

EMOLIZER
A NOVEL SPEECH EMOTION RECOGNITION-BASED
APPROACH TO OPTIMIZE THE CALL
REDISTRIBUTION IN INBOUND CALL CENTERS

Tuwan Shihan Rahim Mahil

A dissertation submitted in partial fulfilment of the requirement for
Bachelor of Science (Honours) degree in Computer Science

Department of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2021

ABSTRACT

Deep learning is a subset of machine learning which facilitates the machines to work independently to simulate some particular tasks without explicit programming or without any human intervention. The researchers are using deep learning algorithms in different domains such as healthcare, robotics, education, and many more because of the tremendous features that deep learning comprise of such as different tuning capabilities, high accurate predictions, etc. Out of all the domains, speech emotion recognition is one of the main areas that the researchers applying deep learning concepts to build more robust applications such as to conduct sentiment analysis, usage of psychological applications, online medical consultation systems, and more.

Moreover, speech emotion recognition is using in call centers to alleviate various problems related to customers as well as call agents. Reducing the customer waiting time of urgent customers is one of the major problems that researchers are trying to resolve. Therefore, they have been devoted several years to find out a suitable solution and finally focused on deep learning despite the effort which was taken based on machine learning and other latest technologies.

When considering the latest efforts which are undertaken to reduce the customer waiting time, those studies reported excellent results but still can be improved. In order to address this issue differently, this research is mainly focused on the hybrid acoustic features in the human voice and a CNN-based architecture. After going through many test scenarios, the work carried out in this study achieved tremendous success for the average number of voice recordings. The system reported 7.232 seconds for 40 voice calls as the overall processing time and this is a good achievement when compared with the existing systems. Furthermore, the implemented EMOLIZER system will be very useful to identify the urgent customers, especially during this COVID-19 pandemic.

Subject descriptors:

Call Centers, Customer Waiting Time Reduction, Hybrid Acoustic Feature

Keywords:

Artificial Intelligence, Deep Learning, Speech Recognition, Speech Emotion Recognition