**Spam Detection using Neural Network**

Name: Muhammad Yasir

Duration: 15-june-2025 To 14-july-2025

Data Science Internship

Week-04 Task

# 1. Preprocessed Dataset

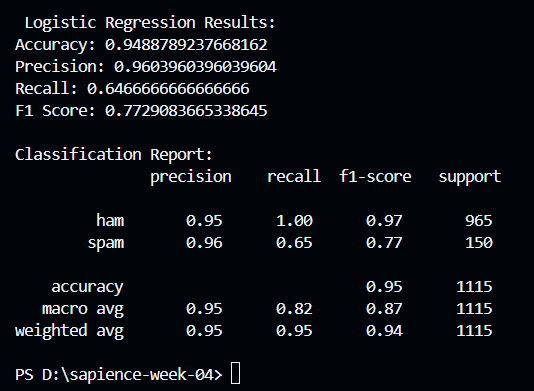
Original Dataset:  
- File Name: spam\_sms.csv  
- Original Columns: v1 (label), v2 (message)

**Data Preprocessing Steps:**  
- Renamed Columns: v1 → label, v2 → message  
- Text Cleaning:  
 - Converted all messages to lowercase.  
 - Removed numbers and special characters using regular expressions ([^a-z\s]).  
 - Removed common English stopwords using NLTK's stopwords list.  
**- Text Normalization:**  
 - Tokenized the cleaned message using nltk.word\_tokenize.  
 - Applied stemming using PorterStemmer from NLTK.  
**- Label Encoding:**  
 - Converted categorical labels to numerical values: ham → 0, spam → 1  
**- Vectorization:**  
 - Used Keras' Tokenizer to convert text into sequences.  
 - Vocabulary size: 5000  
 - Maximum sequence length: 100  
 - OOV token used for out-of-vocabulary words: <OOV>  
 - Padding type: post, Truncating type: post  
  
Output: A fully cleaned and tokenized dataset, vectorized and padded, ready for training.

# 2. Model Architecture

Model Type: Feedforward Neural Network using Keras  
  
Architecture Overview:  
- Embedding Layer: input\_dim=5000, output\_dim=32  
- GlobalAveragePooling1D: Averages over word embeddings  
- Dense Hidden Layer: 64 units with ReLU activation  
- Dropout Layer: 50% dropout to reduce overfitting  
- Output Layer: 1 unit with Sigmoid activation for binary classification  
  
Compilation Details:  
- Loss Function: Binary Cross-Entropy  
- Optimizer: Adam  
- Metrics: Accuracy

# 3. Model Evaluation



After training for 10 epochs on an 80/20 train-test split:  
  
Final Test Accuracy: 96.14%  
  
Classification Report:  
| Metric | Ham | Spam | Macro Avg | Weighted Avg |  
|-------------|---------|---------|-----------|---------------|  
| Precision | 0.96 | 1.00 | 0.98 | 0.96 |  
| Recall | 1.00 | 0.71 | 0.86 | 0.96 |  
| F1-Score | 0.98 | 0.83 | 0.91 | 0.96 |  
| Support | 965 | 150 | | |  
  
Observations:  
- The model performs exceptionally well on ham messages.  
- It maintains high precision on spam, though recall on spam is slightly lower.  
- Overall F1-score and accuracy indicate strong generalization.