



پروژه دوم

هدف: آشنایی با شبکه های عصبی چندلایه پرسپترون و ابزار های آموزش.

کد: پیاده سازی این پروژه را به زبان پایتون انجام دهید؛ در این فعالیت مجاز به استفاده از tensorflow یا pytorch می باشید.

گزارش: ملاک اصلی انجام فعالیت، گزارش آن است و ارسال کد بدون گزارش فاقد ارزش است. برای این فعالیت یک فایل گزارش در قالب pdf تهیه کنید و در آن برای هر سوال، تصاویر ورودی، تصاویر خروجی و توضیحات کامل و جامعی تهیه کنید.

تذکر 1: مطابق قوانین دانشگاه هر نوع کپی برداری و اشتراک کار دانشجویان غیر مجاز بوده و شدیداً برخورد خواهد شد. استفاده از کدها و توضیحات اینترنت به منظور یادگیری بلامانع است، اما کپی کردن غیرمجاز است.

تذکر 2: مجموعه های داده مورد استفاده را به جز در مواردی که صریحاً در صورت سوال ذکر شده باشد، حتماً قبل از استفاده بصورت تصادفی به سه بخش آموزش (۷۰ درصد داده ها)، آزمون (۲۰ درصد داده ها) و اعتبارسنجی (۱۰ درصد داده ها) تقسیم نمایید.

راهنمایی: در صورت نیاز میتوانید سوالات خود را در خصوص پروژه از تدریسارهای درس، از طریق ایمیل زیر یا در گروه تلگرامی بپرسید.

Email: ann.ceit.aut@gmail.com

ارسال: فایل های کد و گزارش خود را در قالب یک فایل فشرده با فرمت StudentID_HW02.zip تا تاریخ ۱۴۰۱/۰۸/۲۵ ارسال نمایید. شایان ذکر است هر روز تاخیر باعث کسر ۱۰٪ نمره خواهد شد.

در این تمرین با عملکرد شبکه های عصبی چند لایه پرسپترون در دسته بندی آشنا می شویم؛ مجموعه داده پیوستی شامل داده های حاصل از سنسورهای محیطی خانه های هوشمند می باشد. برچسب این مجموعه داده بصورت دودویی بوده و نشان می دهد که آیا آژیر خطر آتش سوزی بر اساس داده های دریافتی از سنسور های تعبیه شده فعال گشته است یا خیر. هدف از کار با این مجموعه داده، آموزش مدلی مبتنی بر شبکه های عصبی چندلایه پرسپترون می باشد تا بتوان توسط آن فعال شدن آژیر مذکور را پیش بینی نمود.

(۱) مجموعه داده را بارگزاری کنید و سپس پیش پردازش های مد نظر زیر را اعمال کنید:

الف) تخمین مقادیر گم شده^۱: روش های مختلفی برای تخمین مقادیر گم شده در مجموعه داده های آموزشی ارائه شده است؛ ضمن مطالعه و گزارش راهکار های موجود، به دلخواه مقادیر گم شده در این مجموعه داده را تکمیل کنید. (در گزارش خود طی جداول و نمودار هایی تعداد مقادیر گم شده از هر ویژگی به همراه مقدار تخمینی را گزارش دهید)

¹ Missing values

ب) نرمال سازی: اهمیت نرمال سازی دادگان آموزشی در فرآیند یادگیری مدل های عمیق چندلایه جلورو را بررسی نموده و مجموعه داده مد نظر را نرمال کنید. به نظر شما در صورت عدم نرمال سازی کدامین ویژگی ها میتوانند مشکل ساز باشند و آموزش را با چالش مواجه کنند؟ چرا؟

ج) ویژگی گسسته موجود در مجموعه داده را به ویژگی عددی تبدیل کرده و روش خود را به همراه دلیل انتخاب آن شرح دهید.

د) حذف داده های پرت: در مجموعه داده های آموزشی، داده های پرت به چه داده هایی اطلاق می شود و چه رویکرد هایی برای شناسایی داده های پرت توصیه می شود؟ یکی از رویکرد های توصیه شده مبتنی بر روش های خوشه بندی بدون نظارت نظیر K-Means می باشد؛ چگونگی کارکرد آن را شرح داده و بروی مجموعه داده اعمال کنید. در گزارش خود قید کنید که با چه جزئیات پیاده سازی، چه تعداد داده آموزشی به عنوان داده پرت شناسایی شده و از مجموعه داده کنار گذاشته شد.

