Real-Time Telecommunication Network Management using Data Mining and Machine Learning Techniques

By

Syeda Hajra Farhat Hashmi

MSIT-18 : 205855

Supervisor

Dr. Saad Qaiser

Department of Computing

In

School of Electrical Engineering and Computer Science(SEECS),

National University of Sciences and Technology(NUST),

Islamabad, Pakistan

(Aug 2022)

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis entitled "Real-Time Telecommunication Network Management using Data Mining and Machine Learning Techniques" written by SYEDA HAJRA HASHMI, (Registration No 00000205855), of SEECS has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil 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: 

Name of Advisor: Dr. Saad Qaisar Senior Member IEEE

Date: 15-Jul-2022

HoD/Associate Dean: _____

Date: _____

Signature (Dean/Principal): _____

Date: _____

Approval

It is certified that the contents and form of the thesis entitled "Real-Time Telecommunication Network Management using Data Mining and Machine Learning Techniques" submitted by SYEDA HAJRA HASHMI have been found satisfactory for the requirement of the degree

Advisor : Dr. Saad Qaisar Senior Member IEEE

Signature: 

Date: 15-Jul-2022

Committee Member 1:Dr. Syed Taha Ali

Signature: 

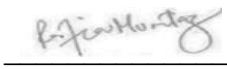
Date: 15-Jul-2022

Committee Member 2:Dr. Arsalan Ahmad

Signature: 

Date: 15-Jul-2022

Committee Member 3:Dr. Rafia Mumtaz

Signature: 

Date: 15-Jul-2022

Dedication:

This thesis is dedicated to my family, friends, their continuous support plays a vital role in thesis completion. And also my respectful supervisor Dr. Saad Qaiser, and thesis committee members Dr. Rafia Mumtaz, Dr. Talha, Dr. Arsalan, without their guidance this could not have come to an end.

Certificate of Originality

I hereby declare that this submission titled "Real-Time Telecommunication Network Management using Data Mining and Machine Learning Techniques" is my own work. To the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics, which has been acknowledged. I also verified the originality of contents through plagiarism software.

Student Name:SYEDA HAJRA HASHMI

Student Signature: 

Acknowledgement:

First of all I would like to thank Allah Almighty for completion of this important journey of my life.

Secondly, I would like to thank my supervisor Dr. Saad Qaiser and respected committee members, who have guided and supported me throughout the process.

And Lastly, I would like to thank my parents, spouse for their support emotionally.

Table of Contents

Abstract:

There are currently many type of industries that requires 24/7 monitoring on different levels for smooth operations. Telecommunication is one of the industry where millions of alarms trigger on daily basis from communication equipment's and needs to be handled within time limit. These monitoring operation is normally handled by human in loop which means huge amount of time wastage. In order to minimize downtime, limit human control over this monitoring, companies have implemented data mining and machine learning techniques that helps in not only proactive monitoring of alarms along with suitable actions and also there is a huge time saved. In this paper we have experimented with some real time telecommunication alarms that are gathered from different telecommunication devices and occurred at different times. We have created a system that can predict future occurrence of an alarm on the specified machine using machine learning technologies. In this paper we have used decision tree classifier in order to classify huge number of data received from devices. We are using it to predict alarms that are to be appeared on a specific device/machine at an specific time stamp.

Keyword- telecommunication network, decision tree classifier

Introduction:

There are a huge number of alerts generate in the network operation process in telecommunication industry. Daily these alarms needs to be monitored and processed by operations team in a telecommunication company. There are multiple devices that contributes in a communication system and each of the device sends in alarms which are in millions in number. These alarms that are generating from devices are of different categories for example major, minor and alarms with other categories. One device failure may cause alarms from other devices as well. That is why the quantity of alarms per day are in millions. These telecommunication companies have different methods of dealings with the faults occuring in the network system. These methods or processes are collectively known as fault management processes that helps in catering a huge number of network faults occuring in the network. Fault management is required to ensure security of network as well as availability, reliability and optimisation if there is any. Fault management is the key element for quality service providance in a telecommunication industry. These faults that are of any category or any level will eventually contribute in degradation of the services which is the key element of the industry. Therefore, the perfect system for fault management is required in these companies. There is a separate department allocated to for detection and rectification of these faults in the network which is called network operations center (NOC). The responsibility of this department includes detection of alarms using different softwares and then rectifying the problem been caused by creating tickets to the field team.

Fault managemnet is responsible for four main tasks that are[1]:

1, Detection of alarm: There are numerous amount of alarms coming in from different communication devices which are been monitored on netwrok mangemnet systems. Fault management system first categorize these alarms as major, minor etc. It helps in identifing the root cause alarm among huge number of alarms.

2, Diagonosis of an alarm: This task involves identifing root cause alarm from all the other alarms.

3, Isolating the root cause alarm: This process isolates the main problem causing alarms from the rest of other alarms.

4, lastly, fixation of the issue: After isolating the main alarm, the issue gets fixed. [1]

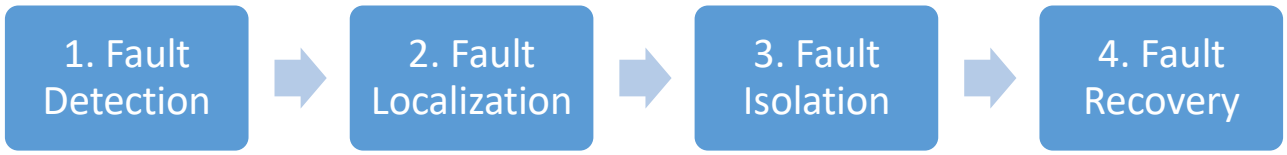


Fig: 1 A generic network fault management process [1]

Previously, these fault management processes were handled manually using human force. But as the increase in the need of communication systems and network systems, these processes are now handled using machines. Data mining and machine learning techniques are used nowadays in order to manage faults in the telecommunication industries. Hence minimizing human control and human error in these crucial systems.

Data mining and machine learning involves number of steps that helps in processing a huge amount of data into a format that can be used further for decision making [2]. First step in the data mining is known as data cleansing in which irrelevant data is removed so that the decision making process gives a more accurate results. This step also involves combining data from multiple source if any. Second step includes further examination of data and transforming it into a suitable format for the data mining algorithm. After this various data mining techniques are applied on the data and pattern evaluation takes place where unique and repetitive patterns are monitored, which helps further in addressing the issue [3].

Now machine learning / data mining not only helps in detection of various alarms, it also helps in alarming the teams beforehand. These tools are used to predict the pattern of alarms occuring on the devices using different or suitable algorithm and then on these predictions, pre-actions can be taken by the teams.

