

*TODO Zahvala*

# Sadržaj

<b>1. Uvod</b>	<b>2</b>
<b>2. Metodologija</b>	<b>3</b>
2.1. Povrati imovine	3
2.2. Statistički modeli povrata	6
2.3. Faktorski modeli	10
2.3.1. Linearni faktorski model	11
2.3.2. Jednofaktorski model povrata	13
2.4. Duboko učenje	13
2.4.1. Povratne neuronske mreže	15
2.4.2. LSTM mreže	17
<b>3. Implementacija</b>	<b>21</b>
3.1. Podatci	21
3.2. Model dubokog učenja	21
<b>4. Rezultati</b>	<b>22</b>
<b>5. Zaključak</b>	<b>23</b>
<b>Literatura</b>	<b>24</b>
<b>Sažetak</b>	<b>26</b>
<b>Abstract</b>	<b>27</b>

# **1. Uvod**

TODO Uvod.

## 2. Metodologija

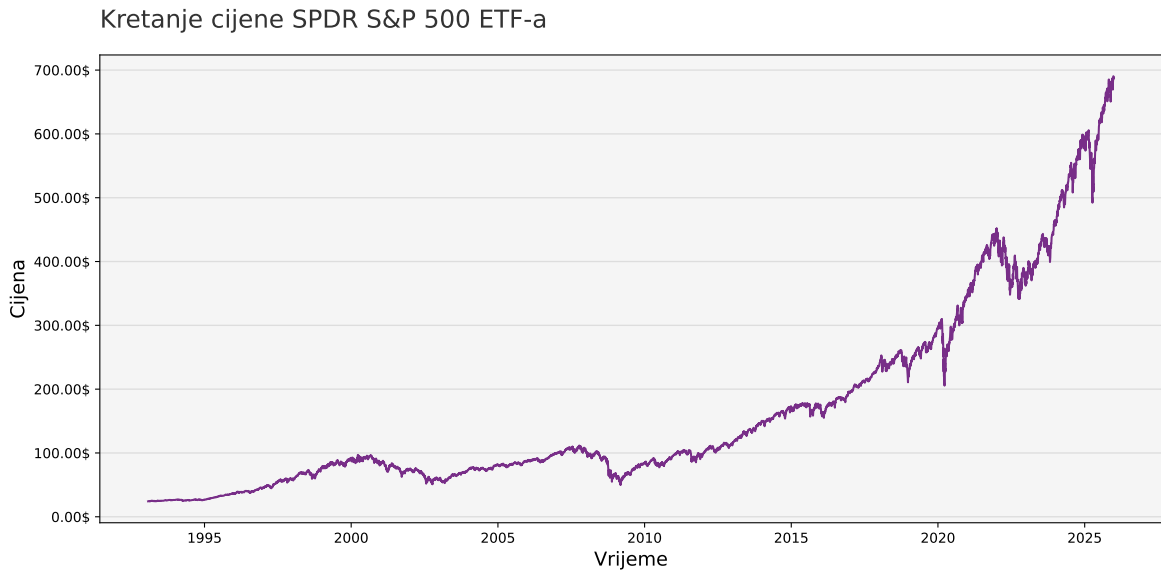
Ovo poglavlje pokriva teoretsku osnovu potrebnu za razumijevanje implementiranog modela dubokog učenja te strukturu podataka na kojima se primjenjuje. Prvo ćemo definirati osnovne varijable koje ćemo se oslanjati u ovom radu. Zatim ćemo proučiti njihova statistička svojstva. Nakon toga razmotrit ćemo faktorske modele koji omogućuju modeliranje naših varijabli. Na kraju ćemo pokriti duboko učenje i specifičnu arhitekturu koja će se koristiti u ovom radu.

### 2.1. Povrati imovine

Vrijednosni papiri predstavljaju financijske instrumente koji potvrđuju određena imovinska ili druga prava njihova vlasnika, poput prava na udio u vlasništvu poduzeća (dionice) ili prava na povrat uložених sredstava uz kamatu (obveznice). Njima se trguje na organiziranim tržištima kapitala, poput burzi, gdje se kupnja i prodaja odvijaju putem ovlaštenih posrednika, a cijene se formiraju na temelju ponude i potražnje. Tržišna cijena vrijednosnog papira u svakom trenutku odražava ravnotežu između kupaca i prodavatelja, pri čemu se svaka realizirana transakcija bilježi kao nova referentna cijena. Tržišni indeksi predstavljaju promjene vrijednosti grupe dionica ili drugih financijskih instrumenata. Služe kao mjerilo performansi tržišta ili određenog segmenta tržišta [1].

Kako bi se omogućila analiza kretanja cijena kroz vrijeme, kontinuirani tok transakcijskih podataka uzorkuje se u diskretnim vremenskim intervalima (npr. minute, sati, dani), čime se dobivaju vremenski nizovi cijena koji služe kao temelj za statističku analizu i modeliranje. Slika 2.1. prikazuje kretanje dnevnih cijena za SPDR S&P 500 ETF (engl. *exchange traded fund*) od njegovog začeća 1993. do kraja 2025. godine.

U ovom radu koristit ćemo povrate umjesto samih cijena imovina. Postoje dva glavna



**Slika 2.1.** Kretanje cijene SPDR S&P 500 ETF-a

razloga za to. Prvi je to što povrati sažimaju kretanja cijena te ih stavljaju na jednaku i usporedivu skalu. Drugi se odnosi na povoljnija statistička svojstva povrata u odnosu na cijene. Postoji više načina za definirati povrate [2].

### Jednostavni (artimetički) povrati

Nek je  $P_t$  cijena imovine u trenutku  $t$ . Ako imovinu posjedujemo od trenutka  $t - 1$  do trenutka  $t$ , ostvareni artimetički povrat dobivamo izrazom:

$$1 + R_t = \frac{P_t}{P_{t-1}} \quad (2.1)$$

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (2.2)$$

Ukoliko imovinu držimo kroz  $T$  perioda ukupni artimetički povrat dobivamo ukupnim množenjem artimetičkih povrata u vremenu [3]:

$$\begin{aligned} 1 + R_{total} &= \frac{P_T}{P_1} = \frac{P_T}{P_{T-1}} \times \frac{P_{T-1}}{P_{T-2}} \times \dots \times \frac{P_2}{P_1} \\ &= (1 + R_T)(1 + R_{T-1}) \dots (1 + R_1) \\ &= \prod_{t=1}^T (1 + R_t). \end{aligned} \quad (2.3)$$

## Kontinuirani (logaritamski) povrati

Prirodni logaritam artimetričkog povrata (2.1) je logaritamski povrat:

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \quad (2.4)$$

gdje  $p_t = \ln(P_t)$ .

