**Report Summary**   
**Dataset**

The dataset contains app reviews with binary sentiment labels (1 = positive, 0 = negative).

**Preprocessing**

Lowercased text

Removed punctuation and stopwords

Used TF-IDF vectorization

**Models Used**

Logistic Regression

Naive Bayes

**Results**

Logistic Regression typically performs better with balanced precision and recall.

Naive Bayes may be slightly faster but can misclassify edge cases.

**Final Thoughts**

Logistic Regression is a good baseline.

Performance can be improved by using deep learning models or tuning hyperparameters.

Further work could include sentiment strength (e.g., scale from 1 to 5).