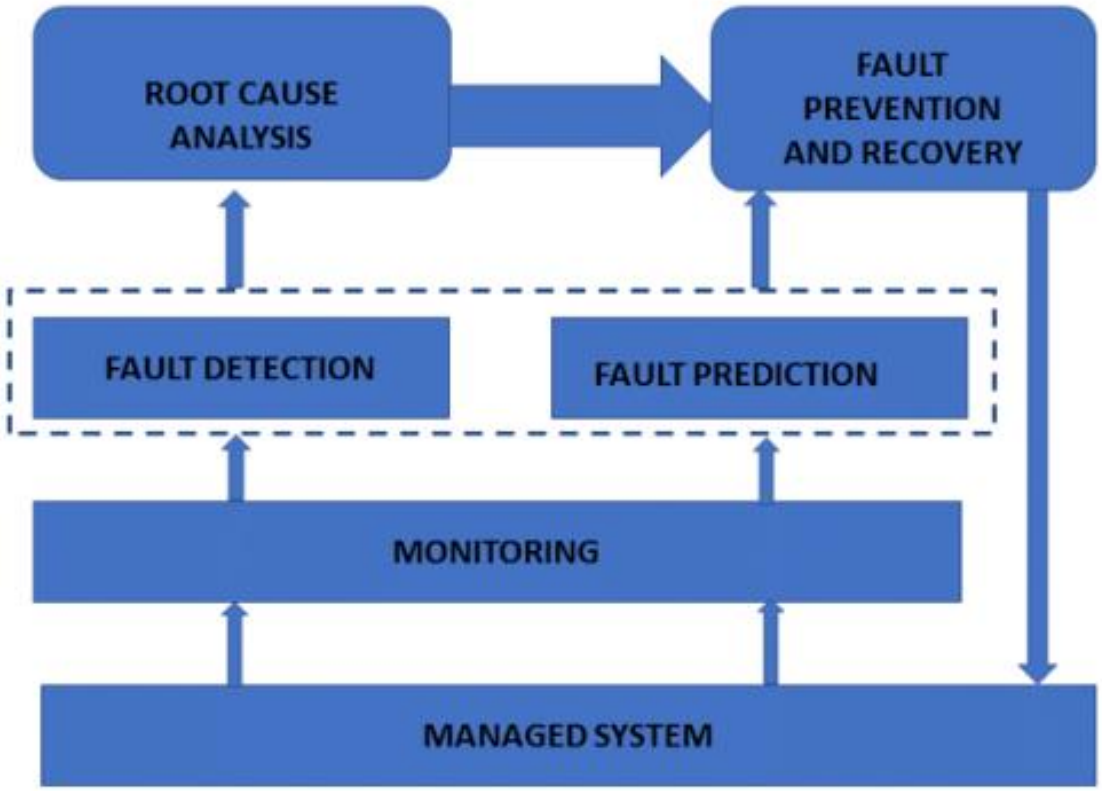


Fig:2 Shows key functions that are included in fault management systema nd how they are related.[4]

Early alerts/warnings can be assigned to the various teams via SMS or Email to their incharge will help in rectifying the problem beforehand. Like for example, if the alarm of battery level on the threshold value is appearing then a alert or warning is to be generated to the field team telling them to be ready for the charging of battery as the device will go down when the battery charge reaches below the threshold value.

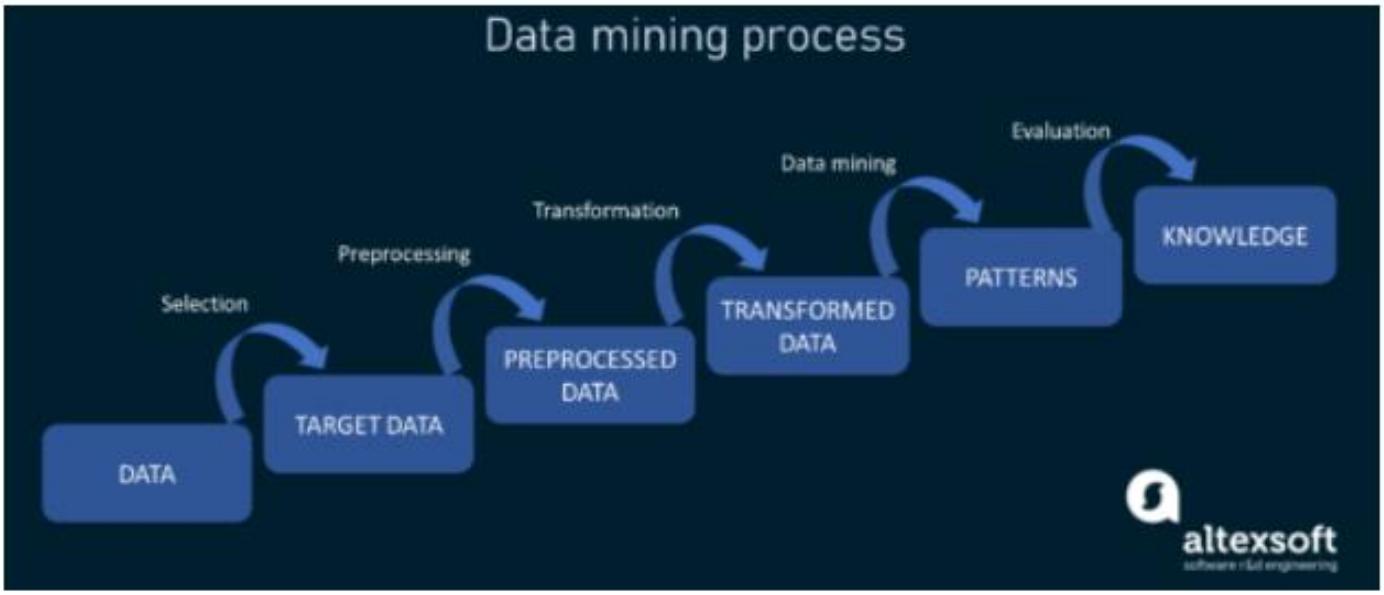


Fig :3 Online illustrating data mining process steps [5].

Literature Review:

There are pattern mining algorithms that have been used to analyze telecommunication alarms data for improving fault management system. A survey on fault management by Mourad Nouioua[6] describes association rules mining based approach in which analyzer defines association between multiple alarm sequesnces. This system however, is not suitable for long alarm sequences. Lozonavu [7] proposed a method to discover alarm pattern using sequential pattern mining algorithm. In this algorithm a relationship graph is constructed between alarms and each relationship is given a weightage which is calculated by confidence measure. These graphs represents a better view of network behavior and discovring of new network elements. Clustering alhorithms are also used to group relevant alarms from a huge set of alarms and then finding the root cause alarm. K-means and K-mediods are two most widely used methods for clustering algorithms. In K-means, centroids are initailaly assigned a value and then all the remaining data is assigned to the nearest centroid. After that recalculation of centroid occurs until centroid does not move[8, 10]. This method may not necessarily finds the optimal configuration as comapre to

global objective function[9]. Whereas, in K-mediods algorithm, this algorithm initially select random K-mediods for the data but on repaet tries to select a better choice. That is why this method is also known as representative object based algorithm[8].

Machine learning approaches are widely used for improving fault management system in the telecommunication industries. Artificial neural network based approaches includes training of algorithm on supervised data so that prediction of desired outputs can take place[11]. Human brain is the inspiration behind these networks.

Hybrid approaches are also been used for prediction of anomalies in telecommunication industries and future forecasting. For example Mirjana Pejic[12] worked on the prediction of churn management by using three stage hybrid approach. This paper have used k-means cluster analyses for identifying market segment in the customer churn dataset and then decision tree algorithm detected the churn ratio.

Methodology:

Data Set:

The data that I have used for this problem is gathered in a month long span. Probes were installed at the telecommuncation devices ends and continuous monitoring on the capturing of the alarms against the device names were held so that all of the data could be collected properly, thus helping in the correct predictions of the attributes.

