

# Predicting the System Failures using Machine Learning Algorithms

Rashmi T V

Assistant Professor, Department of CSE  
East Point College of Engineering And Technology, Bangalore,  
Karnataka, India  
rashmitv.harsha@gmail.com

**Abstract - Quick recuperation stays one of the key difficulties to architects and administrators of vast organized framework. Before recuperation can happen, notwithstanding, one should first identify and analyze the disappointment, so without arrange disappointment expectation it is hard to recoup the failure. Failure in any piece of the system result in slow down the system. Failure Prediction helps in early detection of failure in the system, so that we can avoid the damage caused by the failure in that particular system. By predicting the failure we can avoid loss of data, application services. By prediction we can also decide if the framework will breakdown as indicated by the examination on the assessment of authentic investigation and current conduct of the framework.**

**Keywords :** Machine learning, Decision tree, SVM, Random forest.

## I. INTRODUCTION

Network analysis can offer insight into what is occurring at each critical point of the network. With the development of the Internet, numerous devices are network-connecting and a more complex network structure is formed in which network failures become inevitable. Recognizing and pinpointing potential system disappointments and execution issues as long involved instructed mystery, however a rising age of prescient examination apparatuses guarantees to get more noteworthy exactness systems.

The significant explanation behind facing network failures are shortcomings, mistakes or disposes of in organize gadgets, gadget arrangement changes, operational human blunders and botch gadgets, power outages, server hardware failure, security assaults, for example, foreswearing of administration. Failure of critical application can lead to a few sorts of misfortunes: loss of application service, loss of information or data.

As system frameworks develop in size and many-sided quality it turns out to be progressively unfeasible to look at every segment physically for wellsprings of mistake. Manual determination is tedious, mistake inclined, requires much ability, and doesn't scale. A mechanized approach is fundamental in the event that we need to proceed at the

present advancement rate and enhance framework accessibility. In this way it is beneficial and profitable to assess the status of the system and anticipate the likelihood of system disappointments before executing a critical assignment. Disappointment expectation is a strategy to decide if the framework will breakdown as indicated by the investigation on the assessment of chronicled study and current conduct of the framework [1]. Decision Tree, Random Forest, SVM, Logistic Regression, and Naive Bayes all these methods are used for predicting potential failures in the future. Disappointment expectation contemplate is of uncommon noteworthiness in decreasing the weight of system administration and support, and in limiting misfortune caused by organize disappointment.

Fault or failure prediction will bolster the action of system organization and may help the head on playing out a specific activity to take care of the issue that will occur and it is also important for parts of network management system and is in charge of recognizing and distinguishing shortcomings in the system. Enthusiasm for blame administration has been increased over the past decade due to the developing number of network that have become a critical component of the infrastructure of many organizations in order to avoid making faults and long downtime that costs higher than ever.

### A. Contribution:

To avoid failure in system we use prediction algorithms to predict the failure in advance.

Failure Prediction think about is of mind blowing hugeness in reducing the weight of system administration and support, and in limiting misfortune caused by arrange failure.

## II. Related Work

In 2016 Jiang ZHONG, Weili GUO, Zhenhua WANG proposed Concentrate on arrange disappointment forecast in light of alert logs. To maintain a strategic distance from erratic misfortunes as a result of system disappointment, the unwavering quality of system should be assessed in a few applications schema. This starts the system disappointment expectation inquire about on alarm logs. The algorithms used are Random forest, Ripper, BayesNet, Weibull. The

examination comes to fruition also show that most framework dissatisfaction can be foreseen through exploring system logs with information mining strategy.

In 2017 Yuling Bai, Yunhua Li et.al. Proposed the Here and now forecast of appropriation arrange flaws in light of help vector machine. As the system end of intensity transmission, the dissemination organize (DN) straightforwardly decides the dependability of power vitality supply. To foresee disappointment precisely is vital for expanding the repair productivity of DN. In light of the disappointment information from DN in Beijing, the paper examines here and now DN disappointments expectation and proposes a blame judgment program in view of climate and season factors. Disappointment is dissected to decide its most vital elements. Through help vector machine (SVM) calculation and thinking about the relative meteorological variables, utilizing the arrangement display predicts the quantity of disappointments in DN week by week, and sets up sub-locale grouping gauging model in week recurrence with meteorological impact for DN disappointments expectation.

