

1. پژوهش: شبکه های عصبی بزرگ

هدف این سوال آشنایی با شبکه های عصبی آموزش داده بسیار بزرگ مانند GPT-3 است که طی آن این شبکه و شبکه های مقیاس بزرگ مشابه را مرور کرده و مواردی مانند روش یادگیری، دادگان و کاربردهای آنها را گزارش کنید.

2. پژوهش: روش های 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

- Progressive Learning
- Meta Learning
- Zero Shot Learning
- One Shot Learning

3. پیاده سازی: تحلیل احساس در متن با شبکه LSTM

در این سوال قرار است با استفاده از LSTM مدلی را آموزش دهیم که بتواند احساس مثبت یا منفی مربوط به هر نظر در داده های IMDB را تشخیص دهد. بدین منظور در ابتدا نیاز به آماده سازی داده و سپس طراحی مدل و آموزش و ارزیابی میباشد. مجموعه داده نقد فیلم که معمولاً با نام مجموعه داده IMDB شناخته میشود، شامل 25000 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 آماده سازی و آنالیز داده میباشد. بدین منظور الزم است تمام داده به حروف کوچک تبدیل شوند. عالئم نگارشی را حذف کنید. لیستی از نظرات را بسازید. الزم است متن را به واحدهای مجزا تجزیه کنید (واحدسازی) که برای این کار میتوانید از کاراکتر فاصله و عالئم سجاوندی به عنوان مرز بین کلمات استفاده کنید.

الزم است هر واحد -کلمه- را به یک بردار عددی تبدیل کنید. برای این کار از مدلها و بردارهای آماده موجود مانند Vec2Word ، FastText و یا Bert استفاده کنید. در این کار از بردارهای Bert استفاده کنید. مثال در لینک زیر دریافت کنید <https://github.com/google-research/bert>: حال نوبت به آموزش شبکه میرسد. برای این کار از شبکه ی LSTM استفاده کنید و مدلی بسازید که ورودی آن یک نظر باشد و خروجی آن، بیانگر موافق یا مخالف بودن باشد. برای این منظور الزم است الگوریتم LSTM را به صورت کلی (با امکان تغییر تعداد لایه های مخفی، تعداد نرون های هر لایه، تنظیم تابع فعالساز، تغییر اندازه دسته و ...) پیاده سازی کنید. برای این تمرین میتوانید از کتابخانه ها و ابزارهای موجود مانند تنسورفلو استفاده کنید.

الف) یک شبکه LSTM یک طرفه با $\text{batchsize}=50$ ، یک لایه مخفی با 64 واحد حافظه و تابع فعالساز سیگموید ایجاد کنید و یادگیری آنرا با نرخ 0.01 انجام دهید. شبکه را با بخش Train دادگان آموزش داده و نموداری از روند تغییرات خطای MSE شبکه بر روی این داده ها رسم کنید . 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 .