

Usporedba korištenja tehnika dubokog podržanog učenja i klasične metode za optimizaciju portfelja

Stojsavljević, Marin

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:177025>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-21**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 577

**USPOREDBA KORIŠTENJA TEHNIKA DUBOKOG
PODRŽANOG UČENJA I KLASIČNE METODE ZA
OPTIMIZACIJU PORTFELJA**

Marin Stojsavljević

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 577

**USPOREDBA KORIŠTENJA TEHNIKA DUBOKOG
PODRŽANOG UČENJA I KLASIČNE METODE ZA
OPTIMIZACIJU PORTFELJA**

Marin Stojsavljević

Zagreb, lipanj 2024.

DIPLOMSKI ZADATAK br. 577

Pristupnik: **Marin Stojsavljević (0036514671)**

Studij: Računarstvo

Profil: Računarska znanost

Mentor: doc. dr. sc. Mario Brčić

Zadatak: **Usporedba korištenja tehnika dubokog podržanog učenja i klasične metode za optimizaciju portfelja**

Opis zadatka:

Optimizacija portfelja uključuje strateški odabir kombinacije imovine u investicijskom portfelju kako bi se postigla najbolja moguća ravnoteža između rizika i povrata. To je temeljni koncept u financijama čiji je cilj maksimiziranje očekivanog povrata portfelja za danu razinu rizika ili minimiziranje rizika za danu razinu povrata. Duboko podržano učenje nedavno je privuklo pažnju kao novi pristup optimizaciji portfelja, koristeći svoju sposobnost učenja optimalnih strategija iz povijesnih tržišnih podataka bez oslanjanja na eksplicitne pretpostavke o tržišnoj dinamici. Nasuprot tome, tradicionalne metode poput optimizacije portfelja srednje vrijednosti i varijance koriste statističke procjene povrata imovine i korelacije za izradu portfelja. Ovaj rad će provesti temeljitu usporedbu između metode dubokog podržanog učenja bez modela i metode optimizacije srednje vrijednosti i varijance, pružajući uvid u njihove prednosti, ograničenja i praktična razmatranja u postavkama ulaganja u stvarnom svijetu. Putem ove analize ulagači će moći bolje razumjeti kompromise između modernih tehnika strojnog učenja i utvrđenih financijskih metodologija pri optimizaciji investicijskih portfelja.

Rok za predaju rada: 28. lipnja 2024.

Sadržaj

1. Uvod	2
2. Skup podataka	4
3. Optimizacija srednje vrijednosti i varijance	10
4. Metoda podržanog učenja	17
4.1. Implementacija podržanog učenja za generiranje portfelja dionica	20
4.1.1. Akcije	20
4.1.2. Stanja	20
4.1.3. Nagrade	21
4.1.4. Algoritam učenja	23
4.1.5. Okolina	25
4.1.6. Upotreba alata za implementaciju algoritma podržanog učenja	26
4.1.7. Treniranje i testiranje modela	26
5. Rezultati i rasprava	29
6. Zaključak	51
Literatura	53
Sažetak	56
Abstract	57

1. Uvod

Upravljanje portfeljom je jedna vrlo kompleksna djelatnost i jedno je od ključnih problema u domeni financijskih usluga. Ono predstavlja raspodjelu sredstava na raznoliku imovinu, s ciljem stvaranja što većih povrata uz minimiziranje rizika i operativnih troškova. Unutar portfelja mogu se nalaziti različite klase imovine kao što su gotovina, obveznice, dionice itd.

S jedne strane, pametnom i pažljivom alokacijom udjela među klasama imovine, mogu se postići veliki profiti. Međutim, s druge strane to je vrlo rizična djelatnost, pogotovo ako ulagač pristupa ulaganju neoprezno i špekulativno. Dionice su vrlo osjetljive na krize, pogotovo one globalne, ali zato i mogu donijeti značajno veće prihode od obveznica s kojima su često i uspoređivane. Zato često rezultati ulaganja mogu biti iznenađujući i nepredvidljivi. Prve dionice su izlistane sad već davne 1792. godine na burzi u New Yorku, a danas su jedan od najpopularnijih, ali i najrizičnijih načina ulaganja. S obzirom na popularnost i količini novca koji se nalazi u dionicama, ali i drugim klasama imovine, od davnih dana su ulagači nastojali pronaći i koristiti pametne tehnike kako bi što uspješnije mogli predvidjeti buduće cijene. Jedna od najpopularnijih tradicionalnih tehnika generiranja portfelja je optimizacija portfelja srednje vrijednosti i varijance (engl. *Mean Variance Optimization, MVO*) koju je osmislio američki ekonomist Harry Markowitz 1952. godine [1]. Ona koristi povijesne informacije o vremenskim serijama za procjenu očekivanih povrata imovine, koji se zatim koriste za optimizaciju ulaganja.

U posljednjih nekoliko godina, umjetna inteligencija i njena podgrana strojno učenje su dosegli veliki napredak i popularnost. Stvari koje su prije 10-20 godina izgledale neostvarivo danas su implementirane upravo zahvaljujući umjetnoj inteligenciji. Podržano učenje je grana strojnog učenja koja na poseban način nastoji predvidjeti bu-

duća ponašanja. Ono sadrži agente koji uče u posebno dizajniranoj okolini koja imitira stvarni svijet. Nakon što se agenti utreniraju na takvoj okolini, spremni su za donošenje odluka u stvarnome svijetu. Jedan od najčešćih područja primjene podržanog učenja je autonomna vožnja. U posljednje vrijeme, podržano učenje postaje sve popularnije i u području financija, a posebno kod upravljanja portfeljom.

U ovome radu usporedit će se primjena tradicionalnih metoda za upravljanje portfeljom dionica, točnije metoda optimizacija portfelja srednje vrijednosti i varijance, s tehnikom podržanog učenja čiji se agenti uče na posebnoj okolini koja simulira američko tržište dionica. Obe metode usporedit će se nad istim skupom dionica jednih od najboljih američkih kompanija koje su dio poznatog S&P500 indeksa.

2. Skup podataka

Portfelj je skup financijske imovine koji se može sastojati od mnoštva različitih vrsta predmeta koji sadrže financijsku vrijednost poput dionica, obveznica, novčanih ekvivalenata, ali i drugih vrsta ulaganja kao što su nekretnine, zajednički fondovi i fondovi kojima se trguje na burzi - ETF-ovi (engl. *Exchange-Traded Fund*). Glavni cilj portfelja je postići određeni financijski cilj, bilo da se radi o rastu, prihodu ili očuvanju kapitala. Uz navedene ciljeve jedna od najtežih stvari na koje investitori posebno moraju paziti je rizik. Tradicionalno svaki od navedenih oblika imovina teži većem ili manjem stupnju rizika. Na primjer, dionice se često smatraju klasom imovine koja donosi veći rizik, ali uz veće potencijalne prihode, dok su obveznice s druge strane manje rizične uz prosječno manje prinose.

Dionice su vlasnički udjeli u poduzeću. Sve tvrtke kreću iz nule, međutim kada dosegnu određenu veličinu, mogu se odlučiti postati dioničko društvo. Dio udjela u vlasništvu tvrtke uzme se od osnivača, podijeli se na određen broj dionica te se izlista na burzi po određenoj cijeni po dionici. Burza je visokoorganizirano i specijalizirano tržište gdje se prema strogo utvrđenim pravilima trguje raznim robama i uslugama [2]. Ulagачi kupuju dionice na burzi s očekivanjem da će poduzeća rasti i postati vrijednija, na temelju čega se onda povećava cijena dionica. Zarada na posjedovanju dionica može se ostvariti na 2 načina - isplatom dividende ili prodajom dionice po većoj cijeni nego za koju je kupljena. Dividenda je dio dobiti dioničkog društva koji se isplaćuje vlasnicima dionica. Kada dioničko društvo ostvari dobit ono može odlučiti proslijediti dio dobiti svojim dioničarima. Dio dobiti koji se ne isplati dioničarima, ponovno se ulaže u tvrtku i zove se "zadržana dobit" [3]. Uspješna prodaja dionica zahtjeva veliku dozu znanja, iskustva, ali i samokontrole. Tržište ima tendenciju ponašati se "neracionalno". To znači da se često ne ponaša u skladu s očekivanjima investitora. Zbog toga dionice kompanije

često budu podcijenjene (koštaju malo iako tvrtka ima veliki potencijal i zdrave temelje za rast) ili precijenjene (dionice postanu popularne na tržištu te ih veliki broj investitora kupuje iako nemaju prave temelje za veliki rast u budućnosti). Postoji mnogo metoda s kojima se nastoji predvidjeti buduća cijena dionica. Jedna od najpopularnijih metoda koja se koristi za procjenu prave vrijednosti dionice je omjer cijene i zarade (engl. *price-to-earnings ratio (P/E)*). P/E omjer izračunava se dijeljenjem tržišne vrijednosti cijene po dionici sa zaradom po dionici tvrtke. Visok P/E omjer može značiti da je cijena dionice visoka u odnosu na zaradu i da je moguće precijenjena dok nizak omjer P/E označava da je trenutna cijena dionice niska u odnosu na zaradu te da je stoga dionica moguće potcijenjena. Unatoč postojanju brojnih metrika, studija i istraživanja na tu temu, veliki broj investitora ponaša se hirovito i špekulativno. Ulažu velike količine novca u dionice bez postavljanja unaprijed definirane strategije ulaganja te se nesvjesno izlažu velikom riziku da izgube dio uloženog novca.

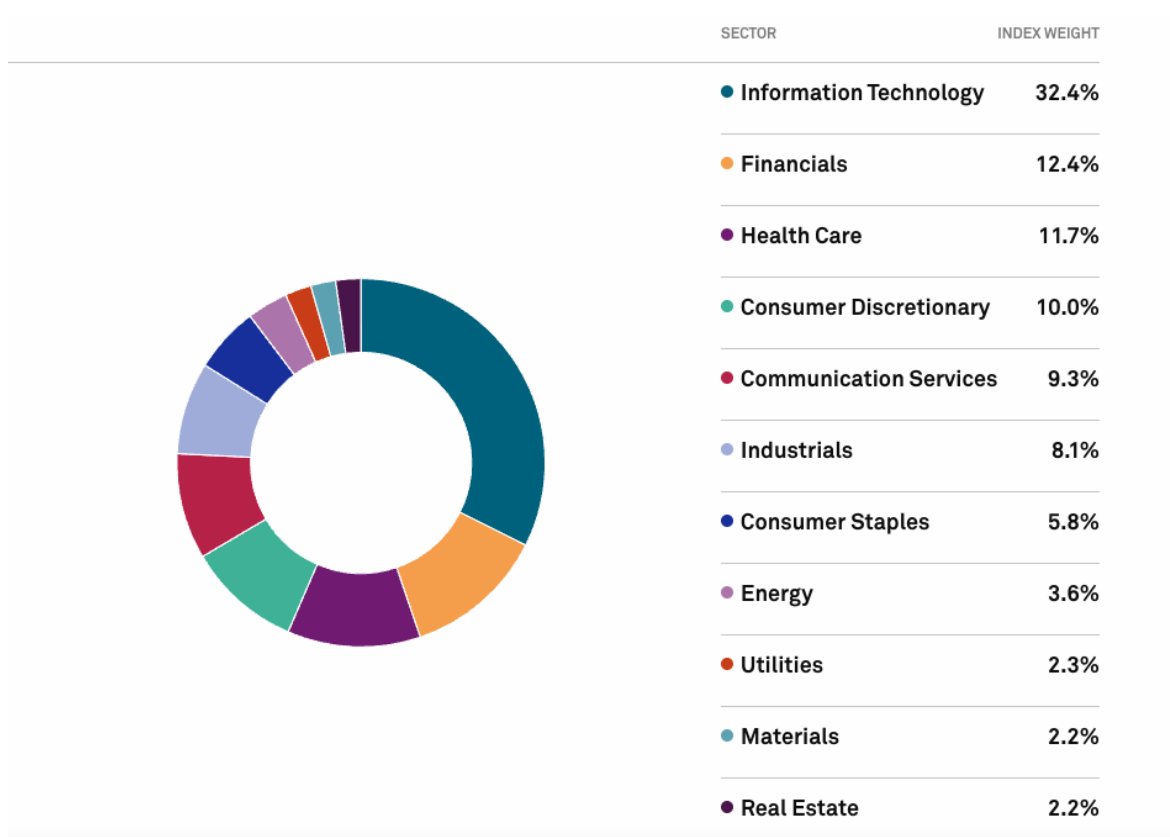
Obveznice su, s druge strane, dužnički instrumenti koje izdaju tvrtke, općine ili vlade u svrhu prikupljanja kapitala. Kada investitor kupi obveznicu, on zapravo posuđuje novac izdavatelju obveznica u zamjenu za periodične isplate kamata i povrat nominalne vrijednosti obveznice po dospeljeću. Dionička društva također mogu izdati obveznice. Razlog tomu je što na taj način mogu doći do kapitala, a da se pritom ne odreknu djela vlasništva, tj. ne mijenjaju vlasničku strukturu. Obveznice se često prikazuju kao suprotnost dionicama. Za razliku od dionica, one imaju manje stope očekivanih prinosa, ali i uz znatno manji rizik. Zato često investitori znaju imati određen dio svojih sredstava uloženi u dionice, a dio u obveznice. Omjer koliko ulažu u jednu, a koliko u drugu klasu imovine ovisi o tome koliko su spremni preuzeti rizika.

