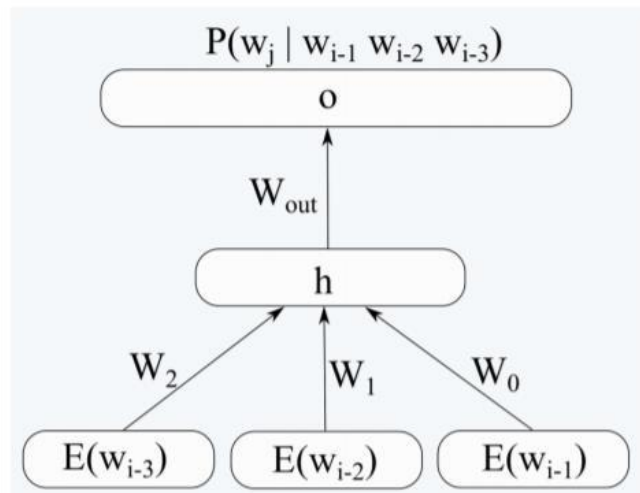


در این تمرین با استفاده از داده هایی که در اختیار شما قرار گرفته است، یک مدل زبانی شبکه عصبی با استفاده از شبکه عصبی feed forward آموزش خواهید داد.



شکل ۱- معماری مربوط به بخش اول

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 می یابد. به طور کلی لایه های مختلف

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

این شبکه به صورت زیر است :

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

توجه: برای ارزیابی خروجی شبکه در حین آموزش از تابع هزینه ی cross-entropy و برای آموزش آن از gradient descent استفاده شود.

$$z = Uh$$

فایل فشرده dataset.zip به همراه این تمرین در اختیار شما قرار می گیرد، که حاوی دو فایل train و test است. برای آموزش شبکه از داده های موجود در فایل

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

train استفاده کرده و perplexity فایل test را برای موارد خواسته

شده گزارش کنید. (در صورت نیاز پیش پردازشهای لازم را اعمال کنید) .

بخش اول

شبکه خواسته شده را پیاده سازی کنید. برای لایه اول می توانید از بردارهای از پیش آموزش داده شده glove.6B.50d استفاده کنید. (تعداد کلمات context ورودی را برابر 4، تعداد نورونهای لایه ی مخفی را برابر 35 و نرخ یادگیری را نیز برابر 0.02 در نظر بگیرید و آزمایشهای زیر را انجام دهید:

(۱) شرح کلی از نحوه پیاده سازی، مقداردهی اولیه به پارامترها، نحوه به روزرسانی پارامترها در هر مرحله و شرط

(۲) خاتمه یادگیری را گزارش کنید.

(۳) نمودار مقادیر perplexity را به ازای iteration های مختلف برای داده های train و test به دست

(۴) آورده و نمودار را تحلیل کنید.

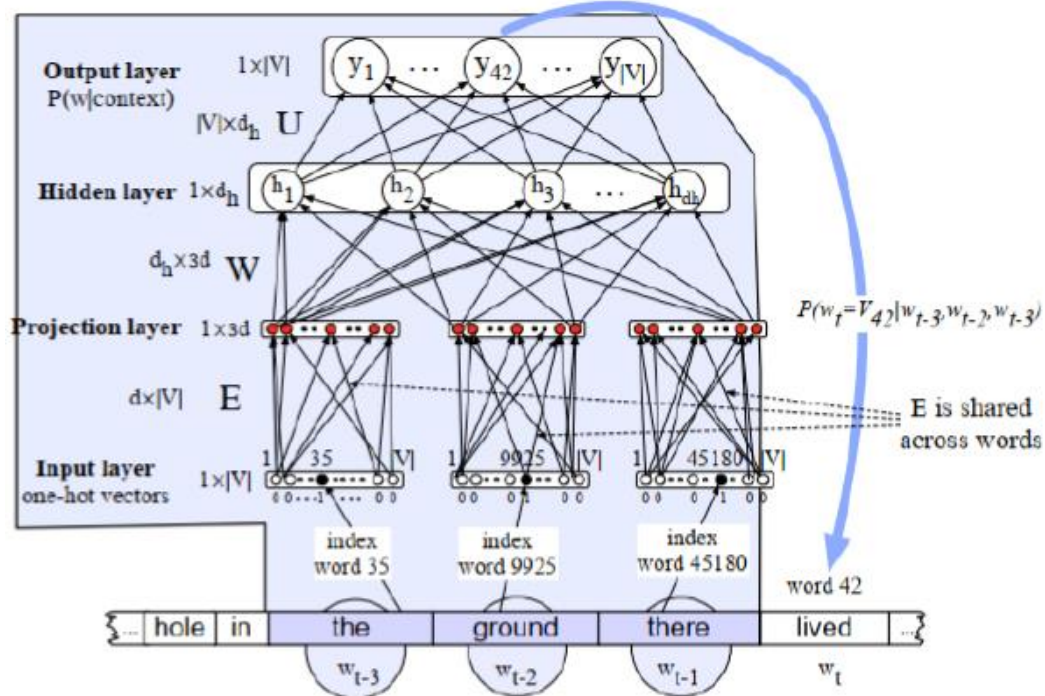
(۵) به ازای تعداد کلمات context مختلف در ورودی (مقادیر ۲ و ۳) دو مدل زبانی جدید آموزش داده perplexity مدل های به

دست آمده را بر روی مجموعه test گزارش کرده و با رسم نمودار این مدل ها را با یکدیگر مقایسه و تحلیل کنید.

- ۶) به ازای نرخ یادگیری 0.01 ، 0.03 و 0.1 مدل را مجددا آموزش داده و perplexity مدل به دست آمده را در هر iteration روی مجموعه test با رسم نمودار تحلیل و مقایسه کنید.
- ۷) تعداد نورونهای لایه مخفی را با مقادیر 50 ، 100 و 150 تغییر دهید، مدل را مجددا آموزش داده و perplexity مدل به دست آمده بر روی مجموعه test را با رسم نمودار تحلیل و مقایسه کنید.

بخش دوم

- 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 می شود. ورودی این شبکه بردارهای one-hot کلمات هستند و یک لایه embedding پس از ورودی اضافه شده که بردار one-hot که در ماتریس E بین کلمات مشترک است ضرب و embedding های مربوط به آن کلمه را می دهد.
- ۱) شبکه فوق را پیاده سازی کنید. بردارهای embedding کلمات را ۵۰ بعدی، تعداد کلمات context ورودی را ۴، تعداد نورونهای لایه مخفی را ۳۵ و نرخ یادگیری را ۰.۲ در نظر بگیرید.
- ۲) نمودار مقادیر perplexity را به ازای iteration های مختلف برای داده های train و test به دست آورده، نمودار را تحلیل کرده و با نتایج قسمت ۲ بخش یک مقایسه کنید.



شکل ۲- معماری مربوط به بخش دوم

بخش سوم: اختیاری

در بخش اول این تمرین در بالا، شبکه feed forward را با یک شبکه RNN جایگزین نموده و آزمایشات را به طور مشابه دوباره انجام دهید.

برای تمام بخشهای فوق گزارش کاملی تهیه نمایید.
می توانید این تمرین را به صورت دو نفره هم انجام دهید.

موفق باشید.