**Dokumentacija koda**

**Uvod:**

Da bi razumjeli moj kod i moju inspiraciju za napravit ga prvo moramo razumjeti osnove dionica i financija.

**Financije u kratko**

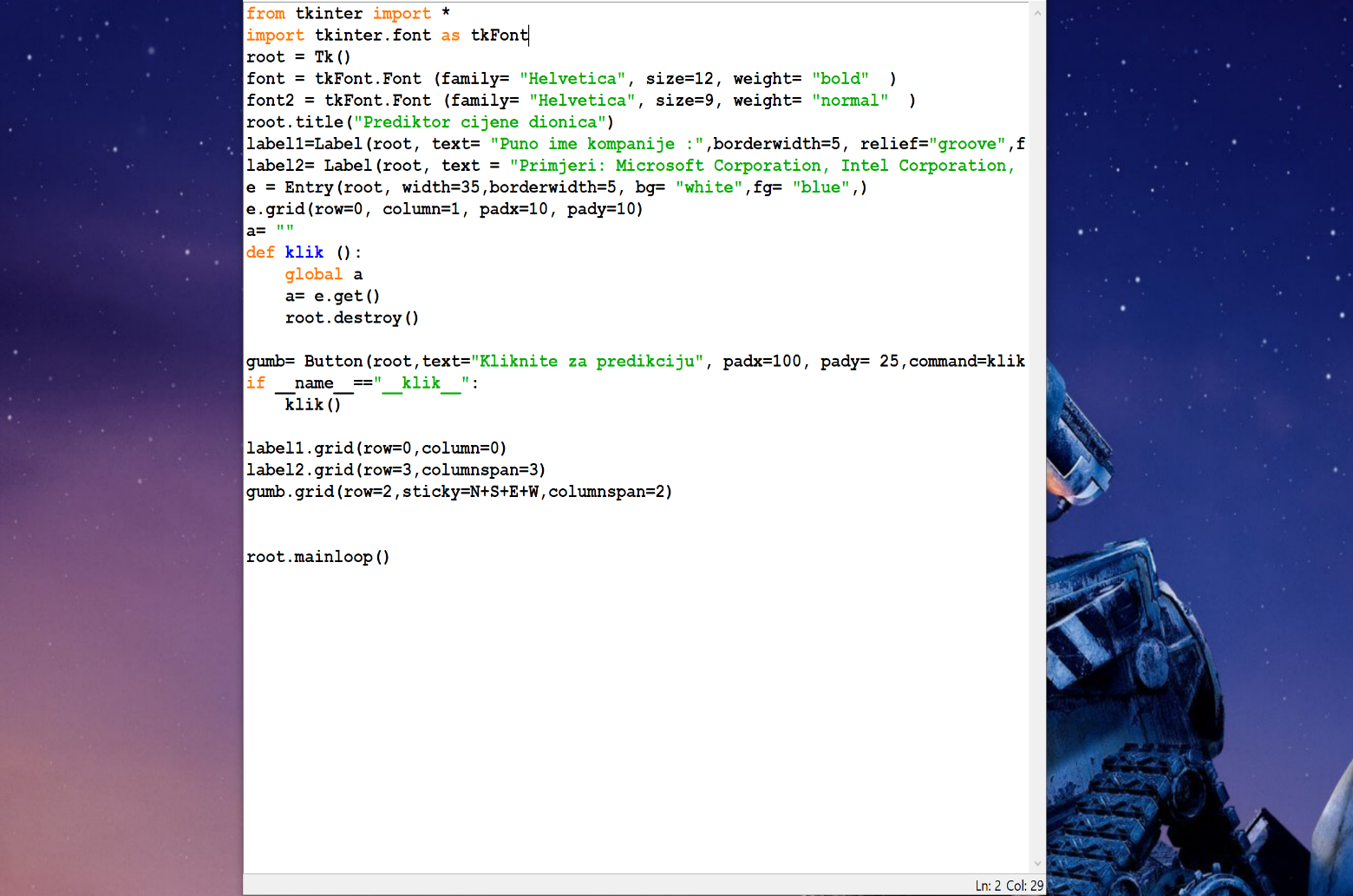
Svakoga dana događa se nešto nama nepoznato, velike kompanije poput Facebooka i Googla još prije mnogo godina počeli su prodavati dionice na njujorškoj burzi. Što to nama znači ? U suštini znači da mi obični ljudi možemo biti vlasnici maloga udjela neke velike kompanije. Od ponedjeljka do petka, od 9 do 16 sata(po američkom istočnom vremenu) investitori kupuju i prodaju dionice. To radi na vrlo jednostavnom principu. Kompanija podjeli svoju vrijednost na puno malih djelova( dionica) te im onda zada početnu cijenu. Ako izađe neka dobra vijest o kompaniji cijena će porasti i oni koji su kupili dionice po manjoj cijeni mogu ih sada prodati po večoj. Tako danas ukratko kompanije i investitori zarađuju novac. Investitori već godinama pokušavaju naći dugotrajni sistem predikcije cijene dionica kako bi zaradili novac. Tako je došla moja inspiracija, odlučio sam napraviti model pomoću strojnog učenja koji bi predvidio cijene dionica. Moj cilj i ideja jest da korisnik upiše ime kompanije za koje ga zanima cijena dionice u budućnosti. Ono što se dogodi sljedeće jest da izađe cijena do sada i predikcija u budućnosti.

