

ABSTRACT

Master thesis: 126 p., 19 fig., 3 tabl., 3 applications, 29 sources

The theme: “Deep neural networks usage in speech recognition tasks”

Master thesis is devoted to the research of speech recognition with recurrent neural networks with long short-term memory and connectionist temporal classification. The paper describes the main approaches to the speech recognition with analysis of their strengths and weaknesses.

Connectionist Temporal Classification algorithm was implemented based on algorithm for training recurrent neural network for speech recognition. Software system for analysis, processing and forecasting of voice and sound data, training, testing and analysis of recurrent neural networks was developed in Python programming language using some additional libraries like PyAudio.

Neural network was trained and tested with TIMIT datasets. The results are analyzed and presented in tables, graphs and illustrations.

The novelty of this work is modification of algorithm for acoustic and linguistic models during finding of exponent for sum of to vectors. Instead of counting exponent, we are sending these vectors inside simple neural network with one hidden layer. This was made to make RNNs work faster and have less deletion errors.

Further research is recommended in such areas: using bigger training datasets, pre-processing for voice and sound data, using recognition cache to make recognition process faster, using different configurations for networks to find optimal solution.

SEQUENCE CLASSIFICATION, SEQUENCE LABELING, SPEECH RECOGNITION, RECURRENT NEURAL NETWORKS, LONG SHORT-TERM MEMORY, CONNECTIONIST TEMPORAL CLASSIFICATION