[[1]](#footnote-0)

Analiza berze i predikcije korišćenjem LSTM, BiLSTM i GRU modela

Aleksa Janjić, diplomirani inženjer elektrotehnike i računarstva i Mihailo Radojević, diplomirani inženjer elektrotehnike i računarstva

***Apstrakt*—Ovaj rad istražuje primenu dubokih neuralnih mreža, tačnije LSTM (eng. Long Short-Term Memory), BiLSTM (eng. Bidirectional Long Short-Term Memory) i GRU (eng. Gated Recurrent Unit) modela, u analizi berzanskog tržišta i predikciji cena finansijskih instrumenata. Uvođenjem ovih složenih neuralnih arhitektura, istražuje se sposobnost modela da identifikuje i modeluje temporalne zavisnosti u vremenskim serijama podataka s ciljem poboljšanja tačnosti predikcija cena akcija. Metodologija rada obuhvata prikupljanje i obradu velikog skupa podataka o berzanskim kretanjima, implementaciju LSTM, BiLSTM i GRU modela kroz duboko učenje, i evaluaciju performansi ovih modela na test skupu podataka. Rezultati istraživanja pružaju uvid u efikasnost različitih neuralnih arhitektura u predviđanju berzanskih kretanja i ukazuju na potencijalne prednosti i ograničenja svakog modela. Ova analiza ima značajnu primenu u donošenju investicionih odluka i upravljanju rizicima na finansijskim tržištima.**

***Ključne reči*—berza; neuralna mreža; duboko učenje; LSTM; BiLSTM; GRU**

# Uvod

Berzansko tržište predstavlja vitalni deo globalne ekonomije, gde se svakodnevno dešavaju kompleksni procesi formiranja cena finansijskih instrumenata, među kojima je ključna završna cena (eng. Closing Price). Ova cena predstavlja poslednju cenu po kojoj je akcija trgovana tokom jednog dana na berzi i ima izuzetan značaj za investitore, analitičare i trgovce u donošenju odluka.

Analiza i predikcija završne cene su od suštinskog značaja za pravilno razumevanje i tumačenje dinamike berzanskog tržišta. Uticaji kao što su ekonomske vesti, geopolitičke tenzije, sezonske promene i mnogi drugi faktori mogu izazvati značajne fluktuacije u cenama akcija. Stoga je razvoj efikasnih modela za predviđanje završne cene postao ključni fokus istraživanja u oblasti finansijske analize.

U ovom istraživanju istražujemo primenu modela rekurentnih neuralnih mreža za analizu berzanskog tržišta i predikciju završne cene svetskih gigantskih kompanija kao što su Amazon, Apple, Microsoft i Google. Kroz korišćenje dubokog učenja, želimo da identifikujemo obrasce i trendove u podacima kako bismo bolje razumeli ponašanje tržišta i poboljšali prediktivne sposobnosti naših modela.

Ovo istraživanje ne samo da doprinosi unapređenju metodologije finansijske analize, već i pruža korisne uvide u dinamiku berzanskog tržišta i proces formiranja cena akcija. Nadamo se da će rezultati ovog rada biti od koristi investitorima, trgovcima i akademskoj zajednici u oblasti finansija i računarstva.

# Skup podataka

U ovom radu korišćena je *pandas\_datareader* biblioteka u programskom jeziku Python koja omogućava jednostavan pristup raznim izvorima finansijskih podataka, uključujući Yahoo Finance, Google Finance, FRED i mnoge druge. Funkcija *pandas\_datareader.get\_data\_yahoo* je posebno dizajnirana da omogući brzo i jednostavno preuzimanje podataka o akcijama sa Yahoo Finance platforme.

Ovaj skup podataka pruža informacije o cenama akcija tokom određenog perioda. To obično uključuje početnu (eng. Open), završnu (eng. Close), najvišu (eng. High) i najnižu (eng. Low) cenu akcija za svaki dan u periodu koji je odabran. Ovaj skup podataka pruža informacije o cenama akcija tokom određenog perioda. To obično uključuje sve pomenute cene akcija za svaki dan u periodu koji je odabran. Podaci obično sadrže vremenski indeksiranje, gde svaki red predstavlja jedan dan trgovanja, a podaci o cenama akcija mogu biti izraženi u lokalnoj valuti kompanije i vremenske oznake se obično odnose na lokalnu vremensku zonu berze na kojoj se akcija trguje.