Razmotrimo zatim ukamaćivanje logaritamskih povrata u vremenu:

$$\begin{aligned} r_{total} &= \ln(1 + R_{total}) = \ln[(1 + R_T)(1 + R_{T-1}) \dots (1 + R_1)] \\ &= \ln(1 + R_T) + \ln(1 + R_{T-1}) + \dots + \ln(1 + R_1) \\ &= r_T + r_{T-1} + \dots + r_1 \\ &= \sum_{t=1}^T r_t \end{aligned} \quad (2.5)$$

Ukamaćivanje logaritamski povrata može se ostvariti zbrajanjem pojedinačnih logaritamskih povrata. To svojstvo zovemo *aditivnost u vremenu*. Logaritamski povrati također posjeduju poželjna statistička svojstva [3]. Za male magnitude vrijedi  $r_t \approx R_t$ . Ova aproksimacija je korisna kada se razmatraju kratki vremenski intervali [1]. Slika 2.2. prikazuje dnevne artimetričke i logaritamske povrate SPDR S&P 500 ETF. Sa slike vidimo da nema značajne razlike između artimetričkih i logaritamskih povrata jer su povrati izračunati na kratkom odnosno dnevnom vremenskom intervalu.

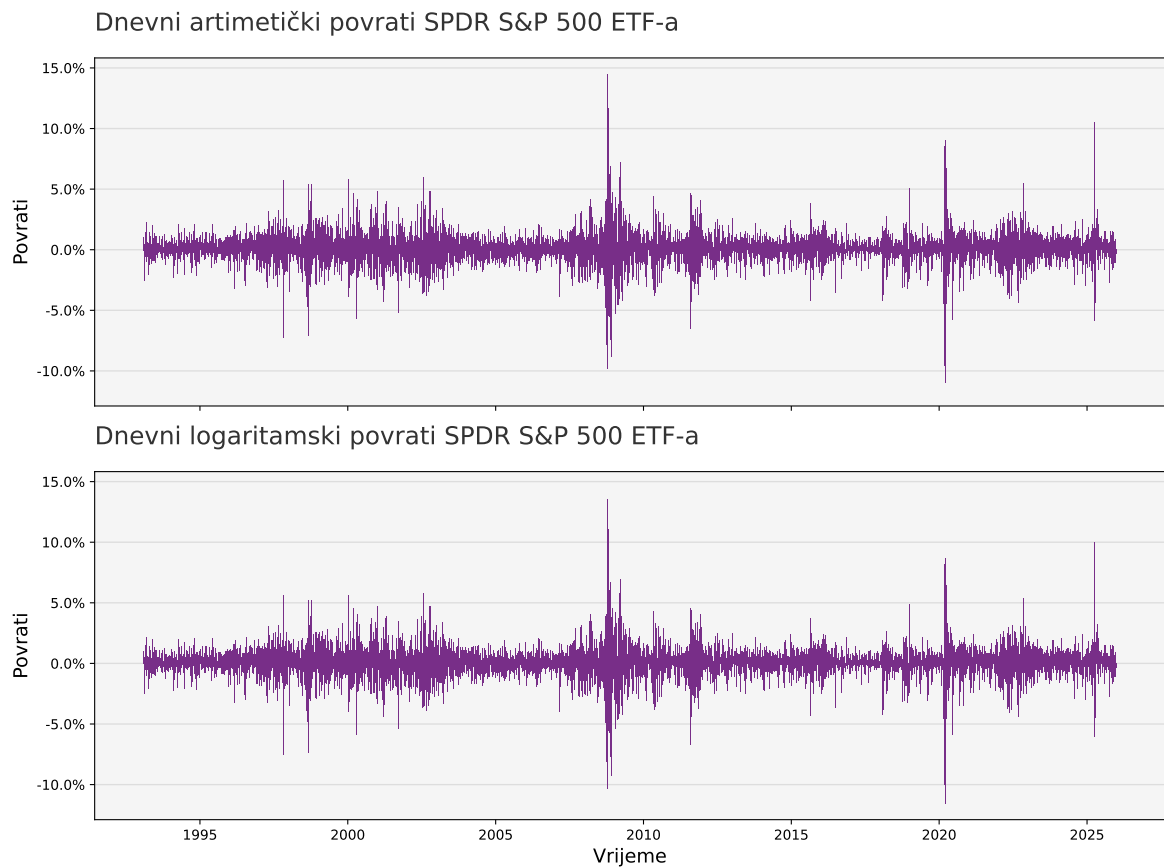
## Povrat portfelja

Artimetrički povrat portfelja koji se sastoji od  $N$  vrijednosnica je otežana artimetrička sredina (engl. *weighted average*) artimetričkih povrata vrijednosnica, gdje je težina (engl. *weight*) pojedine vrijednosnice njezin udio u vrijednosti portfelja [3].

Nek je  $p$  portfelj kojem je  $w_i$  ponder vrijednosnice  $i$ . Tada je artimetrički povrat portfelja  $p$  u trenutku  $t$  dan izrazom:

$$R_{pt} = \sum_{i=1}^N w_i R_{it}, \quad (2.6)$$

gdje je  $R_{it}$  artimetrički povrat vrijednosnice  $i$  u trenutku  $t$ . Ovo svojstvo nazivamo *aditiv-*



**Slika 2.2.** Aritmetički i logaritamski povrati SPDR S&P 500 ETF-a

*nost u prostoru vrijednosnica. Logaritamski povrati ne posjeduju to svojstvo.*

### **Povrati iznad bezrizične kamatne stope**

Povrati iznad bezrizične kamatne stope (engl. *excess return*) predstavlja razliku između ostvarenog povrata određene imovine i referentnog, bezrizičnog povrata, te mjeri dodatnu kompenzaciju koju investitor ostvaruje za preuzimanje rizika. Kao referentni povrat najčešće se koriste prinosi na državne obveznice visoke kreditne kvalitete i kratkog dospjeća, jer se smatra da nose zanemariv rizik [3].

$$R_t^{excess} = R_t - R_f \quad (2.7)$$

## **2.2. Statistički modeli povrata**

Prije nego razmotrimo konkretne statističke modele povrata, potrebno je ukratko opisati osnovne distribucije na kojima će se temeljiti daljnja analiza te načine estimacije njihovih

parametara.

## Normalna distribucija

Normalna distribucija jedan je od najčešće korištenih probabilističkih modela u statistici i financijama zbog svoje matematičke jednostavnosti i dobrih teorijskih svojstava. Ako slučajna varijabla  $X$  slijedi normalnu distribuciju sa sredinom  $\mu$  i varijancom  $\sigma^2$ , tada njezina funkcija gustoće ima oblik [4]:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right). \quad (2.8)$$

Za osmotreni uzorak  $\{x_1, x_2, \dots, x_n\}$ , procjena sredine dana je aritmetičkom sredinom

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (2.9)$$

dok se varijanca procjenjuje izrazom:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2. \quad (2.10)$$

## Lognormalna distribucija

Lognormalna distribucija koristi se za modeliranje slučajnih varijabli čiji je logaritam normalno distribuiran. Ako postoji broj  $a$  takav da  $Y = \ln(X - a)$  prati normalnu distribuciju, slučajna varijabla  $X$  tada slijedi lognormalnu distribuciju. Kako bi to vrijedilo, vjerojatnost da slučajna varijabla  $X$  poprimi vrijednost manju od  $a$  mora biti jednaka nuli. Ako vrijedi  $Y \sim \mathcal{N}(\mu, \sigma^2)$ , tada  $X$  ima gustoću:

$$f(x) = \frac{1}{(x - a)\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(x - a) - \mu)^2}{2\sigma^2}\right], \quad x > a. \quad (2.11)$$

Sredina i varijanca lognormalne distribucije dane su izrazima

$$E[X] = e^{\mu + \frac{\sigma^2}{2}}, \quad \text{Var}[X] = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \quad (2.12)$$

Procjena parametara provodi se logaritamskom transformacijom uzorka, nakon čega se primjenjuju standardne procjene za normalnu distribuciju [4].

