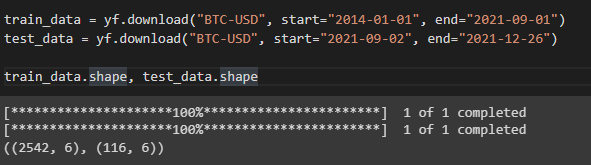
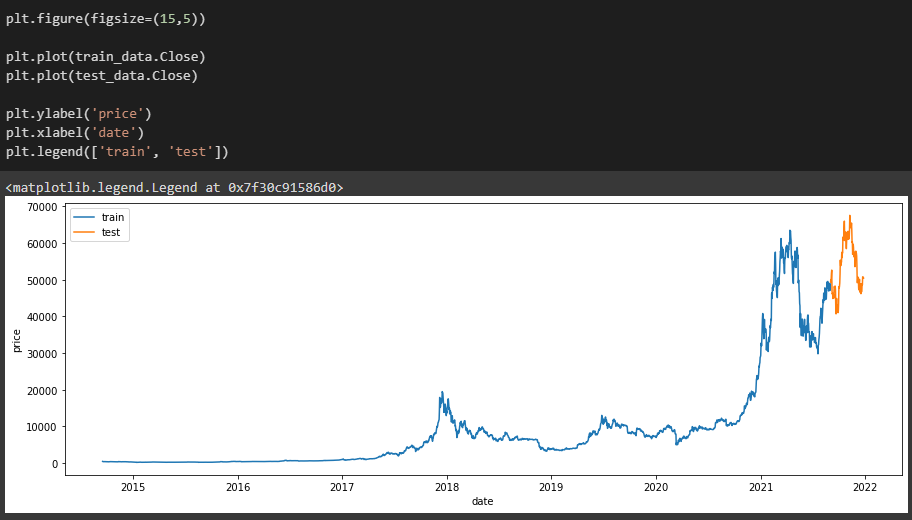
**امیرحسین احمدی 97522292 – تمرین دوازدهم**

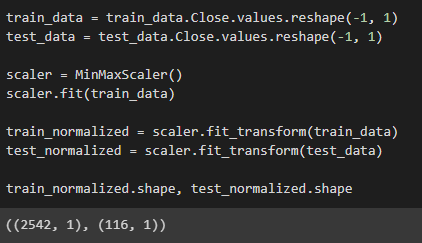
1. برای این سوال ابتدا data frame های مربوط به داده های ترین و تست را همانطور که در سوال گفته شده از 1-1-2014 تا 1-9-2021 برای داده های ترین و از آن به بعد تا الآن را برای تست دانلود میکنیم.



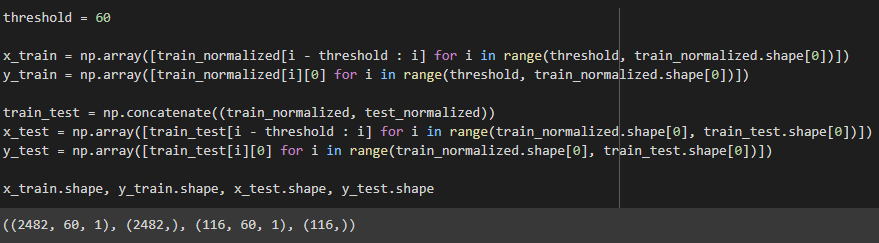
سپس نمودار قیمت زمان بسته شدن (close) را برای کل بازه ی داده های تست و ترین میکشیم.



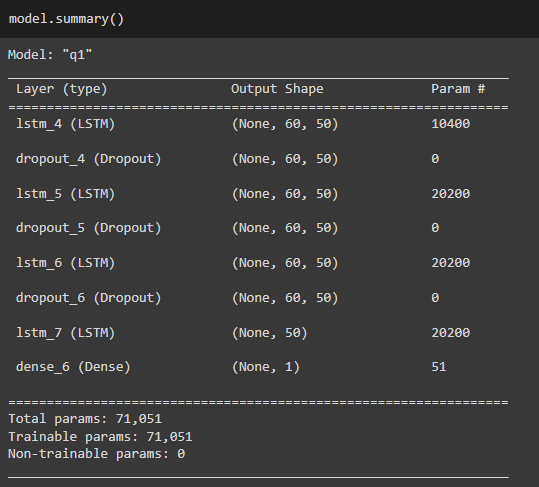
سپس قیمت close را به عنوان داده هایی که میخواهیم روی آن مراحل آموزش را انجام دهیم انتخاب کرده و با MinMaxScaler مقادیر را برای داده های ترین و تست، اسکیل میکنیم.



