the X-matrix of the dataset while the class label was assigned based on the specific disorder type. All the protein targets, drugs and disorders associated with neurodegenerative diseases were labeled as 0 and all the protein targets, drugs and disorders associated with neurological diseases were labeled as 1. There is a difference of 1997 instances between both classes which clearly demonstrates the class-imbalance between two output classes. To prevent the model performance from deterioration, we perform hyperparameters tunning and its special feature i.e., GridSearchCV, short for Grid Search Cross-Validation, which is a hyperparameter tuning technique widely used in machine learning. It helps automate the process of finding the best combination of hyperparameters for a given machine learning model by exhaustively searching through a predefined set of hyperparameter values. The data matrix was of the order 4884×11990. To compile a complete network dataset, total 4884 drugs as instances were taken in each class active against neurological and neurodegenerative disorders having 2048 features in X-matrix (4884×11990). Table 3 shows the class labels, total no. of feature variables, and total number of instances belonging to each class in the complete dataset of Targets-Diseases-Drugs Network model.

Table 3 *Total no. of instances & Features for Target-Disease-Drug Network*

Class Labels	Total no. of Instances	Total no. of Features
1 (Neurological disorders)	3440	11990
0 (Neurodegenerative disorders)	1443	11990

3.11 Multi-variate Artificial Neural network Construction for Drug Repurposing:

We have constructed three networks i.e., targets-disease network, diseases-drugs network and targets-disease-drugs association network and then train these networks by using different machine learning and deep learning models. Now we designed a multi-variate Artificial Neural network to predict the interactions between drugs and target proteins linked to specific diseases. We have utilized two datasets: one encapsulating the Morgan's fingerprints of 4884 drugs as input dataset, and the other containing descriptors of 4884 target proteins along with disease class labels (0 or 1) as output dataset. The distinctiveness of our approach lies in the meticulous correspondence between the rows of these datasets, where each entry in the input dataset (housing drug fingerprints) aligns with the corresponding entry in the output dataset (comprising target protein descriptors and disease class labels). We devised a model capable of comprehending the intricate connections between drug fingerprints, target protein descriptors, and disease classes, ultimately enabling predictions for both target protein descriptors and disease class for a given drug.

In addressing this challenge, we adopted an Artificial Neural Network (ANN) [166] strategy, employing a multi-variate output configuration. We designed our methodology to entail training the ANN to decipher the intricate relationships between drug fingerprints and target protein descriptors, also incorporating information about disease classes. Our model's architecture is meticulously crafted to accept Morgan’s fingerprints of an unidentified drug, which undergo transformation using the RDkit library to form input features for the ANN. Upon processing, our model yielded two outcomes: firstly, the descriptors of the target protein corresponding to the drug's fingerprints, and secondly, the anticipated disease class associated with the drug. The heart

of our approach lies in the concept of a Multi-Output Neural Network. In this framework, our ANN is structured to simultaneously predict multiple outputs, specifically the target protein descriptors and the disease class. Each of these outputs corresponds to a distinct facet of the problem. The neural network is trained on the input drug fingerprints, effectively generating predictions for both descriptors and disease class in a synchronized manner. This leverages the inherent relationships embedded within the aligned datasets.

The alignment of datasets at a granular level ensures that the model comprehends the intricate interplay between drug attributes, target protein properties, and disease classifications. The ANN's structure is systematically optimized via hyperparameter tuning, encompassing the configuration of hidden layers, activation functions, learning rates, and regularization techniques. The model's outputs cater to predicting both target protein descriptors and disease class, presenting a holistic solution for drug-target interaction prediction. After training of our datasets through the utilization of multi-variate ANN i.e., Multi-Output Neural Network. We validate our model by giving the SMILES of an unknown drug. Our innovative architecture predicted the target protein descriptors and disease class simultaneously for the given drug.

3.12 Machine Learning Models Construction & Hyperparameter Optimization:

The quality and acceptability of the machine learning model depends on the data set on which they were trained. The data collected previously was prepared for the purpose of training the models i.e., assigning class labels, treating class imbalance etc. The classification algorithms chosen for the study were Linear and Non-Linear Support Vector Machine (SVM) classifier, Decision Tree classifier, Random Forest classifier, Gradient Boosting Machines (GBM), Multi-Layer Perceptron Neural Network (MLP-NN) and multi-variate Artificial Neural Network (ANN). These models are trained on the three types of networks that we constructed for the dataset’s protein-targets, drugs and neuronal disorders.

3.12.1 Support Vector Machine (SVM) classifier:

Support Vector Machine (SVM) is a widely used supervised machine learning algorithm for classification and regression tasks, particularly effective for complex decision boundaries and high-dimensional data [167]. It is often used for binary classification, where it finds a hyperplane that maximizes the margin between data points of different classes. For linearly separable data, SVM aims to find the hyperplane that can perfectly separate the two classes. The optimal hyperplane is the one that maximizes the distance (margin) between the support vectors of each class. In real-world scenarios, SVM introduces a "soft margin" that allows some misclassification of data points. This is achieved by introducing a regularization parameter (C) that balances the trade-off between maximizing the margin and minimizing misclassification. SVM can handle not only linearly separable data but also nonlinearly separable data in the original feature space. The kernel trick is used to transform the original data into a higher-dimensional space where it might become linearly separable. In the transformed higher-dimensional space, SVM finds a hyperplane that separates the data classes, corresponding to a nonlinear decision boundary when projecting back to the original space. In classification, the SVM algorithm assigns new data points to classes based on their position relative to the learned hyperplane. The choice of hyperparameters like C and the choice of kernel function significantly impacts the model's performance and generalization. SVM is effective in high-dimensional spaces, handles complex decision boundaries, and is less prone to overfitting when properly tuned. Hyperparameter tuning is crucial for optimizing SVM performance, and GridSearchCV is a valuable technique for finding the best combination of hyperparameters. Key hyperparameters include kernel type, regularization parameter, and kernel-specific parameters. GridSearchCV automates the process by searching through a predefined grid of hyperparameter values and systematically trains and evaluates SVM models on various combinations using cross-validation. The training data is split into subsets for training and validation, and the model's performance is evaluated based on a chosen metric. After all iterations, GridSearchCV provides the best cross-validated hyperparameter combination, which is then used to train a final SVM model on the entire training dataset. The model's performance is then assessed on a separate test dataset to evaluate its generalization ability. This automated approach ensures systematic exploration of parameter configurations and optimizes predictive accuracy, enhancing the overall effectiveness of SVM for classification tasks.

3.12.2 Decision Tree classifier:

A Decision Tree classifier is a machine learning algorithm used for classification and regression tasks. It works by partitioning input data into subsets based on different input features, each of which corresponds to a decision node in the tree. The top node, called the root node, represents the entire dataset and is divided into subsets based on one of the input features' values. Each decision node corresponds to a particular feature and a specific threshold value, representing a decision point where data is split into different branches based on whether the feature value is above or below the threshold [168]. The final decision or classification is represented by leaf nodes, which correspond to predicted class labels or regression values. The algorithm selects the best feature and threshold to split the data at each decision node, determined by criteria such as Gini impurity or Mean Squared Error. The process of partitioning data and creating decision nodes is repeated recursively for each subset, and the tree grows deeper as it splits the data into more subsets. Decision Trees are easy to understand and interpret due to their visual representation resembling human decision-making. They can handle both categorical and numerical features and are not sensitive to feature scaling. However, they can be prone to overfitting, especially when the tree is deep and complex, and may not generalize well to unseen data if not pruned properly. Ensemble methods like Random Forest and Gradient Boosting use multiple Decision Trees to make predictions and combine their outputs. Hyperparameter tuning is a crucial aspect of optimizing the performance of a Decision Tree Classifier. The GridSearchCV technique can be used to find the best combination of hyperparameters, such as the maximum depth of the tree, minimum number of samples required to split an internal node, minimum number of samples at a leaf node, the quality of a split, and the maximum number of features considered at each split. By defining a set of possible values for each hyperparameter, the GridSearchCV process involves training and evaluating Decision Tree models using all possible combinations of hyperparameters through cross-validation. The model's performance is assessed using a chosen evaluation metric, and the optimal set of hyperparameters is used to train a final Decision Tree Classifier on the entire training dataset. This automated approach simplifies the process of finding the most suitable hyperparameters and enhances the overall effectiveness of the Decision Tree algorithm for classification tasks.

3.12.3 Random Forest classifier:

The Random Forest classifier is an ensemble machine learning algorithm that combines multiple Decision Trees to improve predictive accuracy and control overfitting. It works by constructing a "forest" of decision trees and aggregating their predictions to make a final classification decision. The algorithm uses bootstrap sampling to randomly select subsets of the original training data, creating slightly different datasets for each tree. Each tree is grown through recursive binary splitting, and the algorithm chooses the best feature and threshold based on the random subset of features considered. After all trees are constructed, new data is fed into the forest, and each tree makes an individual prediction based on its structure [169]. For classification tasks, the class predicted by each tree is considered a "vote," and the final prediction is determined by the majority vote from all the trees. In a classification problem, the class with the most votes becomes the final prediction, while in a regression problem, the final prediction can be the mean or median of the predicted values from all the trees. Random Forests are robust and can handle both categorical and numerical features, making them suitable for a wide range of tasks, including classification and regression. They reduce the variance of the model by averaging out individual errors and random variations in each tree's predictions. They can estimate the generalization error using out-of-bag samples, providing an internal validation measure during training. However, Random Forests may not be as interpretable as individual Decision Trees and could become computationally expensive for large datasets or trees. Parameters such as the number of trees, maximum depth of trees, and the number of features considered at each node can be tuned to optimize the model's performance. Hyperparameter tuning is crucial for optimizing the performance of a Random Forest Classifier. The GridSearchCV technique can be used to explore different parameter combinations, including the number of trees, maximum depth, minimum samples required to split an internal node, minimum samples required to be at a leaf node, and the number of features considered at each split. The approach defines a range of values for each hyperparameter and performs an exhaustive search by training and evaluating Random Forest models with all possible combinations of hyperparameters using cross-validation. The optimal combination of hyperparameters is then used to train a final Random Forest Classifier on the entire training dataset. The model's performance is evaluated on a separate test dataset to gauge its ability to generalize to new data. This automated

approach streamlines the process of finding the most suitable hyperparameters and enhances the overall effectiveness of the Random Forest algorithm for classification tasks.

3.12.4 Gradient Boosting Machines (GBM):

Gradient Boosting Machine (GBM) is an ensemble machine learning algorithm that combines multiple weak learners, typically Decision Trees, to create a strong predictive model. The process starts with creating an initial prediction based on a simple model, such as a single Decision Tree with shallow depth or a constant value. The goal is to iteratively reduce residuals or errors by adding new models. In each iteration, a new weak learner is added to the ensemble, trained to predict the negative gradient of the loss function with respect to the current predictions. The new model's predictions are weighted based on their contribution to reducing errors. The learning rate parameter controls the step size of updates, preventing overfitting by regularizing the process. The ensemble's predictions are updated accordingly, updating the ensemble toward the correct values and reducing residuals. The boosting process continues for a predefined number of iterations or until a stopping criterion is met [170]. The final prediction of the GBM ensemble is the sum of predictions from all individual models in the ensemble, which is often more accurate than that of any single model. GBM is known for its strong predictive power and robustness, handling different types of data and automatically performing feature selection. However, it can be computationally expensive and requires careful tuning of hyperparameters to avoid overfitting. Variants of Gradient Boosting, such as XGBoost and LightGBM, have been developed to optimize performance and speed up training. Hyperparameter tuning is crucial for optimizing GBM performance, and the GridSearchCV technique can automate this process. Key hyperparameters include the number of boosting stages, learning rate, maximum depth of individual trees, and number of features considered at each split. Parameters related to subsampling and regularization can also be tuned to prevent overfitting. To perform hyperparameter tuning using GridSearchCV, a range of values for each hyperparameter is defined, and the algorithm explores all possible combinations through cross-validation. The best combination of hyperparameters is returned, which is then used to train a final GBM model on the entire training dataset. The model's performance is then evaluated on a separate test dataset to assess its generalization ability. This approach automates the search for

suitable parameter values, saving time and reducing the risk of selecting suboptimal configurations.

3.12.5 Multi-Layer Perceptron Neural Network (MLP-NN):

A Multi-Layer Perceptron (MLP) is a feedforward artificial neural network consisting of multiple layers of interconnected neurons designed to process and learn complex patterns in data. It is a foundational architecture in deep learning and is used for tasks such as classification, regression, and feature extraction. MLPs consist of three main types: an input layer, one or more hidden layers, and an output layer. Each layer contains neurons that transform input data using weights and biases [171]. The data flows through the network in a feedforward manner, starting from the input layer and passing through the hidden layers to produce an output in the output layer. Neurons in each layer are connected to neurons in adjacent layers through weighted connections. Common activation functions include ReLU, sigmoid, and tanh, which introduce non-linearity into the network, allowing it to learn complex relationships in data. Each neuron takes the weighted sum of its inputs, multiplied by a weight associated with the connection, and adds a bias term before passing it through the activation function. MLPs are trained using a process called backpropagation, where the network compares its predictions (output) to the actual target values and calculates the error. Optimization algorithms like Gradient Descent are used to update the weights in the direction that minimizes the error. The difference between the predicted output and the true target is quantified using a loss function, with the goal of training to minimize this loss function. MLPs have several hyperparameters that need to be tuned for optimal performance, including the number of hidden layers, the number of neurons in each layer, the choice of activation functions, learning rate, batch size, and regularization strength. They can capture complex non-linear relationships in data and are versatile for various tasks. Hyperparameter tuning is crucial for optimizing MLP performance. Identify the hyperparameters that influence the behavior of your MLP. These might include the number of hidden layers, the number of neurons in each hidden layer, activation functions, learning rate, batch size, optimizer, dropout rate, weight decay, and more. Create a parameter grid that includes different values for the hyperparameters you want to tune. GridSearchCV can automate this process by identifying hyperparameters influencing MLP

behavior, creating a parameter grid with different values, and deciding on the MLP architecture. A function is built to create a MLP model based on these parameters, and GridSearchCV is imported from scikit-learn. The function is fitted on training data, and the best hyperparameters are identified using the best_params_ attribute. A new MLP model is created using these parameters, trained on the full training dataset, and evaluated on a separate test dataset. This process avoids manual trial and error, ensuring the best set of hyperparameters for a specific task. By using GridSearchCV for hyperparameter tuning, the process is automated, avoiding manual trial and error and ensuring the best hyperparameters for specific tasks.

3.12.6 Multi-variate Artificial Neural Network (ANN):

An Artificial Neural Network (ANN) with multi-variate output features, often referred to as a Multi-Output Neural Network, is a type of neural network architecture that can make predictions or classifications involving multiple output variables or features simultaneously. It works by combining input, hidden, and output layers, with neurons in the output layer corresponding to each output feature. The network can include one or more hidden layers between the input and output layers, which enables the network to learn complex relationships within the input data. Activation functions are applied to neurons in the hidden layers and the output layer, introducing non-linearity and helping capture intricate patterns in the data [172]. The output layer contains neurons that correspond to each desired output feature, and the loss function should consider the differences between predicted and actual values for each output feature. Common loss functions for multi-output tasks include Mean Squared Error (MSE) for regression and categorical cross-entropy for classification. During training, the network computes the error between predicted and actual output values using the chosen loss function. Backpropagation is used to propagate the error backward through the network, updating weights and biases to minimize loss. Hyperparameters such as the number of hidden layers, neurons in each layer, activation functions, learning rate, and regularization should be tuned to optimize the network's performance. Multi-Output Neural Networks are advantageous when dealing with tasks involving multiple output variables, as they can capture dependencies and correlations among these variables. They find applications in various domains such as multi-label classification, multi-variate time series prediction, and image

segmentation with multiple classes. Designing a proper architecture for a Multi-Output Neural Network can be complex, especially when dealing with varying ranges and scales of output features. Hyperparameter tuning is a crucial step in optimizing the performance of a Multi-Output Neural Network. It involves determining the number of hidden layers and neurons each layer should contain, choosing appropriate activation functions, adjusting the learning rate, batch size, and choosing an optimizer algorithm. Regularization techniques like L1 or L2 regularization can help prevent overfitting by adding penalty terms to the loss function based on the magnitudes of weights. Adjusting the learning rate during training helps fine-tune weights as the training progresses. Determining the number of epochs to perform and implementing early stopping by monitoring a validation metric can prevent overfitting. Experimenting with different architectures, such as grid search, random search, or Bayesian optimization, can also be helpful. Using automated hyperparameter tuning libraries like scikit-learn's GridSearchCV or RandomizedSearchCV, as well as specialized libraries like Keras Tuner or Optuna, can help systematically explore the hyperparameter space and find the optimal configuration for our model. By optimizing hyperparameters, we can enhance the performance and generalization of our model for our desired task involving multiple output variables.

3.13 Evaluation Methods

3.13.1 Confusion Matrices:

A confusion matrix is commonly used to evaluate the performance of machine learning or any classification model. With the use of the confusion matrix, results may get a geed sense of whether findings are high performing or not, as it provides a tabular representation of the predicted and actual class labels. Typically, a confusion matrix consists of four cells true positives, true negatives, false positives, and false negatives. In this proposed study, targets which are correctly classified as for neurological disorders (in our case, true positives; TP), targets which are correctly classified as for neurodegenerative disorders (in our case, true negatives; TN), targets which are incorrectly classified as for neurological disorders, when the actual class label is for neurodegenerative disorders (in our case, false positives; FP), and targets which are incorrectly classified as for

neurodegenerative disorders, when the actual class label is for neurological disorders (in our case, false negatives; FN), are the elements of the confusion matrix. Predictions that turn out to be false negatives and false positives are the wrong predictions by the model. Because the proposed study presented the classification of neurodegenerative disorders Vs neurological disorders which is a binary classification with two outcomes, we obtained a 2 X 2 matrix. To ensure the generalizability of the models 10-fold Cross validation was performed. Subsequentially, the models were also used for the classification of unknown datasets. By analysing the values of the confusion matrix, various performance parameters such as precision, recall and F1 score etc. were calculated to determine the effectiveness of the model in classifying neurodegenerative disorders and neurological disorders.