**Detaljan opis rada:**

Kod sam podjelio u 3 djela, tako da sam na svakom radio zasebno i nakraju ih sve spojio u jedan veliki.

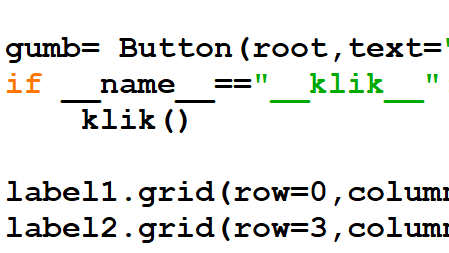
**1.Dio-GUI**

U ovom djelu isprogramirao sam GUI(to jest grafičko sučelje) u kojem korisnik upisuje ime dionice i klikče na gumb koji ga vodi na graf i predikciju. Koristio sam se biblioteku tkinter.

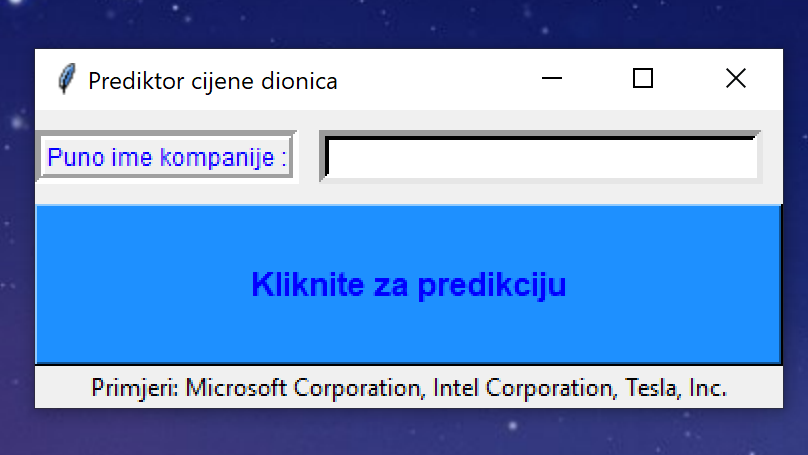


Pomoću *root = TK()* nisam morao cijelo vrijeme pisati tk. nego samo root usvakoj zagradi. Kao i u CSS-u narpavio konfiguraciju teksta(oblik,font,..) i samo je kasnije umetnuo u labele. Label1 i label2 predstavljaju dijelove na GUI ju na kojima će pisati tekst, dok e(Entry) predstavlja dio u kojemu će korisnik upisati informacije. Klik je samo dio u kojem se daje komanda što će se dogoditi kada korisnik klikne gumb. Ono što se dogodi jest da se sačuva input(*e.get*) i makne GUI(*root.destroy*). Gumb mora imati komandu(komanda je klik) i onda sam ga ja još oblikovao i obojao.