Fig: 4 Shows an MSAG present on the locality

As shown in the figure above, these MSAGs produce different alarms that are being captured on different softwares depending on the vendor (i.e Huawei,Nokia,ZTE etc). These alarms are then beautified through a beautifier. Beautification of an alarm involves separation of alarm name, machine name, category of the alarm (that is critical alarm, major alarm, minor alarm etc). Through the process of beautification an alarm is readible by the human as well as actions can be easily taken on these alarms. Each device is connected to it's network managemnet system (separately designed based on the compatible protocols of communication)

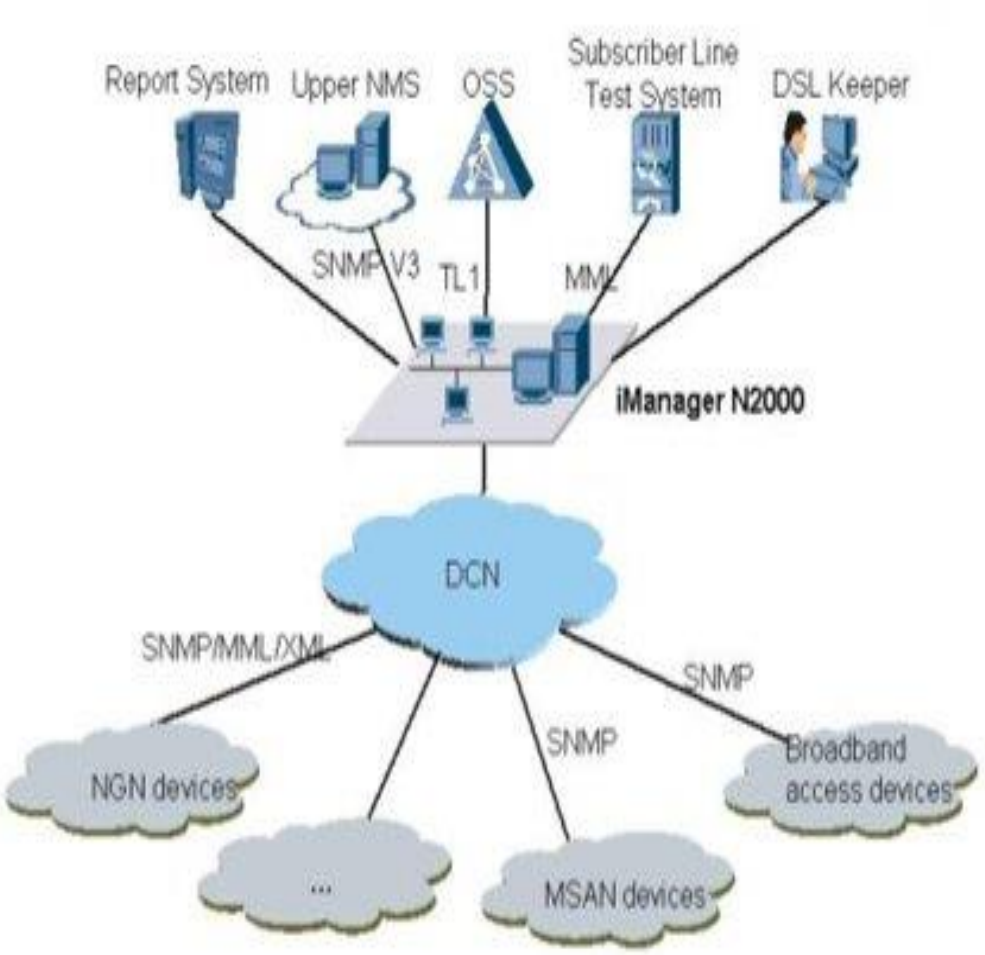


Fig 5: Showing Huawei Network Management system [13]

As shown is the figure above, all the telecommunication devices (NGN, MSAN,broadband access devices) are connected via DCN network with the network mangement system (iManager N2000). The protocl for communication is SNMP or for the NGN devices which are now obsolete SNMP/MML/XML.

Now after the alarms land on the network managemnet system, it gives a visual repesentation of the network and where the fault occurred.

The visual representation of the network on NMS is given below:

Fig 6: Visual representation of a network on NMS.[14]

In this visual representation, once the alarm occurs on any node, it will turn red(showing fault) or any device that goes offline will show as a grey in color (showing a device offline). These NMS will push these alarms to a software named TEMIP(company Huwaei) where each alarm is shown in tabular form for teams to understand alarm properly and take action acordingly.

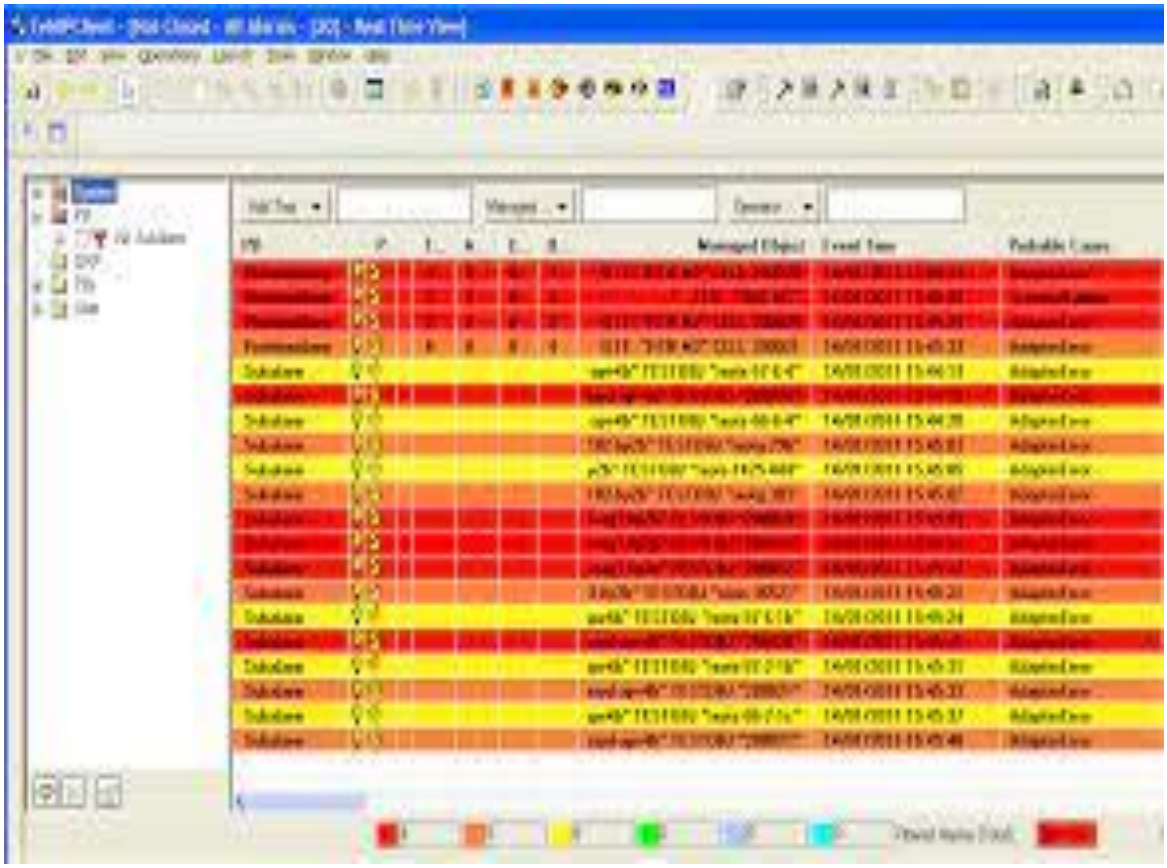


Fig 7: Temip Client preview showing different alarms categorise by the color of alarm.[15]

Network managemnet systems also push data towards the Data bases for stoarge of these alarms. You may retrive this data using SQL code. Or reporting or predictions for future can be eaily made on these data stored in the database.

