

Discrimination Of Metronidazole Complexes



By

Muhammad Abdul Basit

Supervisor

Dr. Tahir Mehmood

Department of Mathematics and Statistics

School of Natural Sciences (SNS)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

August 2022


National University of Sciences & Technology

MS THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: MUHAMMAD ABDUL BASIT, Regn No. 00000330882 Titled: "Analysis of The Large-scale Manufacturing Growth of Pakistan" accepted in partial fulfillment of the requirements for the award of **MS** degree.

Examination Committee Members

1. Name: DR. MUDASSIR IQBAL

Signature: 

2. Name: DR. FIRDOS KHAN

Signature: 

Supervisor's Name: DR. TAHIR MEHMOOD

Signature: 


Head of Department

31/08/2022
Date

COUNTERSIGNED

Date: 31.8.2022


Dean/Principal

Abstract

Diabetic is a chronic human disease. It affects blood glucose in the human body. Imidazole ($C_3N_2H_4$) is an organic compound. Imidazole has many properties such as antituberculosis, antimicrobial anti-inflammatory, antimycotic, antitumor, and antiviral in the human body. By utilizing Fourier transform infrared spectroscopy (FTIR), recent review represent antidiabetic action of inadvance synthesized Metronidazole (*MTZ*) complexes (*MTZ-Benz*, *MTZ-Benz-Cu*, *MTZ-Benz-Cu-Cl₂CHCOOH*, *MTZ*, *MTZ - Cu*, *MTZ - Cu - Cl₂CHCOOH*, *MTZ - Benz - Ag*, *MTZ - Benz - Ag - Cl₂CHCOOH*, *MTZ - Ag* and *MTZ - Ag - Cl₂CHCOOH*) opposed to consequence of IC₅₀ ($\mu\text{g/mL}$) Values of substances for α -glucosidase and α -amylase inhibition assays. Two methods were used in this research named Artificial Neural Network and Partial Least Square. For correcting the spectrum baseline Asymmetric Least Squares were used. In fitted models, Artificial Neural Networks outperform cross-validated root mean square error. Influential wavelengths that explain variance in the antidiabetic activity of Metronidazole compounds were found and mapped against functional groupings using Artificial neural networks.

KEYWORDS: *Artificial neural networks, Antidiabetic, Alpha-glucosidase inhibition, Metronidazole complexes, Partial least squares*

Contents

1	Introduction	1
1.1	Diabetic	1
1.1.1	Type I Diabetes	3
1.1.2	Type II Diabetes	3
1.1.3	Prevention Of Type II Diabetes	4
1.2	Antidiabetic Activities	5
1.2.1	α -Amylase Inhibition Assay	5
1.2.2	α -glucosidase Inhibition Assay	6
1.3	Machine Learning	7
1.3.1	Definition	7
1.3.2	Background Of Machine Learning	8
1.3.3	Uses Of Machine Learning	9
1.3.4	Theory	9
1.4	Supervised And Unsupervised Learning	10
1.4.1	Supervised Learning Models	10
1.5	Objective Of Study	11
2	Literature Review	12
3	Methodology	17

CONTENTS

3.1	Antidiabetic Activity Of Metronidazole	17
3.2	Spectroscopic Experiment	17
3.2.1	Dispersive Infrared Spectrometer (IR)	18
3.2.2	Fourier Transform Infrared Spectrometers	18
3.3	Uses Of Fourier Transform Infrared Spectrometers	19
3.4	Collection Of Data And Statistical Software	19
3.5	Data Preprocessing	20
3.6	Asymmetric Least Squares (ALS)	21
3.7	Iterative Restricted Least Squares (IRLS)	22
3.8	Low-pass FFT Filter (LFF)	22
3.9	Artificial Neural Network	22
3.9.1	Artificial Neural Network And Human Brain	22
3.9.2	Artificial Neural Network Explanation	24
3.10	Artificial Neural Network Algorithm	28
3.11	ANN Model Architectures	36
3.11.1	MultiLayer Perceptron Model	36
3.11.2	Generalized Feedforward ANN Model	37
3.12	Partial Least Square Regression	37
3.12.1	PLS Regression Algorithm	38
3.12.2	Soft-Threshold PLS (St-PLS)	40
3.12.3	Distribution Based Truncation For Variable Selection In PLS (Tr-PLS)	40
4	Results And Discussion	41
4.1	Computations	51
4.2	Conclusions	51

List of Figures

1.1	(a).Difference between type I and type II diabetes.	5
1.2	(a).Supervised And Unsupervised Statistical Learning.	10
3.1	(a).Biological neuron.	23
3.2	(a). An ANN structure. (b). operating method of ANNs.	25
3.3	(a). ANN With One Hidden Layer.	28
3.4	(a). Activation Function.	29
3.5	(a). Binary step function.	30
3.6	(a). Linear Activation Function.	31
3.7	(a). Sigmoid Activation Function.	32
3.8	(a). Tanh Function.	33
3.9	(a). Rectified Linear Unit.	34
3.10	(a). Softmax Function.	35
4.1	Comparison Of Baseline Corrected Spectrum By Using ALS Method And Original Spectrum	42
4.2	Calibration of Alpha-amylase inhibition and Alpha-glucosidase inhibition	43
4.3	Validation of Alpha-amylase inhibition and Alpha-glucosidase inhibition	44

LIST OF FIGURES

4.4	density plot for paramters α -amylase inhibition α -glucosidase inhibition	46
4.5	The influential functional compound assignment and class for α -glucosidase inhibition	47
4.6	The influential functional compound assignment and class for α -amylase inhibition	48
4.7	Alpha amylase inhibition Network	49
4.8	Alpha glucosade inhibition Network	50

List of Tables

1.1	Results of the assay for substances S-1 through S-10's ability to inhibit α -amylase at various doses	6
1.2	Results of the assay for substances S-1 through S-10's ability to inhibit α -glucosidase at various doses	7
3.1	The IC ₅₀ ($\mu\text{g}/\text{mL}$) results Values of substances S-1 to S-10 for assays of α -amylase and α -glucosidase inhibition	18

List of Abbreviations and Symbols

Abbreviations

ANN	Artificial Neural Network
PLS	Partial Least squares
T2DM	Type II Diabetes Metallic
T1DM	Type I Diabetes Metallic
FTIR	Fourier Transform Infrared Spectrometers
ALS	Asymmetric Least Squares
ML	Machine Learning

CHAPTER 1

Introduction

1.1 Diabetic

Glucose plays a crucial role in biological systems as a key source of energy for cellular activity and metabolic intermediates. Diabetes and reduced glucose tolerance are just two of the many illnesses that are linked to improper glucose metabolism.

[1]

The pancreas produces insulin, which facilitates the uptake of glucose from food into your cells for use as fuel. The most prevalent metabolic condition linked to elevated blood sugar in people is diabetes mellitus. Type I diabetes is brought on by a lack of insulin, whereas type II diabetes is brought on by peripheral tissue resistance to insulin and a reduction in insulin secretion from Langerhans islets' beta cells (type II diabetes). [2]

Type I diabetes is assumed to be brought on by autoimmune mechanisms that kill the insulin-producing B cells in the pancreatic islets. Healthy B cells are mistakenly eliminated by active T cells. Autoimmunity is the cause of diabetes I in more than 95% of cases. Less than 5% of cases are idiopathic, and the speed at which pancreatic cells are destroyed varies greatly from person to person. Although it can affect anyone at any age, young children are more likely to contract it, with pre-school years seeing the highest rates of infection. [3]

In terms of clinical traits and etiology, Type I diabetes and type II diabetes are

becoming more difficult to distinguish from one another. There is a theory that both types of diabetes—rapid and slow—will eventually result in insulin resistance since both have beta cell inefficiency. Noninsulin-dependent diabetes formally known as type II diabetes affects people in their mature age of life. It manifests at an adult age. Genes play a part in the establishment of type II diabetes. It affects approximately 90–95% of adults with diabetes.

Any level of hyperinsulinemia at the outset or during pregnancy is referred to as gestational diabetes. There are two forms of GDM: A1GDM, which may be managed with food alone and responds to nutritional therapy, and A2GDM, which requires medication. To properly control the level of sugar in the blood, medicines are used to control the sugar. A problem with the pancreatic β cells or a sluggish effect of β cells' sugar level in the blood is the cause of gestational diabetes. Placental hormone secretion and insulin resistance as a lactogen Growth hormone, prolactin, and progesterone are other hormones linked to the development of gestational diabetes. All of these hormones contribute to elevated blood sugar and insulin resistance during pregnancy. Human is the main hormone linked to increased insulin resistance in this condition. [4]

Since 1985, the disease has increased dramatically, and by 2020, 515 million people will be living with the condition. The population of patients has grown sevenfold over the past 20 years, and if the current trend holds, with an annual rise of seven million patients, by 2030, there will be 585 million patients worldwide. [2]

A chronic condition known as Type II Diabetes Mellitus (T2DM) affects 30 million people in the US, or 9.4% of the total population. [5]

Type II diabetes, which is typically thought to affect people in their middle and older years, is now being discovered earlier in life. Given the rising frequency of the condition, type II diabetes's impediment related to it accounts for 8.4% of all fatalities globally and uses a large number of healthcare expenditures. [6]

1.1.1 Type I Diabetes

Most experts agree that Insulin-producing pancreatic β cells are destroyed as a result of an immune-related, although not directly immune-mediated, process that results in type I diabetes. Age at symptomatic onset was once viewed as a limiting factor for type I diabetes because it was traditionally thought of as an illness that mostly affected adolescents and children. However, this perception has altered over the past ten years. Type I diabetes is among the most prevalent chronic diseases in children, despite the fact that diagnosis is possible at any age. Peaks in presentation occur during or around puberty, between the ages of 5-7. Type I diabetes is somewhat more common in boys and men, but the majority of autoimmune illnesses disproportionately affect women. Type 1 diabetes incidence vary according to the birth month and season. Type 1 diabetes is more frequently discovered in fall and winter, and having been born in the spring is linked to a higher chance of developing the disease. The prevalence and incidence of type I diabetes vary greatly across the world. Finland and Sardinia have the highest rates of type I diabetes. The illness is less common in Venezuela, China, and India. Over the past few decades, type I diabetes has become more common everywhere. Finland, Germany, and Norway, respectively, have seen yearly increases in incidence of 24%, 26%, and 33% . Type I diabetes incidence rates have fluctuated in several nations, however, they have recently plateaued in Sweden. Over the following ten years, the global incidence could double if current rates of occurrence keep rising. Not all age groups have seen a rise in incidence at the same rate; in Europe, children under the age of five have had the biggest increases. [7]

1.1.2 Type II Diabetes

The most prevalent type of diabetes is type II. Worldwide, millions of people have received a Type II diabetes diagnosis, and many more go untreated. If diabetes is untreated or poorly managed, individuals are more likely to experience cardiovascular conditions including a heart attack and stroke. They also face increased chances of losing their eyesight, needing dialysis or a kidney transplant,

and losing their feet and legs owing to blood vessel and nerve damage. [8]

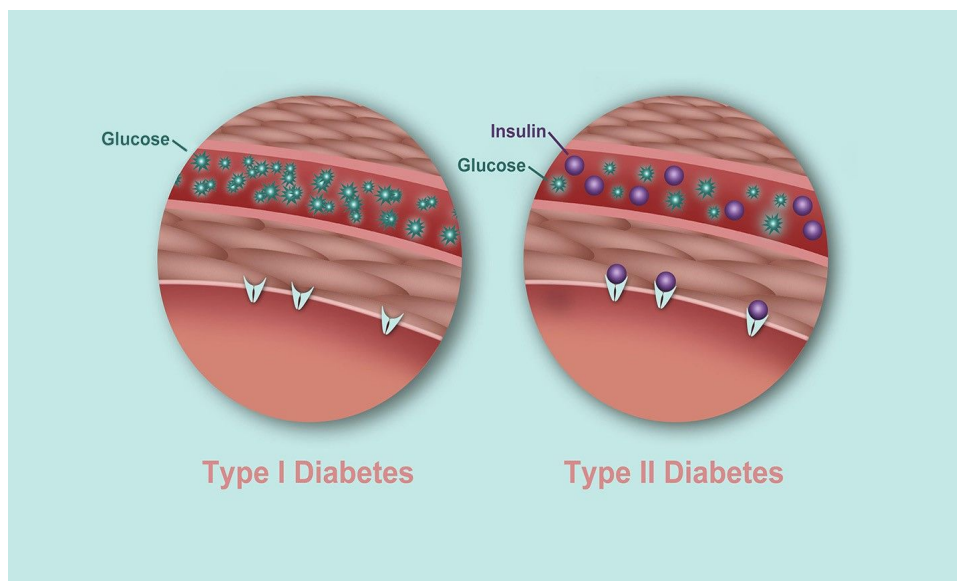
Multiple harmful health problems, such as cardiovascular disease, diabetic retinopathy, diabetic neuropathy, and diabetic nephropathy, can be brought on by T2DM. [9]

Prediabetes is defined by blood glucose levels increased but not complicated. on the other hand, diabetes is a condition in people almost always have before they acquire Type II diabetes. Modern research highlighted this the body may already be suffering some chronic effects of prediabetes, notably in the heart as well as the circulatory system. [10]

1.1.3 Prevention Of Type II Diabetes

People almost always have prediabetes before they acquire type II diabetes, which is defined as blood glucose levels normally increased than usual not now high complicated classified as diabetes. Prediabetes is a significant biological problem. It will be investigated. Scientists in the United States recently conducted a study that proved beyond a doubt that persons with prediabetes can prevent the onset of this disease by altering their diet as well as upping more exercise and other actions. It could even help to return the normal level of sugar in the blood.

The need for lifestyle modifications cannot be overstated. A healthy weight can be maintained, along with a longer lifespan and a lower risk of developing diabetes, by eating a balanced diet and engaging in more physical exercise. The Diabetes Prevention Program (DPP) findings demonstrated that Type II diabetes can be delayed or prevented by weight loss through sensible dietary adjustments and regular exercise. [11]

Figure 1.1: (a).Difference between type I and type II diabetes.

1.2 Antidiabetic Activities

The following α -glucosidase and α -amylase activity to used to determine the in vitro antidiabetic properties.

1.2.1 α -Amylase Inhibition Assay

To carry out this assay, the formerly disclosed technique was used. In the technique, 96-well microtiter transparent plates are employed. The following components were used in the following order to achieve α -amylase inhibitory activity: $40\mu\text{L}$ of 50 mM potassium phosphate buffer, after this $10\mu\text{L}$ of the sample at various concentrations, followed the addition of $10\mu\text{L}$ of α -amylase (0.12 U/mL) and $40\mu\text{L}$ of 0.05% starch. We utilized DMSO in negative power as well as acarbose in positive power. This mixture was incubated on temperature of 50°C for the duration of 30 minutes. To halt the reaction, $20\mu\text{L}$ of 1 M HCl was poured. At the conclusion of the experiment showed, $80\mu\text{L}$ of iodine reagent (5 mM KI and 5 mM I_2) was put in to examine whether starch was present or not. At 540 nm, the absorbance measured.

Formula for percentage (%) inhibitory activity

% inhibitory = $((AC-AS)/AC)*100$ Here **AC** absorbance of the control, and **AS** is the absorbance of the sample.

