



۱. (۱۰٪) [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٪) [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٪) [پیاده‌سازی: فروشنده دوره‌گرد با الگوریتم تکاملی] در مساله فروشنده دوره‌گرد (Traveling Salesman Problem)، هدف یافتن کوتاه‌ترین مسیر برای یک فروشنده است که می‌خواهد از تعداد N شهر عبور کند به گونه‌ای که از تمامی شهرها دقیقاً یک بار عبور کند و به شهر آغازی بازگردد. تعداد کل راه‌حل‌های این مساله که در علوم کامپیوتر، جزو مسائل NP-Hard است، $(N-1)!$ است. در این تمرین می‌خواهیم از الگوریتم ژنتیک برای حل آن استفاده کنیم.

برای ارزیابی روش خود از داده‌های سه مجموعه ali535، d2103 و baysg29 از TSPLIB که به ترتیب شامل ۵۳۵، ۲۱۰۳ و ۲۹ شهر هستند؛ استفاده کنید. داده‌های فوق به همراه توضیحات لازم در آدرس در دسترس می‌باشد.

<https://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/tsp/>

الف) [پاسخ دقیق] برای bayg29 که دارای ۲۹ شهر است، یک برنامه بنویسید که طول تور را برای کلیه حالت‌های مختلف محاسبه کند و بر اساس آن طول تور بهینه را بدست آورید. توجه کنید که این کار زمان‌بر است و توصیه می‌شود کد خود را روی کامپیوتری غیر از لپ‌تاپ خود اجرا کنید (مثلاً می‌توانید از Google CoLab استفاده کنید). زمان رسیدن به پاسخ و زیرساخت مورد استفاده



برای پردازش را بیان کنید. همین زمان را برای دو مجموعه داده دیگر تناسب ببندید و برآورد کنید تعیین پاسخ دقیق آنها چقدر زمان می برد.

الف) [پاسخ تکاملی] مساله را با الگوریتم تکاملی حل کنید. پاسخ برنامه خود را برای سه شهر فوق شامل نمودار بهترین، بدترین و میانگین پاسخ بر حسب نسل، مقدار پارامترهای مورد استفاده، شرایط توقف و پاسخ نهایی گزارش کنید. در این قسمت از روش انتخاب مسابقه‌ای، روش باز ترکیب چرخشی و روش جهش درجی استفاده کنید. جمعیت اولیه را تصادفی انتخاب کنید، مقدار نرخ جهش را ۰.۲٪ و نرخ باز ترکیب را ۰.۹٪ قرار دهید. بهترین پاسخ گزارش شده در TSPLIB برای سه مجموعه ali535, d2103 و bayg29 به ترتیب ۲۰۲۳۹۳، ۸۰۴۵۰ و ۱۶۱۰ است (لینک زیر).

<https://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/STSP.html>

ب) [تاثیر پارامترها] حالت قبل را برای bayg29 با مقادیر مختلفی برای پارامترهای مساله تکرار کنید. نتایج را گزارش کرده و تحلیل خود را از پاسخها بیان کنید.

ج) [حالت نامتقارن] برای حل مساله در حالت نامتقارن (Asymmetric TSP) داده‌های fvtv130 با ۱۳۰ شهر را در نظر بگیرید. در این حالت فاصله (هزینه) رفتن از شهر i به شهر j با فاصله (هزینه) برعکس آن، رفتن از شهر j به شهر i متفاوت است. تفاوت راه حل پیاده سازی شده توسط شما برای بخش قبلی و این حالت چیست. پاسخ تولید شده توسط برنامه خود را به صورت مشابه بخش الف گزارش کنید.

۴. (۵۰٪ اختیاری-نمره اضافی) [پیاده سازی: 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 آماده سازی داده و سپس طراحی مدل و آموزش و ارزیابی می باشد. مجموعه داده نقد فیلم که معمولاً با نام مجموعه داده IMDB شناخته می شود، شامل ۲۵۰۰۰ نقد مثبت و منفی برای آموزش شبکه و تعدادی مشابه برای ارزیابی می باشد. مسئله طراحی شده تشخیص حس مثبت و منفی در هر نظر نوشته شده است. داده مورد نظر از لینک زیر قابل دسترس است.



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

قبل از شروع کار بر روی و طراحی مدل نیاز به آماده سازی و آنالیز داده می باشد. 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 می توانید 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 می رسد. 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 می توانید 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.