

# Detection of dyslexia from child's read speech

Bachelor's thesis  
Author: Richard Martinec  
Supervisor: Ing. Igor Szóke, Ph.D.

## Goals

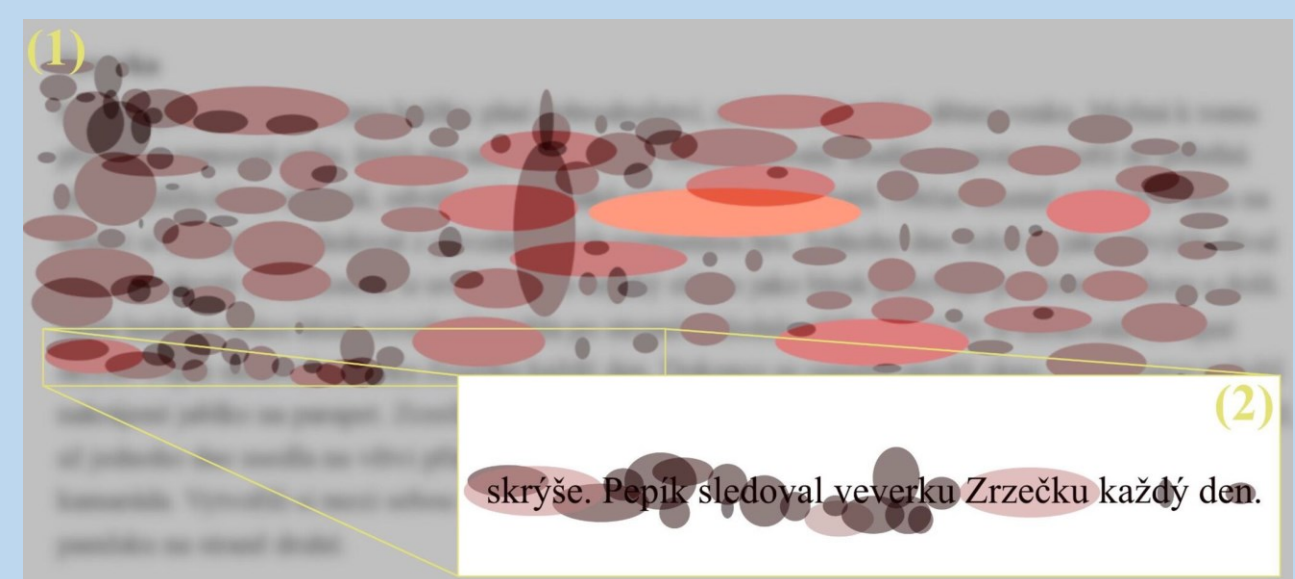
Determine if and **how well can dyslexia be detected** in audio recordings of Czech speech by training multiple classifiers using various approaches.

Create a simple **web application demo** which classifies dyslexia from an audio recording of read speech.



## Dataset

Audio recordings captured in conjunction with **eye-tracking** data by the **Faculty of Arts (MUNI)**.



## Features from embeddings

**HuBERT-Large-FT** – features from hidden states (24+1 layers) used to train a **linear SVM model**. Experimenting with **various feature-preprocessing methods** (e.g. splitting based on word-level alignments).