3.13.2 Accuracy:

One indicator for assessing model performance is accuracy. The accuracy of the classifier is its ability to correctly predict the class labels of instances of different classes (positive and negative class). True positive (TP) are the instances of positive class that were correctly predicted by the classifier. Likewise, the correctly predicted instances of negative class were termed as true negative (TN). False positives (FP) and false negatives (FN) represent the fallacy of the classifier to make correct prediction. FP is an outcome given by the classifier when it inaccurately predicts the positive class while FN occurred due to misclassification of instances of positive class. These four measures were used by all the performance metrices given below. Mathematically, accuracy is defined as.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$$

3.13.3 Precision:

Precision is determined as the ratio of the total number of true positives to the total number of instances predicted as positive.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

3.13.4 Recall:

It is also known as sensitivity. Recall determines the proportion of correctly predicted positive instances out of all actual positive instances. It is calculated by dividing the total number of true positives by the sum of true positives and false negatives.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

3.13.5 F1 Score:

The F1 score is the harmonic mean of precision and recall scores. A higher F1 score indicates a better-quality classifier.

$$\text{F1 Score} = 2 * \frac{\text{Percision} * \text{Recall}}{\text{Percision} + \text{Recall}}$$

The model having all these measures satisfied would be considered as the best model as the results predicted by the model would be reliable enough and then it be used as a generalized model for the prediction on unknown dataset.

Chapter 4

Results and Discussion

4 Results and Discussion

The goal of this study was to collect the data of all the FDA-approved drugs and target proteins associated with the nine neurological and ten neurodegenerative disorders from different online available databases for the development of a comprehensive database. Using the developed database, the construction and integration of three networks i.e., target-disease, disease-drug and target-disease-drug association networks was performed. After networks construction, various machine learning and deep learning models was developed and trained for the classification of the networks with respect to neurological vs neurodegenerative disorders. The ultimate goal of our proposed study was drug repurposing of neurological disorders by making a connection of drug fingerprints with the target descriptors and associated disease. To achieve these goals, a couple of online available target proteins and drugs databases are explored for the data collection and different tools are used for descriptors computation, network construction, Machine Learning/Deep Learning model building and development by undergoing series of different steps. In this chapter, the results of each step of our methodology are presented and discussed to infer from the analysis. In general, our models classify the target proteins and drugs for neurological and neurodegenerative disorders separately. And the main achievement of our proposed work was the prediction of an unknown drug based on their fingerprints for the disease type and specific target descriptors, which is a significant success in the field of drug discovery and precision medicine. Furthermore, several targets and drugs overlap are uncovered between various neuronal disorders. Overall, the results provide a comprehensive view of the drugs and protein targets associated with a specific neuronal disorder type. The results of each step are described in detail in the following paragraphs:

4.1 Data Collection:

4.1.1 Selection of Diseases:

Neurological and Neurodegenerative disorders are selected due to their prevalence and mortality rate worldwide, with the help of literature and different organizations such as World Health

organization (WHO), Global Burden of Diseases (GBD), etc. Nine neurological disorders and ten neurodegenerative disorders were selected, which are shown in Table 4.

Table 4 *List of all selected Neurological & Neurodegenerative disorders*

Serial no.	Neuronal Disorders	Disorder Class
1.	Alzheimer's Disease	Neurodegenerative disorder
2.	Huntington's disease	Neurodegenerative disorder
3.	Prion disease	Neurodegenerative disorder
4.	Down Syndrome	Neurodegenerative disorder
5.	Progressive Supranuclear Palsy	Neurodegenerative disorder
6.	Major Depressive Disorder	Neurodegenerative disorder
7.	Multiple Sclerosis	Neurodegenerative disorder
8.	Parkinson's disease	Neurodegenerative disorder
9.	Dementia	Neurodegenerative disorder
10.	Frontotemporal Lobar Degeneration	Neurodegenerative disorder
11.	Migraine	Neurological disorder
12.	Psychotic disorder	Neurological disorder
13.	Prader-Willi Syndrome	Neurological disorder
14.	Obsessive-compulsive disorder	Neurological disorder
15.	Epilepsy	Neurological disorder
16.	Autism Spectrum Disorder	Neurological disorder
17.	Anxiety disorder	Neurological disorder
18.	Williams-Beuren syndrome	Neurological disorder
19.	Amyotrophic lateral sclerosis	Neurological disorder

4.1.2 Target Data Retrieval:

In the first step of the methodology, a publicly available protein-targets dataset from the two databases are collected for the nine neurological and ten neurodegenerative disorders one by one and then merged all the datasets of all the selected neuronal disorders. The total data of the protein-targets retrieved from STRING database, after merging the targets datasets for each disorder and

the total data of the protein-targets retrieved from Therapeutic Target Database (TTD), after merging the targets datasets of each disorder, by giving the query of each Neurological and Neurodegenerative disorder individually and then merged the dataset of both databases to remove target duplicates. After removing duplicates, the data contained 821 unique targets protein between the protein-protein interaction edges zero to seventy-one. We shortlisted the targets datasets having edges of at least three to seventy-one. The total unique data of target proteins for all the nineteen neuronal disorders remained 236, which is shown in table 5.

Table 5 *List of all the selected Protein-Targets for Neuronal disorders*

Sr. no.	Targets Name	No. of Edges	Serial. no.	No. of Edges	Targets Name
1	MT-ND2	3	**120**	10	CST3
2	HSPG2	3	**121**	10	IL7R
3	ASTN2	3	**122**	10	HLA-DRB1
4	PRRT2	3	**123**	10	PVALB
5	NPAS3	3	**124**	10	GABRB3
6	NDN	3	**125**	10	GABRA5
7	SLC6A4	3	**126**	10	HTR1A
8	SNCA	4	**127**	10	MAOA
9	HTT	4	**128**	10	HTR2A
10	BDNF	4	**129**	11	NALCN
11	APOA1	4	**130**	11	GABRE
12	C1QB	4	**131**	11	VARS
13	CAD	4	**132**	11	KCNS3
14	MARS2	4	**133**	11	SYNJ1
15	KCNMB3	4	**134**	11	SLC2A1
16	PROSC	4	**135**	11	STAT3
17	EPM2A	4	**136**	11	MOG
18	ITPA	4	**137**	11	RELN
19	POMC	4	**138**	11	COMT
20	SLC35A2	4	**139**	11	OXT
21	CASR	4	**140**	12	KCNQ5
22	APOE	4	**141**	12	ADAM10
23	APP	4	**142**	12	IL17A

24	GRN	4	**143**	12	CD40
25	TARDBP	4	**144**	12	IL2RA
26	MAPT	4	**145**	12	IFNB1
27	CTNND2	4	**146**	12	MLXIPL
28	CACNA1H	4	**147**	12	GRM3
29	NCF1	4	**148**	13	ATXN2
30	CACNA1A	4	**149**	13	ARX
31	OCA2	4	**150**	13	PICALM
32	SLC1A1	4	**151**	13	CD80
33	APCS	5	**152**	13	CD86
34	C1QC	5	**153**	13	DYRK1A
35	EEF1A2	5	**154**	14	TH
36	YWHAG	5	**155**	14	ATP6V1A
37	MEF2C	5	**156**	14	KCNT2
38	MT-ND1	5	**157**	14	A2M
39	PARS2	5	**158**	14	CD4
40	TNRC6A	5	**159**	14	NRXN1
41	GSK3A	5	**160**	15	PFN1
42	CACNA1C	5	**161**	15	PNKP
43	HLA-DRB5	5	**162**	15	CACNA2D2
44	SCN1A	5	**163**	15	IDE
45	TRPM8	5	**164**	15	CTSB
46	ATP1A2	5	**165**	15	STX1A
47	OLIG2	5	**166**	15	GRIN2A
48	GABRR1	5	**167**	16	KCNH1
49	AKT1	5	**168**	16	MTOR
50	GABRR3	5	**169**	16	MECP2
51	UBE3A	5	**170**	16	LIMK1
52	DRD4	5	**171**	16	DISC1
53	MAP1B	6	**172**	17	TBK1
54	ATP1A3	6	**173**	17	IL6
55	C1QA	6	**174**	17	SYNGAP1
56	CUX2	6	**175**	17	NSUN5

57	AARS	6	**176**	18	DNAJC6
58	POLG	6	**177**	18	CACNA1E
59	TPP1	6	**178**	18	CNTNAP2
60	GLS	6	**179**	18	FMR1
61	STK39	6	**180**	18	BAZ1B
62	MAPK10	6	**181**	19	TUBA4A
63	DAB1	6	**182**	19	SIGMAR1
64	KCNU1	6	**183**	19	CHRNA2
65	ALDH7A1	6	**184**	20	NEK1
66	MDH1	6	**185**	20	GABRA2
67	SLC6A3	6	**186**	20	CLU
68	TOMM40	6	**187**	20	CHD8
69	OXTR	6	**188**	21	SPG11
70	WBSCR22	6	**189**	21	GRIK1
71	KCNK18	6	**190**	21	SCN2A
72	YWHAZ	7	**191**	21	GRIN2B
73	C3	7	**192**	22	NEFH
74	ATIC	7	**193**	22	SCN3A
75	CSTB	7	**194**	22	KCNB1
76	FUS	7	**195**	23	PLA2G6
77	VCP	7	**196**	23	CACNA1D
78	GRM5	7	**197**	23	SLC25A22
79	DRD2	7	**198**	23	PLCB1
80	CRHR1	7	**199**	23	DNM1
81	KATNAL2	7	**200**	25	MATR3
82	ADNP	7	**201**	25	KCNJ10
83	ELN	7	**202**	25	CRH
84	GABRA4	7	**203**	26	PINK1
85	TAAR6	7	**204**	26	SOD1
86	PIGB	8	**205**	26	PNPO
87	SCARB2	8	**206**	26	NTRK2
88	TSC1	8	**207**	26	SCN9A
89	HCN1	8	**208**	27	DCTN1

90	MDH2	8	**209**	27	OPTN
91	ATP2A2	8	**210**	27	GABRB2
92	ATP1A1	8	**211**	28	GABBR2
93	GIG25	8	**212**	28	SLC1A2
94	GBA	8	**213**	28	CNTN2
95	MBP	8	**214**	29	CLCN2
96	DSCAM	8	**215**	29	CHRNA4
97	NOS1AP	8	**216**	30	STX1B
98	GAD2	8	**217**	30	LGI1
99	FKBP5	8	**218**	31	ATP13A2
100	SLC18A2	9	**219**	31	PARK7
101	PRNP	9	**220**	32	KCNMA1
102	TSC2	9	**221**	32	CACNA1B
103	GRIN2D	9	**222**	37	PCDH19
104	QARS	9	**223**	37	CDKL5
105	GRM7	9	**224**	38	KCND2
106	PPP3CA	9	**225**	38	SCN1B
107	KCNK4	9	**226**	39	SLC12A5
108	TNFRSF1A	9	**227**	39	GTF2I
109	PTEN	9	**228**	40	KCNC1
110	GABRB1	9	**229**	42	SLC6A1
111	NR3C1	9	**230**	42	GABRD
112	CCK	9	**231**	47	KCNT1
113	SNCAIP	10	**232**	48	KCNA2
114	PANK2	10	**233**	49	GRIA2
115	ANG	10	**234**	52	KCNQ3
116	SMN1	10	**235**	52	GABRA1
117	CASK	10	**236**	56	SCN8A
118	CTSD	10	**237**	63	GABRG2
119	GSK3B	10	**238**	71	KCNQ2

The Network of protein targets obtained from STRING database is shown in this figure:

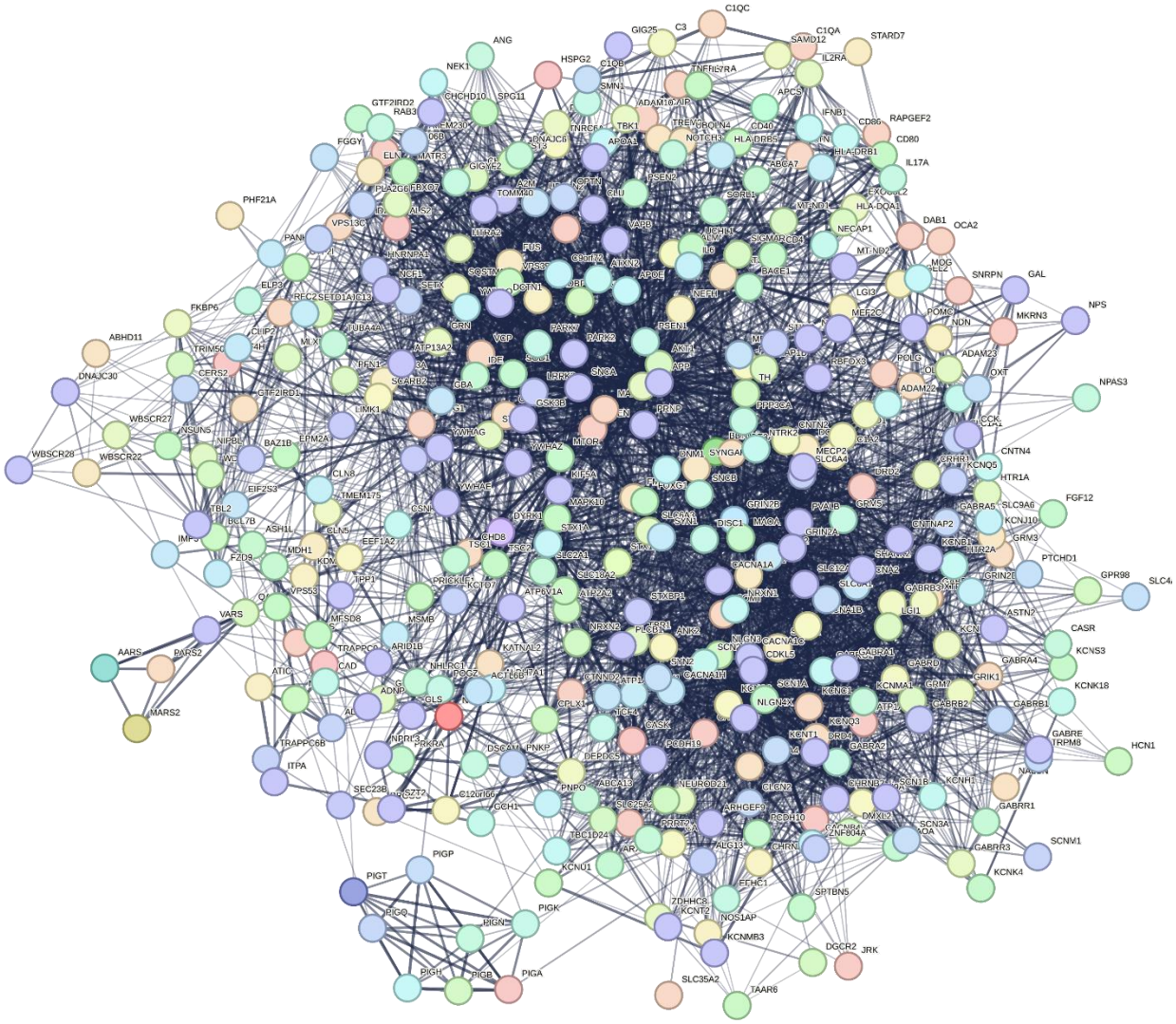

Figure 4.1 *Network of protein targets obtained from STRING database.*

The sequences of all the selected target proteins were obtained from UniProt using the given UniProt IDs and downloaded the data in an Excel file. The data of target protein sequences is shown in table 6.

Table 6 *Protein-Targets Sequences obtained from Uniprot*

Sr. #	Targets Names	Targets Sequences
1.	**KCNU1**	>sp|A8MYU2|KCNU1_HUMAN Potassium channel subfamily U member 1 OS=Homo sapiens OX=9606 GN=KCNU1 PE=1 SV=2

		MNPLAQPVIYSTIFAGTLITALSSHWFFTWVGLEMNMLAFIP VLTKKMNPRSTEAAIKYFLTQATASMILLMAILFNNMLSGQ WTMTNTTNQYSSLMIMMAMAMKLGMAPFHFWVPEVTQG TPLTSGLLLLTWQKLAPISIMYQISPSLNVSLLLTLSILSIMAGS WGGLNQTQLRKILAYSSITHMGWMMAVLPYNPNMTILNLTI YIILTTTAFLLLNLNSSTTTLLLSRTWNKLTWLTPLIPSTLLSL GGLPPLTGFLPKWAIIEEFTKNNSLIIPTIMATITLLNLYFYLRL IYSTSITLLPMSNNVKMKWQFEHTKPTPFLPTLIALTTLLLPIS PFMLMIL
2.	**NCF1**	>sp|P14598|NCF1_HUMAN Neutrophil cytosol factor 1 OS=Homo sapiens OX=9606 GN=NCF1 PE=1 SV=4 MGDTFIRHIALLGFEKRFVPSQHYVYMFLVKWQDLSEKVVY RRFTEIYEFHKTLKEMFPIEAGAINPENRIIPHLPAPKWFDGQ RAAENRQGTLTEYCSTLMSLPTKISRCPHLLDFFKVRPDDLK LPTDNQTKKPETYLMPKDGKSTATDITGPIILQTYRAIANYEK TSGSEMALSTGDVVEVVEKSESGWWFCQMKAKRGWIPASF LEPLDSPDETEDPEPNYAGEPYVAIKAYTAVEGDEVSLLEGE AVEVIHKLLDGWWVIRKDDVTGYFPSMYLQKSGQDVSQAQ RQIKRGAPPRRSSIRNAHSIHQRSRKRLSQDAYRRNSVRFLQ QRRRQARPGPQSPGSPLEEERQTQRSKPQPAVPPRPSADLILN RCSESTKRKLASAV
3.	**TAAR6**	>sp|Q96RI8|TAAR6_HUMAN Trace amine-associated receptor 6 OS=Homo sapiens OX=9606 GN=TAAR6 PE=2 SV=1 MSSNSSLLVAVQLCYANVNGSCVKIPFSPGSRVILYIVFGFGA VLAVFGNLLVMISILHFKQLHSPTNFLVASLACADFLVGVTV MPFSMVRTVESCWYFGRSFCTFHTCCDVAFCYSSLFHLCFISI DRYIAVTDPLVYPTKFTVSVSGICISVSWILPLMYSGAVFYTG VYDDGLEELSDALNCIGGCQTVVNQNWVLTDFLSFFIPTFIM IILYGNIFLVARRQAKKIENTGSKTESSSESYKARVARRERKA AKTLGVTVVAFMISWLPYSIDSLIDAFMGFITPACIYEICCWC

		AYYNSAMNPLIYALFYPWFRKAIKVIVTGQVLKNSSATMNL FSEHI
…	…	……
…	…	……
236.	**C1QC**	>sp|P02747|C1QC_HUMAN Complement C1q subcomponent subunit C OS=Homo sapiens OX=9606 GN=C1QC PE=1 SV=3 MDVGPSSLPHLGLKLLLLLLLLPLRGQANTGCYGIPGMPGLP GAPGKDGYDGLPGPKGEPGIPAIPGIRGPKGQKGEPGLPGHP GKNGPMGPPGMPGVPGPMGIPGEPGEEGRYKQKFQSVFTVT RQTHQPPAPNSLIRFNAVLTNPQGDYDTSTGKFTCKVPGLYY FVYHASHTANLCVLLYRSGVKVVTFCGHTSKTNQVNSGGV LLRLQVGEEVWLAVNDYYDMVGIQGSDSVFSGFLLFPD

4.1.3 Drugs Data Retrieval:

The data of all the FDA-approved, experimental phase, investigational phase, nutraceutical, Illicit and withdrawn drugs was retrieved from multiple databases, such as, DrugCentral, e-Drug3D, ChEMBL and DrugBank by giving the query of all the selected protein-targets one by one and all for all the Neurological and Neurodegenerative disorder individually. Then we merged the dataset of all the databases to remove duplicates of all drugs. Then we keep only the data of all FDA-approved drugs. After shortlisting, the data contained 964 unique drugs as shown in the Table 7. The SMILES of all the selected drugs were obtained from PubChem database using the given drug name and downloaded the data in an Excel file.