The outcomes are more tenable quite far to utilize less characterization when the forecast precision can meet the real prerequisites. Be that as it may, the quantity of DN disappointment factors is various, and numerous elements can't be measured or information can't be made at show, so the gauge exactness is hard to achieve 100%. In any case, with the expansion of information maintenance and information accumulation, the model can include more disappointment factors, at that point the expectation exactness will be enhanced, which will manage DN repair work better and better [12].

Analyzed about Prediction of Failure in Lubricated Surfaces Using Acoustic Time-Frequency Features and Random Forest Algorithm. Algorithm used is Random Forest, Scuffing cause's fatal damage of the surfaces in contact driving to the loss of their functionalities as well as regularly to high reparation costs as well as deferral underway lines to the distinctive grinding administrations was separately researched. The fittingness of machine learning characterization and relapse was contemplated for scraping expectation. The two methodologies were connected independently however can be brought together to expand the expectation time interim of surface disappointment.

In 2014 Jose M. Navarro, Hugo A, et.al. Proposed the Framework nonperformance Prediction through Rare-Events Elastic-Net Logistic Regression. Algorithm used is Logistic Regression, Online abortion forecast. We portray a disconnected prepared strategy we have created keeping in mind the end goal to conjecture online disappointments model creation algorithm is proposed, making use of a blend of versatile net regularization, uncommon occasions expectation procedures and two layers of cross-approval is

an intense instrument for determining failures in a distributed system [14].

### III. Existing system

Decision tree algorithm is commonly used to predict system failure. It is a administered learning calculation. It works for both continuous as well as categorical output variables. The understanding level of Decision Trees calculation is so natural contrasted and other characterization calculations.

Choice tree learning is the change of a choice tree from class-stamped arranging tuples. A choice tree is a flow diagram like architecture, where each inside (non-leaf) center point means an analysis on a characteristic, each branch addresses the aftereffect of an analysis, and each leaf (or terminal) center point holds a class name. The most elevated center point in a tree is the root center.

#### Drawbacks of Decision tree

Decision tree are unstable for prediction. They are regularly generally off base. Figuring can get extremely perplexing, especially if numerous qualities are dubious as well as if numerous results are connected. Not reasonable for expectation of ceaseless trait. Perform ineffectively with numerous class and little information.

### IV. Proposed System

Arbitrary Forest is a gathering of Decision Trees. Irregular Forest is a standout amongst the most prominent and most ground-breaking machine learning calculations. As a rule, the more trees in the woodland the more solid the inside takes after. Identical in the arbitrary woods classifier, the greater the amount of trees in the timberland gives the great exactness happens. The difference between Random forest algorithm also, the choice tree calculation is that in Random Forest, the procedure of finding the root hub and part the component nodes will run randomly.

On the off chance that we need to foresee some value using the decision tree, it will generate some rules based on the past actions, whereas in case of arbitrary woodland calculation arbitrarily chooses perceptions and highlights to construct a few choice trees and after that midpoints the outcomes.

#### Points of interest of Random Forest:

A comparative unpredictable timberland count or the discretionary forest classifier can use for both portrayal and the backslide task. Arbitrary backwoods classifier will

manage the missing characteristics. When we have added trees in the timberland, unpredictable backwoods classifier won't Over fit the model. Random forest Machine Learning Algorithm maintains accuracy even when there is inconsistent data and is simple to use.