My data set represents different type of alarms on different telecommunication devices that may or may not lead to disconnection of the device from the netwrok. Few alarms that tells the condition of battery or related to card etc in the device are not that major alarm, but the alarms that tells the disconnection of device or service downgarde are the alarms that needs to be sent to the field teams for rectification. Because these alarms effects the communication services been provided to the customers.

This data set is captured in the duration of 1 month. Total records of the data set contains approximately 20,000 rows. This data set contains alarms from different devices and it maily include hardware alarms that is if the communication between the devices goes offline. Or the card presnt in the devices for telecommunication goes disconnected. These cards may cause degradation in the services been provide to the customer where as the offline of the device may cause delay in telecommunication services overall. Thus affecting the customers all together. This affecting of customers may cause customer churn in the telecommunication company.

Alarm name is also provided in the dataset in order to know about the nature of an alarm. Second attribute is NE name that means Network element name (or the device name) on which an alarm has occurred. Third attribue is the creation time or the occurrence time of the alrm on the device.

Repition of alarm can tell the faulty device in the network. Device names are also defined in the data set that shows uniqueness to the alarm occuring in a perticular device. Last column tells the nature of alarm wether it is an “Outage” causing alarm or not.

Following are the attributes of the dataset:

1. **NE_Name:** This tells the network element name in the telecommuncation netwrok
2. **Alarm_Name:** This column shows all the alarms appearing on any perticular netwrok device.
3. **CREATIONTIMESTAMP:** The appearance time including date etc of a perticular alarm on the device.
4. **Outage/Non-Outage:** This column tells if an alarm is an outage leading alarm or not leading to an outage

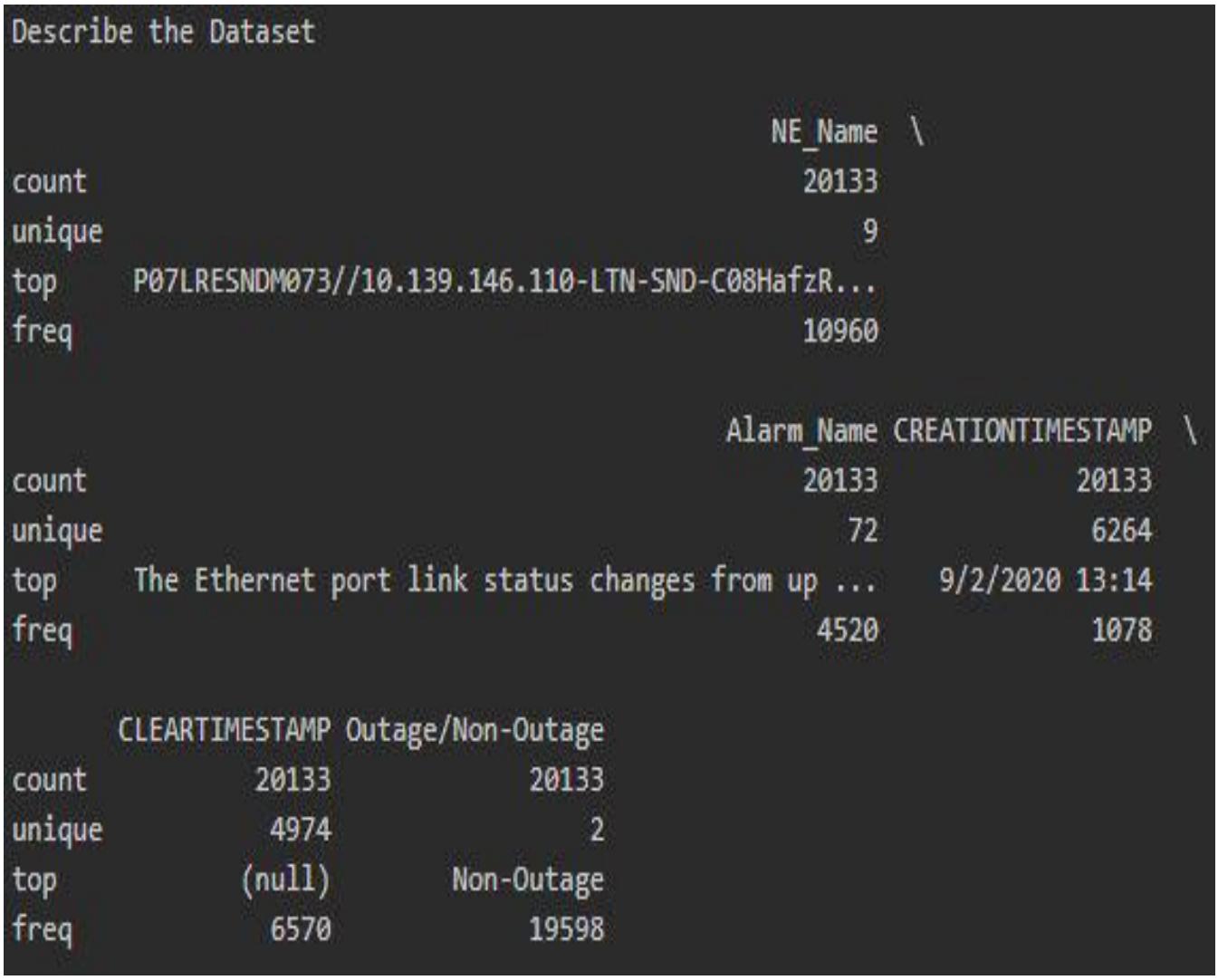
```
First Ten Rows

                                              NE_Name  \
0  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
1  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
2  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
3  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
4  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
5  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
6  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
7  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
8  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...
9  S13KHICTXM009//10.138.1.26-S2-KHI-Chamber OfCo...

                                Alarm_Name CREATIONTIMESTAMP  CLEARTIMESTAMP  \
0      Communication with the device failed    9/1/2020 10:03  9/1/2020 10:06
1  Port link status changed from up to down    9/1/2020 10:05   9/2/2020 7:55
2  Port link status changed from up to down    9/1/2020 10:05   9/2/2020 7:55
3  Port link status changed from up to down    9/1/2020 10:05   9/2/2020 7:55
4  Port link status changed from up to down    9/1/2020 10:05  9/1/2020 12:25
5  Port link status changed from up to down     9/2/2020 6:34   9/2/2020 9:41
6  Port link status changed from up to down     9/2/2020 6:34   9/2/2020 9:41
7  Port link status changed from up to down     9/2/2020 6:34   9/2/2020 9:41
8  Port link status changed from up to down     9/2/2020 6:34   9/2/2020 8:28
9  Port link status changed from up to down     9/2/2020 6:34   9/2/2020 8:46

  Outage/Non-Outage
0        Non-Outage
1        Non-Outage
2        Non-Outage
3        Non-Outage
4        Non-Outage
```

Fig 8: Python code showing ten rows of the Data set.



```
Describe the Dataset

                                                    NE_Name  \
count                                                 20133
unique                                                    9
top     P07LRESNDM073//10.139.146.110-LTN-SND-C08HafzR...
freq                                                  10960

                                                 Alarm_Name CREATIONTIMESTAMP  \
count                                                 20133             20133
unique                                                   72              6264
top     The Ethernet port link status changes from up ...    9/2/2020 13:14
freq                                                   4520              1078

       CLEARTIMESTAMP Outage/Non-Outage
count           20133             20133
unique           4974                 2
top            (null)       Non-Outage
freq             6570             19598
```

Fig:9: Python code showing description of the data set.

