**TEHNIČKO VELEUČILIŠTE U ZAGREBU**

**POLITEHNIČKI SPECIJALISTIČKI DIPLOMSKI STRUČNI STUDIJ**

**SPECIJALIZACIJA INFORMATIKA**



Danijel Klobučar

**PREDVIĐANJE TRŽIŠTA DIONICA**

Zagreb, studeni 2023.

SAŽETAK

Tehnike strojnog učenja mogu se koristiti za predviđanje cijena dionica. Regresijski modeli mogu predvidjeti vrijednost dionice u budućnosti, dok klasifikacijski modeli mogu predvidjeti trend cijena dionica.

Iako su tehnike strojnog učenja korisne, postoji nekoliko izazova u njihovoj primjeni. Tržište dionica je podložno promjenama, a podaci o cijenama dionica mogu biti nepouzdani.

U radu je korišten Google Tensorflow i LSTM neuronskih mreža za predviđanje buduće cijene dionica TSLA.

SADRŽAJ

[1 Uvod 1](#_Toc148948222)

[2 Tehnologije 2](#_Toc148948223)

[2.1 Google TensorFlow i LSTM 2](#_Toc148948224)

[2.2 Ostali alati i biblioteke 3](#_Toc148948225)

[2.2.1 Python: 3](#_Toc148948226)

[2.2.2 Google Colaboratory (Colab): 3](#_Toc148948227)

[2.2.3 Google TensorFlow i Keras: 3](#_Toc148948228)

[2.2.4 Pandas: 3](#_Toc148948229)

[2.2.5 NumPy: 4](#_Toc148948230)

[2.2.6 scikit-learn: 4](#_Toc148948231)

[2.2.7 yahoo\_fin: 4](#_Toc148948232)

[2.2.8 Matplotlib.pyplot: 4](#_Toc148948233)

[3 Postupak 5](#_Toc148948234)

[3.1 Instalacija potrebnih biblioteka u Pythonu 5](#_Toc148948235)

[3.2 Importiranje potrebnih biblioteka 5](#_Toc148948236)

[3.3 Postavljanje neuronske mreže 5](#_Toc148948237)

[3.4 Učitavanje podataka 6](#_Toc148948238)

[3.5 Skaliranje podataka 9](#_Toc148948239)

[3.6 Predobrada podataka 10](#_Toc148948240)

[3.7 Model strojnog učenja 11](#_Toc148948241)

[3.8 Prognoza cijena dionica TSLA 13](#_Toc148948242)

[4 Zaključak 19](#_Toc148948243)

[5 Literatura 20](#_Toc148948244)

*Popis slika:*

[Slika 1. Fluktuacija cijene dionica TSLA 9](#_Toc149063787)

[Slika 2. Usporedba stvarne i predviđene vrijednosti TSLA. 18](#_Toc149063788)

*Popis tablica:*

[Tablica 1. Podaci koji su pročitani za TSLA od datuma 12.10.2020 do 18.10.2023. 7](#_Toc149063798)

[Tablica 2. Tablica s stupcima „close“ i „date“. 8](#_Toc149063799)

[Tablica 3. Rezultati skaliranja podataka 10](#_Toc149063800)

[Tablica 4. Usporedba stvarne i predviđene cijene dionica 15](#_Toc149063801)

# Uvod

U svijetu financija, predviđanje cijena dionica oduvijek je bio izazov koji zaokuplja maštu investitora, istraživača i znanstvenika podataka. Sposobnost predviđanja budućih kretanja cijena potencijalno bi mogla dovesti do značajnih dobitaka, ali nije tajna da je tržište dionica notorno nepredvidljivo.

Tehnike strojnog učenja mogu se koristiti za pomoć u predviđanju cijena dionica. Strojno učenje je grana umjetne inteligencije koja omogućuje računalima da uče iz podataka bez eksplicitnog programiranja. Strojno učenje se može koristiti za prepoznavanje obrazaca u podacima o cijenama dionica, što može dovesti do boljih predviđanja.

Glavni cilj ovog seminarskog rada je razviti model predviđanja koji bi mogao predvidjeti cijene dionica TSLA korporacije TESLA Inc. za naredna tri dana. Yahoo Finanace je korišten za povijesne podatke o dionicama. S ovim podacima napravljena je analiza podataka, predobrade, odabira modela i evaluacije.

Pristup strategiji za rješavanje problema je uključivao nekoliko ključnih koraka: prikupljanje podataka iz pouzdanog izvora, identificiranje relevantnih značajki za obuku, odabir optimalnog modela i fino podešavanje njegovih parametara kako bi se postigla najveća ocjena točnosti. Ova strategija osmišljena je za učinkovito rješavanje zamršenog zadatka predviđanja.

U radu je korišten Google Tensorflow i LSTM neuronskih mreža za predviđanje burzovnog tržišta, a to je ujedno najraširenija tehnika strojnog učenja za predviđanje tržišnih trendova.