Diverzifikacija je jedan od temeljnih pojmova u financijskoj industriji, a posebno u svijetu investiranja. Ona označava raspodjelu uloženog kapitala među različitim klasama imovine, ali i unutar jedne klase imovine također označava raspodjelu po različitim metrikama kao što je sektor industrije ili geografska pripadnost. Diverzifikacija se smatra jednom od ključnih metoda za ublažavanja rizika kod ulaganja te je posebno pogodna za dugoročno investiranje. Dobro diverzificiran portfelj bio bi onaj koji ima veliki udio kapitala raspršen u različite oblike ulaganja kao što su dionice, obveznice, nekretnine, kriptovalute itd. Također, portfelj se može diverzificirati i unutar jedne klase imovine.

Na primjer, udio koji ulažemo u dionice možemo raspodijeliti tako da ulažemo u nekoliko sektora industrije, a ne samo u jedan. Na taj način ublažavamo potencijalan rizik koji bi nastao u slučaju da jedna grana industrije upadne u krizu. Npr. tijekom pandemije koronavirusa početkom 2020. godine, sektor avioindustrije posebno je bio pogođen. S obzirom na to da su diljem svijeta tisuće letova bile obustavljene na nekoliko tjedna, ulagači su zbog neizvjesnosti masovno prodavali svoje udjele u aviokompanijama stoga je njihova cijena pala drastično u samo nekoliko dana. Onaj tko je u to vrijeme imao velike udjele u tim tvrtkama, potencijalno je mogao izgubiti mnogo novaca u slučaju da je prodao svoje udjele tada. S druge strane, neke druge industrije kao što je tehnološka, nisu pretrpjeli velike gubitke. Nekim industrijama kao što je farmaceutska je to čak pogodovalo te su bilježili značajan rast cijene. Također, još jedan oblik diverzifikacije je raspodjela kapitala po geografskom području. Možemo ulagati u tvrtke iz raznih dijelove svijeta kao što su SAD, EU, Azija i ostali. Tako smanjujemo rizik da ako dođe do lokalne krize u nekoj od zemalja na čijoj smo burzi uložili dio kapitala, ne budemo previše pogođeni.

S&P 500 je najpopularniji svjetski investicijski fond. Sastoji se od dionica 500 najrazvijenijih američkih kompanija. Nastao je 1957. godine te se danas smatra referentnom točkom za usporedbu performansa s drugim dioničkim fondovima. Od svog osnutka pa do danas ima prosječan godišnji prinos od 10.28% [4] što ga u prosjeku čini profitabilnijim od velike većine drugih dioničkih fondova. Upravo zbog toga je najpopularniji svjetski fond. Dugoročno gledano osigurava velike prihode uz mali rizik s obzirom na to da odražava snagu američke ekonomije koja je već desetljećima najveća svjetska ekonomija. Naravno, kao i svaki investicijski fond S&P 500 je također podložan globalnim krizama. Najveća kriza u posljednje vrijeme, ona iz 2008. godine, rezultirala je drastičnom padu vrijednosti S&P 500 fonda kad je u jednom trenutku fond izgubio preko 50% svoje vrijednosti. Međutim, do 2013. godine uspio se oporaviti te u narednim godinama nekoliko puta probijati svoju najvišu povijesnu vrijednost [5].

S&P 500 indeks je diverzificiran kroz 11 sektora industrije. Podjela po sektorima po udjelu tržišne kapitalizacije može se vidjeti na slici 2.1. Najveći udio indeksa nalazi se u sektoru informacijske tehnologije kojeg čine neke od najpoznatijih i najuspješnijih svjetskih tvrtki kao što su Microsoft, Nvidia i Apple.



Slika 2.1. Podjela udjela u S&P 500 indeksu po sektorima [6].

U ovome radu, za potrebe usporedbe metode dubokog podržanog učenja s tradicionalnom metodom optimizacije portfelja srednje vrijednosti i varijance koristit će se portfelj koji se u potpunosti sastoji od dionica američkih kompanija koje se nalaze u S&P 500 indeksnom fondu. Međutim, portfelj neće sadržavati svih 500 tvrtki, nego samo 42 kompanije. Razlog tome je jednostavnost u rukovanju manjim portfeljem. Dionice koje se nalaze u portfelju raspodijeljene su u 11 sektora industrije - to su isti oni sektori iz indeksa S&P 500. U tablici 2.1. možemo vidjeti popis kompanija koje se nalaze u portfelju zajedno sa sektorom industrije kojem pripada svaka od kompanija. Najveći broj dionica nalazi se u tehnološkom sektoru - njih sedam. Po četiri kompanije nalaze se u sektorima diskrecijske potrošnje, financija, zdravstva, komunikacijskih usluga te energije. U ostalim sektorima - industriji, materijalima, nediskrecijskoj potrošnji, nekretninama te komunalnim uslugama nalazi se po tri tvrtke.

Dionica	Ime tvrtke	Sektor
TSLA	Tesla Inc	Diskrecijska potrošnja
AMZN	Amazon.com Inc	Diskrecijska potrošnja
F	Ford Moto	Diskrecijska potrošnja
CCL	Carnival Corp	Diskrecijska potrošnja
INTC	Intel Corp	Tehnologija
CSCO	Cisco Systems Inc	Tehnologija
MU	Micron Technology	Tehnologija
MSFT	Microsoft Corp	Tehnologija
AMD	Advanced Micro Devices Inc	Tehnologija
AAPL	Apple Inc	Tehnologija
NVDA	NVIDIA Corp	Tehnologija
LUV	Southwest Airlines Company	Industrija
DAL	Delta Air Lines Inc	Industrija
CAT	Caterpillar Inc	Industrija
MS	Morgan Stanley	Financije
WFC	Wells Fargo & Company	Financije
MA	Mastercard Inc	Financije
JPM	JP Morgan Chase & Company	Financije
PFE	Pfizer Inc	Zdravstvo
DXCM	DexCom Inc	Zdravstvo
EW	Edwards Lifesciences Corp	Zdravstvo
WST	West Pharmaceutical Services Inc	Zdravstvo
SHW	Sherwin-Williams Co	Materijali
FCX	Freeport-Mcmoran Inc	Materijali
NEM	Newmont Mining Corp	Materijali
MNST	Monster Beverage Corp	Nediskrecijska potrošnja
KO	Coca-Cola Company	Nediskrecijska potrošnja
WAL	Walmart Inc	Nediskrecijska potrošnja
NFLX	Netflix Inc	Komunikacijske usluge
GOOGL	Alphabet Cl A	Komunikacijske usluge
META	Meta Platforms Inc	Komunikacijske usluge
EA	Electronic Arts In	Komunikacijske usluge
KMI	Kinder Morgan	Energija
XOM	Exxon Mobil Corp	Energija
HAL	Halliburton Company	Energija
SLB	Schlumberger N.V.	Energija
HST	Host Marriott Financial Trust	Nekretnine
O	Realty Income Corp	Nekretnine
WY	Weyerhaeuser Company	Nekretnine
D	Dominion Energy Inc	Komunalne usluge
NEE	Nextera Energy	Komunalne usluge
CNP	Centerpoint Energy Inc	Komunalne usluge

Tablica 2.1. Popis dionica sa sektorom kojem pripadaju

Za ovaj rad, koristit će se povijesne cijene navedenih dionica u periodu od 1. siječnja 2015. godine do 31. prosinca 2023. godine. Cijene su preuzete s Yahoo finance-a [7] pomoću yfinance Python programske knjižice [8].

3. Optimizacija srednje vrijednosti i varijance

Optimizacija srednje vrijednosti i varijance (engl. *Mean Variance Optimization, MVO*) je matematička metoda raspodjele kapitala u portfelju imovine u svrhu postizanja željenog cilja ulaganja. Postoje 3 moguća cilja ulaganja na temelju kojih postavljamo optimizaciju. To su:

1. Maksimiziranje povrata za danu razinu rizika
2. Postizanje željene stope povrata uz minimiziranje rizika
3. Maksimiziranje povrata generiranih po jedinici rizika

Rizik se obično mjeri volatilnošću portfelja, što zapravo predstavlja varijancu njegove stope povrata. Za dani skup imovine, ovaj proces zahtijeva kao ulazne podatke stope povrata za svaku imovinu, zajedno s njihovim kovarijancijama. Budući da su pravi povrti imovine nepoznati, u praksi se oni procjenjuju ili prognoziraju s pomoću različitih tehnika koje koriste povijesne podatke. Ovaj zadatak se zatim oblikuje kao problem optimizacije, s jednim ili više ciljeva, koji se može riješiti na različite načine [9].

Tipičan postupak rješavanja takvog optimizacijskog problema je da ga postavimo kao problem konveksne optimizacije i generiramo učinkovitu granicu portfelja tako da se niti jedan portfelj ne može poboljšati bez žrtvovanja neke od ključnih metrika - povrata ili rizika. Konveksni optimizacijski problemi su predmeti brojnih matematičkih istraživanja te su stoga poprilično teoretski potkovani. Tipičan konveksni optimizacijski problem se postavlja ovako 3.1

$$\begin{aligned}
& \min_{\mathbf{x}} f(\mathbf{x}) \\
& \text{uz ograničenja } g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\
& Ax = b, \\
& \mathbf{x} \in \mathbb{R}^n, \\
& f(\mathbf{x}), g_i(\mathbf{x}) \text{ konveksne funkcije}
\end{aligned} \tag{3.1}$$

Neka je w vektor težine za skup imovine, μ su očekivani prinosi, rizik portfelja može se opisati kao $w^T \Sigma w$, za matricu kovarijance Σ . Da bismo postigli željenu stopu povrata μ^* , možemo riješiti problem optimizacije portfelja kao što je prikazano na 3.2

$$\begin{aligned}
& \min_w w^T \Sigma w \\
& \text{uz ograničenja } w^T \mu \geq \mu^*, \\
& w_i \geq 0, \\
& \sum_i w_i = 1
\end{aligned} \tag{3.2}$$

μ^* predstavlja granicu portfelja takvu da se portfelj ne može poboljšati bez pogoršanja povrata ili rizika. Još jedna vrlo često upotrebljavana funkcija cilja koja se koristi za rješavanje optimizacije portfelja srednje vrijednosti i varijance je Sharpeov omjer [10], koji mjeri povrat po jedinici rizika. Sharpeov omjer opisuje koliki se višak povrata dobije za svaku dodatnu jedinicu rizika koja se preuzme preko držanja rizičnije imovine. Veći omjer podrazumijeva veći povrat ulaganja u usporedbi s iznosom rizika ulaganja. To je jedna od najčešće korištenih mjera rizika i prinosa koja se koristi u financijama, dijelom zbog svoje jednostavnosti. Sharpeov omjer izračunava se oduzimanjem nerizične stope povrata od očekivane stope povrata, zatim dijeljenjem dobivene brojke sa standardnom devijacijom [11].

Formalno, za portfelj p , Sharpeov omjer definira se po formuli 3.3

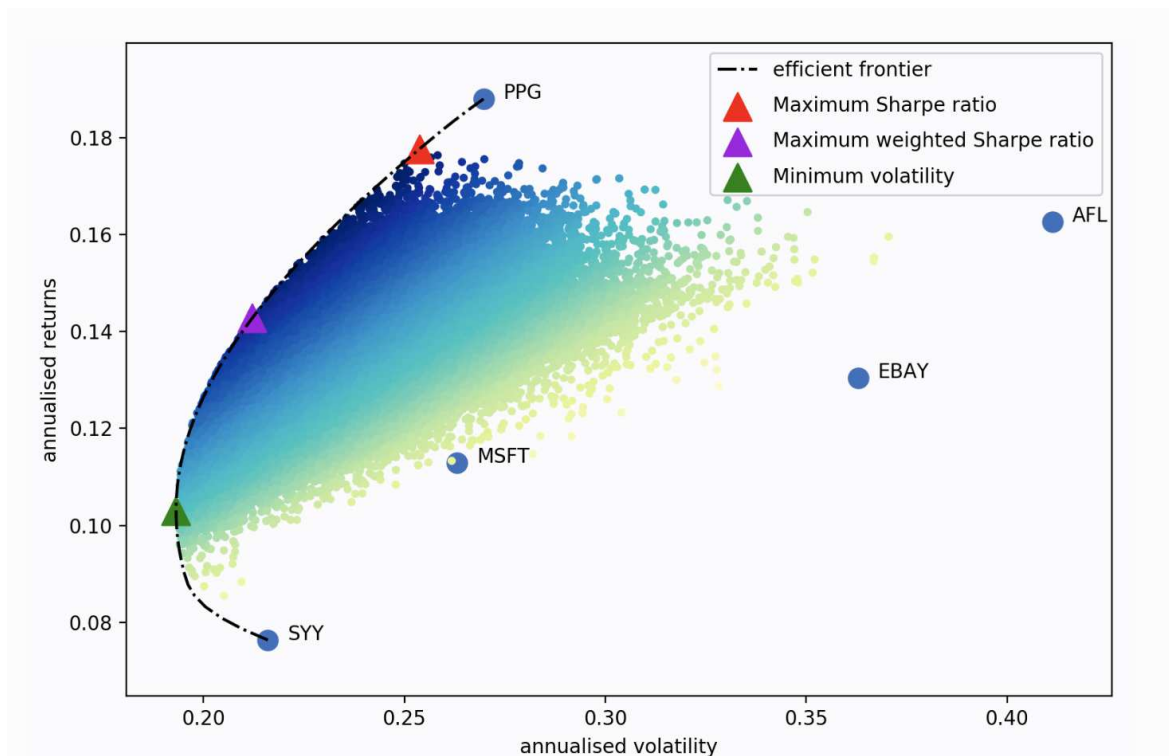
$$\text{Sharpeov omjer}_p = \frac{E[R_p - R_f]}{\sigma_p} \quad (3.3)$$

R_p su prinosi portfelja, σ_p je standardna devijacija tih prinosa, a R_f je konstantna stopa bez rizika. Iako ga je teško riješiti u izravnom obliku 3.4, može se oblikovati kao problem konveksne optimizacije korištenjem supstitucije varijable [12].

$$\max_w \frac{\mu^T w - R_f}{(w^T \Sigma w)^{1/2}} \quad (3.4)$$

Ako mijenjamo ciljani povrat iz formule 3.2, dobit ćemo drugačiji skup težina, odnosno drugačiji portfelj. Skup svih tih optimalnih portfelja naziva se učinkovitom granicom. Svaka točka na dijagramu prikazanom na slici 3.1. (koja je preuzeta s PyPortfolioOpt službene stranice [13]) predstavlja drugačiji mogući portfelj. Tamnija plava boja označava "bolje" portfelje s obzirom na Sharpeov omjer. Isprekidana crna linija predstavlja učinkovitu granicu. Trokutaste oznake na dijagramu predstavljaju najbolje portfelje za različite ciljeve optimizacije. S obzirom na to da je s x-osi predstavljena godišnja volatilitnost, a s y-osi godišnji povrati trokutaste oznake predstavljaju sljedeće stvari:

- Crvena oznaka predstavlja najbolji portfelj s obzirom na Sharpeov omjer
- Zelena oznaka predstavlja najbolji portfelj s obzirom na minimalnu volatilitnost
- Ljubičasta oznaka predstavlja najbolji portfelj s obzirom na utežani Sharpeov omjer



Slika 3.1. Podjela udjela u S&P 500 indeksu po sektorima [6].

