

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



درس یادگیری عمیق و کاربردها تمرین شماره ۴

امیر محمد کریمی

۸۱۰۱۹۴۳۸۳

خرداد ماه ۱۳۹۸

فهرست

بخش ۱ ۳

بخش ۲ ۴

پیوست ۶

مراجع ۷

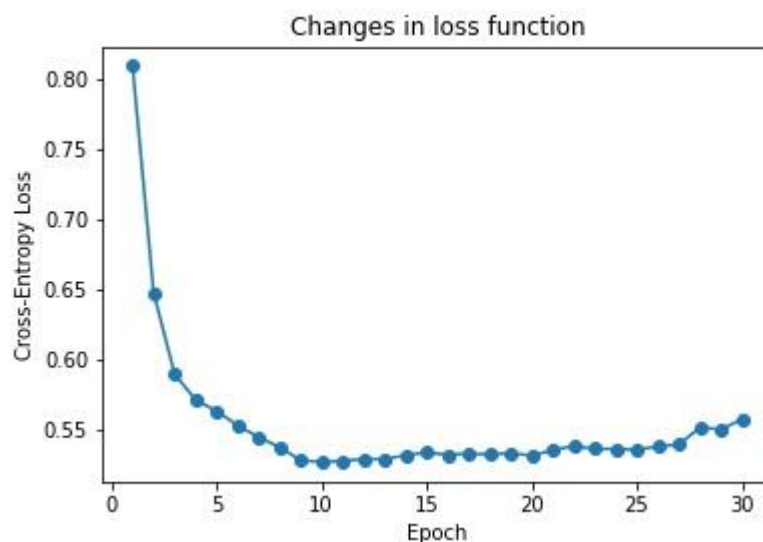
بخش ۱

در این تمرین قصد داریم تا برای تبدیل گفتار به متن روی پایگاه داده‌ی مربوط به صوت ارقام عربی از مدل مخفی مارکوف به همراه یک شبکه‌ی عصبی استفاده کنیم.

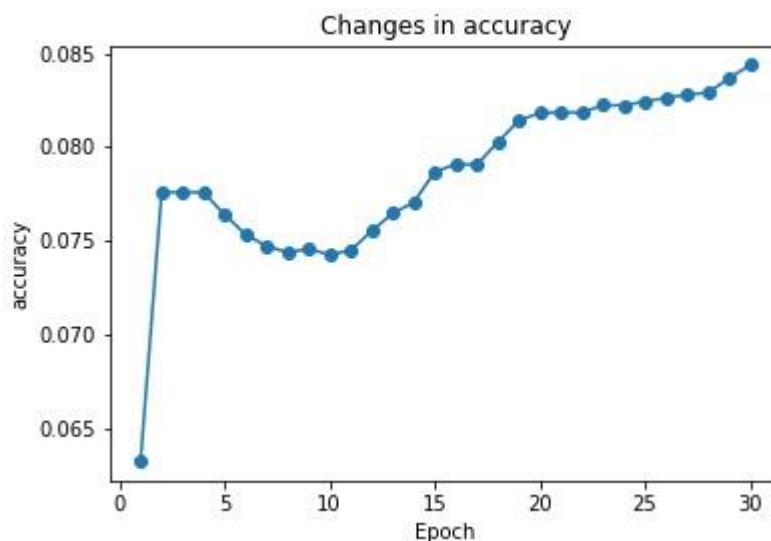
در بخش اول این تمرین، برای هر رقم، یک مدل GMM-HMM داریم. این مدل‌ها را به کمک داده‌های موجود آموزش می‌دهیم و در بخش‌های بعد از احتمالات محاسبه‌شده و خروجی این بخش استفاده خواهیم کرد.

پس از آموزش این ۱۰ شبکه، آن را به کمک داده‌های تست، ارزیابی می‌کنیم و مشاهده می‌کنیم که دقت در حالتی که یک بلاک از داده‌ها را به آن بدهیم، حدود ۴۰ درصد می‌باشد که نزدیک دقت مطلوب می‌باشد.