U ovom radu učitani su podaci o cenama akcija 4 gigantske kompanije Amazon, Apple, Microsoft i Google u periodu od 01.01.2012. do današnjeg datuma kada je rad napisan - 06.02.2024. Fokusirali smo se na predikciju završne cene akcija i učitani skup podataka podeljen je na trenirajući i testirajući skup podataka u odnosu 90:10. Pre podele na trenirajući i testirajući skup podataka, urađeno je skaliranje svih vrednosti atributa iz skupa podataka na interval od 0 do 1.

# Modeli

Za predikciju cena akcija, kao što je već napisano, korišćeni su modeli rekurentnih neuralnih mreža kao što su LSTM (eng. Long Short-Term Memory), BiLSTM (Bidirectional Long Short-Term Memory) i GRU (eng. Gated Recurent Units).

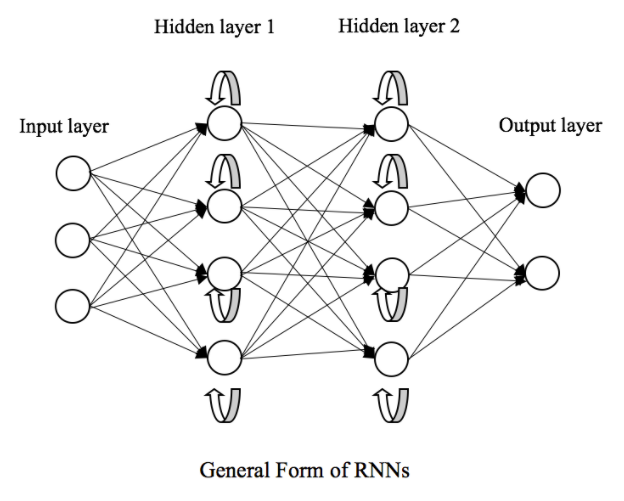
Rekurentne neuralne mreže (eng. Recurrent Neural Networks) su vrsta neuralnih mreža koje su posebno dizajnirane za obradu sekvencijalnih podataka. Za razliku od tradicionalnih neuralnih mreža, koje obrađuju podatke nezavisno od njihovog redosleda, rekurentne neuralne mreže imaju sposobnost da zapamte prethodne informacije i koriste ih u obradi trenutnih podataka u sekvenci. Ovo ih čini izuzetno korisnim za analizu vremenskih serija, prirodne jezike, audio signale, genomske sekvence i druge vrste sekvencijalnih podataka.

Glavna karakteristika rekurentnih neuralnih mreža je unutrašnja petlja koja omogućava prenošenje informacija iz jednog koraka mreže u sledeći. Ovo omogućava rekurentnim neuralnim mrežama da održe stanje ili "memoriju" tokom vremena, što je ključno za obradu sekvencijalnih podataka.

Ova unutrašnja petlja omogućava rekurentnim neuralnim mrežama da prilagode svoj izlaz na osnovu prethodnih koraka u sekvenci.

Međutim, tradicionalne rekurentne neuralne mreže imaju ograničenja u prenošenju informacija na dugim sekvencama podataka. Problemi poput nestajućih i eksplodirajućih gradijenata često se javljaju tokom treniranja rekurentnih neuralnih mreža na dugim sekvencama, što može dovesti do gubitka relevantnih informacija. Zbog toga su razvijene poboljšane verzije rekurentnih neuralnih mreža, kao što su LSTM, BiLSTM i GRU, koje su dizajnirane da prevaziđu ove probleme i omoguće efikasniju obradu dugih sekvenci.

Šematski prikaz rekurentne neuralne mreže je dat na slici 1.



Sl. 1. Generalna forma rekurentne neuralne mreže sa 2 skrivena sloja neurona

**LSTM** modeli su vrsta rekurentnih neuralnih mreža koji su posebno dizajnirani za obradu i modeliranje sekvencijalnih podataka, kao što su vremenske serije, tekstualni podaci i druge vrste sekvencijalnih informacija. LSTM mreže su razvijene kako bi prevazišle problem nestajućih i eksplodirajućih gradijenata koji često nastaju kod tradicionalnih rekurentnih neuralnih mreža tokom treniranja na dugim sekvencama podataka.

Glavna karakteristika LSTM modela su ćelije memorije koje omogućavaju dugoročno pamćenje informacija. Ove ćelije imaju mehanizam kontrole kako bi selektivno zaboravljale ili ažurirale informacije tokom vremena. Ovo omogućava LSTM mrežama da efikasno modeliraju i čuvaju relevantne informacije iz prošlosti, koje su od značaja za predikciju budućih vrednosti.

