Introduction

- **Title**: Qualitative Sentiment Analysis of YouTube Contents based on User Reviews
- **Authors**: Mohammad Rakin Uddin, Nafisa Sadaf Prova, Atik Jawad, Md. Rajib Rayhan, S.M. Rhydh Arnab, and S.M. Hasan Sazzad Iqbal
- **Objective**: Develop a deep learning technique for sentiment analysis on YouTube reviews
- **Model**: Encoder-decoder based Attention model with a squeeze-and-excitation attention layer
- **Dataset**: Nearly 700,000 user comments from 8,000 YouTube channels, categorized into positive, negative, or neutral sentiments

Dataset

- **Size:** 691,400 English comments made by YouTube users
- **Source:** 200 popular YouTube videos
- **Sentiment Categories:** Positive, Negative, Neutral
- **Split:** Training, Testing, and Validation datasets

Preprocessing

- **Steps:** Removal of punctuation and symbols, dropping stop words, tokenizing phrases
- **Tool:** Keras library's "Tokenizer" class

Proposed Model

- **Components:** Embedding layers, Encoder-decoder LSTM model, Attention layer, Fully connected dense layers
- **Process:** Input data is tokenized, passed through an embedding layer, normalized, fed into the encoder-decoder LSTM model, then into the attention layer, flattened, and finally passed through two fully connected dense layers
- **Outcome:** Classification into positive, negative, or neutral sentiments using the softmax activation function

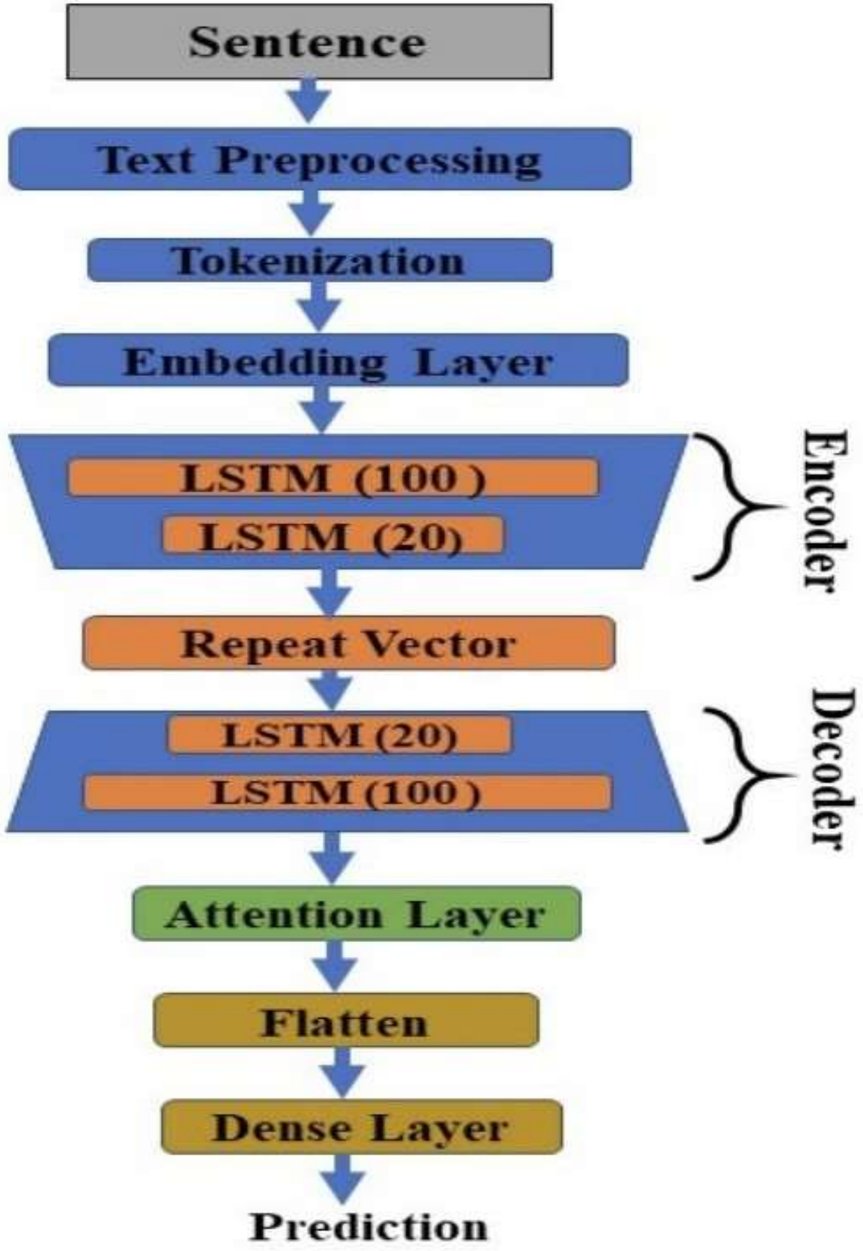

Fig. 1. Schematic diagram of proposed model.

Results

- **Maximum Accuracy: 92.8%**
- **Maximum F1-score: 91.9%**
- **Comparison:** Surpasses several machine learning-based state-of-the-art approaches
- **Impact of Attention Mechanism:** Significant improvement in model accuracy for sentiment analysis