



هدف تمرین:

- ایجاد سیستم بازشناسی گوینده
- آشنایی با بازشناسی گوینده مجموعه بسته (closed-set)
- آشنایی با الگوریتم GMM برای مسئله بازشناسی گوینده

مراحل:

۱. در این تمرین شما با یک مجموعه شامل ۲۰ گوینده (۱۰ گوینده مرد و ۱۰ گوینده زن) کار می کنید. هر گوینده ۱۰ فایل صوتی دارد که ۷ فایل آن به عنوان مجموعه یادگیری برای آن گوینده و ۳ فایل به عنوان مجموعه تست برای آن گوینده در نظر گرفته شده است.
۲. ابتدا با استفاده از الگوریتم VAD پیاده سازی شده در تمارین قبلی، بخش های سکوت را از فایل های صوتی از بین ببرید.
۳. ویژگی های MFCC و LPC مربوط به تمام فایل های صوتی مجموعه یادگیری و تست را ایجاد کنید. توجه کنید که فایل های مجموعه داده، 8 kHz هستند. برای به دست آوردن ضرایب MFCC و LPC، از روش هایی که در تمارین قبل پیاده سازی کرده اید استفاده کنید. برای ویژگی های MFCC، در نهایت با اضافه کردن مشتقات اول و دوم، هر بردار MFCC باید ۳۹ بعدی باشد.
۴. برای هر گوینده یک مدل GMM بر مبنای ویژگی های MFCC و یک مدل GMM بر مبنای ویژگی های LPC آموزش دهید. مدل های GMM ذکر شده باید ۳۲ مولفه ای با کواریانس های قطری باشد. برای مدل GMM می توانید از کتابخانه های موجود استفاده کنید. در صورت استفاده از هر کتابخانه ای، باید پارامترهای مربوطه و دلیل مقدار دهی هر پارامتر را شرح دهید.
۵. در نهایت به ازای هر فایل صوتی در مجموعه داده تست، برچسب گوینده آن را با استفاده از دو مدل ایجاد شده، مشخص کنید. نتیجه باید در دو حالت Top-1 و Top-3 به ازای هر مدل مشخص شود.
 - a. Top-1 به معنای آن است که اگر اولین برچسب برای یک فایل (برچسب صوتی متناظر با بالاترین احتمال) همان برچسب واقعی گوینده باشد، یک شناسایی درست حساب می شود.
 - b. Top-3 به معنای آن است که اگر هر کدام از ۳ برچسب اولیه (متناظر با سه مقدار بیشترین احتمال) برچسب واقعی گوینده بود، یک شناسایی درست حساب می شود.

۶. دو مدل ایجاد شده را از نظر دقت با یکدیگر مقایسه کنید.

۷. علاوه بر گزارش نهایی و کدهای پروژه، باید فایل‌های زیر را نیز ارسال کنید:

a. ماتریس ضرایب MFCC نرمال شده و مشتقات اول و دوم در قالب `mat` یا `joblib`. با نام

MFCC برای فایل هر گوینده در دو مجموعه یادگیری و تست.

b. ماتریس ویژگی‌های LPC در قالب `mat` یا `joblib`. با نام LPC برای فایل هر گوینده در دو

مجموعه داده یادگیری و تست.

c. مدل GMM ذخیره شده برای هر گوینده بر مبنای ویژگی‌های MFCC و همینطور LPC.

- نتیجه نهایی را در قالب یک گزارش PDF و کدهای مربوط به تمرین، به ایمیل ui.speech.course.2021@gmail.com ارسال کنید. توجه کنید که در عنوان ایمیل حتما نام، نام خانوادگی و شماره تمرین (تمرین 4) را قرار دهید و در فایل گزارش نیز نام و نام خانوادگی و شماره دانشجوییتان را ذکر کنید. در غیر اینصورت تکلیف شما دریافت نخواهد شد. همچنین در نظر داشته باشید که کامل بودن توضیحات، مرتب بودن گزارش ارسالی و درستی توضیحات همه در امتیاز نهایی تاثیرگذار هستند.

موفق باشید