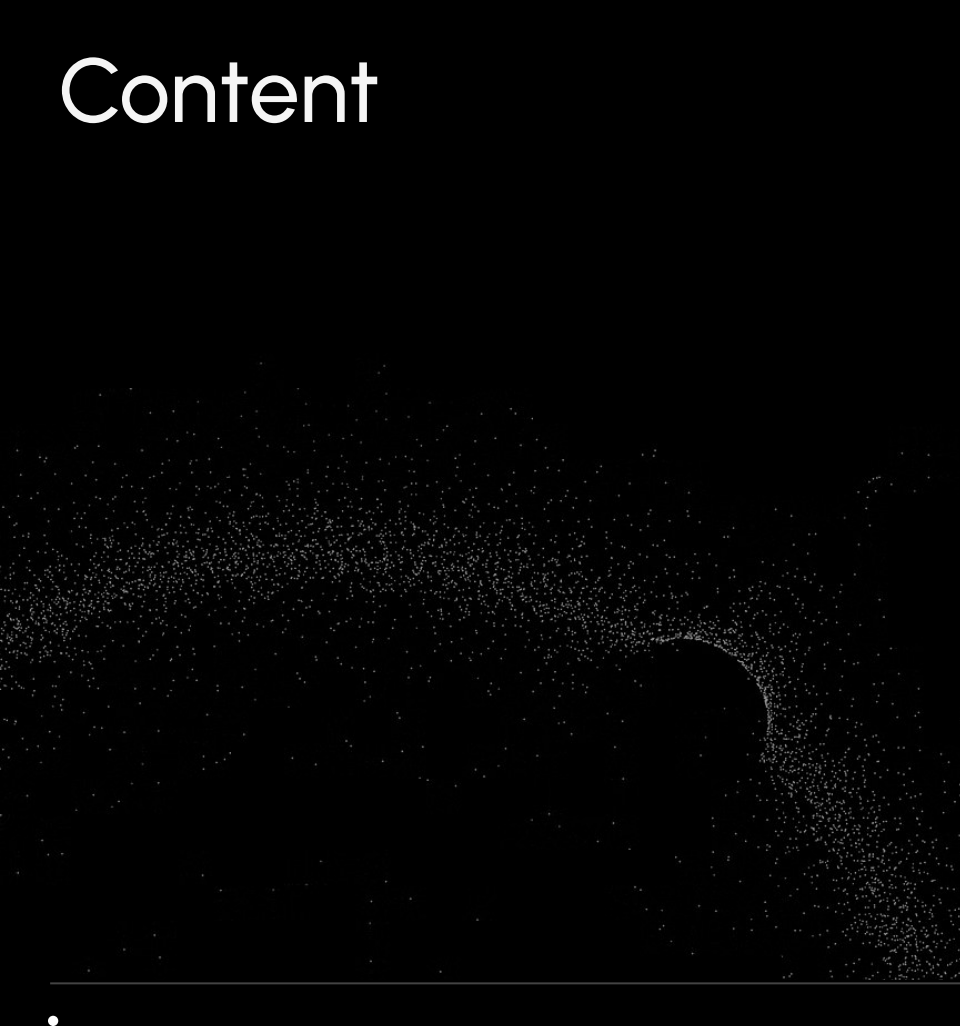


Predicting Software Defects Using Machine Learning Techniques

Presented by

Md Nabil Rahman

Content



1 Research context/background



2 Research Problem/Literature Gap



3 Research Aim and Objectives



4 Research methodology



5 Project planning



6 Project impact and limitations



7 Reference



Research context/background

- Software defects increase cost, delays, and reduce system reliability
- Traditional testing is time consuming and often detects defects late
- To improve the quality of software, early defect prediction is essential.
- Machine learning enables automated prediction using historical metrics
- ML models are used in this study to identify modules that are prone to detection.