# Tehnologije

## Google TensorFlow i LSTM

Googleov TensorFlow je *open-source* okvir za strojno učenje koji je razvila Google Brain ekipa. To je jedan od najpopularnijih i široko korištenih okvira za razvoj i treniranje modela strojnog učenja i dubokog učenja. TensorFlow pruža sveobuhvatnu platformu za različite zadatke u umjetnoj inteligenciji i strojnom učenju, uključujući neuronske mreže, obradu prirodnog jezika, računalni vid i druge. *[12]*

LSTM (*Long Short-Term Memory*) dugoročna kratkoročna memorija je vrsta arhitekture rekurentnih neuronskih mreža (RNN) koja je posebno pogodna za zadatke predviđanja sekvenci. Mreže LSTM su vrsta modela dubokog učenja osmišljena za rad s sekvencijalnim podacima. Imaju sposobnost hvatanja dugoročnih zavisnosti i često se koriste za zadatke kao što su prepoznavanje govora, obrada prirodnog jezika i prognoza vremena. *[11]*

TensorFlow podržava LSTM modele kroz svoju biblioteku Keras. Keras je visokoslojena API za TensorFlow koja olakšava izgradnju i treniranje LSTM modela.

Najčešći primjeri u kojima se TensorFlow i LSTM koriste su:

* Vremenska prognoza: LSTM modeli mogu se koristiti za predviđanje budućih vrijednosti vremenskih varijabli, kao što su temperatura, tlak i vlažnost.
* Prepoznavanje govora: LSTM modeli mogu se koristiti za prepoznavanje govora iz audio zapisa.
* Obrada prirodnog jezika: LSTM modeli mogu se koristiti za razumijevanje i generiranje prirodnog jezika.

Prednosti korištenja TensorFlowa za izgradnju LSTM modela:

* Brzina: TensorFlow je optimiziran za brzu obradu velikih količina podataka.
* Jednostavnost korištenja: Keras API olakšava izgradnju i treniranje LSTM modela.
* Raspon mogućnosti: TensorFlow podržava širok raspon LSTM arhitektura i mogućnosti. *[1]*

## Ostali alati i biblioteke

U seminarskom radu korišteni su sljedeći alati i biblioteke:

### Python:

* Python je visoko-razinski programski jezik poznat po svojoj jednostavnosti i čitljivosti.
* Široko se koristi u razvoju web aplikacija, analizi podataka, umjetnoj inteligenciji, strojnom učenju, automatizaciji i mnogim drugim područjima.
* Python ima bogatu ekosustavu biblioteka i okvira koji ga čine izuzetno svestranim. *[2]*

### Google Colaboratory (Colab):

* Colab je besplatna platforma za rad s Pythonom koju je razvio Google.
* Omogućuje izradu i izvođenje Python bilježnica na *cloudu*, što je posebno korisno za razvoj i dijeljenje projekata.
* Pruža besplatan pristup GPU-ima, što ga čini popularnim alatom za strojno učenje i duboko učenje. *[3]*

### Google TensorFlow i Keras:

* TensorFlow je open-source okvir za strojno učenje razvijen od strane Google Brain tima.
* Keras je API visoke razine TensorFlow platforme. Pruža pristupačno, visokoproduktivno sučelje za rješavanje problema strojnog učenja (ML), s fokusom na moderno duboko učenje. Keras pokriva svaki korak tijeka rada strojnog učenja, od obrade podataka do podešavanja hiperparametara do implementacije. Razvijen je s fokusom na omogućavanje brzog eksperimentiranja.
* TensorFlow se koristi za izradu i treniranje modela strojnog učenja, uključujući duboke neuronske mreže. *[4]*

### Pandas:

* Pandas je popularna biblioteka za analizu i manipulaciju podataka u Pythonu.
* Pruža strukture podataka poput *DataFrame*-a, koje su korisne za obradu i analizu podataka.
* Često se koristi za čišćenje, transformaciju i pripremu podataka prije analize ili strojnog učenja. *[5]*

### NumPy:

* NumPy je biblioteka za numeričke operacije u Pythonu.
* Omogućuje rad s višedimenzionalnim nizovima podataka i pruža širok raspon matematičkih funkcija.
* Ključna je komponenta za znanstveno i inženjersko računanje u Pythonu. *[6]*

### scikit-learn:

* Scikit-learn je biblioteka za strojno učenje u Pythonu.
* Pruža jednostavan pristup različitim algoritmima za klasifikaciju, regresiju, klasteriranje i druge zadatke strojnog učenja.
* Koristi se za izgradnju i evaluaciju modela te za optimizaciju hiperparametara. *[7]*

### yahoo\_fin:

* Python knjižnica za preuzimanje financijskih podataka s Yahoo Financea.
* koristi se za dobivanje informacija o dionicama, indeksima i drugim financijskim podacima. *[8]*

### Matplotlib.pyplot:

* Matplotlib je biblioteka za crtanje grafova i vizualizaciju podataka u Pythonu.
* pyplot je sučelje za Matplotlib koje omogućuje jednostavno stvaranje grafičkih prikaza podataka.
* Koristi se za stvaranje grafova, dijagrama i vizualizacija u raznim analitičkim i istraživačkim projektima. *[9]*

# Postupak

## Instalacija potrebnih biblioteka u Pythonu

Prije nego što počnemo s radom potrebno je prvo instalirati potrebne module. *[12]*



## Importiranje potrebnih biblioteka

Nakon što su moduli uspješno instalirani, potrebno je svaki modul importirati.



