PLAY STORE APP RATING PREDICTION USING CLASSICAL ML MODELS AND ARTIFICIAL NEURAL NETWORK

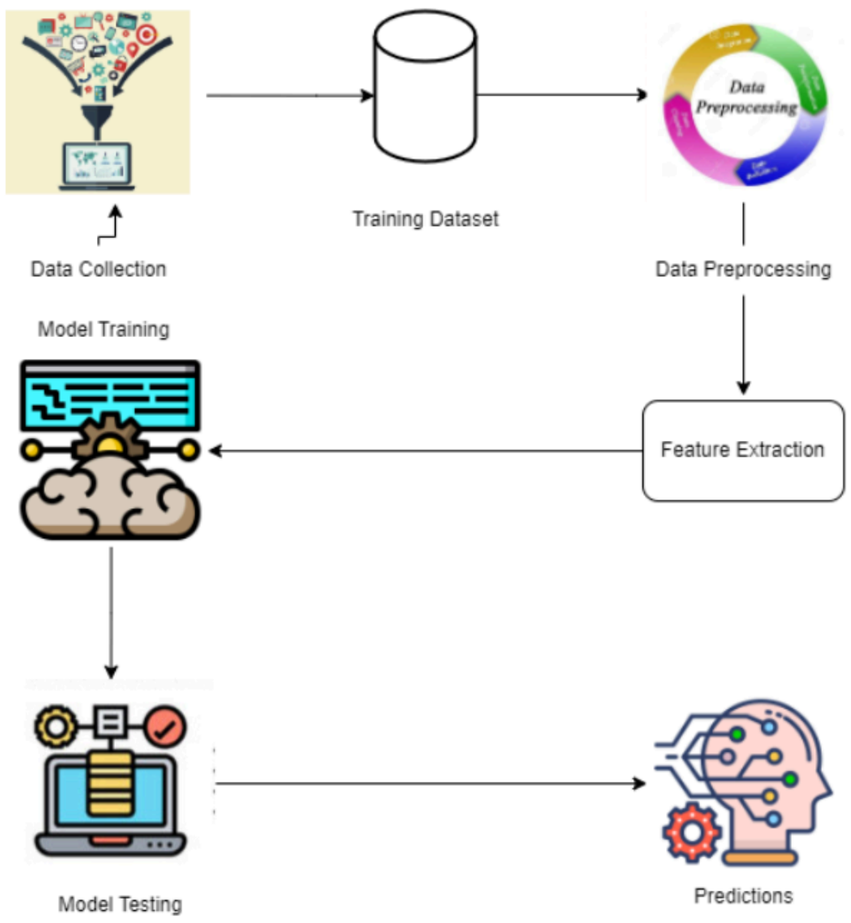
Introduction

App ratings on the Google Play Store significantly influence an app's visibility, user trust, and overall success. However, understanding and predicting these ratings is often challenging due to the complex interplay of factors such as app category, size, price, and user feedback. This study proposes a machine learning framework for analyzing and predicting app ratings using attributes from the Google Play Store. By leveraging advanced algorithms and neural networks, the framework provides accurate predictions and actionable insights for app developers.

Objective

This study aims to develop a machine learning framework for accurately predicting app ratings on the Google Play Store, leveraging app attributes such as reviews, installs, price, and size. The framework focuses on providing actionable insights for app developers, enabling improved app performance, user satisfaction, and market competitiveness.

Architecture

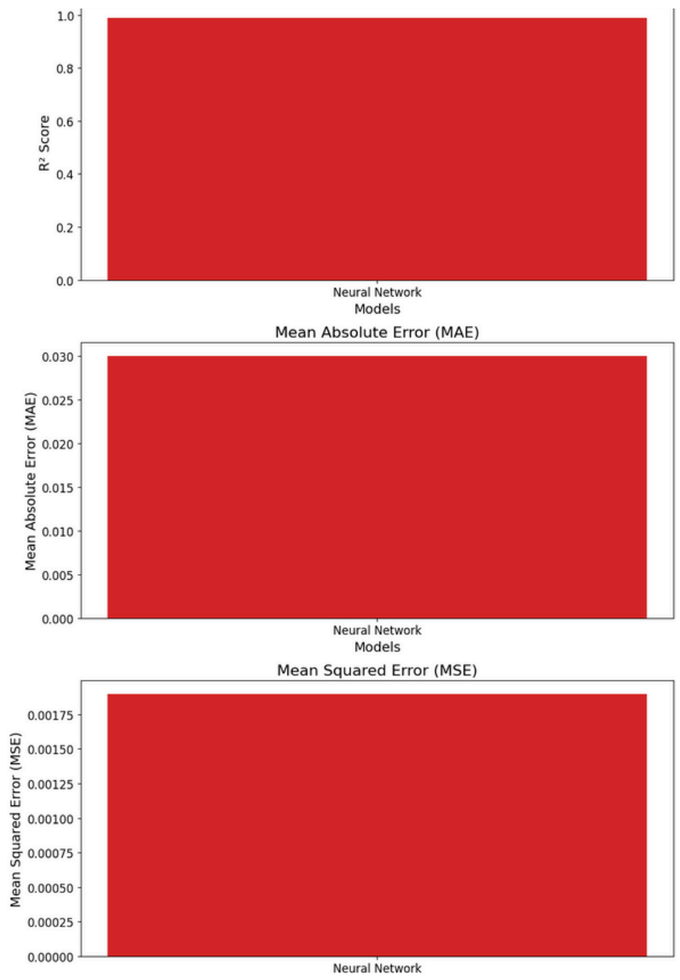


Methodology

- **Data Overview:** A dataset of 10,000 apps with features like reviews, installs, size, price, category, and type.
- **Preprocessing:** Handled missing values using imputation techniques. Scaled numerical features and applied one-hot and label encoding for categorical data.
- **Modeling:** Tested multiple machine learning models (Random Forest, XGBoost, Gradient Boosting, Neural Networks). Used an ensemble approach to combine strengths of different models.
- **Evaluation:** Metrics included R^2, Mean Absolute Error (MAE), and Mean Squared Error (MSE).

Evaluation and Results

Neural Networks performed the best by achieving the highest R^2 score of 0.99 during training with a Mean Absolute Error (MAE) of 0.0303 and a training loss reduced to 0.0019.



Discussion

The Neural Network demonstrated exceptional performance with a high R^2 score of 0.99 during training, making it the most effective model for predicting app ratings. Its ability to capture complex non-linear relationships between app attributes suggests strong potential for real-time rating predictions when applied to new apps. While the Gradient Boosting Regressor performed poorly, the Random Forest and XGBoost models showed reasonable results, although not as strong as the Neural Network. These findings emphasize the importance of model selection for this task and the promise of machine learning in optimizing app performance and enhancing user satisfaction.

Conclusion

This study highlights the potential of machine learning in app rating prediction, achieving robust performance in analyzing and predicting ratings based on app attributes. Insights into app characteristics offer actionable recommendations for developers to optimize app performance and enhance user satisfaction. Future work includes integrating temporal data and exploring advanced deep learning architectures to further improve prediction accuracy and generalization.

University of Engineering and Technology, Lahore

Presented by
Sami Ullah, 2022-CS-143