First we need to import the data to the program, then we need to prepare the data set for the algorithm. Preaparation of data set involves reoving null values and also removing any duplicate entries in the data. These errors may affect the outcome of the algoritm. As have seen previously there are multiple techniques in data mining and machine learning that can be used for prediction purposes. Here we are using decision tree classifier for classification and prediction of the alarms on the machines.

Decision Tree Classifier:

Now let's start to talk about the algorithm that was used in this program. Decision tree classification is a major example of supervised learning. It creates a tree like decision structure of all the possible tests on the attributes. Each branch represents the outcome of the test (which is 'True' or 'False') and each leaf node represents a class label. For this dataset, since we need to predict the occurrence of alarm on the specific node at the specific time that's why decision tree was the suitable choice for such problem.

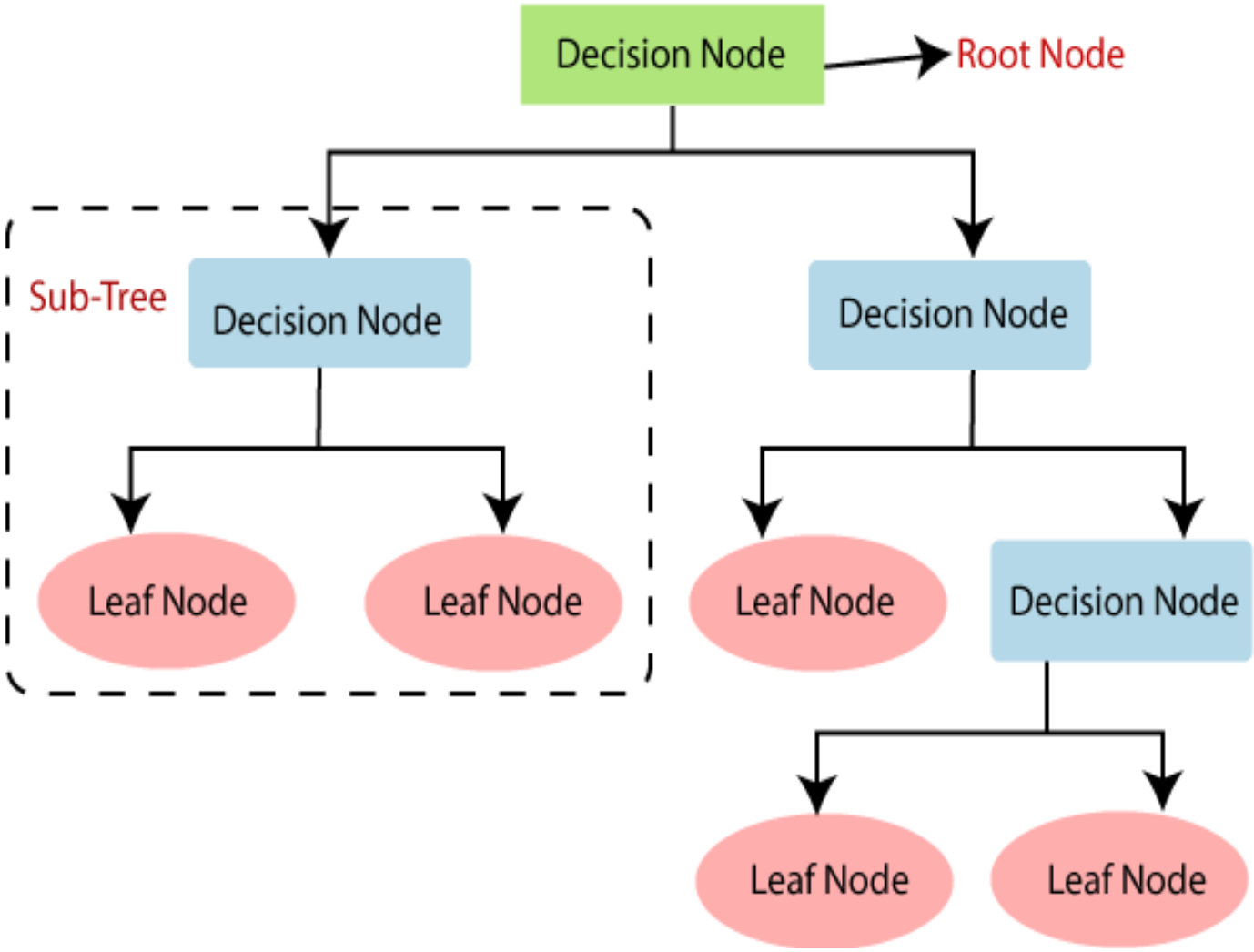


Fig 10: A generic Decision tree work flow diagram.[16]

Decision tree rules have following form:

if (first condition) and (second condition) and (third condition) then outcome.[12]

In the following example the data variables are creation timestamp and network element name and the target variable is alarm name. First we need to divide the data set into two data sets. In the first data set I have the dependent variable (Alarm Name & Nature of alarm) and in the second data set I have the independent variable (Device name), as shown in below code:

```
#creating two separate datasets for inputs and outputs
X= alarms_data['NE_Name']
Y= alarms_data.drop(columns=['NE_Name'])
```

Now, we need to train the classifier with some data and then test the classifier with the required inputs for predictions. We will split the data by using train_test_split method.

```
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,train_size=0.8
```

This method returns a tuple therefore we are splitting the data into training and testing datasets separately. Size=0.8 means 80% of the data is to be used to train the algorithm and 20% of the dataset is then used to test the classifier.

```
model=DecisionTreeClassifier()

model.fit(X_train,Y_train)
```

Above code represents the fitting of the classifier on the training datasets, both data and target variables.

Now we will predict on the testing data set:

```
prediction=model.predict(X_test)
```

After this accuracy of the classifier is also calculated by comparing this output to the Y_test separated earlier. If it's 100% then it means that classifier is working properly. The split of the training and testing datasets also effects the accuracy of the classifier. With decreasing the split percentage (i.e 20%) will decrease accuracy as well.

```
score= model.score(Y_test,prediction)
print(score)
```

For accuracy measurement, both the input arrays need to be 1d arrays, therefore we will transform the 2d array to 1d array by using command:

```
Y_test.values.ravel()
```

Then the score between Y_test and prediction variables are calculated.