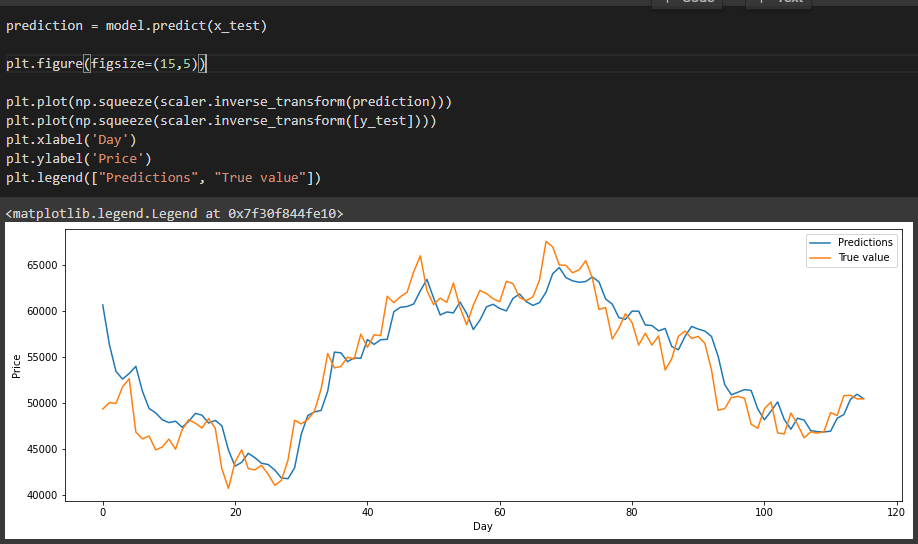
سپس x\_train را به این صورت ساخته که به ازای هر داده، 60 داده قبلی اش را به عنوان فیچرهای آن در نظر میگیریم و داده هایی که 60 داده قبل آن ها را نداریم را در نظر نمیگیریم. برای داده های تست نیز همینطور عمل میکنیم، با این تفاوت که بعضی داده ها 60 روز قبلشان در داده های ترین قرار میگیرد، به همین دلیل ابتدا دو داده را چسبانده و سپس x\_test را محاسبه میکنیم.



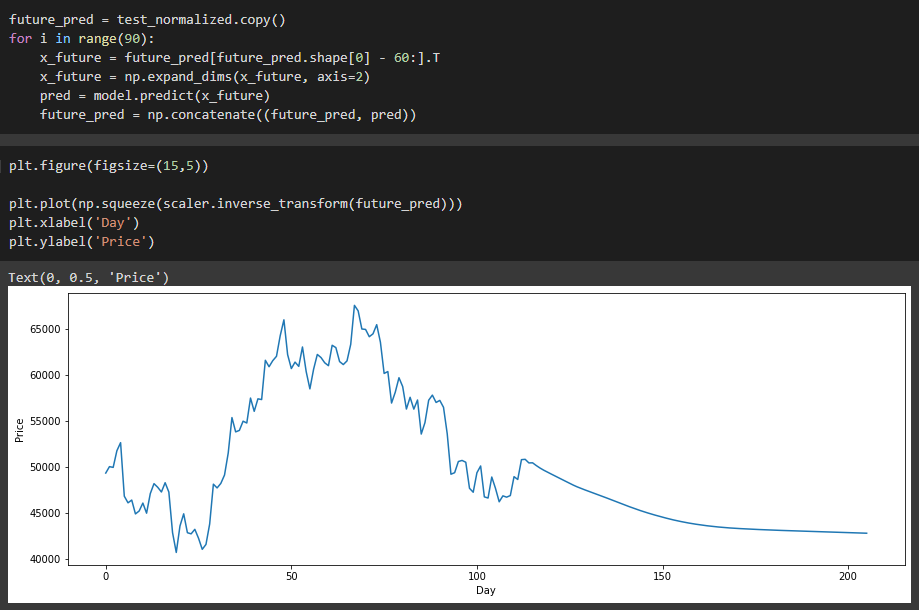
سپس مدل را طبق صورت سوال ساخته و با داده های ترین آموزش میدهیم.