Table 7 *Data of 964 unique FDA-approved drugs for all the neuronal disorders*

Sr. no.	Drugs	Sr. no.	Drugs	Sr. no.	Drugs
1	NADH	**323**	Micronomicin	**645**	Cortisone acetate
2	Metformin	**324**	Fostamatinib	**646**	Desonide
3	Flavin adenine dinucleotide	**325**	Lithium citrate	**647**	Dexamethasone acetate
4	Ubidecarenone	**326**	Halothane	**648**	Desoximetasone

5	Cyclosporine	**327**	Ibutilide	**649**	Difluprednate
6	Haloperidol	**328**	Levamlodipine	**650**	Fluorometholone
7	Fentanyl	**329**	Ranolazine	**651**	Fluticasone
8	Valrubicin	**330**	Atenolol	**652**	Clocortolone
9	Oxygen	**331**	Amiodarone	**653**	Megestrol acetate
10	D-Tyrosine	**332**	Diltiazem	**654**	Flumethasone
11	Isotretinoin	**333**	Dronedarone	**655**	Meprednisone
12	Milnacipran	**334**	Isavuconazole	**656**	Alclometasone
13	Citalopram	**335**	Propiverine	**657**	Amcinonide
14	Clomipramine	**336**	Topiramate	**658**	Diflorasone
15	Escitalopram	**337**	Amoxicillin	**659**	Fluoxymesterone
16	Fluvoxamine	**338**	Clavulanic acid	**660**	Flurandrenolide
17	Paroxetine	**339**	Carbamazepine	**661**	Fluticasone furoate
18	Sertraline	**340**	Lamotrigine	**662**	Hydrocortamate
19	Fluoxetine	**341**	Oxcarbazepine	**663**	Hydrocortisone butyrate
20	Desipramine	**342**	Nitrazepam	**664**	Hydrocortisone cypionate
21	Duloxetine	**343**	Phenacemide	**665**	Hydrocortisone probutate
22	Imipramine	**344**	Clobazam	**666**	Hydrocortisone valerate
23	Nortriptyline	**345**	Levetiracetam	**667**	Loteprednol etabonate
24	Venlafaxine	**346**	Tetracaine	**668**	Medrysone
25	Atomoxetine	**347**	Permethrin	**669**	Prednicarbate
26	Amitriptyline	**348**	Phenazopyridine	**670**	Rimexolone
27	Amoxapine	**349**	Brivaracetam	**671**	Ursodeoxycholic acid
28	Desvenlafaxine	**350**	Chloroprocaine	**672**	Betamethasone phosphate
29	Doxepin	**351**	Dichlorobenzyl alcohol	**673**	Clobetasone
30	Protriptyline	**352**	Pramocaine	**674**	Drospirenone

31	Trimipramine	**353**	Propoxycaine	**675**	Fludrocortisone
32	Cocaine	**354**	Eslicarbazepine	**676**	Gestrinone
33	Phentermine	**355**	Eslicarbazepine acetate	**677**	Hydrocortisone phosphate
34	Vilazodone	**356**	Ethotoin	**678**	Levonorgestrel
35	Mirtazapine	**357**	Lacosamide	**679**	Norethisterone
36	Lumateperone	**358**	Mexiletine	**680**	Segesterone acetate
37	Vortioxetine	**359**	Prilocaine	**681**	Ulipristal
38	Bupropion	**360**	Primidone	**682**	Ulobetasol
39	Risperidone	**361**	Propafenone	**683**	Cyproterone acetate
40	Minaprine	**362**	Rufinamide	**684**	Ethinylestradiol
41	Tramadol	**363**	Safinamide	**685**	Fluprednisolone
42	Trazodone	**364**	Amiloride	**686**	Methyltestosterone
43	Mazindol	**365**	Geraniol	**687**	Paramethasone acetate
44	Methylphenidate	**366**	Hydroxycitronellal	**688**	Prednisolone tebutate
45	Dosulepin	**367**	Linoleic acid	**689**	Caffeine
46	Dextroamphetamine	**368**	Ouabain	**690**	Magnesium oxide
47	Dopamine	**369**	Digoxin	**691**	Sincalide
48	Meperidine	**370**	Bisacodyl	**692**	Cholecystokinin
49	Verapamil	**371**	Rubidium Rb-82	**693**	Ginger
50	Buprenorphine	**372**	Acetyldigitoxin	**694**	Pantothenic acid
51	Ethanol	**373**	Deslanoside	**695**	Pidolic acid
52	Methadone	**374**	Digitoxin	**696**	Risdiplam
53	Morphine	**375**	Bromazepam	**697**	Amikacin
54	Racephedrine	**376**	Chlordiazepoxide	**698**	Anisindione
55	Ribavirin	**377**	Clonazepam	**699**	Hydralazine
56	Clozapine	**378**	Clorazepic acid	**700**	Nabumetone
57	Dexmethylphenidate	**379**	Clotiazepam	**701**	Phenylbutyric acid
58	Dextromethorphan	**380**	Estazolam	**702**	Thiabendazole
59	Mianserin	**381**	Flurazepam	**703**	Streptozocin
60	Aripiprazole	**382**	Midazolam	**704**	Betulinic Acid

61	Benzatropine	**383**	Oxazepam	**705**	Econazole
62	Chlorpheniramine	**384**	Prazepam	**706**	Dapsone
63	Cyclobenzaprine	**385**	Quazepam	**707**	Lapatinib
64	Lorpiprazole	**386**	Temazepam	**708**	Rosuvastatin
65	Loxapine	**387**	Triazolam	**709**	Allopurinol
66	Procaine	**388**	Alprazolam	**710**	Nevirapine
67	Tapentadol	**389**	Diazepam	**711**	Carbimazole
68	Berberine	**390**	Remimazolam	**712**	Efavirenz
69	Olanzapine	**391**	1,2-Benzodiazepine	**713**	Flucloxacillin
70	Quetiapine	**392**	gamma-Aminobutyric acid	**714**	Lumiracoxib
71	Solriamfetol	**393**	Miltefosine	**715**	Methimazole
72	Phenelzine	**394**	Gemcitabine	**716**	Pitavastatin
73	Ziprasidone	**395**	Nelfinavir	**717**	Ticlopidine
74	Copper	**396**	Arsenic trioxide	**718**	Fospropofol
75	Mycophenolic acid	**397**	Anastrozole	**719**	Piperazine
76	Dequalinium	**398**	Fulvestrant	**720**	Zopiclone
77	Gentian violet cation	**399**	Alectinib	**721**	Carisoprodol
78	Ketoconazole	**400**	Alpelisib	**722**	Naratriptan
79	Azacitidine	**401**	Dabrafenib	**723**	Zolmitriptan
80	Chloroxine	**402**	Everolimus	**724**	Methysergide
81	Dactinomycin	**403**	Irinotecan	**725**	Pindolol
82	Daunorubicin	**404**	Lovastatin	**726**	Eletriptan
83	Dipyrithione	**405**	Nintedanib	**727**	Pizotifen
84	Fluorescein	**406**	Sorafenib	**728**	Oxymetazoline
85	Fluspirilene	**407**	Temsirolimus	**729**	Penbutolol
86	Idarubicin	**408**	Vemurafenib	**730**	Tropisetron
87	Leflunomide	**409**	Flumazenil	**731**	Alverine
88	Mitoxantrone	**410**	Etomidate	**732**	Dihydroergocornine
89	Oxybenzone	**411**	Rotigotine	**733**	Gilteritinib
90	Proflavine	**412**	Paliperidone	**734**	Lofexidine

91	Promethazine	**413**	Sulpiride	**735**	Frovatriptan
92	Terazosin	**414**	Methotrimeprazine	**736**	Rizatriptan
93	Tetramethylthiuram monosulfide	**415**	Ropinirole	**737**	Moclobemide
94	Thimerosal	**416**	Cabergoline	**738**	Selegiline
95	Thiram	**417**	Lisuride	**739**	Linezolid
96	Triclocarban	**418**	Asenapine	**740**	Tranylcypromine
97	Vandetanib	**419**	Propiomazine	**741**	Isocarboxazid
98	Histamine	**420**	Aripiprazole lauroxil	**742**	Almotriptan
99	Ketamine	**421**	Buspirone	**743**	Pargyline
100	Norepinephrine	**422**	Epicriptine	**744**	Tedizolid
101	Pramipexole	**423**	Ergoloid mesylate	**745**	Riboflavin
102	Valproic acid	**424**	Flibanserin	**746**	Betahistine
103	Diamorphine	**425**	Iloperidone	**747**	Eravacycline
104	Sumatriptan	**426**	Nandrolone decanoate	**748**	Oxymetholone
105	Capsaicin	**427**	Modafinil	**749**	Procarbazine
106	Chlorpromazine	**428**	Ambrisentan	**750**	Tedizolid phosphate
107	Colchicine	**429**	Amisulpride	**751**	Testosterone cypionate
108	Doxorubicin	**430**	Domperidone	**752**	Testosterone enanthate
109	Gentamicin	**431**	Mesoridazine	**753**	Testosterone undecanoate
110	Indomethacin	**432**	Metoclopramide	**754**	Ubrogepant
111	Levodopa	**433**	Droperidol	**755**	Ginkgo biloba
112	Paclitaxel	**434**	Phenyltoloxamine	**756**	Anethole
113	Pilocarpine	**435**	Promazine	**757**	Mitotane
114	Esketamine	**436**	Benperidol	**758**	Pimavanserin
115	Chondroitin sulfate	**437**	Fenoldopam	**759**	Metergoline
116	Amantadine	**438**	Prochlorperazine	**760**	Epinastine
117	Cyproheptadine	**439**	Betamethasone	**761**	Agomelatine
118	Eltanolone	**440**	Silver cation	**762**	Methylergometrine
119	Hydrocortisone	**441**	Rifampicin	**763**	Valine

120	Hydrocortisone acetate	**442**	Adenosine	**764**	Dalfampridine
121	Hydrocortisone succinate	**443**	Clofarabine	**765**	Guanidine
122	Lithium carbonate	**444**	Penciclovir	**766**	Glucosamine
123	Lorazepam	**445**	Sodium phosphate, monobasic	**767**	Pioglitazone
124	Mepivacaine	**446**	Valaciclovir	**768**	Rosiglitazone
125	Glutamic acid	**447**	Dexamethasone	**769**	Carboxymethylcellulose
126	Baclofen	**448**	Esculin	**770**	Fludeoxyglucose (18F)
127	Bortezomib	**449**	Empagliflozin	**771**	Lipoic acid
128	Corticotropin	**450**	Liraglutide	**772**	Acalabrutinib
129	Curcumin	**451**	Ampicillin	**773**	Amphotericin B
130	Didanosine	**452**	Carindacillin	**774**	Omacetaxine mepesuccinate
131	Donepezil	**453**	Cefamandole nafate	**775**	Pyrimethamine
132	Dronabinol	**454**	Cefixime	**776**	Benzocaine
133	Fluorouracil	**455**	Dicloxacillin	**777**	Tannic acid
134	Flurbiprofen	**456**	Mesalazine	**778**	Entacapone
135	Gemfibrozil	**457**	Icosapent	**779**	Opicapone
136	Glucagon	**458**	Losartan	**780**	Sufentanil
137	Lidocaine	**459**	Imatinib	**781**	Dobutamine
138	Methylprednisolone	**460**	Isoprenaline	**782**	Micafungin
139	Metyrosine	**461**	Quinidine	**783**	Conjugated estrogens
140	Mifepristone	**462**	Artenimol	**784**	Nylidrin
141	Naloxone	**463**	Acetic acid	**785**	Ademetionine
142	Nicardipine	**464**	Betaine	**786**	Alfentanil
143	Nicotine	**465**	Pyridoxine	**787**	Butorphanol
144	Nystatin	**466**	Diethylpropion	**788**	Codeine
145	Propranolol	**467**	Phenmetrazine	**789**	Hydrocodone
146	Propylthiouracil	**468**	Armodafinil	**790**	Hydromorphone

147	Simvastatin	**469**	Disulfiram	**791**	Trihexyphenidyl
148	Sodium chloride	**470**	Benzphetamine	**792**	Wood creosote
149	Topotecan	**471**	Diphenylpyraline	**793**	Ezogabine
150	Vasopressin	**472**	Lisdexamfetamine	**794**	Sorbitol
151	Verteporfin	**473**	Dutasteride	**795**	Erythromycin
152	Vincristine	**474**	Ioflupane	**796**	Hydroquinone
153	Fenofibrate	**475**	Serdexmethylphenidate	**797**	Fludarabine
154	Testosterone	**476**	Acetylcysteine	**798**	Theophylline
155	Zinc acetate	**477**	Carbetocin	**799**	Cetirizine
156	Cholesterol	**478**	Atosiban	**800**	Cladribine
157	Zinc cation	**479**	Oxytocin	**801**	Decitabine
158	Zinc chloride	**480**	Desmopressin	**802**	Flutamide
159	Zinc sulfate	**481**	5-methyltetrahydrofolic acid	**803**	Levothyroxine
160	Infigratinib	**482**	Human immunoglobulin G	**804**	Mercaptopurine
161	Sirolimus	**483**	Pemetrexed	**805**	Garlic
162	Furosemide	**484**	Cupric Chloride	**806**	Pirfenidone
163	Lamivudine	**485**	Raloxifene	**807**	Lactose
164	L-Glutamine	**486**	Clonidine	**808**	Xylose
165	Aspartic acid	**487**	Dexmedetomidine	**809**	Amsacrine
166	Ammonia	**488**	Etoricoxib	**810**	Pyrithione
167	Famotidine	**489**	Oxycodone	**811**	Quinestrol
168	Phosphoric acid	**490**	Remifentanil	**812**	Saquinavir
169	Water	**491**	Rocuronium	**813**	Iron
170	Racemethionine	**492**	Acamprosate	**814**	Toremifene
171	Pyrophosphoric acid	**493**	Perphenazine	**815**	D-Phenylalanine
172	Selenomethionine	**494**	Fluphenazine	**816**	Sapropterin
173	Miconazole	**495**	Brexpiprazole	**817**	Doxycycline
174	Nitrendipine	**496**	Cariprazine	**818**	Iobenguane sulfate
175	Ritodrine	**497**	Lurasidone	**819**	Triamterene

176	Trimebutine	**498**	Bromperidol	**820**	Etidronic acid
177	Potassium	**499**	Quinagolide	**821**	Tiludronic acid
178	Pyridoxal phosphate	**500**	Acetophenazine	**822**	Alendronic acid
179	Trifluoperazine	**501**	Molindone	**823**	Flufenamic acid
180	Azathioprine	**502**	Thiothixene	**824**	Bacitracin
181	Mercaptopurine	**503**	Triflupromazine	**825**	Thrombin
182	Methotrexate	**504**	Zuclopenthixol	**826**	Primaquine
183	Citric acid	**505**	Pipotiazine	**827**	Zidovudine
184	Magnesium	**506**	Rasagiline	**828**	Clindamycin
185	Loperamide	**507**	Ergotamine	**829**	Abacavir
186	Setmelanotide	**508**	Yohimbine	**830**	Atovaquone
187	Afamelanotide	**509**	Dihydro-alpha-ergocryptine	**831**	Carbamide peroxide
188	Galactose	**510**	Maprotiline	**832**	Fosamprenavir
189	Cinacalcet	**511**	Memantine	**833**	Lenalidomide
190	Framycetin	**512**	Tetrabenazine	**834**	Levoleucovorin
191	Etelcalcetide	**513**	Naltrexone	**835**	Maraviroc
192	Strontium chloride	**514**	Dapiprazole	**836**	Pentamidine
193	Pamidronic acid	**515**	Dopexamine	**837**	Stavudine
194	Calcium citrate	**516**	Ergometrine	**838**	Sulfamethoxazole
195	Calcium Phosphate	**517**	Mephentermine	**839**	Trimethoprim
196	Calcium phosphate dihydrate	**518**	Metaraminol	**840**	Folic acid
197	Atorvastatin	**519**	Naphazoline	**841**	Amprenavir
198	Fluvastatin	**520**	Perazine	**842**	Fostemsavir
199	Pravastatin	**521**	Phenoxybenzamine	**843**	Tipranavir
200	Warfarin	**522**	Phentolamine	**844**	Insulin beef
201	Galantamine	**523**	Phenylephrine	**845**	Insulin human
202	Rivastigmine	**524**	Tolazoline	**846**	Insulin lispro
203	Acenocoumarol	**525**	Corticorelin ovine triflutate	**847**	Insulin pork

