


# Accent spotting





[arinaruck Arina Rak](#)  
[Kirili4ik Kirill Gelvan](#)  
[Captaindoggo Timofey Senchenko](#)

[`github.com/Kirili4ik/accent-spotting`](https://github.com/Kirili4ik/accent-spotting)


# Data and EDA

 Dataset


 172


## Speech Accent Archive


Parallel English speech samples from 177 countries


 Rachael Tatman • updated 3 years ago (Version 2)

[Data](#) [Tasks](#) [Notebooks \(6\)](#) [Discussion \(3\)](#) [Activity](#) [Metadata](#)

[Download \(865 MB\)](#) [New Notebook](#) 

 **Usability** 7.6

 **License** CC BY-NC-SA 4.0

 **Tags** health, linguistics, languages

# Data and EDA

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

~ 13 часов аудио



# Data and EDA

native_language	count
tswana	2.0
garifuna	2.0
ganda	2.0
hmong	2.0
filipino	2.0
telugu	2.0
ewe	2.0
kikongo	2.0
chaldean	2.0
luo	2.0
sicilian	1.0
sardinian	1.0
yupik	1.0
sinhalese	1.0
xasonga	1.0
sundanese	1.0
turkmen	1.0
sa'a	1.0
sarua	1.0
temne	1.0
susu	1.0
vakut	1.0

Особенности  
Языковых  
групп

native_language	count
english	579.0
spanish	162.0
arabic	102.0
mandarin	65.0
french	63.0
korean	52.0
portuguese	48.0
russian	48.0
dutch	47.0
turkish	37.0
german	36.0
polish	34.0
north_germanic	33.0
japanese	27.0
macedonian	26.0
farsi	23.0
cantonese	23.0
vietnamese	22.0
amharic	20.0
romanian	20.0
swedish	20.0

# Overfitting ...

- Label smoothing
- Weighted sampling
- Augmentations

CAP = 3

Resample(old\_sr = 44100, new\_sr = 16000)

Stretch(range=(0.8, 1.2))

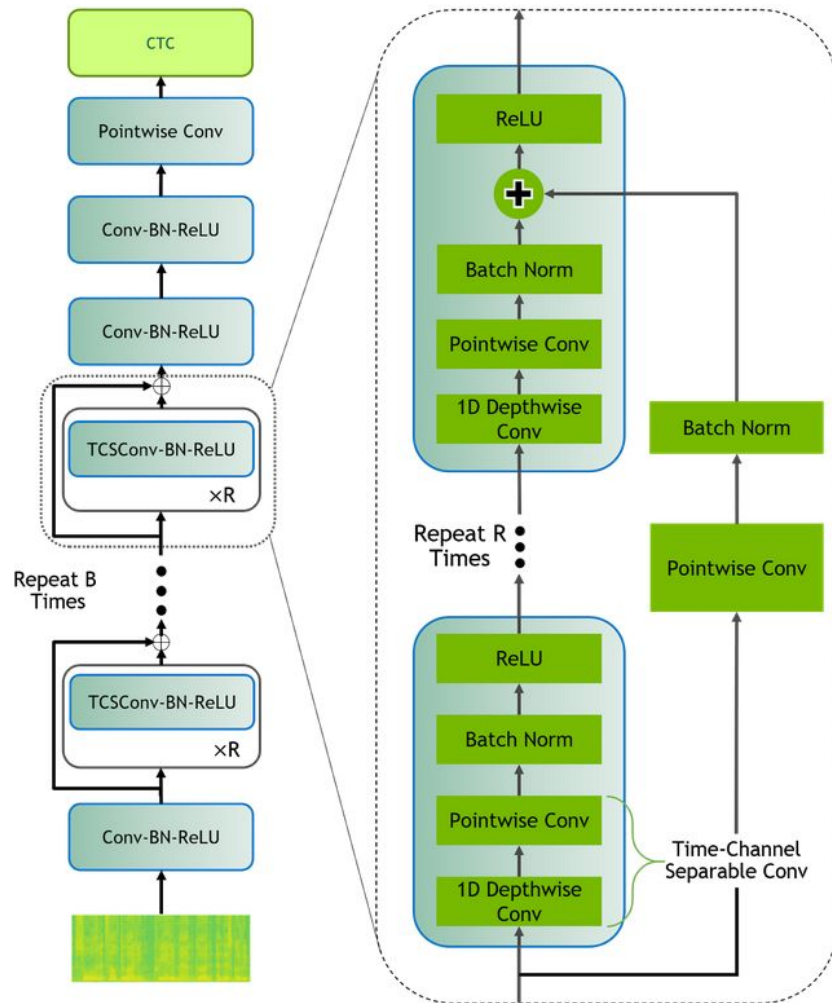
Noise(normal(m=0.0, d=0.0005))

