

Abstract

Most empathetic text-based chatbots recognise emotions in conversation at the utterance level and generate response depending on that emotion without analyzing the overall affect. Humans sense affect of another person through the context of the conversation. Similarly, a chatbot can analyze affect from same. However, it is very challenging to extract context from a single utterance. Different state-of-the-art models for conversational A.I. approaches have been proposed that extract contextual information using multiple utterances and different factors. The aim of this study is to understand can and with what relative performance do state-of-the-art conversational A.I. models detect affect across more than one utterance in contrasting conversations. The research was conducted in two parts. Experiment 1 compares and evaluates the performance of Bc-LSTM, DialogueRNN and TL-ERC in which Bc-LSTM outperforms other models. Experiment 2 evaluates which is the better model, Bc-LSTM or Transformer with pre-trained BERT for emotion classification on EmoContext SemEval2019 dataset in which the new Transformer model outperforms Bc-LSTM by a margin of 3% and a total Micro-average F1 score of 75.