Ključne komponente LSTM ćelije uključuju:

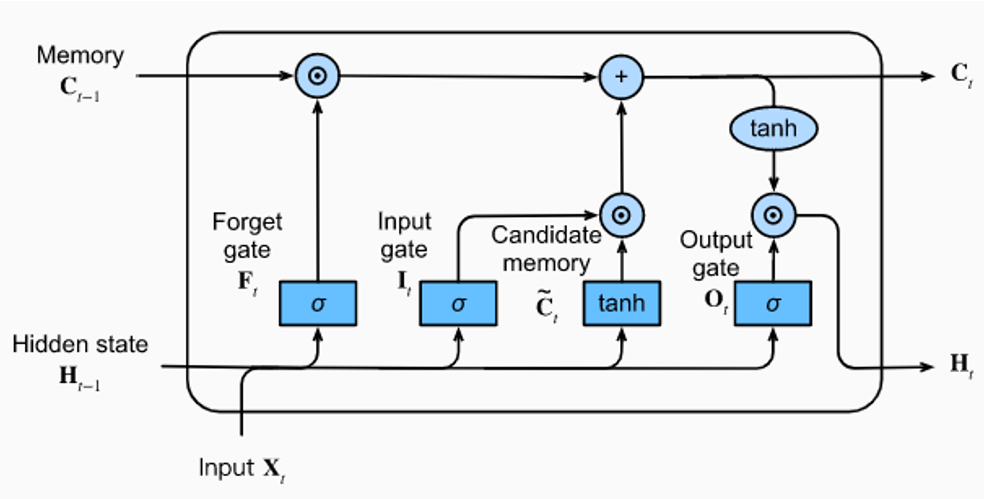
* Stanje ćelije (eng. cell state): Ovo je centralna komponenta koja zadržava informacije iz prošlosti koje su relevantne za trenutni trenutak. Informacije mogu biti dodate, izmenjene ili uklonjene kroz vremenske korake.
* Kapija (eng. gates): LSTM ćelija ima tri kapije koje kontrolišu protok informacija unutar ćelije:

1. Kapija zaboravljanja (eng. Forget gate): Odlučuje koje informacije treba zaboraviti iz stanja ćelije.
2. Ulazna kapija (eng. Input gate): Odlučuje koje nove informacije treba dodati u stanje ćelije.
3. Izlazna kapija (eng. Output gate): Odlučuje koji deo stanja ćelije treba proslediti kao izlaz iz ćelije.

* Stanje skrivenog sloja (eng. hidden state): Ovo je izlaz koji se generiše na osnovu stanja ćelije, ali je modifikovan od strane izlazne kapije. Stanje skrivenog sloja se koristi kao izlaz LSTM ćelije i može se proslediti kao ulaz u naredne vremenske korake ili u druge slojeve mreže.

Prednosti LSTM mreža uključuju sposobnost modeliranja dugoročnih zavisnosti u podacima, stabilnost tokom treniranja na dugim sekvencama podataka i mogućnost efikasnog prenosa informacija kroz vremenske korake. [3]

Šematski prikaz LSTM modela rekurentne neuralne mreže je dat na slici 2.



Sl. 2. Generalna forma LSTM modela rekurentne neuralne mreže

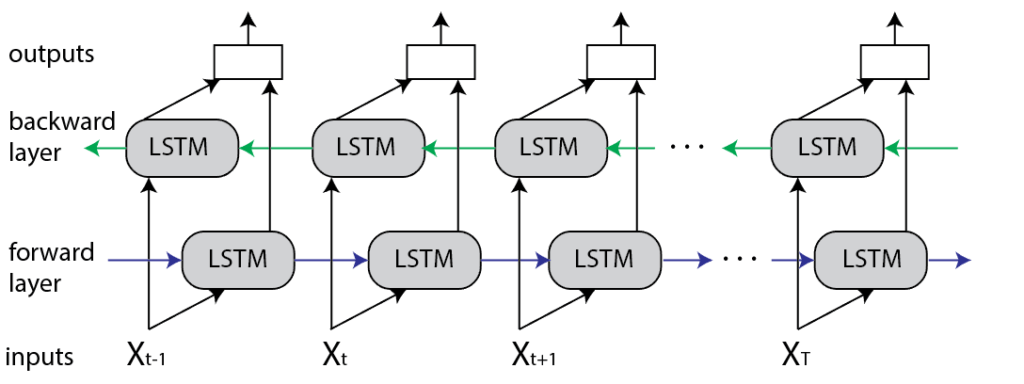
**BiLSTM** model je napredna varijanta LSTM mreže koja dodaje još jedan sloj informacija unutar LSTM arhitekture. Glavna karakteristika BiLSTM modela je sposobnost da istovremeno obrađuje informacije iz prošlosti i budućnosti, što ga čini izuzetno korisnim za analizu sekvencijalnih podataka.