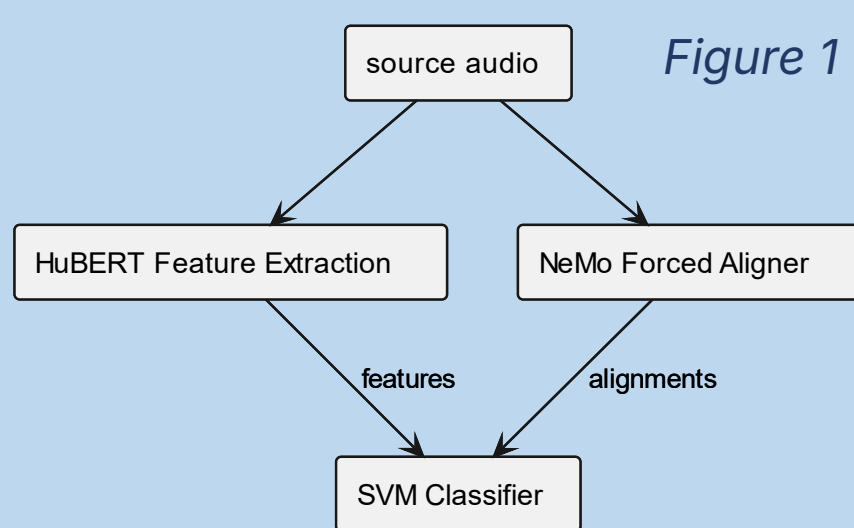


Figure 1

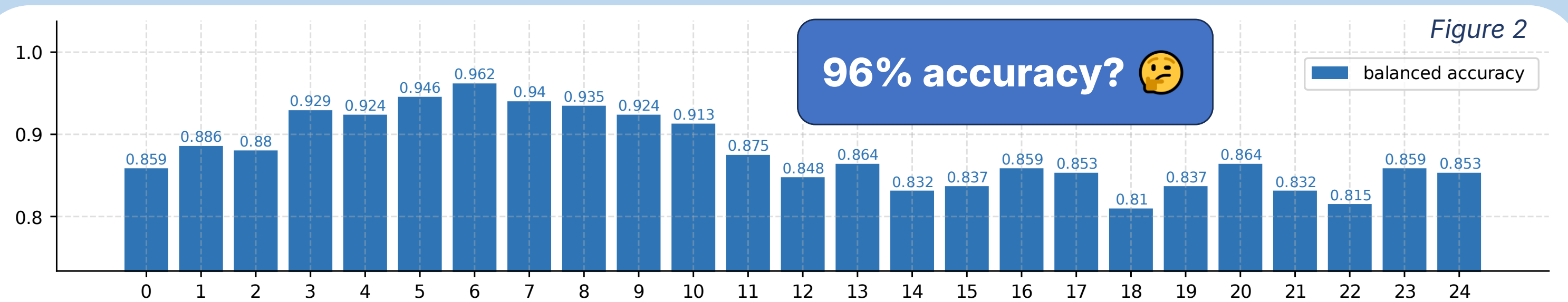


Figure 2

## Features from ASR

**Speech fluency features** from ASR timestamps – partial symptoms of dyslexia:

- Speech pace (letters/s)
- Silence/speech ratio
