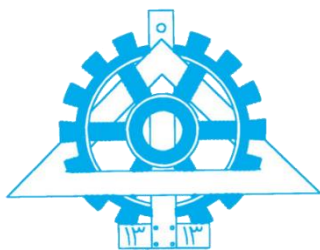


به نام خدا

دانشکده مهندسی برق و کامپیوتر

درس یادگیری عمیق

تمرین سری چهارم



هدف تمرین حاضر این است که شما بتوانید با استفاده از مدل مخفی مارکوف (HMM) و شبکه های عصبی برای کاربرد تبدیل گفتار به متن استفاده کنید.

مجموعه داده های این تمرین مجموعه گفتار ارقام عربی است که شامل ۸۸۰۰ نمونه ارقام ۰ تا ۹ است. هر نمونه بیانگر یک بردار ویژگی شامل ۱۳ ضریب MFCC است.

می توانید این دیتاست را در لینک زیر دریافت نمایید:

<https://archive.ics.uci.edu/ml/datasets/Spoken+Arabic+Digit>

در این تمرین می خواهیم که ابتدا با استفاده از HMM-GMM توسط یک کتابخانه آماده، یک سیستم تشخیص گفتار کلمات (Isolated World Speech Recognition) را پیاده سازی نمایید سپس با دیتاست گفته شده مدل را آموزش دهید. سپس با استفاده از این مدل داده ها را دیکد نمایید، بدین معنی که به ازای هر داده ورودی، خروجی آن برچسب زنی به پنجره های زمانی و تعیین اینکه هر کلمه در چه بازه زمانی قرار گرفته است. سپس با استفاده از خروجی های مدل HMM-GMM یک HMM-DNN آموزش دهید که در آن از یک شبکه Multilayer Perceptron استفاده می کنید. این مدل را پیاده سازی نموده و پس از آموزش دقت آن، زمان همگرایی را گزارش کنید.

از مدل GMM-HMM برای align کردن فریمهای داده ورودی به حالت های مدل مارکوف مخفی استفاده نمایید. در نظر بگیرید به عنوان نمونه داده ورودی شامل ۲۰ بردار ویژگی MFCC و مدل مارکوف مخفی هم ۵ حالت دارد. لذا مرحله alignment موجب می شود که برخی از این فریم ها به حالت اول، برخی به حالت دوم،.. و برخی به حالت آخر اختصاص یابند.

برای شبکه عصبی مدل DNN-HMM، یکبار این شبکه را با یک لایه مخفی و یکبار دیگر با دولایه مخفی پیاده سازی نموده و نتایج را مقایسه کنید.

دقت داشته باشید که برای تابع توزیع حالت مدل کواریانس گوسی بهتر است ماتریس کواریانس قطری در نظر گرفته شود که بحث فرابراز رخ ندهد.

برای پیاده سازی دقت شود که برای جلوگیری از به وجود آمدن مشکل عددی، به جای استفاد از احتمال به صورت معمول، از لگاریتم احتمال استفاده شود، همچنین دقت فرمایید در بروزرسانی ماتریس کواریانس، این ماتریس سینگولار نشود.

در پردازش گفتار از ۳۹ ویژگی MFCC استفاده میشود. برای به دست آوردن سایر ویژگی ها از ۱۳ ویژگی فوق، باید مشتق اول و دوم این ۱۳ ویژگی را به صورت گسسته محاسبه نمایید و در کنار ۱۳ ویژگی ابتدایی قرار دهید. در نتیجه ۳۹ ویژگی حاصل خواهد شد که باید در تمامی تمرین از این ۳۹ ویژگی استفاده نمایید.

می توانید از کتابخانه hmmlearn که برای پایتون پیاده سازی شده است استفاده کنید.

[/https://hmmlearn.readthedocs.io/en/stable](https://hmmlearn.readthedocs.io/en/stable)

#### نکات:

- توجه کنید که نیمی از نمره تمرین مربوط به گزارش می باشد. لازم به ذکر است که رعایت اصول نگارشی حائز اهمیت می باشد.
- در صورتی که امکان اجرای کد بر روی سیستم خود را ندارید می توانید از google colab استفاده نمایید.
- در صورتی وجود هرگونه اشکال در اجرای کد نمره صفر برای این تمرین لحاظ خواهد شد. در صورتی که از notebook jupyter استفاده می کنید دقت کنید که کد شما به صورت cell به cell اجرا شود.
- گزارش تمرین را حتما به صورت PDF ارسال نمایید.
- کدهای تمرین را به همراه گزارش تمرین در سایت درس آپلود نمایید.
- نحوه نامگذاری تمرین براساس studentnumber\_homeworknumber.pdf باشد .
- زبان پیاده سازی این تمرین python می باشد.
- هر گونه پرسش پیرامون تمرین را با ایمیل های [raminnakhli@gmail.com](mailto:raminnakhli@gmail.com) و [esmaeilfarahng@gmail.com](mailto:esmaeilfarahng@gmail.com) مکاتبه فرمایید.