U ovome radu optimizacija srednje vrijednosti i varijance koristit će upravo Sharpeov omjer kao ciljnu funkciju jer omogućuje optimizaciju povrata portfelja prilagođenih riziku, a da ne moramo specificirati eksplicitne brojke za minimalne očekivane povrate ili maksimalnu toleranciju rizika.

Za implementaciju metode optimizacije srednje vrijednosti i varijance, koristi se programska biblioteka PyPortfolioOpt [13] za programski jezik Python [14]. Optimizacija srednje vrijednosti i varijance zahtijeva dvije stvari - očekivane povrate imovine (μ iz formule 3.2) i matricu kovarijance (Σ iz formule 3.2) odnosno, model rizika koji kvantificira rizik imovine. PyPortfolioOpt pruža metode za procjenu obje stvari. Očekivani povrati imovine modeliraju se s pomoću modula `expected_returns`, dok se rizik imovine modelira modulom `risk_models`. Uz postojeće modele za spomenute stvari, PyPortfolioOpt također podržava i korisnike koji bi željeli koristiti svoje vlastite modele.

Očekivane povrate je u praksi prilično teško sa sigurnošću znati. Stoga, najbolje što možemo učiniti je doći do procjena, na primjer ekstrapolacijom povijesnih podataka. Upravo ta procjena očekivanih budućih povrata predstavlja najčešći uzrok grešaka u optimizaciji srednje varijance. Postupak optimizacije je dobar i pruža snažna matematička

jamstva, ali samo u slučaju točnih ulaznih podataka. Upravo zbog toga, procjena očekivanih povrata ponekad može učiniti više štete nego koristi. Za procjenu očekivanih povrata za metodu optimizacije portfelja srednje vrijednosti i varijance u ovome radu koristi se metoda srednjih povijesnih povrata preko metode `expected_returns.mean_historical_return` koja je već implementirana u programskoj knjižici `PyPortfolioOpt`. Kao ulaz metode predaje se skup povijesnih cijena dionica za određeni vremenski period. Navedena metoda koristi geometrijsku sredinu za izračun.

Uz očekivane povrate, optimizacija srednje vrijednosti i varijance zahtijeva i model rizika kako bi mogla kvantificirati rizik imovine. Najčešće korišteni model rizika je matrica kovarijance. Ta matrica opisuje volatilitnost imovine te njihovu međuovisnost. To je vrlo važno jer je jedno od načela diverzifikacije da se rizik može smanjiti stvaranjem mnogo nekoreliranih procjena - korelacija je samo normalizirana kovarijanca. Na mnogo načina, predmet modela rizika daleko je važniji od očekivanih prinosa jer je povijesna varijanca općenito mnogo dosljednija statistika od srednjih povijesnih prinosa. Problem je, međutim, u tome što u praksi nemamo pristup matrici kovarijance, na isti način kao što nemamo ni pristup očekivanim prinosima. Jedino što možemo napraviti u ovom slučaju je kreirati procjene na temelju povijesnih podataka. Najjednostavniji pristup je izračunavanje matrice kovarijance uzorka na temelju povijesnih povrata, ali relativno novija istraživanja pokazuju da postoje mnogo robusniji statistički procjenitelji matrice kovarijance.

Nepriistrana, ali često loše procijenjena kovarijanca uzorka može se kombinirati sa strukturiranim procjeniteljem F koristeći formulu 3.5 gdje δ označava konstantu skupljanja:

$$\hat{\Sigma} = \delta F + (1 - \delta)S \quad (3.5)$$

Ovaj postupak naziva se skupljanjem jer se može smatrati "skupljanjem" matrice kovarijance uzorka prema drugom procjenitelju, što se u skladu s tim naziva cilj skupljanja koji može biti znatno pristran, ali uz malu pogrešku procjene [15]. Postoji mnogo mogu-

ćih opcija za cilj, a svaka će rezultirati različitom optimalnom konstantom skupljanja δ . PyPortfolioOpt nudi nekoliko metoda skupljanja pomoću metode `risk_models.CovarianceShrinkage`, a za ovaj rad koristit će se Ledoit-Wolf metoda skupljanja. Postoji nekoliko varijacija Ledoit-Wolf metode, a ovdje se koristi smanjenje konstantne varijance gdje je za cilj postavljena dijagonalna matrica sa srednjom vrijednosti varijanci imovine na dijagonalama te nulama postavljenim na svim ostalim pozicijama u matrici. Rezultantnu matricu kovarijacije postavljamo da bude pozitivna semi-definitna uz postavljanje negativnih vlastitih vrijednosti na nulu te ponovnu izgradnju matrice.

Kako bi se izgenerirao skup optimalnih portfelja, koristi se metoda `efficient_frontier.EfficientFrontier` iz PyPortfolioOpt programske knjižice. Toj metodi predajemo matricu kovarijacije i očekivane povrate dionica te isto tako definiramo raspon težina udjela pojedine dionice u portfelju. On je ograničen između 0 i 1 gdje 0 označava da u portfelju ne sadržimo nikakvih udjela u toj dionici dok 1 označava da se cijeli naš investirani udio nalazi u toj jednoj dionici. Također, s pomoću funkcije `efficient_frontier.EfficientFrontier` možemo definirati i funkcije ograničenja. Jedna takva funkcija koja je ovdje korištena je L2 regularizacija koja služi kako bi se portfelj više diverzificirao. Tako se povećava broj dionica u kojima postoji udio uložene novca te se također jednoličnije raspodjeljuju resursi. Regularizacija se kontrolira parametrom γ . Što je veći navedeni parametar, veća je regularizacija te će vrijednosti biti raspršene ravnopravnije među dionicama, dok manji iznos γ parametra označava da će veća količina biti investirana u manji broj dionica. Kao funkciju cilja postavljamo već spominjani Sharpeov omjer kojeg nastojimo maksimizirati.

Nakon što dobijemo željene udjele u portfelju za pojedini dan, moramo se pobrinuti da portfelj sadrži udio u dionicama u skladu s njihovim cijenama. Ono što mi dobijemo rješavanjem optimizacijskog problema je skup udjela u dionicama u obliku težina koji su ograničeni na raspon $[0, 1]$. Te težine moramo prilagoditi stvarnim cijenama dionica u datom trenutku. Na primjer, ako je cijena dionice tvrtke Apple u danu 70\$, dok je naš ukupan budžet 1000\$ uz dobivenu težinu od 0.1 (10%) za tu dionicu nakon izračuna, moramo odlučiti hoćemo li sadržavati jednu dionicu Applea u našem portfelju ($1 * 70\$$) ili dvije ($2 * 70\$$). Možemo kupovati samo cjelobrojne vrijednosti dionica.

Kako bi riješili navedeni problem, upotrebljavala se metoda `DiscreteAllocation.greedy_portfolio()` iz programske knjižice `PyPortfolioOpt` [13]. Metoda se izvršava u dvije runde. U prvoj rundi kupujemo onoliko dionica koliko možemo za svaku dionicu bez prekoračenja željene težine. U primjeru Applea, $100 / 70 = 1,43$, stoga kupujemo 1 dionicu po cijeni od 70 \$. Nakon ponavljanja istog postupka kroz sve dostupne dionice, ostat će nam višak novca budući da smo uvijek zaokruživali na manje.

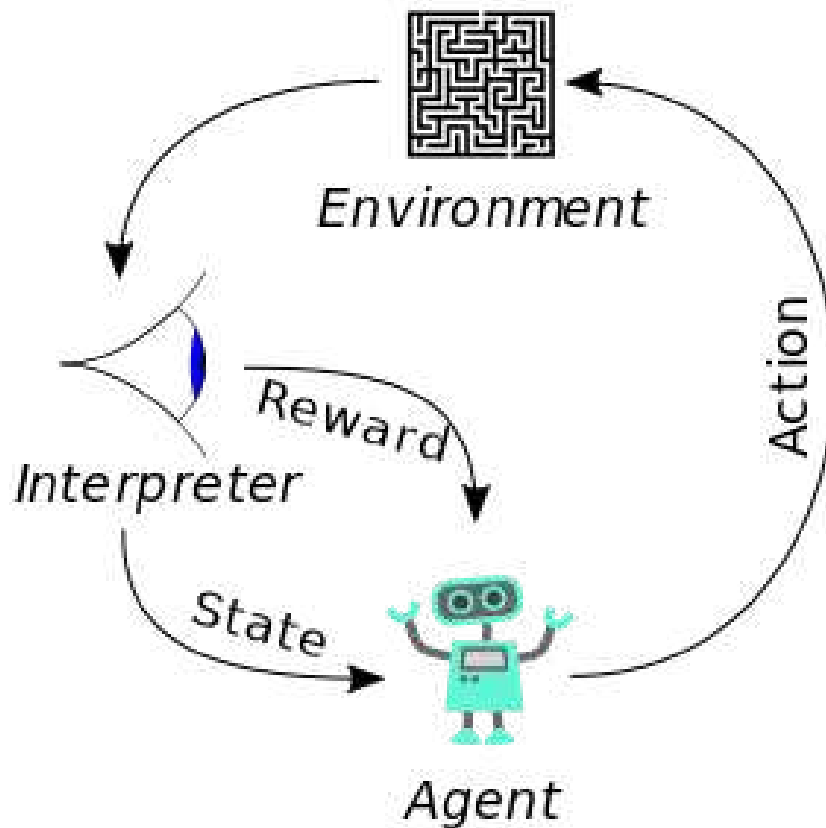
U drugom krugu izračunavamo koliko trenutne težine odstupaju od originalnih težina za svaku dionicu. Željeli smo da Apple čini 10% portfelja, ali kupili smo samo 70\$ vrijedne dionice Applea, tako da postoji odstupanje od $0,1 - 0,07 = 0,03$. Neke dionice će imati veće odstupanje od željene težine pa ćemo prvo otkupiti upravo njih. Zatim ponavljamo postupak, uvijek kupujući dionice imovine čija je trenutna težina najudaljenija od idealne težine. Iako ovaj algoritam ne jamči optimalno rješenje, predstavlja vrlo zgodnu metodu za rješavanje problema raspodjele udjela u dionice na temelju težina.

4. Metoda podržanog učenja

Podržano učenje (engl. *Reinforcement Learning, RL*) je područje strojnog učenja koje se bavi proučavanjem i razvojem načina na koji inteligentni agent poduzima radnje u dinamičnom okruženju kako bi maksimizirao kumulativnu nagradu prelazeći između stanja okoline. Takvo učenje jedna je od tri osnovne paradigme strojnog učenja, uz nadzirano učenje (engl. *supervised learning*) i učenje bez nadzora (engl. *unsupervised learning*) [16].

Podržano učenje razlikuje se od učenja pod nadzorom po tome što ne sadrži oznake za svaku stavku u skupa podataka za treniranje i testiranje te po tome što ne treba eksplicitno ispravljati radnje koje nisu optimalne. Umjesto toga, fokus je na pronalaženju ravnoteže između istraživanja neistraženog teritorija i iskorištavanja trenutačnog znanja s ciljem maksimiziranja dugoročne nagrade [17].