در نهایت داده های تست را به مدل داده و آن ها را به همراه لیبل اصلیشان روی نمودار میبریم. قبل این کار نیز با استفاده از scaler قیمت ها را به حالت عادیشان برمیگردانیم.

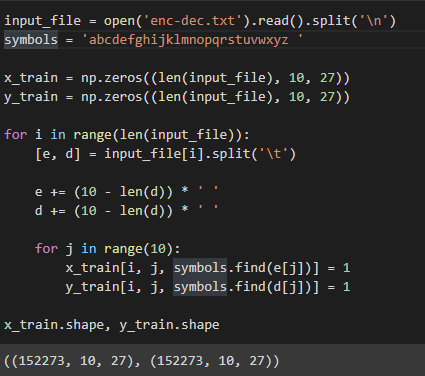


حال برای پیشبینی 90 روز آینده، روزها را یکی یکی جلو رفته و از داده های قبلی و داده های پیشبینی شده قبلی برای آن استفاده میکنیم و در نهایت آن را روی نمودار میبریم.

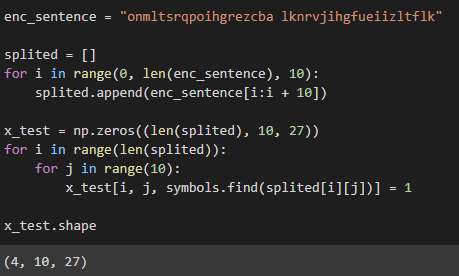


در ارتباط با threshold در نظر گرفته شده برای فیچرها که در کد برابر با 60 در نظر گرفته شده، هر چقدر این threshold بیشتر شود، فیچرهای بیشتری داریم و احتمالن به مدل کمک میکند تا پیشبینی بهتری داشته باشد، همچنین با توجه به ترتیبی بودن آموزش در شبکه های بازگشتی، با زیاد کردن تعداد فیچرها، سرعت به شدت پایین میاید. همچنین اگر threshold را خیلی خیلی زیاد کنیم ممکنه باعث شود که دیتاهای قدیمی که دیگر زیاد ربطی به داده های جدید ندارند استفاده شوند و مدل به اشتباه بیوفتد.

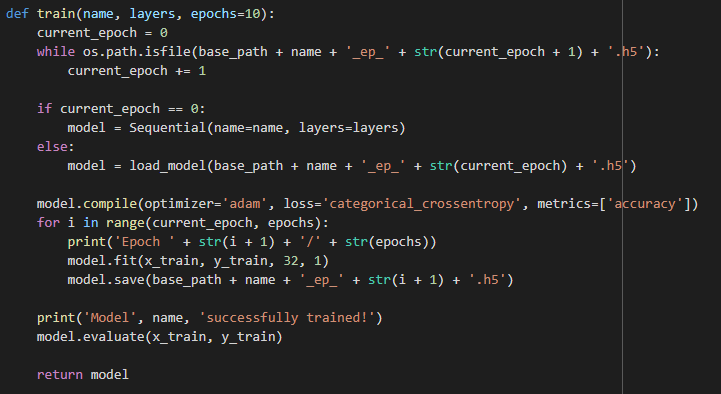
1. برای این سوال در ابتدا فایل enc-dec.txt را گرفته و x\_train و y\_train را از آن استخراج میکنیم. اگر هر کدام از کلمات نیز کمتر از 10 حرف داشت به آن اسپیس اضافه میکنیم و آن ها را به صورت one hot در میاوریم.