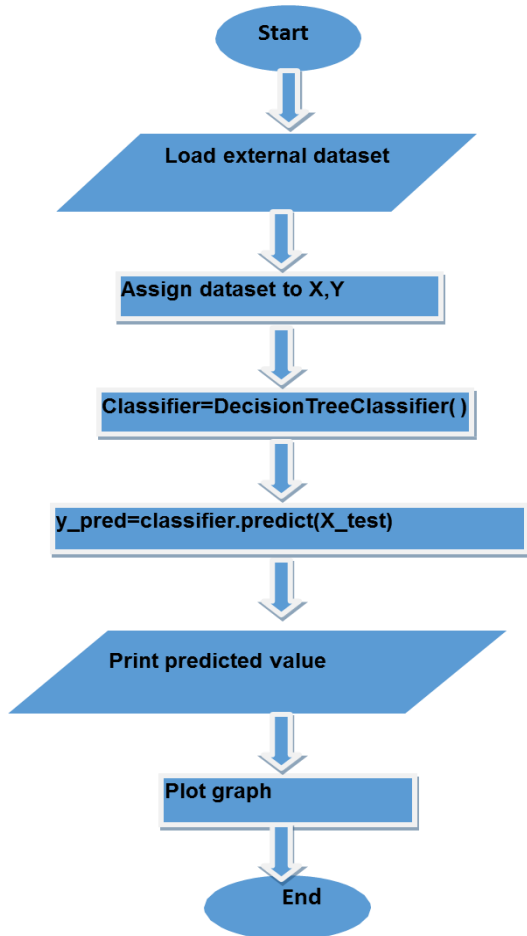


Fig 1: Flowchart for Decision Tree model

SVM is a directed learning technique that takes a gander at information and sorts it into one of two classifications. In a SVM information focuses are spoken to as focuses in space such that focuses from various classifications are isolated by a plane. New information are mapped into a similar space and their area in respect to the plane is utilized to anticipate which classifications each point has a place, with the plane being alluded to as the choice limit ( i.e., deciding to which class the information has a place ).

For the situation where the choice limit should be non-direct( i.e., where classes can't be isolated by a straight line), SVM's also have the ability to project space through a non-linear function, lifting the data to a space with a higher dimension where a linear boundary does separate classes.

**Experimental Results**

This section contains the modules Decision Tree, Random Forest, Support Vector Machine.

**Scenario 1: Decision Tree**

**Confusion Matrix**  $\begin{bmatrix} 86 & 1 \\ 0 & 6 \end{bmatrix}$

	Precision	Recall	F1 Score	Support
0	1.00	0.99	0.99	87
1	0.86	1.00	0.92	6
Avg/Total	0.99	0.99	0.99	93

**Scenario 2: Random Forest**

**Confusion Matrix**  $\begin{bmatrix} 107 & 2 \\ 1 & 6 \end{bmatrix}$

	Precision	Recall	F1 Score	Support
0	0.99	0.98	0.99	109
1	0.75	0.86	0.80	7
Avg/Total	0.98	0.97	0.97	116