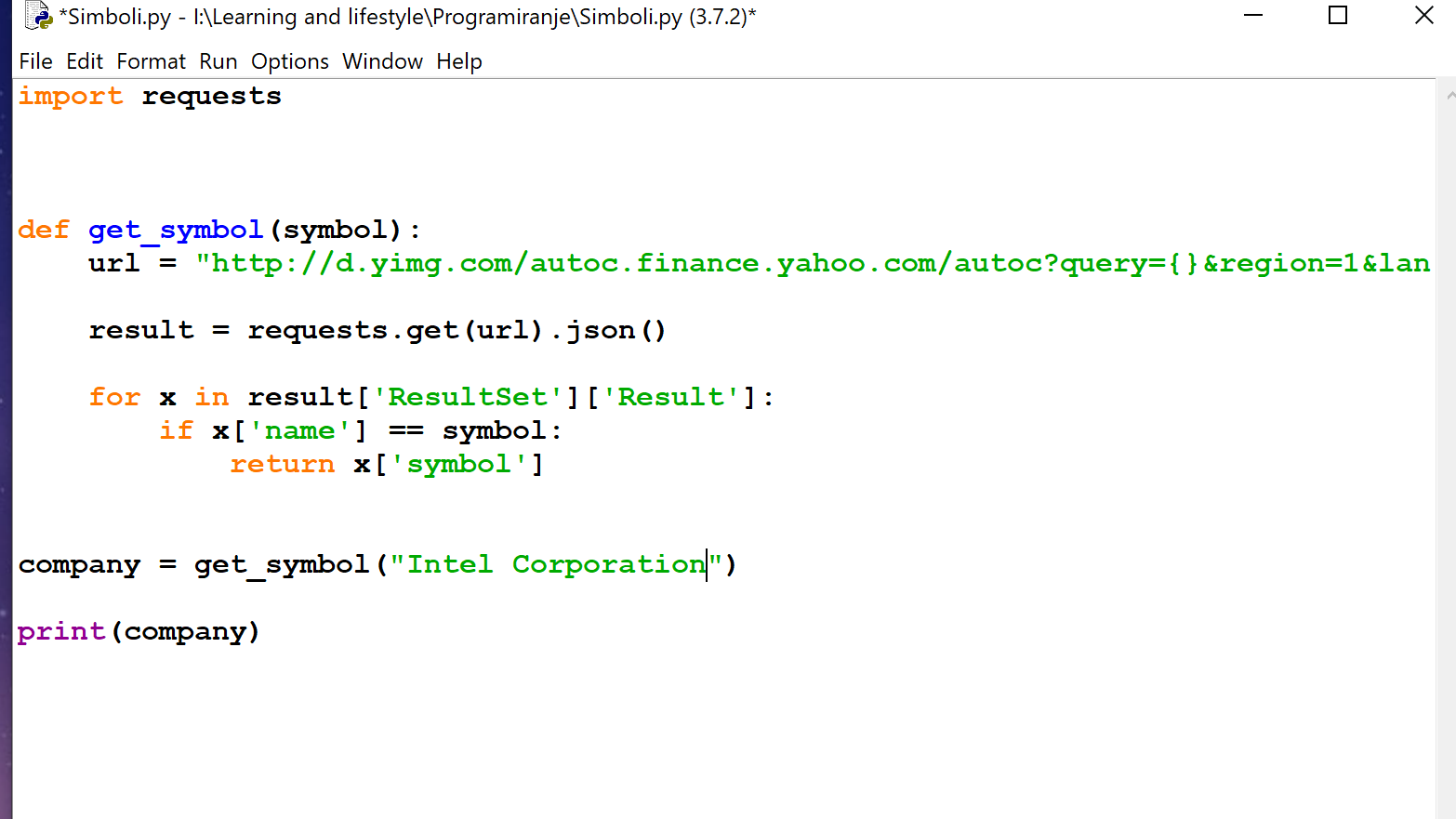
Ovaj dio na slici dolje je bitan jer on zapravo nalaže da je klik glavna funkcija pa onda se taj dio koda prvi izvršava i na taj način nemamo kasnije problema s varijablama koje nisu definirane(jer u kliku dajemo varijabli a string).