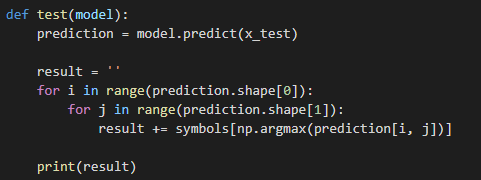
سپس جمله ی داده شده در صورت سوال برای تست را گرفته و 10 کلمه 10 کلمه جدا میکنیم و آن را نیز به صورت one hot در میاوریم و در x\_test نگه میداریم.



در ادامه تابع train را تعریف میکنیم که ورودی آن یک اسم، لایه های مدل مورد نظر و تعداد ایپوک ها برای train است. هر بار که train صدا زده میشود چک میکند که مدلی با اسم مدنظر قبلن train شده است (در drive موجود است) یا خیر. در صورت وجود داشتن بالاترین ایپوک آن را لود کرده (برای هر ایپوک هر مدل آن را سیو میکنیم) و در صورت موجود نبود مدل را با اسم و لایه های داده شده میسازیم. سپس آن را از ایپوکی که ترین شده تا آخرین ایپوک ترین کرده و به ازای هر ایپوک مدل را سیو میکنیم. در آخر نیز تابع evaluate را صدا میزنیم تا نتیجه نهایی مدل را ببینیم.

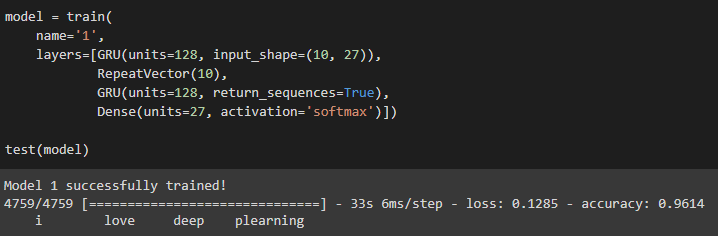


در کنار تابع train تابع test را نیز تعریف کرده که x\_test را به مدل داده و جمله خروجی مدل را چاپ میکند.



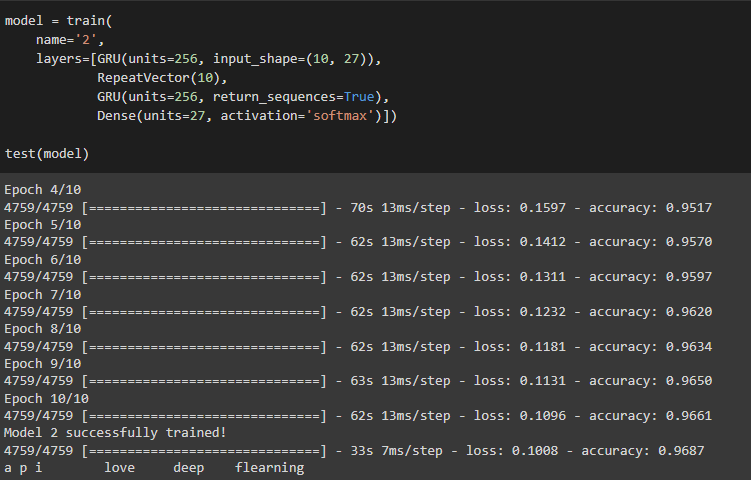
حال مدل های مختلف با لایه های GRU را امتحان کرده تا ببینیم کدام نتیجه دلخواه را به ما میدهد. ساختار کلی مدل ها به این صورت است که در ابتدا یک لایه GRU به عنوان many to one داریم، سپس تعدادی لایه GRU به عنوان many to one و در آخر نیز یک لایه dense با فعالساز softmax و 27 یونیت (تعداد حروف الفبا + اسپیس) برای انتخاب حرف مورد نظر داریم.

مدل اولی که امتحان کرده ایم یک لایه GRU میانی با 128 یونیت دارد.



