

پروژه آنتی بیوتیک فاز دوم :

در این فاز از پروژه داده های مورد نیاز برای پیش بینی مصرف آنتی بیوتیک در سال های آینده را توانستیم پیدا کنیم و بر اساس آنها توانستیم حالت های مختلفی که بهترین خروجی را میدهد ایجاد کنیم .

تا در فاز سوم بتوانیم از آن استفاده کرده و پیش بینی نهایی را انجام بدهیم .

در فایل اول ایجاد شده در این فاز:

(Antibiotic_project_2_DataPreparation)

به این نتیجه رسیدیم که باید از دوز مصرفی ماهانه برای پیش بینی استفاده شود. داده ها را گرفته و تغییراتی در سطر ها و ستون های آن ایجاد کردیم .

و داده های گم شده ایی که وجود داشت را نیز پیدا کردیم.

از طرفی همسان سازی داده ها را نیز انجام دادیم تا داده ها را بتوانیم در مدل خود قرار بدهیم.

در فایل دوم ایجاد شده از این فاز :

(Antibiotic_project_2_withoutnormalizer)

داده های آماده شده از فایل قبلی را که در اختیار داشتیم داده های گم شده را پیدا کردیم و تمام سطر هایی که داده گم شده در آن ها وجود داشت را نیز حذف کردیم و همچنین داده های نویزی که همان داده های منفی بودند را پیدا کردیم و مقدار آن ها را برابر با صفر قراردادیم تا بتوانیم مدل بهتری را ایجاد کنیم.

از طرفی برای ایجاد مدل شبکه عصبی و از انجایی که داده ها به صورت سری زمانی بودند می بایست تاریخچه ایی برای داده های خود ایجاد کنیم

بنابراین تاریخچه را با گرفتن داده های یک سال گذشته و شیفت دادن توانستیم ایجاد کنیم .

حال داده هایی که در تاریخچه قرار گرفته بودند را نیز با استفاده از شبکه عصبی مدل سازی کردیم و در شبکه عصبی از دو لایه پنهان به ترتیب با ۵۰ و ۱۰۰ نورون ایجاد کردیم .

اما همچنان داده های خود را نرمال نکرده ایم برای قرار دادن در شبکه عصبی و با استفاده از

معیار RMSE توانستیم خطا شبکه عصبی را پیدا کنیم که عدد

۵۰۵۸۰۴.۷۸ به دست آمد . و توانستیم نمودار مربوط به پیش بینی را هم ایجاد کنیم . و تفاوت

مقدار اصلی و پیش بینی شده با شبکه عصبی را مشاهده کردیم.

در فایل سوم ایجاد شده از این فاز :

(Antibiotic_project_2_StandardScaler)

تنها تفاوت ایجاد شده در این فایل با فایل قبلی نرمال کردن داده ها در شبکه عصبی بود که با

استفاده از نرمال کردن داده ها خطا کاهش پیدا کرد اما برای تبدیل شدن داده های نرمال شده به

داده اصلی خطای خیلی بزرگی ایجاد شد

پس به این نتیجه رسیدیم در این پروژه بهتر است از نرمال کردن داده ها صرف نظر شود.

نکته قابل توجه ایی که وجود داشت :

به دلیل کمبود داده ها و نبودن هیچگونه کلید خارجی و ارتباط بین بخش های مختلف از داده ها

پیش بینی برای ماشین کمی با مشکل روبه ور شد و تنها راهی که وجود داشت کاهش خطا ایجاد

شده چه با استفاده از تغییرات در داده ها و چه با استفاده از تغییر در شبکه عصبی بود بنابراین

خطای ایجاد شده که به نسبت مقادیر بزرگ داده ها نیز خطای زیادی نمی باشد به دلیل کمبود

داده بود.