204	Ritonavir	**526**	Budesonide	**848**	Biotin
205	Ganciclovir	**527**	Telavancin	**849**	Urea
206	Irbesartan	**528**	Manganese	**850**	Trastuzumab deruxtecan
207	Lorazepam	**529**	Butabarbital	**851**	Nitric Oxide
208	Prednisone	**530**	Butalbital	**852**	Lysine
209	Triamcinolone	**531**	Butobarbital	**853**	Dofetilide
210	Lutein	**532**	Flunitrazepam	**854**	Doxazosin
211	Dimercaprol	**533**	Meprobamate	**855**	Ebastine
212	Aluminium	**534**	Methylphenobarbital	**856**	Glasdegib
213	Aluminium phosphate	**535**	Pentobarbital	**857**	Pazopanib
214	Aluminum acetate	**536**	Secobarbital	**858**	Dasatinib
215	Deferoxamine	**537**	Talbutal	**859**	Entrectinib
216	Florbetaben (18F)	**538**	Thiopental	**860**	DL-alpha-Tocopherol
217	Florbetapir (18F)	**539**	Apalutamide	**861**	Vitamin A
218	Flutemetamol (18F)	**540**	Eszopiclone	**862**	Iron sucrose
219	Tromethamine	**541**	Ganaxolone	**863**	Binimetinib
220	Hydroxychloroquine	**542**	Glutethimide	**864**	Tocopherol
221	Propofol	**543**	Ketazolam	**865**	Allantoin
222	Clotrimazole	**544**	Lormetazepam	**866**	Balsalazide
223	Hydrogen peroxide	**545**	Medroxyprogesterone acetate	**867**	Nedocromil
224	Tamoxifen	**546**	Stiripentol	**868**	Oxyquinoline
225	Estradiol	**547**	Taurine	**869**	Rosin
226	Cysteine	**548**	Phenobarbital	**870**	Ifosfamide
227	Bifonazole	**549**	Cenobamate	**871**	Levofloxacin
228	Chenodeoxycholic acid	**550**	Cyclophosphamide	**872**	Metronidazole
229	Dexibuprofen	**551**	Epirubicin	**873**	Minocycline
230	Diclofenac	**552**	Fluorouracil	**874**	Cabazitaxel

231	Estrone	**553**	Tacrolimus	**875**	Picropodophyllin
232	Fluconazole	**554**	Acetylsalicylic acid	**876**	Griseofulvin
233	Ibuprofen	**555**	Ivabradine	**877**	Ixabepilone
234	Progesterone	**556**	Albumin human	**878**	Trastuzumab emtansine
235	Riluzole	**557**	Ciclopirox	**879**	Vinblastine
236	Adenine	**558**	Trichlormethiazide	**880**	Vinflunine
237	Prasterone	**559**	Almitrine	**881**	Vinorelbine
238	Docetaxel	**560**	Bretylium	**882**	Pentazocine
239	Lansoprazole	**561**	Etacrynic acid	**883**	Noscapine
240	Menadione	**562**	Magnesium acetate	**884**	Mecamylamine
241	Dihydroergotamine	**563**	Magnesium gluconate	**885**	Carbamoylcholine
242	Racepinephrine	**564**	Potassium acetate	**886**	Decamethonium
243	Flortaucipir F-18	**565**	Potassium cation	**887**	Doxacurium
244	Eribulin	**566**	Potassium gluconate	**888**	Metocurine
245	Ethambutol	**567**	Potassium sulfate	**889**	Mivacurium
246	Acitretin	**568**	Magnesium cation	**890**	Pancuronium
247	Aminohippuric acid	**569**	Artemether	**891**	Pipecuronium
248	Amrinone	**570**	Eplerenone	**892**	Tubocurarine
249	Apomorphine	**571**	Lumefantrine	**893**	Vecuronium
250	Carbidopa	**572**	Diazoxide	**894**	Cisatracurium
251	Carboplatin	**573**	Ambroxol	**895**	Atracurium
252	Cefaclor	**574**	N-acetyl-alpha-D-glucosamine	**896**	Biperiden
253	Cisplatin	**575**	Beta-D-Glucose	**897**	Gallamine triethiodide
254	Coenzyme M	**576**	Acarbose	**898**	Atracurium besylate
255	Dantrolene	**577**	Migalastat	**899**	Succinylcholine
256	Dexketoprofen	**578**	Mirtazapine	**900**	Levallorphan
257	Dihydroergocristine	**579**	Hydrocortisone	**901**	Quinidine barbiturate
258	Doconexent	**580**	Ascorbic acid	**902**	Lubiprostone
259	Esomeprazole	**581**	Rifaximin	**903**	Miglustat
260	Felodipine	**582**	Sodium sulfate	**904**	Amylmetacresol

261	Gefitinib	**583**	Dolasetron	**905**	Etidocaine
262	Indigotindisulfonic acid	**584**	Granisetron	**906**	Lidoflazine
263	Itraconazole	**585**	Glimepiride	**907**	Gliquidone
264	Masoprocol	**586**	Repaglinide	**908**	Tolbutamide
265	Melphalan	**587**	Glipizide	**909**	Stearic acid
266	Methyldopa	**588**	Glyburide	**910**	Etoposide
267	Methylene blue	**589**	Carbon dioxide	**911**	Palmitic Acid
268	Niclosamide	**590**	Deutetrabenazine	**912**	Lansoprazole
269	Nifedipine	**591**	Valbenazine	**913**	Arginine
270	Nitazoxanide	**592**	Deserpidine	**914**	Chlorpropamide
271	Oxytetracycline	**593**	Rose bengal	**915**	Glymidine
272	Pantoprazole	**594**	Isometheptene	**916**	Tolazamide
273	Rabeprazole	**595**	Propylhexedrine	**917**	Minoxidil
274	Sulfasalazine	**596**	Rescinnamine	**918**	Nicorandil
275	Triclabendazole	**597**	Orphenadrine	**919**	Hydrocortisone
276	Trimetrexate	**598**	Glycine	**920**	Edaravone
277	alpha-Linolenic acid	**599**	Fluciclovine (18F)	**921**	Benzoyl peroxide
278	Cholecalciferol	**600**	Guaifenesin	**922**	alpha-Tocopherol acetate
279	Oxitriptan	**601**	Magnesium acetate tetrahydrate	**923**	alpha-Tocopherol succinate
280	Zonisamide	**602**	Magnesium carbonate	**924**	Flavin mononucleotide
281	Isradipine	**603**	Huperzine A	**925**	Larotrectinib
282	Flunarizine	**604**	Felbamate	**926**	Crizotinib
283	Manidipine	**605**	D-Serine	**927**	Sodium oxybate
284	Nilvadipine	**606**	Pimecrolimus	**928**	Fingolimod
285	Cinnarizine	**607**	Voclosporin	**929**	Choline
286	Bioallethrin	**608**	Capecitabine	**930**	Varenicline
287	Cannabidiol	**609**	Erlotinib	**931**	Atropine

288	Enflurane	**610**	Methylprednisolone hemisuccinate	**932**	Amobarbital
289	Spironolactone	**611**	Prednisolone	**933**	Estradiol acetate
290	Ergocalciferol	**612**	Prednisolone acetate	**934**	Estradiol benzoate
291	Amlodipine	**613**	Prednisolone phosphate	**935**	Estradiol cypionate
292	Clevidipine	**614**	Abiraterone	**936**	Estradiol dienanthate
293	Lacidipine	**615**	Binimetinib	**937**	Estradiol valerate
294	Nisoldipine	**616**	Copanlisib	**938**	Physostigmine
295	Celecoxib	**617**	Dacomitinib	**939**	Scopolamine
296	Ethosuximide	**618**	Encorafenib	**940**	Homatropine methylbromide
297	Gabapentin	**619**	Enzalutamide	**941**	Ipratropium
298	Gabapentin enacarbil	**620**	Idelalisib	**942**	Methscopolamine
299	Methsuximide	**621**	Niraparib	**943**	Oxybutynin
300	Oxatomide	**622**	Olaparib	**944**	Pregnenolone
301	Paramethadione	**623**	Oxaliplatin	**945**	Tropicamide
302	Phensuximide	**624**	Palbociclib	**946**	Pentolinium
303	Pregabalin	**625**	Rucaparib	**947**	Trimethaphan
304	Trimethadione	**626**	Selumetinib	**948**	Chlorzoxazone
305	Pimozide	**627**	Talazoparib	**949**	Bendroflumethiazide
306	Nimodipine	**628**	Temozolomide	**950**	Cromoglicic acid
307	Drotaverine	**629**	Tirbanibulin	**951**	Hydrochlorothiazide
308	Magnesium sulfate	**630**	Venetoclax	**952**	Gliclazide
309	Menthol	**631**	Vorinostat	**953**	Indapamide
310	Phenytoin	**632**	gamma-Hydroxybutyric acid	**954**	Nateglinide
311	Ziconotide	**633**	Zaleplon	**955**	Chlorothiazide
312	Adefovir dipivoxil	**634**	Fluticasone propionate	**956**	Carbon monoxide

313	Tenofovir disoproxil	**635**	Ciclesonide	**957**	Eptifibatide
314	Telbivudine	**636**	Clobetasol propionate	**958**	Hesperidin
315	Zolpidem	**637**	Fluocinolone acetonide	**959**	Disopyramide
316	D-Threonine	**638**	Fluocinonide	**960**	Bumetanide
317	Desflurane	**639**	Beclomethasone dipropionate	**961**	Finasteride
318	Sevoflurane	**640**	Deflazacort	**962**	Tiagabine
319	Isoflurane	**641**	Mometasone furoate	**963**	Cyclothiazide
320	Halothane	**642**	Flunisolide	**964**	Meclofenamic acid
321	Perampanel	**643**	Piracetam	**965**	Methohexital

4.2 Descriptors Formation:

Once the data of the protein-targets, drugs and diseases was collected the descriptors were computed for proteins and drugs to train the various machine learning and deep learning models.

4.2.1 Protein Descriptors Computation:

The sequence-based descriptors of all the selected protein-targets were computed using web interphase of ProtrWeb. 9921 descriptors were calculated for all the 238 unique protein-targets including amino acid composition, dipeptide and tripeptide composition, C/T/D (Composition/Transition/Distribution), conjoint triad, sequence-order coupling number, Quasi-sequence-order descriptors, pseudo-amino acid, and amphiphilic pseudo-amino acid composition. The overview of the excel file of target sequence descriptors is shown in table 8 in appendix.

4.2.2 Drugs Fingerprints Computation:

For the drugs dataset, Morgan fingerprints were generated which is the best molecular fingerprint used for drug discovery purposes. The SMILES of all the drugs were extracted from PubChem which were used as input for the calculation of the Morgan fingerprints. The ALLChem package of RDKit library was imported in Python for the generation of Morgan fingerprints. Morgan fingerprints for all the drugs are shown in the table 9 in appendix.

4.3 Targets-Diseases Network Construction:

After extracting the data of all the targets protein for ten neurodegenerative disorders and nine neurological disorders from STRING and Therapeutic Target Database, it was saved in an excel file and then give the class label for each target class with respect to the disorder type. The targets-diseases network was constructed in such a way that there is an association of target-protein with the specific disorder type. The targets-diseases network is shown in table 10 in appendix.

4.4 Diseases-Drugs Network Construction:

By using the collected data of the FDA-approved drugs associated with specific protein-targets and for all the selected disorders using different online available databases, as mentioned in the methodology chapter, we constructed the Diseases-Drugs network for association of the 10 neurodegenerative disorders and nine neurological disorders with the drugs. The diseases-drugs network was constructed in such a way that there is an association of each FDA-approved drug with the specific disorder type. The data of the Diseases-Drugs Network saved in an csv file, shown in table 11 in appendix.

4.5 Targets-Diseases-Drugs Network Construction:

Targets-Diseases-Drugs Network is a comprehensive network that combines diverse data types to understand the intricate relationships between neurological and neurodegenerative disorders and their potential therapeutic interventions, which is constructed by combining the two networks: the Targets-Diseases Network and the Diseases-Drugs Network. The network provides an encompassing perspective, offering insights into the triadic relationship among protein targets, disorders, and the FDA-approved drugs. Each facet of the network is represented using distinct techniques, with protein targets characterized by descriptors and drugs represented by binary-encoded Morgan fingerprints. The final constructed Targets-Diseases-Drugs association network is shown in the table 12 in appendix.

4.6 Classification models for Targets-Diseases Network Model training & prediction:

Different classification models were trained to predict the disease association of the protein with respect to neurological and neurodegenerative disorders. Out of 4884 target proteins, 3440 came out to be associated with neurological disorders. While 1443 came out to be associated with neurodegenerative disorders. The models were built on the protein descriptors dataset used as the X-matrix while the class label was assigned based on the specific disorder type. All the protein targets associated with neurodegenerative diseases were labeled as 0 and all the protein targets associated with neurological diseases were labeled as 1. For training the model, first the dataset was splitted into 80% training set and 20% testing set and then perform hyperparameters tunning through the GridSearchCV, short for Grid Search Cross-Validation, which is a hyperparameter tuning technique widely used in machine learning. It automated the process of finding the best combination of hyperparameters for a given machine learning model by exhaustively searching through a predefined set of hyperparameter values. The data matrix was of the order 4884×9921.

First, the Classification model was built by Support Vector Machine on the Target-Diseases association Network. Then a parameters grid was defined for the SVM model training such as ‘C’:

[0.1, 1, 10], ‘gamma’: [0.1, 0.01, 0.001], and ‘kernel’: ['rbf'] for the hyperparameters tunning of the model parameters. To achieve the best model performance for the Targets-Diseases Network by using SVM, a special hyperparameter tuning technique i.e., GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'C': 1, 'gamma': 0.1, and 'kernel': 'rbf'. After training the model, the model achieved 91.7% Accuracy, 91.3% Precision, 97.3% Recall and 94.2% F1 Score on the test set.

Another Classification model was built by using Decision Tree Classifier on the Target-Diseases association Network. A parameters grid was defined for the decision tree model training such as 'max_depth': [None, 5, 10, 15], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4], and 'max_features': [None, 'sqrt', 'log2'] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'max_depth': 15, 'max_features': 'log2', 'min_samples_leaf': 1, and 'min_samples_split': 5. After training the model, the model achieved 90.8% Accuracy, 91.4% Precision, 95.7% Recall and 93.5% F1 Score on the test set.

Another Classification model was built by using Random Forest Classifier on the Target-Diseases association Network. A parameters grid was defined for the random forest model training such as 'n_estimators': [100, 200, 300], 'max_depth': [None, 5, 10], 'min_samples_split': [2, 5], 'min_samples_leaf': [1, 2], and 'max_features': ['sqrt', 'log2'] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'max_depth': None, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 5, and 'n_estimators': 200. After training the model, the model achieved 91.6% Accuracy, 91.7% Precision, 96.7% Recall and 94.1% F1 Score on the test set.

Gradient Bossting Machine (GBM) model was built on the Target-Diseases association Network dataset for classification. A parameters grid was defined for the GBM model training such as 'n_estimators': [100, 200], 'learning_rate': [0.01, 0.1], 'max_depth': [3, 5], 'min_samples_split': [2, 5], and 'min_samples_leaf': [1, 2] for the hyperparameters tunning of the model parameters.

GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'learning_rate': 0.1, 'max_depth': 3, 'min_samples_leaf': 1, 'min_samples_split': 5, and 'n_estimators': 100. After training the model, the model achieved 91.8% Accuracy, 91.6% Precision, 97.2% Recall and 94.3% F1 Score on the test set.

At the end, another Classification model was built by using MLP classifier on the Target-Diseases association Network. 80-20 train-test split was done for training and testing the MLP neural network model. The model learned by the following sets of parameters such as hidden_layer_sizes= (50, 100), activation='relu', solver='adam', alpha=0.01, and learning_rate = 'adaptive'. After training the model, the model achieved 64.6% Accuracy, 68.8% Precision, 90.6% Recall and 79.7% F1 Score on the test set. Figure 4.2 shows a comparison between the performance measures such as accuracy, recall, precision and F1-Score on the test set among all the trained machine learning/deep learning models on Target-Disease Network.

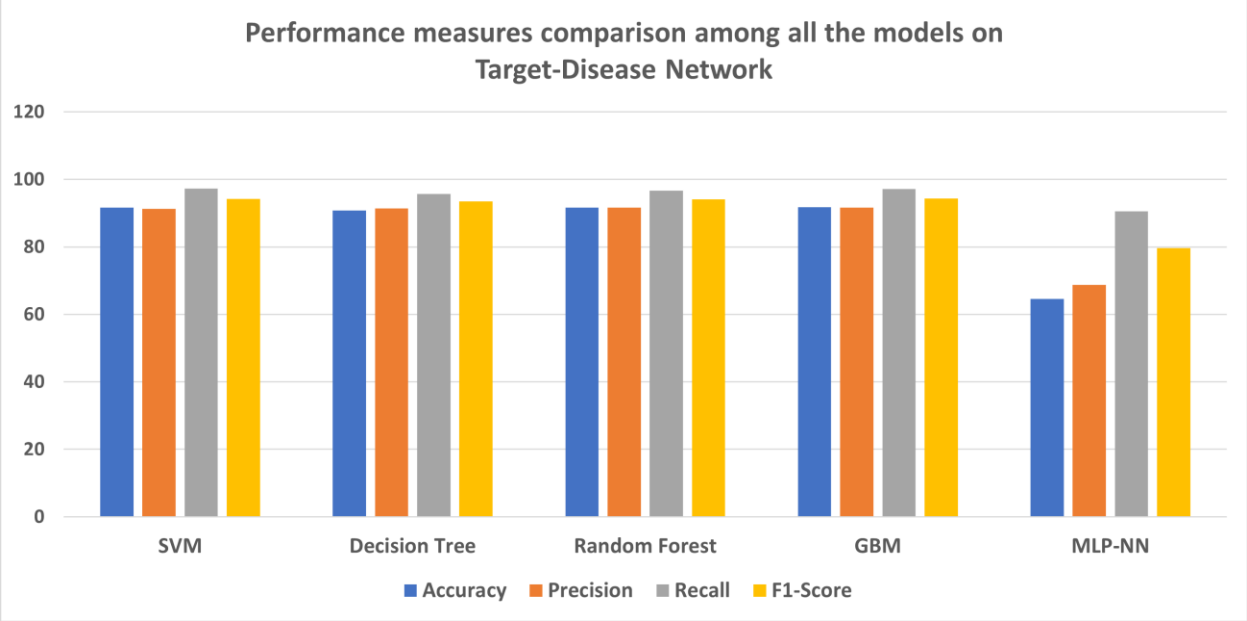

Figure 4.2 *Performance measures comparison among all the models on Target-Disease Network*

4.7 Classification models for Diseases-Drugs Network Model training & prediction:

There were different classification models which were trained for the prediction of the drugs against targets-protein associated with neurodegenerative disorders and with neurological disorders. The labels were given to the drugs based on the activation/inhibition data collected from different online available databases. Out of 4884 drugs, 3440 came out to be active against targets associated with neurological disorders while the other 1443 were active against targets associated with neurodegenerative disorders. The models were built on the morgan fingerprints dataset used as the X-matrix while the class label was assigned based on the specific disorder type. For training the model, first the dataset was splitted into 80% training set and 20% testing set and then perform hyperparameters tunning through the GridSearchCV. It automated the process of finding the best combination of hyperparameters for a given machine learning model by exhaustively searching through a predefined set of hyperparameter values. The data matrix was of the order 4884×2048.