Kod standardnih LSTM mreža, informacije se propagiraju unapred kroz vremenske korake, omogućavajući mreži da zapamti i koristi informacije iz prošlosti. Međutim, BiLSTM modeli proširuju ovu ideju tako što uvode još jedan sloj ćelija koji propagira informacije unazad kroz vremenske korake. Ovo omogućava modelu da istovremeno razmatra kontekst iz prošlosti i budućnosti pri donošenju odluka o trenutnom koraku u sekvenci.

BiLSTM sadrži unaprednu LSTM petlju koja radi na isti način kao i standardna LSTM mreža, obrađujući informacije iz prošlosti kako bi generisala izlaz za trenutni korak. Paralelno sa unaprednom petljom, BiLSTM model koristi drugi set LSTM ćelija u petlji koja propagira informacije unazad kroz vremenske korake. Ovaj deo modela analizira kontekst iz budućnosti kako bi dopunio informacije iz prošlosti. Izlazi iz oba dela BiLSTM-a se kombinuju kako bi se generisao konačni izlaz za trenutni korak u sekvenci. Ova kombinacija omogućava modelu da uzme u obzir širi kontekst informacija prilikom donošenja odluka.

Prednosti BiLSTM modela su brojne, a neke od njih su razmatranje konteksta iz prošlosti i budućnosti, sposobnost modeliranja dugoročnih zavisnosti i efikasna obrada sekvencijalnih podataka. BiLSTM modeli su široko korišćeni u oblastima kao što su mašinsko prevođenje, analiza sentimenta, prepoznavanje entiteta u tekstu, prepoznavanje govora, analiza vremenskih serija i mnoge druge primene gde je važno uzeti u obzir kontekst iz oba smera u sekvenci. [4]

Šematski prikaz BiLSTM modela rekurentne neuralne mreže je dat na slici 3.



Sl. 3. Generalna forma BiLSTM modela rekurentne neuralne mreže

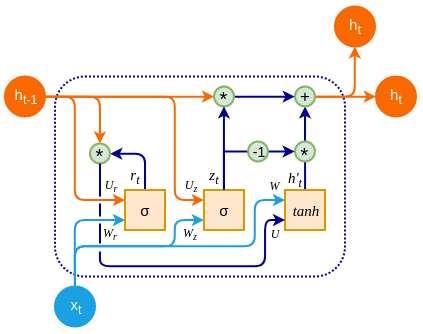
**GRU** je još jedan tip rekurentnih neuralnih mreža koji je dizajniran za efikasnu obradu sekvencijalnih podataka. GRU je sličan LSTM modelu, ali je dizajniran da bude jednostavniji i da zahteva manje parametara za treniranje. Ovaj model je popularan izbor u oblastima gde je potrebna efikasna obrada sekvencijalnih podataka, kao što su prirodni jezik, vremenske serije i druge.

Glavna karakteristika GRU-a su dva ključna mehanizma, poznata kao kapija za ažuriranje (eng. Update gate) i kapija za poništavanje (eng. Reset gate), koji omogućavaju modelu da selektivno ažurira svoja interna stanja tokom vremena. Ovi mehanizmi omogućavaju GRU modelu da efikasno obrađuje dugoročne zavisnosti u podacima, smanjujući istovremeno broj parametara u odnosu na LSTM modele.

GRU funkcioniše tako što kapija za ažuriranje kontroliše koliko informacija će biti ažurirano u skladu sa novim podacima. Ako je vrednost ažuriranja bliska 1, to znači da će se većina informacija zadržati iz prethodnog koraka, dok će se manje informacija uzeti iz novih podataka. Ako je vrednost bliska 0, znači da će se više pažnje posvetiti novim podacima. Pored toga, kapija za resetovanje kontroliše koliko informacija iz prethodnog stanja će biti "zaboravljeno" pri obradi novih podataka. Ako je vrednost zaboravljanja bliska 1, to znači da će se veći deo prethodnog stanja zadržati, dok će se manje zaboraviti. Ako je vrednost bliska 0, to znači da će se više pažnje posvetiti novim podacima, a manje prethodnom stanju. Kombinacijom informacija iz prethodna dva mehanizma, GRU model generiše finalni izlaz za trenutni korak u sekvenci.