Svi ovi djelovi s gridom definiraju veličinu i koliko će prostora svaki segment uzimati na GUI – u. Kada pokrenemo kod dogodi se sljedeće:



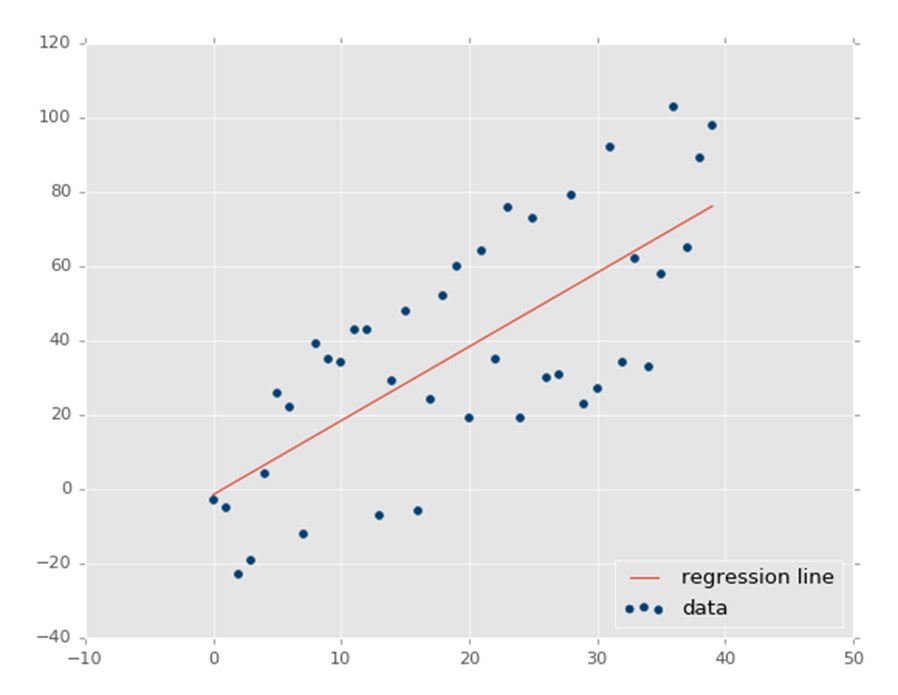
**2.Dio-Simboli**



U ovom djelu korisnikov input se uspoređuje s listom kompanija na API -u ( to jest na stranici). Kada se nađe isto ime kompanije „Kod“ uzima simbol te kompanije s te stranice i sačuva ga kao string. Dakle pomoću biblioteke requests smo se spojili na