



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



پروژه نهایی یادگیری ماشین

دکتر ابوالقاسمی-دکتر توسلی پور

بهار ۱۴۰۱

فهرست مطالب

۱	فصل ۱: مقدمه
۲	فصل ۲: تعریف مسئله و شرح پروژه
۲	۱.۲ تعریف کلی مسئله
۳	۲.۲ جمع آوری داده
۴	۱.۲.۲ جملات
۵	۳.۲ گزارش اولیه
۶	۴.۲ گروه بندی
۷	۵.۲ تمیز سازی داده ها و استخراج ویژگی
۸	۶.۲ طبقه بندی
۹	۷.۲ خوشه بندی
۱۰	فصل ۳: بارم بندی
۱۱	فصل ۴: گزارش کار
۱۳	فصل ۵: نکات پایانی

فصل ۱

مقدمه

به طور کلی هدف از این پروژه انجام صفر تا صد یک پروژه واقعی یادگیری ماشین است. در این پروژه شما با چالش هایی روبه رو خواهید شد که بعضاً در دیگر پروژه های دانشگاهی یادگیری ماشین با آن روبه رو نبوده اید. به عنوان مثال در اکثر پروژه ها داده (بعضاً پیش پردازش شده) در اختیار شما قرار می گرفت اما شما در این پروژه بخش جمع آوری داده را خواهید داشت.

امید است در پایان پس از انجام پروژه توانایی شما به عنوان یک مهندس در ارتباط با جمع آوری داده ها، تحلیل و پردازش آن ها و طبقه بندی و خوشه بندی افزایش پیدا کرده و با دید وسیع تری نسبت به حوزه یادگیری ماشین، در این راه قدم بگذارید.

فصل ۲

تعریف مسئله و شرح پروژه

۱.۲ تعریف کلی مسئله

هدف نهایی این پروژه طبقه‌بندی و خوشه‌بندی صدا بر اساس احساسات آن است. ابتدا از شما خواسته شده است در زمان مشخص شده، داده را برای هر گروه از احساسات، به صورت توضیح داده شده، جمع‌آوری کنید. در ادامه لازم است توضیح نسبتاً کوتاهی در ارتباط با چالش‌ها و تفاوت‌های این نوع داده در غالب یک گزارش اولیه گردآوری کنید. پس از تمیزسازی داده آن‌ها را طبقه‌بندی و خوشه‌بندی می‌کنید. در انتها هم لازم است گزارشی کامل از تحلیل نتایج خود بنویسید.

۲.۲ جمع آوری داده

در مرحله اول لازم است هر نفر به صورت **انفرادی** برای هر کدام از ده جمله داده شده با احساسات:

۱- عصبانیت

۲- خوشحالی

۳- ناراحتی

۴- خنثی

از دو نفر خانم و دو نفر آقا داده‌ی صدا جمع آوری کنید. دقت کنید که هر کدام از دسته‌ها (خانم و آقا) باید یک نفر با سن زیر سی سال و دیگری با سن بیشتر از سی باشد. برای یک‌دست بودن داده صوت‌ها را در فرمت wav و با نام‌گذاری به شکل زیر قرار دهید:

studentID_participantID_textID_emotionID.wmv

برای مثال شرکت‌کننده سوم در حال خواندن جمله دهم با احساس ترس است. فایل صدای آن با فرض شماره دانشجویی ۸۱۰۱XXXX با این نام ذخیره می‌شود:

۸۱۰۱XXXX_۳_۱۰_۳.wmv

هم‌چنین در یک فایل CSV جدول برچسب‌ها را با ستون‌های زیر تهیه کنید و در کنار فایل‌های صوتی قرار دهید:

emotionID textID، sex، age، participantID، name، voice

در نهایت همه فایل‌ها را در یک پوشه در گوگل درایو آپلود کرده و لینک آن را در محل تحویل قرار دهید.

در صورتی که برای بیش از چهار نفر شرکت‌کننده داده ارسال کنید، نمره‌ی اضافه به شما تعلق می‌گیرد.

دقت داشته باشید که اگر در حین ضبط صوت محیط را ساکت نگه دارید در مرحله‌ی بعد تمیزسازی داده برای شما راحت‌تر خواهد بود که این یکی از چالش‌های انجام پروژه است: هزینه‌ی جمع‌آوری داده در مقایسه با هزینه‌ی پیش‌پردازش مورد نیاز. هم‌چنین طول صدا بیشتر از **۳ ثانیه** نشود. مهلت ارسال داده حداکثر تا پایان روز **۱۹ فروردین** است.