The first Classification model was built by Support Vector Machine on the Diseases-Drugs association Network. Then a parameters grid was defined for the SVM model training such as 'C': [0.1, 1, 10], 'gamma': [0.1, 0.01, 0.001], and 'kernel': ['rbf'] for the hyperparameters tunning of the model parameters. To achieve the best model performance for the Diseases-Drugs Network by using SVM, a special hyperparameter tuning technique i.e., GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'C': 10, 'gamma': 0.1, and 'kernel': 'rbf'. After training the model, the model achieved 76.2% Accuracy, 78.1% Precision, 91.8% Recall and 84.4% F1 Score on the test set.

Another Classification model was built by using Decision Tree Classifier on the Diseases-Drugs association Network. A parameters grid was defined for the decision tree model training such as 'max_depth': [None, 5, 10, 15], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4], and 'max_features': [None, 'sqrt', 'log2'] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'max_depth': None, 'max_features': None, 'min_samples_leaf': 1, and

'min_samples_split': 10. After training, the model achieves 74.8% accuracy, 79.7% precision, 85.8% recall, and 82.7% F1 score on the test set.

Another Classification model was built by using Random Forest Classifier on the Diseases-Drugs association Network. A parameters grid was defined for the random forest model training such as 'n_estimators': [100, 200, 300], 'max_depth': [None, 5, 10], 'min_samples_split': [2, 5], 'min_samples_leaf': [1, 2], and 'max_features': ['sqrt', 'log2'] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'max_depth': None, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 5, and 'n_estimators': 100. After training, the model achieves 76.6% accuracy, 78.7% precision, 91.3% recall, and 84.5% F1 score on the test set.

Gradient Bossting Machine (GBM) model was built on the Diseases-Drugs association Network dataset for classification. A parameters grid was defined for the GBM model training such as 'n_estimators': [100, 200], 'learning_rate': [0.01, 0.1], 'max_depth': [3, 5], 'min_samples_split': [2, 5], and 'min_samples_leaf': [1, 2] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'learning_rate': 0.1, 'max_depth': 5, 'min_samples_leaf': 2, 'min_samples_split': 2, and 'n_estimators': 100. After training, the model achieves 74.6% accuracy, 76.8% precision, 91.2% recall, and 83.4% F1 score on the test set.

At the end, another Classification model was built by using MLP classifier on the Diseases-Drugs association Network. The model learned by the following sets of parameters such as hidden_layer_sizes= (50, 100), activation='relu', solver='adam', alpha=0.01, and learning_rate = 'adaptive'. After training the model, the model achieved 75.1% Accuracy, 78.7% Precision, 88.4% Recall and 83.5% F1 Score on the test set. Figure 4.3 shows a comparison between the performance measures such as accuracy, recall, precision and F1-Score on the test set among all the trained machine learning/deep learning models on Disease-Drug Network.

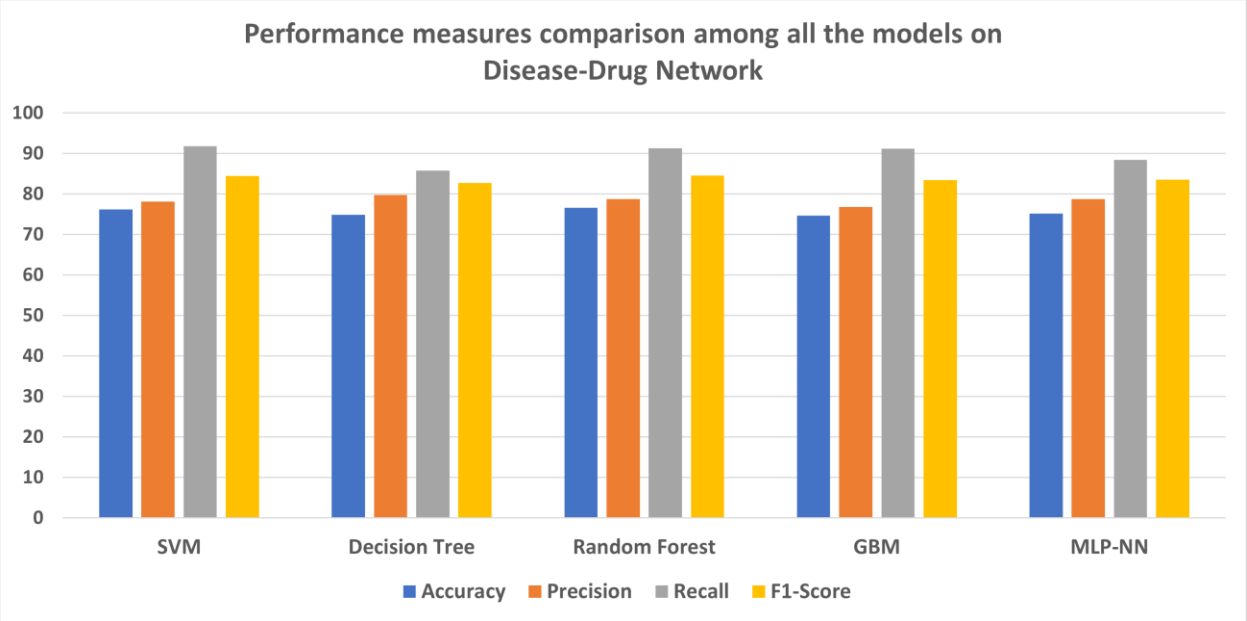

Figure 4.3 *Performance measures comparison among all the models on Disease-Drug Network*

4.8 Classification models for Targets-Diseases-Drugs association Network Model training & prediction:

There were different classification models which were trained on the combined Targets-Diseases-Drugs Network to predict the disease association with the protein targets as well as their drugs with respect to neurological and neurodegenerative disorders. The dataset of Targets-Diseases-Drugs association Network for training was built by combining the datasets of the two networks i.e., Targets-Diseases Network and Diseases-Drugs Network. Out of total 4884 number of instances in the network, 3440 target proteins and drugs came out to be associated with neurological disorders. While 1443 target proteins and drugs came out to be associated with neurodegenerative disorders. The models were built on the complete network dataset having 9921 protein descriptors, 2048 morgan's fingerprints of all the drugs, one hot encoding of the nineteen neuronal disorders, drugs activity and no. of protein edges were used as the X-matrix of the dataset while the class label was assigned based on the specific disorder type. To achieve the best performance by the models, GridSearchCV was used for model training. The data matrix was of the order 4884×11990. To compile a complete network dataset, total 4884 drugs as instances were

taken in each class active against neurological and neurodegenerative disorders having 2048 features in X-matrix (4884×11990).

The first Classification model was built by Support Vector Machine on the Targets-Diseases-Drugs association Network. A parameters grid was defined for the SVM model training such as ‘C': [0.1, 1, 10], 'gamma': [0.1, 0.01, 0.001], and 'kernel': ['rbf'] for the hyperparameters tunning of the model parameters. For achieving the best model performance for the Targets-Diseases-Drugs Network by using SVM, a special hyperparameter tuning technique i.e., GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'C': 10, 'gamma': 0.01, and 'kernel': 'rbf'. After training the model, the model achieved 99.2% Accuracy, 99.2% Precision, 99.7% Recall and 99.4% F1 Score on the test set.

Another Classification model was built by using Decision Tree Classifier on the Targets-Diseases-Drugs association Network. A parameters grid was defined for the decision tree model training such as 'max_depth': [None, 5, 10, 15], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4], and 'max_features': [None, 'sqrt', 'log2'] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'max_depth': 15, 'max_features': None, 'min_samples_leaf': 1, and 'min_samples_split': 2. After training the model, the model achieved 99.0% Accuracy, 99.8% Precision, 98.8% Recall and 99.3% F1 Score on the test set.

Random Forest Classifier was trained on the Targets-Diseases-Drugs association Network. A parameters grid was defined for the random forest model training such as 'n_estimators': [100, 200, 300], 'max_depth': [None, 5, 10], 'min_samples_split': [2, 5], 'min_samples_leaf': [1, 2], and 'max_features': ['sqrt', 'log2'] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'max_depth': None, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 5, and 'n_estimators': 300. After training the model, the model achieved 95.5% Accuracy, 94.9% Precision, 98.9% Recall and 96.9% F1 Score on the test set.

Gradient Bossting Machine (GBM) model was built on the Targets-Diseases-Drugs association Network dataset for classification. A parameters grid was defined for the GBM model training such as 'n_estimators': [100, 200], 'learning_rate': [0.01, 0.1], 'max_depth': [3, 5], 'min_samples_split': [2, 5], and 'min_samples_leaf': [1, 2] for the hyperparameters tunning of the model parameters. GridSearchCV was applied on the defined parameters grid. The model automatically learned by applying different sets of parameters and ultimately achieved the best performance. The best parameters set was 'learning_rate': 0.1, 'max_depth': 5, 'min_samples_leaf': 2, 'min_samples_split': 5, and 'n_estimators': 100. After training the model, the model achieved 99.3% Accuracy, 99.5% Precision, 99.4% Recall and 99.4% F1 Score on the test set.

At the end, another Classification model was built by using MLP classifier on the Targets-Diseases-Drugs association Network. The model learned by the following sets of parameters such as hidden_layer_sizes= (50, 100), activation='relu', solver='adam', alpha=0.01, and learning_rate = 'adaptive'. After training the model, the model achieved 70.6% Accuracy, 75.4% Precision, 86.1% Recall and 80.6% F1 Score on the test set. Figure 4.4 shows a comparison between the performance measures such as accuracy, recall, precision and F1-Score on the test set among all the trained machine learning/deep learning models on Target-Disease-Drug Network.

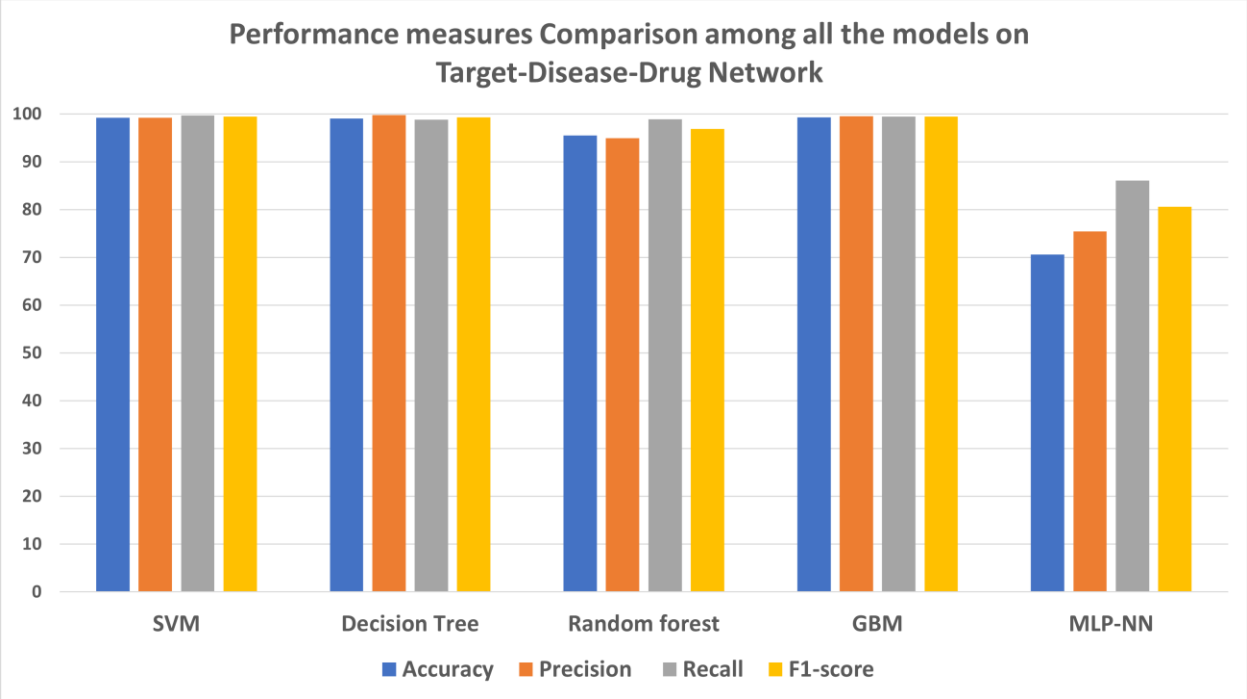

Figure 4.4 *Performance measures Comparison among all the models on Target-Disease-Drug Network*

4.9 Multi-variate Artificial Neural network Construction for Drug Repurposing:

After the construction of three networks i.e., targets-disease network, diseases-drugs network and targets-disease-drugs association network, different machine learning and deep learning models were trained on these networks and achieved good accuracy. A multi-variate Artificial Neural network was designed to predict the interactions between drugs and target proteins linked to specific diseases. Two main datasets were utilized for this purpose: one encapsulating the Morgan's fingerprints of 4884 drugs as input dataset, and the other containing descriptors of 4884 target proteins along with disease class labels (0 or 1) as output dataset. Then the model was trained using Artificial Neural Network (ANN) strategy, employing a multi-variate output configuration. The resultant model can comprehend the intricate connections between drug fingerprints, target protein descriptors, and disease classes, ultimately enabling predictions for both target protein descriptors and disease class for a given drug.

4.9.1 Model's Performance:

Designing a proper architecture for a Multi-Output Neural Network can be complex, especially when dealing with varying ranges and scales of output features. The alignment of datasets at a granular level ensures that the model comprehends the intricate interplay between drug attributes, target protein properties, and disease classifications. The ANN's structure was systematically optimized via hyperparameter tuning, encompassing the configuration of the different number of hidden layers, neurons in each layer, activation functions, batch size, and learning rate, to optimize the network's performance. Learning rate was adjusted during training to fine-tune the weights as the training progresses. Hyperparameters were optimized with different values to enhance the performance and generalization of our model for our desired task involving multiple output variables. But it still needs more hyperparameters tunning for giving best model performance.

Figure 4.32 shows the performance of our model with comparison to varying hyperparameters and predictions on Targets descriptors and disease class.

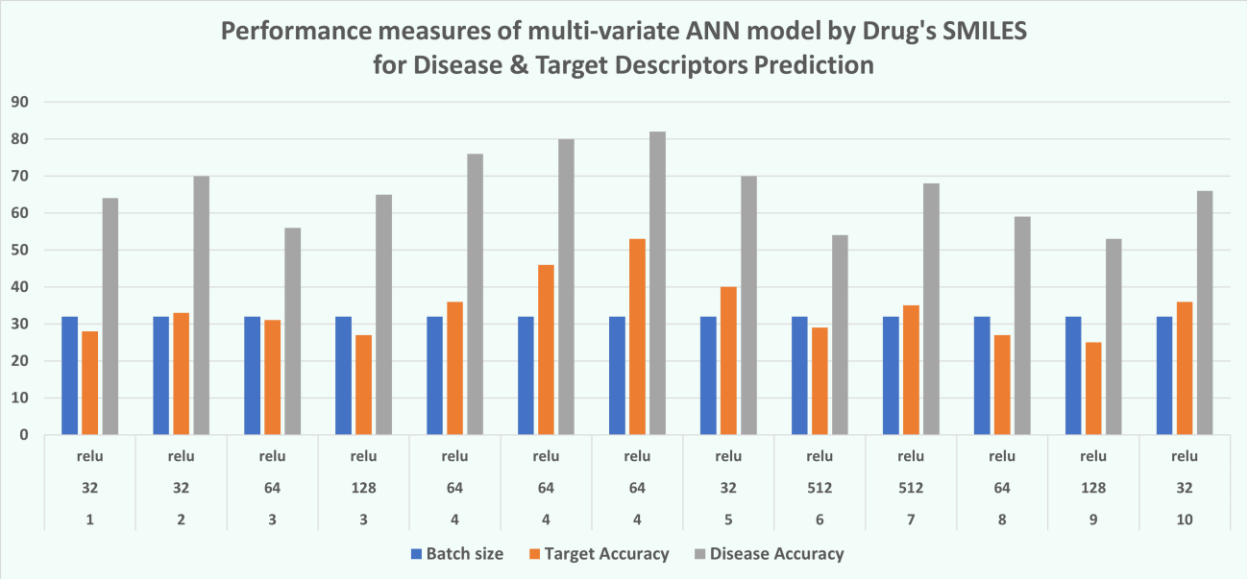

Figure 4.5 *Performance measures of multi-variate ANN model*

4.9.2 Input Unknown drug for Validation.

Our model was validated by giving the SMILES of an unknown drug. It automatically computes the morgan's fingerprints from the SMILES of that unidentified drug, using the RDkit library to form input features for the ANN. The model's outputs cater to predicting both target protein descriptors and disease class, presenting a holistic solution for drug-target interaction prediction. Our innovative model architecture predicted the target protein descriptors and disease class simultaneously for the given drug as shown in Figure 4.6.

```
Enter the SMILES string of the drug: CN(C)CCCN1C2=CC=CC=C2CCC3=CC=CC=C31
```

```
Enter the SMILES string of the drug: CN(C)CCCN1C2=CC=CC=C2CCC3=CC=CC=C31
1/1 [==============================] - 0s 18ms/step
Predicted Target Descriptors: [[0.07569497 0.05344455 0.03320792 ... 0.00024519 0.00086426 0.00060904]]
Predicted Disease Class: [[0]]
21/21 [==============================] - 0s 7ms/step - loss: 1.1508 - dense_5_loss: 0.3263 - dense_6_loss: 0.8246
Test Accuracy for Target Descriptors: 0.53835757386236412
Test Accuracy for Disease class: 0.8245556354522705
```

Figure 4.6 *multi-variate ANN model Validation*

Chapter 5

Conclusion

5 Conclusion

The protein plays a critical role in governing essential biological processes and intricate signaling cascades in a cell. Perturbations in protein activity, manifested as Loss-of-Function and Gain-of-Function alterations, have profound implications for a diverse spectrum of ailments. In this study, Network pharmacology approach was employed to investigate the targets association and overlap in various neurological and neurodegenerative disorders. Three networks were built between the protein-targets, FDA-approved drugs and neuronal disorders. The datasets of all the networks were categorized into two classes i.e., neurological disorders and neurodegenerative disorders.