Table 1.1: Results of the assay for substances S-1 through S-10's ability to inhibit α -amylase at various doses

Sample Codes	S/No	800 μ g	400 μ g	200 μ g	100 μ g	50 μ g	25 μ g	12.5 μ g	6.25 μ g
MTZ Benz (S-1)	S-1	85.908	76.65	73.98	62.93	53.76	40.49	45.76	21.43
MTZ Benz-Cu (S-2)	S-2	87.012	81.11	74.89	64.24	55.81	42.21	37.23	23.11
MTZ Benz-Cu-Cl ₂ CHCOOH (S-3)	S-3	82.87	76.67	69.95	57.87	50.82	42.48	30.34	19.24
MTZ (S-4)	S-4	84.22	78.56	71.59	59.1	52.22	44.81	32.22	21.47
MTZ-Cu (S-5)	S-5	83.89	73.45	58.23	48.56	39.76	29.34	20.42	12.78
MTZ-Cu-Cl ₂ CHCOOH (S-6)	S-6	83.12	76.11	64.72	53.12	45.12	39.12	25.1	16.89
MTZ Benz-Ag (S-7)	S-7	80	75	61.5	49.4	37.2	40.2	26.2	15.4
MTZ Benz-Ag-Cl ₂ CHCOOH (S-8)	S-8	82	76.2	62.2	52.1	36.5	41.7	27.1	15.8
MTZ-Ag (S-9)	S-9	81	75.7	61.9	51.3	38.2	41.1	26.8	16.1
MTZ-Ag-Cl ₂ CHCOOH (S-10)	S-10	83	77	64	54.3	35.4	42.1	27.3	16.5

By utilizing the table curve methodology calculation of IC₅₀ was measured.

1.2.2 α -glucosidase Inhibition Assay

Formerly described methodology used for carry out this α -glucosidase inhibitory activity, with only minor alterations. The following chemicals were added to this reaction mixture in a 96-well plate in the following order: 25 μ L p-nitrophenyl β -D-glucopyranoside (p-NPG) (20 mM) were added, after this 69 μ L phosphate buffer (50 mM) put in, after it, the inclusion of pH 6.8 and 1 μ L α -glucosidase enzyme (3 U/mL). 5 μ L of material finally concentrated of 800, 400, 200, 100, 50, 25, 12.5 and 6.25 μ g/mL put in. Once more, DMSO was used for negative control and acarbose utilized for positive control. This reaction mixture, was formed in a 96 wells plate, was incubated at a temperature of 37°C for 30 minutes. With adding 100 μ L of NaHCO₃ (0.5 mM), this process stopped. By utilizing a microplate reader (BioTek Elx-800, USA), absorbance at 405 nm was measured.

Table 1.2: Results of the assay for substances S-1 through S-10's ability to inhibit α -glucosidase at various doses

<i>Sample Codes</i>	<i>S/No</i>	<i>800 μg</i>	<i>400μg</i>	<i>200μg</i>	<i>100μg</i>	<i>50μg</i>	<i>25μg</i>	<i>12.5μg</i>	<i>6.25μg</i>
MTZ Benz (S-1)	S-1	94.54	92.31	92.13	89.09	84.41	70.34	67.9	54.78
MTZ Benz-Cu (S-2)	S-2	\pm 96.22	\pm 95.01	93.81	90.24	86.02	72.14	69.23	65.1
MTZ Benz-Cu-Cl ₂ CHCOOH (S-3)	S-3	98.44	96.55	93.66	90.44	85.98	78.79	69.21	67.34
MTZ (S-4)	S-4	99.17	97.01	95.55	91.11	88.96	83.11	72.37	68.04
MTZ-Cu (S-5)	S-5	87.19	83.12	71.03	65.1	57.8	51.69	44.32	31.09
MTZ-Cu-Cl ₂ CHCOOH (S-6)	S-6	89.12	81.01	75	67.11	55.21	49.11	40.81	41.09
MTZ Benz-Ag (S-7)	S-7	88	80	74	72	58	50	39	40
MTZ Benz-Ag-Cl ₂ CHCOOH (S-8)	S-8	86	81	73	78	56	52	38	37
MTZ-Ag (S-9)	S-9	92	83	78	81	57.6	51	42	34
MTZ-Ag-Cl ₂ CHCOOH (S-10)	S-10	90	81.5	81	79.2	57.2	53	40	38

1.3 Machine Learning

1.3.1 Definition

The study of computer algorithms that improve over time as a result of experience and data is known as machine learning (ML). It consider a part related to Artificial Intelligence (AI). Without the requirement for explicit programming, to make predictions or judgments, machine learning algorithms construct a model using training data.

Information technology, statistics, probability, artificial intelligence, psychology, neuroscience, and many more fields all present a vast field called machine learning. With machine learning, issues can be easily resolved by creating a model that accurately represents a chosen dataset. By teaching computers to emulate the human brain, machine learning has expanded the study of statistics and transformed it into a comprehensive discipline that generates basic statistical computational theories of learning processes.

Making algorithms that let computers learn is the core of machine learning. Finding statistical patterns or other types of data patterns is learning process. Basic goal of creating machine learning algorithms was to be able to simulate how hu-

mans learn certain activities. These algorithms can also provide information on how challenging learning is in various situations.

Machine learning and Big Data computing technologies are being developed in a different way today than they were in the past. The capacity to automatically apply a range of complicated mathematical calculations to massive data has recently been established in machine learning, and this allows for a much faster calculation of the results. Today, numerous machine learning algorithms have been created, updated, and improved.

Programming using adaptation is fairly common. It is utilized in machine learning programs that have the ability to detect models, learning with previous error, abstract recent information by data, enhance accuracy as well as productiveness of this processing as well as result. more ever, multidimensional data are exit in different type of application domains is functioned in machine learning. [12]

1.3.2 Background Of Machine Learning

Pascal and Leibniz created machines that could mimic human addition and subtraction in the seventeenth century, which is when machine learning first emerged. The phrase "machine learning" was first used in contemporary times by IBM's Arthur Samuel, who also showed that computers could be taught to play checkers. Following this, Rosenblatt created the perceptron in 1958, one of the first neural network topologies. Werbos' creation of the multilayer perceptron (MLP) in 1975 was a major advance. Following this were the creation of support vector machines by Cortes and Vapnik and decision trees by Quinlan in 1986. Distributed multilayered learning algorithms have recently been developed under the umbrella of deep learning. When creating classifiers or other predictors, these algorithms can develop effective representations of the data that make it simpler to retrieve relevant information.

1.3.3 Uses Of Machine Learning

In AI research domain, machine learning has extreme significance. Although it is impossible to describe an intelligent system without the ability to learn, the former intelligent system typically lacks this ability. It has been applied to numerous branches of artificial intelligence (AI), including computer vision, natural language understanding, additional reasoning, intelligent robots, and models recognition. Several examples of particular applications include search engines, biological diagnosis, fraud detection of credit card, investigation of stock market, DNA sequencing sequences, handwriting and voice matching recognition, games method, as well as robotics utilization.

1.3.4 Theory

In the 1950s, the belief that human cognition could be demonstrated by machines gave rise to the term "artificial intelligence." Artificial intelligence (AI), according to Jerrold S. Maxmen, will bring in the twenty-first century.

ML which regarded as subclass of artificial intelligence (AI), illustrates the empirical acquisition associated with human intellect and can study and improve its evaluation using computer algorithms. With repetitions and algorithm changes, the computer may take input and estimate a result. Thus, to improve its capacity to forecast future events, ground truth is continuously altered while attempts are made to compare the results with the algorithm's outputs. [13]

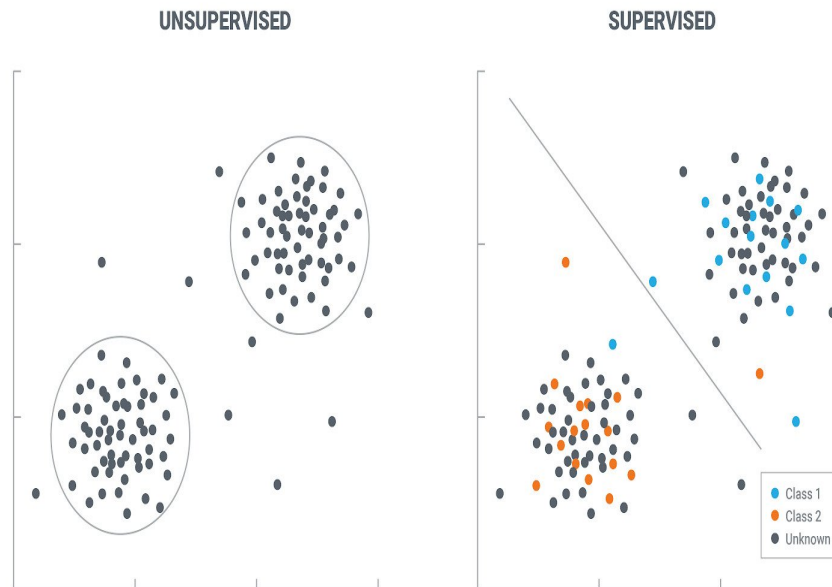
Data interpretation is supported by machine learning in a wide range of sectors, including science, medicine, the economy, policymaking, etc. Equations in mathematics, statistical analysis, and computer programming are used to implement machine learning. Machine learning is categories as under:

1. supervised machine learning, and
2. unsupervised machine learning.

1.4 Supervised And Unsupervised Learning

Numerous techniques for understanding data are referred to collectively as "statistical learning." These tools can be categorized into supervised or unsupervised. The general aim of supervised statistical learning is to build a statistical model for calculating or forecasting an output based on one or more inputs. A range of fields, including business, medicine, astrophysics, and policy making, are affected by this kind of problem. Unsupervised statistical learning has inputs but no supervised outputs, but we can nevertheless derive correlations and patterns from such data.

Figure 1.2: (a). Supervised And Unsupervised Statistical Learning.



1.4.1 Supervised Learning Models

Two types of supervised learning models are regression and classification models. Regression models are used when the response variable is made up of continuously actual values, for instance time, money, intensity, length, etc. Estimating the relationship between the numerical value data of an outcome variable and

several explanatory variables is useful. The categorization model, on the other hand, is a sort of guided learning in which the response variable is categorized like ‘child’ or ‘adult’, ‘True’ or ‘False’, ‘male’ or ‘female’, and binary values ‘0’ or ‘1’. Examples from the real world include test scorecard prediction, sentiment analysis, and light detection.

1.5 Objective Of Study

1. To discover functional groups for the analysis based on previously produced metronidazole compounds’ antidiabetic activity. The antidiabetic effect of metronidazole at 50% concentration levels against inhibitors of α -glucosidase and α -amylase was modelled using the FTIR spectrum.
2. To compare the performance of partial least squares (PLS) regression and artificial neural network models in the α -glucosidase and α -amylase inhibition of already synthesized Metronidazole complexes at 50% level of concentration.

CHAPTER 2

Literature Review

The examination of the literature is a means of organising to determine whether statistical approaches or gaps or vacuums on a certain issue need to be filled. It has been openly or methodically examined in the past. A thorough analysis of the research methods applied in the literature on a given issue is also provided in this chapter.

Karel Dieguez-Santana et al. (2019) reported the α -glucosidase inhibitor using LDA and decision tree.[14]

Karel DiŽguez-Santana et al. (2017) reported the α -glucosidase inhibitor using different machine learning technique. [15]

Singh and Lakshmiganthan (2017) used various algorithms on different dataset. Applied Naïve Bayesian, Random Forest (RF), KNN and used evaluation techniques like K-fold Cross-Validation. Using K-fold Cross-Validation, the method achieved 64.47% accuracy.[16]

Singh et al. (2017) employed Nave Bayes, function-based multilayer perceptrons, as well as decision tree-based random forests. To extract trustworthy and useful features from the dataset, the correlation approach was utilized for feature extraction. The author demonstrated that the Nave Bayes algorithm outperformed than random forest and multilayer perceptron.[17]

An adaptive neural network approach was utilized by Smith et al. (1988) to create associative models. It displayed a 76% accuracy with a random train-test split,

meaning that 576 randomly selected instances were trained and 192 instances were tested. [18]

For classification, Khashei et al. (2012) used ANN, support vector machines, K-nearest neighbour, quadratic discriminant analysis, as well as linear discriminant analysis and demonstrated 80% efficiency using a randomly train-test divide..[19]

For effective early prediction, Naz and Ahuja (2020) developed a multilayer feed forward network technique. The model's accuracy in analysing diabetes was 98.07%. [20]

A diabetes risk forecasting method by using a new upgraded deep neural networks approach was put forth by Zhou et al. (2020), and it is capable of both predicting and identifying future cases of the disease. [21]

With an accuracy of 97.5%, Alharbi and Alghahtani (2019) created a mixed model by utilizing a genetic algorithm (GA) as well as extreme learning machine algorithm. [22]

Rakesh Motka and Viral Parmar developed the comparative study for classifying the diabetic patient. [23]

To teach the network how to detect the illness pattern, Sapon et al. [24] used 250 diabetic patients, both sexes, ranging in age from 25 to 78.. The patients were both male and female. The Bayesian regulation algorithm outperformed the Broyden-Fletcher-Goldfarb-Shanno (BFGS), Quasi-Newton, as well as Levenberg-Marquardt algorithms in terms of diabetes prediction among three algorithms. In this study, the Bayesian regulatory algorithm demonstrated a strong correlation between obtained value and original value (i.e. 0.99579) with a precise forecasting of 88.8%, confirming the validation that this method is suitable for accurately predicting diabetes.

In 2012, Choubey et al. [25] used a genetic algorithm (GA) and naive Bayes ap-

proaches to forecast diabetes in women aged 21–78 years for the purpose of dataset classification. There were 768 incidents in total. First, the classification procedure for PIDD was carried out using naive Bayes, as well as a genetic algorithm (GM) was employed for addition and subtraction of variable from the sample. It boosted ROC and classification accuracy while comparatively reducing computing cost and time. The most accurate result and superior ROC when compared to other approaches are highlighted by the accuracy comparison of ROC, genetic algorithms, and naive bayes on the Pima Indian Diabetes Dataset.

Kayaer & Yildirim [26] used MLP neural networks, radial basis function (RBF), as well as general regression neural networks (GRNNs) on the Pima Indians Diabetes data in 2003. In comparison to the training data, the Levenberg-Marquardt training method has been shown to produce the best results. Even when using all of the underlying variables, RBF's accuracy was not superior to MLP's. Using test data, the GRNNs achieved the best outcome (i.e., 80.21 percent). This method demonstrates a superb and sensible choice in order to properly classify diabetic data.

In order to forecast the severity of diabetes in 700 randomly chosen instances, Florez et al. [27] used the software R in 2016. 2.952 was the MSE. The relatively high-risk groupings factors which contribute in diabetes were then identified. The original model includes overall number of pregnancies (PRG), salivary plasma concentration (PLASMA), blood pressure (BP), body mass index (BODY), as well as pedigree function for diabetes (PEDIGREE). According to their coefficients, PRG, PLASMA, BODY, as well as PEDIGREE have different effects on forecasting diabetes. The MSE for many factors has a higher likelihood of containing diabetes and is 3.21068. Therefore, PRG, PLASMA, BODY, and PEDIGREE are the last variables.

In 2018, Chawan [28] performed research with the goal of creating a system that can predict diabetes by combining the finding of various machine learning approaches to accurately and early diagnose diabetes in patients. The study uses SVM and Logistic Regression, to predict diabetes. They came to the conclusion

that SVM performed better than logistic regression.

In 2019, Kaur and Kumari [29] conducted studies using five distinct diabetes detection models which are multifactor dimensionality reduction (MDR), K-Nearest Neighbor (k-NN), linear kernel support vector machine (SVM-linear), as well as radial basis kernel support vector machine (SVM-RBF). With the use of the Boruta wrapper technique, the dataset's features were chosen. The observed measurements showed that Each of the methods performed well, with the SVM-linear model offering particularly strong precision and accuracy of 0.88 and 0.89, respectively. Based on the study's findings, it can be said that k-NN as well as linear kernel support vector machine (SVM-linear) are major diabetes forecasting methodologies with the highest accuracy.