API(yahoo finance). Ova for petlja uspoređuje input s ostalim kompanija na toj stranici i kada se poistovjete returna njezin simbol. Kasnije se samo pozovemo funciju za neku kompaniju(ja sam ovdje stavio Intel) i ona nam da njezin simbol.

**3.Dio-model**

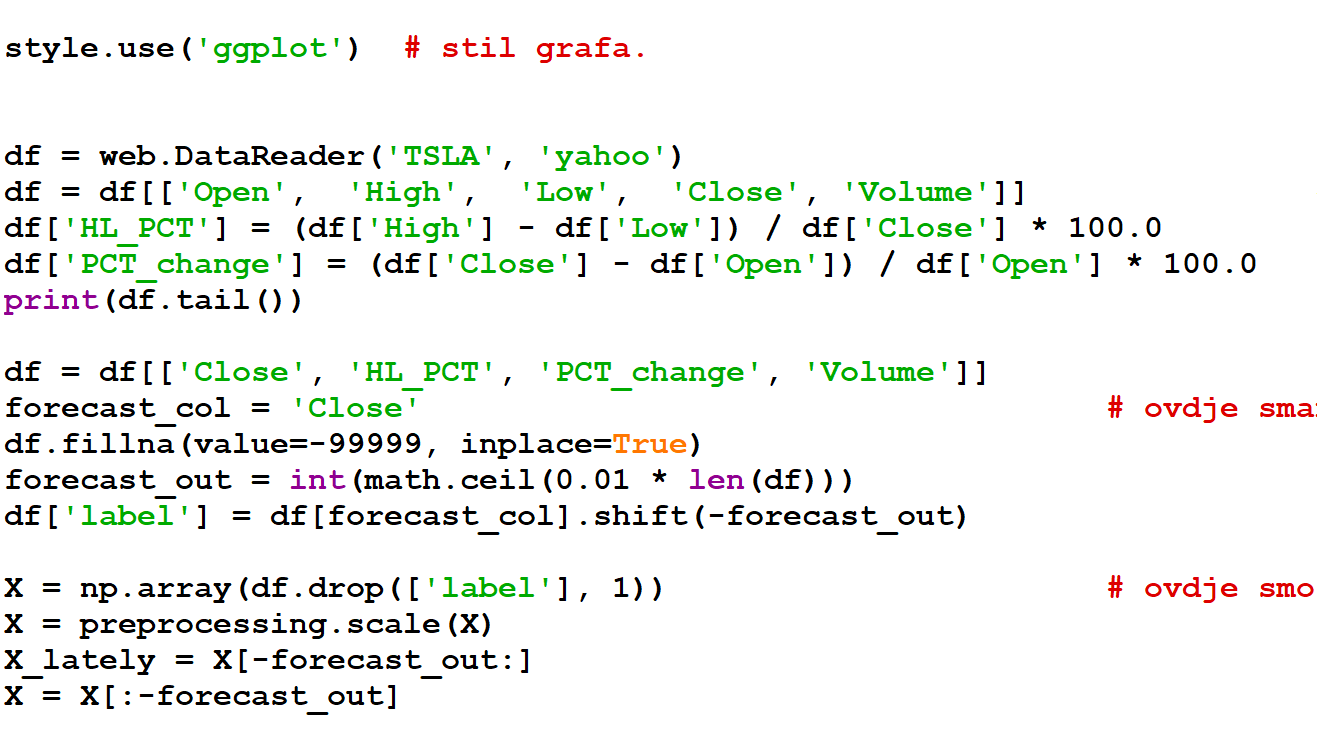
Ovaj dio je najkompleksniji jer ovdje se događa predikcija cijene i cijeloukupni proces izdavanja grafa. Da bi razumjeli 3. dio moramo prvo razumjeti regresiju.



