



به نام خدا

تمرین سوم درس یادگیری عمیق

دکتر رشاد حسینی



هدف تمرین حاضر طراحی و پیاده‌سازی مدل مخفی مارکف برای شناسایی حرکت قلم (تصویر) و شناسایی ارقام (صوت) است.

جزئیات مجموعه داده‌های مورد استفاده در این تمرین به شرح ذیل است:

- a. مجموعه داده اول شامل دنباله حرکت قلم برای ۵ حرف صدادار انگلیسی است. این مجموعه داده شامل ۵ فایل متناظر با حروف می‌باشد. هر داده در هریک از فایل‌ها شامل دنباله‌ای از موقعیت‌های قلم (x,y) در صفحه مختصات و در زمان t است. در هر یک از فایل‌های مربوط به حروف، داده‌های موقعیت زوج را به عنوان داده‌های train و داده‌های موقعیت فرد را به عنوان داده‌های test برای آن حرف در نظر بگیرید. دیتاست حروف به پیوست تمرین ارسال شده است.
- b. مجموعه داده دوم یک مجموعه گفتار ارقام عربی است که شامل ۸۸۰۰ نمونه برای ارقام ۰ تا ۹ است. هر نمونه بیانگر یک بردار ویژگی شامل ۱۳ ضریب MFCC است که در پردازش گفتار متداول است. می‌توانید این دیتاست را از لینک زیر دریافت نمایید:

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

- (۱) مراحل ۲ الی ۵ را برای هریک از مجموعه‌های داده به طور مجزا انجام دهید:
- (۲) ابتدا GMM-HMMهای مجزا برای هریک از حروف در مجموعه داده اول (یا رقم در مجموعه داده دوم) به روش Baum-Welch آموزش دهید. ساختار مدل را به دو روش Belkin و ساختار معمول (ماتریس transition میان stateها sparse نباشد)، در نظر بگیرید و مراحل ۳ الی ۵ را برای هریک از ساختارها انجام دهید. لازم به ذکر است در مورد هر ساختار تعداد stateهای مخفی و تعداد componentهای مدل مخلوط بایستی به عنوان hyper-parameter تنظیم شوند. همچنین شرط توقف در مورد EM حین آموزش ذکر شود.

۳) از طریق محاسبه احتمال مشاهدات بر مبنای الگوریتم Forward در هریک از GMM-HMM ها و در نهایت تعیین GMM_HMM متناظر با بیشترین احتمال، ماتریس confusion و مقدار accuracy برای داده‌های train و test را گزارش نمایید.

۴) الگوریتم Viterbi را بر روی ساختار موازی حاصل از GMM-HMM های فوق اجرا و ماتریس confusion و مقدار accuracy برای داده‌های train و test را گزارش نمایید.

۵) نتایج قسمت ۳ و ۴ را تحلیل و مقایسه کنید.

نکات:

- گزارش‌های ارسالی حتما به فرمت PDF و مطابق با نمونه گزارش ارسالی در صفحه درس باشد.
- نحوه نامگذاری فایل تمرین براساس studentnumber_homeworknumber.pdf باشد.
- مهلت ارسال این تمرین ۱۳۹۷/۰۲/۲۱ است.
- لازم به ذکر است در این تمرین زبان پیاده‌سازی برای پیاده‌سازی این تمرین Python است و مجاز به استفاده از کتابخانه‌های مربوط به HMM نمی‌باشید.
- در صورت هرگونه سوال پیرامون تمرین با ایمیل‌های razavi_f@yahoo.com و hadiasheri@gmail.com مکاتبه بفرمایید.

موفق باشید.