Kumar [30] assessed the effectiveness of the machine learning algorithms for diabetes mellitus forecasting. Support vector machines, artificial neural networks, logistic regression, classification trees, as well as K-nearest neighbour are among algorithms employed by systems. The performance of the system is assessed using the evaluation methods: receiver operating characteristic (roc) curve, F1 measure, false positive rate, sensitivity, negative predictive value, and accuracy, specificity, sensitivity, as well as precision. The system employing Logistic Regression has the best performance of 78% and a rate of misinterpretation of 0.22%. Utilizing Naive Bayes as well as Logistic Regression, the greater accuracy and negative prediction are 82% and 73%, accordingly. The dataset is partitioned using tenfold cross-validation.

In order to classify diabetes in Iran, Heydari et al. [31] compared various classification systems. The artificial neural network, Bayesian network, decision tree, support vector machine, as well as closest neighbours are methods that have been utilises the system. The artificial neural network-based system which performed best with a 97.44 percent prediction performance. The accuracy of the support vector machine, 5-nearest neighbour, decision tree, as well as Bayesian network are 81.19, 90.85, 95.03, and 91.60 percent, respectively. The dataset for the system

consists of 2536 cases that were examined for type 2 diabetes.

A prediction framework for diabetes mellitus was put forth by Ashiquzzaman et al. [32] utilizing a deep learning methodology, with the dropout method being used to reduce overfitting. A dropout layer follows each of the two completely connected layers. A single node uses the output layer to find the conclusion. The system's maximum accuracy was 88.41% when it was used to analyze the Pima Indian diabetes dataset.

By employing ANN and PLS, a gap in the literature regarding the antidiabetic effect of metronidazole complexes was discovered.

Future studies on the use of ANN and PLS will take on a new perspective as a result of this study. In a similar vein, the contrast between the two models is also likely to be intriguing.

CHAPTER 3

Methodology

3.1 Antidiabetic Activity Of Metronidazole

The data consists of the various metronidazole complexes that were synthesized (*MTZ-Benz*, *MTZ-Benz-Cu*, *MTZ-Benz-Cu-Cl₂CHCOOH*, *MTZ*, *MTZ-Cu*, *MTZ-Cu-Cl₂CHCOOH*, *MTZ-Benz-Ag*, *MTZ-Benz-Ag-Cl₂CHCOOH*, *MTZ-Ag* and *MTZ-Ag-Cl₂CHCOOH*). This dataset is acquired from the study [33]. Here Response variable consists of values of α -glucosidase inhibition assay and α -amylase inhibition assay of the above-synthesized complexes at 50% concentration level. The data set of metronidazole complexes at a 50 % concentration level is provided in Table 4.1, along with the results of tests to see how well they inhibit α -amylase and α -glucosidase against it. For instance, at a 50% concentration level of metronidazole benzoate (*MTZ Benz*), α -glucosidase showed a 38.4, and alpha-amylase showed a 5.9 inhibition level.

3.2 Spectroscopic Experiment

There are two groups of infrared spectrometers.

- 1: dispersive infrared spectrometer (IR).
- 2: Fourier transform infrared spectrometers (FTIR).

Table 3.1: The IC₅₀ ($\mu\text{g}/\text{mL}$) results Values of substances S-1 to S-10 for assays of α -amylase and α -glucosidase inhibition

<i>Sample Codes</i>	<i>α-amylase inhibition</i>	<i>α-glucosidase inhibition</i>
<i>MTZ Benz (S-1)</i>	38.4	5.9
<i>MTZ Benz-Cu (S-2)</i>	44.8	4.8
<i>MTZ Benz-Cu-Cl₂CHCOOH (S-3)</i>	43.98	4.65
<i>MTZ (S-4)</i>	42.5	4.52
<i>MTZ-Cu (S-5)</i>	125.6	24.7
<i>MTZ-Cu-Cl₂CHCOOH (S-6)</i>	98.7	25.6
<i>MTZ Benz-Ag (S-7)</i>	65.2	7.22
<i>MTZ Benz-Ag-Cl₂CHCOOH (S-8)</i>	58.6	16.4
<i>MTZ-Ag (S-9)</i>	77.3	25.3
<i>MTZ-Ag-Cl₂CHCOOH (S-10)</i>	58.2	144

3.2.1 Dispersive Infrared Spectrometer (IR)

A grating type monochromator is used in a dispersive IR spectrometer to divide the light from a polychromatic source into various spectral constituents, which are then each individually detected by a detector. Because just a small percentage of the radiation is measured at a time due to this cumbersome sampling process, the signal intensity is quite low.

3.2.2 Fourier Transform Infrared Spectrometers

An interferometer is used in an FTIR spectrometer to produce an interferogram. When compared to dispersive instruments, FTIR instruments have faster sample rates and superior signal to noise ratios since all wavelengths are examined concurrently.

3.3 Uses Of Fourier Transform Infrared Spectrometers

like a method for the identification and quantification of organic substances, including chemical bonds, and organic content, Fourier Transform Infrared (FTIR) has been created (e.g., protein, carbohydrate, and lipid). This kind of analysis can be applied to characterise materials that are in the shape of liquids, solutions, pastes, powders, films, fibres, as well as gases. The FTIR analysis process involves exposing materials to infrared (IR) radiation. The sample's molecules' atomic vibrations are then affected by the IR radiations, which causes a specific energy to be absorbed and/or transmitted. As a result, the FTIR can be used to identify certain chemical vibrations present in the material.

Attenuated Total Reflection-Fourier Transformation Infrared Spectroscopy (ATR-FTIR) was utilized to perform ultraviolet spectrum tests within mid-IR (4000-550 cm^{-1} wavenumber) range. Attenuated Total Reflection-Fourier Transformation Infrared spectroscopy (ATR-FTIR) is a form of spectroscopy in which infrared light is injected into a prism at an angle greater than the critical angle for internal reflection. It utilizes a typical high linearity room temperature sensor (standard) as well as a UATR Diamond ATR (Single Reflection). [34]

3.4 Collection Of Data And Statistical Software

The data collection experiments are carried out in the School of Natural Sciences(SNS), National University of Sciences and Technology (NUST), Islamabad, Pakistan. By using R all calculations as well as analysis are performed. <https://www.R-project.org/>. Calculations are done using the R-package 'neuralnet'.

3.5 Data Preprocessing

Any acquired spectra that are measured by infrared spectrometry or other methods may include the intended signal as well as undesired components like noise and background. As a result, eliminating the baseline from the spectra is a crucial step in doing either a qualitative or quantitative study of the spectra. [35]

Baseline correction is a component of data preparation. It is expected that the zero-point baseline of FTIR spectra will be a smooth line, although this is frequently not the case due to linear or nonlinear additions. Tuning parameters exist for baseline correction procedures. Here, the objective process for selecting baseline correction procedure [23] is chosen rather than the subjective approaches. According to the algorithm,

1. The levels of all the parameters that will be examined are chosen for each baseline correction procedure.
2. The level baseline is adjusted for each parameter using the corresponding algorithm.
3. Through the use of PLS, corrected baseline spectral data are modelled with antibiotic activity responses and their validated predictive accuracy is evaluated.
4. The most effective parameter levels are chosen for every baseline correction algorithm.
5. The most effective baseline correction algorithm with the greatest capacity for prediction is chosen.

For prediction capability, the cross-validated root mean squared error (RMSE) is commonly utilized. Here is a quick explanation of the three probable baseline algorithms that were taken into consideration:

1. Asymmetric least squares (ALS).

2. Iterative Restricted Least Squares (IRLS), and
3. Lowpass FFT filter (LFF).

3.6 Asymmetric Least Squares (ALS)

In the ALS method put out by Eilers and Boelen, a smooth line (b) lower than the experimental data (X) is sought after by penalising the first derivative of the latter. Instrument technique baseline issues can be identified by a superimposed signal made up of a string of signal peaks that are either all positive or all negative. To obtain a baseline estimate, this method employs a smoother with an asymmetric weighting of deviations.

By placing a limit on the 2^{nd} derivative, the asymmetric least squares (ALS) [36] technique smoothes predictor variables with significant error by using the least square method.

$$Y = \sum W_i (X_i - b_i)^2 + \lambda \sum (\Delta^2 b_i)$$

where b_i estimated baseline, X_i is original spectrum, W_i are asymmetric residual weights, Δ^2 is the second derivative of estimated baseline. The algorithm has two parameters to tune: λ is the smoothing parameter and W is the weight. where W_i is the weighting vector, is the penalty parameter for the second derivative, and the total is over all x-axis points. A basic ALS strategy involves changing the W_i values to p if $X_i > b_i$ and to $1-p$ otherwise in order to induce the asymmetry with a parameter p ($0 < p < 1$).

One issue with the ALS method is that the $X_i - b_i$ values under the peaks can grow to be so enormous as to produce inflections in the projected baseline b_i . There have been various proposed improvements to the algorithm to enhance the performance of the ALS technique. The proposal to multiply the weights p by $\exp(-X_i - b_i/k)$ for $X_i > b_i$, where k is proportional to the distinctive peak height, proved to be the most successful. In this approach, the amplitude devalues the peak influence. The objective of the current work is to provide a peak screening technique that is more effective and can improve baseline estimation quantitatively.

The first and second derivatives of the original smoothed spectrum serve as the foundation for this screening \mathbf{X}_i . The random smooth lines were added to the peaks in order to measure the performance of various methods, and the difference between the predicted and estimated baselines was measured. [37]

3.7 Iterative Restricted Least Squares (IRLS)

Iterative Restricted Least Squares (IRLS) [38], which has regular baseline suppression and regressions with a second derivative limit, is the main smoothing technique. Here $S = f(X_i, b_i, \lambda_1, \lambda_2, p_i)$, Here function ‘ S ’ consists of five variables,

where X_i is the value of original spectrum , b_i is the estimated baseline, λ_1 is the second derivative constraint for primary smoothing, λ_2 is the second derivative constraint for secondary smoothing and p is the residual weights.

3.8 Low-pass FFT Filter (LFF)

Fast Fourier Transform filtering is the foundation of the low-pass FFT filter (LFF) [39], a technique for eliminating baselines. Here $S = f(X_i, b_i, S, H)$, where X_i is original spectrum , b_i estimated baseline, S is steepness of filter curve, and H is half-way point of filter curve.

3.9 Artificial Neural Network

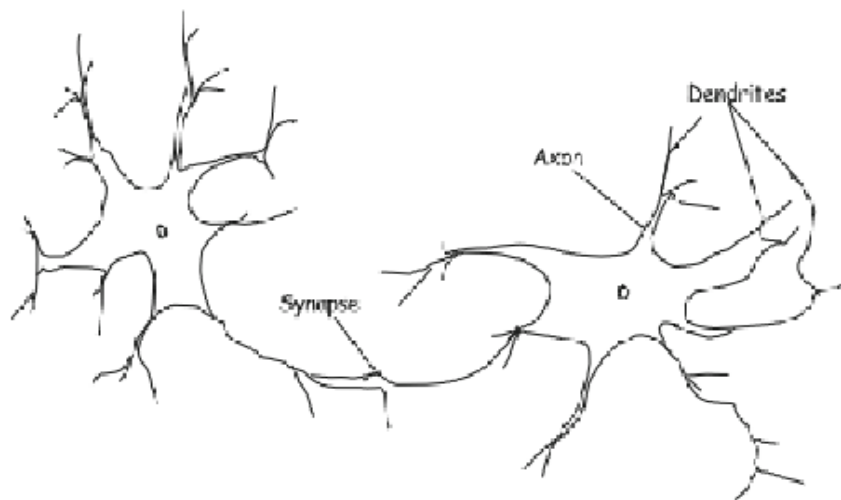
3.9.1 Artificial Neural Network And Human Brain

In layman’s terms, a computational model called an Artificial Neural Network (ANN) with biological influences that consists of processing units (known as neurons) and connections between them with coefficients (weights) connected to the connections. The neural structure is made out of these connections, and training

and recall algorithms are coupled to this structure. Because of the connections established between the neurons, neural networks are known as connectionist models.

In his description of neural networks (NNs), Deboeck and Kohonen [40] used the terms non-linear, multi-layered, parallel regression approaches to refer to a group of mathematical techniques that can be applied to signal processing, forecasting, and clustering. Neural network modelling is described as being similar to fitting a line, plane, or hyperplane across a set of data points. Any data set can be fitted with a line, plane, or hyper plane to establish a representation of the data on a smaller scale or to indicate any potential links between the user-selected inputs and outputs.

Figure 3.1: (a).Biological neuron.



However, the following are the primary qualities that are thought of and described as shared roles in both natural and man-made networks:

1. Adaptation and learning;
2. Generalization
3. Significant parallelism
4. Robustness.

5. Associative storage of information.
6. Processing of spatial and temporal information.

3.9.2 Artificial Neural Network Explanation

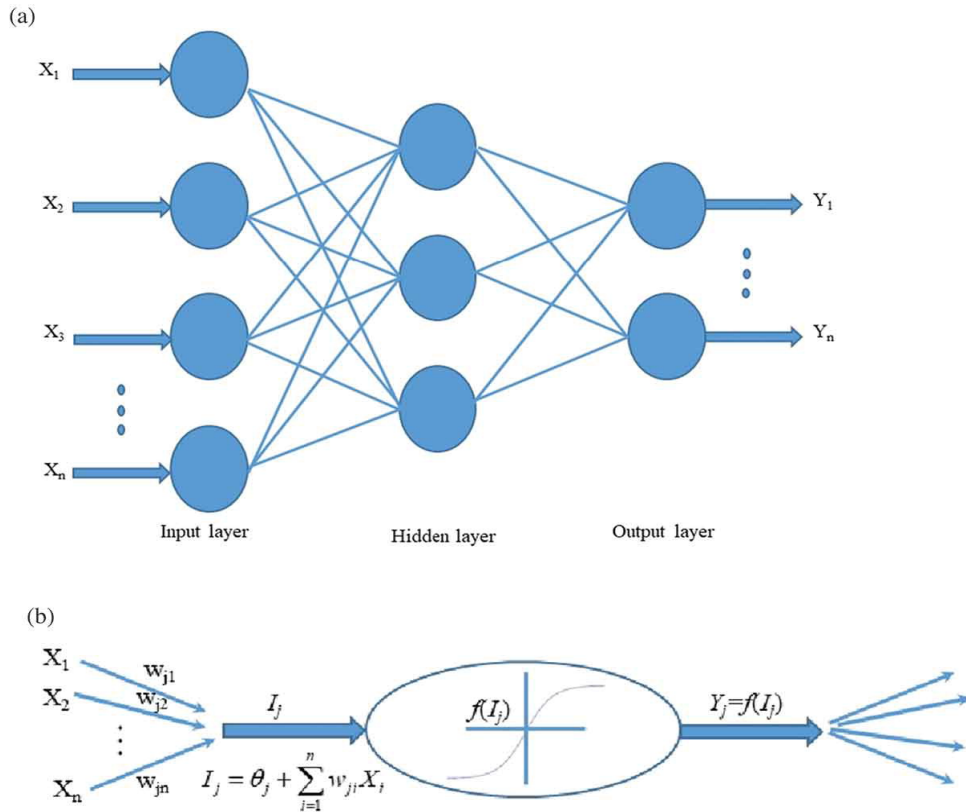
Artificial neural networks (ANN), often known as neural networks, are models that draw inspiration from the human brain and have the potential to explain complicated issues. They are also the technology that is most widely used in the medical sciences to calculate binary responses. Because of their internal control, flexibility, and ease of use, neural network models are distribution-free data mining techniques that are successfully employed in a variety of study areas, including prediction, pattern recognition, classification, clustering, forecasting, and so forth. ANNs are computer architectures that mimic the structure of the human brain. It is constructed by several "neurons" (also known as "nodes") that are arranged in layers. [41]