For the classification of protein-targets and FDA-approved drugs for each neuronal disorder, five machine learning models were trained on the three different networks, the Decision Tree, Random Forest and Gradient Boosting Classifiers equally emerged as the optimal models for predicting the disease association of a given protein-target and drug. (Specifically in relation to neurological and neurodegenerative disorders). About 91% accuracy was achieved on the classification of Target-Disease Network, about 76% accuracy was achieved on the classification of Disease-Drug Network and about 99% accuracy was achieved on the test set of Target-Disease-Drug Network, which is an excellent model performance.

Our results provide a comprehensive view of the protein-targets association with the specific neurological and neurodegenerative disorders. Our study also revealed the target overlap among multiple neuronal disorders. We have developed a multi-variate Artificial Neural Network (ANN) to predict drug-target interactions linked to specific diseases. Our model was trained using a multi-variate output configuration, enabling predictions for both target protein descriptors and disease class for a given drug. The model's structure was optimized via hyperparameter tuning, encompassing configuration of hidden layers, activation functions, learning rates, and regularization techniques. Our multi-variate ANN model predicts multiple outputs, specifically disease class with 82% accuracy and target protein descriptors with 53% accuracy, each corresponding to a distinct facet of the problem, which provides a holistic solution for drug-target interaction prediction.

In conclusion, this study contributes to the database development for all the FDA-approved drugs and protein targets associated with many neurological and neurodegenerative disorders. We performed classification of all the protein-targets and drugs between neurological and neurodegenerative disorders. Our main contribution is the development of a model to predict target protein descriptors and disease class for CNS active agents. This will offer potential avenues for the development of new therapeutics for various un-treated disorders and ultimately pave a new insight into the personalized treatment strategies.

5.1 Future aspects:

Despite the insights provided by this study, there are several future directions that could be explored. First, multi label classification among the various neurological and neurodegenerative disorders could be performed based on the constructed database of all the protein targets and FDA-approved drugs. Secondly, our model's classification performance could be validated by clinical data to confirm the significance of our models in terms of targeted therapy approach. Additionally, the performance of our proposed model could be enhanced by further hyperparameters tuning and large computational resources. Lastly, our proposed model could be used as a helping hand in personalized treatment in hospitals. Therefore, future studies should aim to validate and extend our findings and explore new avenues to provide deeper insights into the biology of neuroscience and facilitate the development of more effective treatments for patients.

Chapter 6

References

6 References

1. Kennedy, D.P. and R. Adolphs, *The social brain in psychiatric and neurological disorders.* Trends Cogn Sci, 2012. **16**(11): p. 559-72.
2. Checkoway, H., J.I. Lundin, and S.N. Kelada, *Neurodegenerative diseases.* IARC scientific publications, 2011(163): p. 407-419.
3. Ryan, N.S., et al., *Magnetic resonance imaging evidence for presymptomatic change in thalamus and caudate in familial Alzheimer's disease.* Brain, 2013. **136**(5): p. 1399-1414.
4. Iyer, P.M., C.P. Doherty, and P. Bede, *Imaging Biomarkers in Neurodegenerative Conditions.* Neurodegenerative Disorders: A Clinical Guide, 2016: p. 13-27.
5. Burke, T., et al., *Huntington's disease.* Neurodegenerative disorders: A clinical guide, 2016: p. 167-179.
6. Mattson, M.P., *Apoptosis in neurodegenerative disorders.* Nature reviews Molecular cell biology, 2000. **1**(2): p. 120-130.
7. Floyd, R.A., *Antioxidants, oxidative stress, and degenerative neurological disorders.* Proceedings of the Society for Experimental Biology and Medicine, 1999. **222**(3): p. 236-245.
8. Holt, C.E., K.C. Martin, and E.M. Schuman, *Local translation in neurons: visualization and function.* Nature structural & molecular biology, 2019. **26**(7): p. 557-566.
9. Kennedy, D.P. and R. Adolphs, *The social brain in psychiatric and neurological disorders.* Trends in cognitive sciences, 2012. **16**(11): p. 559-572.
10. Lohoff, F.W., *Overview of the genetics of major depressive disorder.* Current psychiatry reports, 2010. **12**: p. 539-546.
11. Tiller, J.W., *Depression and anxiety.* The Medical Journal of Australia, 2013. **199**(6): p. S28-S31.
12. Grande, I., et al., *Bipolar disorder.* The Lancet, 2016. **387**(10027): p. 1561-1572.
13. Tsuang, M., *Schizophrenia: genes and environment.* Biological psychiatry, 2000. **47**(3): p. 210-220.
14. Hirokawa, N., S. Niwa, and Y. Tanaka, *Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease.* Neuron, 2010. **68**(4): p. 610-638.
15. Feigin, V.L., et al., *GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015.* Lancet Neurol, 2017. **16**(11): p. 877-97.
16. Charlson, F., et al., *New WHO prevalence estimates of mental disorders in conflict settings: a systematic review and meta-analysis.* The Lancet, 2019. **394**(10194): p. 240-248.
17. Collaborators, G.M.D., *Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019.* The Lancet Psychiatry, 2022. **9**(2): p. 137-150.
18. Deuschl, G., et al., *The burden of neurological diseases in Europe: an analysis for the Global Burden of Disease Study 2017.* The Lancet Public Health, 2020. **5**(10): p. e551-e567.
19. Organization, W.H., *Neurological disorders: public health challenges*. 2006: World Health Organization.

20. Wareham, L.K., et al., *Solving neurodegeneration: Common mechanisms and strategies for new treatments.* Molecular neurodegeneration, 2022. **17**(1): p. 23.
21. Seeley, W.W., et al., *Neurodegenerative diseases target large-scale human brain networks.* Neuron, 2009. **62**(1): p. 42-52.
22. Qin, L., et al., *Systemic LPS causes chronic neuroinflammation and progressive neurodegeneration.* Glia, 2007. **55**(5): p. 453-462.
23. Simpson, D.S. and P.L. Oliver, *ROS generation in microglia: understanding oxidative stress and inflammation in neurodegenerative disease.* Antioxidants, 2020. **9**(8): p. 743.
24. Hashimoto, M., et al., *Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases.* Neuromolecular medicine, 2003. **4**: p. 21-35.
25. Xie, A., et al., *Shared mechanisms of neurodegeneration in Alzheimer's disease and Parkinson's disease.* BioMed research international, 2014. **2014**.
26. Orsini, M., et al., *Amyotrophic lateral sclerosis: new perpectives and update.* Neurology international, 2015. **7**(2): p. 5885.
27. Snowden, S.G., et al., *Neurotransmitter imbalance in the brain and Alzheimer's disease pathology.* Journal of Alzheimer's Disease, 2019. **72**(1): p. 35-43.
28. Chen, Z., et al., *Epigenetic regulation of synaptic disorder in Alzheimer's disease.* Frontiers in Neuroscience, 2022. **16**: p. 888014.
29. Rao, R.V. and D.E. Bredesen, *Misfolded proteins, endoplasmic reticulum stress and neurodegeneration.* Current opinion in cell biology, 2004. **16**(6): p. 653-662.
30. Bayer, T.A. and O. Wirths, *Intracellular accumulation of amyloid-Beta-a predictor for synaptic dysfunction and neuron loss in Alzheimer's disease.* Frontiers in aging neuroscience, 2010. **2**: p. 1359.
31. Sharma, M., et al., *Intranasal exposure to low-dose rotenone induced alpha-synuclein accumulation and Parkinson's like symptoms without loss of dopaminergic neurons.* Neurotoxicity research, 2022. **40**(1): p. 215-229.
32. Rogers, J.T., et al., *Alzheimer's disease drug discovery targeted to the APP mRNA 5′ untranslated region.* Journal of Molecular Neuroscience, 2002. **19**: p. 77-82.
33. Bartzokis, G., *Age-related myelin breakdown: a developmental model of cognitive decline and Alzheimer's disease.* Neurobiology of aging, 2004. **25**(1): p. 5-18.
34. Lp, R., *Amyotrophic lateral sclerosis.* N Engl J Med, 2001. **344**: p. 1688-1700.
35. Whitwell, J.L., et al., *Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics.* Brain, 2012. **135**(3): p. 794-806.
36. Rohrer, J.D., et al., *Distinct profiles of brain atrophy in frontotemporal lobar degeneration caused by progranulin and tau mutations.* Neuroimage, 2010. **53**(3): p. 1070-1076.
37. Dyment, D.A., G.C. Ebers, and A.D. Sadovnick, *Genetics of multiple sclerosis.* The Lancet Neurology, 2004. **3**(2): p. 104-110.
38. Brockmann, K. and D. Berg, *The significance of GBA for Parkinson's disease.* Journal of inherited metabolic disease, 2014. **37**: p. 643-648.
39. Sacuiu, S., *Dementias.* Handbook of clinical neurology, 2016. **138**: p. 123-151.
40. Mead, S., *Prion disease genetics.* European Journal of Human Genetics, 2006. **14**(3): p. 273-281.
41. Dickson, D.W., R. Rademakers, and M.L. Hutton, *Progressive supranuclear palsy: pathology and genetics.* Brain Pathology, 2007. **17**(1): p. 74-82.

42. Zeydan, B. and K. Kantarci, *What Is Neurodegeneration?*, in *Molecular Imaging of Neurodegenerative Disorders*. 2023, Springer. p. 1-13.
43. Moguilner, S., et al., *Multimodal neurocognitive markers of frontal lobe epilepsy: insights from ecological text processing.* NeuroImage, 2021. **235**: p. 117998.
44. Thijs, R.D., et al., *Epilepsy in adults.* The Lancet, 2019. **393**(10172): p. 689-701.
45. Lin, S.-H., L.-T. Lee, and Y.K. Yang, *Serotonin and mental disorders: a concise review on molecular neuroimaging evidence.* Clinical Psychopharmacology and Neuroscience, 2014. **12**(3): p. 196.
46. Sutherland, H.G. and L.R. Griffiths, *Genetics of migraine: insights into the molecular basis of migraine disorders.* Headache: The Journal of Head and Face Pain, 2017. **57**(4): p. 537-569.
47. Heckers, S., et al., *Structure of the psychotic disorders classification in DSM-5.* Schizophrenia research, 2013. **150**(1): p. 11-14.
48. Lieberman, J.A. and M.B. First, *Psychotic disorders.* New England Journal of Medicine, 2018. **379**(3): p. 270-280.
49. Geschwind, D.H., *Genetics of autism spectrum disorders.* Trends in cognitive sciences, 2011. **15**(9): p. 409-416.
50. Zohar, J. and H. Westenberg, *Anxiety disorders: a review of tricyclic antidepressants and selective serotonin reuptake inhibitors.* Acta Psychiatrica Scandinavica, 2000. **101**: p. 39-49.
51. Cassidy, S.B. and D.J. Driscoll, *Prader–willi syndrome.* European journal of human genetics, 2009. **17**(1): p. 3-13.
52. Patterson, D., *Molecular genetic analysis of Down syndrome.* Human genetics, 2009. **126**: p. 195-214.
53. Morris, M., Colleen A and P.D. Mervis, Carolyn B, *Williams syndrome and related disorders.* Annual review of genomics and human genetics, 2000. **1**(1): p. 461-484.
54. Moshé, S.L., et al., *Epilepsy: new advances.* The Lancet, 2015. **385**(9971): p. 884-898.
55. Feinstein, A., *The neuropsychiatry of multiple sclerosis.* The Canadian Journal of Psychiatry, 2004. **49**(3): p. 157-163.
56. Sutherland, H.G., C.L. Albury, and L.R. Griffiths, *Advances in genetics of migraine.* The journal of headache and pain, 2019. **20**(1): p. 1-20.
57. McKeith, I., et al., *Dementia with Lewy bodies.* The Lancet Neurology, 2004. **3**(1): p. 19-28.
58. Ogilvie, A.D., et al., *Polymorphism in serotonin transporter gene associated with susceptibility to major depression.* The Lancet, 1996. **347**(9003): p. 731-733.
59. Geddes, J.R. and D.J. Miklowitz, *Treatment of bipolar disorder.* The lancet, 2013. **381**(9878): p. 1672-1682.
60. Hollander, E., *Treatment of obsessive-compulsive spectrum disorders with SSRIs.* The British Journal of Psychiatry, 1998. **173**(S35): p. 7-12.
61. Talkowski, M.E., E.V. Minikel, and J.F. Gusella, *Autism spectrum disorder genetics: diverse genes with diverse clinical outcomes.* Harvard review of psychiatry, 2014. **22**(2): p. 65.
62. Butler, M.G., J.L. Miller, and J.L. Forster, *Prader-Willi syndrome-clinical genetics, diagnosis and treatment approaches: an update.* Current pediatric reviews, 2019. **15**(4): p. 207-244.

63. Roizen, N.J. and D. Patterson, *Down's syndrome.* The Lancet, 2003. **361**(9365): p. 1281-1289.
64. Jurado, L.A.P., et al., *A duplicated gene in the breakpoint regions of the 7q11. 23 Williams-Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK.* Human molecular genetics, 1998. **7**(3): p. 325-334.
65. Grodzicki, W. and K. Dziendzikowska, *The role of selected bioactive compounds in the prevention of Alzheimer's disease.* Antioxidants, 2020. **9**(3): p. 229.
66. Adam, O.R. and J. Jankovic, *Symptomatic treatment of Huntington disease.* Neurotherapeutics, 2008. **5**(2): p. 181-197.
67. Riviere, M., et al., *An analysis of extended survival in patients with amyotrophic lateral sclerosis treated with riluzole.* Archives of neurology, 1998. **55**(4): p. 526-528.
68. Karapetyan, Y.E., et al., *Unique drug screening approach for prion diseases identifies tacrolimus and astemizole as antiprion agents.* Proceedings of the National Academy of Sciences, 2013. **110**(17): p. 7044-7049.
69. Stamelou, M. and G. Höglinger, *A review of treatment options for progressive supranuclear palsy.* CNS drugs, 2016. **30**: p. 629-636.
70. Diehl-Schmid, J., et al., *Imaging frontotemporal lobar degeneration.* Current neurology and neuroscience reports, 2014. **14**: p. 1-11.
71. Geschwind, D.H., et al., *Dementia and neurodevelopmental predisposition: cognitive dysfunction in presymptomatic subjects precedes dementia by decades in frontotemporal dementia.* Annals of neurology, 2001. **50**(6): p. 741-746.
72. Cheng, F., et al., *Network-based approach to prediction and population-based validation of in silico drug repurposing.* Nature communications, 2018. **9**(1): p. 2691.
73. Truong, T.T., et al., *Repurposing drugs via network analysis: Opportunities for psychiatric disorders.* Pharmaceutics, 2022. **14**(7): p. 1464.
74. Hopkins, A.L., *Network pharmacology: the next paradigm in drug discovery.* Nature chemical biology, 2008. **4**(11): p. 682-690.
75. Urbina, F., A.C. Puhl, and S. Ekins, *Recent advances in drug repurposing using machine learning.* Current Opinion in Chemical Biology, 2021. **65**: p. 74-84.
76. Hopkins, A.L., *Network pharmacology.* Nature biotechnology, 2007. **25**(10): p. 1110-1111.
77. Fang, J., et al., *Network pharmacology-based study on the mechanism of action for herbal medicines in Alzheimer treatment.* Journal of ethnopharmacology, 2017. **196**: p. 281-292.
78. Gopal, J., V.S. Prakash Sinnarasan, and A. Venkatesan, *Identification of repurpose drugs by computational analysis of disease–gene–drug associations.* Journal of Computational Biology, 2021. **28**(10): p. 975-984.
79. Singh, A., N. Thakur, and A. Sharma. *A review of supervised machine learning algorithms*. in *2016 3rd international conference on computing for sustainable global development (INDIACom)*. 2016. Ieee.
80. Choudhary, R. and H.K. Gianey. *Comprehensive review on supervised machine learning algorithms*. in *2017 International Conference on Machine Learning and Data Science (MLDS)*. 2017. IEEE.
81. Hahne, F., et al., *Unsupervised machine learning.* Bioconductor case studies, 2008: p. 137-157.
82. Adorf, C.S., et al., *Analysis of self-assembly pathways with unsupervised machine learning algorithms.* The Journal of Physical Chemistry B, 2019. **124**(1): p. 69-78.

83. Zhou, Z.-H. and Z.-H. Zhou, *Semi-supervised learning.* Machine Learning, 2021: p. 315-341.
84. Sutton, R.S., *Introduction: The challenge of reinforcement learning*, in *Reinforcement learning*. 1992, Springer. p. 1-3.
85. Weiss, K., T.M. Khoshgoftaar, and D. Wang, *A survey of transfer learning.* Journal of Big data, 2016. **3**(1): p. 1-40.
86. Dietterich, T.G. *Ensemble methods in machine learning*. in *International workshop on multiple classifier systems*. 2000. Springer.
87. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning.* nature, 2015. **521**(7553): p. 436-444.
88. Zhao, K. and H.-C. So, *Using drug expression profiles and machine learning approach for drug repurposing.* Computational methods for drug repurposing, 2019: p. 219-237.
89. Bossy-Wetzel, E., R. Schwarzenbacher, and S.A. Lipton, *Molecular pathways to neurodegeneration.* Nature medicine, 2004. **10**(Suppl 7): p. S2-S9.
90. Okouchi, M., et al., *Neuronal apoptosis in neurodegeneration.* Antioxidants & redox signaling, 2007. **9**(8): p. 1059-1096.
91. Glass, C.K., et al., *Mechanisms underlying inflammation in neurodegeneration.* Cell, 2010. **140**(6): p. 918-934.
92. Teleanu, R.I., et al., *Neurotransmitters—Key Factors in Neurological and Neurodegenerative Disorders of the Central Nervous System.* International Journal of Molecular Sciences, 2022. **23**(11): p. 5954.
93. Gautam, R. and M. Sharma, *Prevalence and diagnosis of neurological disorders using different deep learning techniques: a meta-analysis.* Journal of medical systems, 2020. **44**(2): p. 49.
94. Palmer, A.M., *The role of the blood brain barrier in neurodegenerative disorders and their treatment.* Journal of Alzheimer's Disease, 2011. **24**(4): p. 643-656.
95. Heneka, M.T., J.J. Rodríguez, and A. Verkhratsky, *Neuroglia in neurodegeneration.* Brain research reviews, 2010. **63**(1-2): p. 189-211.
96. Wolfe, M.S., *The role of tau in neurodegenerative diseases and its potential as a therapeutic target.* Scientifica, 2012. **2012**.
97. Liao, Y.-Z., J. Ma, and J.-Z. Dou, *The role of TDP-43 in neurodegenerative disease.* Molecular Neurobiology, 2022. **59**(7): p. 4223-4241.
98. Xi, Z., et al., *Investigation of c9orf72 in 4 neurodegenerative disorders.* Archives of neurology, 2012. **69**(12): p. 1583-1590.
99. Soto, C., *Unfolding the role of protein misfolding in neurodegenerative diseases.* Nature Reviews Neuroscience, 2003. **4**(1): p. 49-60.
100. Yin, Y. and Z. Wang, *ApoE and neurodegenerative diseases in aging.* Aging and Aging-related diseases: Mechanisms and interventions, 2018: p. 77-92.
101. Noble, W., et al., *The importance of tau phosphorylation for neurodegenerative diseases.* Frontiers in neurology, 2013. **4**: p. 83.
102. Krizanac-Bengez, L., M.R. Mayberg, and D. Janigro, *The cerebral vasculature as a therapeutic target for neurological disorders and the role of shear stress in vascular homeostatis and pathophysiology.* Neurological research, 2004. **26**(8): p. 846-853.
103. Kanwar, J.R., B. Sriramoju, and R.K. Kanwar, *Neurological disorders and therapeutics targeted to surmount the blood–brain barrier.* International journal of nanomedicine, 2012: p. 3259-3278.