۱.۲.۲ جملات

برای این بخش ده جمله که معمولا با هر کدام از احساسات گفته شده می‌توان بیان کرد، ذکر شده‌اند:

۱- سلام بر شما

۲- دستتون درد نکنه

۳- زحمت کشیدید

۴- خیلی خوشحال شدم

۵- کاملاً مشخصه

۶- حق با شماست

۷- حال شما چگونه

۸- به لطف شما

۹- قطعاً همین‌طور

۱۰- واقعاً عالی شد

۳.۲ گزارش اولیه

در این مرحله لازم است به صورت گروهی گزارشی تهیه کنید (حداقل دو صفحه) و در آن راجع به تشخیص احساس در صدا اطلاعات جمع‌آوری کنید. در این گزارش شما باید نکاتی را در ارتباط با تاثیر تفاوت‌های فردی مانند سن و جنسیت در تشخیص این احساسات به همراه چالش‌ها و سختی‌های مسئله و جمع‌آوری داده ذکر کنید. مهم این است با جست‌وجو قبل از انجام پروژه نسبت به داده‌ای که قرار است با آن‌ها کار کنید، دید پیدا کنید. لزومی ندارد این بررسی‌ها فنی و از دید یادگیری ماشین انجام شود بلکه باید روان بوده و یک کاربر عادی هم بتواند با خواندن متن شما این تفاوت‌ها را درک کرده و با مسئله آشنا شود. مهلت ارسال این گزارش حداکثر تا پایان روز ۲۶ فروردین است.

۴.۲ گروه بندی

شما می‌توانید به صورت انفرادی و یا گروه‌های حداکثر تا ۴ عضو فعالیت داشته باشید. فعالیت به صورت گروهی نمره مثبت دارد. اسامی افراد گروه را یک نفر به نمایندگی حداکثر تا پایان روز ۲۳ اسفند در محل تحویل گروه‌بندی ارسال کند. دقت کنید که با وجود گروه‌بندی بخش جمع‌آوری داده که در ادامه توضیح داده خواهد شد به صورت انفرادی باید انجام شود.

۵.۲ تمیز سازی داده ها و استخراج ویژگی

حال پس از جمع آوری داده لازم است با پردازش اولیه داده آن‌ها را آماده برای مراحل بعدی پروژه کنید. در این مرحله دقت داشته باشید با توجه به نوع داده‌ی خود بهترین روش‌ها را انتخاب کرده تا بتوانید دقت بالاتری در مراحل بعدی بدست آورید. توجه کنید که بسته به ساکت بودن و یا نبودن محیط ممکن است مجبور به استفاده از الگوریتم‌های **کاهش نویز** از صوت شوید. به منظور استفاده از داده خام مراحل قبل، نیاز به استخراج ویژگی از فایل‌های صوتی و تبدیل آن‌ها به **بردار ورودی** مناسب برای مدل‌ها خواهید داشت. برای استخراج ویژگی‌های صوت از **دوروش** مختلف استفاده کنید و برای مقایسه این دوروش در بخش‌های بعد نمودار **ROC** برای آن‌ها رسم نمایید.

۲.

۶.۲ طبقه بندی

مراحل طبقه بندی در زیر آمده است:

۱- داده را به دو دسته آموزش و تست تقسیم کنید. داده تست باید حداقل **۳۰ درصد** کل داده باشد. برای جداسازی داده تست دقت کنید که مواردی همچون وجود حداقل یک نفر که هیچ صدایی از آن در داده آموزش نبوده است، رعایت کنید.

۲- در این مرحله باید با استفاده از روش های یادگیری ماشین که در طول ترم آموخته اید **طبقه بندی را برای احساسات و جنسیت** انجام دهید. در این مرحله حداقل از **۳ روش جداگانه** استفاده کرده و نتایج را باهم مقایسه و تحلیل کنید. دقت داشته باشید چیزی که در این جا اهمیت بسیار بالایی دارد صرفا کد نیست بلکه تحلیل نمودار ها داده ها و مقایسه ی روش های مختلف است.

توجه کنید در این قسمت برای استفاده از شبکه های عصبی فقط قادر به استفاده از شبکه ی **MLP**^۱ هستید و از RNN^۲ و CNN^۳ نمی توانید استفاده کنید.