## Postavljanje neuronske mreže

Postavke neuronske mreže:

* **'*N\_STEPS*'** je broj dana u našem prozoru, odnosno koliko duboko će ići neuronska mreža. Postavljeno je na 7, što odgovara tjednu vremenskom periodu, odnosno analizirati će se podaci u tjednim intervalima.
* Niz ***'LOOKUP\_STEPS*'** sadrži listu brojeva koji odgovaraju danima za predviđanje cijena dionica u budućnosti. Kao primjer, imamo tri dana:
  + jedan dan — sutra
  + dva dana — prekosutra
  + tri dana — za tri dana
  + Za predviđanje od četiri ili pet dana, dodajte odgovarajuće brojeve [1, 2, 3, 4, 5].
* ***'STOCK'*** sadrži simbol dionice koja će se analizirati. U ovom slučaju, simbol je 'TSLA', što označava dionicu kompanije Tesla, Inc. (TSLA).
* ***'date\_now'*** nam omogućuje da vidimo koja je cijena dionice upravo u tom trenutku, a koristi datum u formatu 'YYYY-MM-DD'
* **'date\_3\_years\_back'** početna točka za analizu podataka koja se proteže unatrag 3 godine, odnosno 1104 dana te nam omogućuje da vidimo kolika je cijena dionice tada bila za traženu kompaniju. *[12]*



## Učitavanje podataka

Podaci se učitavaju iz Yahoo Financija. U ovom primjeru dobivamo 1104 stupaca što predstavlja zadnje 3 godine s intervalom od 1 dana.



Za export podataka u excel tablicu koristimo *pd* iz *pandas* i sljedeći kod.



Tablica 1. Podaci koji su pročitani za TSLA od datuma 12.10.2020 do 18.10.2023.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **open** | **high** | **low** | **close** | **adjclose** | **volume** | **ticker** |
| **2020-10-12 00:00:00** | 147,3333282 | 149,5800018 | 146,1933289 | 147,4333344 | 147,4333344 | 116373300 | TSLA |
| **2020-10-13 00:00:00** | 147,7833405 | 149,6300049 | 145,5333405 | 148,8833313 | 148,8833313 | 103391100 | TSLA |
| **2020-10-14 00:00:00** | 149,9266663 | 155,3000031 | 149,1166687 | 153,7666626 | 153,7666626 | 143639100 | TSLA |
| **2020-10-15 00:00:00** | 150,1033325 | 152,1900024 | 147,5 | 149,6266632 | 149,6266632 | 107017200 | TSLA |
| **.....** | ..... | ..... | ..... | ..... | ..... | ..... | TSLA |
| **2023-10-13 00:00:00** | 258,8999939 | 259,6000061 | 250,2200012 | 251,1199951 | 251,1199951 | 102073800 | TSLA |
| **2023-10-16 00:00:00** | 250,0500031 | 255,3999939 | 248,4799957 | 253,9199982 | 253,9199982 | 88917200 | TSLA |
| **2023-10-17 00:00:00** | 250,1000061 | 257,1799927 | 247,0800018 | 254,8500061 | 254,8500061 | 93562900 | TSLA |
| **2023-10-18 00:00:00** | 252,6999969 | 254,6300049 | 242,0800018 | 242,6799927 | 242,6799927 | 121459300 | TSLA |

Prema zadanim postavkama dohvaćeni su svi stupci. Zbog zatvorenosti burze subotom i nedjeljom nedostaju datumi 14. i 15.10.2023. na kraju tablice. Stupac „close“ prikazuje tržišnu cijenu na zatvaranju, dok stupac „open“ prikazuje tržišnu cijenu na otvaranju. Nije potrebno uključiti sve te stupce u model strojnog učenja, stoga ćemo zadržati samo stupac „close“ i dodati stupac date.

