1. **What is polarity detection?**
   It is a binary classification task that represents an important dowel in most sentiment analysis applications.
2. **What are the drawbacks of RNN? how can this be overcomed? What is the advantage of using Attention mechanism?**
   RNN suffers from vanishing and exploding gradients when there are long-term dependencies in the input data. To address this problem, long short-term memory LSTM and gated recurrent unit GRU are introduced. LSTM solves the problem via input, forget, and output gates whereas GRU addresses the problem by a reset gate and an update gate. Attention mechanism is used to focus on the important parts of context by assigning different weights. It is used to improve DNNs by letting them know where to focus for learning. Since it has low training time and uses parallel computation, attention mechanism is used to focus on the important areas of text.
3. **What was the evaluation criteria used for comparing the models?**
   Firstly, the models are compared for sentiment analysis with long reviews. Secondly, a similar comparison is made with short tweets. The proposed model is also compared with a stacking method that aggregates the results obtained by all algorithms. ROC curve is used to compare the models using different thresholds. AUC is compared to evaluate the performance of the model.
4. **What gates are used in LSTM? What are their features?**
   Input gate, forget gate, output gate. The input gate decides what relevant information can be added from the current step, and the output gates finalize the next hidden state.The forget gate decides which information needs attention and which can be ignored. The output gate determines the value of the next hidden state. This state contains information on previous inputs.
5. **Why performance of ABCDM was better on review datasets than twitter dataset?**
   Twitter datasets contain a small number of words and ABCDM does not yield significant improvement using short comments, because the first feature extraction layer in this model is an RNN-based network which is designed to capture long dependencies.