Voice Emotion Recognition with Audio Data Analysis and Machine Learning Algorithms

Chavgova, Gabriela (School: Model High School of Mathematics "Akademik Kiril Popov")

Emotions are a key aspect of being human. Yet our technological tools largely ignore that. It is reasonable to expect that the next stage of personal assistants can also recognize and react to one's emotional state too. We hypothesize that the human voice carries enough information for accurate voice emotion recognition and present machine learning (ML) algorithms as a solution. Taking into account psychological and machine learning research, we conduct a number of experiments constructing various ML models in order to distinguish 8 emotions (neutral, anger, disgust, fear, calm, joy, sadness, and surprise). We use 4 databases containing audio recordings of these emotions. We perform data analysis and neural network predictions to classify voice recordings. We have achieved 93-94% accuracy of the female model, 66-67% of the male one, and 87% of the gender-neutral algorithm. The male data was not enough for good generalization. A key direction to improve these results would be to find more well-classified data. We discovered that gender-specific algorithms generally achieve higher accuracy and better generalization. We also found that some specific emotions are confused for others. This project would be favorable for both machines and humans. It has the potential to be used in psychology (psychodiagnostics, social psychology) as well as in the advancement of human-machine interactions.

Awards Won:
Second Award of $2,000