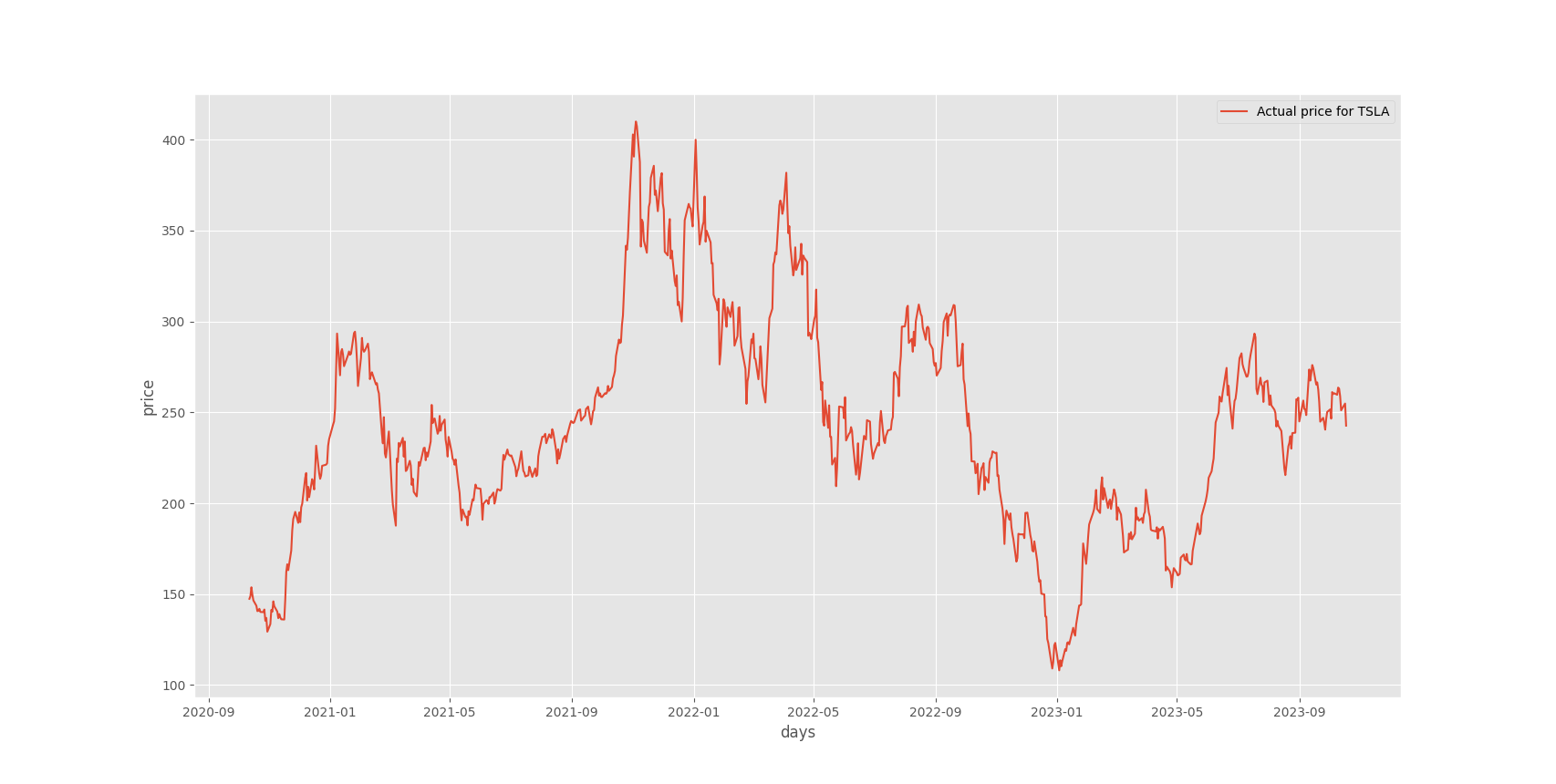


Tablica 2. Tablica s stupcima „close“ i „date“.

|  |  |  |
| --- | --- | --- |
|  | **close** | **date** |
| **2020-10-12 00:00:00** | 147,4333 | 2020-10-12 00:00:00 |
| **2020-10-13 00:00:00** | 148,8833 | 2020-10-13 00:00:00 |
| **2020-10-14 00:00:00** | 153,7667 | 2020-10-14 00:00:00 |
| **2020-10-15 00:00:00** | 149,6267 | 2020-10-15 00:00:00 |
| **2020-10-16 00:00:00** | 146,5567 | 2020-10-16 00:00:00 |
| **.....** | **.....** | **.....** |
| **2023-10-12 00:00:00** | 258,87 | 2023-10-12 00:00:00 |
| **2023-10-13 00:00:00** | 251,12 | 2023-10-13 00:00:00 |
| **2023-10-16 00:00:00** | 253,92 | 2023-10-16 00:00:00 |
| **2023-10-17 00:00:00** | 254,85 | 2023-10-17 00:00:00 |
| **2023-10-18 00:00:00** | 242,68 | 2023-10-18 00:00:00 |

Dobivene podatke prikazat ćemo na dijagramu pomoću sljedećeg koda.





Slika 1. Fluktuacija cijene dionica TSLA

Grafikon prikazuje fluktuaciju cijene dionica TSLA od zadnjih 3 godine. Dijagram se temelji na stupcu "close", što je tržišna cijena na zatvaranju.

## Skaliranje podataka

Da bi se poboljšale performanse LSTM neuronske mreže, potrebno je skalirati podatke iz stupca Close.

Skaliranje podataka je važna tehnika prije treniranja modela strojnog učenja, jer omogućuje modelu da bolje uči iz podataka i da daje bolje predviđanja.

Za skaliranje podataka koristimo klasu *MinMaxScaler* iz biblioteke *sklearn.preprocessing.* Ova klasa skalira sve vrijednosti u zadanom intervalu, u ovom slučaju između 1 i 0.

Za to se koristi metoda *fit\_transform* koja sve vrijednosti skalira između 1 i 0 (gdje 1 predstavlja najveću vrijednost iz stupca „close“, a 0 najmanju).