GRU modeli su jednostavniji od LSTM modela i zahtevaju manje parametara za treniranje, što ih čini bržim i manje sklonim preobučavanju (eng. overfitting). Zahvaljujući svojoj jednostavnoj arhitekturi, GRU modeli su često brži u treniranju i inferenciji u poređenju sa složenijim LSTM modelima. Iako su jednostavniji od LSTM modela, GRU modeli i dalje su u stanju da efikasno modeliraju i koriste dugoročne zavisnosti u sekvencijalnim podacima. [5]

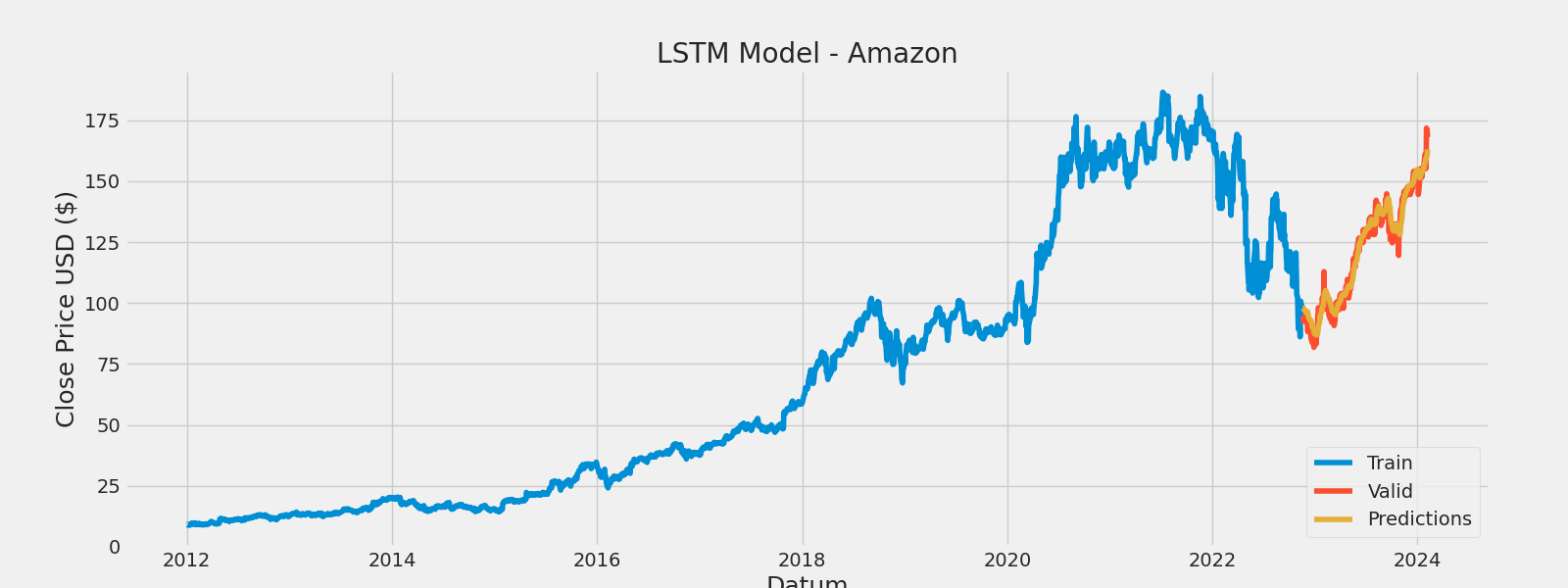
Šematski prikaz GRU modela rekurentne neuralne mreže je dat na slici 4.