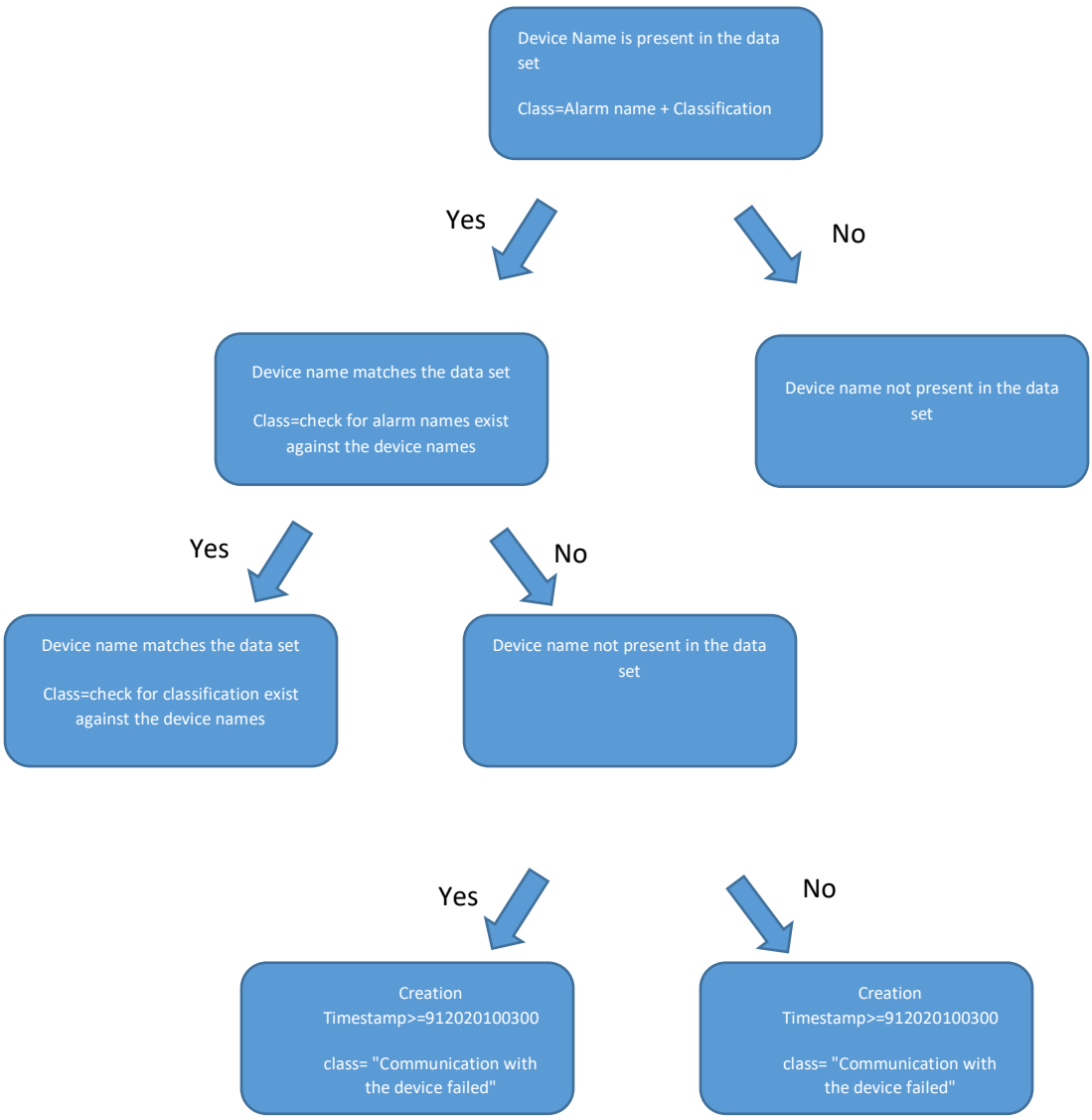


Fig 11: Flow diagram of decision tree working for this example.

Results:

As you can see from earlier topic, since we had two data variables as an input for the algorithm therefore use of simple algorithm like decision tree for prediction in our program was a good decision. It successfully predicted alarms on specific machine at specific timestamp. With more data attributes and other algorithm we can build more advance version of this alarm prediction program.

Beforehand prediction of alarms on specific nodes/machines will turn out to be a major help in telecommunication industries where a slight decrease in the service may result in the customer churn for the company. Field teams may be given alert to prepare for any type of alarm that is to occur in a particular machine.

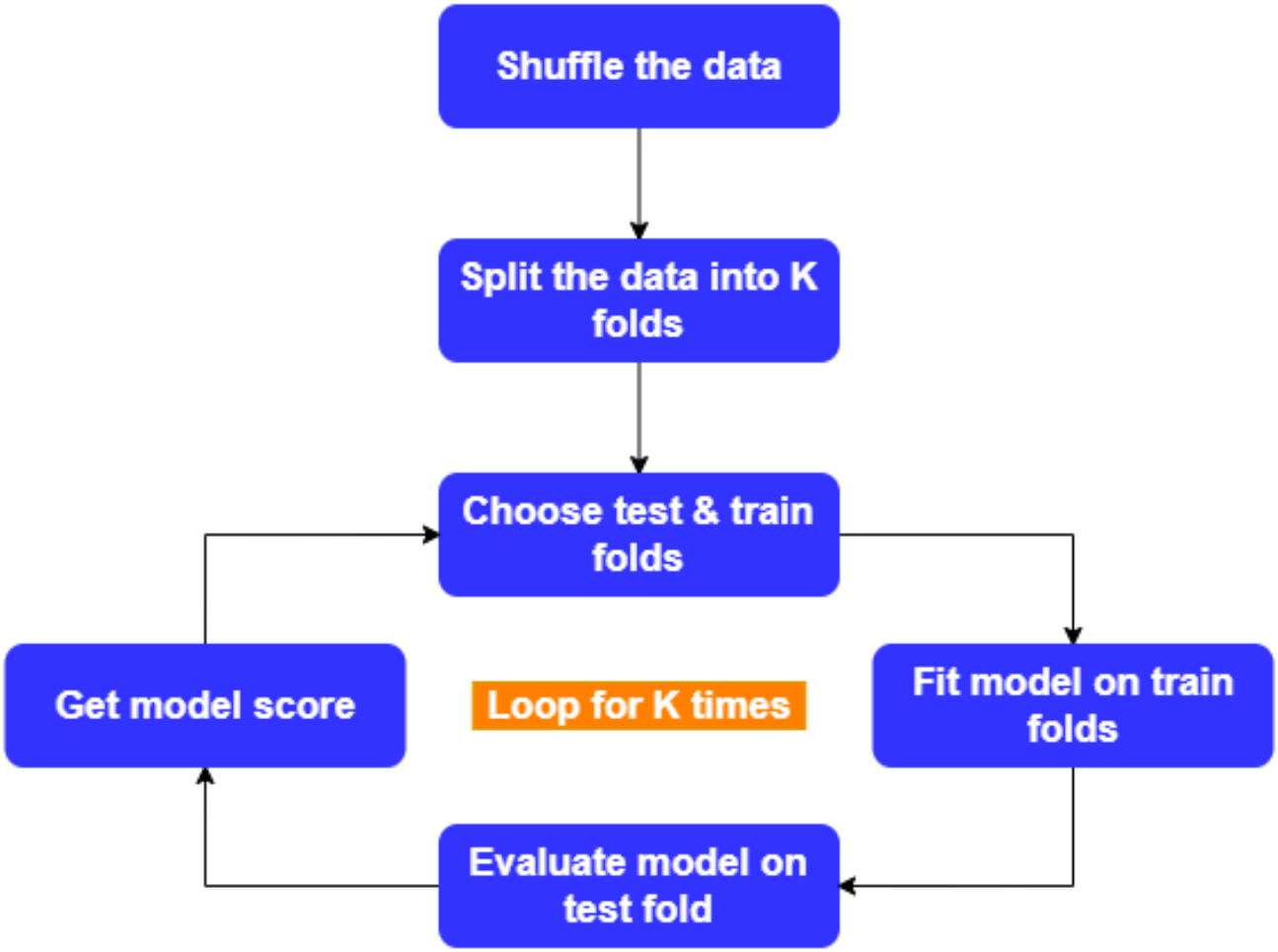
An alert system may be integrated with this prediction outputs that will generate SMS or emails to the respective in charge of the areas where machines are placed so that early actions are to be taken to avoid any problem. One more advantage of such program is that outage alarms may also be controlled via these predictions. Outage alarms are caused when the device/machine completely shut down causing a major delay in the communication services provided to the customers, thus resulting in unhappy customers. These outages alarm when appearing on a device means that device is faulty and quick action is required to rectify the problem.