## Studentova t distribucija

Studentova t distribucija predstavlja generalizaciju normalne distribucije s dodatnim parametrom broja stupnjeva slobode  $\nu$ , koji kontrolira debljinu repova. Za manje vrijednosti  $\nu$  distribucija ima izraženije repove, čime omogućuje robusnije modeliranje ekstremnih vrijednosti.

Gustoća Studentove t distribucije sa sredinom  $\mu$ , varijancom  $\sigma^2$  i  $\nu$  stupnjeva slobode ima oblik:

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi\sigma^2}} \left(1 + \frac{(x - \mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}. \quad (2.13)$$

Za procjenu parametara t distribucije koriste se metode procjene najveće izglednosti (engl. *maximum likelihood estimation*) [5].

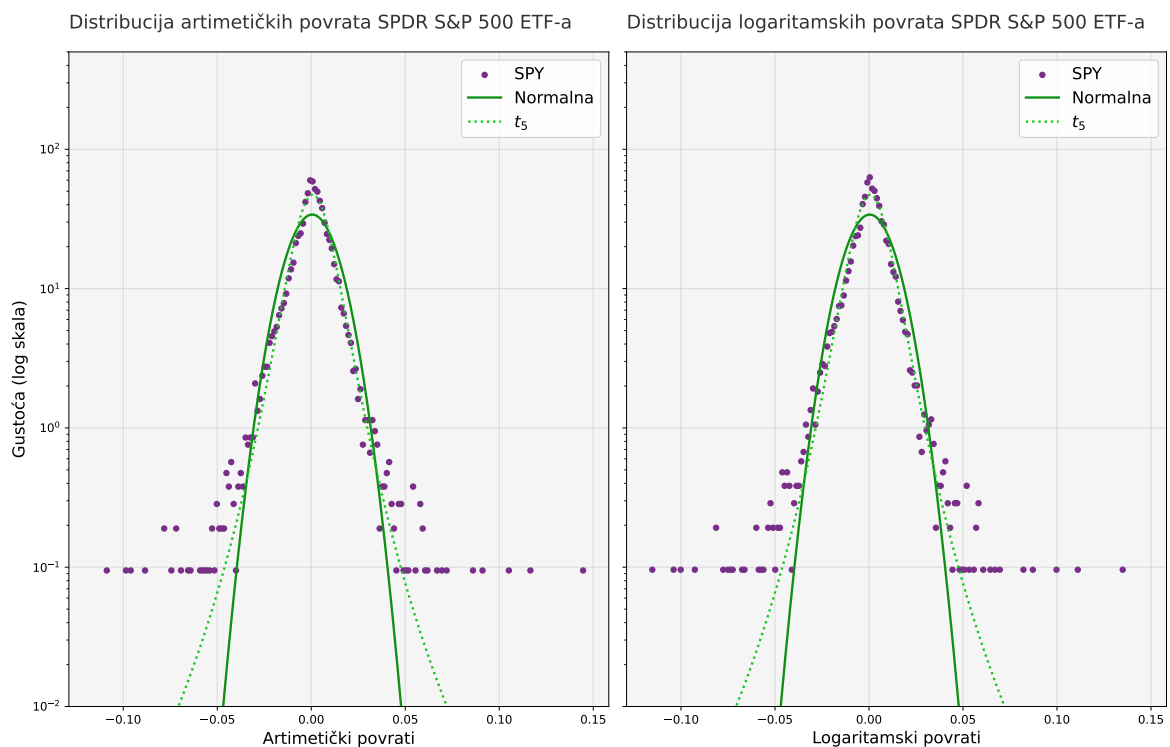
## Statistička svojstva povrata

U statističkoj analizi financijskih vremenskih nizova često se polazi od pretpostavke da su aritmetički povrati nezavisne i jednako distribuirane slučajne varijable s normalnom distribucijom, konstantnom sredinom i varijancom, čime se značajno pojednostavljuje teorijska obrada i izvođenje analitičkih rezultata. Međutim, takva pretpostavka suočava se s nizom ograničenja: aritmetički povrati imaju donju granicu od  $-1$ , dok normalna distribucija nema ograničenja po realnoj osi, višeperiodni povrati ne zadržavaju normalnu distribuciju zbog multiplikativne prirode (2.3), a empirijski podaci često pokazuju odstupanja od normalnosti [6].

Alternativno možemo pretpostaviti da su logaritamski povrati nezavisno i identično normalno distribuirani sa sredinom  $\mu$  i varijancom  $\sigma^2$ . S tom pretpostavkom impliciramo da su aritmetički povrati nezavisno i identično lognormalno distribuirani. Ovaj pristup ima povoljnija matematička svojstva jer su zbrojevi logaritamskih povrata kroz više razdoblja također normalno distribuirani (2.5), a pritom se prirodno zadovoljava donja granica aritmetičkog povrata (2.11). Unatoč tim prednostima, ni pretpostavka lognormalnosti u potpunosti ne opisuje realna tržišna kretanja, budući da stvarni financijski

povrati često pokazuju deblje repove distribucije i veću učestalost ekstremnih vrijednosti nego što to predviđa normalna distribucija [3].

Kako bismo bolje obuhvatili ta empirijska svojstva, možemo koristiti studentovu  $t$ -distribuciju. Slika 2.3. prikazuje distribuciju gustoće vjerojatnosti dnevnih artimetričkih i logaritamskih povrata SPDR S&P 500 ETF-a. Također su prikazane distribucije gustoće vjerojatnosti normalne distribucije i studentove  $t$ -distribucije s pet stupnjeva slobode <sup>1</sup>. Parametri  $\mu$  i  $\sigma^2$  za normalnu i  $t_5$ -distribuciju procijenjeni su iz artimetričkih povrata za lijevi graf i logaritamskih povrata za desni graf. Sa slike vidimo kako  $t_5$ -distribucija zbog svojih težih repova vijernije modelira distribuciju povrata od normalne distribucije.