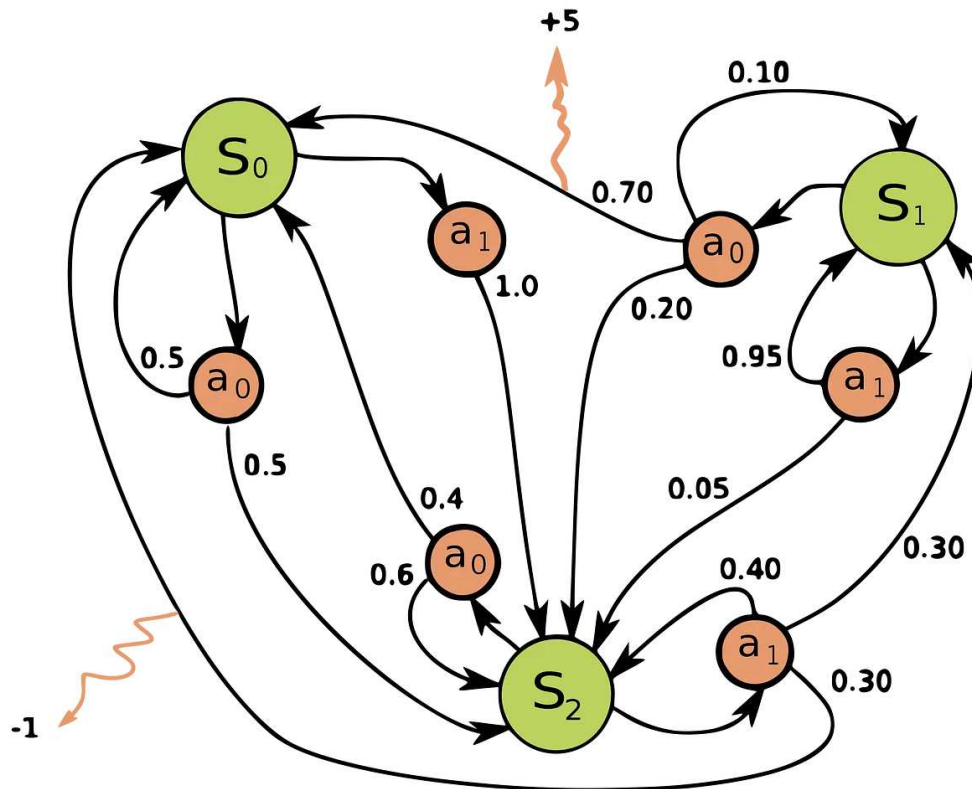
Zbog svoje općenitosti i robusnosti, podržano učenje proučava se u mnogim područjima, kao što su teorija igara, teorija kontrole, operacijska istraživanja, teorija informacija, optimizacija temeljena na simulaciji, sustavi s više agenata, inteligencija roja i statistika. U područjima operacijskog istraživanja i kontrole, podržano učenje naziva se aproksimativno dinamičko programiranje ili neurodinamičko programiranje. Problemi od interesa za podržano učenje također su proučavani u teoriji optimalne kontrole, koja se uglavnom bavi postojanjem i karakterizacijom optimalnih rješenja i algoritmima za njihovo točno izračunavanje, a manje učenjem ili aproksimacijom, osobito u nedostatku matematičkog modela okoline [17].



Slika 4.1. Tipičan prikaz podržanog učenja gdje agent izvršava akcije u okruženju za koje dobije nagradu te na temelju toga prelazi iz jednog stanja u drugo. [18].

Okruženje ili okolina koja se koristi za učenje ponašanja agenta u podržanom učenju se obično formalizira s pomoću Markovljevog procesa odlučivanja (engl. *Markov Decision Process, MDP*). Markovljev proces odlučivanja se sastoji od četvorke (S, A, P_a, R_a) , u kojoj:

- S označava skup stanja
- A je skup akcija
- $P_a(s, s') = \Pr(S_{t+1} = s' \mid S_t = s, A_t = a)$ je vjerojatnost da će akcija a u stanju s u trenutku t dovesti do stanja s' u trenutku $t + 1$
- $R_a(s, s')$ je neposredna nagrada primljena za prelazak iz stanja s u stanje s' , kao posljedica akcije a



Slika 4.2. Primjer jednostavnog Markovljevog procesa odlučivanja s tri stanja (zeleni krugovi) i dvije akcije (narančasti krugovi), s dvije nagrade (narančaste strelice) [19].

Rješenje za Markovljev proces odlučivanja je politika π koja specificira radnje $\pi(e)$ koje će donositelj odluke izabrati kada je u stanju s . Cilj je odabrati politiku π koja će maksimizirati očekivani diskontirani zbroj nagrada u potencijalno beskonačnom horizontu, što je prikazano formulom 4.1 [20].

$$E \left[\sum_{t=0}^{\infty} \gamma^t R_{a_t}(s_t, s_{t+1}) \right] \quad (4.1)$$

Područje dubokog podržanog učenja (engl. *Deep reinforcement Learning, DRL*) iskorištava potencijal dubokog učenja korištenjem neuronskih mreža kao aproksimatora funkcija za procjenu vrijednosti odnosa stanja i akcija ili za učenje preslikavanja politika π . Ove su tehnike postigle veliki uspjeh u industriji igara, robotici, kontinuiranoj kontroli i

financijama. Upravo su financije ovdje glavni fokus i to je jedan od glavnih razloga zašto je podržano učenje odabrano za generiranje portfelja dionica.

4.1. Implementacija podržanog učenja za generiranje portfelja dionica

Za potrebe ovoga rada, implementirana je metoda podržanog učenja čija je zadaća naučiti agenta da što bolje može izgenerirati portfelj dionica. Kako bi agent mogao učiti, potrebno mu je pružiti povijesne cijene dionica i omogućiti mu da donosi akcije an temelju stanja u kojem se nalazi. Cilj ove implementacije je imati agenta podržanog učenja koji može donositi kvalitetne odluke nad okolinom koja se sastoji od podataka koje agent nije imao na raspolaganju prilikom treniranja. Implementacija izrađena za potrebe ovog rada inspirirana je implementacijom opisanom u radu [9] uz određene preinake.

4.1.1. Akcije

Akcije su modelirane kao vektor dimenzije n , gdje n predstavlja broj dionica koje su na raspolaganju agentu za ulaganje. Agent može i ne mora imati udio u pojedinoj dionici. Formalno, to možemo zapisati kao vektor težina $\mathbf{w} = [w_1, \dots, w_n]$, gdje vrijedi $\sum_{i=1}^N w_i = 1$ i $0 \leq w_i \leq 1$. Ako pojedina dionica i ima težinu $w_i = 0$ u vektoru težina \mathbf{w} , to označava da agent nema nikakvih uloga u toj dionici. Ako s druge strane je $w_i = 1$, to označava da je cjelokupni udio raspoloživog novca investiran u tu jednu dionicu i .

4.1.2. Stanja

Cijenu dionice u trenutku t možemo označiti s P_t . Jednostavni povrat dionice za jedan period definiramo kao $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$. Zatim, to možemo preformulirati kao $\frac{P_t}{P_{t-1}} = R_t + 1$. Logaritam povrata za jedan period označavamo formulom 4.2

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(R_t + 1). \quad (4.2)$$

U ovome slučaju jedan period označava jedan dan. Za modeliranje stanja uzimamo

logaritam povrata na temelju cijena dionica na kraju dana. Logaritam povrata dionice tijekom perioda retrospektive T možemo preformulirati kao $r_t = [r_{t-1}, r_{t-2}, \dots, r_{t-T}]$. Za potrebe ovog rada, koristit ćemo period retrospektive od $T = 60$ dana. Stanje će sadržavati n klasa imovine koje predstavljaju n dionica kompanija koja su dostupna za ulaganje. Jedno stanje okoline u trenutku t definiramo kao matricu dimenzije $[n \times (T + 1)]$ kao što je prikazano na formuli 4.3

$$S_t = \begin{bmatrix} w_1 & r_{1,t-1} & \cdots & r_{1,t-T} \\ w_2 & r_{2,t-1} & \cdots & r_{2,t-T} \\ \vdots & \vdots & \ddots & \vdots \\ w_n & r_{n,t-1} & \cdots & r_{n,t-T} \end{bmatrix} \quad (4.3)$$

Prvi stupac matrice 4.3 označava vektor težina w u trenutku t .

4.1.3. Nagrade

S obzirom na to da koristimo tehniku podržanog učenja za raspodjelu portfelja, želimo funkciju nagrađivanja koja pomaže u optimizaciji za povrate prilagođene riziku. Sharpeov omjer je najčešće korištena mjera za to, međutim, nije prikladan za postavke online učenja budući da je definiran tijekom vremenskog razdoblja T . Kako bi se riješio navedeni problem, koristi se diferencijalni Sharpeov omjer D_t [21] koji predstavlja povrate prilagođene riziku u svakom vremenskom koraku t i za koji je utvrđeno da daje dosljednije povrate od maksimiziranja profita [21]. Stoga agent kojemu je cilj maksimizirati svoje buduće nagrade modelirane pomoću diferencijabilnog Sharpeovog omjera uči kako optimizirati za povrate prilagođene riziku.

Sharpeov omjer možemo definirati tijekom razdoblja od t povrata R_t , s pomoću procjena prvog i drugog trenutka distribucije povrata kao što prikazuje 4.4 formula.

$$S_t = \frac{A_t}{K_t(B_t - A_t^2)^{1/2}} \quad (4.4)$$

Pojedini elementi formule 4.4 definirani su kao 4.5

$$\begin{aligned}
 A_t &= \frac{1}{t} \sum_{i=1}^t R_i \\
 B_t &= \frac{1}{t} \sum_{i=1}^t R_i^2 \\
 K_t &= \left(\frac{t}{t-1} \right)^{1/2}
 \end{aligned} \tag{4.5}$$

K_t označava faktor normalizacije. A i B se mogu rekurzivno procijeniti pomoću vremenskog normalizacijskog faktora η . Diferencijabilni Sharpeov omjer možemo dobiti proširenjem S_t na prvi red u η što je prikazano formulom 4.6

$$S_t \approx S_{t-1} + \eta D_t|_{\eta=0} + O(\eta^2) \tag{4.6}$$

Diferencijabilni Sharpeov omjer D_t možemo dobiti formulom označenom na 4.7

$$D_t = \frac{\partial S_t}{\partial \eta} = \frac{B_{t-1} \Delta A_t - \frac{1}{2} A_{t-1} \Delta B_t}{(B_{t-1} - A_{t-1}^2)^{3/2}} \tag{4.7}$$

Pojedini elementi formule 4.7 definirani su kao 4.8

$$\begin{aligned}
 A_t &= A_{t-1} + \eta \Delta A_t \\
 B_t &= B_{t-1} + \eta \Delta B_t \\
 \Delta A_t &= R_t - A_{t-1} \\
 \Delta B_t &= R_t^2 - B_{t-1}
 \end{aligned} \tag{4.8}$$

U formuli 4.8 inicijaliziraju se $A_0 = B_0 = 0$. Parametar η odabran je kao $\eta \approx 252$ zato što jedna godina ima otprilike 252 radna dana tijekom kojih se može trgovati dionicama.

4.1.4. Algoritam učenja

Algoritmi podržanog učenja mogu se podijeliti u dvije kategorije:

1. algoritmi koji se temelje na modelu (engl. *model-based*)
2. algoritmi bez modela (engl. *model-free*)

Kod algoritama koji se temelje na modelu, agent podržanog učenja ima pristup modelu okoline, dok kod algoritama bez modela agent mora naučiti donositi odluke unutar okoline samostalno bez modela. Algoritmi bez modela nastoje naučiti ishode svojih radnji kroz prikupljanje iskustva putem algoritama kao što su politika gradijenta (engl. *Policy Gradient*), Q-učenje (engl. *Q-Learning*), itd. Takav algoritam će pokušati akciju više puta i prilagoditi svoju politiku ili strategiju na temelju ishoda njegovog djelovanja nad okolinom kako bi se optimizirale nagrade koje dobije za svoje postupke. Razliku između algoritama koji se temelje na modelu i onih bez modela možemo objasniti i na sljedeći način. Ako agent može predvidjeti nagradu za neku akciju prije nego što je stvarno izvede i time planira što treba učiniti, algoritam se temelji na modelu. S druge strane, ako agent stvarno treba provesti radnju da vidi što se događa i uči iz toga, onda nema modela.

Metode optimizacije politike usredotočene su na politiku π koja je funkcija koja preslikava agentova stanja u akcije koje će poduzimati s obzirom na stanja u kojima se nalazi, a jedne od najčešćih su metode gradijenta politike koje često igraju vitalnu ulogu u optimizaciji politike. Kada agent uči i zapravo ne zna koje radnje daju najbolji rezultat u odgovarajućim stanjima, on to čini izračunavanjem gradijenata politike. Jedna od najpopularnijih takvih optimizacijskih metoda je PPO (engl. *Proximal Policy Optimization*) [22]. U usporedbi s drugim algoritmima, tri glavne prednosti PPO-a su jednostavnost, stabilnost i učinkovitost uzorka. PPO je klasificiran kao metoda gradijenta politike za obuku agentove mreže politika. Za treniranje prave mreže pravila, PPO poduzima malo ažuriranje pravila u smislu veličine koraka, tako da agent može pouzdano doći do optimalnog rješenja. Prevelik korak može usmjeriti politiku u pogrešnom smjeru, stoga ima male mogućnosti oporavka, dok premali korak smanjuje ukupnu učinkovitost. Poslje-

dično, PPO implementira funkciju isječka koja sprječava da ažuriranje politike agenta bude preveliko ili premalo [22].

Ukratko, PPO algoritam funkcionira na sljedeći način [23]:

1. Metode gradijenta politike: PPO se temelji na metodama gradijenta politike, koje izravno optimiziraju funkciju politike koja preslikava stanja u akcije. Ovo je u surotnosti s metodama koje procjenjuju funkcije vrijednosti.
2. Ciljna funkcija: PPO ima za cilj maksimizirati očekivanu kumulativnu nagradu dobivenu interakcijom s okolinom. To se obično postiže maksimiziranjem surogat funkcije cilja koja aproksimira poboljšanje politike.
3. Odrezani zamjenski (surogat) cilj: Jedna od ključnih značajki PPO-a je odrezani zamjenski cilj, koji sprječava velika ažuriranja pravila koja bi mogla dovesti do katastrofalnih ishoda. Smanjujući omjer između vjerojatnosti radnji prema novoj i staroj politici, PPO osigurava da ažuriranje politike ostane unutar sigurnog raspona.
4. Više epoha i mini-serijskih ažuriranja: PPO obično uključuje više epoha interakcije s okolinom, tijekom kojih se putanje prikupljaju. Te se putanje zatim koriste za izračun zamjenske (surogat) funkcije cilja, koja se optimizira korištenjem mini-serijskih ažuriranja.
5. Procjena funkcije vrijednosti: PPO često uključuje procjenu funkcije vrijednosti kako bi se smanjila varijanca u procjenama gradijenta. To pomaže stabilizirati trening i poboljšati učinkovitost uzorka.
6. Paralelizacija: PPO se može paralelizirati kako bi se ubrzao trening prikupljanjem putanja iz više instanci okruženja istovremeno.