- Audio signal-to-noise ratio

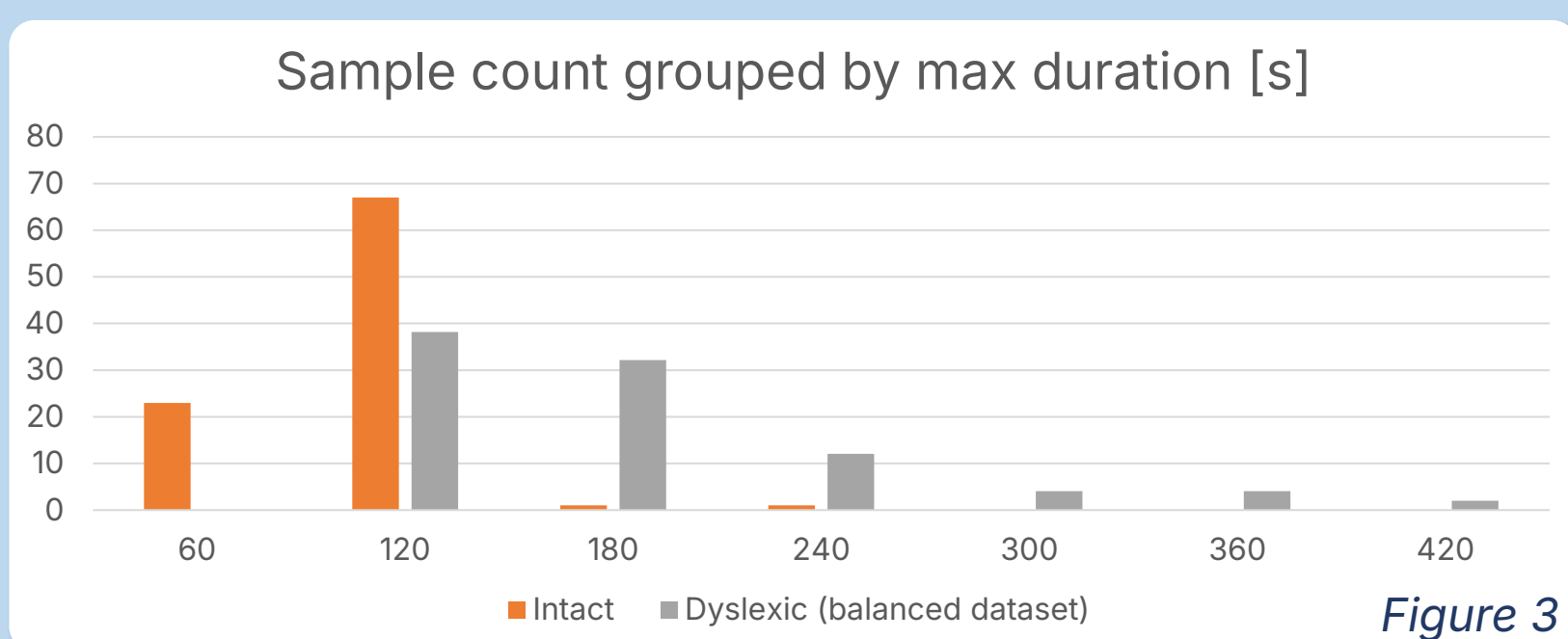


Figure 3

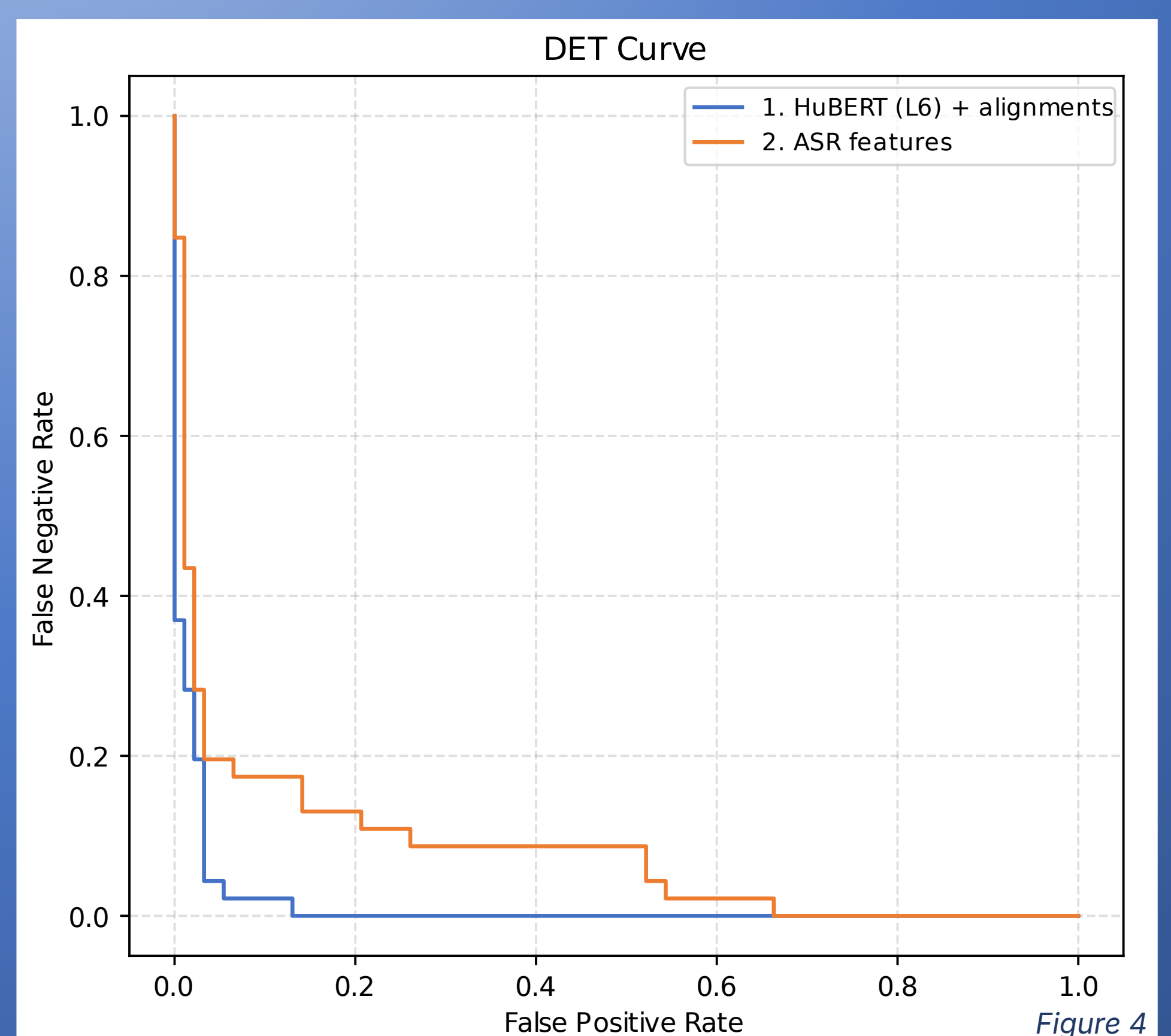
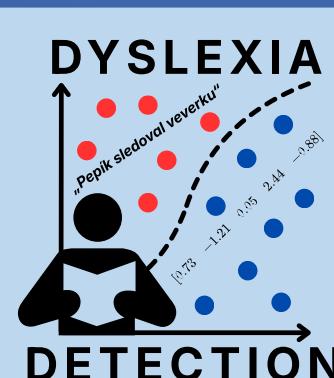


Figure 4

## Results demo



**Web demo** showing a selection of dyslexia classifying SVM models.