**Slika 2.3.** Distribucija dnevnih artimetričkih i logaritamskih povrata SPY ETF-a

<sup>1</sup>Biramo pet stupnjeva slobode kako bi imali konačna prva četiri momenta distribucije. Za više informacija pogledati [5]

## Slučajni vektori

Nek je  $\mathbf{X} = (X_1, \dots, X_p)$  slučajni vektor  $p$  slučajnih varijabli. Tada su vektor sredina i kovarijacijska matrica dani izrazima:

$$E(\mathbf{X}) = \boldsymbol{\mu}_X = [E(X_1), \dots, E(X_p)]^T \quad (2.14)$$

$$\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}_X = E[(\mathbf{X} - \boldsymbol{\mu}_X)(\mathbf{X} - \boldsymbol{\mu}_X)^T], \quad (2.15)$$

uz uvjet da dana očekivanj postoje. Neka su  $\{\mathbf{x}_1, \dots, \mathbf{x}_T\}$  realizacije slučajnog vektora  $\mathbf{X}$ . Tada su uzoračka sredina i kovarijacijska matrica dane izrazima [3]:

$$\hat{\boldsymbol{\mu}}_x = \frac{1}{T} \sum_{t=1}^T \mathbf{x}_t, \quad \hat{\boldsymbol{\Sigma}}_x = \frac{1}{T-1} \sum_{t=1}^T (\mathbf{x}_t - \hat{\boldsymbol{\mu}}_x)(\mathbf{x}_t - \hat{\boldsymbol{\mu}}_x)^T. \quad (2.16)$$

Ako pretpostavimo da slučajni vektor  $\mathbf{X}$  dolazi is multivariatne normalne distribucije s vektorom sredina  $\boldsymbol{\mu}$  i kovarijaciskom matricom  $\boldsymbol{\Sigma}$ , tada je funkcija izglednosti uzorka (engl. *likelihood function*) dana jednadžbom [7]:

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{x}_1, \dots, \mathbf{x}_T) = (2\pi)^{-\frac{pT}{2}} |\det(\boldsymbol{\Sigma})|^{-\frac{T}{2}} \exp\left(-\frac{1}{2} \sum_{t=1}^T (\mathbf{x}_t - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_t - \boldsymbol{\mu})\right). \quad (2.17)$$

S obzirom da taj oblik funkcije izglednosti nije najprikladniji za računanje računalom iskazat ćemo i njezin logaritamski oblik (engl. *log-likelihood function*) te ćemo zamijeniti argument  $\boldsymbol{\Sigma}$  s  $\boldsymbol{\Sigma}^{-1}$ :

$$\begin{aligned} \ln \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1}; \mathbf{x}_1, \dots, \mathbf{x}_T) &= \\ &= \frac{1}{2} \left( -pT \ln(2\pi) + T \ln[\det(\boldsymbol{\Sigma}^{-1})] - \sum_{t=1}^T (\mathbf{x}_t - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_t - \boldsymbol{\mu}) \right). \end{aligned} \quad (2.18)$$

## 2.3. Faktorski modeli

U analizi povrata imovina često se primjenjuju multivarijatne statističke metode kako bi se proučilo ponašanje i međusobna povezanost većeg broja vrijednosnica unutar portfelja. Međutim, modeliranje velikog broja vremenskih nizova povrata dovodi do visokodimenzionalnih i kompleksnih modela koji su složeni za interpretaciju i primjenu u praksi.

Empirijska istraživanja pokazuju da se povrati različitih vrijednosnica često kreću na sličan način, što nas upućuje da postoje neki zajednički faktori koji utječu na njihovo kretanje. Primjerice, u razdobljima gospodarske krize pad aktivnosti cijelog gospodarstva obično rezultira istodobnim padom cijena većine vrijednosnica, dok rast određenog sektora, poput tehnološkog, često dovodi do rasta cijena većine dionica unutar tog sektora [8].

Te činjenice omogućuju faktorskim modelima da objasne kretanje većeg broja povrata pomoću ograničenog broja zajedničkih faktora te da pojednostavne njihovu analizu. Postoje tri vrste faktorskih modela [2]. Prvi su *markoekonomske faktorske modele* koji se fokusiraju na varijable kao što su rast BDP-a, kamatne stope, stopa inflacije i slično. U ovakvom su modelu faktori osmotrivi pa se model može estimirati linearnom regresijom. Drugi su *fundamentalni faktorske modele* koji koriste podatke o poduzećima kako bi konstruirali svoje faktore. Treći su *statistički faktorske modele* čiji su faktori neosmotrivi latentne varijable koje se estimiraju iz podataka [3]. Ovaj rad će se fokusirati prvu vrstu odnosno na *makroekonomske faktorske modele*.

### 2.3.1. Linearni faktorski model

Pretpostavimo da imamo  $p$  imovina kroz  $T$  vremenskih trenutaka. Neka je  $r_{it}$  povrat imovine  $i$  u trenutku  $t$ . Opća forma faktorskog modela dana je izrazom:

$$r_{it} = \alpha_i + \beta_{i1}f_{1t} + \cdots + \beta_{im}f_{mt} + \epsilon_{it}, \quad t = 1, \dots, T; \quad i = 1, \dots, p, \quad (2.19)$$

gdje je  $\alpha_i$  konstanta,  $\{f_{jt} | j = 1, \dots, m\}$  su  $m$  zajedničkih (sistemskih) faktora,  $\beta_{ij}$  su koeficijenti imovine  $i$  uz faktor  $j$ , a  $\epsilon_{it}$  je specifičan (idiosinkratski) faktor imovine  $i$ .

Za faktor  $\mathbf{f}_t = (f_{1t}, \dots, f_{mt})^T$  vrijedi:

$$E(\mathbf{f}_t) = \boldsymbol{\mu}_f, \quad (2.20)$$

$$\text{Cov}(\mathbf{f}_t) = \boldsymbol{\Sigma}_f, \quad m \times m \text{ matrica}, \quad (2.21)$$

dok je idiosinkratski faktor  $\epsilon_{it}$  modeliran bijelim šumom koji nije koreliran sa sistemskim

faktorima  $f_{jt}$  i drugim idiosinkratskim faktorima. Dakle pretpostavljamo:

$$E(\epsilon_{it}) = 0 \quad \text{za sve } i, t \quad (2.22)$$

$$\text{Cov}(f_{jt}, \epsilon_{is}) = 0 \quad \text{za sve } i, j, t, s \quad (2.23)$$

$$\text{Cov}(\epsilon_{it}, \epsilon_{js}) = \begin{cases} \sigma_i^2, & \text{ako } i = j \text{ i } t = s, \\ 0, & \text{inače.} \end{cases} \quad (2.24)$$

Jednadžbu (2.19) možemo zapisati i u matričnom obliku za svih  $p$  imovina u trenutku  $t$ :

$$\mathbf{r}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} \mathbf{f}_t + \boldsymbol{\epsilon}_t, \quad t = 1, \dots, T, \quad (2.25)$$

gdje je  $\mathbf{r}_t = (r_{1t}, \dots, r_{pt})^\top$  vektor povrata,  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)^\top$  vektor konstanti,  $\boldsymbol{\beta} = [\beta_{ij}]$  je  $p \times m$  matrica koeficijenata, a  $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{pt})^\top$  je vektor idiosinkratskih faktora čija je kovarijacijska matrica  $\text{Cov}(\boldsymbol{\epsilon}_t) = \boldsymbol{\Psi} = \text{diag}\{\sigma_1^2, \dots, \sigma_k^2\}$   $p \times p$  dijagonalna matrica. Kovarijacijsku matricu povrata  $\mathbf{r}_t$  možemo računati sljedećom formulom:

$$\text{Cov}(\mathbf{r}_t) = \boldsymbol{\beta} \boldsymbol{\Sigma}_f \boldsymbol{\beta}^\top + \boldsymbol{\Psi}. \quad (2.26)$$