Za implementaciju algoritma podržanog učenja u ovom radu, koristi se upravo PPO metoda.

4.1.5. Okolina

Okolina nad kojom se agent uči djelovanju sadrži sve povijesne cijene dionica potrebne za učenje agenta. Kako bi modelirali stanja u kojima se agent može nalaziti, bilo je potrebno izračunati logaritamske povrate završnih cijena dana u kojima se moglo trgovati dionicama.

Okruženje služi kao imitacija tržišta na temelju povijesnih cijena dionica. Isto tako, služi i kao posrednik i mjenjačnica - u svakom vremenskom koraku obrađuje akcije agenta i rebalansira portfelj koristeći najnovije cijene te izračunatu novu alokaciju portfelja. Okolina je zadužena da obavijesti agenta o promjeni dana te ga obavještava o novim cijenama dionica oblikovanih u obliku stanja. Također, okolina je dužna obavijestiti agenta o nagradi u obliku diferencijabilnog Sharpeovog omjera na temelju akcija koje agent napravi. Za potrebe ovog rada pretpostavlja se da nema transakcijskih troškova u okruženju te je rebalansiranje portfelja trenutno, bez vremenskog odmak. Iako to nije realistična slika stvarnog stanja na tržištu s obzirom na to da te dvije stavke znatno utječu na stvarne odluke kod trgovanja dionicama, okolina je tako dizajnirana radi jednostavnosti i demonstrativne svrhe ovog rada. Prije nego što bi uopće mogli razmišljati o upotrebi metode podržanog učenja u stvarnome svijetu, potrebno je na kvalitetan način modelirati transakcijske troškove.

Kako bi mogli u svakom trenutku t izračunati vrijednost portfelja $vrijednost_portfelja_t$, koristimo formulu 4.9

$$vrijednost_portfelja_t = \sum_i P_{i,t} \cdot udio_{i,t-1} + c_{t-1} \quad (4.9)$$

U formuli 4.9, kako bi izračunali vrijednost portfelja u trenutku t , zbrajamo umnoške cijene dionice i u trenutku (danu) t ($P_{i,t}$) i udjela koje agent ima u trenutnom portfelju za dionicu i u trenutku $t - 1$ ($udio_{i,t-1}$ - broj dionica kompanije i). Tome zbroju na kraju dodajemo još i količinu investiranu u gotovini u trenutku c_{t-1} .

4.1.6. Upotreba alata za implementaciju algoritma podržanog učenja

Za implementaciju algoritma podržanog učenja, korištena je programska knjižica otvorenog koda RLlib [24] koja je napravljena za programski jezik Python [14]. RLlib nudi podršku za visoko distribuirana radna opterećenja podržanog učenja na razini proizvodnje, dok istovremeno održava unificirana i jednostavna aplikacijska programska sučelja (engl. *Application Programming Interface, API*) za veliki izbor industrijskih aplikacija. Pomoću RLliba mogu se trenirati agenti u postavci s jednim ili više agenata, isključivo iz izvanmrežnih skupova podataka ili pomoću eksterno povezanih simulatora. RLlib se koristi u mnogim industrijama kao što su kontrola klime, industrijska kontrola, proizvodnja i logistika, financije, igre, automobili, robotika, dizajn brodova i mnoge druge.

4.1.7. Treniranje i testiranje modela

Za treniranje i testiranje metode podržanog učenja, korištene su cijene 42 dionice opisanih u poglavlju 2. u rasponu od početka 2013. do kraja 2023. godine - ukupno 11 godina. Podržano učenje na kraju se sastoji od 5 nezavisnih agenata koji su istrenirani nad 5 različitih vremenskih razdoblja podataka. Svaki agent koristio je ukupno pet godina za treniranje, jednu za validaciju i jednu za testiranje na način na koji je prikazano u tablici 4.1.

Trening	Validacija	Test
2013 - 2017	2018	2019
2014 - 2018	2019	2020
2015 - 2019	2020	2021
2016 - 2020	2021	2022
2017 - 2021	2022	2023

Tablica 4.1. Period treniranja, validacije i testiranja za agente podržanog učenja

U tablici 4.2. prikazane su vrijednost hiperparametara korištenih kod implementacije PPO algoritma s RLlib programskom knjižicom.

Parametar	Vrijednost parametra
gamma	0.9
lambda	0.9
stopa učenja	0.0003
clip	0.25
broj epoha	16
veličina serije	128

Tablica 4.2. Hiperparametri korišteni za PPO algoritam

Gamma je faktor popusta koji određuje koliko buduće nagrade vrijede u usporedbi s neposrednim nagradama - niža vrijednost znači da agent daje prioritet kratkoročnim nagradama više nego dugoročnim nagradama. Lambda je parametar generalizirane procjene prednosti (engl. *Generalized Advantage Estimation, GAE*), koji uravnotežuje pristranost i varijancu u procjeni prednosti. Stopa učenja diktira veličinu ažuriranja parametara modela tijekom obuke. Također, ona kontrolira koliko brzo ili sporo model uči - politika agenta se ažurira u manjim koracima, što može dovesti do stabilnijeg, ali sporijeg učenja. Klip parametar je parametar izrezivanja koji se koristi u PPO algoritmu za ograničavanje promjene pravila tijekom svakog koraka ažuriranja. Osigurava da nova politika ne odstupa previše od stare politike. Broj epoha definira broj ponavljanja stohastičkog gradijenta po seriji podataka. Drugim riječima, kontrolira koliko se puta skupljena serija podataka propušta kroz optimizator tijekom svakog koraka obuke. Veličina serije određuje veličinu skupa iskustava (stanje, akcija, nagrada, sljedeće stanje) uzorkovanih iz okruženja koja se koristi za svako ažuriranje obuke. Manja veličina serije dovodi do češćih ažuriranja s manje podataka po jednom ažuriranju, što čini treniranje modela šumovitijim, ali bržim u smislu ažuriranja. Veće veličine serije dovode do stabilnijih ažuriranja, ali je svako ažuriranje računalno skuplje i rjeđe.

PPO implementacija također koristi i duboki model. Iako je ovo podržano učenje bez modela, ovdje se model koristi za drugu svrhu od one za koju se koristi model kod podržanog učenja s modelom. U ovome slučaju, kod PPO algoritma koristimo neuronsku mrežu koja predviđa politiku (skup mogućih akcija iz trenutnog stanja) te vrijednost portfelja s obzirom na trenutno stanje. Kod podržanog učenja koje se bazira na modelu, model predviđa samu okolinu.

Parametar	Vrijednost parametra
broj slojeva	2
broj neurona u sloju	32
aktivacijska funkcija	tanh

Tablica 4.3. Parametri modela neuronske mreže

Parametri modela koji su korišteni prikazani su u tablici 4.3. U izrađenoj implementaciji model je neuronska mreža s 2 skrivena sloja u kojoj svaki sloj ima 32 neurona. Kao aktivacijska funkcija koristi se funkcija *tanh*. Također, isti slojevi neuronske mreže koriste se i za politiku i za funkciju vrijednosti što omogućava brže učenje modela.

5. Rezultati i rasprava

Metoda optimizacije srednje vrijednosti i varijance te metoda dubokog podržanog učenja testirane su na periodu od 5 godina, od 2019. do 2023. godine, uključivo. Na početku testiranja za obje metode imali smo na raspolaganju 100000\$. Obje metode imale su opciju investirati u dionice 42 kompanije opisane u poglavlju 2. Udio u pojedinim dionicama u trenutku t prikazan je preko vektora težina $\mathbf{w} = [w_1, \dots, w_n]$, gdje vrijedi $\sum_{i=1}^N w_i = 1$ i $0 \leq w_i \leq 1$. Nakon što svaka pojedina metoda izgenerira vektor težina, težine se preoblikuju s pomoću metode

`DiscreteAllocation.greedy_portfolio()` iz programske knjižice `PyPortfolioOpt` [13] kako bi se zaokružile na cjelobrojne vrijednosti s obzirom na to da cijenu dionica. Tako se brinemo da u svakom trenutku t imamo samo cjelobrojnu količinu pojedinih dionica kompanija u portfelju.

Za metodu podržanog učenja trenirano je pet nezavisnih agenata gdje je svaki od njih korišten za jednu godinu testiranja performansa. Svaki agent koristi pet godina za treniranje te jednu godinu za validaciju (evaluaciju) prije nego što je spreman za testiranje.

Metoda optimizacije srednje vrijednosti i varijance koristi period od posljednjih 60 dana prije dana t kako bi izračunala svoje dvije najbitnije stavke za izračun rezultata - očekivane povrate i model rizika. Za period izračuna uzeto je točno 60 dana zato što se jednako toliko dana koristi za implementaciju stanja kod podržanog učenja.

Za prikaz rezultata, korištena je programska knjižica `Pyfolio` [25]. Ona se koristi za prikaz raznih metrika i rezultata financijskih skupova podataka u obliku tablica i dijagrama. U ovome radu korištena je metoda `pyfolio.create_returns_tear_sheet()`. Kako bi mogli prikazati rezultate s pomoću navedene metode, potrebno je pružiti tri skupa podataka. Prvi sadrži dnevne povrate di-

onica kompanija. Drugi skup podataka za svaki dan i svaku dionicu sadrži udio novaca koliko je uloženo u svaku od njih. Posljednji, treći skup sadrži preslikavanja kompanije u sektore kojima pripadaju.

Tablica 5.1. prikazuje usporedbu metrika izračunatih na temelju rezultata metode optimizacije srednje vrijednosti i varijance (MVO) i metode podržanog učenja (RL).

Metrika	MVO	RL
Godišnji prinos	20,645%	21,569%
Kumulativni prinosi	154,644%	163,484%
Godišnja volatilitnost	31,248%	24,817%
Sharpeov omjer	0,76	0,91
Calmar omjer	0,37	0,55
Stabilnost	0,77	0,81
Maksimalno povlačenje	-55,724%	-39,021%
Omega omjer	1,15	1,19
Sortino omjer	1,04	1,28
Nagib	-1,13	-0,64
Kurtoza	11.42	13.29
Omjer repa	1,09	1,08
Dnevna rizična vrijednost	-3,843%	-3,037%

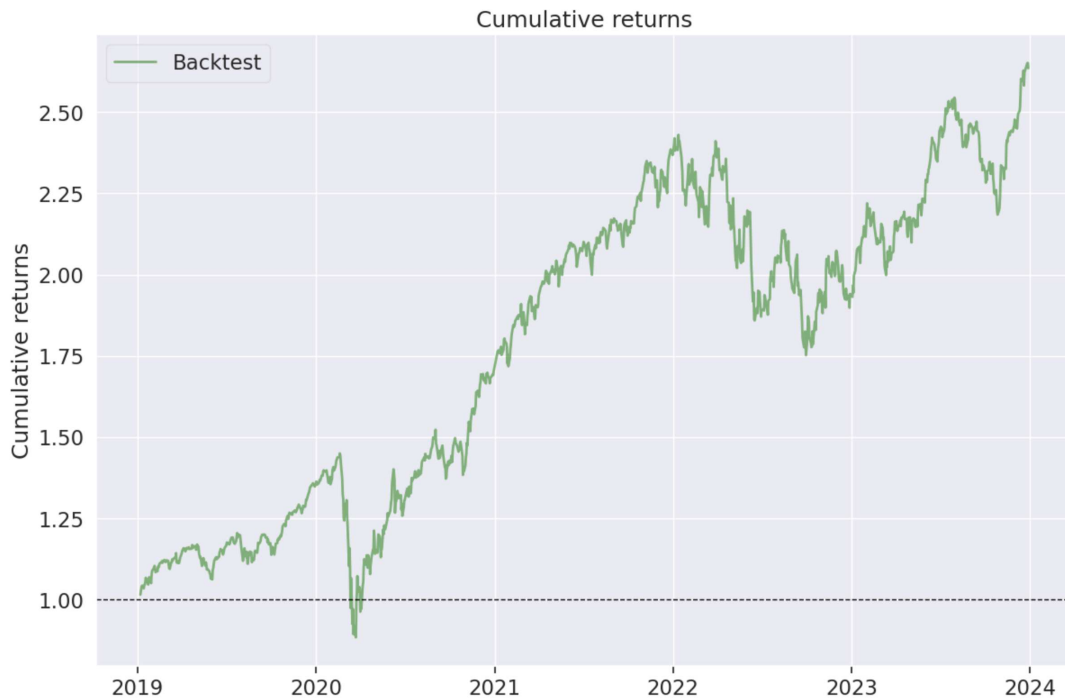
Tablica 5.1. Usporedba metrika rezultata metode optimizacije srednje vrijednosti i varijance (MVO) i metode podržanog učenja (RL)

