Model Development Phase Template

Date	26 November 2024
Team ID	SWTID1726490119
Project Title	Toxic Comment Classification for Social Media
Maximum Marks	5 Marks

Model Selection Report

For the toxic comment classification project, different deep learning architectures such as ANNs (Artificial Neural Networks), CNNs (Convolutional Neural Networks), LSTMs (Long Short-Term Memory networks), and BiLSTMs (Bidirectional LSTM) were evaluated. The evaluation focused on factors like performance, model complexity, and computational requirements to select the most suitable model for classifying toxic comments.

Model Selection Report:

Model	Description
Simple Neural Network (ANN)	The ANN model is a feedforward neural network used for classification tasks. It utilizes an embedding layer to transform text into dense vectors, followed by a dense hidden layer with ReLU activation. The output layer uses a sigmoid activation function to predict multiple binary labels (toxic, severe_toxic, obscene, threat, insult, identity_hate). The model is trained using binary cross-entropy loss and the Adam optimizer.
Convolutional Neural Network (CNN)	The CNN model is designed to capture local features in the text by applying convolutional layers over the embedded text data. It consists of multiple convolutional layers with varying filter sizes followed by max-pooling. After the convolution layers, the output is flattened and passed through a dense layer with ReLU activation. The model is trained using

	binary cross-entropy loss with the Adam optimizer. Dropout is used for regularization.
Long Short-Term Memory (LSTM)	The LSTM model uses a recurrent neural network architecture that is capable of learning long-term dependencies in sequence data. It incorporates an embedding layer for text representation and an LSTM layer to capture the temporal dependencies in the data. Dropout regularization is applied to avoid overfitting. The output layer uses sigmoid activation for multi-label classification. The model is compiled with binary cross-entropy loss and the Adam optimizer.
Bi-directional LSTM (BiLSTM)	The BiLSTM model extends the LSTM by processing the sequence in both forward and backward directions, allowing the model to capture context from both past and future inputs. Like the LSTM, it uses an embedding layer, followed by a bidirectional LSTM layer. Dropout regularization is applied, and the output layer uses sigmoid activation for multi-label classification. The model is compiled using binary cross-entropy loss and the Adam optimizer.