

گزارش پروژه کارشناسی

از بهمن ماه ۱۴۰۰ تا شهریور ۱۴۰۲

قوام الدین سلیمانی

دانشکده مهندسی برق و کامپیوتر
دانشگاه کاشان

استاد راهنما : سرکار خانم دکتر فرشته دهقانی

فهرست مطالب

۲	۹ بهمن تا ۲۲ بهمن ۱۴۰۰
۳	۲۸ اسفند ۱۴۰۰ تا ۵ فروردین ۱۴۰۱
۴	۳ دی تا ۱۶ دی ۱۴۰۱
۵	۲۰ دی تا ۸ بهمن ۱۴۰۱
۶	۲ اردیبهشت تا ۱۵ اردیبهشت ۱۴۰۲
۷	۶ خرداد تا ۱۹ خرداد
۸	تیر ۱۴۰۲
۹	۲۵ مرداد تا ۸ شهریور ۱۴۰۲
۱۰	۹ شهریور تا ۱۵ شهریور ۱۴۰۲

۹ بهمن تا ۲۲ بهمن ۱۴۰۰

- مطالعه مقاله و ترجمه
ابتدا مقاله ی جدید که قرار بود مطالعه شود ، بررسی و سپس ترجمه کردم و بخش هایی از اون رو که متوجه نمیشدم یادداشت برداری و درموردش سرچ کردم و از نتایج استفاده کردم.
- تماشای ویدئو های مربوط به الگوریتم های داده کاوی و مدل سازی
مدل های مختلفی برای تحلیل داده ها و پیش بینی ها وجود دارد . جهت یادآوری درس داده کاوی و مفاهیم آن ، ویدئو ها را تماشا و نکات را یادداشت برداری کردم.
- دیتاست
پس از جستجو در مورد دیتاست مورد بحث در مقاله ، آن را در kaggle یافتم .
- آغاز یادگیری Latex
نوشتن پایان نامه و اسلاید سمینار با Latex برایم اهمیت ویژه ای داشت . پس زمانی صرف آن کردم.
- Data clearing
داده هایی که در دیتاست وجود دارد حاوی مقادیر زیادی از NaN هستند و روش هایی برای استاندارد سازی آن وجود دارد.
- مدل سازی برای رگرسیون لجستیک
جهت این کار لازم است متغیر های Dummy ایجاد کنیم . چون داده های ما اسمی هستند و در مدل رگرسیون خطی آن ها را به کار نگرفته ایم . سپس از sklearn کمک میگیریم.

۲۸ اسفند ۱۴۰۰ تا ۵ فروردین ۱۴۰۱

- ایجاد قالب گزارش
- در ادامه یادگیری Latex ابتدا دنبال قالب های آماده بودم اما به دلایل مختلف تصمیم گرفتم یک قالب را بر اساس نمونه های موجود خودم طراحی کنم.
- ادامه مطالعه مقاله و ترجمه
- بخش دیگری از مقاله باقی مانده بود و با مرور بخش قبلی ، بخش جدید را ادامه دادم.
- مدل سازی جهت RF
- مطابق مراحل که برای رگرسیون لجستیک استفاده شد بر اساس مستندات موجود در sklearn انجام گشت.
- ادامه یادگیری Latex
- بخش های زیادی از Latex همچنان برایم ابهام داشت و به مرور های زیادی برمیخورم و باید بیشتر جستجو میکردم.
- آغاز نوشتن مطالبی جهت استفاده در پایان نامه
- از نکات یادداشت شده و استفاده از مقاله ی مربوطه برای نوشتن پایان نامه استفاده کردم . همچنین درمورد نکاتی درباره پایان نامه مناسب تحقیق کردم.