در این بخش، قصد داریم تا یک DNN با یک لایه مخفی و همچنین دو لایه مخفی را جایگزین GMM در کنیم و مدل را مجدداً آموزش دهیم. برای این کار، شبکه را به این صورت تعریف می‌کنیم که ورودی آن ۹ فریم ۳۹ تایی (یک فریم اصلی و ۴ فریم قبل و بعد از آن) و خروجی شبکه، احتمال هر state شبکه‌ی مارکوف به ورودی می‌باشد (با توجه به این که ۱۰ کلاس داریم و هر HMM دارای ۶ حالت است، بنابراین خروجی شبکه ۶۰ تایی می‌باشد). سپس این خروجی را به کمک قانون بیز به احتمال مشاهده به شرط حالت تبدیل می‌کنیم و به شبکه‌ی HMM می‌دهیم. پس از آموزش شبکه‌ی DNN، نمودار دقت و خطای آن به صورت زیر به دست می‌آید:

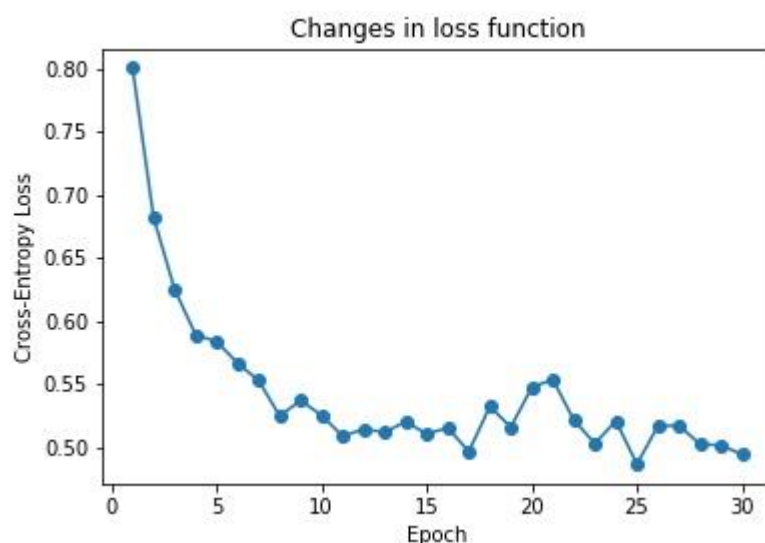


شکل ۱. میزان خطای شبکه با یک لایه مخفی در epoch ۳۰

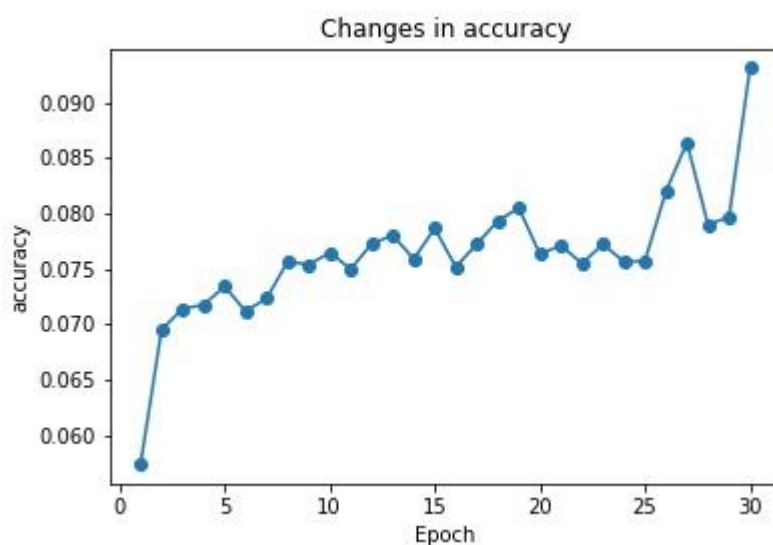


شکل ۲. میزان دقت شبکه با یک لایه مخفی در epoch ۳۰

سلام در حالتی که شبکه با دو لایه‌ی مخفی را برای این قسمت در نظر بگیریم، خروجی دقت و خطای آن به صورت زیر در می‌آید:



شکل ۳. میزان خطای شبکه با دو لایه‌ی مخفی در epoch ۳۰



شکل ۴. میزان دقت شبکه با دو لایه‌ی مخفی در epoch ۳۰

با مقایسه‌ی این شبکه با یک و دو لایه‌ی مخفی مشاهده می‌کنیم که شبکه با دو لایه‌ی مخفی اندکی بهتر عمل می‌کند و سرعت همگرایی آن نیز اندکی بیشتر است.

پیوست 1: روند اجرای برنامه

برای اجرای پروژه لازم است تا در ابتدا فایل‌های مربوط به دیتاست را در پوشه‌ای با نام Arabic-digits قرار دهید و سپس notebook آپلود شده در پوشه‌ی Codes را به صورت cell به cell اجرا کنید.

[1] Pytorch.org. (2019). *torch.nn — PyTorch master documentation*. [online] Available at: <https://pytorch.org/docs/stable/nn.html> [Accessed 20 May 2019].