Zatim jednadžbu (2.25) možemo zapisati u sljedećem obliku:

$$\mathbf{r}_t = \boldsymbol{\xi} \mathbf{g}_t + \boldsymbol{\epsilon}_t, \quad (2.27)$$

gdje je  $\mathbf{g}_t = (1, \mathbf{f}_t^\top)^\top$ , a  $\boldsymbol{\xi} = [\boldsymbol{\alpha}, \boldsymbol{\beta}]$  je  $p \times (m + 1)$  matrica. Transponiramo li prethodnu jednadžbu i grupiramo podatke za svih  $T$  trenutaka dobivamo:

$$\mathbf{R} = \mathbf{G} \boldsymbol{\xi}^\top + \mathbf{E}, \quad (2.28)$$

gdje je  $\mathbf{R}$   $T \times p$  matrica povrata čiji je  $t$ -ti redak  $\mathbf{r}_t^\top$ ,  $\mathbf{G}$  je  $T \times (m + 1)$  matrica čiji je  $t$ -ti redak  $\mathbf{g}_t^\top$ ,  $\mathbf{E}$  je  $T \times p$  matrica idiosinkratskih faktora čiji je  $t$ -ti redak  $\boldsymbol{\epsilon}_t^\top$  [3].

S obzirom da nas zanimaju *makroekonomski faktorski modeli* čiji su faktori  $\mathbf{f}_t$  osmotrivi, jednadžba (2.28) ima oblik višestruke multivariatne linearne regresije. Zbog toga

parametre modela možemo estimirati metodom najmanjih kvadrata [9]:

$$\hat{\xi}^\top = \begin{bmatrix} \hat{\alpha}^\top \\ \hat{\beta}^\top \end{bmatrix} = (\mathbf{G}^\top \mathbf{G})^{-1} (\mathbf{G}^\top \mathbf{R}), \quad (2.29)$$

odakle su  $\alpha$  i  $\beta$  lako dostupni. Rezidualne, odnosno povrate idiosinkratskih faktora možemo lako dobiti koristeći formulu (2.28):

$$\hat{\mathbf{E}} = \mathbf{R} - \mathbf{G} \hat{\xi}^\top. \quad (2.30)$$

### 2.3.2. Jednofaktorski model povrata

Jedan od najpoznatijih *makroekonomskih faktorskih modela* koristi povrat tržišta kao faktor koji utječe na sve vrijednosnice:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \epsilon_{it} \quad i = 1, \dots, p; \quad t = 1, \dots, T, \quad (2.31)$$

gdje je  $r_{it}$  povrat vrijednosnice  $i$  iznad bezrizične kamatne stope, a  $r_{mt}$  povrat tržišta iznad bezrizične kamatne stope. Kod modeliranja dionica, za povrat tržišta  $r_{mt}$  uzima se povrat nekog tržišnog indeksa (npr. S&P 500) iznad bezrizične kamatne stope. Koeficijenti modela  $\alpha_i$  i  $\beta_i$  procjenjuju se metodom najmanjih kvadrata (2.29).

Ovaj rad fokusirat će se upravo na jednofaktorski model povrata. Cilj ovog rada biti će ispitati može li model dubokog učenja, iz prozora povijesnih povrata  $\mathbf{R}_H = \{\mathbf{r}_t\}_{t=1}^k$ , procijeniti koeficijente  $\alpha_i$  i  $\beta_i$ , koji će bolje odgovarati budućem prozoru povrata  $\mathbf{R}_F = \{\mathbf{r}_t\}_{t=k+1}^T$ , nego procjena koeficijenata  $\alpha_i$  i  $\beta_i$  koju možemo dobiti metodom najmanjih kvadrata na istom povijesnom prozoru povrata  $\mathbf{R}_H$  za  $1 \leq k < T$ . Način na koji ćemo mjeriti koliko dobro procjena koeficijenata  $\alpha_i$  i  $\beta_i$  odgovara budućem prozoru povrata  $\mathbf{R}_F$  biti će detaljnije objašnjen u potpoglavlju 3.2.

## 2.4. Duboko učenje

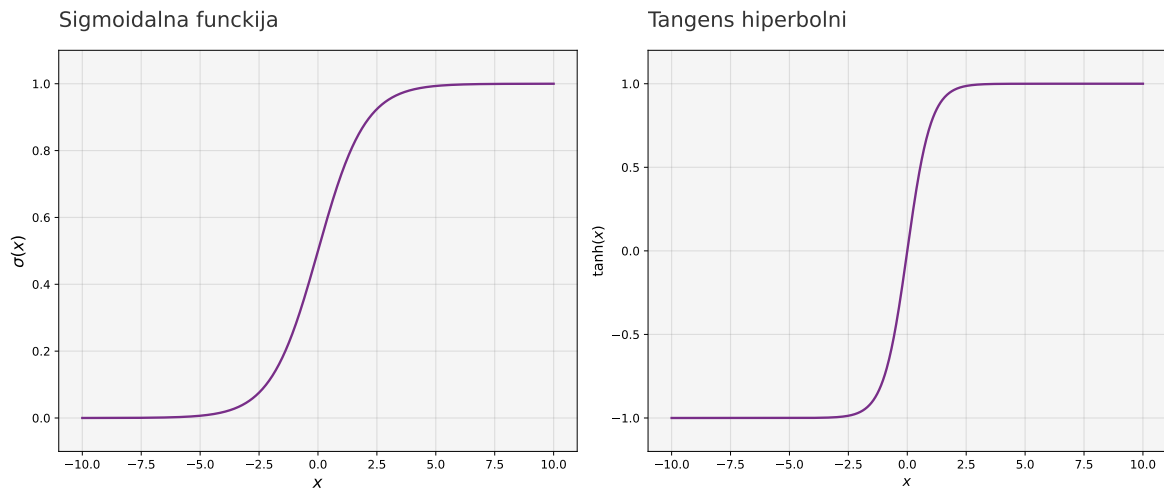
Duboko učenje predstavlja podpodručje strojnog učenja koje se ističe u rješavanju problema s visokom dimenzionalnošću podataka, kao što su računalni vid, obrada prirodnog

jezika, financijska te slične složene domene. Temeljna ideja dubokog učenja je izgradnja hijerarhijskih, složenih reprezentacija podataka, koje se dobivaju primjenom uzastopnih nelinearnih transformacija modeliranih pomoću višeslojnih neuronskih mreža. Osnovni element svake neuronske mreže je umjetni neuron. Za zadani ulazni vektor  $\mathbf{x} = (x_1, \dots, x_n)^T$ , izlaz neurona definiramo jednažbom:

$$h = f(\mathbf{w}^T \mathbf{x} + b) \quad (2.32)$$

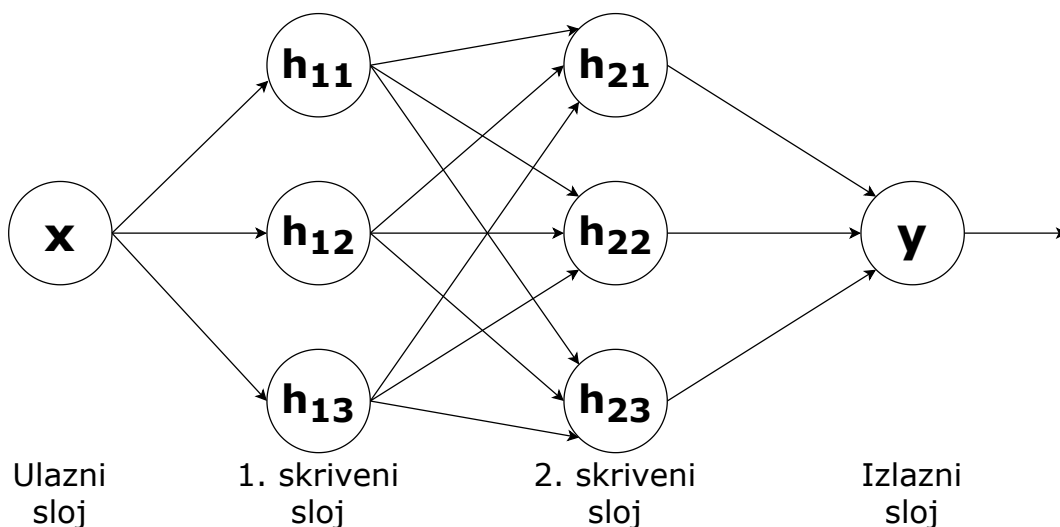
gdje je  $\mathbf{w} = (w_1, \dots, w_n)^T$  vektor težina neurona koji određuje doprinos pojedine komponente ulaznog vektora, a skalar  $b$  omogućuje dodatni pomak linearne kombinacije ulaza. Aktivacijska funkcija  $f$  uvodi nelinearnost u model, čime se omogućuje aproksimacija složenih i nelinearnih odnosa u podacima [10].

Neke od najčešće korištenih aktivacijskih funkcija su sigmoidalna funkcija i hiperbolni tangens, slika 2.4. Obje su nelinearne, monotonno rastuće i kontinuirano diferencijabilne, što je ključno svojstvo u procesu učenja neuronskih mreža metodama koje se temelje na gradijentnom spustu. Sigmoidalna funkcija ima kodomen u intervalu  $\langle 0, 1 \rangle$ , dok hiperbolni tangens poprima vrijednosti u intervalu  $\langle -1, 1 \rangle$ , pri čemu obje funkcije imaju karakterističan S-oblik.



**Slika 2.4.** Primjer aktivacijskih funkcija

Umjetne neurone zatim možemo organizirati u slojeve. Prvi, odnosno ulazni, sloj, zatim proizvoljan broj skrivenih te na kraju izlazni sloj. Time dobivamo osnovnu arhitekturu neuronske mreže, slika 2.5.



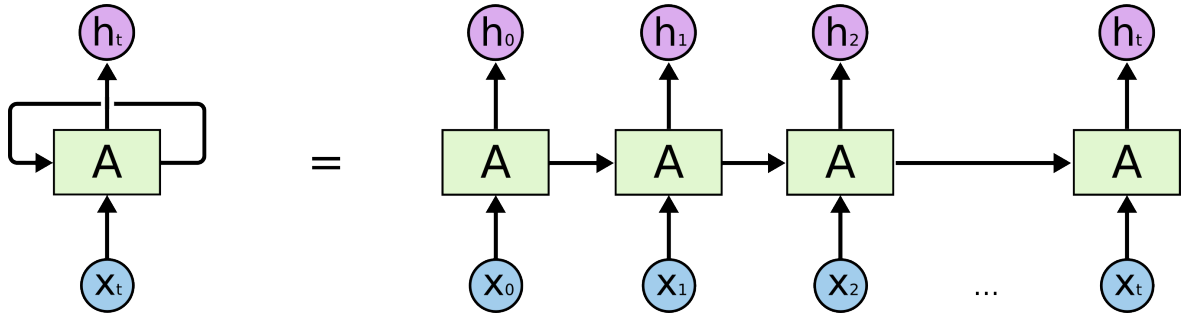
**Slika 2.5.** Jednostavna neuronska mreža

Među najznačajnijim arhitekturama dubokih modela ističu se duboke unaprijende mreže, konvolucijske neuronske mreže te povratne neuronske mreže. Svaka od navedenih arhitektura prilagođena je specifičnim vrstama podataka i problemima. Odabir odgovarajuće arhitekture ovisi o prirodi i strukturi odabranih podataka.

### 2.4.1. Povratne neuronske mreže

Dijeljenje parametara jedna je od ranijih ideja u strojnom učenju i statističkom modeliranju, koja omogućuje proširivanje modela na primjere različitih oblika te generalizaciju na podatke varijabilne duljine, osobito u kontekstu sekvencijalnih podataka. Ovakav se princip koristi u skrivenim Markovljevim modelima, gdje se isti skup parametara, poput matrice prijelaza  $P(\mathbf{s}_t \mid \mathbf{s}_{t-1})$ , koristi u svakom vremenskom koraku. Na taj način model ne ovisi o apsolutnoj poziciji u sekvenci, već isključivo o odnosu između susjednih stanja, što omogućuje modeliranje sekvenci proizvoljne duljine i dijeljenje statističke snage kroz različite vremenske korake. Povratne neuronske mreže (engl. *recurrent neural networks*, *RNN*) preuzimaju i generaliziraju ovu ideju u okviru dubokog učenja. One su prilagođene za obradu sekvencijalnih podataka tako da u svakom vremenskom koraku primjenjuju iste težine i istu transformaciju nad ulazom i prethodnim skrivenim stanjem, pri čemu skriveno stanje  $\mathbf{h}_t$  ovisi o trenutnom ulazu  $\mathbf{x}_t$  i stanju  $\mathbf{h}_{t-1}$  [11].

Slika 2.6. prikazuje dijagram jednostavne povratne neuronske mreže i njezin razmotani (engl. *unrolled*) oblik kroz vremenske korake. Povratnu neuronsku mrežu moguće



**Slika 2.6.** Jednostavna povratna neuronska mreža [12]

je promatrati kao višestruke kopije iste neuronske mreže, pri čemu svaka kopija prosljeđuje informaciju svojoj sljedbenici u idućem vremenskom koraku. U razmotanom prikazu jasno se vidi da se u svakom koraku primjenjuje ista RNN ćelija A s istim parametrima, dok skriveno stanje predstavlja viđeni dio slijeda koja se prenosi kroz vrijeme. Na taj način model zadržava informaciju o prethodnim ulazima te postupno akumulira kontekst, što mu omogućuje modeliranje ovisnosti unutar sekvence, neovisno o njezinoj duljini [12].

