

Machine Learning Course By Dr. Nazerfard

CE5501 | Fall 2023

Teaching Assistants

AMIRHOESSION BABYIAN

دانشجو: نیما پری فرد

شماره دانشجویی: 402131017

ایمیل: nima.parifard@aut.ac.ir

فهرست ارائه نهایی پروژه درس یادگیری ماشین

[بخش 1 بیان چالش های پیش آمده: 2](#_Toc158835997)

[بخش 2 شبیه سازی و ترسیم گراف مترو: 3](#_Toc158835998)

[بخش 3 ساخت دیتاست: 6](#_Toc158835999)

[بخش 4 تربیت مدل یادگیری ماشین برای خروجی هر ایستگاه: 7](#_Toc158836000)

[بخش 5 تربیت مدل یادگیری ماشین برای ورودی هر ایستگاه: 7](#_Toc158836001)

[بخش 6 تربیت مدل یادگیری ماشین برای افراد حاضر در مترو: 8](#_Toc158836002)

[بخش 7 تربیت مدل یادگیری ماشین برای برای تخمین نرخ پواسون: 8](#_Toc158836003)

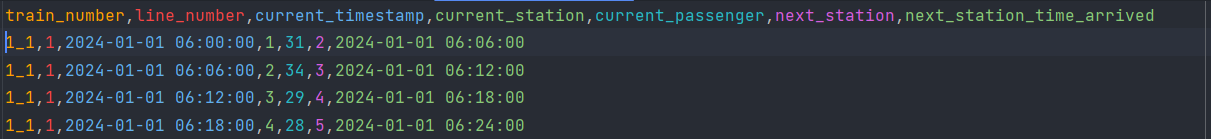
[بخش 8 ارزیابی: 9](#_Toc158836004)

لینک گیت هاب پروژه: <https://github.com/nimaiparifard/cursed-metro-ml-final-project->

# بخش 1 بیان چالش های پیش آمده:

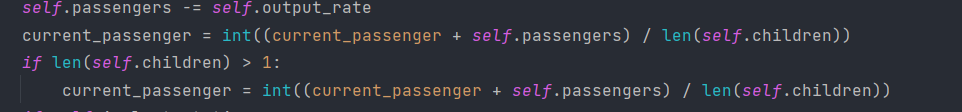
1. اینکه در تایم های اول صبح که هنوز اولین قطار ها به ایستگاه ها نرسیده خروجی از ان ایستگاه نداریم.

راه حل: من با استفاده یک دیتاست که در آن مشخص کردم هر قطار چه ساعتی به ایستگاه می رسد مشکل خود به خود می شود. و واضح چون در آن ساعت کسی وارد نشده



1. در ایستگاه های مشترک خروجی ها چگونه تقسیم شوند

راه حل: در کلاس بحث شد که به دوقسمت تقسیم می شوند من با استفاده از Children گزاشتن روی ایستگاه ها خروجی را تقسیم بر تعداد بچه های هر ایستگاه می کردم.



1. نرخ خروجی هر قطار در یک ایستگاه نباید بیش تر از افراد داخل ان قطار شوند

راهحل: با استفاده از همان فایل csv من در لحظه افراد حاضر در هر قطار را داشتم و یک شرط گزاشتم که خروجی از افراد حاضر در قطار بیش تر نشود.

A screen shot of a computer code

Description automatically generated

1. وقتی قطار به ایستگاه اخر می رسد باید همه افرادی که داخل هستند خارج شوند

A black background with white text

Description automatically generated

وقتی نرخ خروجی را می خواستم تعیین کنم اول چک می کردم ببینم در ایستگاه اخر هست یا خیر

1. در تایم اخر بعد از اینکه در بعضی ایستگاه ها اخرین قطار رد شد دیگر ورودی در آن ایستگاه ها نداریم و مترو بسته می شود

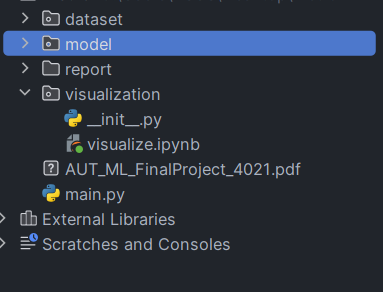
با استفاده از همان dataset می توان اینکار به طور اتوماتیک حل می شود.