Regresija je oblik takozvanog nadziranog učenja u kojem mi stroju pokažemo podatke i njihove ishode kako bi ono naučilo te s obzirom na to pretpostavilo kretanje u budućnosti. Ovakav algoritam može se koristi na svakakvim polima znanosti, no ni on nije u potpunosti točno jer uvijek postoji odstupanje od pravila.

Strojno učenje grana je umjetne inteligencije koja se bave oblikovanjem algoritama koji svoju učinkovitost poboljšavaju na temelju empirijskih podataka. Strojno učenje jedno je od danas najaktivnijih i najuzbudljivijih područja računarske znanosti, ponajviše zbog brojnih mogućnosti primjene koje se protežu od raspoznavanja uzoraka i dubinske analize podataka do robotike, računalnog vida, bioinformatike i računalne lingvistike. Kolegij obuhvaća dva osnovna pristupa strojnom učenju: nadzirano učenje (klasifikacija i regresija) i nenadzirano učenje (grupiranje i smanjenje dimenzionalnosti).

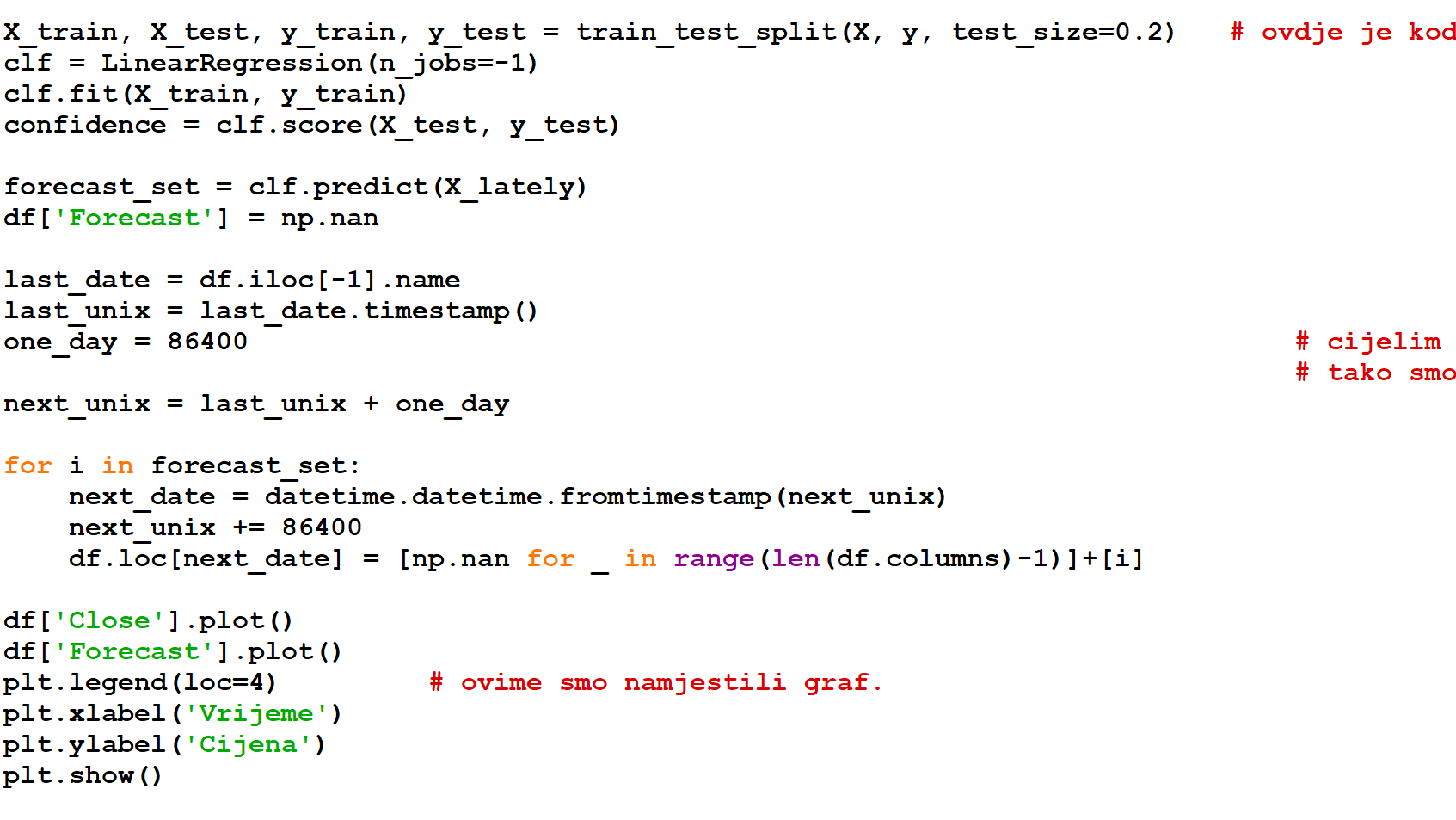


Dakle dataframe je nešto kao tablica i pomoću njega možemo spakirati puno informacija na jedno mjesto.

U ovom djelu smo definirali promjenu u postotku. Podatke smo uzeli s yahoo financa te smo ih pohranili za koristiti.

U drugom smanjujemo količinu podataka te nepotrebene podatke mičemo tako da su veliki negativan broj te ih onda stroj smatra kao nevažnim.To je važno jer određuje vrijeme koje je prihvatljivo za ovaj kod. Jer kada bi uzeli svaku sekundu cijene bilo bi nemoguće napraviti predikciju(barem ne s mojim računalom)

U trečem djelu smo samo podatake pretvorili u array kako bi ih lakše mogili koristit za predikciju.

