FAULT PREDICTION USING NAIVE BAYES CLASSIFIER BASED LEARNING MODEL

MINI PROJECT REPORT

SUBMITTED BY:

DHARANIDHARAN K

DHARSHINI S

UNIVERSITY REGISTER NUMBER:

910022104037

910022104007

Under the guidance of

DR.PROFESSOR A.MUTHUMARI

in partial fulfilment of requirement for award of Degree of

BACHELOR OF ENGINEERING

IN

Computer Science and ENGINEERING

Anna University Regional Campus, Madurai

Anna University: Chennai 600 025

June 2024

Anna University Regional Campus, Madurai

Department of Computer Science and Engineering

Declaration

I undersigned, hereby declare that the Mini project titled "**Fault Prediction Using Naive Bayes Classifier Based Learning Model**" submitted in partial fulfilment for the award of Degree of Engineering of Anna University Regional Campus, Madurai is a bonafide record of work done by me under the guidance of **Dr. A. Muthumari of Department Of Computer Science And Engineering in Anna University Regional Campus Madurai**. This report has not been previously published for the award of any degree, diploma, or similar title of any University.

Date:09/06/2024

Dharanidharan K (910022104037)

Dharshini S (910022104007)

Anna University Regional Campus, Madurai

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the report titled "**Fault Prediction using Naive Bayes Classifier Based Learning Model**" being submitted **by** ***Dharanidharan K (910022104037) and Dharshini S (910022104007)***, in partial fulfilment of the requirements for the award of the of Degree of Engineering, is a bonafide record of the project work done by ***Dharanidharan K 910022104037 and Dharshini S 910022104007*** of **Department** ***of Computer Science And Engineering &*** **Anna University Regional Campus Madurai**

Dr.Professor A.Muthumari

Assistant Professor

Acknowledgements

Through this acknowledgement, I express my sincere gratitude towards all those who helped me in this project, which has been a significant learning experience.

I extend my heartfelt appreciation to my project partner, whose dedication, hard work, and collaboration were pivotal to the success of this project. Your insights, support, and teamwork have made this endeavor a rewarding journey.

Special thanks to our project guide, **Dr. A. Muthumari, Assistant Professor of Department of Computer Science and Engineering, Anna University Regional Campus, Madurai** for his invaluable efforts in coordinating our work and providing guidance in the right direction. Your mentorship has been instrumental in our progress.

I also wish to acknowledge the encouragement and support received from our classmates and friends. Without their assistance, our project would not have reached its full potential.

Lastly, I offer my sincere love and gratitude to my parents and all others who have supported us throughout this project. Your unwavering support has been a constant source of motivation.

Dharanidharan K (910022104037)

Dharshini S (910022104007)

Title: Fault Prediction in Power Supply using Naive Bayes Classifier and Decision Tree Algorithm.

Introduction:

It is important to predict faults so that the power supply systems remain stable and reliable. We propose a method of predicting faults based on machine learning where we will pay more attention to Naive Bayes classifier and decision tree algorithm. Naive Bayes is a simple classical approach while decision tree algorithm has been found to have higher accuracy in prediction of faults for power supply systems.

Problem Statement:

Power supply systems experience many different types of faults like voltage sags/swells, line outages and equipment failures. These faults can result into power loss, damage of equipment or pose a threat to safety. Manual inspection and reactive maintenance are not only time-consuming but also ineffective. Therefore, automated fault prediction methods should be used to identify potential problems before they occur.

Work Flow:

Data Collection: Collect historical data about power supply systems which includes voltage levels, currents readings, status of equipments among others record if there are any faults during this period.

Data Preprocessing: Clean the data by filling in missing values, removing outliers and noise.

Selection of Features: Select features that are highly correlated with occurrence of faults e.g., voltage sags/swells, changes in equipment status etc.

Model Training: Use prepared dataset to train Naive Bayes Classifier and Decision Tree Algorithm models.

Model Evaluation: Check accuracy, F1 score, precision, recall and sensitivity when evaluating the performance of these models.

Prediction of Faults: Based on the trained models predict potential power supply system failures.

Objective:

The study aims at creating a reliable fault prediction model for power supply systems with high accuracy using different machine learning algorithms. This is because decision tree algorithm has been seen to be more accurate compared to naive bayes classifier in our case.

Planned Algorithm:

Naive Bayes Classifier:

Uses Bayes' theorem, with predictors all assumed to be independent of one another.

Even though it's simple, Naive Bayes works well in practice; this is especially true when dealing primarily with categorical data types.

Decision Tree Algorithm:

Divides the input space into regions and assigns a label based on which class occurs most often among training instances within that region.

They are easy to interpret and can effectively handle numerical as well as categorical inputs.

Algorithm Integration:

These two algorithms – Naive Bayes Classifier and Decision Tree Algorithm – form an integrated fault prediction framework. While they learn from the same data set, Decision Tree Algorithm is preferred for its higher accuracy in fault prediction task.

Result Consideration:

Performance Measures:

Accuracy - the percentage of correctly predicted cases

F1 score – Harmonic mean between precision and recall thus balancing false positives and false negatives

Precision: The ability not to provide false alarms i.e {True Positives}/{(True Positives + False Positives)}

Recall: The ability not to miss positives i.e {True Positives}/{(True Positives + False Negatives)}

Sensitivity or True Positive Rate which indicates how many actual positives were identified correctly vis a vis all potential positives that might have been found This would be useful for finding rare events contained within noisy backgrounds such as signal detection problems in multiple hypothesis testing procedures where we have many hypotheses being tested simultaneously against some observed data

Comparative Analysis:

Compare Naive Bayes Classifier with Decision Tree Algorithm based on evaluation metrics; explore trade-offs between accuracy and interpretability.

Visualization:

Visualize decision boundaries and classification results so as to understand model behaviors better.

Pros:

When it comes to predicting faults in power supply systems, the Decision Tree Algorithm is superior to the Naive Bayes Classifier in terms of accuracy.

The Naive Bayes Classifier is efficient in terms of computation and it is robust against irrelevant features.

Cons:

In some cases, the assumption of independence among predictors by Naive Bayes may not always be true.

When dealing with complex data sets, Decision Trees may overfit easily.

Usage:

It can be used to forecast failures in power supply systems so as to prevent any downtime and ensure reliability.

Conclusion and Future Scope:

Thus we see that when it comes to accuracy of fault prediction in power supply systems, decision tree algorithm outperforms naive bayes classifier. Future works can include looking at ensemble methods as well as hybrid models for better performance. Moreover integrating real time data streams with more advanced anomaly detection techniques would make these predictions systems stronger against faults effectiveness.