

## تمرین تشخیص کلمات صوتی با استفاده از شبکه عصبی

در زیر لیست ۲۰ کلمه پرکاربرد آمده است:

### لیست کلمات برای تشخیص:

۱. بله
۲. نه
۳. سلام
۴. خداحافظ
۵. لطفا
۶. تشکر
۷. ببخشید
۸. کمک
۹. توقف
۱۰. برو
۱۱. چپ
۱۲. راست
۱۳. بالا
۱۴. پایین
۱۵. شروع
۱۶. پایان
۱۷. باز
۱۸. بسته
۱۹. روشن
۲۰. خاموش

## توضیحات تمرین:

۱. **تعداد تکرارها:** هر کلمه را باید حداقل پنج بار ضبط کنید تا تنوع کافی در داده‌ها داشته باشید.
۲. **شرایط ضبط:** داده‌ها را در شرایط مختلف ضبط کنید تا شبکه عصبی بتواند با نویزها و تفاوت‌های محیطی بهتر کنار بیاید.
۳. **فرمت فایل‌ها:** فایل‌ها را با فرمت WAV و با نرخ نمونه‌برداری ثابت (مثلاً ۱۶ کیلوهرتز) ذخیره کنید. (طبق آموزش سر کلاس)
۴. **برچسب‌گذاری:** هر فایل صوتی باید به درستی برچسب‌گذاری شود تا شبکه عصبی بتواند به درستی آموزش ببیند.
۵. **افزایش داده‌ها: (Data Augmentation)** از هر فایل صوتی ۱۰ نسخه افزایش یافته (augmented) تهیه کنید تا داده‌های بیشتری برای آموزش داشته باشید. (طبق آموزش سر کلاس)

## ساختار پوشه‌ها برای ذخیره فایل‌های صوتی:

یک ساختار پوشه‌ای مناسب برای ذخیره داده‌های صوتی می‌تواند به صورت زیر باشد:

Speech Recognition

data/

balee/

balee\_01.wav

balee\_02.wav

...

nah/

nah\_01.wav

nah\_02.wav

...

salam/

salam\_01.wav

salam\_02.wav

...

...

## نامگذاری فایل‌ها:

برای هر کلمه، فایل‌ها را با فرمت زیر نامگذاری کنید:

کلمه\_شماره.wav

مثال:

balee\_01.wav

balee\_02.wav

...

## تعداد ۱۰ نسخه افزایش یافته از هر فایل صوتی:

برای هر فایل صوتی، ۱۰ نسخه افزایش یافته (augmented) ایجاد کنید و آنها را در همان پوشه ذخیره کنید.

نامگذاری فایل‌های افزایش یافته به صورت زیر باشد:

کلمه\_aug\_01.wav

کلمه\_aug\_02.wav

...

مثال:

balee\_01\_aug\_01.wav

balee\_01\_aug\_02.wav

...

## فرمت فایل‌ها:

همه فایل‌های صوتی باید با فرمت WAV و نرخ نمونه‌برداری ۱۶ کیلوهرتز باشند.

### مثال:

data/

balee/

balee\_01.wav

balee\_01\_aug\_01.wav

balee\_01\_aug\_02.wav

...

balee\_02.wav

balee\_02\_aug\_01.wav

balee\_02\_aug\_02.wav

...

nah/

nah\_01.wav

nah\_01\_aug\_01.wav

nah\_01\_aug\_02.wav

...

nah\_02.wav

nah\_02\_aug\_01.wav

nah\_02\_aug\_02.wav

...

...

این ساختار و توضیحات به شما کمک می‌کند تا داده‌های صوتی را به صورت منظم و قابل دسترس نگه دارید و بتوانید به راحتی آنها را در نرم افزار بارگذاری و پردازش کنید.

در نهایت شما می‌بایست تمام داده‌ها (صوت‌های ضبط شده) + کدهای نوشته شده را در یک ریپازیتوری گیت هاب آپلود و لینک آن را ارسال کنید.

موفق باشید

محمد زارع