EMAIL SPAM CLASSIFIER USING NLP & MACHINE LEARNING



METHODOLOGY



Data Cleaning

Removed irrelevant columns, null values, and duplicate entries to ensure data quality.

Your paragraph text



Exploratory Data Analysis (EDA)

Analyzed message lengths, word/sentence counts to understand data distribution.



Text Preprocessing

Cleaned and prepared text for modeling using standard NLP techniques.



Lowercasing & Tokenization

Converted text to lowercase and split it into individual words (tokens).



Model Training

Trained and tested 3 classifiers: GaussianNB, MultinomialNB, and BernoulliNB.

Conclusio: The project successfully builds a robust spam detection system using NLP and Naive Bayes classifiers, achieving high accuracy and precision. The trained model is ready to be deployed on a website interface for real-time email classification.n