PitchShift(low=-3, high=3)

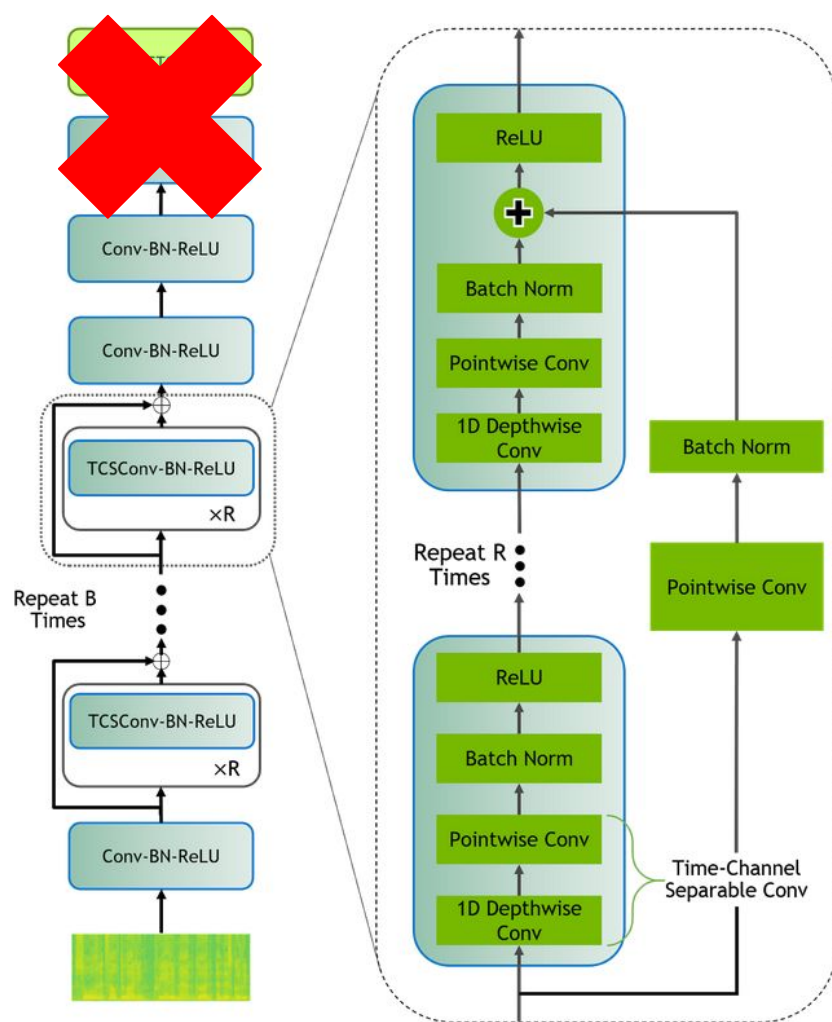
Volume(gain\_range=(0, 10))

Fade(fade\_len = 16000)

