



۱. (۱۰٪ اختیاری-نمره اضافی) [پژوهش: شبکه‌های عصبی بزرگ] هدف این سوال آشنایی با شبکه‌های عصبی آموزش داده بسیار بزرگ مانند GPT-3 است که طی آن این شبکه و شبکه‌های مقیاس بزرگ مشابه را مرور کرده و مواردی مانند روش یادگیری، دادگان و کاربردهای آنها را گزارش کنید.

۲. (۲۰٪ اختیاری-نمره اضافی) [پژوهش: روش‌های دیگر یادگیری] روش‌های یادگیری زیر را مرور کرده و علاوه بر بیان ایده آنها، روش کار و نمونه کاربردهایی از هر کدام را بیان کنید.

الف) Progressive Learning

ب) Meta Learning

ج) Zero Shot Learning

د) One Shot Learning

۳. (۷۰٪ اختیاری-نمره اضافی) [پیاده‌سازی: تحلیل احساس در متن با شبکه LSTM] در این سوال قرار است با استفاده از شبکه‌ی حافظه‌ی کوتاه مدت ماندگار (LSTM) مدلی را آموزش دهیم که بتواند احساس مثبت یا منفی مربوط به هر نظر در داده‌های IMDB را تشخیص دهد. بدین منظور در ابتدا نیاز به آماده‌سازی داده و سپس طراحی مدل و آموزش و ارزیابی می‌باشد.

مجموعه داده نقد فیلم که معمولاً با نام مجموعه داده IMDB شناخته می‌شود، شامل ۲۵۰۰۰ نقد مثبت و منفی برای آموزش شبکه و تعدادی مشابه برای ارزیابی می‌باشد. مسئله طراحی شده تشخیص حس مثبت و منفی در هر نظر نوشته شده است. داده مورد نظر از لینک زیر قابل دسترسی است.

<http://ai.stanford.edu/~amaas/data/sentiment>

قبل از شروع کار بر روی و طراحی مدل نیاز به آماده‌سازی و آنالیز داده می‌باشد. بدین منظور لازم است تمام داده به حروف کوچک تبدیل شوند. علائم نگارشی را حذف کنید. لیستی از نظرات را بسازید. لازم است متن را به واحدهای مجزا تجزیه کنید (واحدسازی) که برای این کار می‌توانید از کاراکتر فاصله و علائم سجاوندی به عنوان مرز بین کلمات استفاده کنید.



لازم است هر واحد (کلمه) را به یک بردار عددی تبدیل کنید. برای این کار از مدل‌ها و بردارهای آماده موجود مانند Word2Vec، FastText، و یا Bert استفاده کنید. در این کار از بردارهای Bert استفاده کنید. مثلا در لینک زیر دریافت کنید:

<https://github.com/google-research/bert>

حال نوبت به آموزش شبکه می‌رسد. برای این کار از شبکه‌ی LSTM استفاده کنید و مدلی بسازید که ورودی آن یک نظر باشد و خروجی آن، بیانگر موافق یا مخالف بودن باشد. برای این منظور لازم است الگوریتم LSTM را به صورت کلی (با امکان تغییر تعداد لایه‌های مخفی، تعداد نرون‌های هر لایه، تنظیم تابع فعال‌ساز، تغییر اندازه دسته (Batch) و...) پیاده‌سازی کنید. برای این تمرین می‌توانید از کتابخانه‌ها و ابزارهای موجود مانند تنسورفلو استفاده کنید.

الف) یک شبکه LSTM یک طرفه با $batch\ size = 50$ ، یک لایه مخفی با ۶۴ واحد حافظه و تابع فعال‌ساز سیگموئید ایجاد کنید و یادگیری آن را با نرخ ۰.۰۱ انجام دهید. شبکه را با بخش Train دادگان آموزش داده و نموداری از روند تغییرات خطای MSE شبکه بر روی این داده‌ها رسم کنید. درستی (Accuracy) مدل آموزش یافته را روی داده آزمون گزارش کنید.

ب) شبکه بخش الف را دوطرفه کنید و نتایج را در این حالت بدست آورد و با حالت یک طرفه مقایسه کنید.

ج) در شبکه قسمت ب، تعداد لایه‌ها را به دو لایه افزایش داده و تعداد واحدهای حافظه هر لایه را با انجام چند آزمایش برای مقادیر مختلف بدست آورید تا بر اساس آن ساختار بهتر شناسایی شود. برای هر کدام از مقادیر بررسی شده، تعداد تکرار آموزش و مقدار درستی را روی دادگان آزمون گزارش کنید.