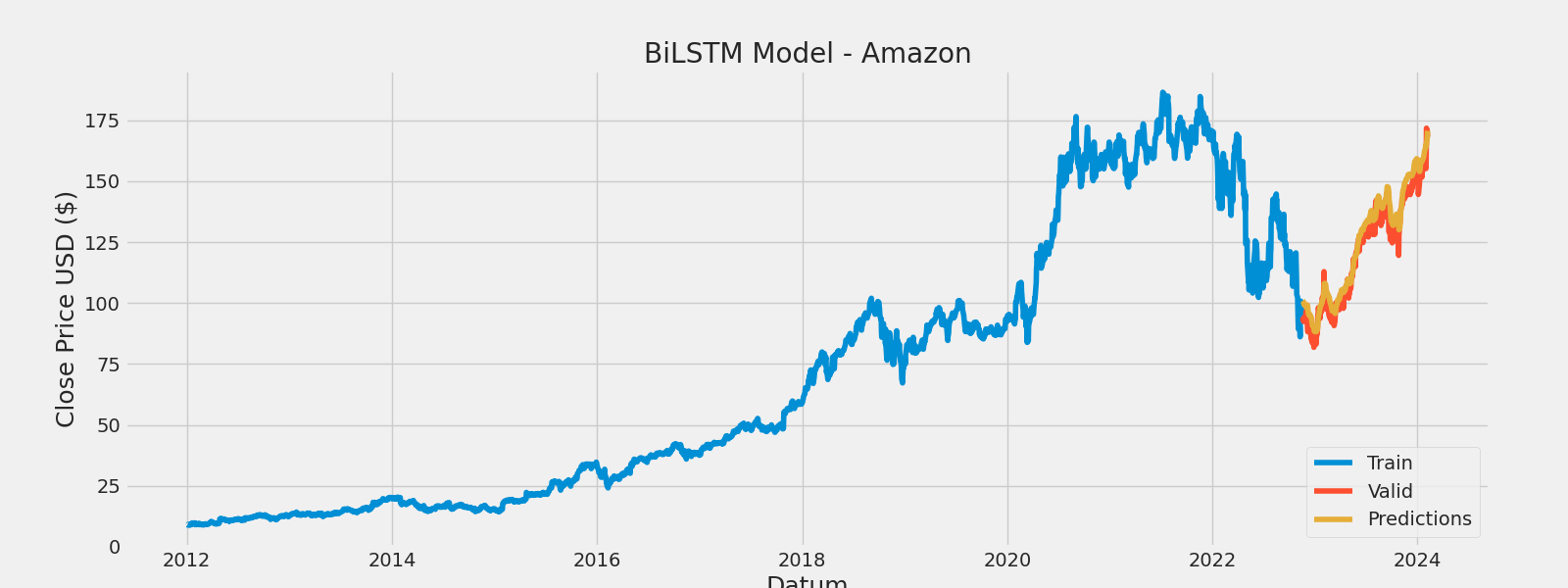
Sl. 4. Generalna forma GRU modela rekurentne neuralne mreže

Početna ideja bila je da se u ovom radu kreiraju tri modela rekurentnih neuralnih mreža (LSTM, BiLSTM i GRU) sa istom arhitekturom neuralne mreže kako bismo mogli da uporedimo ponašanje sva tri modela, s obzirom da sva tri modela imaju različitu kompleksnost u internoj strukturi, principu funkcionisanja i implementaciji. Izazov je takođe uskladiti arhitekturu mreža tako da ona ne bude ni suviše jednostavna ni suviše kompleksna zbog poznatih problema podobučavanja i preobučavanja. Arhitektura modela rekurentnih mreža u ovom radu se sastoji od 4 sloja - ulazni sloj ima 128 veštačkih neurona i ReLU (eng. Rectified Linear Unit) aktivacionu funkciju, na njega se nastavlja sledeći sloj mreže sa 64 veštačka neurona i istom aktivacionom funkcijom, ali za svaki model pojedinačno, drugačija je implementacija ova dva sloja neuralnih mreža. Nakon ova dva sloja, slede dva potpuno povezana, odnosno “gusta” (eng. Dense) sloja, jedan sloj sa 25 veštačkih neurona i jedan, poslednji sloj koji će generisati predikcije sa jednim veštačkim neuronom. Ovakva arhitektura modela rekurentnih neuralnih mreža određena je empirijski.

# Rezultati

U ovom radu su upoređeni LSTM, BiLSTM i GRU modeli rekurentnih neuralnih mreža istih arhitektura u cilju analize i predikcije završnih cena akcija na berzi iz Yahoo Finance dostupnog skupa podataka za četiri svetske kompanije Amazon, Apple, Microsoft i Google. Za svaku kompaniju, nakon učitavanja podataka iz baze od datuma 01.01.2012. do 06.02.2024. godine, urađena je normalizacija podataka i podela na trenirajući i testirajući skup. Modeli su trenirani na trenirajućem skupu koristeći Adam optimizator koji minimizuje srednju kvadratnu funkciju greške tokom samo jedne epohe obučavanja i za veličinu paketa (eng. Batch Size) ulaznih podataka jednak 1. Nakon procesa treniranja, performanse modela su evaluirane na testirajućem skupu podataka preko korena srednje kvadratne greške završne cene akcija na celom testirajućem skupu. U nastavku će biti prikazane vremenske zavisnosti cene od interesa tokom vremena za kompaniju Amazon na slikama 5, 6 i 7.