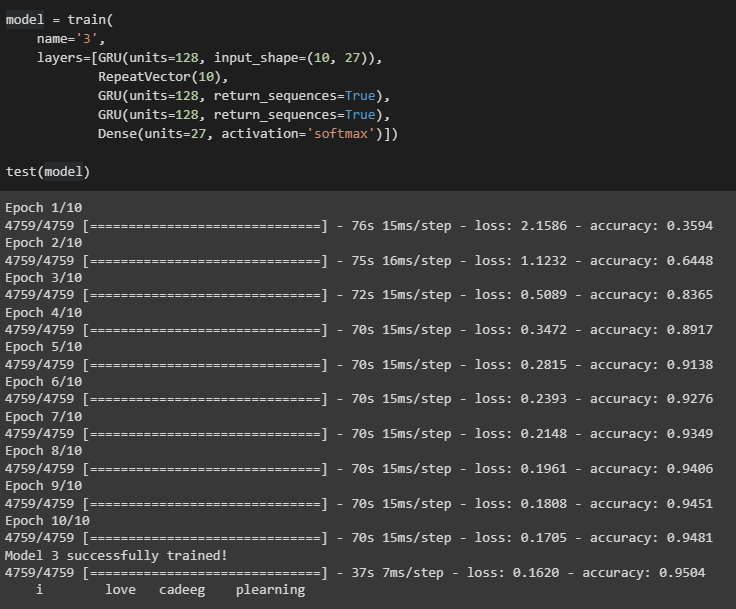
همانطور که میبیند مدل دقت خوبی دارد و پیشبینی خوبی داشته و فقط اسپیس اول learning را p تشخیص داده است.

سپس تعداد یونیت های لایه های GRU را به 256 افزایش میدهیم.



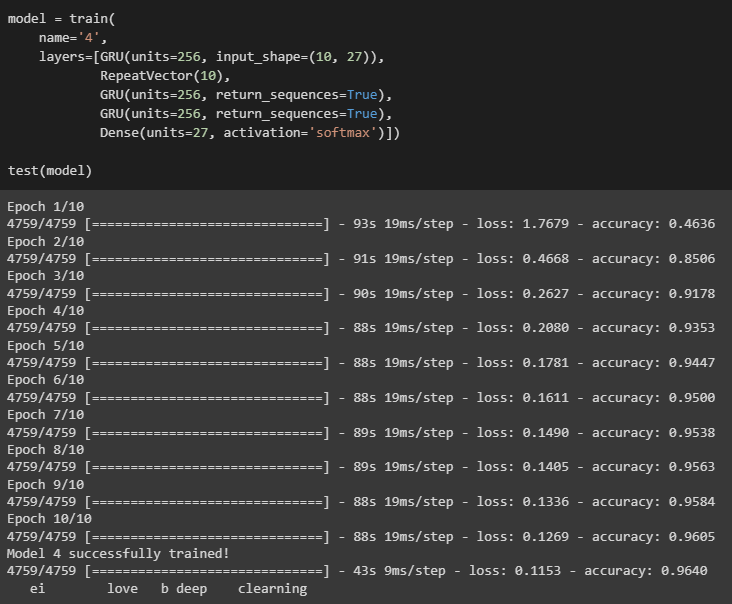
مشاهده میکنیم که دقت بسیار بهتر شده که با توجه به یونیت های بیشتر طبیعی است که مدل زودتر همگرا شود. ولی میبینیم که خروجی تست خوبی ندارد که احتمالن نشان از overfit شدن مدل دارد.

برای مدل سوم یونیت ها را به 128 برگردانده و در عوض یک لایه gru میانی دیگر اضافه میکنیم.



مشاهده میشود که مدل دیرتر از دو مدل قبلی همگرا شده و دقت کمتری دارد و نتیجه تست خوبی هم ندارد و به نظر میاد که underfit است و نیاز به ایپوک های بیشتر دارد.

برای مدل چهارم یونیت های مدل سوم را افزایش دادیم تا سریع تر همگرا شود.



همانطور که میبینید مانند مدل 2 دقت خوبی دارد ولی overfit شده و روی داده تست دقت خوبی نداشته است.