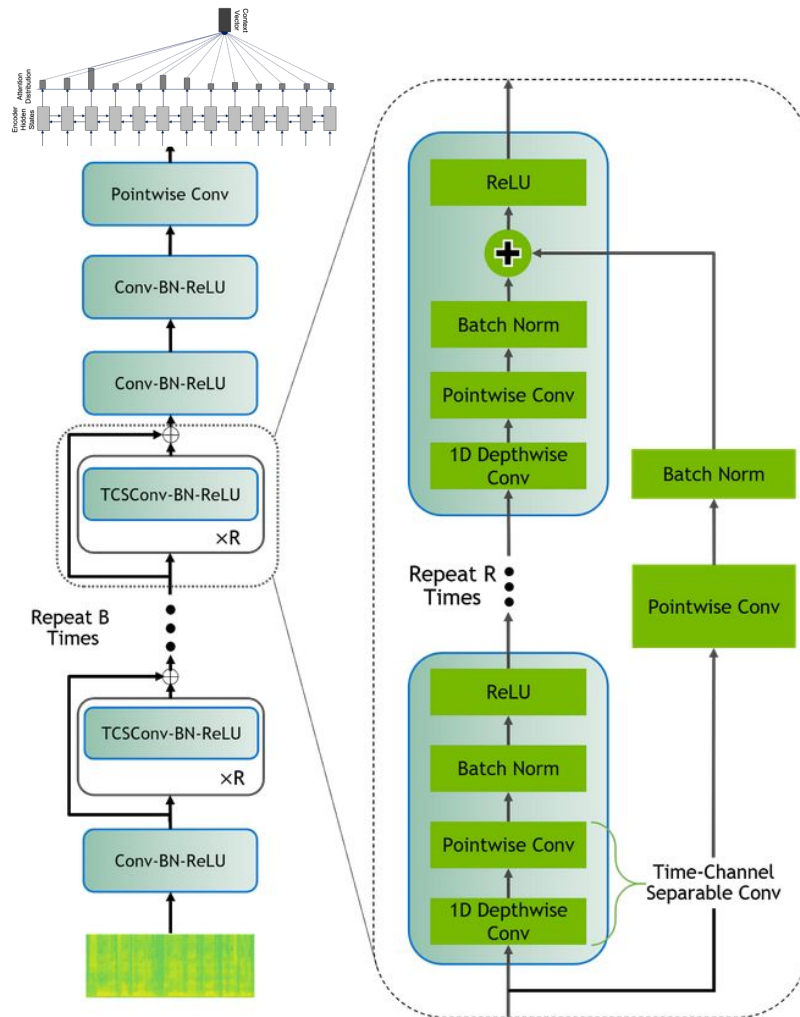
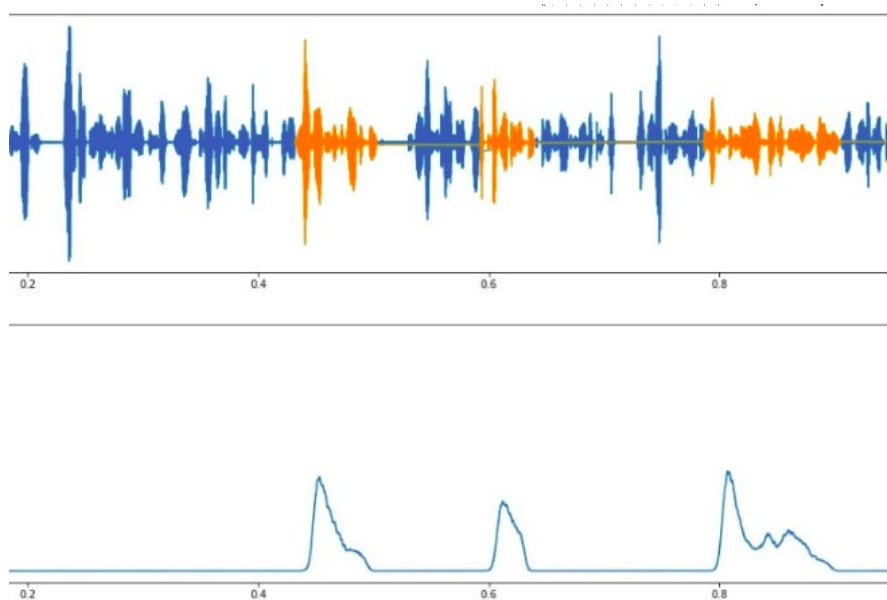
# Architecture