Godišnji prinos ili povrat predstavlja prosječni godišnji prinos portfelja u proteklom razdoblju. Vidimo da je metoda podržanog učenja tu bila za skoro 1% bolja. **Kumulativni prinosi** predstavljaju koliki je ukupni prinos portfelja na kraju 2023. godine u odnosu na početak investiranja. Tu također vidimo da je podržano učenje imalo bolje rezultate uz skoro 10% veće prinose, što bi na temelju početnog uloga značilo nešto manje od 10000\$ veće zarade. **Godišnja volatilitnost** mjeri standardnu devijaciju prinosa portfelja na godišnjoj osnovi, pokazujući stupanj varijacije u prinosima. Želimo da taj broj bude što je moguće manji jer on označava koliki je potencijalni rizik ulaganja. Vrijednost za podržano učenje je osjetno manja nego za metodu optimizacije srednje vrijednosti i

varijance. Međutim, treba napomenuti da i vrijednost od 24,817% kolika je za podržano učenje isto predstavlja relativno visok stupanj rizika. Kao što je već objašnjeno u ovome radu, **Sharpeov omjer** predstavlja mjeru povrata prilagođenu riziku te što je željeno ponašanje da on bude što je moguće veći. **Calmar omjer** uspoređuje godišnji prinos portfelja s njegovim maksimalnim povlačenjem, odnosno ukazuje na povrat portfelja u odnosu na njegov najveći gubitak tijekom razdoblja testiranja. Želimo da njegova vrijednost bude što je moguće veća, stoga možemo primijetiti da podržano učenje ima bolji calmar omjer. **Stabilnost** je metrika koja odražava dosljednost prinosa portfelja te je njen veći iznos poželjan. **Maksimalno povlačenje** predstavlja najveći uočeni gubitak od vrha do dna tijekom razdoblja testiranja. S obzirom na to, želimo da je taj postotak što je moguće manji, tako da je opet podržano učenje imalo bolje rezultate. **Omega omjer** je mjera vjerojatnosti povrata iznad određenog praga u usporedbi s vjerojatnošću povrata ispod tog praga. Stoga, veći iznos predstavlja bolji rezultat. **Sortino omjer** je varijacija Sharpeovog omjera koja samo kažnjava negativnu volatilnost, odnosno fokusira se na negativna odstupanja. Što je iznos sortino omjera veći, to bolje. **Nagib** je metrika koja mjeri asimetriju distribucije povrata. Negativan nagib ukazuje da portfelj ima tendenciju češće imati negativne povrate što je slučaj kod obje metode, međutim kod metode optimizacije srednje vrijednosti i varijanci on je znatno lošiji. **Kurtoza** mjeri ekstremna odstupanja distribucije povrata. Željeno ponašanje ovisi od slučaja do slučaja, ali je generalno poželjno da nema toliko velika odstupanja, što u ovom slučaju označava da metoda optimizacije srednje vrijednosti i varijance ima bolji rezultat. **Omjer repa** koristi se za procjenu rizika ekstremnih gubitaka u odnosu na ekstremne dobitke. Veća vrijednost predstavlja bolji rezultat, stoga je ovo druga metrika u kojoj metoda optimizacije srednje vrijednosti i varijance ima za nijansu bolji rezultat od podržanog učenja. **Dnevna rizična vrijednost** predstavlja maksimalni očekivani gubitak na određenoj razini pouzdanosti u razdoblju od jednog dana. Dnevni iznos od -3,037% implicira da postoji određena vjerojatnost da bi portfelj mogao izgubiti do tog postotka u jednom danu što znači da je i u ovom slučaju podržano učenje imalo bolje rezultate.

Na temelju svih vrijednosti metrika iz tablice 5.1., možemo zaključiti da je podržano učenje imalo osjetno bolje rezultate. Jedine dvije metrike u kojima je ono imalo lošije performanse u odnosu na metodu optimizacije srednje vrijednosti i varijance je kod kurtoza metrike te omjera repa.

U nastavku ćemo ući detaljnije u usporedbu metode optimizacije srednje vrijednosti i varijance s podržanim učenjem na testnom skupu podataka uz pomoć raznih dijagrama.

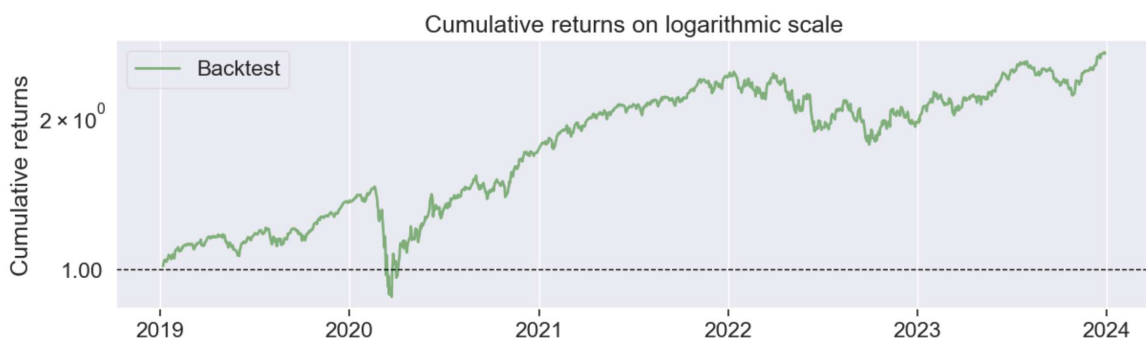


Slika 5.1. Graf koji prikazuje kumulativne prinose portfelja generiranog tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

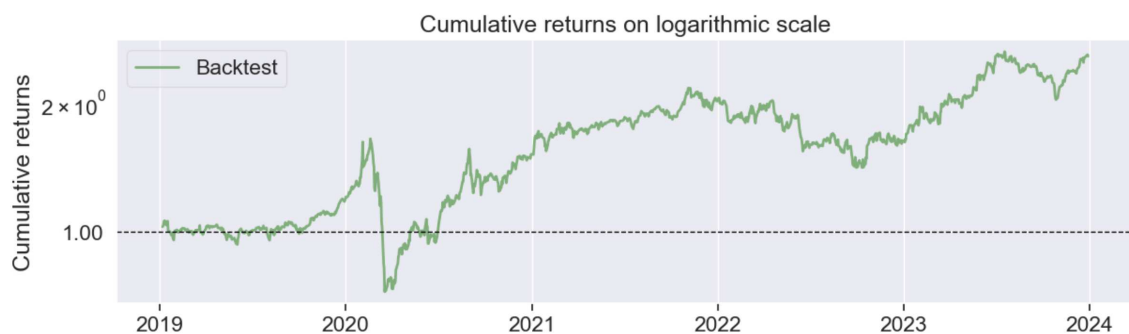


Slika 5.2. Graf koji prikazuje kumulativne prinose portfelja generiranog metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.1. i 5.2. prikazuju kumulativne prinose portfelja generiranog tehnikom podržanog učenja, odnosno metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine. Grafovi imaju vrlo sličan izgled. Možemo uočiti vremenske periode kad su obje metode imali dobre ili loše prinose. Jedini period kad su i jedna i druga metoda imale negativne kumulativne prinose je početkom 2020. godine u doba pandemije koronavirusa. Tada je prisutan strelovit pad iz kojeg se tehnika podržanog učenja nešto brže oporavila. Kod metode optimizacije srednje vrijednosti i varijance možemo vidjeti prisutnu osjetno veću volatilnost kroz veći pad i porast krivulje grafa u istim vremenskim periodima.

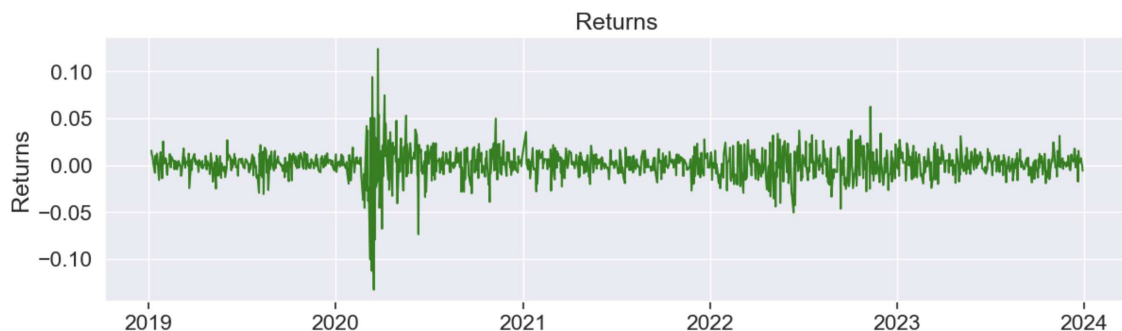


Slika 5.3. Graf koji prikazuje logaritam kumulativnih prinosa portfelja generiranog tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

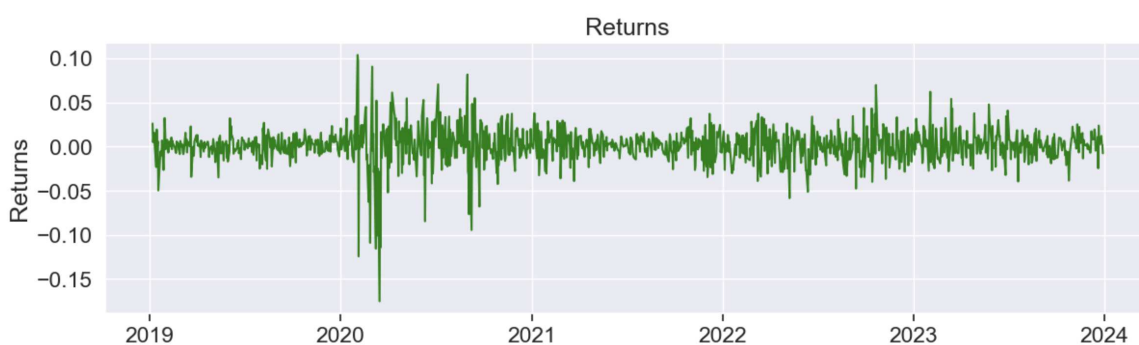


Slika 5.4. Graf koji prikazuje logaritam kumulativnih prinosa portfelja generiranog metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.3. i 5.4. prikazuju logaritam kumulativnih prinosa portfelja generiranog tehnikom podržanog učenja, odnosno metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine. Grafovi su vrlo sličnima onima od prije u kojima su prikazani kumulativni prinosi bez logaritma. Vrijednost ovih dijagrama je u tome što daju bolji uvid u postotne promjene prinosa s obzirom na to da su skokovi umjereniji.

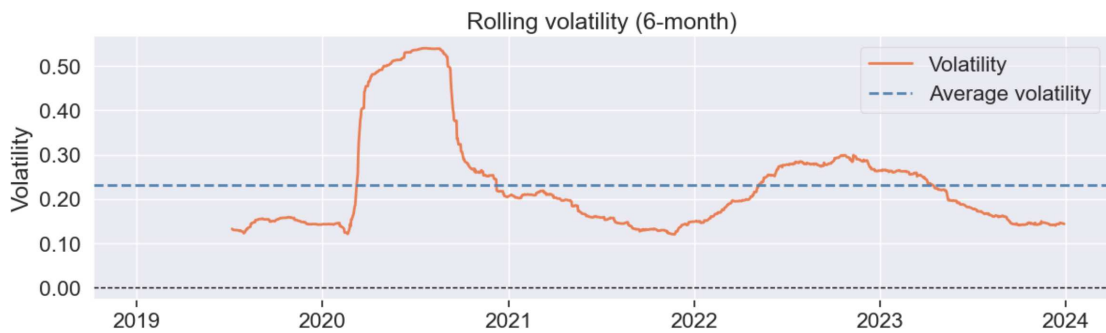


Slika 5.5. Prikaz dnevnih povrata portfelja generiranog tehnikom podržanog učenja u periodu od 2019. do 2023. godine.



Slika 5.6. Prikaz dnevnih povrata portfelja generiranog metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.3. i 5.4. prikazuju dijagrame dnevnih povrata portfelja generiranog tehnikom podržanog učenja, odnosno metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine. Ovaj graf prikazuje kolika je dnevna volatilitnost portfelja kroz vrijeme. Možemo uočiti da donji graf, onaj koji prikazuje metodu optimizacije srednje vrijednosti i varijance, ima veće dnevne skokove što indicira da je ta metoda volatilnija. Međutim, možemo primijetiti da je najveći pozitivni dnevni skok bio kod portfelja izgeneriranog tehnikom podržanog učenja (u periodu velike volatilitnosti početkom 2020. godine). To smo zapravo mogli primijetiti u tablici 5.1. kroz kurtoza metriku koja mjeri upravo to najviše dnevno pozitivno odstupanje. S druge strane, najviše dnevno negativno odstupanje možemo primijetiti kod metode optimizacije srednje vrijednosti i varijance te iznosi nešto više od 15%.

