ALPEN-ADRIA UNIVERSITÄT KLAGENFURT

Machine Learning for spoken commands classification

Thomas Marquet

February 9, 2021

Summary

Motivation

- First step to make it's own Jarvis at home (or at least a less than intelligent coffee machine)
- It can be used afterwards to be attacked in the context of my PhD Thesis

Learning Task

The Learning task is about the classification of 20 one word spoken commands plus being able to recognize unlabeled words and silence. Therefore the output is on 22 labels. The following words

are the targets :

```
labels = ["yes", "no", "up", "down", "left","right", "on", "off", "stop", "go", "zero"
          , "one", "two", "three", "four","five", "six", "seven", "eight", "nine",'silence','unknown']
```

Those words are used to represent the unknown label :

```
unknown_labels = ["bed", "bird", "cat", "dog", "happy", "house", "marvin", "sheila","tree","wow"]
```

Related Work

Data Analysis

Model : MLP

Model : CNN

Model : Small CNN

Model : LSTM

Model : LSTM + convolutions

Live Demo

Conclusion