Model Development Phase Template

Date	July 2024
Team ID	740096
Project Title	The Language Of Youtube: A Text Classification Approach To Video Descriptions
Maximum Marks	5 Marks

Model Selection Report

In the model selection report for future deep learning and computer vision projects, various architectures, such as CNNs or RNNs, will be evaluated. Factors such as performance, complexity, and computational requirements will be considered to determine the most suitable model for the task at hand.

Model Selection Report:

Model	Description
Logistic Regression Classifier	The logistic Regression can be used for classifier tasks not for predicting continuous values like carbon monoxide levels. It provides probabilistic predictions, making it easy to understand and implement, while performing well with large datasets and requiring less computational power compared to more complex models.
Random Forest Classifier	The Random Forest Classifier is ideal because it combines the strength of multiple decision trees, offering high accuracy, robust performance against overfitting, and the ability to handle large and complex datasets. The method is inheartly for classification tasks. Which means we need to convert the prediction of carbon monoxide levels into a classification problem.

Decision Tree Classifier	The Decision Tree Classifier is chosen due to its ability to handle non-linear relationships, interpretability in decision-making processes, and robustness in handling diverse types of data.
Naïve Bayes	• **Description**: A probabilistic classifier based on Bayes' theorem with strong (naive) independence assumptions between the features. **Use** • **Case**: Commonly used for text classification tasks like spam detection and sentiment analysis due to its simplicity and effectiveness.