The ANN has been extensively used in contemporary epidemiological research, particularly as a prediction tool, for the same objective. [42]

The potential for prospective ignition sensing is identified by the perception of NN models, which may improve model performance and raise estimate accuracy. [43]

The artificial neural network (ANN) model mimics human brain functions using computation and mathematics. A computational model called an artificial neural network (ANN) is made up of numerous processing elements that accept inputs and give responses based on their activation functions. [44]

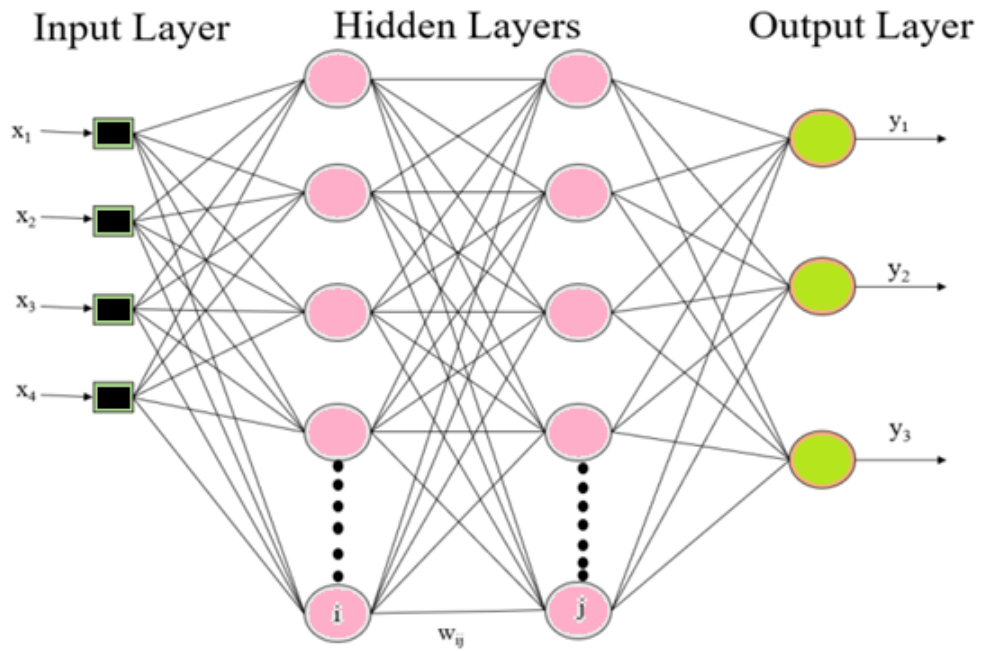
Figure 3.2: (a). An ANN structure. (b). operating method of ANNs.



$$Y_i = f(\theta_i + \sum_{j=1}^n w_{ij} X_j)$$

where: Y_i is output, f is the transfer function, θ_i is the bias, n is the number of neurons, w_{ij} is the connection weights and X_j is the input variable. [45]

The connections that have been built between the various processing units and the pertinent parameters within the neural network architecture are what drive the global behavior of these neurons. Each link connecting the neurons in successive layers has a weight. The weight w_{ij} in the network represents the strength of the link between the i th neuron in one layer and the j th neuron in the following layer. A neural network's structure consists of one "input" layer, one or more "hidden" layers, and one "output" layer. The complexity of the system under consideration determines how many hidden layers there are and how many neurons there are in each of these layers. A typical ANN architecture with two hidden layers is shown in Figure 3-1.



1. According to the equation below, a weighted sum is first calculated, and then the bias term β_o is added to this sum:

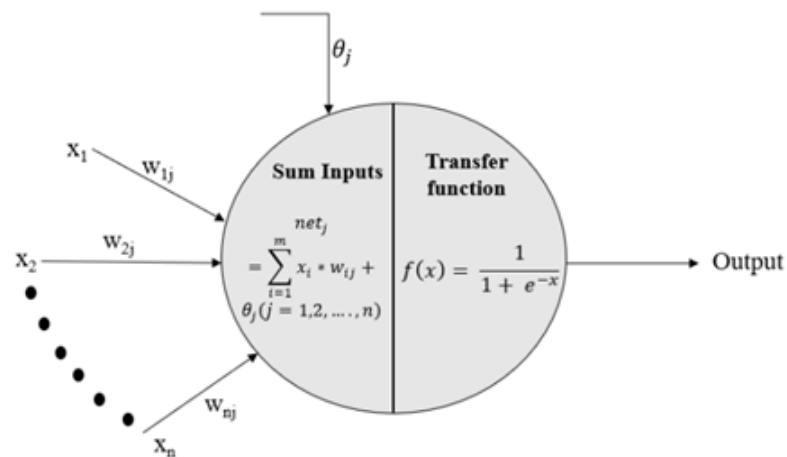
$$(3.9.1)$$

$$Y = \beta_o + \sum_{j=0}^n W_j X_j$$

2. A mathematical "transfer function" is used to transform Y . All network inputs and outputs are normalized using this function to a small range of values. As opposed to, say, having a range of values between 50 and 500, this enables the neural network to spot patterns and trends in the data more effectively. This can be done using a variety of transfer functions. The following equation, however, illustrates how a sigmoid function is employed in this attempt for the reasons outlined in section "Model Development":

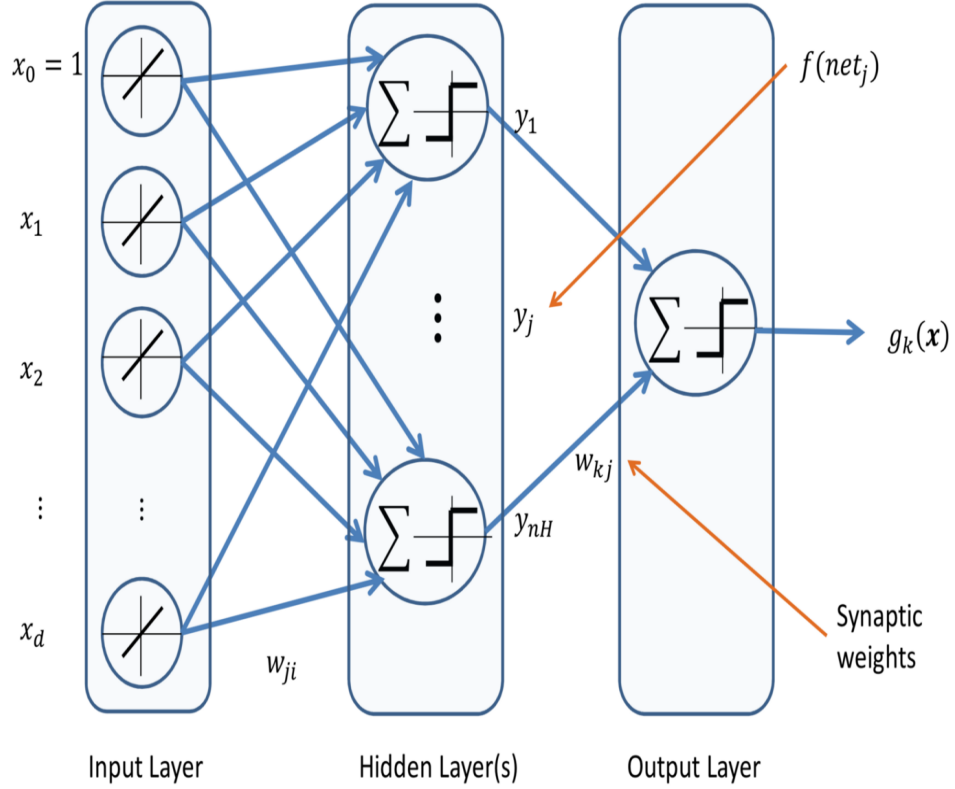
$$f(x) = 1/(1 + e^{-x}) \quad (3.9.2)$$

3. The neuron in the subsequent layer receives the outcome in the end. After creating a neural network for a specific application, training must be carried out using the initial weights, which were picked at random. These models are well-liked because they are adaptable and learn by identifying patterns in the incoming data. There are two methods for training a network: supervised training and unsupervised training. The desired output must be provided along with the inputs during supervised training to maximize the network weights and identify the optimal set of weights that results in the output with the least amount of error. Function approximation, regression analysis, time series prediction, and classification tasks like pattern and sequence recognition are some of the areas where models with supervised training can be used. Unsupervised training, on the other hand, is useful when a model must interpret the inputs on its own and defines the structure of "unlabeled" data, or data that has not been classified. Applications of unsupervised learning include, but are not limited to, clustering and anomaly detection .[46]



3.10 Artificial Neural Network Algorithm

Figure 3.3: (a). ANN With One Hidden Layer.



1. Feed-forward neural network is:

$$net_j = \sum_{i=0}^d X_i W_{ij} + \beta_{jo} = \sum_{i=0}^d X_i W_{ij}$$

$$Y_j = f(net_j)$$

similarly we find net_k as

$$net_k = \sum_{j=1}^{nH} Y_j W_{kj} + \beta_{ko}$$

So total output Z_k can be find as

$$Z_k = f(net_k)$$

Hence one formula for finding output of neural network for one hidden layer is

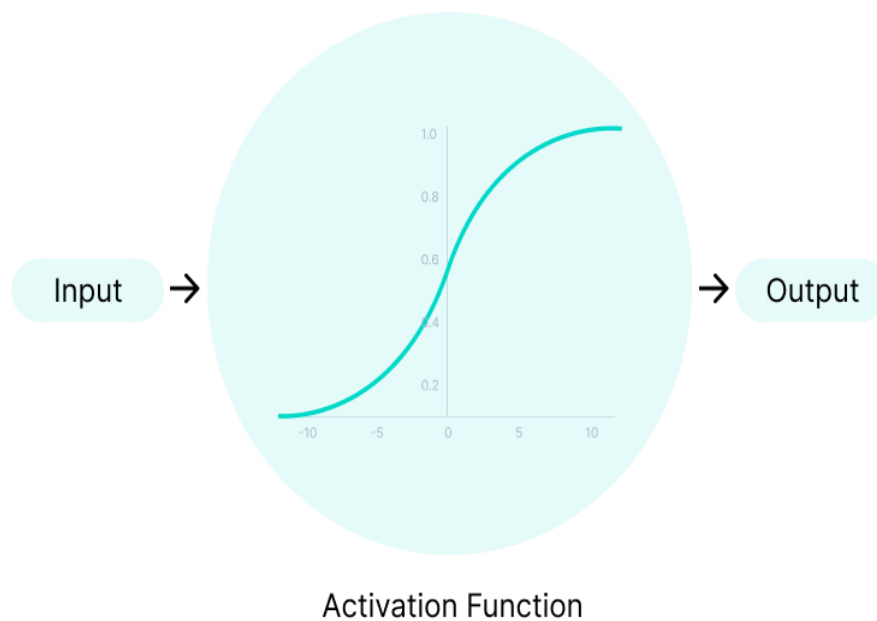
$$Z_k = f\left(\sum_{j=1}^{nH} W_{kj} f\left(\sum_{i=0}^d X_i W_{ij} + \beta_{jo}\right) + \beta_{ko}\right)$$

Here f denote the activation function, nH denote the number of perceptrons in the hidden layers and β_o are the bias in the neural network. Some of the activation function illustrated below.

2. *Activation function is:*

Whether a neuron should be activated or not is determined by an activation function. This means that it will determine, using more straightforward mathematical processes, whether or not the neuron's input to the network is significant throughout the prediction process. The main purpose of activation function is that provide non linearity to the neural network.

Figure 3.4: (a). Activation Function.



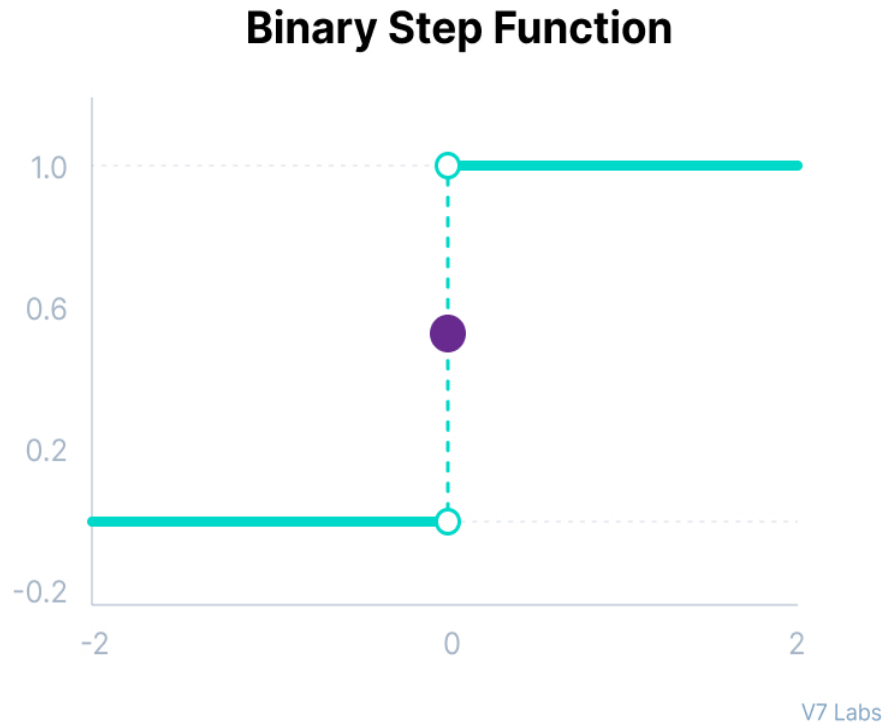
V7 Labs

3. *Types of neural network activation function:*

- (a) *Binary step function:* The binary step function is depend on the threshold that decides whether or not a neuron should be triggered. The neuron is activated if the input exceeds a threshold; if not, it is deacti-

vated, which inhibits its output from being transmitted to the following hidden layer.

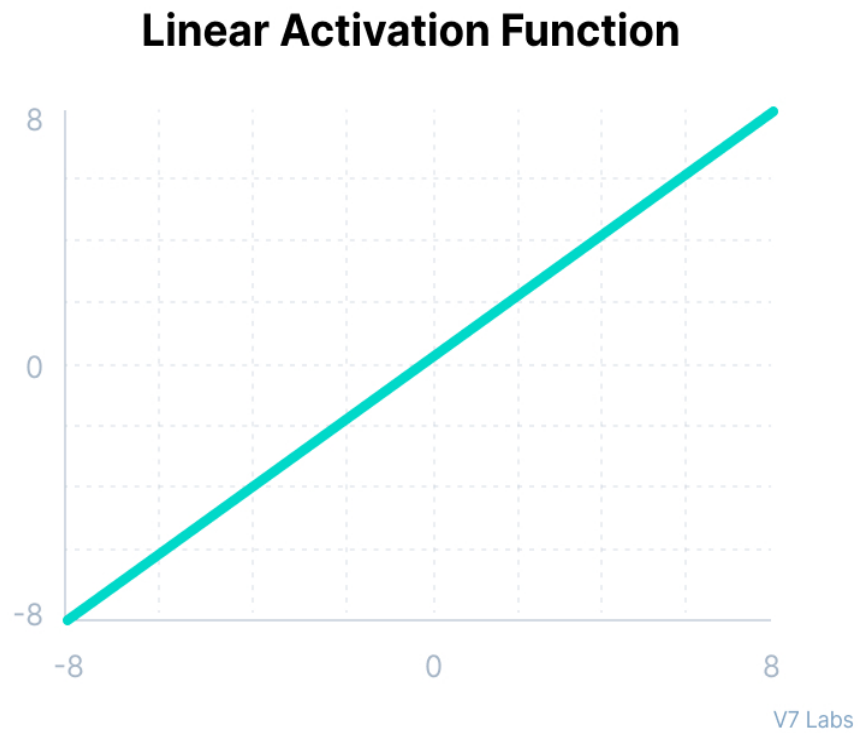
Figure 3.5: (a). Binary step function.