Tablica 3. Rezultati skaliranja podataka

|  |  |  |  |
| --- | --- | --- | --- |
|  | **close** | **date** | **scaled\_close** |
| **2020-10-12 00:00:00** | 147,4333 | 2020-10-12 00:00:00 | 0,130298922 |
| **2020-10-13 00:00:00** | 148,8833 | 2020-10-13 00:00:00 | 0,135102304 |
| **2020-10-14 00:00:00** | 153,7667 | 2020-10-14 00:00:00 | 0,151279238 |
| **2020-10-15 00:00:00** | 149,6267 | 2020-10-15 00:00:00 | 0,137564728 |
| **2020-10-16 00:00:00** | 146,5567 | 2020-10-16 00:00:00 | 0,127394813 |
| **.....** | **.....** | **.....** | **.....** |
| **2023-10-12 00:00:00** | 258,87 | 2023-10-12 00:00:00 | 0,499453391 |
| **2023-10-13 00:00:00** | 251,12 | 2023-10-13 00:00:00 | 0,473780089 |
| **2023-10-16 00:00:00** | 253,92 | 2023-10-16 00:00:00 | 0,483055615 |
| **2023-10-17 00:00:00** | 254,85 | 2023-10-17 00:00:00 | 0,486136437 |
| **2023-10-18 00:00:00** | 242,68 | 2023-10-18 00:00:00 | 0,445821026 |

Maksimalna cijena dionice TSLA dana 4.11.2021 iznosila je $409.97, što je 1 na skali, dok je najmanja cijena dionice TSLA dana 3.1.2023 iznosila $108.09, što je 0 na skali od 0 do 1.

## Predobrada podataka

U tablici se trenutno nalaze podaci koji su preliminarno učitani i skalirani za model strojnog učenja.

Sada je potrebno pripremiti te podatke za sljedeći postupak, jer je početni cilj predvidjeti cijenu dionica za naredna tri dana. To znači da se stupci moraju pomaknuti za broj dana za koje se želi predvidjeti cijena i u skladu s tim pripremiti podaci za model.

U ovom slučaju, stupac „close“ je ciljni stupac i pomaknuti će se i spremiti rezultati u stupac „future“. Programski kôd za to je sljedeći:



Ova skripta treba još izračunati posljednju sekvencu, koja sadrži konačni prozor podataka. Ta se sekvenca koristi za predviđanje budućih cijena.

LSTM-ovi X i Y nizovi su posljednje komponente skripte. Polje X sadrži niz na određenom koraku, a polje Y sadrži ciljnu cijenu na tom koraku.

## Model strojnog učenja

Sad je potrebno uspostaviti model strojnog učenja i njegove postavke. U ovom pristupu koristi se LSTM neuronska mreža i sekvencijalni model. Model ima nekoliko LSTM i Dropout slojeva, kao i Dense sloj.

Parametar „x\_train“ je niz podataka, pripremljenih za model po kojem će se vršiti obuka. Parametar „y\_train“ pruža odgovor na obučavanje na temelju ciljane kolone 'close'. LSTM slojevi 1 ima 60 neurona, dok sloj 2 ima 120 neurona. Između LSTM slojeva postoji skriveni sloj s 30% izbačenih neurona. Posljednji korak modela je gusto sloj s 20 neurona i jednim neuronom. Na samome kraju imamo Dense sloj s 20 neurona kao i samo jedan neuron što je rezultat na zadnjem koraku modela.



Model će trenirati tijekom 80 epoha (sekvenci računanja) s veličinom serije u svakoj epohi jednakom 8. U ovom primjeru, koristi se *'mean\_squared\_error*' kao parametar funkcije gubitka i '*adam*' optimizator, koji je najpopularniji u zadacima predviđanja cijena dionica.

Dobiveni rezultati podučavanja modela će se sažeti i vratiti programu za predviđanja vrijednosti dionica na burzi.

Ispis modela procesa obuke:



Prema ispisu gubitak je oko 0,0035 ili 3%.

## Prognoza cijena dionica TSLA

Sada imamo sve potrebne podatke i možemo prognozirati cijene TSLA dionica za naredna 3 dana.

Sljedeći programski kôd inicijalizira niz predviđenih cijena za sljedeća tri dana od dana pokretanja skripte. U petlji, na svakom koraku (za svaki dan iz niza dana [1, 2, 3]), podaci se pripremaju i model se trenira.



Za pokretanje ove petlje može proći duže vrijeme.

Nakon završetka prethodnog procesa, dobiva se statistika izvedenog procesa.

Layer (type) Output Shape Param #

=================================================================

lstm\_4 (LSTM) (None, 7, 60) 14880

dropout\_4 (Dropout) (None, 7, 60) 0

lstm\_5 (LSTM) (None, 120) 86880

dropout\_5 (Dropout) (None, 120) 0

dense\_4 (Dense) (None, 20) 2420

dense\_5 (Dense) (None, 1) 21

=================================================================

Total params: 104201 (407.04 KB)

Trainable params: 104201 (407.04 KB)

Non-trainable params: 0 (0.00 Byte)

Krajnji rezultat se dobiva provjerom ispravnosti kalkulacija i izvlačenjem rezultata iz niza.

**if** **bool(**predictions**)** **==** **True** **and** **len(**predictions**)** **>** 0**:**

predictions\_list **=** **[str(**d**)+**'$' **for** d **in** predictions**]**

predictions\_str **=** ', '**.**join**(**predictions\_list**)**

message **=** f'{STOCK} prediction for upcoming 3 days ({predictions\_str})'

**print(**message**)**

Dobiveni predviđeni rezultati na dan 20.10.2023 za sljedeća tri dana su:

TSLA prediction for upcoming 3 days (225.57$, 228.67$, 221.99$)

Mogućnost određivanja vrijednosti cijena za sve prethodne dane (unazad 3 godine s kojima raspolažemo) omogućuje nam da procijenimo točnost pretpostavki koje je model napravio.

