

TODO Zahvala

Sadržaj

| | |
|-------------------------------------|-----------|
| 1. Uvod | 2 |
| 2. Metodologija | 3 |
| 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 | 12 |
| 2.4. LSTM Model | 12 |
| 3. Podatci | 13 |
| 4. Rezultati | 14 |
| 5. Zaključak | 15 |
| Literatura | 16 |
| Sažetak | 17 |
| Abstract | 18 |
| A: The Code | 19 |

1. Uvod

TODO Uvod.

2. Metodologija

Ovo poglavlje pokriva teoretsku osnovu potrebnu za razumijevanje implementiranog modela dubokog učenja te samu arhitekturu modela. Prvo ćemo definirati varijable pomoću kojih ćemo kasnije oblikovati naše podatke. Zatim ćemo proći osnove faktorskih modela i razmotriti kako jednofaktorski model može objasniti povrate imovina. Na kraju ćemo opisati sam model dubokog učenja, njegovu arhitekturu i funkcije cilja koje se koriste u ovom radu. !TODO Revidirat nakon ostatka poglavlja!

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 CROBEX indeksa u zadnjih 5 godina.

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



Slika 2.1. Kretanje cijene CROBEX indeksa

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 ukaćivanjem 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 CROBEX indeksa u zadnjih 5 godina. 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. Artimetrički i logaritamski povrati CROBEX indeksa

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

Višak povrata

Višak povrata (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 pri-nosi na državne obveznice visoke kreditne kvalitete i kratkog dospijeca, 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 μ i varijancom σ^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 ν , koji kontrolira debljinu repova. Za manje vrijednosti ν distribucija ima izraženije repove, čime omogućuje robusnije modeliranje ekstremnih vrijednosti.

Gustoća Studentove t distribucije sa sredinom μ , varijancom σ^2 i ν stupnjeva slobode ima oblik:

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

Zbog složenog oblika funkcije gustoće, ne postoji izravan način za izračun njenih parametara. Za procjenu parametara t distribucije koriste se metode procjene najveće izglednosti (engl. *maximum likelihood estimation*) [5].

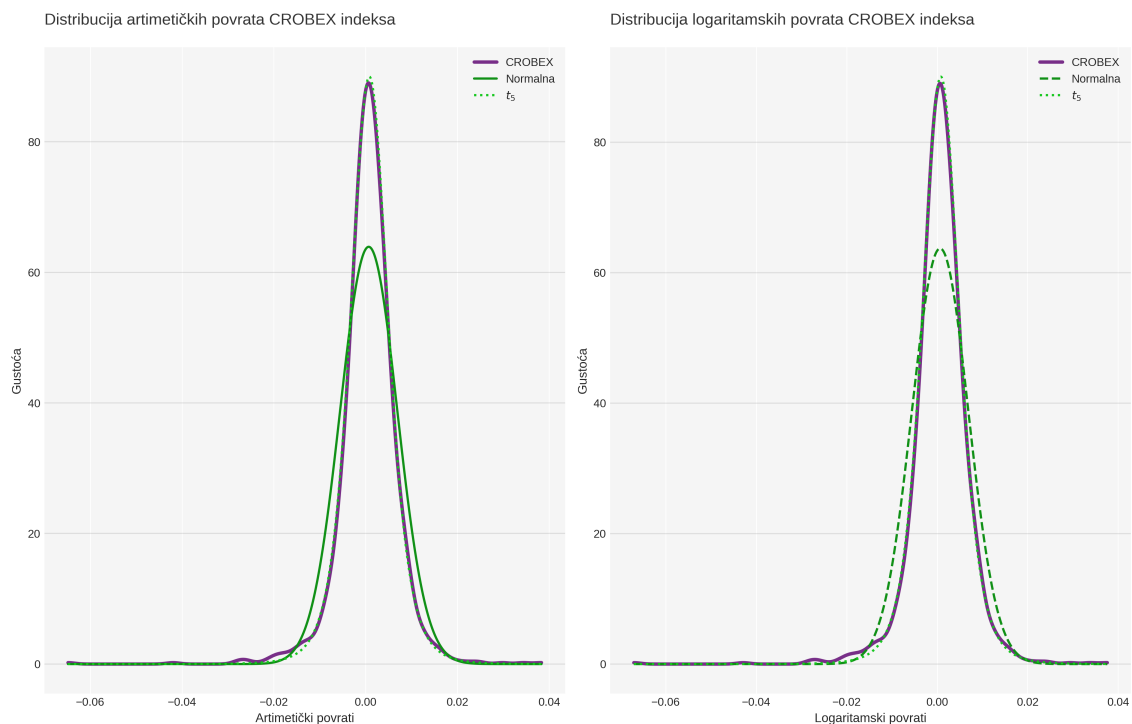
Statistička svojstva povrata

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.

Alternativno možemo pretpostaviti da su logaritamski povrati nezavisno i identično normalno distribuirani sa sredinom μ i varijancom σ^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 log-normalnosti 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 aritmetičkih i logaritamskih povrata CROBEX indeksa u zadnjih 5 godina. Također je prikazana distribucija gustoće vjerojatnosti normalne distribucije i studentove t -distribucije s pet stupnjeva slobode ¹. Parametri μ i σ^2 za normalnu i t_5 -distribuciju procijenjeni su iz aritmetičkih povrata za lijevi graf i logaritamskih povrata za desni graf. Sa slike vidimo kako t_5 -distribucija puno vijernije modelira distribuciju povrata. Također vidimo da su, zbog male magnitude dnevnih povrata, aritmetički i logaritamski povrati gotovo identično distribuirani.



Slika 2.3. Distribucija dnevnih aritmetičkih i logaritamskih povrata CROBEX indeksa

¹Biramo pet stupnjeva slobode kako bi imali konačna prva četiri momenta distribucije. Za više informacija pogledaj [5, Chapter 28.]

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 [6]:

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

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 [7].

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

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

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

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

dok je idiosinkratski faktor ϵ_{it}

2.3.2. Jednofaktorski model povrata

TODO Objašnjenje jednofaktorskog modela

Procjena parametara jednofaktorskog modela

MLE estimacija parametara pomoću OLS

2.4. LSTM Model

TODO Objasniti arhitekturu našeg modela, različiti pristupi TOOD Možda će biti potrebno raščlaniti po funkcijama cilja

3. Podatci

TODO Odakle nam podatci, kako su procesirani, povezat s definiranim varijablama

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] 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>
- [7] S. Begušić i Z. Kostanjčar, “Modeli povrata imovina”, Prezentacija s kolegija Analitika financija 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

Privitak A: The Code

TODO Appendix ako bude potrebe