It has the following mathematical representation:

$$Y = \begin{cases} 0, & \text{if } X < 0 \\ 1, & \text{if } X \geq 0 \end{cases} \quad (3.10.1)$$

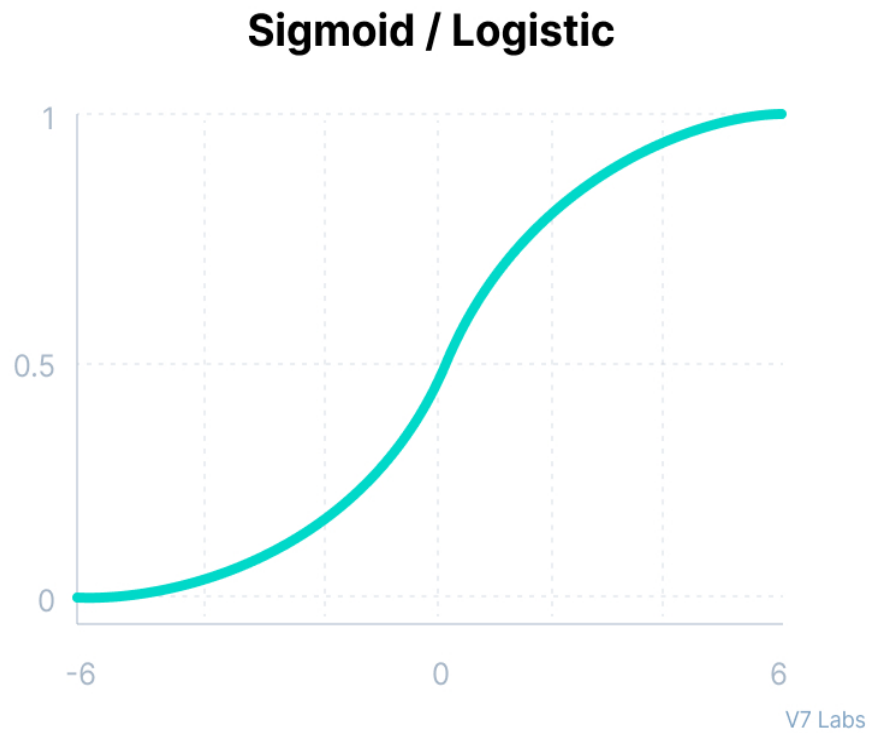
- (b) *Linear activation function*: The activation is proportionate to the input in a linear activation function, also referred to as "no activation" or the "identity function". The function just spits out the value it was given, doing nothing to the weighted sum of the input.

Figure 3.6: (a). Linear Activation Function.

It has the following mathematical representation:

$$Y = \begin{cases} X \end{cases} \quad (3.10.2)$$

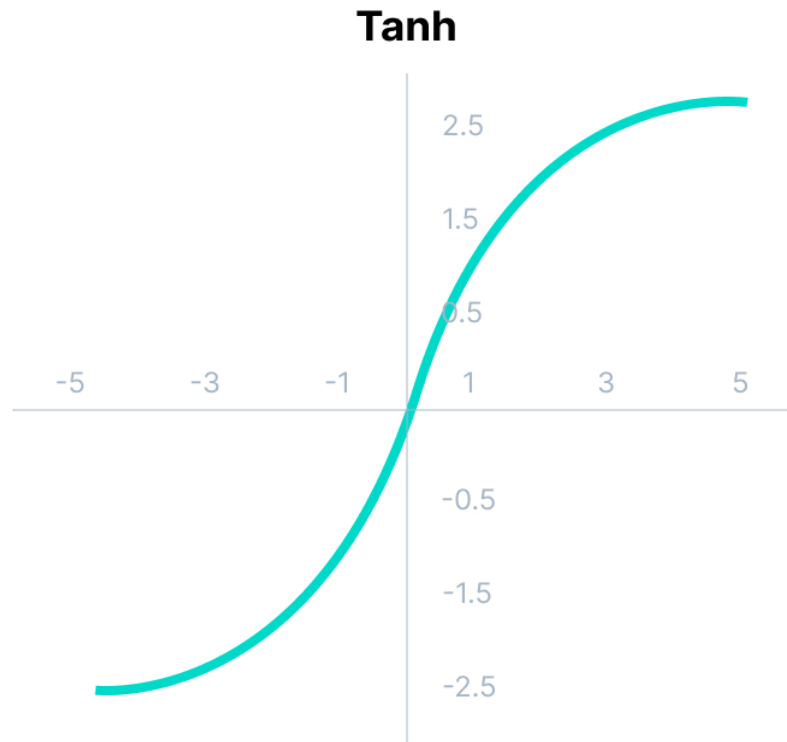
- (c) *Sigmoid/logistic activation function*: Any real value may be used as an input for this function, which returns values between 0 and 1.

Figure 3.7: (a). Sigmoid Activation Function.

It has the following mathematical representation:

$$Y = \left\{ \frac{1}{1 + e^{-X}} \right. \quad (3.10.3)$$

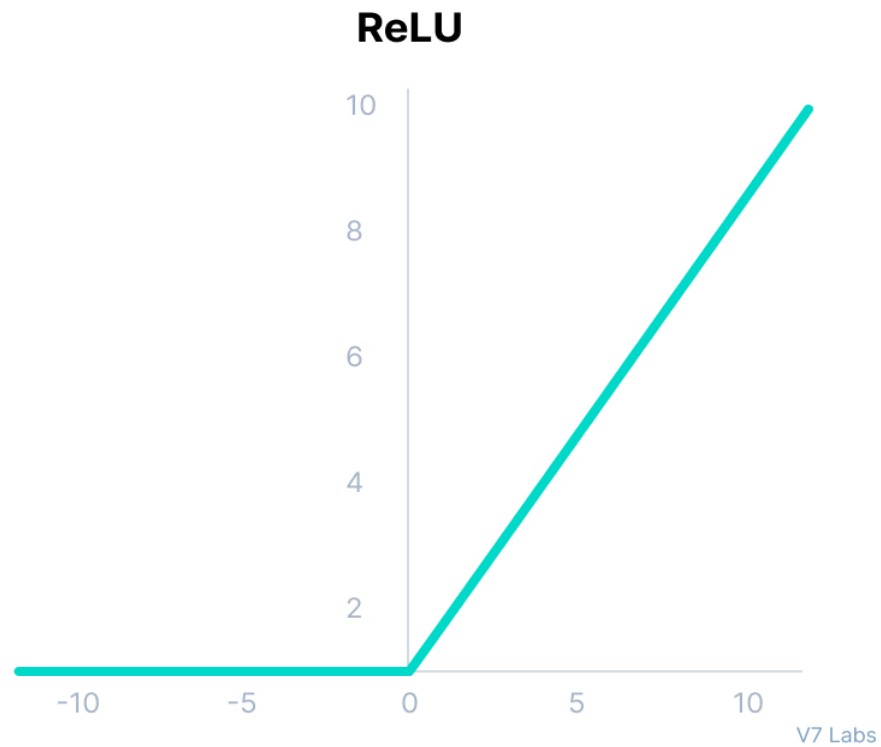
- (d) *Tanh function*: With a variation in output range of -1 to 1, the sigmoid/logistic activation function and the tanh function are strikingly similar, and they even have the same S-shape.

Figure 3.8: (a). Tanh Function.

It has the following mathematical representation:

$$Y = \left\{ \frac{(e^X - e^{-X})}{(e^X + e^{-X})} \right. \quad (3.10.4)$$

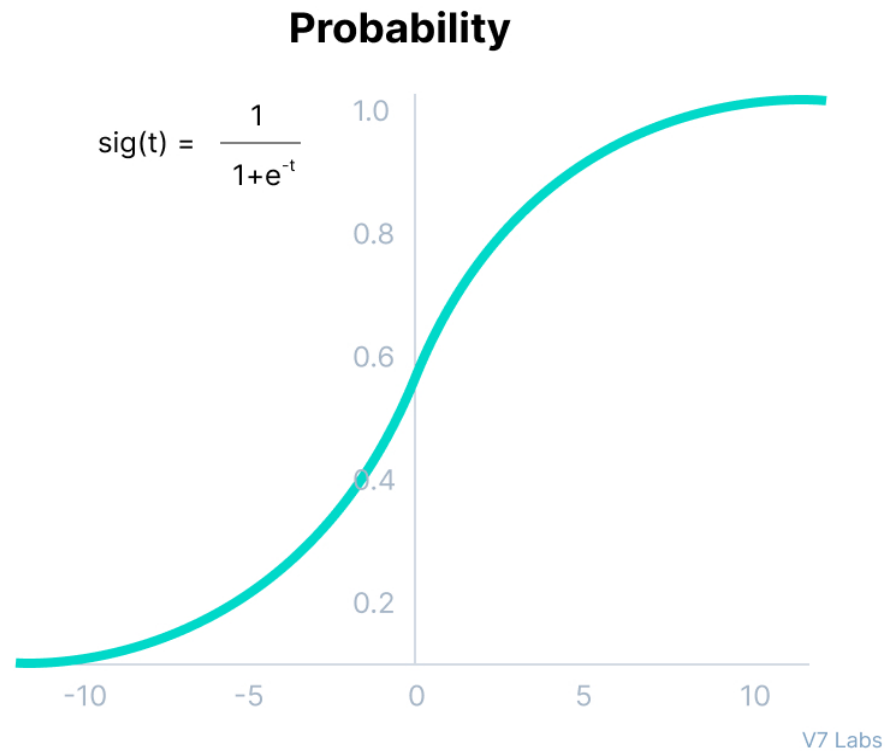
- (e) *Rectified Linear Unit (ReLU)*: The fundamental issue here is that not all of the neurons are activated simultaneously by the ReLU function. If the output of the linear transformation is less than 0 will the neurons become inactive.

Figure 3.9: (a). Rectified Linear Unit.

It has the following mathematical representation:

$$Y = \begin{cases} \max(0, X) \end{cases} \quad (3.10.5)$$

- (f) *Softmax function*: The result of the sigmoid function, which reflect probability, was somewhere between 0 and 1.

Figure 3.10: (a). Softmax Function.

$$Y = \text{Softmax}(x_i) = \left\{ \frac{\exp(x_i)}{\sum_j \exp(x_j)} \right\} \quad (3.10.6)$$

Here we discuss only few important activation function, many of them are not discuss here.[47]

4. *Loss function:*

$$\epsilon = \frac{1}{2} \sum_{k=1}^c (Y_k - Z_k)^2$$

Here ϵ denote the loss function, c denote total output, Y is the given output of our data, and Z is the predicted output.

With regard to, this loss function must be optimized. The hidden weights and inputs. We can now update the weights for the best cost function by computing the partial derivative of the cost function for each weight.

so we can find change in weights as

$$\Delta W = -\eta \frac{\partial \epsilon}{\partial W}$$

and for W_{pq} is

$$\Delta W_{pq} = -\eta \frac{\partial \epsilon}{\partial W_{pq}}$$

similarly, for W_{kj} is

$$\Delta W_{kj} = -\eta \frac{\partial \epsilon}{\partial W_{kj}} = -\eta \frac{\partial \epsilon}{\partial net_k} * \frac{\partial net_k}{\partial W_{kj}}$$

Here $net_k = \sum_{j=1}^{nH} Y_j W_{kj} + \beta_{ko}$

$$\Delta W_{kj} = -\eta \left(\frac{\partial \epsilon}{\partial Z_k} * \frac{\partial Z_k}{\partial net_k} \right) * Y_j,$$

Here $Z_k = f(net_k)$

$$\Delta W_{kj} = \eta (Y_k - Z_k) * f'(net_k) * Y_j$$

Similarly we find all other weights.

When we have the updated weights from error function as a result of the back propagation algorithm, we can modify weights as follows:

$$W_{p+1} = W_p + \Delta W$$

This is artificial neural network algorithm. [48]

3.11 ANN Model Architectures

3.11.1 MultiLayer Perceptron Model

A feedforward artificial neural network with at least three layers of nodes is known as an MLP. Each neuron utilizes a non-linear activation function, whether in the hidden or output layer. The fact that MLP has numerous layers and uses a non-linear activation function separates it from conventional linear perceptrons and enables these networks to discern between data that cannot be linearly separated. Regression analysis can be used to create mathematical models using MLPs, which are universal function approximators. A wide range of modeling applications, including pattern classification, prediction, and function approximation, are ideally

suited for these networks.

The classification of data into distinct classes is what pattern classification is all about. When the present and historical trends are known, prediction relates to anticipating time series data, whereas function approximation entails modeling the relationship between the variables. [49]

3.11.2 Generalized Feedforward ANN Model

A GFF neural network is an ANN in which, unlike recurrent neural network models, the unit connections do not form a cycle. [50]

The data in this network, which was the original and most basic version of ANN, flows from the input nodes to the output node through the hidden nodes in only one direction, forward. In contrast to MLP, which is built on perceptrons, this network's architecture uses a generalized shunting neuron (GSN) model as its fundamental computational unit. The GFF neural network architecture can carry out a variety of tasks, including complicated pattern classification problems, dynamical modeling, time series forecasting, pattern recognition, and data mining, thanks to these shunting neurons' capacity for constructing complex, nonlinear decision boundaries. [51]

3.12 Partial Least Square Regression

A large range of techniques known as partial least squares (PLS) is used to model relationships between sets of observed quantities using latent variables. By maximizing the covariance between two various blocks of the original variables, partial least squares define additional components. PLS accomplishes this by first projecting measured data onto a more condensed latent space and fitting the model using this revised dataset. Regarding the assumptions that PLS needs to be supported by data for them to be applicable, such hypotheses are not as restrictive. PLS does show promise as a multivariate strategy for reducing dataset collinearity. PLS also addresses the issue that emerges when the number of people is substan-

tially less than the cardinality of the variables. Furthermore, because PLS does not require that data come from normal or well-known distributions, it has minimal requirements for residual distribution.

The iterative Ordinary Least Squares (OLS) procedure is used to estimate the coefficients of a system of simultaneous equations. H. Wold, an econometrician, improved his Fixed Point technique in 1966 to create PLS. [52] This approach was replaced by the Nonlinear Iterative Partial Least Squares algorithm, which computes principal components and canonical correlations using an iterative series of simple and multiple ordinary least squares regressions. [53] The General PLS technique, which is based on Nonlinear Iterative Partial Least Squares, first appeared as an iterative method for locating latent variables around the end of 1977. There are various methods for extracting latent vectors, and each one results in a distinct PLS version. For instance, PLS Mode A refers to the original Wold technique, but PLS1 and PLS2 are the most popular variations. Additionally, certain PLS variations simulate relationships among more than two groups. Since the 1980s, Wold's son has added diagnostic interpretation while simplifying the PLS method. [54] [55] [56] Because of these latter developments, the PLS approach is now a widely used tool for scientific data analysis that guarantees computational simplicity and ease of application even for very large datasets. PLS is capable of performing numerous statistical tasks, including regression and classification, but it may also be used as a descriptive tool. Furthermore, nonlinear PLS and dimension reduction can also be defined.

Today, PLS is a very effective technique in a variety of scientific and professional settings.