Jednostavni povratni model možemo definirati sljedećim jednadžbama:

$$\mathbf{h}_t = \tanh(\mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h), \quad (2.33)$$

$$\mathbf{o}_t = \mathbf{W}_{hy}\mathbf{h}_t + \mathbf{b}_o, \quad (2.34)$$

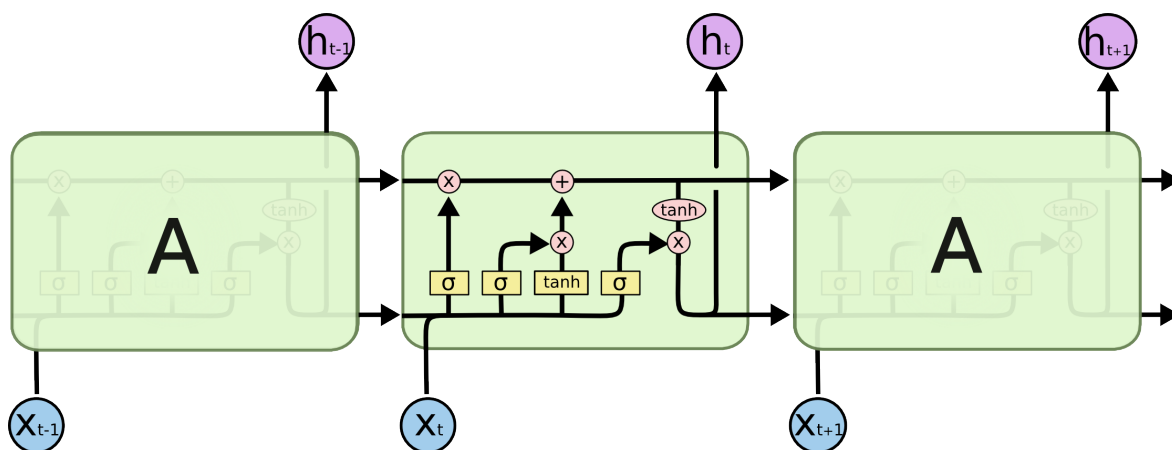
gdje matrica  $\mathbf{W}_{xh}$  projicira ulaz u prostor reprezentacije stanja, matrica  $\mathbf{W}_{hh}$  modelira evoluciju stanja dok matrica  $\mathbf{W}_{hy}$  projicira stanje na prostor predikcija. Vektori  $\mathbf{b}_h$  i  $\mathbf{b}_o$  omogućuju linearni pomak [13].

Međutim, pri učenju dugoročnih zavisnosti u povratnim neuronskim mrežama pojavljuju se značajni matematički problemi. Gradijenti koji se propagiraju kroz velik broj vremenskih koraka imaju tendenciju eksponencijalnog smanjivanja (problem nestajućeg gradijenta) ili, rjeđe, eksponencijalnog rasta (problem eksplodirajućeg gradijenta), što otežava stabilnu optimizaciju modela. Čak i ako pretpostavimo da su parametri stabilni, utjecaj udaljenih vremenskih koraka postaje eksponencijalno manji u odnosu na nedavne, zbog čega mreža teško uči dugoročne ovisnosti [11].

## 2.4.2. LSTM mreže

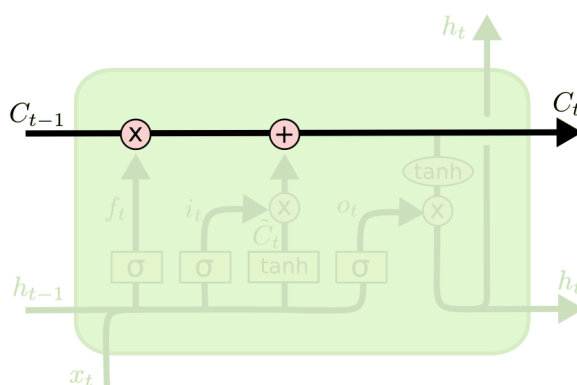
U okviru ovog rada koristit ćemo posebnu vrstu povratnih neuronskih mreža koje nazivamo povratna ćelija s dugoročnom memorijom (engl. *long short-term memory*, *LSTM*). LSTM mreže dizajnirane su kako bi izbjegle probleme nestajućeg i eksplodirajućeg gradijenta te omogućile učenje dugoročnih ovisnosti u sekvencijalnim podacima. Zbog toga se često primjenjuju u modeliranju vremenskih nizova [12].

Razmotana LSTM mreža ima strukturu lančanih ćelija karakterističnu za povratne neuronske mreže, no njene ponavljajuće ćelije imaju složeniju unutarnju arhitekturu. Umjesto jednog neuronskog sloja, LSTM ćelija sastoji se od četiri sloja koji imaju specifičnu interakciju. Ključni element LSTM arhitekture jest stanje ćelije  $C_t$ , koje predstavlja



**Slika 2.7.** Arhitektura LSTM mreže [12]

internu memorijsku komponentu modela. Ono se propagira kroz vremenske korake koristeći samo linearne transformacije, čime se omogućuje stabilniji prijenos informacija i ublažava problem nestajućeg gradijenta. Za razliku od standardnog skrivenog stanja



**Slika 2.8.** Tok stanja ćelije  $C_t$  [12]

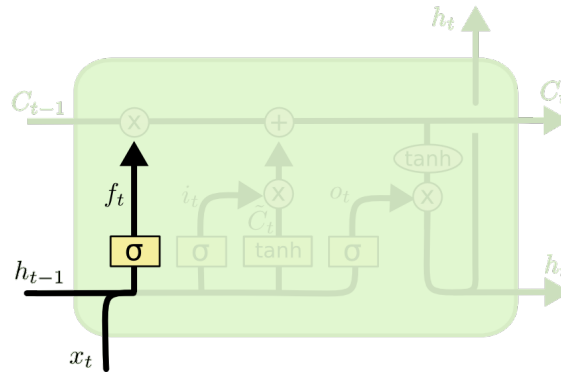
u jednostavnim RNN mrežama, stanje ćelije eksplicitno je regulirano posebnim mehanizmom vrata (engl. *gate*). LSTM arhitektura koristi tri vrste vrata: vrata zaboravljanja, ulazna vrata te izlazna vrata.

## Vrata zaboravljanja

Prvi korak u LSTM mreži se odnosi se na određivanje koje informacije iz prethodnog stanja ćelije  $\mathbf{C}_{t-1}$  treba izbaciti. Ta se odluka donosi pomoću vrata zaboravljanja, koja kao ulaz primaju prethodno skriveno stanje  $\mathbf{h}_{t-1}$  i trenutni ulaz  $\mathbf{x}_t$ . Rezultat je vektor  $\mathbf{f}_t$  koji računamo na sljedeći način:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f). \quad (2.35)$$

Pošto vrata zaboravljanja koriste sigmoidu kao aktivacijsku funkciju,  $\mathbf{f}_t \in [0, 1]^n$ . Elementi stanja  $\mathbf{C}_t$  na pozicijama na kojima je vektor  $\mathbf{f}_t$  blizu nuli bit će zaboravljeni, dok će se oni na pozicijama na kojima je vektor  $\mathbf{f}_t$  blizu jedinici propustiti u iduće stanje.