# Execute model for the whole history range

Last\_3\_y\_df **=** init\_df**.**copy**()**

y\_predicted **=** model**.**predict**(**x\_train**)**

y\_predicted\_transformed **=** np**.**squeeze**(**scaler**.**inverse\_transform**(**y\_predicted**))**

first\_seq **=** scaler**.**inverse\_transform**(**np**.**expand\_dims**(**y\_train**[:**6**],** axis**=**1**))**

last\_seq **=** scaler**.**inverse\_transform**(**np**.**expand\_dims**(**y\_train**[-**3**:],** axis**=**1**))**

y\_predicted\_transformed **=** np**.**append**(**first\_seq**,** y\_predicted\_transformed**)**

y\_predicted\_transformed **=** np**.**append**(**y\_predicted\_transformed**,** last\_seq**)**

Last\_3\_y\_df**[**f'predicted\_close'**]** **=** y\_predicted\_transformed

Tablica 4. Usporedba stvarne i predviđene cijene dionica

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **close** | **date** | **scaled\_close** | **predicted\_close** | **Razlika** |
| **2020-10-12 00:00:00** | 147,4333344 | 2020-10-12 00:00:00 | 0,130298922 | 140,2100067 | 4,90% |
| **2020-10-13 00:00:00** | 148,8833313 | 2020-10-13 00:00:00 | 0,135102304 | 140,093338 | 5,90% |
| **2020-10-14 00:00:00** | 153,7666626 | 2020-10-14 00:00:00 | 0,151279238 | 141,5599976 | 7,94% |
| **2020-10-15 00:00:00** | 149,6266632 | 2020-10-15 00:00:00 | 0,137564728 | 135,3399963 | 9,55% |
| **2020-10-16 00:00:00** | 146,5566711 | 2020-10-16 00:00:00 | 0,127394813 | 136,9433289 | 6,56% |
| **2020-10-19 00:00:00** | 143,6100006 | 2020-10-19 00:00:00 | 0,117633424 | 129,3466644 | 9,93% |
| **2020-10-20 00:00:00** | 140,6466675 | 2020-10-20 00:00:00 | 0,107816837 | 150,8685455 | 7,27% |
| **2020-10-21 00:00:00** | 140,8800049 | 2020-10-21 00:00:00 | 0,108589811 | 150,499649 | 6,83% |
| **2020-10-22 00:00:00** | 141,9299927 | 2020-10-22 00:00:00 | 0,112068089 | 151,9382172 | 7,05% |
| **2020-10-23 00:00:00** | 140,2100067 | 2020-10-23 00:00:00 | 0,106370318 | 151,4535675 | 8,02% |
| **2020-10-26 00:00:00** | 140,093338 | 2020-10-26 00:00:00 | 0,105983832 | 151,0520935 | 7,82% |
| **2020-10-27 00:00:00** | 141,5599976 | 2020-10-27 00:00:00 | 0,110842412 | 152,1712189 | 7,50% |
| **2020-10-28 00:00:00** | 135,3399963 | 2020-10-28 00:00:00 | 0,090237512 | 147,8763733 | 9,26% |
| **......** | **......** | **......** | **......** | **......** |  |
| **2021-05-21 00:00:00** | 193,6266632 | 2021-05-21 00:00:00 | 0,283322834 | 200,9826355 | 3,80% |
| **2021-05-24 00:00:00** | 202,1466675 | 2021-05-24 00:00:00 | 0,311546918 | 207,359848 | 2,58% |
| **2021-05-25 00:00:00** | 201,5633392 | 2021-05-25 00:00:00 | 0,309614536 | 208,9780731 | 3,68% |
| **2021-05-26 00:00:00** | 206,3766632 | 2021-05-26 00:00:00 | 0,325559558 | 212,36203 | 2,90% |
| **.....** | **.....** | **.....** | **.....** | **.....** |  |
| **2023-09-28 00:00:00** | 246,3800049 | 2023-09-28 00:00:00 | 0,458077998 | 247,0254211 | 0,26% |
| **2023-09-29 00:00:00** | 250,2200012 | 2023-09-29 00:00:00 | 0,470798693 | 251,6622314 | 0,58% |
| **2023-10-02 00:00:00** | 251,6000061 | 2023-10-02 00:00:00 | 0,475370213 | 253,8735352 | 0,90% |
| **2023-10-03 00:00:00** | 246,5299988 | 2023-10-03 00:00:00 | 0,45857488 | 250,1608582 | 1,47% |
| **2023-10-04 00:00:00** | 261,1600037 | 2023-10-04 00:00:00 | 0,507039467 | 259,9609985 | 0,46% |
| **2023-10-05 00:00:00** | 260,0499878 | 2023-10-05 00:00:00 | 0,503362335 | 262,2340698 | 0,84% |
| **2023-10-06 00:00:00** | 260,5299988 | 2023-10-06 00:00:00 | 0,504952459 | 262,3969421 | 0,72% |
| **2023-10-09 00:00:00** | 259,6700134 | 2023-10-09 00:00:00 | 0,502103599 | 261,34198 | 0,64% |
| **2023-10-10 00:00:00** | 263,6199951 | 2023-10-10 00:00:00 | 0,515188642 | 263,8675537 | 0,09% |
| **2023-10-11 00:00:00** | 262,9899902 | 2023-10-11 00:00:00 | 0,513101634 | 264,1273193 | 0,43% |
| **2023-10-12 00:00:00** | 258,8699951 | 2023-10-12 00:00:00 | 0,499453391 | 260,7994995 | 0,75% |
| **2023-10-13 00:00:00** | 251,1199951 | 2023-10-13 00:00:00 | 0,473780089 | 253,7746735 | 1,06% |
| **2023-10-16 00:00:00** | 253,9199982 | 2023-10-16 00:00:00 | 0,483055615 | 254,0262909 | 0,04% |
| **2023-10-17 00:00:00** | 254,8500061 | 2023-10-17 00:00:00 | 0,486136437 | 254,8500061 | 0,00% |
| **2023-10-18 00:00:00** | 242,6799927 | 2023-10-18 00:00:00 | 0,445821026 | 242,6799927 | 0,00% |
| **2023-10-19 00:00:00** | 220,1100006 | 2023-10-19 00:00:00 | 0,371053769 | 220,1100006 | 0,00% |

