

مقدمه

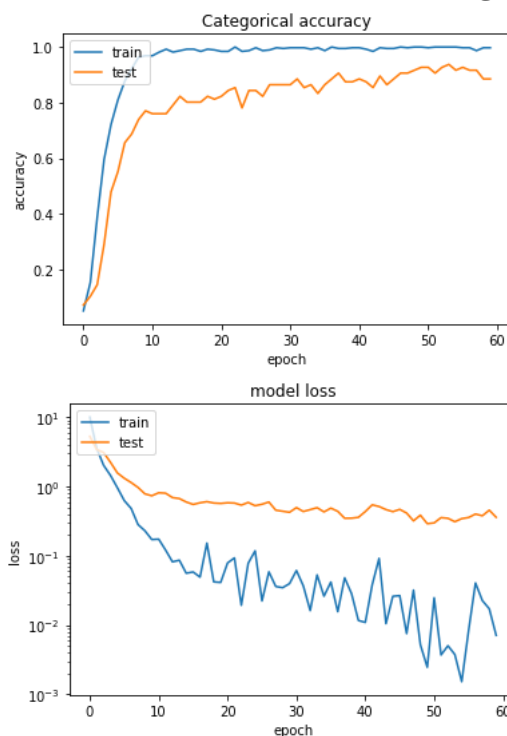
در این پروژه قرار بود که یک سری فرمان های صوتی را به کمک یک شبکه عصبی با حجم اندک طبقه بندی کنیم.

مراحل انجام

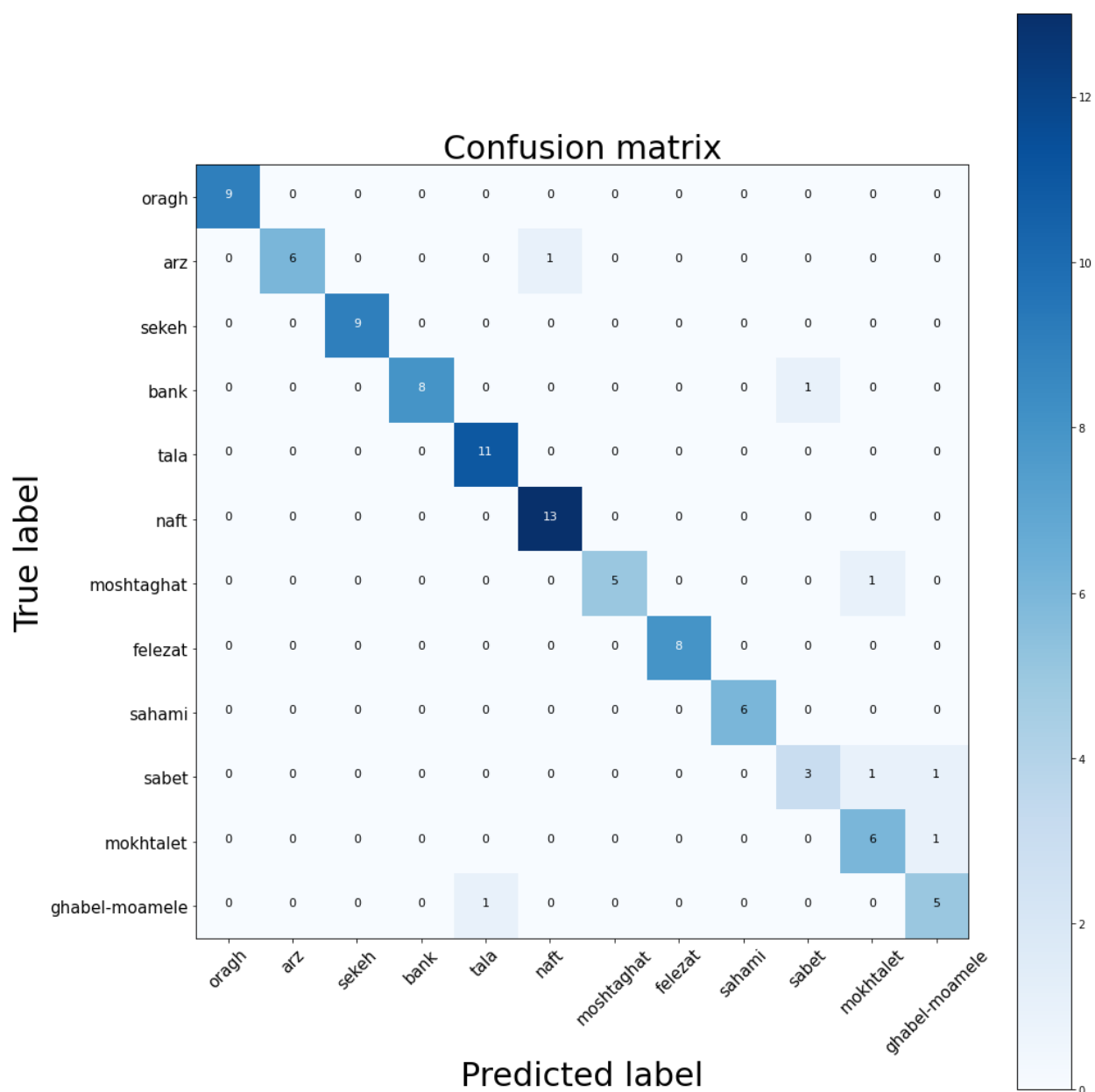
باتوجه به اینکه حجم دیتا خیلی زیاد نبود و میخواستیم که بهترین نتایج را بگیریم، به سراغ مدل های از پیش آموزش دیده رفتیم. در این مسیر با مدل های attention که در این ریپو (<https://github.com/douglas125/SpeechCmdRecognition>) معرفی و آموزش دیده بود برخورد کردم و همین مدل را به عنوان مدل اولیه کار در نظر گرفتم. برای آموزش این مدل ابتدا باید فایل های صوتی را به فرمت npy در می آوردم تا با کد های این ریپو کار انجام بشود به این منظور در فایل audioUtils تابعی برای این کار اضافه کردم و تبدیل را انجام دادم. کد های مربوط به آموزش شبکه را در فایل Speech_Recog_Demo.ipynb تغییر دادم تا مناسب کار خودمان شود. و در انتها آموزش را انجام دادم.

ارزیابی مدل

تغییر درصد های صحت و لاس در طی فرایند آموزش به صورت زیر است:



همچنین ماتریس کانفیوژن به صورت زیر شد :



نحوه اجرای مدل

همانطور که در ریپازیتوری <https://github.com/rezakarbasi/SpeechCmdRecognition> عنوان شده است،

باید با دستور زیر نتیجه را برای یک فایل صوتی به دست آورد :

```
python inference.py 'PATH/T0/AUDIO/FILE'
```