۳- تحلیل نتایج و رسم نمودار خطا و دقت طبقه بندی.

^۱ Multi Layer Perceptron
^۲ Recurrent Neural Network
^۳ Convolutional Neural Network

۷.۲ خوشه بندی

در اینجا هم همانند قسمت قبلی باید با استفاده از روش هایی که در درس آموخته اید با انتخاب حداقل دو روش خوشه بندی داده را دسته بندی کنید. این کار را برای تعداد خوشه های ۲، ۴، ۱۰ و تعداد کل ها participant انجام داده و تحلیل کنید. هم چنین شباهت داده ی درون یک خوشه و تفاوت بین خوشه و دلایلی که برخی داده در یک خوشه قرار گرفته اند باید بررسی دقیق شوند. تحلیل و گزارش در این بخش از اهمیت بالایی برخوردار است.

فصل ۳

بارم بندی

از ۱۰۰	نمره دهی
۵	گروه بندی
۱۵	جمع آوری داده ها
۲۰	پیش پردازش داده ها
۲۰	خوشه بندی
۲۰	طبقه بندی
۲۰	گزارش کار نهایی

فصل ۴

گزارش کار

همان‌طور که قبلاً هم گفته شد، علاوه بر کد درست، گزارش کار مفصل و توضیح و تحلیل درست داده و نمودارها از اهمیت بسیار بالایی برخوردار است. سعی کنید تمام نکات قابل ذکر در انجام پروژه را در گزارش کار ذکر کنید. دقت داشته باشید که گزارش کار اولیه حداقل دو صفحه‌ای که برای توضیحات تفاوت‌ها و چالش‌ها به همراه اسامی اعضای گروه که مهلت آن تا پایان ۲۶ فروردین است، به عنوان مقدمه گزارش کار نهایی است. گزارش نهایی هم همراه با کد و تحلیل موارد خواسته‌شده تا انتهای روز ۲۷ خرداد فرصت دارد. در زیر نکاتی را متذکر می‌شویم که حتماً باید در گزارش کار نهایی ذکر شود:

- روش‌هایی که برای قسمت پیش پردازش استفاده کردید و توضیح مختصر نحوه کارکرد هر کدام از آن‌ها
- تحلیل‌هایی که در بخش‌های خوشه‌بندی با تعداد خوشه‌هایی متفاوت کردید
- مدل‌هایی که برای طبقه‌بندی استفاده کردید را مختصراً توضیح دهید و درباره چرایی انتخاب آن‌ها نیز توضیح دهید
- استفاده از روش‌های دیگر تدریس شده در درس مانند روش‌های کاهش بعد و ... در جهت بهبود کارایی مدل بلامانع است ولی حتماً باید در گزارش کار ذکر شود.
- برای هر کدام از مدل‌هایی که آموزش داده اید، `precision, recall, f1-score, error` را در گزارش قرار دهید

گزارش کار می‌بایست مرتب و بخش‌بندی شده باشد و تحلیل‌های دقیق داشته باشد. همچنین نوشتن گزارش کار با \LaTeX تا ده درصد نمره امتیازی دارد.

فصل ۵

نکات پایانی

- طبق تقویم آموزشی موعد تحویل نهایی تا انتهای هفته ۱۶ ام درس خواهد بود. با توجه به فشردگی در هفته های آخر با برنامه ریزی درست سعی به انجام پروژه به بهترین نحو ممکن بکنید.
- هیچگونه شباهتی در انجام این پروژه بین افراد مختلف پذیرفته نمی شود. در صورت کشف هرگونه تقلب مطابق قوانین درس با افراد خاطی برخورد خواهد شد.
- استفاده از مراجع با ارجاع به آنها بلامانع است. اما در صورتی که گزارش شما ترجمه عینی از آن ها باشد، یا از گزارش افراد دیگر استفاده کرده باشید کار شما تقلب محسوب می شود.
- بعد از مطالعه ی کامل و دقیق این توضیحات، در صورتی که سوالی در مورد پروژه داشتید بهتر است در فروم درس مطرح کنید تا بقیه از آن استفاده کنند، در غیر این صورت یا در گروه تلگرامی مطرح کنید یا به طراحان پروژه ایمیل بزنید.

Elahe.bvakili97@ut.ac.ir

sh.vassef@ut.ac.ir

banafshehkarimian@ut.ac.ir

m.karimi76@ut.ac.ir