104. Castellani, R.J., et al., *Antioxidant protection and neurodegenerative disease: The role of amyloid-β and tau.* American Journal of Alzheimer's Disease & Other Dementias®, 2006. **21**(2): p. 126-130.
105. Stoessl, J., *Potential therapeutic targets for Parkinson's disease.* Expert opinion on therapeutic targets, 2008. **12**(4): p. 425-436.
106. Malik, R., et al., *Overview of therapeutic targets in management of dementia.* Biomedicine & Pharmacotherapy, 2022. **152**: p. 113168.
107. Islam, M., et al., *Therapeutic strategies for tauopathies and drug repurposing as a potential approach.* Biochemical Pharmacology, 2022. **198**: p. 114979.
108. Atif, M., et al., *Targeting adenosine receptors in neurological diseases.* Cellular Reprogramming, 2021. **23**(2): p. 57-72.
109. Langeh, U. and S. Singh, *Targeting S100B protein as a surrogate biomarker and its role in various neurological disorders.* Current Neuropharmacology, 2021. **19**(2): p. 265-277.
110. Beglopoulos, V. and J. Shen, *Gene-targeting technologies for the study of neurological disorders.* Neuromolecular medicine, 2004. **6**: p. 13-30.
111. Devara, D., Y. Choudhary, and S. Kumar, *Role of MicroRNA-502-3p in Human Diseases.* Pharmaceuticals, 2023. **16**(4): p. 532.
112. Altar, C.A., M.P. Vawter, and S.D. Ginsberg, *Target identification for CNS diseases by transcriptional profiling.* Neuropsychopharmacology, 2009. **34**(1): p. 18-54.
113. Zhang, S., et al., *CHIP as a therapeutic target for neurological diseases.* Cell death & disease, 2020. **11**(9): p. 727.
114. Kweon, H.-J. and B.-C. Suh, *Acid-sensing ion channels (ASICs): therapeutic targets for neurological diseases and their regulation.* BMB reports, 2013. **46**(6): p. 295.
115. Takahashi, K., J.B. Foster, and C.-L.G. Lin, *Glutamate transporter EAAT2: regulation, function, and potential as a therapeutic target for neurological and psychiatric disease.* Cellular and Molecular Life Sciences, 2015. **72**: p. 3489-3506.
116. Alavi, M.S., et al., *The potential of CYP46A1 as a novel therapeutic target for neurological disorders: An updated review of mechanisms.* European Journal of Pharmacology, 2023: p. 175726.
117. Olloquequi, J., et al., *Protein tyrosine phosphatase 1B (PTP1B) as a potential therapeutic target for neurological disorders.* Biomedicine & Pharmacotherapy, 2022. **155**: p. 113709.
118. He, W., et al., *The Potential of NLRP3 Inflammasome as a Therapeutic Target in Neurological Diseases.* Molecular Neurobiology, 2023. **60**(5): p. 2520-2538.
119. Marogianni, C., et al., *The role of C9orf72 in neurodegenerative disorders: a systematic review, an updated meta-analysis, and the creation of an online database.* Neurobiology of aging, 2019. **84**: p. 238. e25-238. e34.
120. Hyman, S.E., *Neurotransmitters.* Current biology, 2005. **15**(5): p. R154-R158.
121. Muir, J.L., *Acetylcholine, aging, and Alzheimer's disease.* Pharmacology Biochemistry and Behavior, 1997. **56**(4): p. 687-696.
122. Meltzer, H.Y., *Role of serotonin in depression.* Annals of the New York Academy of Sciences, 1990.
123. Huang, D., et al., *Mechanism of Rhizoma Coptidis in epilepsy with network pharmacology.* Allergologia et Immunopathologia, 2022. **50**(3): p. 138-150.
124. Bezhentsev, V., et al., *Identification of potential drug targets for treatment of refractory epilepsy using network pharmacology.* Journal of bioinformatics and computational biology, 2018. **16**(01): p. 1840002.

125. Jethalia, V., et al., *Analysing the role of Saraswatarishta in the treatment of neurological disorders based on network pharmacology.* Neuroscience Research Notes, 2021. **3**(5): p. 23-35.
126. Casas, A.I., et al., *From single drug targets to synergistic network pharmacology in ischemic stroke.* Proceedings of the National Academy of Sciences, 2019. **116**(14): p. 7129-7136.
127. Sharma, R., et al., *Role of Shankhpushpi (Convolvulus pluricaulis) in neurological disorders: An umbrella review covering evidence from ethnopharmacology to clinical studies.* Neuroscience & Biobehavioral Reviews, 2022: p. 104795.
128. Wang, S., et al., *Icariin, an up-and-coming bioactive compound against neurological diseases: network pharmacology-based study and literature review.* Drug design, development and therapy, 2021: p. 3619-3641.
129. Qu, S.-Y., et al., *Analysis of antidepressant activity of huang-lian jie-du decoction through network pharmacology and metabolomics.* Frontiers in pharmacology, 2021. **12**: p. 619288.
130. Li, Y., et al., *Anti-inflammatory and immunomodulatory effects of baicalin in cerebrovascular and neurological disorders.* Brain Research Bulletin, 2020. **164**: p. 314-324.
131. Mao, X.-Y., *Drug repurposing in neurological diseases: Opportunities and challenges.* Drug Repurposing-Hypothesis, Molecular Aspects and Therapeutic Applications, 2020: p. 137-144.
132. Shakkour, Z., et al., *Drug repurposing in neurological disorders: implications for neurotherapy in traumatic brain injury.* The Neuroscientist, 2021. **27**(6): p. 620-649.
133. Paranjpe, M.D., A. Taubes, and M. Sirota, *Insights into computational drug repurposing for neurodegenerative disease.* Trends in pharmacological sciences, 2019. **40**(8): p. 565-576.
134. Fagan, S.C., *Drug repurposing for drug development in stroke.* Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 2010. **30**(7P2): p. 51S-54S.
135. Lee, H.-M. and Y. Kim, *Drug repurposing is a new opportunity for developing drugs against neuropsychiatric disorders.* Schizophrenia research and treatment, 2016. **2016**.
136. Chen, R., et al., *Machine learning for drug-target interaction prediction.* Molecules, 2018. **23**(9): p. 2208.
137. Qin, L., et al., *Drug Repurposing for Newly Emerged Diseases via Network-based Inference on a Gene-disease-drug Network.* Molecular Informatics, 2022. **41**(9): p. 2200001.
138. Zhang, H., et al., *Learning multi-scale heterogenous network topologies and various pairwise attributes for drug–disease association prediction.* Briefings in Bioinformatics, 2022. **23**(2): p. bbac009.
139. D'Souza, S., K. Prema, and S. Balaji, *Machine learning models for drug–target interactions: current knowledge and future directions.* Drug Discovery Today, 2020. **25**(4): p. 748-756.
140. Redkar, S., et al., *A machine learning approach for drug-target interaction prediction using wrapper feature selection and class balancing.* Molecular informatics, 2020. **39**(5): p. 1900062.
141. Jamali, A.A., et al., *DrugMiner: comparative analysis of machine learning algorithms for prediction of potential druggable proteins.* Drug discovery today, 2016. **21**(5): p. 718-724.

142. Wang, W., et al., *Drug repositioning by integrating target information through a heterogeneous network model.* Bioinformatics, 2014. **30**(20): p. 2923-2930.
143. Kapetanovic, I., *Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach.* Chemico-biological interactions, 2008. **171**(2): p. 165-176.
144. Macalino, S.J.Y., et al., *Role of computer-aided drug design in modern drug discovery.* Archives of pharmacal research, 2015. **38**: p. 1686-1701.
145. Mering, C.v., et al., *STRING: a database of predicted functional associations between proteins.* Nucleic acids research, 2003. **31**(1): p. 258-261.
146. Chen, X., Z.L. Ji, and Y.Z. Chen, *TTD: therapeutic target database.* Nucleic acids research, 2002. **30**(1): p. 412-415.
147. Consortium, U., *UniProt: a hub for protein information.* Nucleic acids research, 2015. **43**(D1): p. D204-D212.
148. Ursu, O., et al., *DrugCentral: online drug compendium.* Nucleic acids research, 2016: p. gkw993.
149. Pihan, E., et al., *e-Drug3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design.* Bioinformatics, 2012. **28**(11): p. 1540-1541.
150. Gaulton, A., et al., *ChEMBL: a large-scale bioactivity database for drug discovery.* Nucleic acids research, 2012. **40**(D1): p. D1100-D1107.
151. Wishart, D.S., et al., *DrugBank: a comprehensive resource for in silico drug discovery and exploration.* Nucleic acids research, 2006. **34**(suppl_1): p. D668-D672.
152. Fornito, A., A. Zalesky, and M. Breakspear, *The connectomics of brain disorders.* Nature Reviews Neuroscience, 2015. **16**(3): p. 159-172.
153. Stovner, L., et al., *Neurological disorders in the G lobal B urden of D isease 2010 study.* Acta Neurologica Scandinavica, 2014. **129**: p. 1-6.
154. Kohl, M., S. Wiese, and B. Warscheid, *Cytoscape: software for visualization and analysis of biological networks.* Data mining in proteomics: from standards to applications, 2011: p. 291-303.
155. Xiao, N., et al., *protr/ProtrWeb: R package and web server for generating various numerical representation schemes of protein sequences.* Bioinformatics, 2015. **31**(11): p. 1857-1859.
156. Morgan, H.L., *The generation of a unique machine description for chemical structures-a technique developed at chemical abstracts service.* Journal of chemical documentation, 1965. **5**(2): p. 107-113.
157. Kim, S., et al., *PubChem in 2021: new data content and improved web interfaces.* Nucleic acids research, 2021. **49**(D1): p. D1388-D1395.
158. Landrum, G., *RDKit: open-source cheminformatics http://www. rdkit. org.* Google Scholar There is no corresponding record for this reference, 2016. **3**(8).
159. Zelle, J.M., *Python programming: an introduction to computer science*. 2004: Franklin, Beedle & Associates, Inc.
160. Massiris, M.M., et al. *Python implementation of local intervoxel-texture operators in neuroimaging using Anaconda and 3D Slicer environments*. in *2017 XLIII Latin American Computer Conference (CLEI)*. 2017. IEEE.
161. Zastre, M. *Jupyter notebook in CS1: An experience report*. in *Proceedings of the Western Canadian Conference on Computing Education*. 2019.
162. Giorgi, F.M., C. Ceraolo, and D. Mercatelli, *The R language: an engine for bioinformatics and data science.* Life, 2022. **12**(5): p. 648.

163. Racine, J.S., *RStudio: a platform-independent IDE for R and Sweave*. 2012, JSTOR.
164. Singh, D. and B. Singh, *Investigating the impact of data normalization on classification performance.* Applied Soft Computing, 2020. **97**: p. 105524.
165. Mangkunegara, I.S. and P. Purwono. *Analysis of DNA sequence classification using SVM model with Hyperparameter Tuning Grid Search CV*. in *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*. 2022. IEEE.
166. Agatonovic-Kustrin, S. and R. Beresford, *Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research.* Journal of pharmaceutical and biomedical analysis, 2000. **22**(5): p. 717-727.
167. Awad, M., et al., *Support vector machines for classification.* Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers, 2015: p. 39-66.
168. Song, Y.-Y. and L. Ying, *Decision tree methods: applications for classification and prediction.* Shanghai archives of psychiatry, 2015. **27**(2): p. 130.
169. Svetnik, V., et al., *Random forest: a classification and regression tool for compound classification and QSAR modeling.* Journal of chemical information and computer sciences, 2003. **43**(6): p. 1947-1958.
170. Natekin, A. and A. Knoll, *Gradient boosting machines, a tutorial.* Frontiers in neurorobotics, 2013. **7**: p. 21.
171. Singh, J. and R. Banerjee. *A study on single and multi-layer perceptron neural network*. in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. 2019. IEEE.
172. Regev, N., I. Yoffe, and D. Wulich. *Classification of single and multi propelled miniature drones using multilayer perceptron artificial neural network*. in *International Conference on Radar Systems (Radar 2017)*. 2017. IET.

7 Appendix

Table 8: *Target Sequences Descriptors*

Target	A	..	AA	...	AAA	...	CID	..	BHAR	..	CHA	...	CHO(	...	BIG(	...	DA'	...	hyc	...	norr	...	pol	...	cha	...	sec	...	sol	...	pro	...	VS1	...	Gra	..	Xc2	..	Pc1	...	Pc2.H
MT-NI	0	..	0	...	0	...	0	..	0.04	..	0	...	0.07	...	0	...	0	...	0	...	0.4	...	0	...	0	...	0	...	0	...	0	...	0	...	0	..	0	..	16	...	0
HSPG2	0	..	0	...	0	...	0	..	0.04	..	0	...	0.07	...	0	...	0	...	0	...	0.4	...	0	...	0	...	0	...	0	...	0	...	0	...	0	..	0	..	16	...	0
ASTN2	0	..	0	...	0	...	0	..	0.04	..	0	...	0.07	...	0	...	0	...	0	...	0.4	...	0	...	0	...	0	...	0	...	0	...	0	...	0	..	0	..	16	...	0
PRRT2	0	..	0	...	0	...	0	..	0.04	..	0	...	0.07	...	0	...	0	...	0	...	0.4	...	0	...	0	...	0	...	0	...	0	...	0	...	0	..	0	..	16	...	0
NPAS3	0	..	0	...	0	...	0	..	0.07	..	0.2	...	0.19	...	0.2	...	-0	...	0	...	0.5	...	0	...	0.1	...	0	...	0	...	0	...	1	...	0	..	0	..	340	...	-0
NDN	0	..	0	...	0	...	0	..	0.07	..	0.2	...	0.19	...	0.2	...	-0	...	0	...	0.5	...	0	...	0.1	...	0	...	0	...	0	...	1	...	0	..	0	..	340	...	-0
SLC6A	0	..	0	...	0	...	0	..	0.07	..	0.1	...	0.1	...	0.1	...	0	...	0	...	0.4	...	0	...	0.1	...	0	...	0	...	0	...	1	...	0	..	0	..	81	...	0
SNCA	0	..	0	...	0	...	0	..	0.19	..	0.3	...	0.22	...	0.3	...	-0	...	0	...	0.5	...	0	...	0.1	...	0	...	0	...	0	...	1	...	0	..	0	..	31	...	0
HTT	0	..	0	...	0	...	0	..	0.14	..	0.3	...	0.29	...	0.3	...	0	...	0	...	0.5	...	0	...	0.1	...	0	...	0	...	0	...	1	...	0	..	0	..	61	...	0
BDNF	0	..	0	...	0	...	0	..	0.05	..	0	...	0.01	...	0	...	0	...	0	...	0.4	...	0	...	0.1	...	1	...	0	...	0	...	0	...	0	..	0	..	34	...	-0
APOA	0	..	0	...	0	...	0	..	0.05	..	0	...	0.01	...	0	...	0	...	0	...	0.4	...	0	...	0.1	...	1	...	0	...	0	...	0	...	0	..	0	..	34	...	-0
C1QB	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
CAD	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
MARS	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
KCNM	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
PROS(	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
EPM2A	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
ITPA	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
POMC	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
SLC35A	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
CASR	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
APOE	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
APP	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
GRN	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0
TARDB	0	..	0	...	0	...	0	..	0.01	..	0.1	...	0.06	...	0	...	-0	...	0	...	0.4	...	0	...	0.1	...	0	...	1	...	0	...	0	...	0	..	0	..	38	...	0

Table 9 *Morgan Fingerprints of FDA-approved Drugs*

Drugs	Bit_1	Bit_2	Bit_3	…	Bit_13	Bit_14	Bit_15	…	Bit_2039	Bit_2040	Bit_2041	Bit_2047	Bit_2048
NADH	0	0	0	…	0	0	0	…	0	0	0	0	0
Metformin	0	0	0	…	0	0	0	…	0	0	0	0	0
Flavin adenine d	0	1	0	…	0	0	0	…	0	0	0	0	0
Ubidecarenone	0	0	0	…	0	0	0	…	0	0	0	0	0
Cyclosporine	0	1	0	…	0	0	0	…	0	0	1	0	0
Haloperidol	0	0	0	…	0	0	0	…	0	0	0	0	0
Fentanyl	0	0	0	…	0	0	0	…	0	0	0	0	0
Valrubicin	0	0	0	…	0	0	0	…	0	0	0	0	0
Oxygen	0	0	0	…	0	0	0	…	0	0	0	0	0
D-Tyrosine	0	1	0	…	0	0	0	…	0	0	0	0	0
Isotretinoin	0	0	0	…	0	0	0	…	0	0	0	0	0
Milnacipran	1	0	0	…	0	0	0	…	0	0	0	0	0
Citalopram	0	0	0	…	0	0	0	…	0	0	0	0	0
Clomipramine	0	0	0	…	0	0	0	…	0	0	0	0	0
Escitalopram	0	0	0	…	0	0	0	…	0	0	0	0	0
Fluvoxamine	0	0	0	…	0	1	0	…	0	0	0	0	0
Paroxetine	0	0	0	…	0	0	0	…	0	0	0	0	0
Sertraline	0	0	0	…	0	0	0	…	0	0	0	0	0
Fluoxetine	0	1	0	…	0	0	0	…	1	0	0	0	0
Desipramine	0	0	0	…	0	0	0	…	0	0	0	0	0
Duloxetine	0	1	0	…	0	0	1	…	1	0	0	0	0
Imipramine	0	0	0	…	0	0	0	…	0	0	0	0	0
Nortriptyline	0	0	0	…	0	0	0	…	0	0	0	0	0
Venlafaxine	0	1	1	…	0	0	0	…	0	0	0	0	0
Atomoxetine	0	1	0	…	0	0	0	…	1	0	0	0	0
Amitriptyline	0	0	0	…	0	0	0	…	0	0	0	0	0
Amoxapine	0	0	0	…	0	0	0	…	0	0	0	0	0
Desvenlafaxine	0	1	1	…	0	0	0	…	0	0	0	0	0