۳ دی تا ۱۶ دی ۱۴۰۱

- مرور و مطالعه و یادآوری فعالیت های گذشته
به علت وقفه طولانی در حین پروژه مشغول به مرور موارد قبلی شدم.
- مصور سازی داده ها
مطابق بخش هایی از مقاله و جهت تهیه برخی گزارشات این بخش یکی از بخش های مهم است که باید انجام شود و معمولاً قبل از مدل سازی با الگوریتم های مختلف انجام می شود .
- ایجاد قالب مناسب برای beamer و ایجاد اسلاید ها
برای ارائه سمینار لازم است قالب مربوطه در Latex ایجاد شود و از PowerPoint پرهیز میکنم. پس یک Draft از ساختار کلی و به صورت آزمایشی و با جنبه آموزشی در beamer ایجاد کردم.
- ادامه پایان نامه
با تمرکز بر فصل یک بخش هایی تکمیل شد و فصل های بعدی مطالبی به عنوان نمونه و به صورت Draft اضافه شد.
- اصلاح نوت بوک
نوت بوکی که ایجاد کردم به هم ریختگی هایی داشت و نیاز بود در بخش هایی توضیحاتی بیفزایم. از جمله CleanCoding و برخی ها Markdown
- مطالعه در مورد نحوه عملکرد الگوریتم ها
برای نوشتن الگوریتم های مورد نظر لازم است یک مرور کلی بر نحوه عملکرد آنها داشته باشیم تا بتوانیم عملکردشان را درک کنیم و در پایان نامه نیز ذکر کنیم.
- تکمیل گزارشات

۲۰ دی تا ۸ بهمن ۱۴۰۱

- مرور و مطالعه و یادآوری فعالیت های گذشته
- تنظیم فونت و سائز و اصلاح برخی شماره گذاری ها
- اعمال cross validation
- فصل مقدمه تا حدودی تکمیل و اصلاح شد
- افزودن رفرنس های فصل ۱
- تغییر ساختار فصل های پایان نامه
- ادامه تکمیل فصل ۲ شامل ۴ الگوریتم اصلی
- مطالعه درمورد نحوه عملکرد ریاضیاتی هر الگوریتم
- پیش نویس فصل کارهای پیشین و بخشی از ۲ فصل پایانی
- یادگیری رسم برخی نمودار ها در Latex
- تکمیل گزارشات

۲ اردیبهشت تا ۱۵ اردیبهشت ۱۴۰۲

- مرور و مطالعه و یادآوری فعالیت های گذشته
- نوشتن روش عملکرد الگوریتم های naive bayes و adaboost
- تکمیل بخش کار های پیشین
- تکمیل بخش روش های داده کاوی در فصل ۲
- تکمیل بخش روش های اندازه گیری صحت مدل

۶ خرداد تا ۱۹ خرداد

- ایجاد بخش های مربوط به بخش اینترنت اشیا Arduino , Node-Red , Flask و راه اندازی آنها
- تکمیل بخش اینترنت اشیا در پایان نامه
- اصلاح به هم ریختگی فاصله نوشته ها در Latex

تیر ۱۴۰۲

- خرابی هارد دیسک لپ تاپ به علت بدسکتوری (۳ میلیون تومان خسارت)
- از دست رفتن اطلاعات و انجام مجدد موارد تغییرات در پایان نامه از بهمن ۱۴۰۱ تاکنون به علت ثبت نکردن commit برای آن:

۲۵ مرداد تا ۸ شهریور ۱۴۰۲

- مرور ، مطالعه و یادآوری درمورد بخش های مربوط به فصل ۲
- اصلاح بخش پیش پردازش داده ها در فصل ۲ و تکمیل موارد مربوط به آن
- تکمیل بخش لغت نامه
- ایجاد بخش پیاده سازی در پایان نامه
- تکمیل بقیه الگوریتم های یادگیری ماشین در فصل ۲
- وارد کردن نمودار های مختلف در بخش نتایج
- اصلاح مدل های cross validation و رسم نهایی نمودار ها و به دست آوردن نتایج نهایی
- تکمیل بخش نتایج در پایان نامه
- ساخت اسلاید های Beamer
- اصلاح و تکمیل فصل پایانی جمع بندی و کارهای آتی
- اصلاح و تکمیل چکیده
- بررسی کلی موارد
- تکمیل گزارشات

۹ شهریور تا ۱۵ شهریور ۱۴۰۲

- اصلاح روش feature selection
- تغییر پیدا کردن متغیر ها و رسم نمودار ها و تغییرات کلی در ساختار پروژه
- حذف روش dummy variables و استفاده از encoding به جای آن
- افزودن Resampling (SMOTE for Over sampling)
- مقایسه حالت های resample شده و resample نشده
- اصلاحات نگارشی در پایان نامه
- تغییر node-red
- حذف روش رگرسیون لجستیک از روش های سنجش به دلیل استفاده از آن در FE
- افزودن عنوان ماشین لرنینگ افزایشی به بخش کارهای آتی

۱۶ شهریور تا ۲۰ شهریور ۱۴۰۲

- اصلاح فونت
- اصلاح افعال
- افزودن Recall & Precision و نمودارها به پایان نامه و اسلاید ها