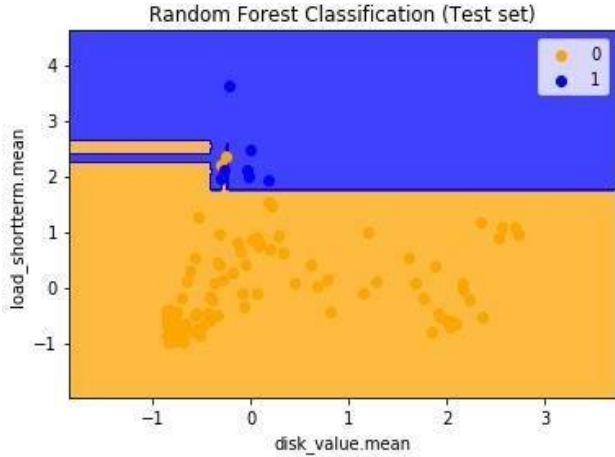


Fig 2 : Linear Graph for Random Forest

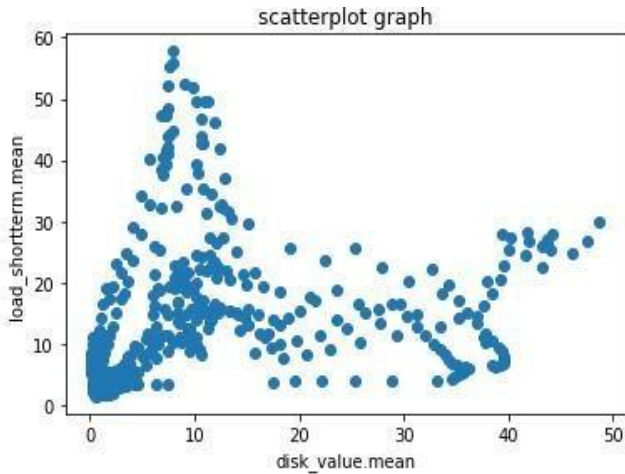


Fig 3: Scatter Plot for Random Forest

Scenario 3: SVM

Confusion Matrix  $\begin{bmatrix} 107 & 2 \\ 2 & 5 \end{bmatrix}$

	Precision	Recall	F1 Score	Support
0	0.98	0.98	0.98	109
1	0.71	0.71	0.71	7
Avg/Total	0.97	0.97	0.97	116

The above graph represents the Scatter Plot for SVM. X-axis represents the set of disk\_value.mean Y-axis represents the load\_shortterm.mean of our dataset.

V. CONCLUSION

In this undertaking we have played out the present investigation on Datasets. By applying predictive algorithms on the Dataset we can anticipate the failure in the system. Failure in the part of the system may damage the entire system process, hence predicting the failure in advance may help the system for their normal execution. So for prediction, we use prediction calculations like Choice Tree, Irregular Forest and Bolster Vector Machine algorithms. All these algorithms classify the datasets into two classes in light of the achievement and failure of the system, among these algorithms Random Forest Algorithm gives more accuracy.

In future we are planned to predict the failure in system in early stage. We are wanting to build up a web application that gives information about the failure. This project is useful for predicting a network failure in the system, so that we can avoid failure in the system at early stage.

REFERENCES

- [1] Zhiqiangcai, Weitaosi, Shubinsi and Shudong, Sun. Modeling of failure prediction Bayesian network with divide-and-conquer principle, North western polytechnic university, China 2014.
- [2] Sahana, D. S., and L. Girish. "Automatic drug reaction detection using sentimental analysis." International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 (2015).
- [3] Mike Chen, Alice X. Zheng, Jim Lloyd, Michael I. Jordan, Eric Brewer University of California at Berkeley and eBay Inc. Failure diagnosis using decision trees, 2015.
- [4] Jihad Ali, Rehanullah khan, Nasir Ahmad, Imran maqsood computer science engineering, UET Peshwar, Pakistan. Random forest and decision trees, 2012.
- [5] Yuling Bai, Yunhua Li School of Automation Science and Electrical Engineering Beihang University Beijing, China, Yongmei Liu, Zhao Ma Power Distribution Research Department China Electric Power Research Institute Beijing, China. Short-term Prediction of Distribution Network Faults Based on Support Vector Machine, 2017.
- [6] L. G. (2019). "Anomaly Detection in NFV Using Tree-Based unsupervised Learning Method". International Journal of Science, Technology, Engineering and Management - A VTU Publication, 1(2), Retrieved from <http://ijesm.vtu.ac.in/index.php/IJESM/article/view/232>
- [7] Rashmi, T. V., and Keshava Prasanna. "Load Balancing As A Service In Openstack-Liberty." International Journal of Scientific & Technology Research 4.8 (2015): 70-73.
- [8] ZHONG, Weili GUO, Zhenhua WANG Study on network failure prediction based on alarm logs, 2016.
- [9] Yuling Bai, Yunhua Li, Beijing, china, Yongmei Liu, Zhao Ma Short-term prediction of distribution network faults based on support vector machine, 2017.