Sl. 5. Zavisnost cene akcija kompanije Amazon tokom vremena - LSTM

Sl. 6. Zavisnost cene akcija kompanije Amazon tokom vremena - BiLSTM

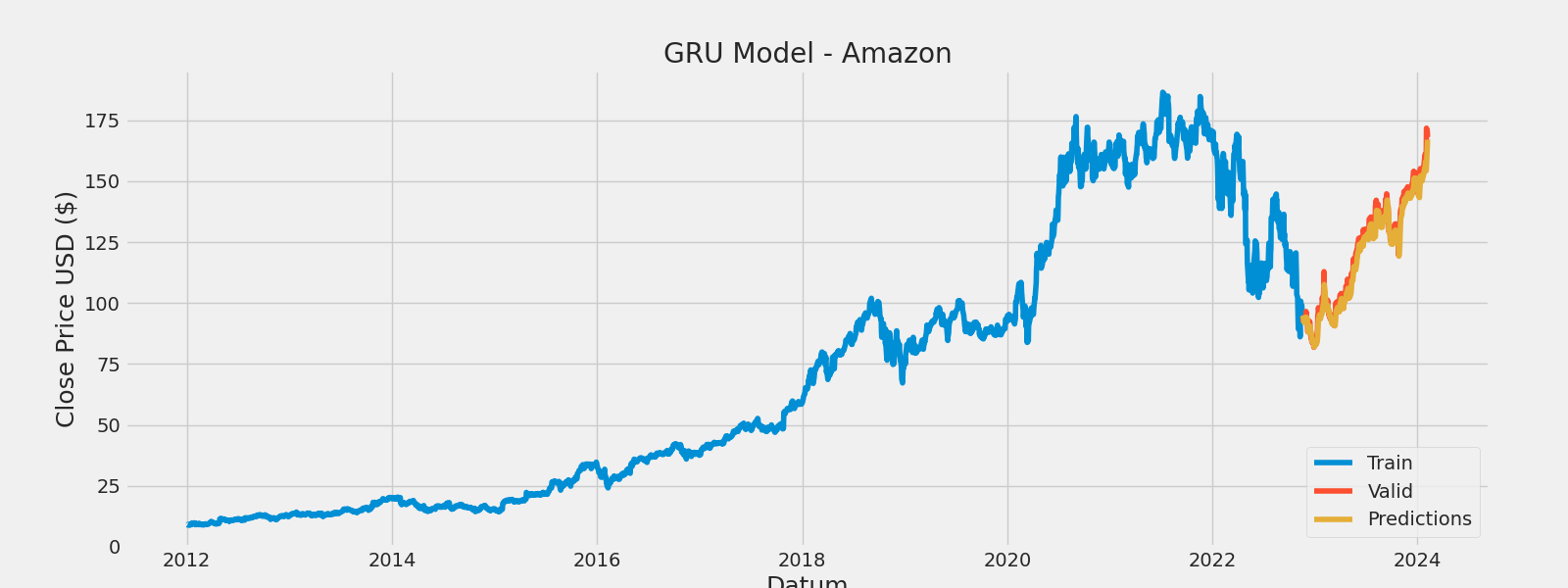
Sl. 7. Zavisnost cene akcija kompanije Amazon tokom vremena - GRU

TABELA I

Koren srednje kvadratne greške na testirajućem skupu [USD $]

|  | **LSTM** | **BiLSTM** | **GRU** |
| --- | --- | --- | --- |
| **Amazon** | 4.014 | 5.188 | 3.363 |
| **Apple** | 5.418 | 7.287 | 3.145 |
| **Microsoft** | 7.858 | 9.595 | 5.589 |
| **Google** | 3.957 | 8.123 | 3.512 |

U prethodnoj tabeli 1 možemo videti koliko iznose koreni srednje kvadratne greške na testirajućim skupovima podataka za sve skupove podataka, odnosno za sve dostupne podatke kompanija i za svaki prediktivni model rekurentnih neuralnih mreža koji su korišćeni u ovom radu. S obzirom da su ove vrednosti izražene u jedinicama američkih dolara, u narednoj tabeli 2, kao dodatak informativnosti dobijenih rezultata prikazani su odnosi tih korena srednje kvadratne greške na testirajućim skupovima podataka i srednjih vrednosti testirajućih skupova podataka za svaku kompaniju izraženi u procentima. Ti procenti potvrđuju ono što se može videti na slikama iznad konkretno za kompaniju Amazon, ali i za ostale kompanije, a to je da se najbolji rezultati postižu za GRU model koji je najjednostavniji od pomenuta tri modela. Nešto slabiji rezultati se postižu korišćenjem LSTM modela, dok BiLSTM model vrši najlošiju predikciju, naročito za podatke kompanije Google, gde odnos korena srednje kvadratne greške i srednje vrednosti testirajućeg skupa podataka iznosi čak 6.81%.

