**Harnessing LSTM Networks for Advanced Text Classification in Natural Language Processing**

**Abstract:**

Long Short-Term Memory (LSTM) networks for advanced text classification in Natural Language Processing (NLP). The project focuses on classifying text messages into categories such as 'spam' and 'ham'. Key steps include data visualization, text preprocessing, model creation, training, evaluation, and prediction. Data visualization is performed using seaborn and matplotlib to explore the distribution of text message categories. Text preprocessing involves converting text to lowercase, stripping punctuation, and tokenizing words. The LSTM model is built using Keras, with word embeddings generated via the Tokenizer and Doc2Vec models. The model's performance is evaluated through various metrics, including accuracy and confusion matrices. Hyperparameter tuning is applied to optimize the model's results. Additionally, it includes steps for saving the trained model and demonstrating its prediction capabilities on new, unseen text messages. This approach highlights the robustness of LSTM networks in managing sequence data for text classification tasks, providing a solid foundation for further enhancements in NLP applications.

**Keywords:**

NLP, LSTM Networks, Text Classification, Spam Detection, Keras, Word Embeddings, Doc2Vec, Text Preprocessing, Data Visualization, Sequence Data, Hyperparameter Tuning, Accuracy, Confusion Matrix, Prediction