Tablica prikazuje da je model u početku imao višu grešku preciznosti u predviđanju TSLA dionice od korporacije TESLA Inc. na američkoj burzi NASDAQ, ali kasnije je ta greška bila sve manja, odnosno model je bio precizniji.

Za lakši pregled podataka, grafikon se može ponovno napraviti kako bi se usporedile stvarne i predviđene vrijednosti.

# Execute model for the whole history range

last\_3\_y\_df **=** init\_df**.**copy**()**

y\_predicted **=** model**.**predict**(**x\_train**)**

y\_predicted\_transformed **=** np**.**squeeze**(**scaler**.**inverse\_transform**(**y\_predicted**))**

first\_seq **=** scaler**.**inverse\_transform**(**np**.**expand\_dims**(**y\_train**[:**6**],** axis**=**1**))**

last\_seq **=** scaler**.**inverse\_transform**(**np**.**expand\_dims**(**y\_train**[-**3**:],** axis**=**1**))**

y\_predicted\_transformed **=** np**.**append**(**first\_seq**,** y\_predicted\_transformed**)**

y\_predicted\_transformed **=** np**.**append**(**y\_predicted\_transformed**,** last\_seq**)**

last\_3\_y\_df **[**f'predicted\_close'**]** **=** y\_predicted\_transformed

# Add predicted results to the table

date\_now **=** dt**.**date**.**today**()**

date\_tomorrow **=** dt**.**date**.**today**()** **+** dt**.**timedelta**(**days**=**1**)**

date\_after\_tomorrow **=** dt**.**date**.**today**()** **+** dt**.**timedelta**(**days**=**2**)**

last\_3\_y\_df**.**loc**[**date\_now**]** **=** **[**predictions**[**0**],** f'{date\_now}'**,** 0**,** 0**]**

last\_3\_y\_df**.**loc**[**date\_tomorrow**]** **=** **[**predictions**[**1**],** f'{date\_tomorrow}'**,** 0**,** 0**]**

last\_3\_y\_df**.**loc**[**date\_after\_tomorrow**]** **=** **[**predictions**[**2**],** f'{date\_after\_tomorrow}'**,** 0**,** 0**]**

# Result chart

plt**.**style**.**use**(**style**=**'ggplot'**)**

plt**.**figure**(**figsize**=(**16**,**10**))**

plt**.**plot**(**last\_3\_y\_df **[**'close'**][-**150**:].**head**(**147**))**

plt**.**plot**(**last\_3\_y\_df **[**'predicted\_close'**][-**150**:].**head**(**147**),** linewidth**=**1**,** linestyle**=**'dashed'**)**