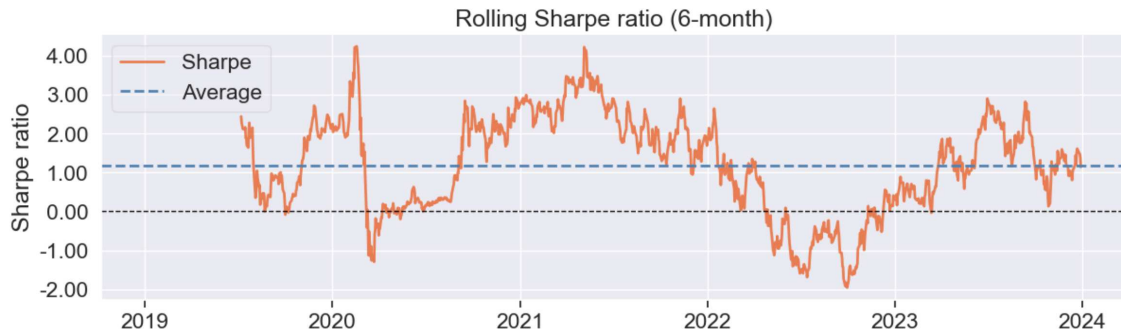


Slika 5.7. Prikaz kotrljajuće volatilnosti u šestomjesečnom periodu portfelja generiranog tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

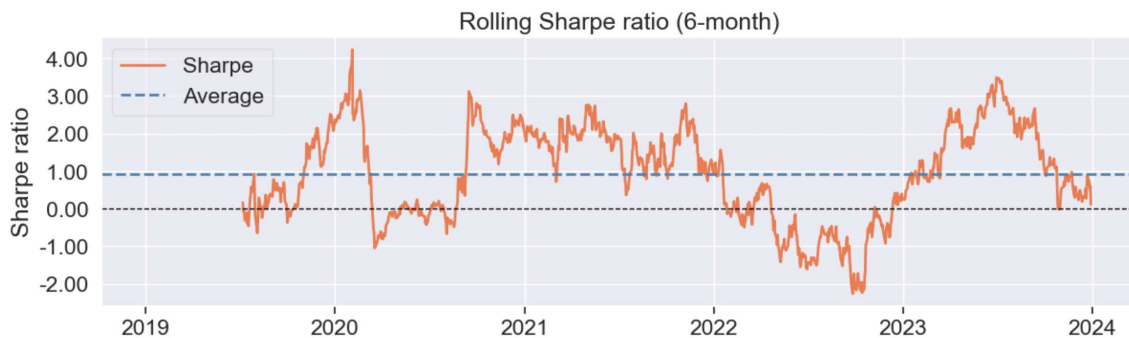


Slika 5.8. Prikaz kotrljajuće volatilnosti u šestomjesečnom periodu metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Na slikama 5.3. i 5.4. možemo vidjeti prikaz kotrljajuće volatilnosti u šestomjesečnom periodu portfelja generiranog tehnikom podržanog učenja te metodom optimizacije srednje vrijednosti i varijance. Graf prikazuje šestomjesečnu pokretnu volatilnost. To znači da se u svakoj točki na dijagramu volatilnost izračunava na temelju povrata u prethodnih šest mjeseci. Kako se prozor kreće prema naprijed, najstariji dan tog šestomjesečnog prozora se uklanja iz izračuna, a dodaje se onaj najnoviji. Ako usporedimo navedeni graf za promatrane dvije metode, možemo primijetiti osjetno veću volatilnost kod metode optimizacije srednje vrijednosti i varijance. Prosječna volatilnost koja je prikazana isprekidanom plavom linijom kod metode optimizacije srednje vrijednosti i varijance iznosi 0.3, dok kod podržanog učenja ona iznosi oko 0.23. Unatoč tome, možemo primijetiti slične trendove rasta i pada volatilnosti kod obje metode, gdje je posebno volatiln period bio početkom 2020. godine zbog COVID-19 pandemije.

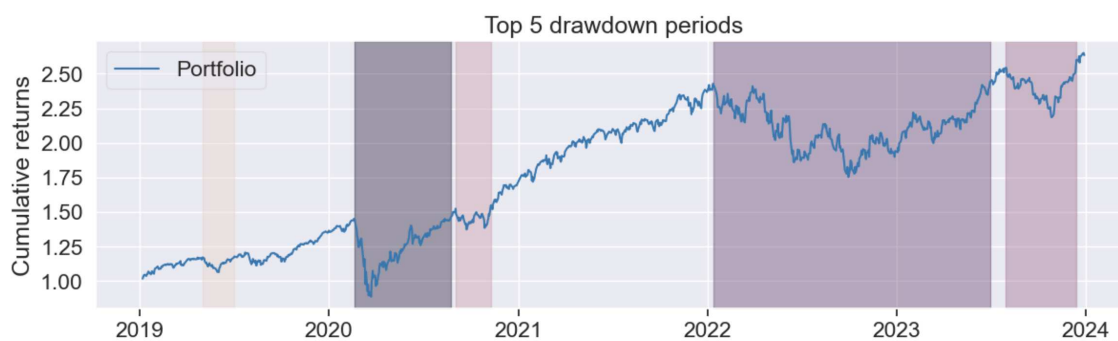


Slika 5.9. Prikaz kotrljajućeg Sharpeovog omjera u šestomjesečnom periodu portfelja generiranog tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

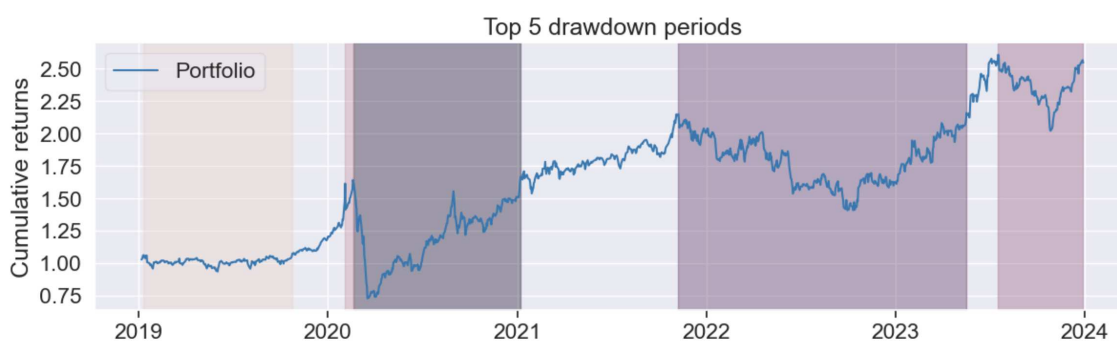


Slika 5.10. Prikaz kotrljajućeg Sharpeovog omjera u šestomjesečnom periodu metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Grafovi prikazani na slikama 5.9. i 5.10. prikazuju kotrljajući Sharpeov omjer u šestomjesečnom periodu portfelja generiranog tehnikom podržanog učenja te metodom optimizacije srednje vrijednosti i varijance. Kotrljajući Sharpeov omjer koristi povrate od posljednjih šest mjeseci za izračun Sharpeovog omjer na sličan način na koji se računa kotrljajuća volatilnost. Ako usporedimo ove dvije metode s obzirom na kotrljajući Sharpeov omjer, možemo primijetiti kako podržano učenje ima bolje rezultate. Prosječan Sharpeov omjer u šestomjesečnom periodu iznosi nešto više od jedan, dok kod metode optimizacije srednje vrijednosti i varijance iznosi točno jedan.



Slika 5.11. Prikaz pet vremenskih razdoblja koja su imala najveće povlačenje, odnosno gubitak za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.



Slika 5.12. Prikaz pet vremenskih razdoblja koja su imala najveće povlačenje, odnosno gubitak za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.11. i 5.12. prikazuju po pet najdužih vremenskih perioda gubitka za svaku od metoda koje uspoređujemo. Preciznije rečeno, period povlačenja predstavlja vremenski period otkad portfelj počne gubiti na svojoj vrijednosti pa sve dok ne povрати svoju najveću vrijednost, odnosno onu s kojom počinje period povlačenja. Na grafovima možemo vidjeti da se periodi povlačenja dobrim djelom preklapaju, odnosno da se pojavljuju u približno isto vrijeme za obje metode. Kod obje metode, najduži period povlačenja predstavlja cijela 2022. godina i prvi dio 2023. godine, s tim da je za metodu optimizacije srednje vrijednosti i varijance on ipak par mjeseci duži. Taj period zapravo oslikava sliku tržišta iz 2022. godine koje je globalno bilo u padu. Na primjer, najpoznatiji svjetski in-

deksni fond S%P500, u kojem se također nalaze i svih 42 kompanije čije dionice imaju ove metode na raspolaganju, imao je u 2022. godini gubitak od 18%.

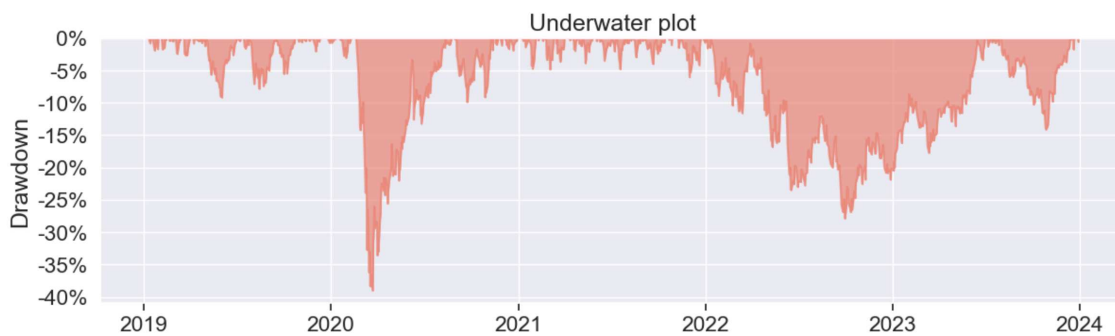
Povlačenje	Datum vrhunca	Datum doline	Datum oporavka	Trajanje
39.02%	19.02.2020	23.03.2020	24.08.2020	134
27.89%	12.01.2022	30.09.2022	30.06.2023	383
14.15%	31.07.2023	27.10.2023	13.12.2023	98
9.91%	02.09.2020	23.09.2020	09.11.2020	49
9.19%	03.05.2019	03.06.2019	03.07.2019	44

Tablica 5.2. Opis najvećih perioda povlačenja za podržano učenje

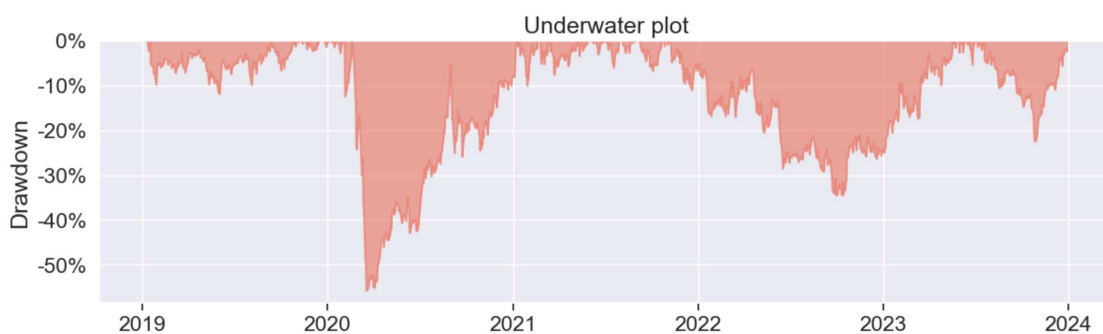
Povlačenje	Datum vrhunca	Datum doline	Datum oporavka	Trajanje
55.72%	19.02.2020	18.03.2020	07.01.2021	232
34.52%	08.11.2021	30.09.2022	18.05.2023	399
22.56%	18.07.2023	27.10.2023	NaT	NaN
12.45%	04.02.2020	05.02.2020	19.02.2020	12
11.93%	11.01.2019	03.06.2019	25.10.2019	206

Tablica 5.3. Opis najvećih perioda povlačenja za metodu optimizacije srednje vrijednosti i varijance

Ono što prikazuju slike 5.11. i 5.12. možemo detaljnije vidjeti u tablicama 5.2. i 5.3. Primjećujemo da su periodi povlačenja za podržano učenje u pravilu kraći i imaju manje gubitke. Kod metode optimizacije srednje vrijednosti i varijance možemo primjetiti da jedan takav period nije završio s krajem 2023. godine kad je testiranje gotovo. Podatci iz tablice nam također potvrđuju da se periodi najdužih povlačenja s najvećim gubiticima u dobroj mjeri preklapaju kod obje metode, ali su ti periodi duži i imaju veće gubitke kod metode optimizacije srednje vrijednosti i varijance.

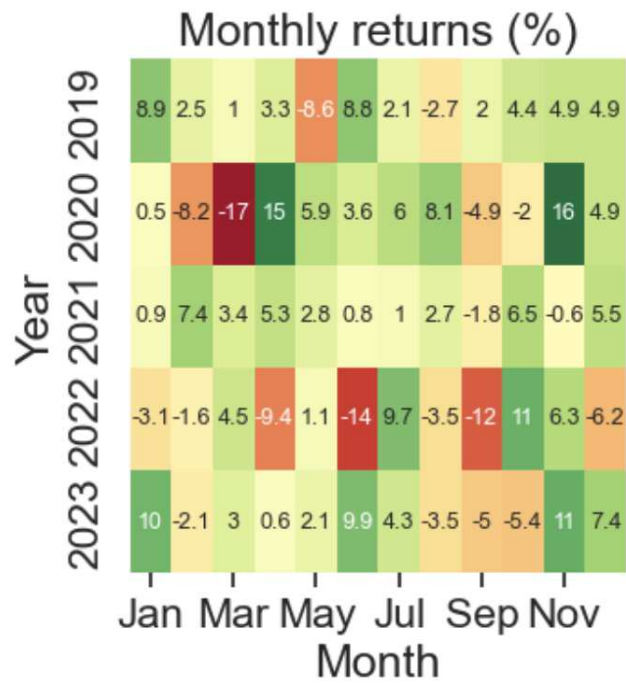


Slika 5.13. Prikaz podvodnog grafa za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

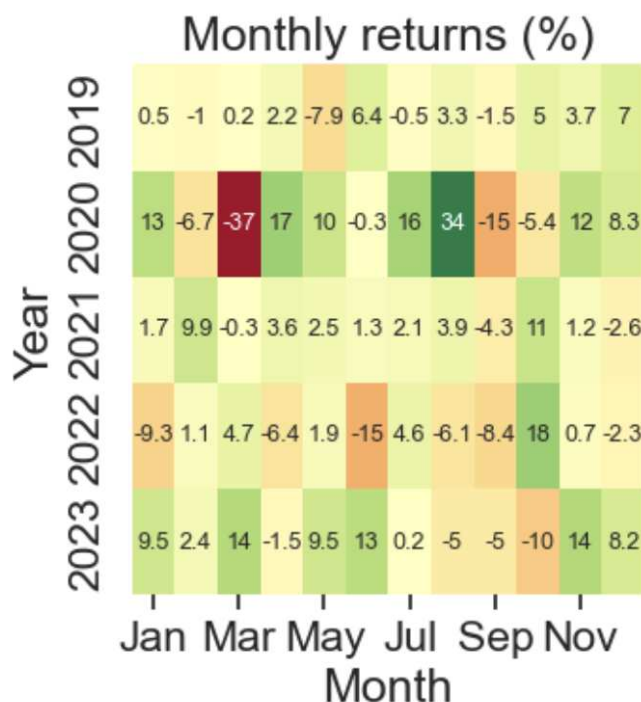


Slika 5.14. Prikaz podvodnog grafa za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Na slikama 5.13. i 5.14. prikazani su podvodni grafovi za portfelje generirane tehnikom podržanog učenja te metodom optimizacije srednje vrijednosti i varijance. Podvodni graf ili dijagram prikazuje gubitke portfelja - što ima više crveno obojene površine na grafu, to znači da je bilo više gubitaka. Možemo vidjeti da tu definitivno metoda optimizacije srednje vrijednosti i varijance prednjači u obojenosti. Također, možemo vidjeti da je i maksimalni zabilježen gubitak veći za metodu optimizacije srednje vrijednosti (55,724%) i varijance u odnosu na tehniku podržanog učenja (39,021%) što smo mogli i vidjeti kroz metriku maksimalnog povlačenja u tablici 5.1. Također, možemo primijetiti da se obojenost podvodnog grafa i grafa koji prikazuje pet najvećih razdoblja povlačenja dobrim djelom poklapa za iste vremenske periode.