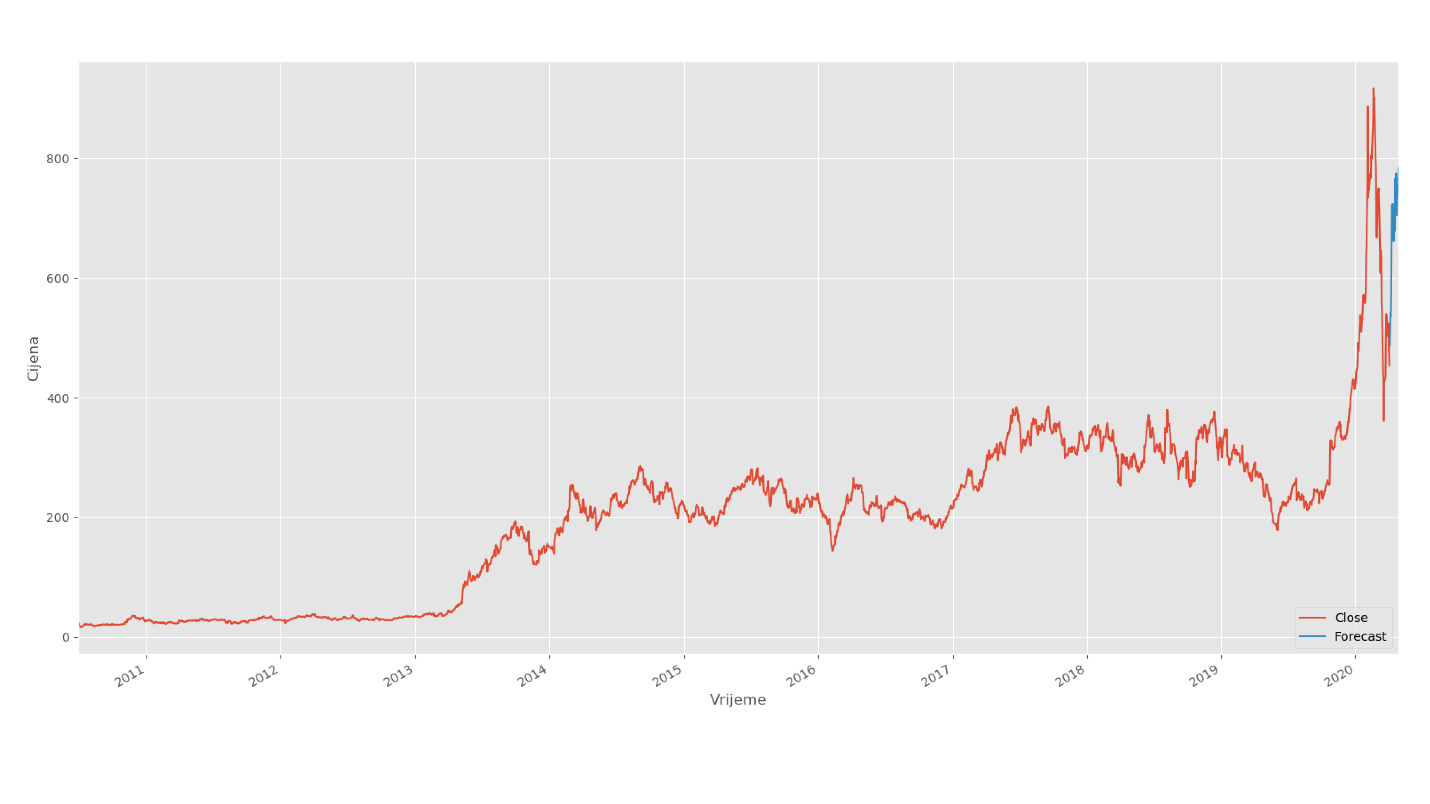


U prvom djelu model počinje trenirati s podacima iz prošlosti koju su mu dane u onom labelu(arrayu). Korisit regresijski algoritam(uz pomoć biblioteke) i onda testira i traži povezanost. Nadalje stvaramo *df Forcast* koji će nam biti cijene u budućnosti.

U drugom i trećem djelu definiramo vrijeme i kako će predvidjeti cijenu. Dajemo danu vrijednost u sekundama i onda pomoću for petlje radimo definiramo kolkiko će vremena u budućnost predviđati.

U četvrtom djelu stavramo graf i dva djela(close(cijene do sada) i forcast(cijene u budućnosti)).

Rezultat tog koda je sljedeći(ovdje je primjer cijena dionice Tesle):

Crveni dio prikazuje cijenu do sada, a plavi pokazuje cijenu u budućnosti.

**Tehničke informacije:**

Zadatak sam ispogramirao u Pythonu i koristio sam se s mnogo biblioteka. Ovdje je lista svih biblioteka koje sam koristio.

Koristio sam dosta modula pa ću ih samo navesti ovdje sve te objasniti njihove mogućnosti:

pandas - tablično slaganje podataka(data frame).

datetime - vremenski periodi, u našem slučaju možemo točno odrediti period od kojega gledamo cijenu dionice.

matplotlib.pyplot- grafovi.

numpy-modul koji nam je pomogao okom matmeatike i pohranjivanja podataka u obliku arraya.

sklearn-najpoznatiji modul za strojno učenje, on nam je dopustio naparvti model na osnovu učenja.

Tkinter- biblioteka za GUI.