ه) مجموعه داده را به صورت تصادفی به سه مجموعه داده آموزشی، مجموعه داده آزمون و مجموعه داده اعتبار سنجی تفکیک نمایید.

۲) آیا مجموعه داده مفروض جداپذیر خطی می باشد یا خیر؟ چرا؟ برای پاسخ به این پرسش بایستی یک شبکه عمیق جلورو طراحی کرده و طبق نتایج حاصل از آن بتوانید جداپذیر خطی بودن/نبودن را نتیجه بگیرید. معماری شبکه طراحی شده را توضیح داده و بیان کنید چرا این شبکه می تواند مشخص کند که مجموعه داده جدا پذیر خطی می باشد یا خیر؟

۳) در طراحی شبکه های عمیق جلورو، تعیین مقدار بهینه برای تعداد لایه های مخفی و تعداد نورون های هر یک از اهمیت بسزایی برخوردار است بگونه ای عملکرد شبکه نهایی به این دو وابسته است. با سعی و خطا یک شبکه عمیق جلورو با عملکرد مطلوب طراحی کنید. گراف شبکه نهایی را به همراه نمودار خطا و صحت پیش بینی^۲ به ازای پیش روی در تکرار های آموزش در گزارش خود بیان کنید. ضمناً نتایج سعی و خطای خود را در قالب یک جدول با جزئیات کامل گزارش کرده و چگونگی دستیابی به معماری شبکه نهایی را توضیح دهید.

۴) چرا نوع تابع فعالیت نورون های شبکه عمیق جلورو مهم است؟ با تکرار آموزش شبکه حاصل در سوال (۳) به ازای سه نوع تابع فعالیت مختلف نتایج را با آن قسمت مقایسه کنید. آیا نتایج نهایی و روند آموزشی متفاوت است؟ چرا؟

۵) انتخاب مقادیر بزرگ یا کوچک برای اندازه دسته^۳ آموزشی در هر تکرار چه تاثیری در روند آموزش شبکه عمیق دارد؟ با تکرار آموزش شبکه حاصل در سوال (۳) به ازای سه مقدار مختلف برای اندازه دسته آموزشی، تاثیر این پارامتر را بررسی نمایید. علاوه بر ماتریس های درهم ریختگی، میتوانید از تفسیر نمودار های خطا و صحت پیش بینی استفاده نمایید.

^۲ Accuracy

^۳ Batch Size

۵) مقصود از بیش برآزش در آموزش شبکه عمیق جلورو چیست؟ به چه روش هایی و با چه استدلال هایی می توان ادعا نمود که یک شبکه به آن دچار شده است؟ با چه تکنیک هایی می توان بیش برآزش را ایجاد یا از آن اجتناب کرد؟ حال، شبکه ای طراحی کنید که در آن بیش برآزش بروی مجموعه داده ی آموزشی رخ دهد؛ بیش برآزش را در با توضیح نمودار های خطا، صحت پیش بینی و ماتریس های درهم ریختگی (یکی به ازای داده های آزمون و دیگری به ازای داده های آموزش) نشان دهید؛ ضمناً گراف شبکه ی حاصل و روند طراحی آن را در گزارش خود شرح دهید.

۶) منظور از تعمیم پذیری شبکه عمیق جلورو چیست؟ چرا تعمیم پذیر بودن یک شبکه از اهمیت بسازی برخوردار است؟ چگونه میتوان شبکه ای طراحی نمود که تعمیم پذیر باشد؟ در طراحی شبکه عمیق تعمیم پذیر بایستی به چه نکات توجه نمود؟ با توجه به چه معیار هایی می توان تعمیم پذیر بودن مدل را نمایش داد؟ هم اکنون یک شبکه عمیق جلورو طراحی کنید که در آن تعمیم پذیری حاصل شده باشد؛ وجود تعمیم پذیری در شبکه را با توضیح نمودار های خطا، صحت پیش بینی و ماتریس های درهم ریختگی (یکی به ازای داده های آزمون و دیگری به ازای داده های آموزش) نشان دهید؛ ضمناً گراف شبکه ی حاصل و روند دستیابی به آن را در گزارش خود شرح دهید.

توجه: برای آموزش شبکه های عمیق می توانید از منابع و بسترهای سخت افزاری برخط رایگان نظیر Google Colab یا Kaggle استفاده نمایید.

موفق باشید