plt**.**plot**(**last\_3\_y\_df **[**'close'**][-**150**:].**tail**(**4**))**

plt**.**xlabel**(**'days'**)**

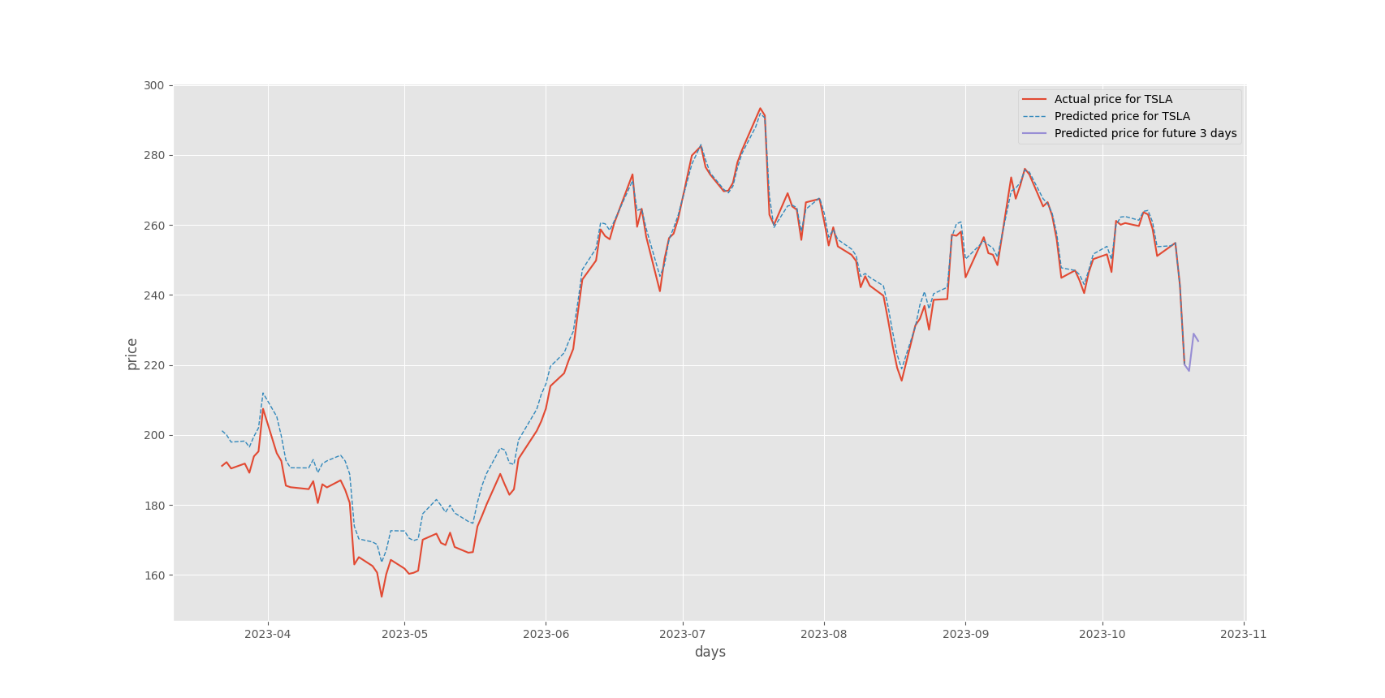
plt**.**ylabel**(**'price'**)**

plt**.**legend**([**f'Actual price for {STOCK}'**,**

f'Predicted price for {STOCK}'**,**

f'Predicted price for future 3 days'**])**

plt**.**show**()**



Slika 2. Usporedba stvarne i predviđene vrijednosti TSLA.

# Zaključak

U svijetu predviđanja tržišta dionica, uspjeh se mjeri sposobnošću korištenja uvida temeljenih na podacima za donošenje informiranih odluka. Iako je model LSTM pokazao hvalevrijedan učinak, ključno je priznati složenost tržišta dionica. Čimbenici poput ekonomskih pokazatelja, geopolitičkih događaja i raspoloženja na tržištu samo su neke od varijabli koje utječu na cijene dionica, pridonoseći njihovoj inherentnoj nepredvidljivosti.

U ovom svijetu, uspjeh se mjeri sposobnošću korištenja uvida temeljenih na podacima za donošenje informiranih odluka. Iako je model LSTM pokazao dobar rezultat, važno je priznati složenost tržišta dionica. Čimbenici poput ekonomskih pokazatelja, geopolitičkih događaja i raspoloženja na tržištu samo su neki od faktora koji utječu na cijene dionica, pridonoseći njihovoj inherentnoj nepredvidljivosti.

U ovom seminarskom radu prikazano je kako je relativno lako dohvatiti povijesne podatke, obraditi ih i na temelju njih istrenirati model koji će dati predviđenu vrijednost na temelju dosadašnjeg ponašanja. Taj model je stekao dobru pouzdanost kada su u pitanju dosadašnja događanja, ali ne i za nova. Naime, kod novih događanja na tržištu dionica, veliku ulogu igraju sami ljudi koji su izuzetno nepredvidljivi. Stoga, nije dovoljno imati podatke o cijeni zatvaranja dionica kako bi se iste moglo predvidjeti.

Umjetna inteligencija, iako vrlo napredna, još uvijek nije dostigla razinu pouzdanosti koja se od nje očekuje u ovom sektoru. Naime, ista se koristi kao alat za savjetovanje, no rijetko se kada dopušta da ista direktno trguje na otvorenom tržištu.

-----------------------

Danijel Klobučar

# Literatura

[1] web stranica, <https://pieriantraining.com/tensorflow-lstm-example-a-beginners-guide/>, (19.10.2023.)

[2] web stranica, <https://en.wikipedia.org/wiki/Python_(programming_language)>, (19.10.2023.)

[3] web stranica, <https://colab.google/>, (18.10.2023.)

[4] web stranica, <https://www.tensorflow.org/guide/keras>, (19.10.2023.)

[5] web stranica, <https://pandas.pydata.org/>, (18.10.2023.)

[6] web stranica, <https://numpy.org/>, (18.10.2023.)

[7] web stranica, <https://scikit-learn.org/stable/>, (18.10.2023.)

[8] web stranica, <https://pypi.org/project/yahoo-fin/>, (18.10.2023.)

[9] web stranica, <https://matplotlib.org/stable/tutorials/pyplot.html>, (18.10.2023.)

[10] web stranica, <https://medium.com/@matthew1992/tensorflow-and-lstm-neural-networks-for-stock-market-forecasting-b47b36d2358d> , (18.10.2023.)

[11] web stranica, <https://www.analyticsvidhya.com/blog/2021/10/machine-learning-for-stock-market-prediction-with-step-by-step-implementation/>, (18.10.2023.)

[12] web stranica,<https://medium.com/@dmytrosazonov/how-to-predict-stock-market-using-google-tensorflow-and-lstm-neural-network-81ccc41a22a8> , (18.10.2023.)