Assignment No 6

Name : Samiksha More

PRN : 22420127

Roll : 381070

Problem Statement: Sentiment Analysis using LSTM Network or GRU

Objective

The objective is to build a sentiment analysis system that can classify text (such as reviews, tweets, or comments) into positive, negative, or neutral sentiments using Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) networks. The aim is to leverage their ability to capture long-term dependencies in sequential text data for improved accuracy over traditional machine learning models.

Theory

Sentiment Analysis is the process of identifying and categorizing emotions expressed in text. Since text is sequential, simple feedforward networks cannot capture dependencies between words.

* Recurrent Neural Networks (RNNs): Designed for sequence data but suffer from vanishing/exploding gradient problems.
* LSTM (Long Short-Term Memory): Uses input, forget, and output gates to retain long-term dependencies and overcome vanishing gradients.
* GRU (Gated Recurrent Unit): A simplified variant of LSTM with fewer gates, faster to train, and effective for most tasks.
* Word Embeddings: Represent text in vector form (Word2Vec, GloVe, or embeddings layer).
* Output Layer: Typically uses Softmax (for multiclass) or Sigmoid (for binary sentiment classification).

Methodology

1. Dataset Preparation:
   * Collect text data (e.g., IMDB reviews, Twitter dataset).
   * Preprocess text (tokenization, lowercasing, removing stopwords, punctuation).
   * Convert text to numerical form using embeddings.
   * Split into training, validation, and testing sets.
2. Model Design (LSTM/GRU Architecture):
   * Input layer: accepts word embeddings.
   * LSTM/GRU layers: capture sequential dependencies and contextual meaning.
   * Dense layers: map learned features to output space.
   * Output layer: sigmoid for binary classification or softmax for multiclass.
3. Training:
   * Optimize using Adam/SGD.
   * Loss function: Binary Cross-Entropy (binary) or Categorical Cross-Entropy (multiclass).
   * Train over multiple epochs with backpropagation through time.
4. Evaluation:
   * Test on unseen text.
   * Metrics: Accuracy, Precision, Recall, F1-score.
5. Prediction:
   * Input text → Preprocessing → Embedding → LSTM/GRU → Sentiment classification.

Advantages

* Captures long-term word dependencies in text.
* Handles varying input lengths effectively.
* Outperforms traditional machine learning approaches (Naive Bayes, SVM).
* GRU is computationally faster and requires fewer parameters than LSTM.

Limitations

* Requires large labeled datasets for effective training.
* Computationally expensive compared to traditional models.
* Sensitive to noisy text (e.g., sarcasm, spelling errors).
* Training time can be long for large corpora.

Applications

* Social Media Monitoring: Analyzing user sentiments on Twitter, Facebook, Instagram.
* Customer Feedback: Understanding product/service reviews.
* Business Intelligence: Brand reputation management.
* Politics & Public Opinion: Sentiment tracking in elections.
* Healthcare: Analyzing patient feedback on treatments.

Working / Algorithm

1. Input raw text is collected and preprocessed.
2. Words are converted into embeddings (dense numerical vectors).
3. The embedding sequence is fed into LSTM/GRU layers.
4. Hidden states capture context and sentiment cues from sequences.
5. The final hidden state is passed to fully connected layers.
6. Output layer predicts sentiment (positive/negative/neutral).
7. Model is trained using backpropagation through time (BPTT).

Conclusion

LSTM and GRU-based models are highly effective for sentiment analysis because they capture contextual meaning across sequences of words. While LSTMs provide more control with multiple gates, GRUs offer faster training with comparable accuracy. Both are widely applied in real-world applications like product review classification, social media monitoring, and customer sentiment tracking.