Research Problem/Literature Gap

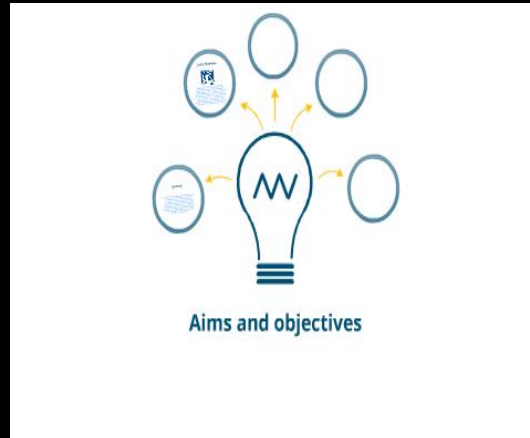
- Limited Industrial and Real-Time Validation,
- Risk of Overfitting,
- Limited Dataset Generalization
- Model Robustness and Stability
- cost- sensitive learning
- Synthetic Data Dependency (SMOTE)

Review paper: [1] "A Novel Approach to Improve Software Defect Prediction Accuracy Using Machine Learning"

- The research methodology for this paper was experimental, comparative and Quantitative Research Approach.
- Supervised machine learning was used.
- The primary contribution is feature selection increases the accuracy of defect prediction.
- Logistic Regression/J48 dataset dependant is the best-performing model.

Research Aim and Objectives

The purpose of this study is to find out how the effectiveness of machine learning techniques in predicting software defects.



- To review existing literature on software defect prediction.
- To implement selected models using a public dataset.
- To evaluate and compare model performance using standard metrics.

Research methodology

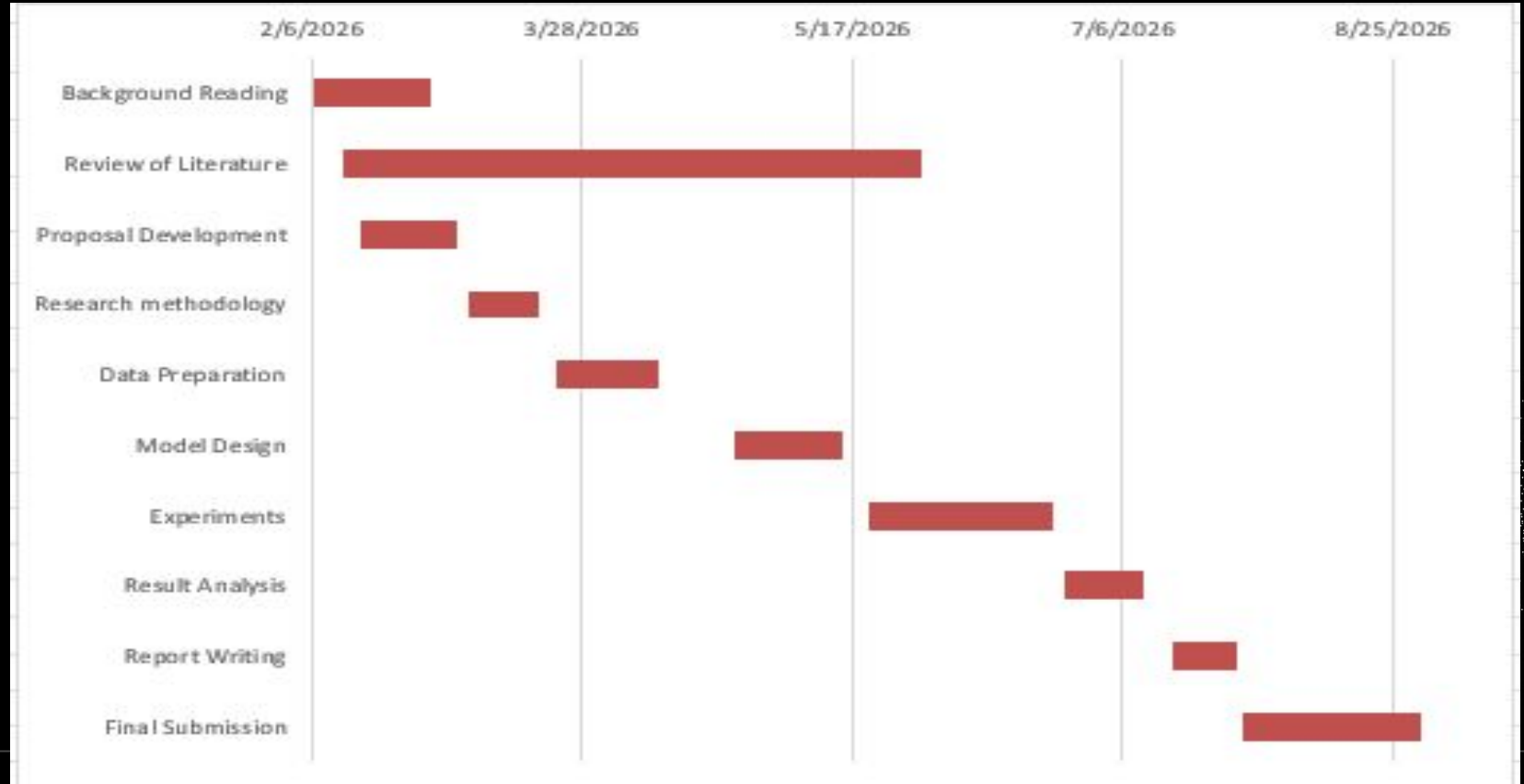
- The research methodology for this study will be experimental and quantitative research approach.