3.12.1 PLS Regression Algorithm

There are various PLS algorithms, [57] where orthogonal scores PLSR [58] is seen to be most important. Here independent data matrix $X_{(c,r)}$ uses the equation $Y = \beta_o + X\beta + \epsilon$ to explain the variation in response $Y_{(r,1)}$. Here β_o and β are the PLS regression parameters and ϵ is the error term. With scaled $X_0 = X - \mathbf{1}\bar{x}'$ and

$Y_0 = Y - 1\bar{y}$, the PLS algorithm runs for N components that is $a = 1, 2, \dots, N$ by computing

1. *Loading weights* by:

$$w_a = X_{a-1}' Y_{a-1}$$

Here weights denote the covariance of Y_{a-1} with X_{a-1} , The length of normalized loading weights is equal to 1 by

$$w_a \leftarrow w_a / \|w_a\|$$

2. *Score vector* z_a by:

$$z_a = X_{a-1} w_a$$

3. *X-loadings* q_a by regressing the data matrix in X_{a-1} on the score vector:

$$q_a = X_{a-1}' \frac{z_a}{z_a' z_a}$$

Similarly, Y-loading p_a by

$$p_a = Y_{a-1}' \frac{z_a}{z_a' z_a}$$

4. *Deflated* X_{a-1} and Y_{a-1} as:

$$X_a = X_{a-1} - z_a q_a'$$

$$Y_a = Y_{a-1} - z_a p_a$$

5. If $a < N$ return to 1.

Loading weights, scores and loadings are assembled as $W = [w_1, w_2, \dots, w_N]$, $Z = [z_1, z_2, \dots, z_N]$, $Q = [q_1, q_2, \dots, q_N]$ and $P = [p_1, p_2, \dots, p_N]$. These results in $\hat{\beta} = W(Q'W)^{-1}P$ with $\hat{\beta}_o = \bar{y} - \bar{x}\hat{\beta}$.

Variable selection, which is necessary for consistent parameter estimate and prediction, is not carried out by standard PLS. Additionally, it helps enhance comprehension of the fitted model. Thus, it is now an important component of PLS modeling. Soft-threshold PLS (St-PLS) and distribution-based truncation for variable selection in PLS (Tr-PLS) are two potential PLS variants.

3.12.2 Soft-Threshold PLS (St-PLS)

Soft-thresholding step in the PLS algorithm (St-PLS) is inspired by the nearest shrunken centroid approach. The St-PLS algorithm modifies the loading-weights at each step of iterative PLS as

1. Scaling:

$$w_i \leftarrow w_i / \max_l |w_{i,l}|, \text{ for } l = 1, \dots, n$$

2. Soft-thresholding:

$$w_{i,l} \leftarrow \text{sign}(w_{i,l})(|w_{i,l}| - \delta)_+, \text{ for } l = 1, \dots, n \text{ and some } \delta \in [0, 1]. \text{ Here } (\dots)_+ \text{ means } \max(0, \dots)$$

3. Normalizing:

$$w_i \leftarrow w_i / \|w_i\|$$

The degree of thresholding is a shrinkage $\delta \in [0, 1)$ parameter in St-PLS, which is a bigger δ result that leads to a smaller number of selected variables. Cross-validation tuning is required for the best model fitting.

3.12.3 Distribution Based Truncation For Variable Selection In PLS (Tr-PLS)

For the variable selection, Liland et al. (2013) [59] proposed truncating the PLS loading weights \mathbf{w} depending on distributions. Tr-PLS modifies the loading weights as part of the sequential PLS algorithm at each step to:

- 1) loading weights are sorted ' w ' as ' w_n '
- 2) Create a confidence interval based on the median of ' w_n ', which is based on a threshold pr .
- 4) Labels outliers as informative contributors and inliers as noise.
- 5) Truncate inlier.

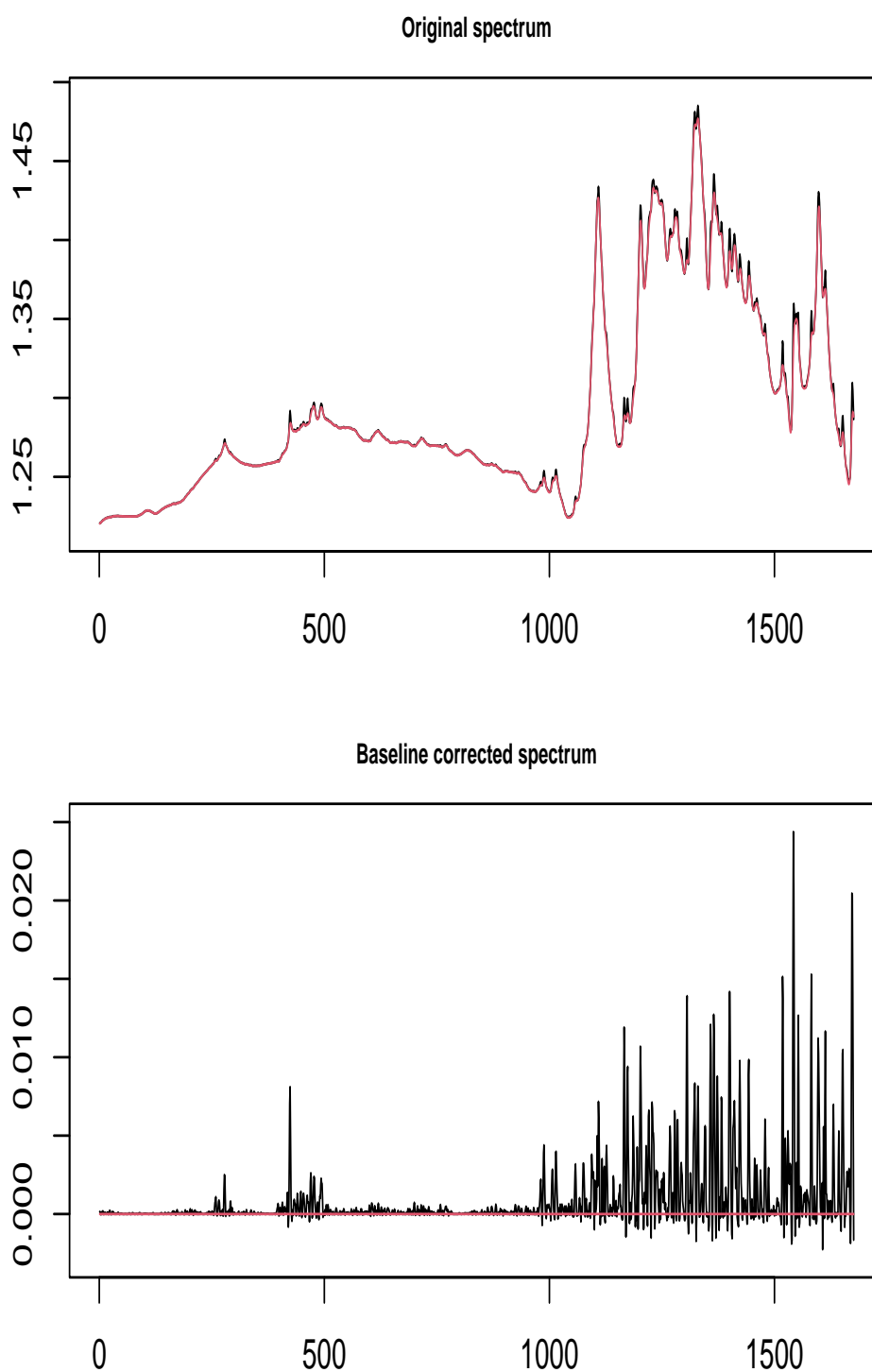
CHAPTER 4

Results And Discussion

Performing the calculation for S-1 to S-10 complexes of Metronidazole. The Antidiabetic activity of these complexes is calculated for the concentration level of these complexes at 50% level of α -amylase and α -glucosidase inhibition. Table:1 shows the concentration level at 50% alpha-amylase and alpha-glucosidase inhibition. The synthesized Metronidazole complexes samples were analyzed using the ATR-FTIR spectroscopy method. FTIR spectrum baseline correction at zero is necessary for detecting the influential spectrum peaks. For the spectrum baseline correction, the ALS method is used. Weight p and the smoothing parameter (λ) are the two parameters of ALS for tuning.

The spectrum baseline is corrected by using the ALS optimal parameters in which choose λ is 0.01, which are used in further analysis. A comparison of baseline-corrected spectrum and original spectrum is presented in figure 2.

Figure 4.1: Comparison Of Baseline Corrected Spectrum By Using ALS Method And Original Spectrum



To make infrared spectral measurements, we employed ATR-FTIR. X-axis: Dis-

play the wavenumber in (cm^{-1}). The above part of the figure, shows the original spectrum of the FTIR spectrum. The below part of the figure shows the baseline corrected spectrum which reduces the effect of such drifts. For the baseline correction, Asymmetric Least Squares (ALS) method is used. The ALS approach is employed for spectrum baseline correction. For tuning the original spectrum, two parameters are used: weight and smoothing parameter. Here λ is the smoothing parameter and W is the weight. These parameter are optimized. The spectrum baseline is adjusted using the ALS optimum parameters before being used in analysis.

Root Mean Square Error (RMSE) is the standard deviation of the errors. The RMSE represents the degree of dispersion of these residuals.

Figure 4.2: Calibration of Alpha-amylase inhibition and Alpha-glucosidase inhibition

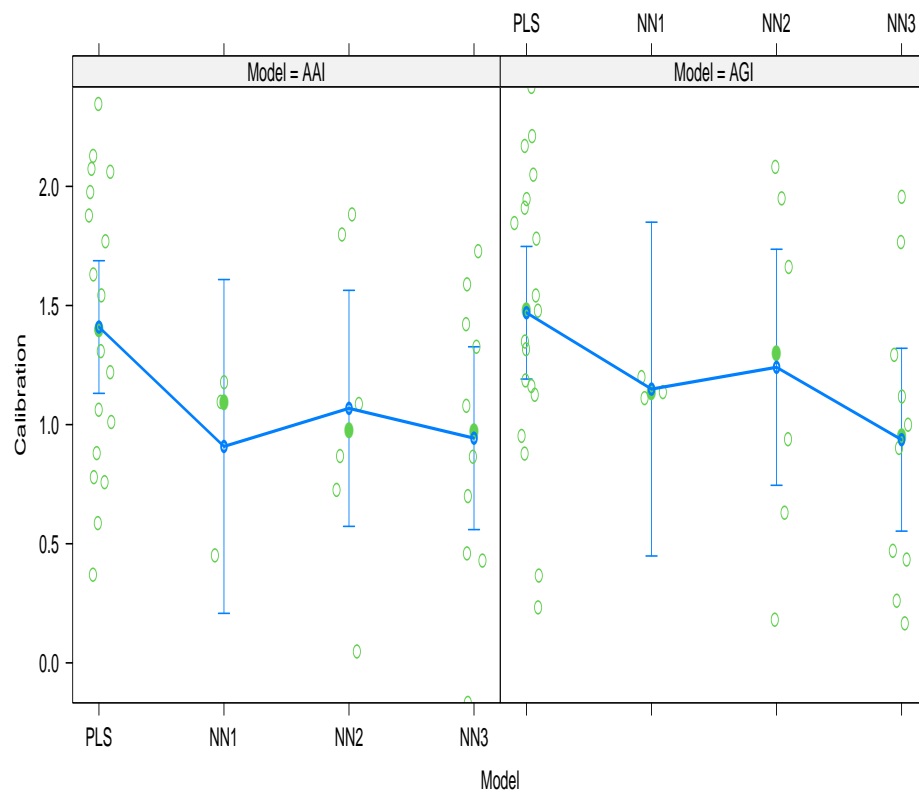
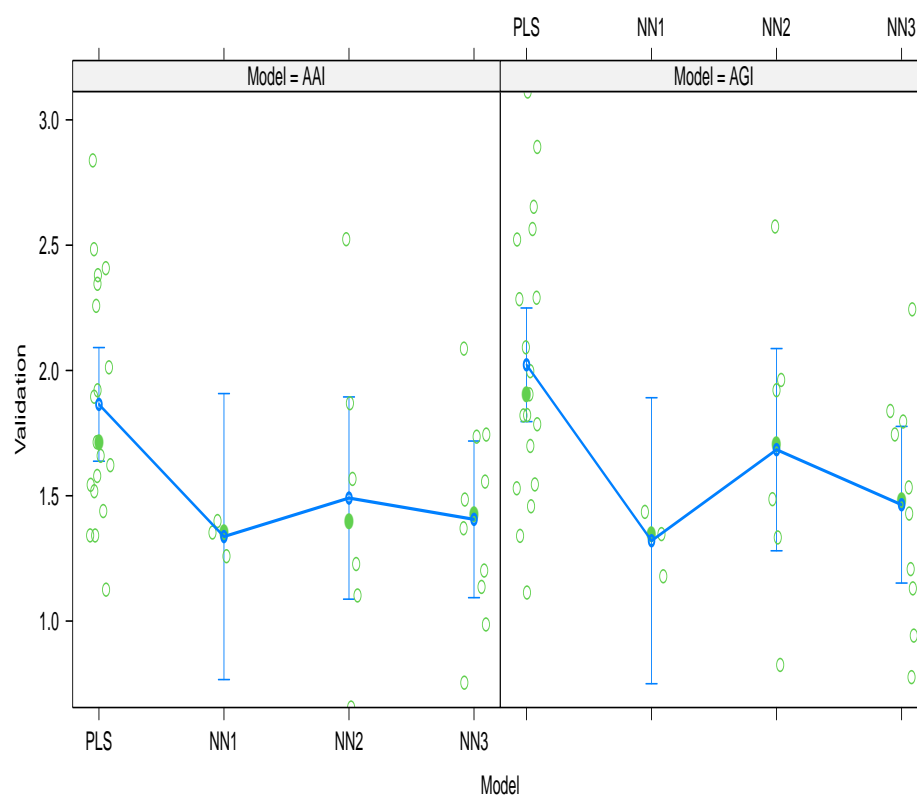


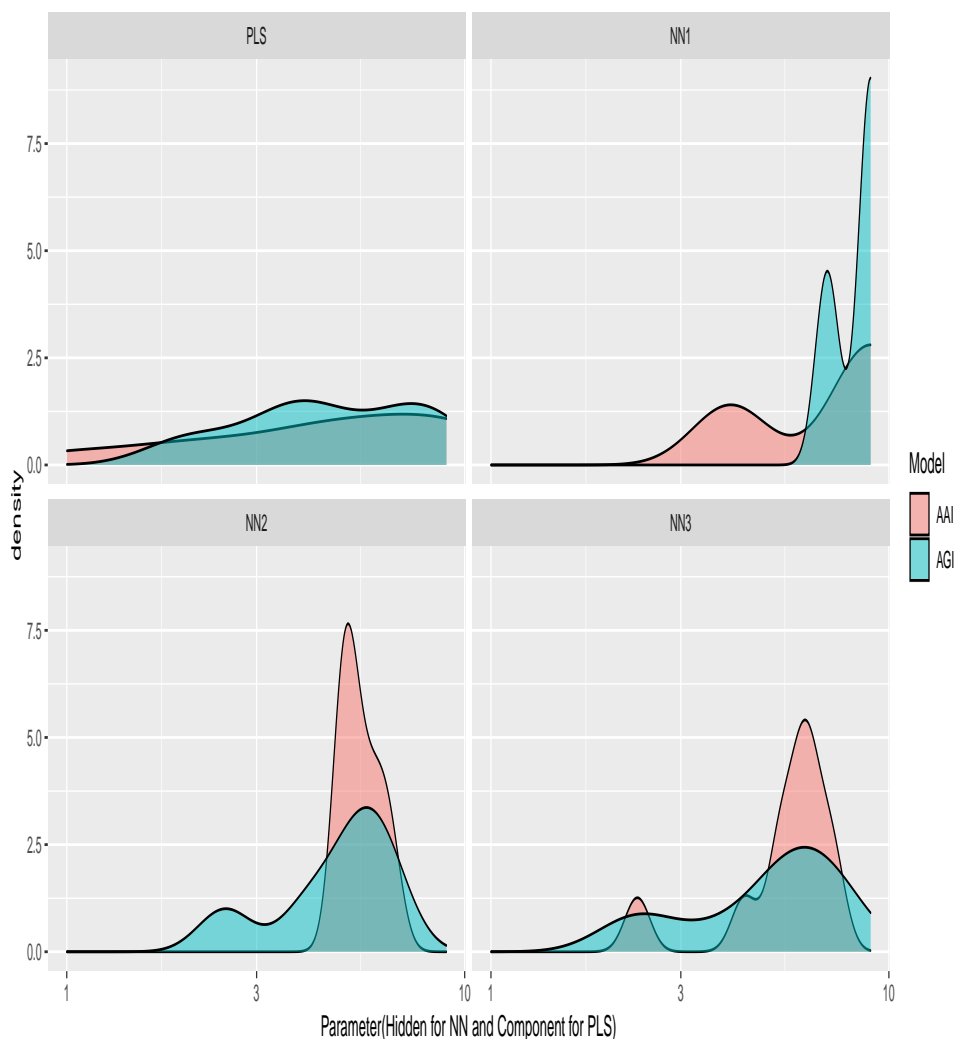
Figure 4.3: Validation of Alpha-amylase inhibition and Alpha-glucosidase inhibition

For the analysis of α -amylase and α -glucosidase inhibition at 50% concentration level of metronidazole complexes, we use two methods PLS and ANN. ANN in this work trained for one, two, and three hidden layers. each hidden layers use a maximum of 10 neurons and optimize. After the optimization of hidden neurons, find the root mean square error for PLS and ANN for all three hidden layers. In the above figure, Y-axis shows the root mean square error, and X-axis shows PLS and ANN for one, two, and three hidden layers. This figure shows that for the calibration dataset ANN show better performance for one hidden layer in the case of α -amylase inhibition and for α -glucosidase inhibition ANN shows the least root mean square error for three hidden layers. PLS shows the greater root mean square error (RMSE). ANN with two hidden layer also show higher RMSE. But ANN with one hidden layer show greater variability in RMSE, so ANN with three

hidden layers is best choice for this model.

