



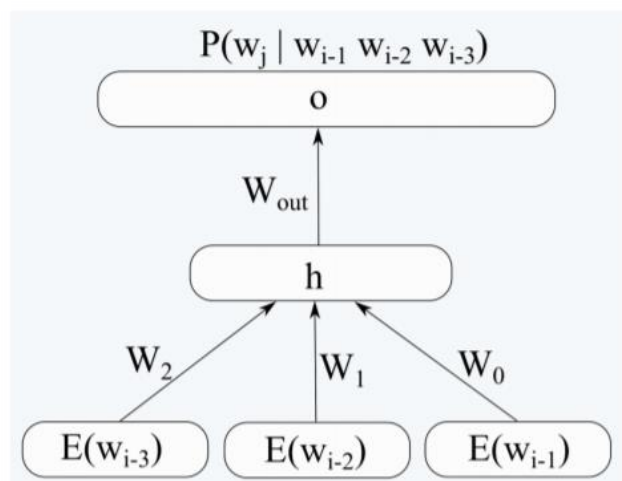
به نام پروردگار
پردازش زبان طبیعی، تمرین عملی شماره ۳
موعد تحویل:



زینب زارعی (zeinabzare73@gmail.com)

مقدمه:

در این تمرین با استفاده از داده‌هایی که در اختیار شما قرار گرفته است، یک مدل زبانی شبکه عصبی با استفاده از شبکه عصبی feed forward آموزش خواهیم داد.



ورودی این شبکه در لایه اول بردارهای embedding کلمات است و در خروجی به هریک از کلمات موجود در مجموعه لغات، احتمال اینکه آن کلمه، کلمه بعدی باشد، اختصاص می‌یابد. به طور کلی لایه‌های مختلف این شبکه به صورت زیر است: (برای ارزیابی خروجی شبکه در حین آموزش از تابع هزینه cross-entropy و برای آموزش آن از gradient descent استفاده شود).

$$e = (Ex_1, Ex_2, \dots, Ex)$$

$$h = \sigma(We + b)$$

$$z = Uh$$

$$y = \text{softmax}(z)$$

فایل فشرده dataset.zip به همراه این تمرین قرار داده شده، که حاوی دو فایل train و test است. برای آموزش شبکه از داده‌های موجود در فایل train استفاده کرده و perplexity فایل test را برای موارد خواسته شده گزارش کنید. (در صورت نیاز پیش پردازش‌های لازم را اعمال کنید).

سوال ۱:

الف) شبکه خواسته شده را پیاده سازی کنید. برای لایه اول می‌توانید از بردارهای از پیش آموزش داده شده ¹ [glove.6B.50d](#) استفاده کنید. تعداد کلمات context ورودی را برابر ۴، تعداد نورون‌های لایه مخفی را برابر ۳۵ و نرخ یادگیری را نیز برابر ۰.۰۲ در نظر بگیرید و آزمایش‌های زیر را انجام دهید.

ب) شرح کلی از نحوه پیاده‌سازی، مقداردهی اولیه به پارامترها، نحوه بروزرسانی پارامترها در هر مرحله و شرط خاتمه یادگیری را گزارش کنید.

ج) نمودار مقادیر perplexity را به ازای iteration های مختلف برای داده‌های train و test به دست آورده و نمودار را تحلیل کنید.

د) به ازای تعداد کلمات context مختلف در ورودی (۲،۳) دو مدل زبانی جدید آموزش داده، perplexity مدل‌های به دست آمده را بر روی مجموعه test گزارش کرده و با رسم نمودار این مدل‌ها را با یکدیگر مقایسه و تحلیل کنید.

ه) به ازای نرخ یادگیری ۰.۰۱، ۰.۰۳ و ۰.۱ مدل را مجدداً آموزش داده و perplexity مدل به دست آمده را در هر iteration بر روی مجموعه test با رسم نمودار تحلیل و مقایسه کنید.

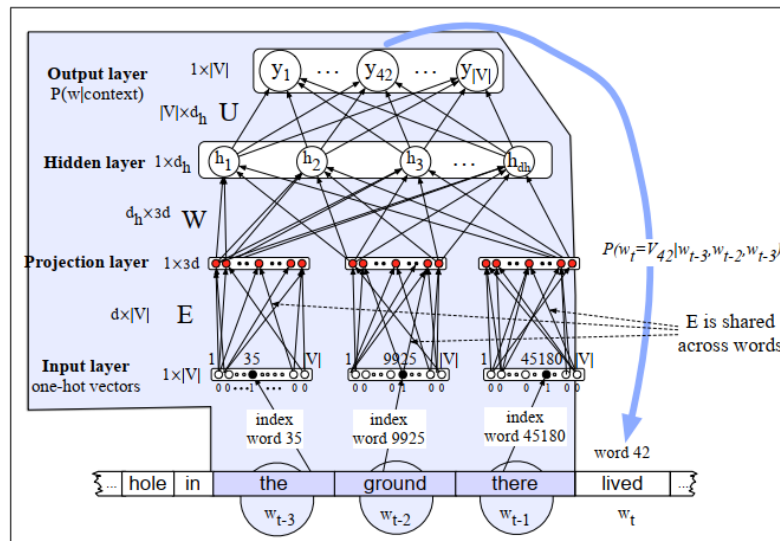
ز) تعداد نورون‌های لایه مخفی را با مقادیر ۵۰، ۱۰۰ و ۱۵۰ تغییر دهید، مدل را مجدداً آموزش داده و perplexity مدل به دست آمده بر روی مجموعه test را با رسم نمودار تحلیل و مقایسه کنید.

سوال ۲:

در سوال قبل از بردارهای از پیش آموزش داده شده برای ورودی به شبکه استفاده شد. در این سوال آموزش بردارهای کلمات داخل خود شبکه انجام می‌شود. ورودی این شبکه بردارهای one-hot کلمات هستند. و یک

¹ pretrained

لایه embedding پس از ورودی اضافه شده که بردار one-hot کلمات در ماتریس E که بین کلمات مشترک است ضرب و embedding های مربوط به آن کلمه را می‌دهد.



الف) شبکه فوق را پیاده سازی کنید. بردارهای embedding کلمات را ۵۰ بعدی، تعداد کلمات context ورودی را ۴، تعداد نوروں‌های لایه مخفی را ۳۵ و نرخ یادگیری را ۰.۲ در نظر بگیرید.

ب) نمودار مقادیر perplexity را به‌ازای iteration های مختلف برای داده‌های train و test به دست آورده، نمودار را تحلیل کرده و با نتایج قسمت ج سوال ۱ مقایسه کنید.

لطفاً به قواعد حل تمرین که در CECM قرار داده شده است توجه کنید.