TABELA II

Količnik korena srednje kvadratne greške na testirajućem skupu i srednje vrednosti cena akcija na testirajućem skupu [%]

|  | **LSTM** | **BiLSTM** | **GRU** |
| --- | --- | --- | --- |
| **Amazon** | 3.313 | 4.282 | 2.775 |
| **Apple** | 3.173 | 4.268 | 1.842 |
| **Microsoft** | 2.490 | 3.041 | 1.771 |
| **Google** | 3.317 | 6.810 | 2.944 |

# Zaključak

Dobijeni rezultati prikazani u prethodnom delu nam pokazuju da su kreirani LSTM, BiLSTM i GRU modeli rekurentnih neuralnih mreža sa jednostavnom arhitekturom u stanju da daju veoma dobre predikcije na skupovima podataka koji su vrlo dinamični. GRU model, kao najjednostavniji od pomenuta tri modela generiše najverodostojnije predikcije od ova tri modela, a BiLSTM kao najkompleksniji model postiže najlošije performanse na testirajućem skupu podataka, što se može uočiti direktno na slikama zavisnosti cena akcija kompanija, kao i numerički i analitički upoređujući količnike korean srednje kvadratne greške i srednje vrednosti testirajućeg skupa podataka svake kompanije pojedinačno. Unapređenje ovog rada može biti kreiranje nekih novih i kompleksnijih arhitektura za svaki model neuralnih mreža pojedinačno u cilju dodatnog smanjivanja količnika korena srednje kvadratne greške i srednje vrednosti testirajućeg skupa podataka. Takođe, u ovom radu se nismo bavili određivanjem optimalne vrednosti paketa ulaznih podataka, što takođe može da bude nastavak ovog rada. Na samom kraju, ali ne i najmanje bitno, obučavanje modela korišćenih u ovom radu trajalo je samo kroz jednu epohu, s obzirom da je obučavanje velikog trenirajućeg skupa, kao što je bio slučaj u ovom radu, vremenski veoma zahtevno, stoga obučavanje modela povećanjem broja epoha obučavanja bi doprinelo boljim rezultatima na testirajućem skupu, a uzimajući u obzir upravo tu jednu epohu obučavanja, postignuti rezultati u ovom radu su veoma zadovoljavajući svih korišćenih modela rekurentnih neuralnih mreža na svim ispitivanim skupovima podataka kompanija.

Literatura

1. Stock Market Analysis + Prediction using LSTM, <https://www.kaggle.com/code/faressayah/stock-market-analysis-prediction-using-lstm/notebook>
2. Intro to Recurrent Neural Network LSTM | GRU, <https://www.kaggle.com/code/thebrownviking20/intro-to-recurrent-neural-networks-lstm-gru>
3. Long Short Term Memory Networks Explanation, <https://www.geeksforgeeks.org/long-short-term-memory-networks-explanation/>
4. Bidirectional LSTM in NLP, <https://www.geeksforgeeks.org/bidirectional-lstm-in-nlp/>
5. Gated Recurrent Unit Networks, <https://www.geeksforgeeks.org/gated-recurrent-unit-networks/>
6. <https://opendatascience.com/understanding-the-mechanism-and-types-of-recurring-neural-networks/>
7. <https://medium.com/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c>
8. <https://www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm>
9. <https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2>

Abstract

This paper explores the application of deep neural networks, specifically LSTM (Long Short-Term Memory), BiLSTM (Bidirectional Long Short-Term Memory), and GRU (Gated Recurrent Unit) models, in analyzing the stock market and predicting the prices of financial instruments. Introducing these complex neural architectures aims to investigate the model's ability to identify and model temporal dependencies in time series data with the goal of improving the accuracy of stock price predictions. The research methodology involves collecting and processing a large dataset of stock movements, implementing LSTM, BiLSTM, and GRU models through deep learning, and evaluating the performance of these models on a test dataset. The research results provide insights into the effectiveness of different neural architectures in predicting stock movements and highlight potential advantages and limitations of each model. This analysis has significant implications for making investment decisions and managing risks in financial markets.

**Stock market analysis and prediction using LSTM, BiLSTM and GRU models**

Aleksa Janjić i Mihailo Radojević

1. Mihailo Radojević - Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: [rm233209m@student.etf.bg.ac.rs](mailto:rm233209m@student.etf.bg.ac.rs))

   Aleksa Janjić - Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e.mail: [ja233085m@student.etf.bg.ac.rs](mailto:ja233085m@student.etf.bg.ac.rs)) [↑](#footnote-ref-0)