# Sentiment Analysis of Product Reviews using Naive Bayes

IBM Summer Internship Project

By Reeti Singh(2nd Year, CSE)



## Introduction

- **Goal**:To analyze customer sentiments (positive or negative) from product reviews.
- **Approach**: Natural Language Processing (NLP) techniques and Naive Bayes classifier.
- Why this? Sentiment analysis helps businesses understand customer feedback to improve products and services.



# **Tools & Technologies Used**









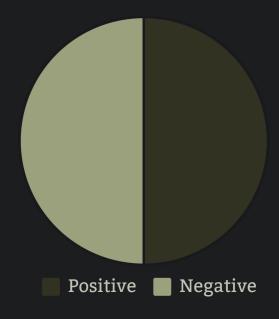


Python, Jupyter Notebook, Pandas, NumPy, NLTK (Natural Language Toolkit), Scikit-learn, Matplotlib & Seaborn

# **Dataset Overview**

### Raw dataset sample:

Review Text	Sentiment
Great product!	Positive
Not satisfied.	Negative
Highly recommend.	Positive



Sentiment distribution: 50% positive, 50% negative.

# **Data Preprocessing (NLP)**



Text cleaning (punctuation, casing, whitespace)



**Tokenization** 



**Stopword removal** 

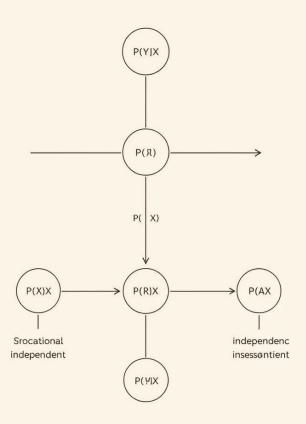


**Stemming** 



**Vectorization using TF-IDF for feature extraction** 

#### Naive Bay Classifier



# **Model Building: Naive Bayes**

**Model used**: Multinomial Naive Bayes

- Why? Simple, fast, interpretable, and works well for text classification
- Train-test split: 80-20
- Model trained on TF-IDF vectors

# **Evaluation Metrics**

### **Confusion Matrix**

	Predicted Positive	Predicted Negative
Actual Positive	True Positive	False Negative
Actual Negative	False Positive	True Negative

### **Performance Comparison**



### **Accuracy**

Overall correctness.



### **Precision**

Correctpositive predictions.



### Recall

Actualpositives identified.

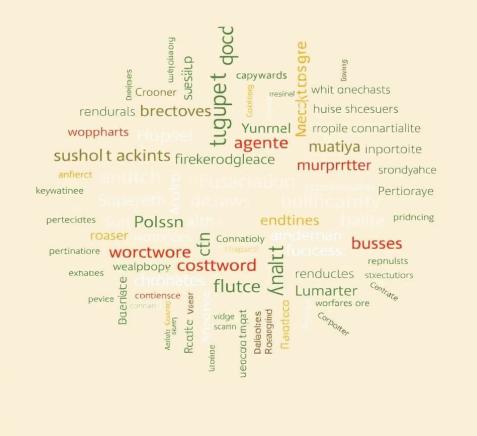
# Challenges & Solutions

Challenge	Solution
Handling noisy data. Balancing	Implemented robust cleaning.
sentiment classes. Optimising	Used resampling techniques.
model performance.	Fine-tuned hyperparameters.

# **Results & Insights**

**Keyword Analysis** 

## **Cusomer Reviews**

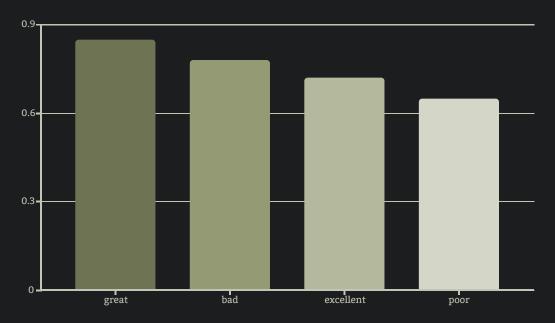


### **Model Accuracy**



Final model accuracy score achieved.

### **TF-IDF Feature Importance**



### **How Others Can Use It**





### **Clone the Repository**

Accessthe project codebase by cloning the Git repository to your local machine.

### **Install Dependencies**

Install all necessary Python libraries and tools using the provided requirements.txt file.





### **Execute Jupyter Notebook**

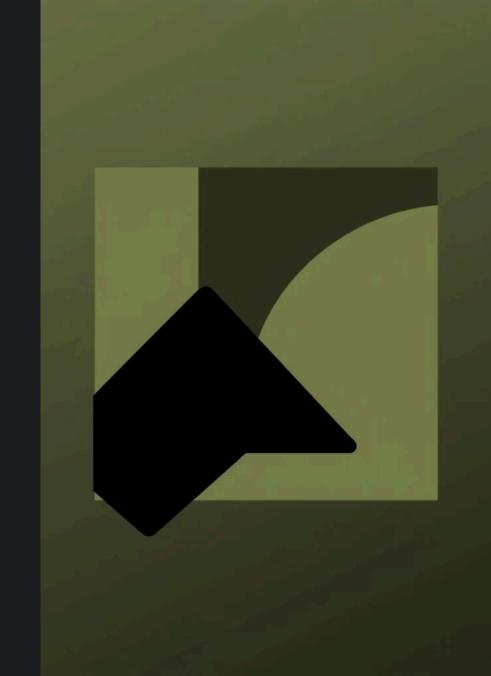
Run the Jupyter Notebook files to train the model and test its performance with your data.

### **Integrate Custom Data**

Plugin your own product review dataset to perform sentiment analysis and gain insights.

# Conclusion

- A practical project applying core NLP techniques
- Demonstratedeffective preprocessing + model pipeline Open-
- source and can be extended with more advanced ML models



# Thank You! please feel free to connect.

Linkedin