# Sequence classification

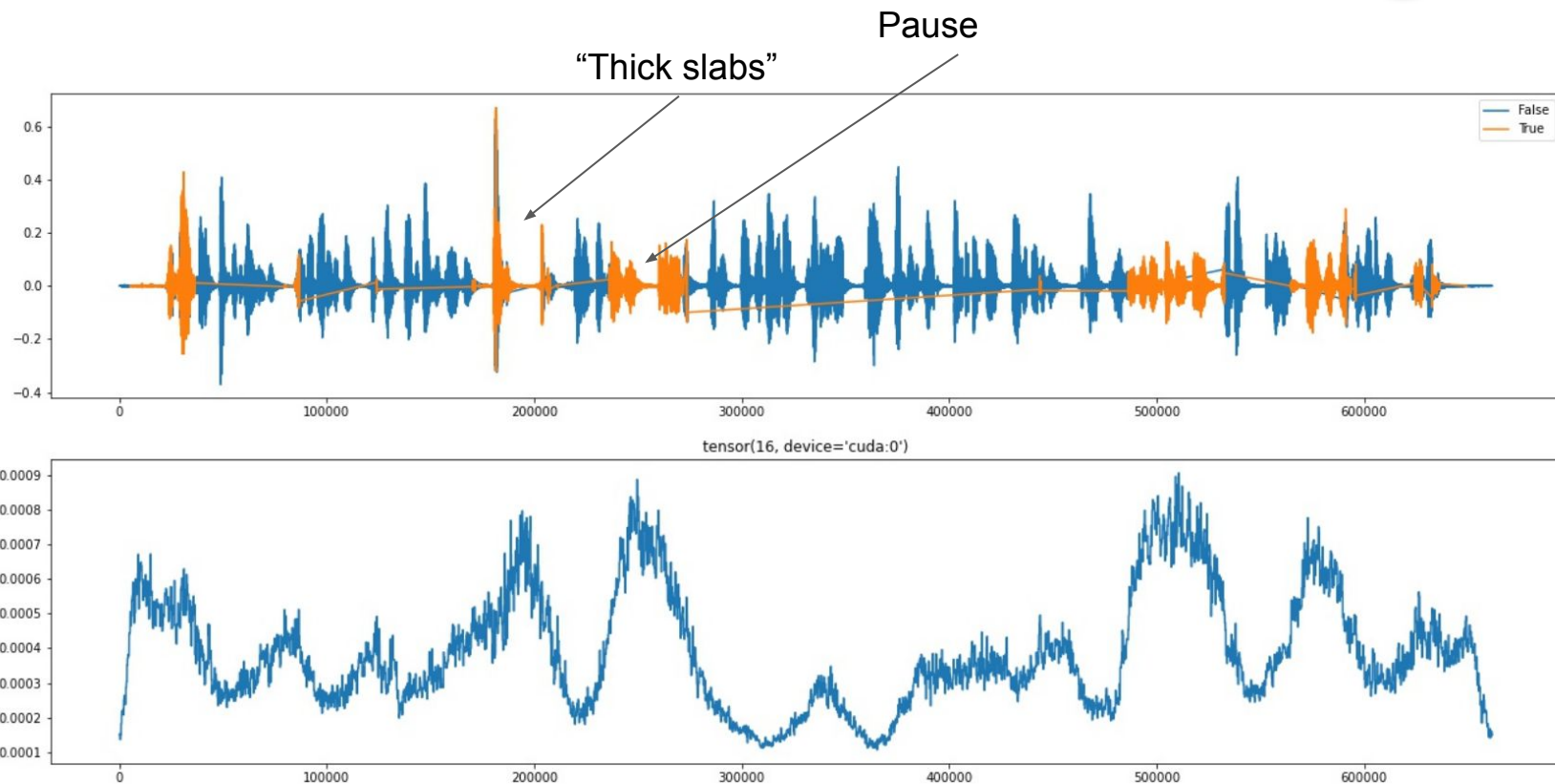


# Attention for accent spotting

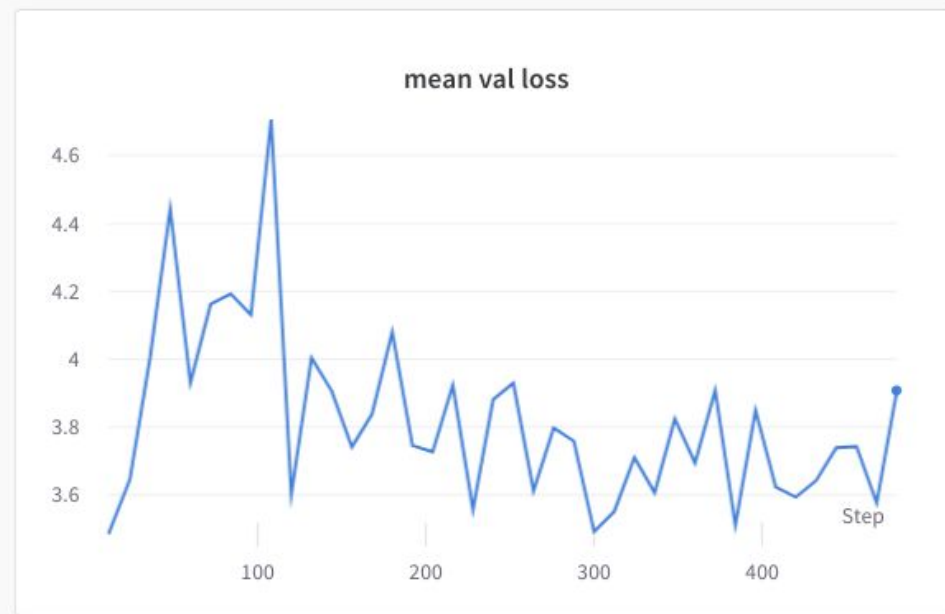
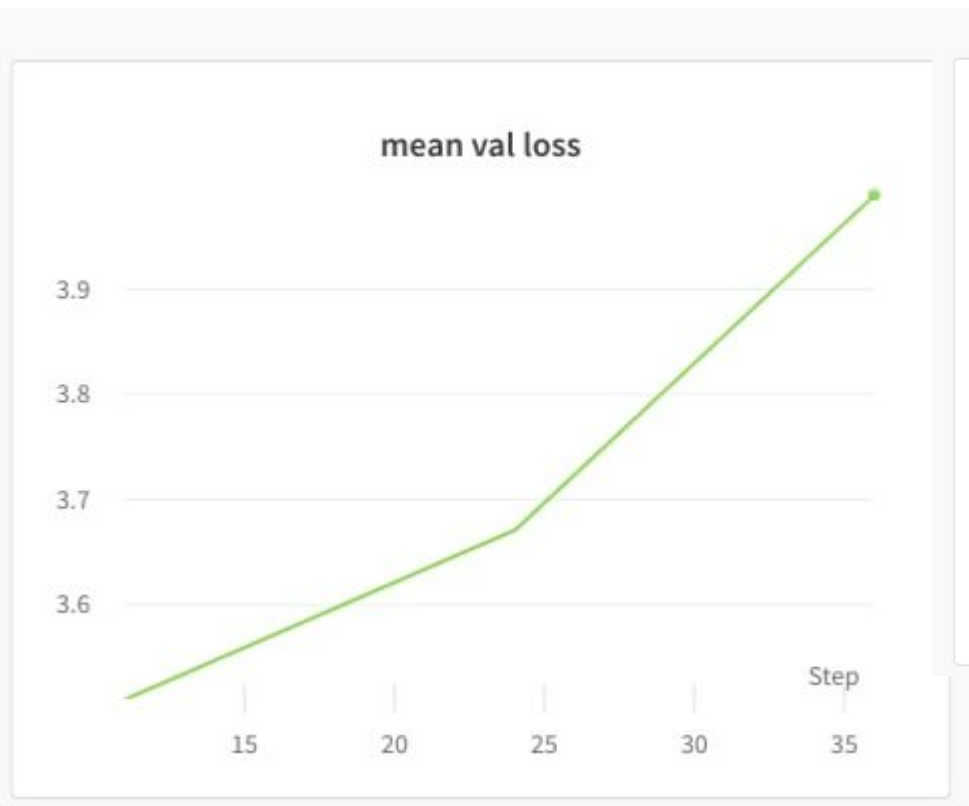




# Pictures of attention



# Learning curves



But binary accuracy ~ 0.89 !

# Future ideas



gettyimages

guvendemir

Thank you for ***attention***