

Deep Learning: Project 2

Project plan

Team members: Dora Doljanin & Filip Pavičić

Project topic

Speech commands classification with recurrent neural networks

Description of the problem

We want to solve the problem of identifying speech commands in voice recordings. We will use the Speech Commands Data Set v0.01, which consists of 1 second labelled clips of voice commands, to classify commands into one of 10 classes (yes, no, up, down, left, right, on, off, stop, go), plus silence or unknown.

Project goals

The purpose of this project is to solve the problem of voice command classification using deep learning and recurrent neural networks. The goal is to test and compare the performances of different network architectures by conducting various experiments.

Expected results

- Project plan
- Report
- Application code
- PowerPoint Presentation

Tasks

1. Explore technical literature, theoretical background, pre-trained models, and related work
2. Familiarize with the dataset
3. Choose the technologies and network architectures to use
4. Split the original training dataset on training and validation datasets
5. Create a simple model which should work significantly better than the random classifier
6. Discover various noise cancelling techniques and their influence on the model
7. Create a model with a separate network for recognising „silence“ and „unknown“ classes

8. Investigate the influence of parameters change on the obtained results
9. Analyse and compare the results
10. Draw a conclusion
11. Present the obtained results in a report and a PowerPoint presentation

Requirements

- At least one of the network architectures should be Long short-term memory (LSTM)

Deadline

The deadline to accomplish all project goals is 10th of May 2022.