In the figure 4.3, shows the performance of PLS and ANN with one, two, and three hidden layers for Validation dataset. For both α -amylase inhibition and α -glucosidase inhibition PLS and ANN with two hidden layers show greater RMSE. both α -amylase inhibition and α -glucosidase inhibition ANN with one hidden layer show better performance than others. But again, ANN with one hidden layer shows greater variability than ANN with one hidden layers so, we consider ANN with three hidden layers for further analysis.

A density plot is a visual representation of the distribution of a numeric variable. Figure 5 shows the density plot of parameters of ANN and PLS for α -amylase inhibition and α -glucosidase inhibition.

Figure 4.4: density plot for paramters α -amylase inhibition α -glucosidase inhibition

The above figure shows a density plot of PLS and ANN for one, two, and three hidden layers. In this plot, peach color shows the density of α -amylase inhibition and aqua color shows the density of α -glucosidase inhibition. In the figure 4.4, X-axis shows the parameters for ANN and PLS and Y-axis shows the density. PLS shows the smoothness and belongs to almost uniform distribution. For α -amylase inhibition, PLS shows more uniformity than α -glucosidase inhibition assay. ANN with first and second hidden layers sometimes shows no density and sometimes shows greater density. ANN with first and second hidden layers shows negative skewness in the graph. ANN with a third hidden layer shows approximately normality in the graph for α -amylase inhibition and α -glucosidase inhibition.

Figure 4.5 & 4.6 shows the influential coefficient of the best-fitted model for each. The functional compound assignment and class for each model are also mentioned.

Figure 4.5: The influential functional compound assignment and class for α -glucosidase inhibition

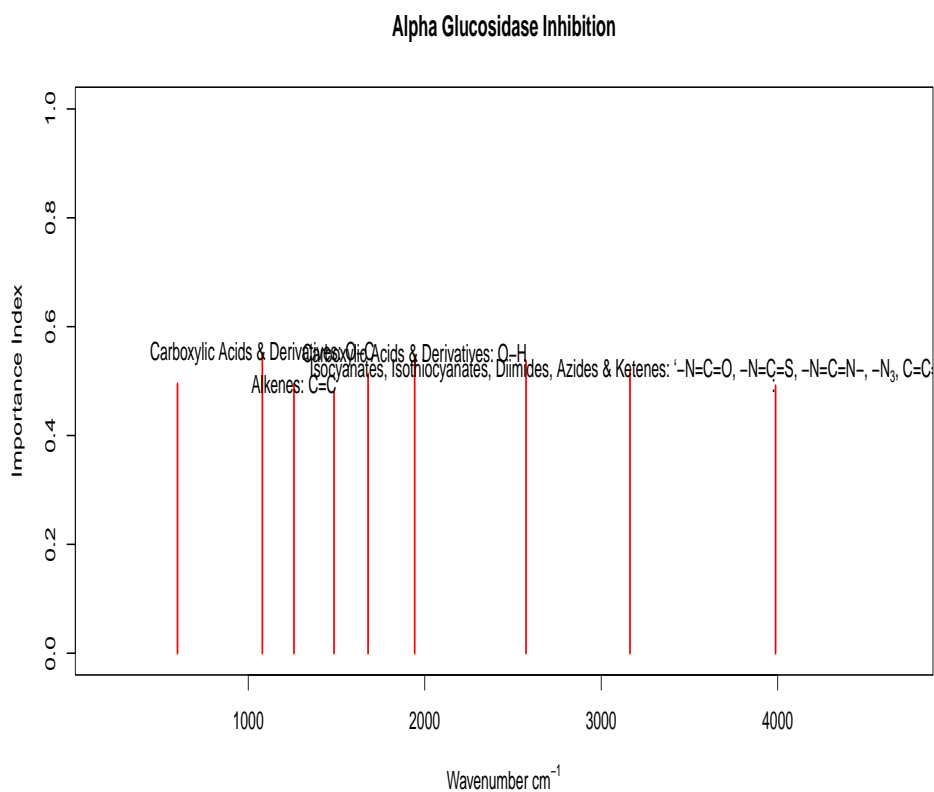
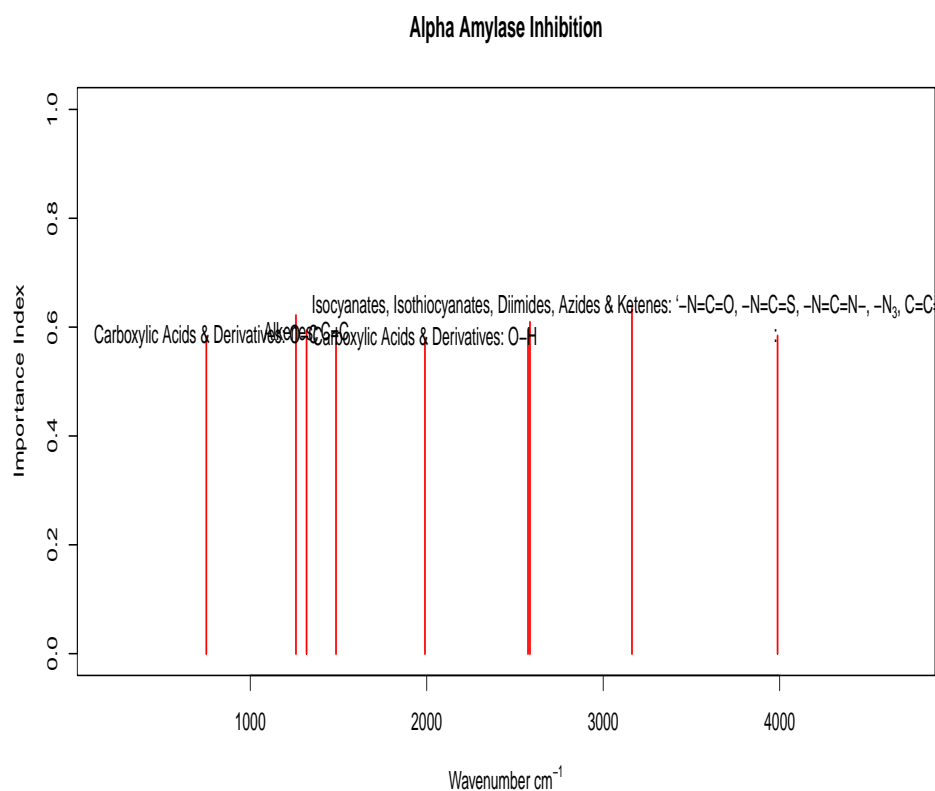


Figure 4.6: The influential functional compound assignment and class for α -amylase inhibition



In figure 4.5, X-axis show wavenumber in cm^{-1} and Y-axis show the importance of Index, Red line show the influential of functional compounds. For prediction of antidiabetic activity of metronidazole complexes against α -glucosidase inhibition the influential wavenumber corresponds to Carboxylic Acids & Derivatives functional compounds $O-H$, Carboxylic Acids & Derivatives functional compounds $O-C$, Alkenes functional compounds $C \equiv C$, and Isocyanates, Isothiocyanates, Diimides, Azides & Ketenes functional compounds $-N \equiv C \equiv O$, $-N \equiv C \equiv S$, $-N \equiv C \equiv N-$, $-N_3$, $C \equiv C \equiv O$

For prediction of antidiabetic activity of metronidazole complexes against α -amylase inhibition the influential wavenumber corresponds to Carboxylic Acids & Derivatives functional compounds $O-H$, Carboxylic Acids & Derivatives functional compounds $O-C$, Alkenes functional compounds $C \equiv C$, and Isocyanates,

Isothiocyanates, Diimides, Azides & Ketenes functional compounds $-N \equiv C \equiv O$, $-N \equiv C \equiv S$, $-N \equiv C \equiv N-$, $-N_3$, $C \equiv C \equiv O$ These two graphs show the antidiabetic activity of metronidazole complexes against alpha-amylase inhibition and alpha-glucosidase inhibition have the same influential functional compounds.

Figure 4.7: Alpha amylase inhibition Network

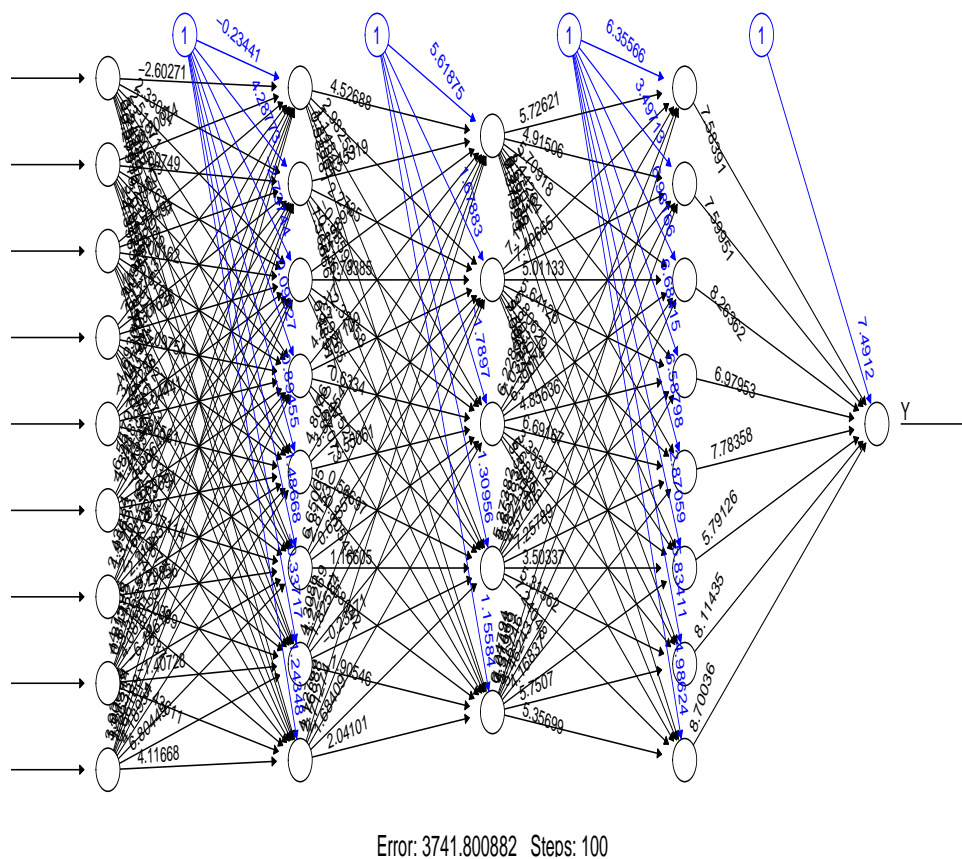
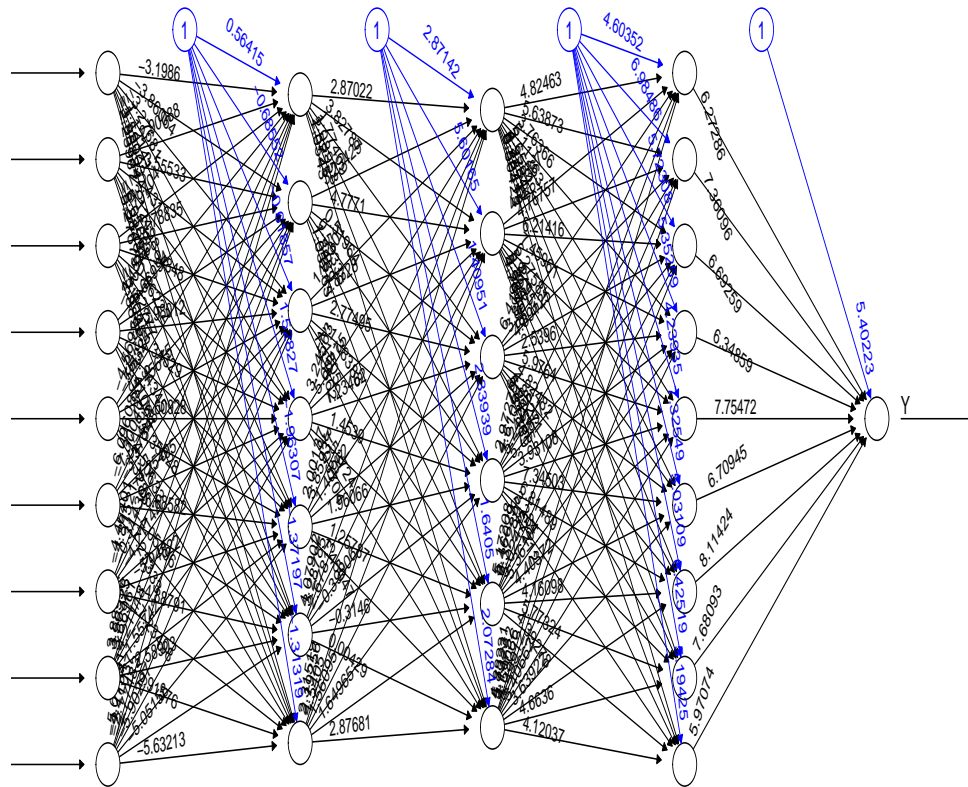


Figure 4.8: Alpha glucosade inhibition Network



Error: 3741.800882 Steps: 93

The artificial neural network's parameters are tuned by selecting three hidden layers. For α -amylase inhibition, the first hidden layer of ANN consists of eight neurons, the second hidden layer consist of five neurons, and the third hidden layer consists of again eight hidden layer and for α -glucosidase inhibition, the first hidden layer consists of seven neurons, second hidden layer consist of six neurons and third hidden layer consist of nine hidden layers. The above figure shows the value of hidden neurons and weights for alpha-amylase inhibition and alpha-glucosidase inhibition. Weights for α -amylase inhibition assay are -0.23441, 5.61875, and 6.35566 for one, two, and three hidden layers respectively. Similarly weights for α -glucosidase inhibition are 0.56415, 2.87142, and 4.60352 for one, two, and three hidden layers respectively. The optimized structure of ANN for α -amylases inhibition and α -glucosidase inhibition is shown in figures 4.7 & 4.8.