10-fold Cross Validation:

10 fold cross validation process means input data is split into 10 groups of separate data inputs instead of just two groups. For each time one group is been used for testing while all the remaining will be used for training of the algorithm.[19]

Following steps are included for 10-fold validation process:

1. First it is required to shuffle all the data present so that any order can be removed from it.
2. Split your data into 10 small groups
3. Now take one fold for testing and remaining all the folds for training of algorithm
4. Train the algorithm and test on that one fold and note down the results
5. Now repeat whole process by using the next fold.
6. And at the end sum up all the score and get the mean score.



Following are the results:

Accuracy(1)=0.668

Accuracy(2)=0.706

Accuracy(3)=0.713

Accuracy(4)=0.678

Accuracy(5)=0.688

Accuracy(6)=0.682

Accuracy(7)=0.669

Accuracy(8)=0.669

Accuracy(9)=0.663

Accuracy(10)=0.664

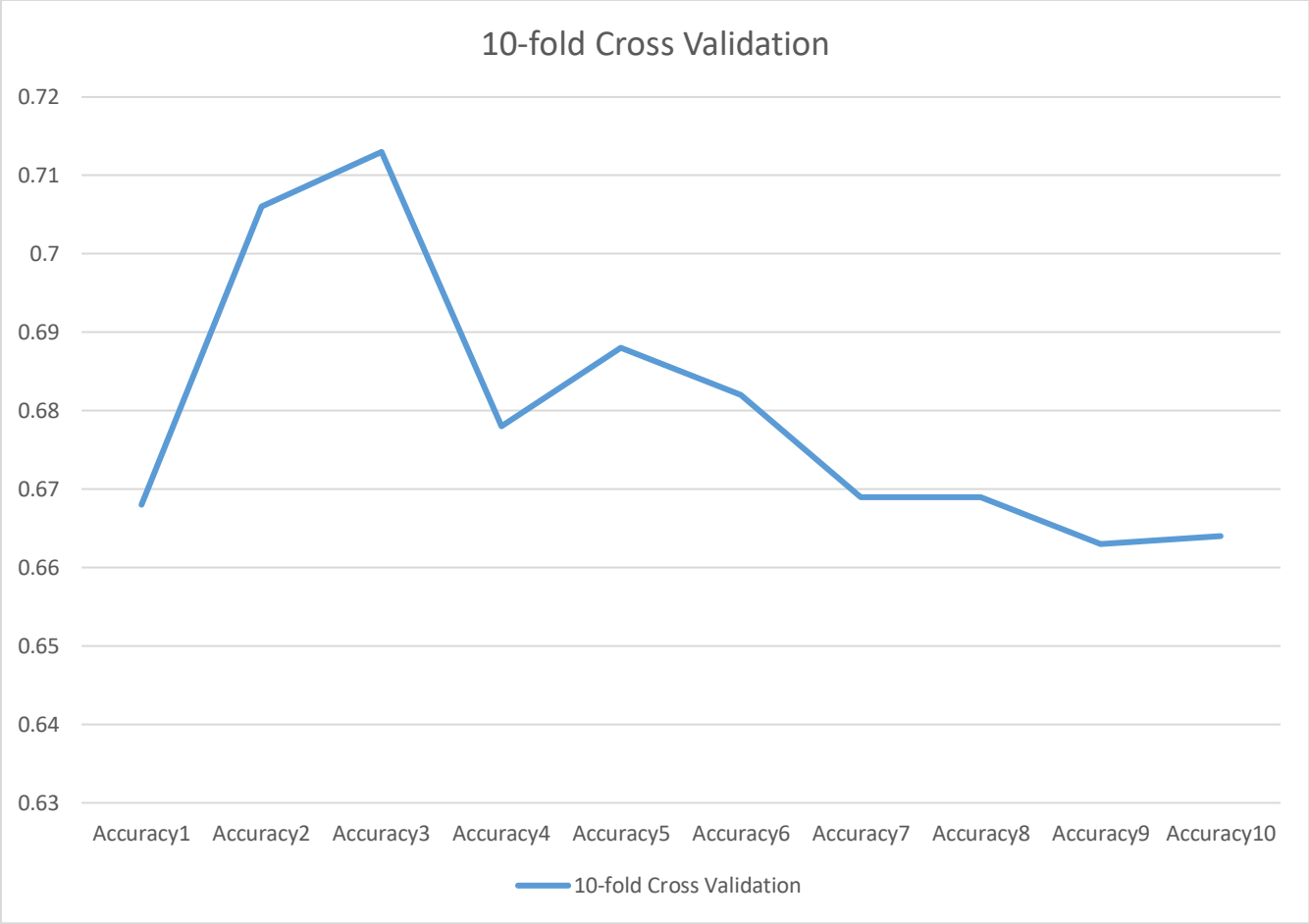


Table 1: Showing Accuracy trend for 10-fold data analysis

It is observed that there is a consistency in the accuracy for all the 10 fold data used. Since only one fold is used for testing while the remaining were used for the training of data.

Now lets observe accuracy measurements of our algorithm for difference proportions of data used for training and testing.

If we train algorithm with 80% of the data and 20% is kept for testing then the accuracy is :

Accuracy=0.998

If we train the agorithm with 20% of the data and 80% is been used for testing then the accuracy is:

Accuracy=0.421

If we train the algorithm with 60% of the data is been used for training and 40% of the data is been used for testing, then accuracy is:

Accuracy=0.489

If we train the algorithm with 40% of the data is been used for training and 60% of the data is been used for testing, then accuracy is:

Accuracy=0.411

If we train the algorithm with 50% of the data is been used for training and 50% of the data is been used for testing, then accuracy is:

Accuracy=0.444

Let’s see the graph below:

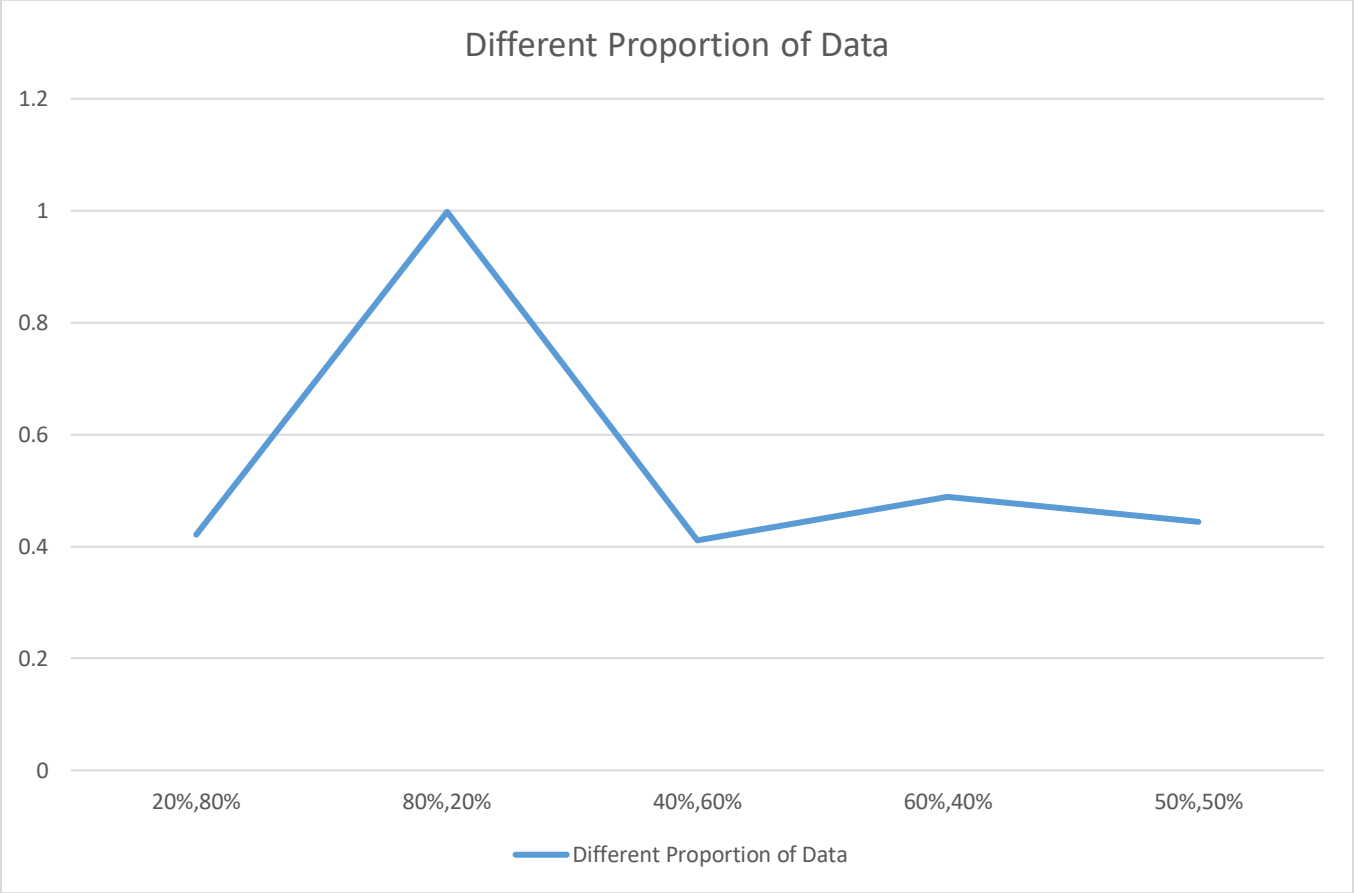


Table 2: Showing trend between accuracy and different proportions of training & testing data sets

Comparison of Decision Tree Algorithm with other Algorithms:

	Decision Tree	Support Vector Machine	Naïve Bayes	K-Nearest Neighbor Algorithm	Artificial Neural Network Algorithm
Computational Complexity	Computational complexity is not high	Complex calculation when there are many class labels	Sensitive to pre-processing of data input	Well suited for multimodel classes	Requires high processing time if neural network is large
Easy to understand output	Output is easy to understand and to interpret(i.e using tree diagram)	Suitable when the sample size is smaller than the number of dimensions	Easy to extend to multi-class classification problem	Time to find the nearest neighbors in large training data set can be excessive	Difficult to know how many neurons and layers are necessary
Numerical and categorical data	It can easily handle numerical and categorical data	Not usually emplyed for continuous numerical variables, mostly for categorical variables	only for categorised data	It is senitiv eto noisy or irrelevant attributes	Applicable to wide range of problems in real life
Fast, efficiency	Efficiant and easy to implement	When the number of features larger than the number of samples,it is crucial to choose suitable Kernel function	Easy to implement	Performance of algorithm depends on the number of dimensions used	Learning can be slow

Fig 12: A comparison of different algorithms with Decision tree algorithm [17, 18]

Discussion:

We have only worked on one of the use case that predicting the alarms that are to be appearing on a specific machine at specific time. There are several use cases that this project may be extended on. For example, algorithm can be used to calculate life span of a machine so that early actions could be taken by the company to replace the machine in order to have uninterrupted communication.

Faulty cards installed at the machine could also be predicted against specific alarm, if we will provide it along with data set.

Conclusion:

In a nutshell, a decion tree is used to predict the alarms that are to be appeared on a specific machine and on a specific time stamp.

Future Work:

There are several use cases that this project may be extended on. For example, algorithm can be used to calculate life span of a machine so that early actions could be taken by the company to replace the machine in order to have uninterrupted communication.

Faulty cards installed at the machine could also be predicted against specific alarm, if we will provide it along with data set.

References:

[1]: A Survey of Machine Learning for Network Fault Management Mourad Nouioua, Philippe Fournier-Viger, Ganghuan He, Farid Nouioua, Zhou Min

[2]: B. M. Patil, D. Toshniwal, and R. C. Joshi, (2009) “Predicting burn patient survivability using decision tree in weka environment”, IEEE International Advance Computing Conference, pp1353- 1356.

[3]: P. K. Srimani, and K. Balaji, (2014) “A comparative study of different classifiers on search engine based educational data”, International Journal of Conceptions on computing and Information Technology, Vol. 2, pp 6-11.

[4]: Cloud Mantra, Fault Detection Using Machine Learning Techniques, February 28, 2020 in Machine Learning

[5]: altexsoft, Data Science vs Machine Learning vs AI vs Deep Learning vs Data Mining: Know the Differences, 26,Jan, 2021

[6]: A Survey of Machine Learning for Network Fault Management, Mourad Nouioua, Philippe Fournier-Viger, Ganghuan He, Farid Nouioua, Zhou Min

[7]: Lozonavu, M., Vlachou-Konchylaki, M., Huang, V.: Relation discovery of mobile network alarms with sequential pattern mining. In: 2017 International Conference on Computing, Networking and Communications (ICNC), pp. 363–367. IEEE (2017)

[8]: T. Velmurugan, A State of Art Analysis of Telecommunication Data by k-Means and k-Medoids Clustering Algorithms, (2018)

[9]: Dharmarajan, A. and Velmurugan, T. (2016) Effi-ciency of k-Means and k-Medoids Clustering Algorithms Using Lung Cancer Dataset. Int. Journal of Data Mining Techniques and Applications, 5, 150-156.

[10]: Jocelyn Edinio ZackoGbadoubissaacAdo Adamou AbbaAribcAbdelhak MouradGueroui, Efficient k-means based clustering scheme for mobile networks cell sites management

[11]: Journal: International Journal of Advanced Computer Science and Applications(IJACSA), Vol. 10, No. 9, 2019, Customers Churn Prediction using Artificial Neural Networks (ANN) in Telecom Industry, Yasser Khan, Shahryar Shafiq, Abid Naeem, Sabir Hussain, 2019

[12]: Journal of Risk and Financial Management, 2021, Churn Management in Telecommunications: Hybrid Approach Using Cluster Analysis and Decision Trees, Mirjana Peji´c Bach * , Jasmina Pivar and Božidar Jakovi´c

[13]: Wikipedia, Decision Tree,

[14]: SdxCentral studios, Huawei Network Management System (NMS), March 27, 2020

[15]: Planet networking and communication, Intelliget Network Managemnet System

[16]: TeMIP Client User's Guide, Unified Correlation Analyzer for EBC Problem Detection,HP, hp Unified Correlation Analyzer,2015, Edition 1.0

[17]: Java T point, Decision Tree Classification Algorithm

[18]: Research gate, Functional and structural biomarkers of cognitive outcomes after brain tumor resection, Sep 2020

[19]: Sagar S. Nikam*, A Comparative Study of Classification Techniques in Data Mining Algorithms, 2015

[20]: Satish Gunjal, K Fold Cross Validation, Quality Tech Tutorials