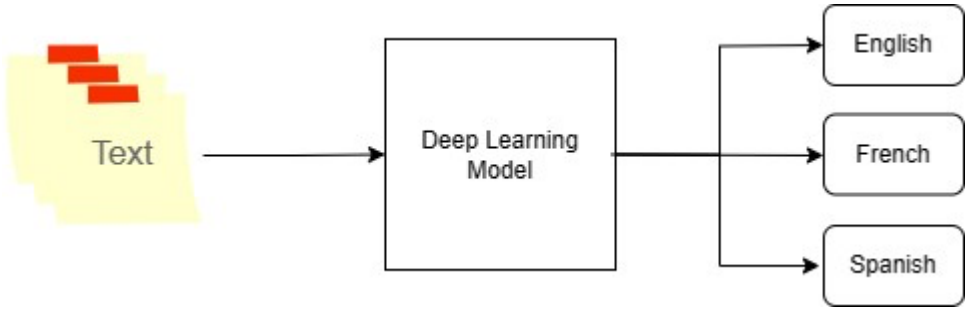
Title

●LingualSense: Deep Learning for Language Detection Across Texts

Problem Description

- **Objective**: To build a model that can automatically identify the language of a given text. Language identification is essential for various applications, including machine translation, multilingual document tracking, and electronic devices (e.g., mobiles, laptops).



Outcomes

- A trained deep learning model capable of predicting the language of input text.
- Improved accuracy in language detection for multilingual documents.

Modules for Implementation

- **Data Preprocessing**:
  - Cleaning and tokenization of text data.
  - Handling unwanted symbols, numbers, and special characters.
- **Feature Extraction**:
  - Representing text data as numerical features (e.g., word embeddings, TF-IDF vectors).
- **Model Architecture**:
  - Using deep learning techniques (e.g., Convolutional Neural Networks, Recurrent Neural Networks) for language identification.
- **Training and Evaluation**:
  - Splitting the dataset into training and validation sets.
  - Training the model on labeled data.
  - Evaluating model performance using accuracy, precision, recall, and F1-score.

Milestones (8 Weeks)

- **Week 1-2: Data Preparation and Exploration**
    - Collect and preprocess the language detection dataset.
    - Explore data statistics and visualize language distributions.
- **Week 3-4: Feature Extraction and Model Selection**
    - Implement feature extraction techniques (e.g., word embeddings).
    - Choose an appropriate deep learning architecture for language identification.
- **Week 5-6: Model Training and Tuning**
    - Train the selected model on the preprocessed data.
    - Optimize hyperparameters (e.g., learning rate, batch size).
- **Week 7: Model Evaluation and Documentation**
    - Evaluate model performance using validation data.
    - Document model architecture, training process, and results.
- **Week 8: Presentation and Final Documentation**
    - Prepare a presentation summarizing the project.
    - Create detailed documentation covering all aspects of the project.

6. Evaluation Criteria

- **Accuracy**: Achieve a minimum accuracy of 90% on the validation set.
- **Documentation Quality**: Well-organized and comprehensive documentation. ● **Presentation**: Clear and concise presentation of the project.