# بخش 2 شبیه سازی و ترسیم گراف مترو:

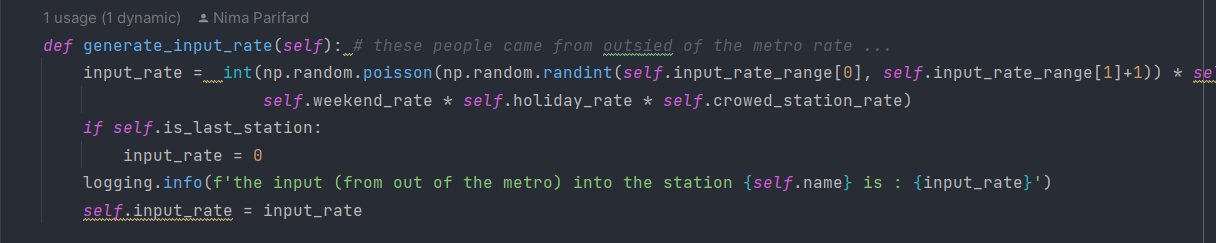
شبیه سازی با حل مشکلات قبل به دست آمد اگر فایل metro.py اجر کنید شبیه سازی نشان داده می شود.

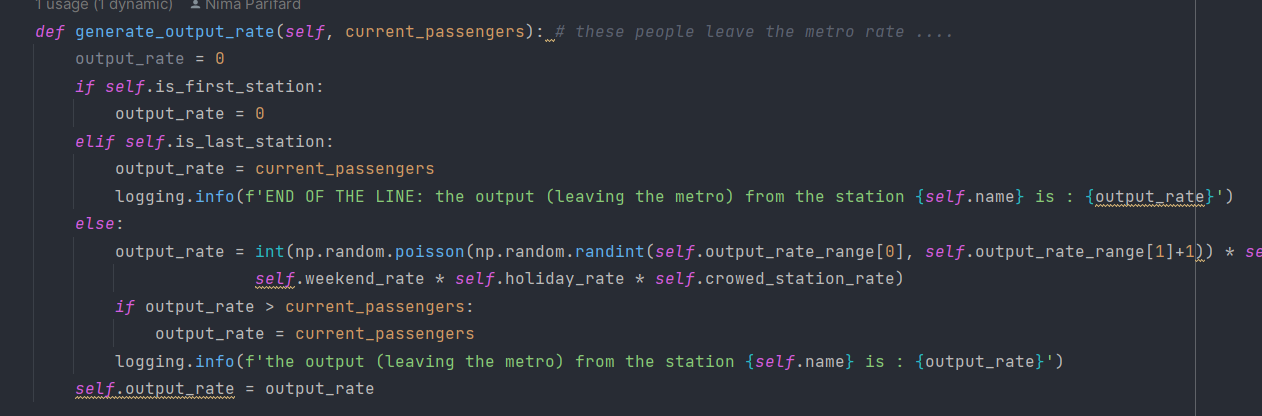
پروژه بنده دارای سه ماژول dataset که شبیه سازی در آن انجام شده است. دارای سه فایل مهم است. Station.py که عملیات هایی که برای هر ایستگاه نیاز است انجام می دهیم. در metro.py جریان شبیه سازی انجام می شود. درنهایت یک فایل csv به ما می دهد که داده های شبیه سازی در آن بارگزاری شده است. و یک train\_status\_creation.py که فایل حرکت قطار و زمان رسیدنشون با ایتسگاه های متخلف در یک روز بارگزاری می کند.

در ماژول مدل 4 فایل جوپیتر وجود دارد که هر کدام مدل های ورودی و خروجی و پواسون و افراد حاظر در ایستگاه را پیش بینی می کند.



در کلاس Station سه تابع مهم وجود دارد یکی برای تولید نرخ ورودی و تولید نرخ خروجی و همین طور تابعی برای اینکه جریان مسافر را تولید کند.