Table 10 *Target-Disease Network*

Targets	Targets_Sequences	No._of_Edges	Class_Label
MT-ND2	MNPLAQPVIYSTIFAGTLITALSSHWFFTWVGLEMNMLA	3	Neurodegenerative dis
HSPG2	MGWRAAGALLLALLLHGRLLAVTHGLRAYDGLSLPEDIET	**3**	Neurodegenerative dis
ASTN2	MAAAGARLSPGPGSGLRGRPRLCFHPGPPPLLPLLLLFLLL	3	Neurological disorder
PRRT2	MAASSSEISEMKGVEESPKVPGEGPGHSEAETGPPQVLA	3	Neurological disorder
NPAS3	MAPTKPSFQQDPSRRERITAQHPLPNQSECRKIYRYDGIY	3	Neurological disorder
NDN	MSEQSKDLSDPNFAAEAPNSEVHSSPGVSEGVPPSATLA	3	Neurological disorder
SLC6A4	METTPLNSQKQLSACEDGEDCQENGVLQKVVPTPGDKV	3	Neurological disorder
SNCA	MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGV	4	Neurodegenerative dis
HTT	MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQ	4	Neurodegenerative dis
BDNF	MTILFLTMVISYFGCMKAAPMKEANIRGQGGLAYPGVRT	4	Neurodegenerative dis
APOA1	MKAAVLTLAVLFLTGSQARHFWQQDEPPQSPWDRVKD	4	Neurodegenerative dis
C1QB	MMMKIPWGSIPVLMLLLLLGLIDISQAQLSCTGPPAIPGIP	**4**	Neurodegenerative dis
CAD	MAALVLEDGSVLRGQPFGAAVSTAGEVVFQTGMVGYP	4	Neurological disorder
MARS2	MLRTSVLRLLGRTGASRLSLLEDFGPRYYSSGSLSAGDDAC	4	Neurological disorder
KCNMB3	MDFSPSSELGFHFVAFILLTRHRTAFPASGKKRETDYSDGI	4	Neurological disorder
PROSC	MWRAGSMSAELGVGCALRAVNERTFGENYVQELLEKAS	4	Neurological disorder
EPM2A	MRFRFGVVVPPAVAGARPELLVVGSRPELGRWEPRGAV	4	Neurological disorder
ITPA	MAASLVGKKIVFVTGNAKKLEEVVQILGDKFPCTLVAQK	4	Neurological disorder
POMC	MPRSCCSRSGALLLALLLQASMEVRGWCLESSQCQDLTT	4	Neurological disorder
SLC35A2	MAAVGAGGSTAAPGPGAVSAGALEPGTASAAHRRLKYI	4	Neurological disorder
CASR	MAFYSCCWVLLALTWHTSAYGPDQRAQKKGDIILGGLFP	4	Neurological disorder
APOE	MKVLWAALLVTFLAGCQAKVEQAVETEPEPELRQQTEW	4	Neurodegenerative dis
APP	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMFCG	4	Neurodegenerative dis
GRN	MWTLVSWVALTAGLVAGTRCPDGQFCPVACCLDPGGA	4	Neurodegenerative dis
TARDBP	MSEYIRVTEDENDEPIEIPSEDDGTVLLSTVTAQFPGACGL	4	Neurodegenerative dis
MAPT	MAEPRQEFEVMEDHAGTYGLGDRKDQGGYTMHQDQE	4	Neurodegenerative dis
CTNND2	MFARKPPGAAPLGAMPVPDQPSSASEKTSSLSPGLNTSM	4	Neurological disorder
CACNA1H	MTEGARAADEVRVPLGAPPPGPAALVGASPESPGAPGP	4	Neurological disorder
NCF1	MGDTFIRHIALLGFEKRFVPSQHYVYMFLVKWQDLSEKV	4	Neurological disorder
CACNA1A	MARFGDEMPARYGGGGSGAAAGVVVGSGGGRGAGGS	4	Neurological disorder
OCA2	MHLEGRDGRRYPGAPAVELLQTSVPSGLAELVAGKRRLP	4	Neurological disorder
SLC1A1	MGKPARKGCEWKRFLKNNWVLLSTVAAVVLGITTGVLV	4	Neurological disorder
APCS	MNKPLLWISVLTSLLEAFAHTDLSGKVFVFPRESVTDHVN	5	Neurodegenerative dis
C1QC	MDVGPSSLPHLGLKLLLLLLLLPLRGQANTGCYGIPGMPG	**5**	Neurodegenerative dis
EEF1A2	MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEK	5	Neurological disorder
YWHAG	MVDREQLVQKARLAEQAERYDDMAAAMKNVTELNEPL	5	Neurological disorder
MEF2C	MGRKKIQITRIMDERNRQVTFTKRKFGLMKKAYELSVLCD	5	Neurological disorder
MT-ND1	MPMANLLLLIVPILIAMAFLMLTERKILGYMQLRKGPNVV	5	Neurological disorder
PARS2	MEGLLTRCRALPALATCSRQLSGYVPCRFHHCAPRRGRR	5	Neurological disorder
TNRC6A	MRELEAKATKDVERNLSRDLVQEEEQLMEEKKKKKDDKK	5	Neurological disorder

Table 11 *Disease-Drug Network*

Drugs	Drugs_SMILES	Status	Class_Label
NADH	C1C=CN(C=C1C(=O)N)[C@H]2[C@@H]([C@(	Approved	Neurodegenerative disorder
Metformin	CN(C)C(=N)N=C(N)N	Approved	Neurodegenerative disorder
Flavin adenine d	CC1=CC2=C(C=C1C)N(C3=NC(=O)NC(=O)C3=	Approved	Neurodegenerative disorder
Ubidecarenone	CC1=C(C(=O)C(=C(C1=O)OC)OC)C/C=C(\C)/C	Approved	Neurodegenerative disorder
Cyclosporine	CC[C@H]1C(=O)N(CC(=O)N([C@H](C(=O)N[	Approved	Neurodegenerative disorder
Haloperidol	C1CN(CCC1(C2=CC=C(C=C2)Cl)O)CCCC(=O)C	Approved	Neurodegenerative disorder
Fentanyl	CCC(=O)N(C1CCN(CC1)CCC2=CC=CC=C2)C3=	Approved	Neurological disorder
Valrubicin	CCCCC(=O)OCC(=O)[C@]1(C[C@@H](C2=C((	Approved	Neurological disorder
Oxygen	O=O	Approved	Neurological disorder
D-Tyrosine	C1=CC(=CC=C1C[C@H](C(=O)O)N)O	Approved	Neurological disorder
Isotretinoin	CC1=C(C(CCC1)(C)C)/C=C/C(=C/C=C/C(=C\C(	Approved	Neurological disorder
Milnacipran	CCN(CC)C(=O)[C@@]1(C[C@@H]1CN)C2=C(	Approved	Neurological disorder
Citalopram	CN(C)CCCC1(C2=C(CO1)C=C(C=C2)C#N)C3=C	Approved	Neurological disorder
Clomipramine	CN(C)CCCN1C2=CC=CC=C2CCC3=C1C=C(C=C3	Approved	Neurological disorder
Escitalopram	CN(C)CCC[C@@]1(C2=C(CO1)C=C(C=C2)C#N	Approved	Neurological disorder
Fluvoxamine	COCCCC/C(=N\OCCN)/C1=CC=C(C=C1)C(F)(F	Approved	Neurological disorder
Paroxetine	C1CNC[C@H]([C@@H]1C2=CC=C(C=C2)F)CC	Approved	Neurological disorder
Sertraline	CN[C@H]1CC[C@H](C2=CC=CC=C12)C3=CC(=	Approved	Neurological disorder
Fluoxetine	CNCCC(C1=CC=CC=C1)OC2=CC=C(C=C2)C(F)(	Approved	Neurological disorder
Desipramine	CNCCCN1C2=CC=CC=C2CCC3=CC=CC=C31	Approved	Neurological disorder
Duloxetine	CNCC[C@@H](C1=CC=CS1)OC2=CC=CC3=CC=	Approved	Neurological disorder
Imipramine	CN(C)CCCN1C2=CC=CC=C2CCC3=CC=CC=C31	Approved	Neurological disorder
Nortriptyline	CNCCC=C1C2=CC=CC=C2CCC3=CC=CC=C31	Approved	Neurological disorder
Venlafaxine	CN(C)CC(C1=CC=C(C=C1)OC)C2(CCCCC2)O	Approved	Neurological disorder
Atomoxetine	CC1=CC=CC=C1O[C@H](CCNC)C2=CC=CC=C2	Approved	Neurological disorder
Amitriptyline	CN(C)CCC=C1C2=CC=CC=C2CCC3=CC=CC=C3	Approved	Neurological disorder
Amoxapine	C1CN(CCN1)C2=NC3=CC=CC=C3OC4=C2C=C(	Approved	Neurological disorder
Desvenlafaxine	CN(C)CC(C1=CC=C(C=C1)O)C2(CCCCC2)O	Approved	Neurological disorder
Doxepin	CN(C)CC/C=C/1\C2=CC=CC=C2COC3=CC=CC=	Approved	Neurological disorder
Protriptyline	CNCCCC1C2=CC=CC=C2C=CC3=CC=CC=C13	Approved	Neurological disorder
Trimipramine	CC(CN1C2=CC=CC=C2CCC3=CC=CC=C31)CN((	Approved	Neurological disorder
Cocaine	CN1[C@H]2CC[C@@H]1[C@H]([C@H](C2)O	Approved	Neurological disorder
Phentermine	CC(C)(CC1=CC=CC=C1)N	Approved	Neurological disorder
Vilazodone	C1CN(CCN1CCCCC2=CNC3=C2C=C(C=C3)C#N	Approved	Neurological disorder
Mirtazapine	CN1CCN2C(C1)C3=CC=CC=C3CC4=C2N=CC=C	Approved	Neurological disorder
Lumateperone	CN1CCN2[C@H]3CCN(C[C@H]3C4=C2C1=CC	Approved	Neurological disorder
Vortioxetine	CC1=CC(=C(C=C1)SC2=CC=CC=C2N3CCNCC3)	Approved	Neurological disorder
Bupropion	CC(C(=O)C1=CC(=CC=C1)Cl)NC(C)(C)C	Approved	Neurological disorder
Risperidone	CC1=C(C(=O)N2CCCCC2=N1)CCN3CCC(CC3)C	Approved	Neurological disorder
Minaprine	CC1=CC(=NN=C1NCCN2CCOCC2)C3=CC=CC=	Approved	Neurological disorder

Table 12 *Target-Disease-Drug Network*

Targets	No._of_Edges	Diseases	Drugs	Status	Disorder_Class
MT-ND2	3	Alzheimers Disease	NADH	Approved	Neurodegenerative disorde
MT-ND2	3	Alzheimers Disease	Metformin	Approved	Neurodegenerative disorde
MT-ND2	3	Alzheimers Disease	Flavin adenine	Approved	Neurodegenerative disorde
MT-ND2	3	Alzheimers Disease	Ubidecarenone	Approved	Neurodegenerative disorde
HSPG2	3	Alzheimers Disease	Cyclosporine	Approved	Neurodegenerative disorde
HSPG2	3	Alzheimers Disease	Haloperidol	Approved	Neurodegenerative disorde
ASTN2	3	Migraine	Fentanyl	Approved	Neurological disorder
PRRT2	3	Migraine	Valrubicin	Approved	Neurological disorder
NPAS3	3	Psychotic disorder	Oxygen	Approved	Neurological disorder
NDN	3	Prader-Willi Syndrom	D-Tyrosine	Approved	Neurological disorder
NDN	3	Prader-Willi Syndrom	Isotretinoin	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Milnacipran	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Citalopram	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Clomipramine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Escitalopram	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Fluvoxamine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Paroxetine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Sertraline	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Fluoxetine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Desipramine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Duloxetine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Imipramine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Nortriptyline	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Venlafaxine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Atomoxetine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Amitriptyline	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Amoxapine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Desvenlafaxine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Doxepin	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Protriptyline	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Trimipramine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Cocaine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Phentermine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Vilazodone	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Mirtazapine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Lumateperone	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Vortioxetine	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Bupropion	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Risperidone	Approved	Neurological disorder
SLC6A4	3	Obsessive-compulsiv	Minaprine	Approved	Neurological disorder

Targets	No._of_Edges	Diseases	Drugs	Status	Disorder_Class
BDNF	4	Huntingtons disease	Vasopressin	Approved	Neurodegenerative disorde
BDNF	4	Huntingtons disease	Verteporfin	Approved	Neurodegenerative disorde
BDNF	4	Huntingtons disease	Vincristine	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Fenofibrate	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Copper	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Testosterone	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Zinc acetate	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Cholesterol	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Zinc cation	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Zinc chloride	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Zinc sulfate	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Infigratinib	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Ethanol	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Sirolimus	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Furosemide	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Glucagon	Approved	Neurodegenerative disorde
APOA1	4	Prion disease	Lamivudine	Approved	Neurodegenerative disorde
C1QB	4	Prion disease	Zinc cation	Approved	Neurodegenerative disorde
C1QB	4	Prion disease	Zinc acetate	Approved	Neurodegenerative disorde
C1QB	4	Prion disease	Zinc chloride	Approved	Neurodegenerative disorde
C1QB	4	Prion disease	Zinc sulfate	Approved	Neurodegenerative disorde
CAD	4	Epilepsy	L-Glutamine	Approved	Neurological disorder
CAD	4	Epilepsy	Aspartic acid	Approved	Neurological disorder
CAD	4	Epilepsy	Ammonia	Approved	Neurological disorder
CAD	4	Epilepsy	Famotidine	Approved	Neurological disorder
CAD	4	Epilepsy	Phosphoric aci	Approved	Neurological disorder
CAD	4	Epilepsy	Water	Approved	Neurological disorder
CAD	4	Epilepsy	Glutamic acid	Approved	Neurological disorder
MARS2	4	Epilepsy	Racemethionir	Approved	Neurological disorder
MARS2	4	Epilepsy	Phosphoric aci	Approved	Neurological disorder
MARS2	4	Epilepsy	Pyrophosphori	Approved	Neurological disorder
MARS2	4	Epilepsy	Selenomethio	Approved	Neurological disorder
KCNMB3	4	Epilepsy	Miconazole	Approved	Neurological disorder
KCNMB3	4	Epilepsy	Nitrendipine	Approved	Neurological disorder
KCNMB3	4	Epilepsy	Procaine	Approved	Neurological disorder
KCNMB3	4	Epilepsy	Ritodrine	Approved	Neurological disorder
KCNMB3	4	Epilepsy	Trimebutine	Approved	Neurological disorder
KCNMB3	4	Epilepsy	Potassium	Approved	Neurological disorder
PROSC	4	Epilepsy	Pyridoxal phos	Approved	Neurological disorder
EPM2A	4	Epilepsy	Chlorpromazin	Approved	Neurological disorder

Targets	No._of_Edges	Diseases	Drugs	Status	Disorder_Class
SLC6A3	6	Dementia	Metyrosine	Approved	Neurodegenerative disorder
SLC6A3	6	Dementia	Phenelzine	Approved	Neurodegenerative disorder
TOMM40	6	Dementia	1,2-Benzodiazep	Approved	Neurodegenerative disorder
TOMM40	6	Dementia	Cholesterol	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Carbetocin	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Sodium chloride	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Atosiban	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Oxytocin	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Desmopressin	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Estradiol	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Progesterone	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Testosterone	Approved	Neurodegenerative disorder
OXTR	6	Major Depressive D	Vasopressin	Approved	Neurodegenerative disorder
WBSCR22	6	Williams-Beuren sy	5-methyltetrahy	Approved	Neurological disorder
KCNK18	6	Migraine	Desflurane	Approved	Neurological disorder
KCNK18	6	Migraine	Enflurane	Approved	Neurological disorder
KCNK18	6	Migraine	Sevoflurane	Approved	Neurological disorder
KCNK18	6	Migraine	Isoflurane	Approved	Neurological disorder
KCNK18	6	Migraine	Halothane	Approved	Neurological disorder
YWHAZ	7	Parkinson's disease	Zolpidem	Approved	Neurodegenerative disorder
C3	7	Prion disease	Clozapine	Approved	Neurodegenerative disorder
C3	7	Prion disease	Zinc chloride	Approved	Neurodegenerative disorder
C3	7	Prion disease	Zinc sulfate	Approved	Neurodegenerative disorder
C3	7	Prion disease	Copper	Approved	Neurodegenerative disorder
C3	7	Prion disease	Human immuno	Approved	Neurodegenerative disorder
C3	7	Prion disease	Zinc acetate	Approved	Neurodegenerative disorder
C3	7	Prion disease	Zinc cation	Approved	Neurodegenerative disorder
ATIC	7	Epilepsy	Methotrexate	Approved	Neurological disorder
ATIC	7	Epilepsy	Pemetrexed	Approved	Neurological disorder
ATIC	7	Epilepsy	Water	Approved	Neurological disorder
CSTB	7	Epilepsy	Cupric Chloride	Approved	Neurological disorder
GRN	7	Frontotemporal Der	Tamoxifen	Approved	Neurodegenerative disorder
GRN	7	Frontotemporal Der	Estradiol	Approved	Neurodegenerative disorder
GRN	7	Frontotemporal Der	Cysteine	Approved	Neurodegenerative disorder
TARDBP	7	Frontotemporal Der	Bifonazole	Approved	Neurodegenerative disorder
TARDBP	7	Frontotemporal Der	Chenodeoxycho	Approved	Neurodegenerative disorder
TARDBP	7	Frontotemporal Der	Dexibuprofen	Approved	Neurodegenerative disorder
TARDBP	7	Frontotemporal Der	Diclofenac	Approved	Neurodegenerative disorder
TARDBP	7	Frontotemporal Der	Estrone	Approved	Neurodegenerative disorder
TARDBP	7	Frontotemporal Der	Fluconazole	Approved	Neurodegenerative disorder