Model Development Phase Template

|  |  |
| --- | --- |
| Date | 03 October 2024 |
| Team ID | LTVIP2024TMID24818 |
| Project Title | Toxic Comments Classification For Social Media |
| Maximum Marks | 6 Marks |

**Model Selection Report**

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Description** | **Hyperparameters** | **Performance Metric (e.g., Accuracy, F1 Score)** |
| Logistic  Regression | Logistic regression is a statistical method used for binary classification that models the relationship between a dependent variable and one or more independent variables using a logistic function. It estimates the probability that a given input belongs to a particular category by transforming linear combinations of inputs through the sigmoid function. | - | Accuracy score = 95% |
| LinearSVC Model | LinearSVC is a linear Support Vector Machine (SVM) model used for classification tasks, particularly effective for large datasets. It finds the optimal hyperplane that maximizes the margin between different classes, focusing on linear decision boundaries. With efficient training on high-dimensional data, LinearSVC is well-suited for text classification and other scenarios where linear separability is possible. | - | Accuracy score = 95% |
| Long Short-Term Memory (LSTM) Model | LSTM is a type of recurrent neural network (RNN) designed to learn long-term dependencies and patterns in sequential data. It uses specialized memory cells and gating mechanisms to manage the flow of information, allowing it to retain context over long sequences. LSTMs are widely used in applications such as natural language processing, time series prediction, and speech recognition. | - | Accuracy score = 97% |