- Machine learning models Random Forest, Decision Tree, and Logistic Regression will be implemented using Jupyter Notebook.

- Publicly accessible software defect databases, such the NASA PROMISE repository will be used.

- Accuracy, precision, recall, and F1-score will be used to assess the model's performance.

Project planning



Project impact

- Finding the best machine learning model to predict software defects
- Showing how better software engineering practices can be supported by defect prediction

Project limitations

- Lack of categorized data
- high-quality historical data
- extremely unbalanced datasets
- overfitting to training data

References:

- [1] IQRA MEHMOOD, SIDRA SHAHID, HAMEED HUSSAIN SHAFIQ AHMAD , INAYAT KHAN , SHAHID RAHMAN ,NAJEEB ULLAH, (Member, IEEE), AND SHAMSUL HUDA. "A Novel Approach to Improve Software Defect Prediction Accuracy Using Machine Learning," in IEEE Access, vol. 11, pp. 63579-63597, 2023, doi: 10.1109/ACCESS.2023.3287326
- [2] M. Begum, M. H. Shuvo, I. Ashraf, A. A. Mamun, J. Uddin and M. A. Samad, "Software Defects Identification: Results Using Machine Learning and Explainable Artificial Intelligence Techniques," in IEEE Access, vol. 11, pp. 132750-132765, 2023, doi: 10.1109/ACCESS.2023.3329051
- [3] T. S. Kumar and B. Booba, "A Systematic Study on Machine Learning Techniques for Predicting Software Faults," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Hassan, India, 2021, pp. 133-136, doi: 10.1109/MysuruCon52639.2021.9641090.
- [4] S. B. Mohite, P. Semwal, J. Anand, M. M. Sati, A. Singh and N. Tripathi, "Applying Machine Learning to Predict Software Defects," 2024 International Conference on Advances in Computing, Communication and Materials (ICACCM), Dehradun, India, 2024, pp. 1-5, doi: 10.1109/ICACCM61117.2024.11059225.
- [5] C. Anjali, J. P. M. Dhas and J. A. P. Singh, "A Study on Predicting Software Defects with Machine Learning Algorithms," 2022 International Conference on Intelligent Innovations in Engineering and Technology (ICIET), Coimbatore, India, 2022, pp. 195-198, doi: 10.1109/ICIET55458.2022.9967593.



Thank you