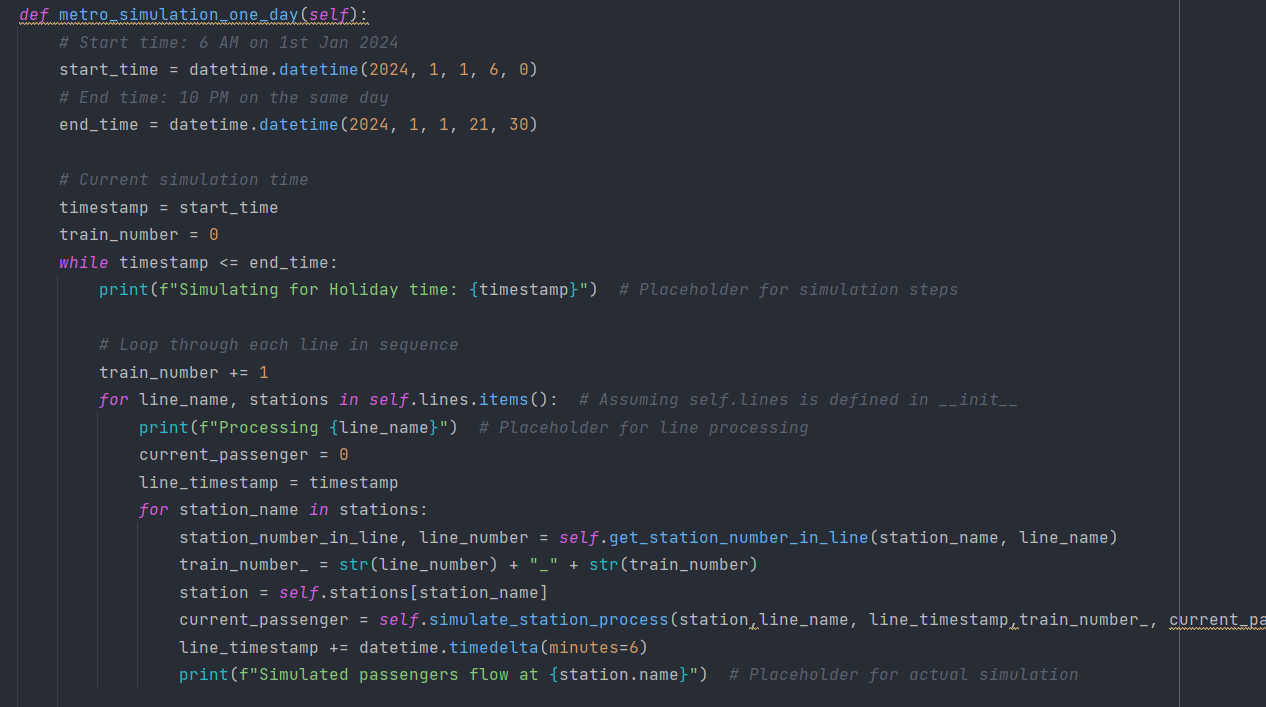




A screen shot of a computer program

Description automatically generated

در کلاس Metro اول از همه تمام ایستگاه با استفاده از کلاس station تمام ایستگاه ها اول مقداردهی می شوند. و بعد با تابع metro\_simulation\_one\_day شبیه سازی انجام می دهیم.



گراف پیاده سازی مترو:

A screenshot of a computer

Description automatically generated

# بخش 3 ساخت دیتاست:

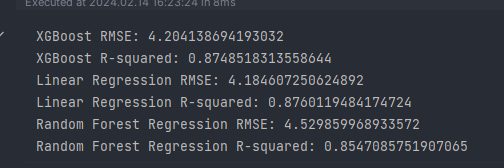
برای ساخت دیتاست من خودم ویژگی های دیگری اضافه کردم مثل اینکه آیا در ساعت شلوغی هست یا خیر. روز تعطیل است یا خیر و روز اخر هفته است یا خیر شماره خط ایستگاه و اینکه ایا جزو ایستگاه های شلوغ هست یا خیر. ایستگاه های شلوغ بعضی ایستگاه ها هستند که خودم مشخص کردم.

A screen shot of a computer

Description automatically generated

# بخش 4 تربیت مدل یادگیری ماشین برای خروجی هر ایستگاه:

اول از همه به علت کم بودن ویژگی ها از یک کرنل دو بعدی برای بالا بردن تعداد ویژگی ها و بعد آن ها استفاده کردم. بعد سه تا مدل رگرسیون خطی و xgboost و random forest استفاده کردم و نتایج زیر به دست آمد.



# بخش 5 تربیت مدل یادگیری ماشین برای ورودی هر ایستگاه:

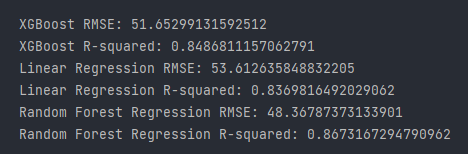
اول از همه به علت کم بودن ویژگی ها از یک کرنل دو بعدی برای بالا بردن تعداد ویژگی ها و بعد آن ها استفاده کردم. بعد سه تا مدل رگرسیون خطی و xgboost و random forest استفاده کردم و نتایج زیر به دست آمد.

A screenshot of a computer

Description automatically generated

# بخش 6 تربیت مدل یادگیری ماشین برای افراد حاضر در مترو:

اول از همه باید دیتاست آماده می کردم که در هر لحظه چند نفر داخل ایستگاه هستند. بعد برای بقیه مدل ها مانند دو روش قبلی عمل کردم.



# بخش 7 تربیت مدل یادگیری ماشین برای برای تخمین نرخ پواسون:

برای نرخ پواسون از مدل های یادگرفته شده استفاده کردم میانگین در دیتاست برای هر ایستگاه محاسبه کردم و به عنوان ورودی مدل دادم و در نهایت خروجی را دریافت کردم. به عنوان نرخ پواسون معرفی شده است.

# بخش 8 ارزیابی:

برای ارزیابی از دو معیار rmse و r2 استفاده کردم که در شکل برای هر مدل این معیاری ها را مشاهده می کنید. معیار r2 به خاطر اینکه بین صفر و یک قرار می گیرد حس بهتری به ما می دهد نسبت به سایر مدل ها.