Project Report: Email Spam Detection with Machine Learning

Objective

The aim of this project was to build a machine learning model that can detect spam emails (junk mail) based on their content. Spam emails often include cryptic messages, scams, and even phishing attempts. Identifying such emails can significantly improve email security and user experience.

Tools & Technologies Used

- **Language**: Python
- **IDE**: Visual Studio Code (VSCode)
- **Libraries**:
    - `pandas` for data handling
    - `sklearn` for model training and evaluation
    - `TfidfVectorizer` for text vectorization
    - `MultinomialNB` (Naive Bayes) for classification
    - `pickle` for model saving

Dataset

- **Source**: Kaggle - SMS Spam Collection Dataset
- **Structure**:
    - `v1`: Label (ham or spam)
    - `v2`: Message text

Sample Data:

css
CopyEdit

```
label      message

ham        Go until jurong point, crazy.. Available only ...

spam       Free entry in 2 a wkly comp to win FA Cup fina...
```

-

Class Distribution:

yaml
CopyEdit

```
ham: 4825

spam: 747
```

-

Methodology

1. **Data Preprocessing**:
    - Loaded and cleaned the dataset.
    - Converted `ham` and `spam` labels to binary (0 = ham, 1 = spam).
    - Vectorized text using `TfidfVectorizer` with stopwords removed.
2. **Model Training**:
    - Split the dataset into training and test sets (80/20).
    - Trained a `Multinomial Naive Bayes` model on the vectorized data.
3. **Model Evaluation**:
    - Predicted the test set.
    - Evaluated using Accuracy, Precision, Recall, and F1 Score.

Results

- **Accuracy**: `96.86%`

- **Classification Report**:

Label	Precision	Recall	F1-score	Support
Ham	0.96	1.00	0.98	965
Spam	1.00	0.77	0.87	150
Overall Accuracy			**0.97**	1115

-

- The model performs very well overall, especially for detecting **ham** messages. There's slight room for improvement in **spam recall**, which could be addressed with further tuning or data augmentation.

Output Files

- `spam_model.pkl`: Trained Naive Bayes model

- `vectorizer.pkl`: TF-IDF vectorizer used for transforming input text

Conclusion

This project successfully demonstrates the use of **machine learning to classify email messages** as spam or ham based on their content. The model shows high accuracy and is ready for deployment in real-world applications such as email filters.