4.1 Computations

R programme is used for calculations, modelling, and figures. ². For baseline correction R package 'baseline' ³ and for ANN model fitting R package, 'neuralnet' is used.

4.2 Conclusions

The purpose of this work is to identify functional groups for the characterization of the antidiabetic activity of complexes of metronidazole that have already been synthesized. ASL is effective in correcting baselines. The PLS and ANN were used to model the antidiabetic behavior of Metronidazole complexes using the FTIR spectrum, where the antidiabetic activity was tested against α -amylases inhibition and α -glucosidase inhibition. The potent group of the practical compound is also plotted. ANN shows better performance than PLS.

Research is also required to evaluate and assess current discoveries for medicinal applications. Additionally, in future research, we sought to locate the group of significant wavenumbers that may more accurately describe the metronidazole complexes.

²R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.Rproject.org>.

³Liland, K. H., Almøy, T. & Mevik, B.-H. Optimal choice of baseline correction for multivariate calibration of spectra. *Appl. Spectrosc.* 64, 1007–1016 (2010).

References

- [1] Jialin Wen, Na Li, Dan Li, Minmin Zhang, Yangwen Lin, Zhenping Liu, Xiaofeng Lin, and Lingling Shui. Cesium-doped graphene quantum dots as ratiometric fluorescence sensors for blood glucose detection. *ACS Applied Nano Materials*, 4(8):8437–8446, 2021.
- [2] Sarhang Hasan Azeez, Sarwar Nawzad Jafar, Zahra Aziziaran, Le Fang, Ahang Hasan Mawlood, and Muhammed Furkan Ercisli. Insulin-producing cells from bone marrow stem cells versus injectable insulin for the treatment of rats with type i diabetes. *Cellular, Molecular and Biomedical Reports*, 1(1):42–51, 2021.
- [3] Unai Galicia-Garcia, Asier Benito-Vicente, Shifa Jebari, Asier Larrea-Sebal, Haziq Siddiqi, Kepa B Uribe, Helena Ostolaza, and César Martín. Pathophysiology of type 2 diabetes mellitus. *International journal of molecular sciences*, 21(17):6275, 2020.
- [4] Saher Mahmood Jwad and Hawraa Yousif AL-Fatlawi. Types of diabetes and their effect on the immune system. *Journal of Advances in Pharmacy Practices (e-ISSN: 2582-4465)*, pages 21–30, 2022.
- [5] Centers for Disease Control, Prevention, et al. Diabetes report card 2014. atlanta, ga: Centers for disease control and prevention, us dept of health and human services; 2015, 2015.
- [6] Natalie Nanayakkara, Andrea J Curtis, Stephane Heritier, Adelle M Gadowski, Meda E Pavkov, Timothy Kenealy, David R Owens, Rebecca L

REFERENCES

- Thomas, Soon Song, Jencia Wong, et al. Impact of age at type 2 diabetes mellitus diagnosis on mortality and vascular complications: systematic review and meta-analyses. *Diabetologia*, 64(2):275–287, 2021.
- [7] Mark A Atkinson, George S Eisenbarth, and Aaron W Michels. Type 1 diabetes. *The Lancet*, 383(9911):69–82, 2014.
- [8] Giulio Maria Pasinetti, Jun Wang, Shanee Porter, and Lap Ho. Caloric intake, dietary lifestyles, macronutrient composition, and alzheimer’disease dementia. *International journal of Alzheimer’s disease*, 2011, 2011.
- [9] World Health Organization. *World health statistics 2016: monitoring health for the SDGs sustainable development goals*. World Health Organization, 2016.
- [10] AL DePaula, ALV Macedo, N Rassi, S Vencio, CA Machado, BR Mota, LQ Silva, A Halpern, and V Schraibman. Laparoscopic treatment of metabolic syndrome in patients with type 2 diabetes mellitus. *Surgical endoscopy*, 22(12):2670–2678, 2008.
- [11] Jacob M Haus, Thomas PJ Solomon, Christine M Marchetti, John M Edmison, Frank Gonzalez, and John P Kirwan. Free fatty acid-induced hepatic insulin resistance is attenuated following lifestyle intervention in obese individuals with impaired glucose tolerance. *The Journal of Clinical Endocrinology & Metabolism*, 95(1):323–327, 2010.
- [12] Vladimir Nasteski. An overview of the supervised machine learning methods. *Horizons. b*, 4:51–62, 2017.
- [13] J Matthew Helm, Andrew M Swiergosz, Heather S Haeberle, Jaret M Karnuta, Jonathan L Schaffer, Viktor E Krebs, Andrew I Spitzer, and Prem N Ramkumar. Machine learning and artificial intelligence: definitions, applications, and future directions. *Current reviews in musculoskeletal medicine*, 13(1):69–76, 2020.

REFERENCES

- [14] Karel Diéguez-Santana, Oscar M Rivera-Borroto, Amilkar Puris, Hai Pham-The, Huong Le-Thi-Thu, Bakhtiyor Rasulev, and Gerardo M Casañola-Martin. Beyond model interpretability using lda and decision trees for α -amylase and α -glucosidase inhibitor classification studies. *Chemical biology & drug design*, 94(1):1414–1421, 2019.
- [15] Karel Dieguez-Santana, Oscar M Rivera-Borroto, Amilkar Puris, Huong Le-Thi-Thu, Gerardo M Casanola-Martin, et al. A two qsar way for antidiabetic agents targeting using α -amylase and α -glucosidase inhibitors: model parameters settings in artificial intelligence techniques. *Letters in Drug Design & Discovery*, 14(8):862–868, 2017.
- [16] Asmita Singh, Malka N Halgamuge, and Rajasekaran Lakshmganthan. Impact of different data types on classifier performance of random forest, naive bayes, and k-nearest neighbors algorithms. *International Journal of Advanced Computer Science and Applications*, 8(12), 2017.
- [17] DAAG Singh, E Jebamalar Leavline, and B Shanawaz Baig. Diabetes prediction using medical data. *Journal of Computational Intelligence in Bioinformatics*, 10(1):1–8, 2017.
- [18] Jack W Smith, James E Everhart, WC Dickson, William C Knowler, and Robert Scott Johannes. Using the adap learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings of the annual symposium on computer application in medical care*, page 261. American Medical Informatics Association, 1988.
- [19] Mehdi Khashei, Saeede Eftekhari, and Jamshid Parvizian. Diagnosing diabetes type ii using a soft intelligent binary classification model. *Review of Bioinformatics and Biometrics*, 1(1):9–23, 2012.
- [20] Huma Naz and Sachin Ahuja. Deep learning approach for diabetes prediction using pima indian dataset. *Journal of Diabetes & Metabolic Disorders*, 19(1):391–403, 2020.

REFERENCES

- [21] Huaping Zhou, Raushan Myrzashova, and Rui Zheng. Diabetes prediction model based on an enhanced deep neural network. *EURASIP Journal on Wireless Communications and Networking*, 2020(1):1–13, 2020.
- [22] Abir Alharbi and Munirah Alghahtani. Using genetic algorithm and elm neural networks for feature extraction and classification of type 2-diabetes mellitus. *Applied Artificial Intelligence*, 33(4):311–328, 2019.
- [23] Rakesh Motka, Viral Parmarl, Balbindra Kumar, and AR Verma. Diabetes mellitus forecast using different data mining techniques. In *2013 4th International Conference on Computer and Communication Technology (ICCCT)*, pages 99–103. IEEE, 2013.
- [24] Muhammad Akmal Sapon, Khadijah Ismail, and Suehazlyn Zainudin. Prediction of diabetes by using artificial neural network. In *Proceedings of the 2011 International Conference on Circuits, System and Simulation, Singapore*, volume 2829, page 299303, 2011.
- [25] Dilip Kumar Choubey, Sanchita Paul, Santosh Kumar, and Shankar Kumar. Classification of pima indian diabetes dataset using naive bayes with genetic algorithm as an attribute selection. In *Communication and computing systems: proceedings of the international conference on communication and computing system (ICCCS 2016)*, pages 451–455, 2017.
- [26] Kamer Kayaer, Tulay Yildirim, et al. Medical diagnosis on pima indian diabetes using general regression neural networks. In *Proceedings of the international conference on artificial neural networks and neural information processing (ICANN/ICONIP)*, volume 181, page 184, 2003.
- [27] Jose C Florez, Kathleen A Jablonski, Nick Bayley, Toni I Pollin, Paul IW de Bakker, Alan R Shuldiner, William C Knowler, David M Nathan, and David Altshuler. Tcf7l2 polymorphisms and progression to diabetes in the diabetes prevention program. *New England Journal of Medicine*, 355(3):241–250, 2006.

REFERENCES

- [28] Tejas N Joshi and Pramila M Chawan. Logistic regression and svm based diabetes prediction system. *International Journal For Technological Research In Engineering*, 5, 2018.
- [29] Harleen Kaur and Vinita Kumari. Predictive modelling and analytics for diabetes using a machine learning approach. *Applied computing and informatics*, 2020.
- [30] Ashok Kumar Dwivedi. Analysis of computational intelligence techniques for diabetes mellitus prediction. *Neural Computing and Applications*, 30(12): 3837–3845, 2018.
- [31] Mahmoud Heydari, Mehdi Teimouri, Zainabohoda Heshmati, and Seyed Mohammad Alavinia. Comparison of various classification algorithms in the diagnosis of type 2 diabetes in iran. *International Journal of Diabetes in Developing Countries*, 36(2):167–173, 2016.
- [32] Akm Ashiquzzaman, Abdul Kawsar Tushar, Md Islam, Dongkoo Shon, Kichang Im, Jeong-Ho Park, Dong-Sun Lim, Jongmyon Kim, et al. Reduction of overfitting in diabetes prediction using deep learning neural network. In *IT convergence and security 2017*, pages 35–43. Springer, 2018.
- [33] Bushra Rafique, Kainat Shafique, Sabahat Hamid, Saima Kalsoom, Muhammad Hashim, Bushra Mirza, Laila Jafri, and Mudassir Iqbal. Novel copper complexes of metronidazole and metronidazole benzoate: Synthesis, characterization, biological and computational studies. *Journal of Biomolecular Structure and Dynamics*, pages 1–16, 2021.
- [34] Tahir Mehmood, Mudassir Iqbal, and Rabia Hassan. Prediction of antibacterial activity in ionic liquids through ftir spectroscopy with selection of wavenumber by pls. *Chemometrics and Intelligent Laboratory Systems*, 206: 104124, 2020.
- [35] Xianchun Shen, Shubin Ye, Liang Xu, Rong Hu, Ling Jin, Hanyang Xu, Jianguo Liu, and Wenqing Liu. Study on baseline correction methods for the

REFERENCES

- fourier transform infrared spectra with different signal-to-noise ratios. *Applied Optics*, 57(20):5794–5799, 2018.
- [36] Paul HC Eilers and Hans FM Boelens. Baseline correction with asymmetric least squares smoothing. *Leiden University Medical Centre Report*, 1(1):5, 2005.
- [37] Vitaly I Korepanov. Asymmetric least-squares baseline algorithm with peak screening for automatic processing of the raman spectra. *Journal of Raman Spectroscopy*, 51(10):2061–2065, 2020.
- [38] Kristian Hovde Liland, Trygve Almøy, and Bjørn-Helge Mevik. Optimal choice of baseline correction for multivariate calibration of spectra. *Applied spectroscopy*, 64(9):1007–1016, 2010.
- [39] Ahmet K Atakan, WE Blass, and DE Jennings. Elimination of baseline variations from a recorded spectrum by ultra-low frequency filtering. *Applied Spectroscopy*, 34(3):369–372, 1980.
- [40] Guido Deboeck and Teuvo Kohonen. *Visual explorations in finance: with self-organizing maps*. Springer Science & Business Media, 2013.
- [41] Sonali B Maind, Priyanka Wankar, et al. Research paper on basic of artificial neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(1):96–100, 2014.
- [42] M Akhil Jabbar, BL Deekshatulu, and Priti Chandra. Classification of heart disease using artificial neural network and feature subset selection. *Global Journal of Computer Science and Technology Neural & Artificial Intelligence*, 13(3):4–8, 2013.
- [43] Hossein Hakimpoor, Khairil Anuar Bin Arshad, Huam Hon Tat, Naser Khani, and Mohsen Rahmandoust. Artificial neural networks’ applications in management. *World applied sciences journal*, 14(7):1008–1019, 2011.

REFERENCES

- [44] Y-S Park and S Lek. Artificial neural networks: Multilayer perceptron for ecological modeling. In *Developments in environmental modelling*, volume 28, pages 123–140. Elsevier, 2016.
- [45] Abiodun Ismail Lawal and Musa Adebayo Idris. An artificial neural network-based mathematical model for the prediction of blast-induced ground vibrations. *International Journal of Environmental Studies*, 77(2):318–334, 2020.
- [46] Wan-dong Hong, Xiang-rong Chen, Shu-qing Jin, Qing-ke Huang, Qi-huai Zhu, and Jing-ye Pan. Use of an artificial neural network to predict persistent organ failure in patients with acute pancreatitis. *Clinics*, 68:27–31, 2013.
- [47] Activation functions in neural networks. <https://www.v7labs.com/blog/neural-networks-activation-functions>, . Accessed: 2022-07-19.
- [48] Introduction to artificial neural networks — explanation, formulation & derivation. <https://towardsdatascience.com/introduction-to-artificial-neural-networks-5036081137bb>, . Accessed: 2019-12-04.
- [49] Matt W Gardner and SR Dorling. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15):2627–2636, 1998.
- [50] Andreas Zell. *Simulation neuronaler netze*, volume 1. Addison-Wesley Bonn, 1994.
- [51] Ganesh Arulampalam and Abdesselam Bouzerdoum. A generalized feedforward neural network architecture for classification and regression. *Neural networks*, 16(5-6):561–568, 2003.
- [52] Herman Wold. Estimation of principal components and related models by iterative least squares. *Multivariate analysis*, pages 391–420, 1966.
- [53] Herman Wold. Nonlinear iterative partial least squares (nipals) modelling: some current developments. In *Multivariate analysis-III*, pages 383–407. Elsevier, 1973.

REFERENCES

- [54] Svante Wold. Nonlinear partial least squares modelling ii. spline inner relation. *Chemometrics and Intelligent Laboratory Systems*, 14(1-3):71–84, 1992.
- [55] Svante Wold, Nouna Kettaneh-Wold, and Bert Skagerberg. Nonlinear pls modeling. *Chemometrics and intelligent laboratory systems*, 7(1-2):53–65, 1989.
- [56] Svante Wold, Arnold Ruhe, Herman Wold, and WJ Dunn, Iii. The collinearity problem in linear regression. the partial least squares (pls) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, 5(3):735–743, 1984.
- [57] Tahir Mehmood, Solve Sæbø, and Kristian Hovde Liland. Comparison of variable selection methods in partial least squares regression. *Journal of Chemometrics*, 34(6):e3226, 2020.
- [58] Svante Wold, Harold Martens, and Herman Wold. The multivariate calibration problem in chemistry solved by the pls method. In *Matrix pencils*, pages 286–293. Springer, 1983.
- [59] Kristian Hovde Liland, Martin Høy, Harald Martens, and Solve Sæbø. Distribution based truncation for variable selection in subspace methods for multivariate regression. *Chemometrics and Intelligent Laboratory Systems*, 122:103–111, 2013.