**Slika 2.9.** Vrata zaboravljanja [12]

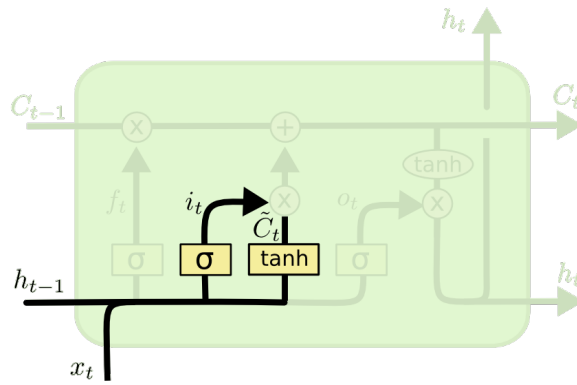
## Ulazna vrata

U idućem koraku moramo odrediti koju ćemo informaciju iz ulaza  $\mathbf{x}_t$  dodati stanju ćelije. Ovo radimo u dva koraka:

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i), \quad (2.36)$$

$$\tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_C \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_C). \quad (2.37)$$

Vektor  $\mathbf{i}_t$  određuje koje informacije je potrebno ažurirati, a sloj koji koristi tangens hiperbolni računa doprinose  $\tilde{\mathbf{C}}_t$  iz ulaza  $\mathbf{x}_t$  koje ćemo dodati stanju ćelije  $\mathbf{C}_{t-1}$ .



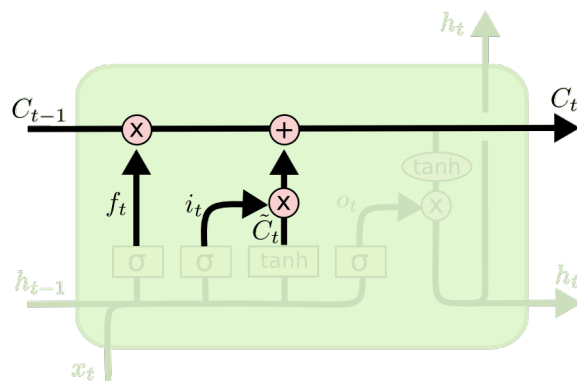
**Slika 2.10.** Ulazna vrata [12]

## Ažuriranje stanja ćelije

Zatim stanje ćelije zatim ažuriramo na sljedeći način:

$$\mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{C}}_t. \quad (2.38)$$

gdje je  $\odot$  oznaka za Hadamartov umnožak. Množimo staro stanje s  $\mathbf{f}_t$  kako bi zaboravili podatke koje smo odredili ranije. Zatim dodamo umnožak  $\mathbf{i}_t \odot \tilde{\mathbf{C}}_t$  koji predstavlja odabrane doprinose ulaza.



**Slika 2.11.** Ažuriranje stanja ćelije [12]

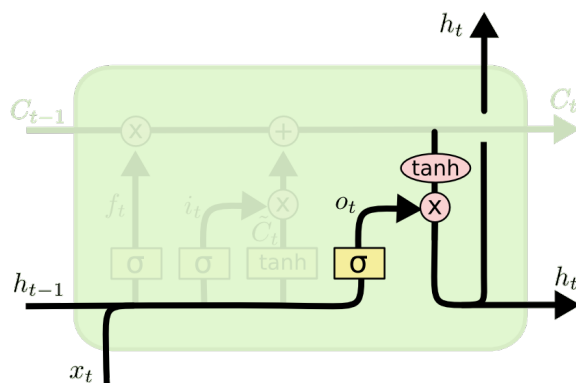
## Izlazna vrata

Na kraju, izlaz ćelije će se temeljiti na novom stanju ćelije  $\mathbf{C}_t$ . Prvo ćemo pomoću izlaznih vrata  $\mathbf{o}_t$  odrediti koje djelove stanja želimo ćemo zadržati na izlazu. Potom ćemo stanje ćelije prvo propustit kroz tangens hiperbolni prije nego ga pomnožimo s izlaznim

vratima:

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o), \quad (2.39)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{C}_t) \quad (2.40)$$



**Slika 2.12.** Izlazna vrata [12]

## **3. Implementacija**

### **3.1. Podatci**

TODO Objasnjenje odakle nam dolaze podatci, kako su procesirani te povezat s definiranim varijablama

### **3.2. Model dubokog učenja**

TODO Objasniti arhitekturu našeg modela TOOD Objasniti različite funkcije cilja

## 4. Rezultati

TODO Rezultati i rasprava

## 5. Zaključak

TODO Zaključak

## Literatura

- [1] S. Begušić i Z. Kostanjčar, “Financijska tržišta i instrumenti”, Prezentacija s kolegija Analitika financija na FER-u, 2025.
- [2] J. Y. Campbell, A. W. Lo, i A. C. MacKinlay, *The Econometrics of Financial Markets*, 2. izd. Princeton University Press, 1997.
- [3] R. S. Tsay, *Analysis of financial time series*, 3. izd. John Wiley & Sons, Inc., 2010.
- [4] N. L. Johnson, S. Kotz, i N. Balakrishnan, *Continuous Univariate Distributions*, 2. izd. John Wiley & Sons, Inc., 1994., sv. 1.
- [5] —, *Continuous Univariate Distributions*, 2. izd. John Wiley & Sons, Inc., 1994., sv. 2.
- [6] R. Cont, “Empirical properties of asset returns: stylized facts and statistical issues”, *Quantitative Finance*, sv. 1, br. 2, str. 223–236, 2001. <https://doi.org/10.1080/713665670>
- [7] M. Taboga, “Multivariate normal distribution - maximum likelihood estimation”, Lectures on probability theory and mathematical statistics. Kindle Direct Publishing. Online appendix., 2021., pristupljeno 13. 2. 2026. [Mrežno]. Adresa: <https://www.statlect.com/fundamentals-of-statistics/multivariate-normal-distribution-maximum-likelihood>
- [8] S. Begušić i Z. Kostanjčar, “Modeli povrata imovina”, Prezentacija s kolegija Analitika financija na FER-u, 2025.
- [9] R. A. Johnson i D. W. Wichern, *Applied multivariate statistical analysis*, 5. izd. Pearson Education, 2002.

- [10] J. Krapac i S. Šegivć, “Unaprijedni duboki modeli”, Prezentacija s kolegija Duboko učenje 1 na FER-u, 2025.
- [11] I. Goodfellow, Y. Bengio, i A. Courville, *Deep Learning*. MIT Press, 2016., <http://www.deeplearningbook.org>.
- [12] C. Olah, “Understanding lstm networks”, 2015., pristupljeno 13. 2. 2026. [Mrežno]. Adresa: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- [13] M. Tutek, P. Bevandić, J. Šarić, i S. Šegvić, “Jednostavni povratni modeli”, Prezentacija s kolegija Duboko učenje 1 na FER-u, 2025.

## Sažetak

**TBD**

Marko Miljković

Unesite sažetak na hrvatskom.

**Ključne riječi:** prva ključna riječ; druga ključna riječ; treća ključna riječ

## Abstract

**TBD**

Marko Miljković

Enter the abstract in English.

**Keywords:** the first keyword; the second keyword; the third keyword