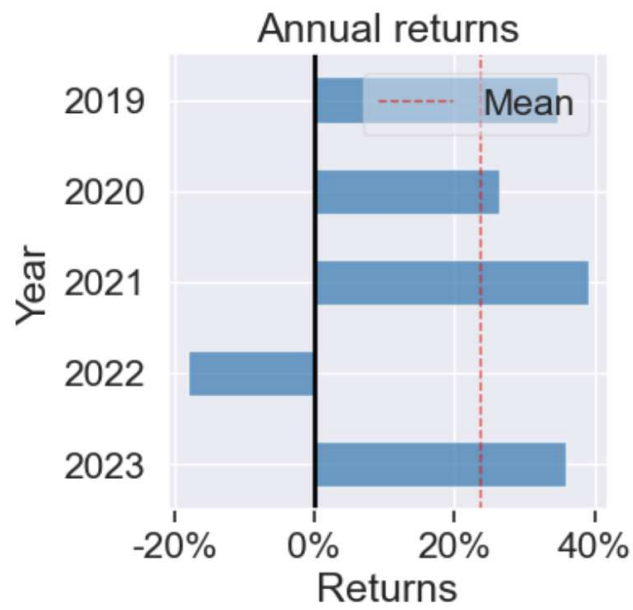


Slika 5.15. Prikaz toplinske karte za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

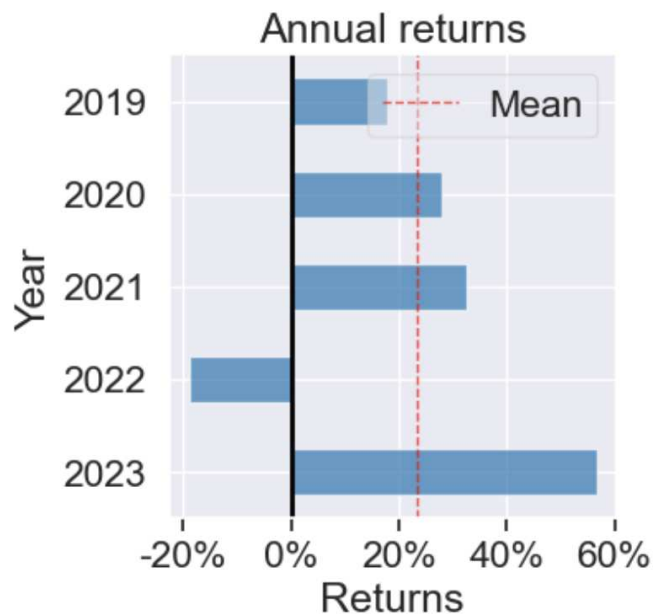


Slika 5.16. Prikaz toplinske karte za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.15. i 5.16. prikazuju toplinsku kartu za dane dvije metode. Toplinska karta prikazuje prinose portfelja po mjesecima i godinama gdje su zelenom bojom obojeni najveći prinosi, dok su crvenom označeni najveći gubitci. Možemo uočiti da je toplinska karta za podržano učenje generalno zelenija od karte za metodu optimizacije srednje vrijednosti i varijance. Također, možemo primijetiti da je jednoličnija svijetlo-žuta boja na toplinskoj karti za metodu optimizacije srednje vrijednosti i varijance posljedica toga što ta metoda ima dvije velike iznimke - najveći dobitak iznosi 34% (kolovoz, 2020.), dok najveći gubitak iznosi -37% (ožujak, 2020). Ostali mjeseci u svim godinama se ne približavaju tim dvjema krajnjim vrijednostima, stoga je to razlog jednoličnoj obojenosti te toplinske karte.

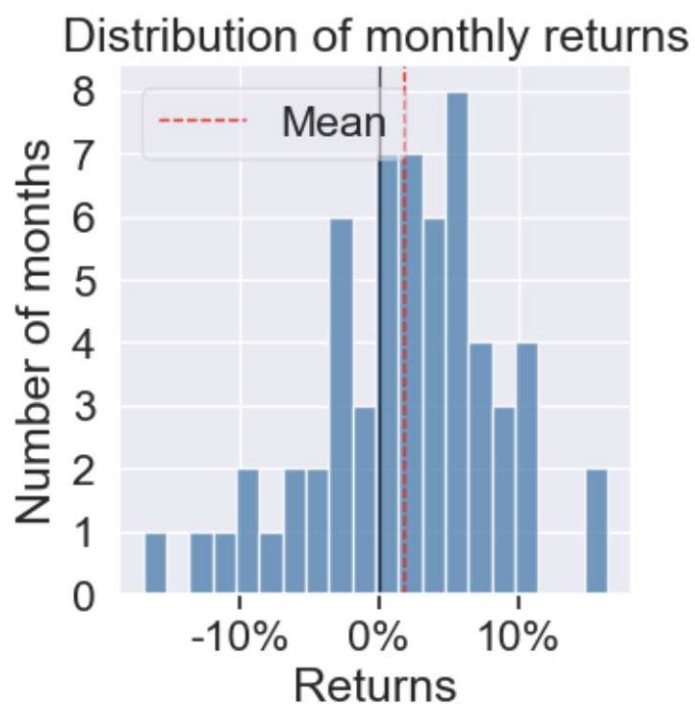


Slika 5.17. Graf koji prikazuje godišnji prinos portfelja zajedno sa srednjom vrijednosti za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

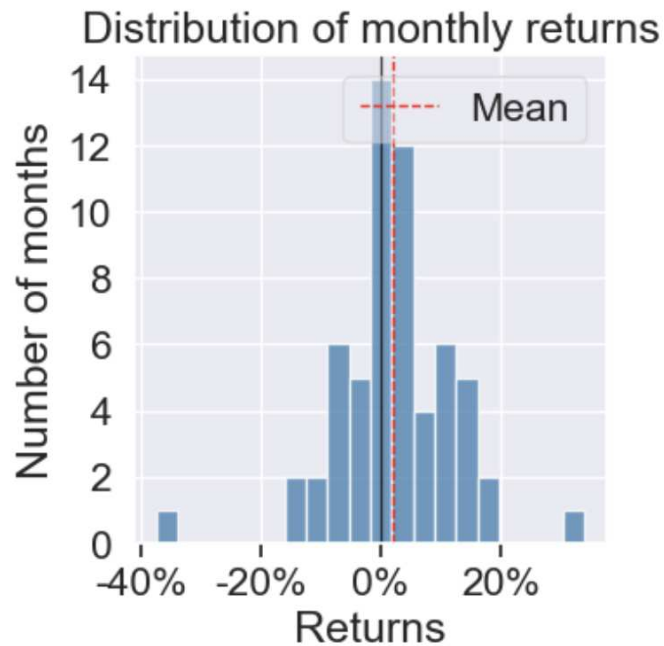


Slika 5.18. Graf koji prikazuje godišnji prinos portfelja zajedno sa srednjom vrijednosti za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Grafovi prikazani na slikama 5.17. i 5.18. pružaju uvid u godišnji prinos portfelja za svaku godinu testiranja. Možemo primijetiti da obje metode imaju četiri godine s pozitivnim godišnjim prinosom te jednu godinu s negativnim, 2022. godinu. Te je godine najpoznatiji svjetski indeksni fond S&P500, u kojem se također nalaze i svih 42 kompanije čije dionice imaju ove metode na raspolaganju, imao gubitak od 18%. Stoga ne čudi da su obje metode imale loše rezultate te godine. Grafovi također prikazuju i prosječnu vrijednost godišnjih prinosa koja je nešto veća za podržano učenje.

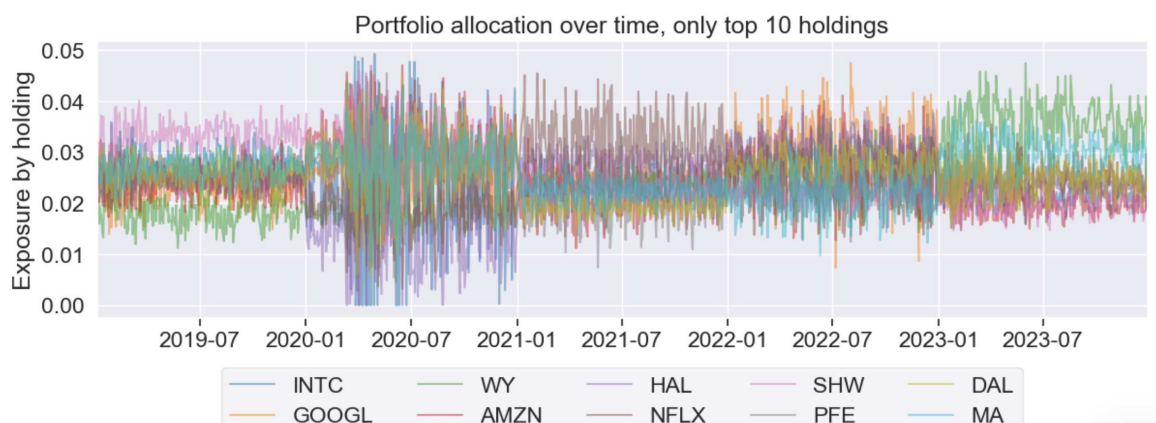


Slika 5.19. Graf koji prikazuje mjesečni prinos portfelja zajedno sa srednjom vrijednosti za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.

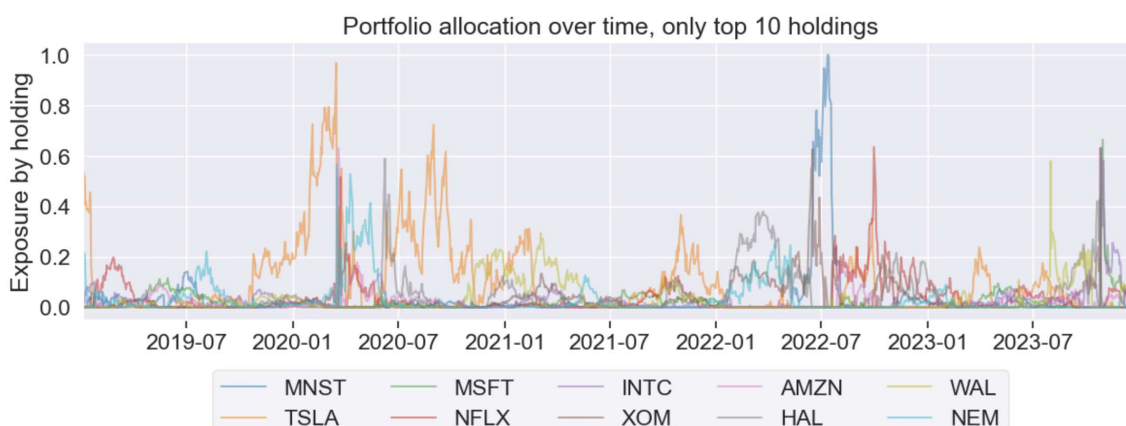


Slika 5.20. Graf koji prikazuje mjesečni prinos portfelja zajedno sa srednjom vrijednosti za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.19. i 5.20. prikazuju mjesečne prinose portfelja generiranog tehnikom podržanog učenja te metodom optimizacije srednje vrijednosti i varijance. Možemo primijetiti da je prosječan prinos portfelja generiranog podržanim učenjem veći od prosječnog mjesečnog prinosa druge metode. Na slici 5.16. možemo primijetiti dvije velike iznimke prinosa kod metode optimizacije srednje vrijednosti i varijance koje su se također mogle vidjeti i u toplinskoj karti prikazanoj na 5.16.



Slika 5.21. Prikaz udjela deset kompanija s najvećim udjelima u portfelju generiranog tehnikom podržanog učenja u periodu od 2019. do 2023. godine.



Slika 5.22. Prikaz udjela deset kompanija s najvećim udjelima u portfelju generiranog metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Grafovi prikazani na slikama 5.21. i 5.22. pokazuju možda i najbolju sliku kolika je zapravo stvarna razlika u volatilnosti između dvije metode i zašto tehnika podržanog učenja ima značajno bolje rezultate. Na dijagramima su prikazani grafovi udjela u deset kompanija u kojima je pojedina metoda imala najveće udjele. Vidimo da se ti udjeli znatno preklapaju kod podržanog učenja te u niti jednom trenutku nijedna dionica ne prelazi 5% udjela u portfelju. S druge strane, metoda optimizacije srednje vrijednosti i varijance ima tendenciju ulagati ogromne iznose u jako mali skup dionica. Čak možemo primijetiti da je u dva navrata metoda imala skoro 100% udjela u jednoj kompaniji. U 2020. godini to je bila tvrtka Tesla Inc (TSLA), dok je u 2022. godini to bila tvrtka Monster

Beverage Corp (MNST).

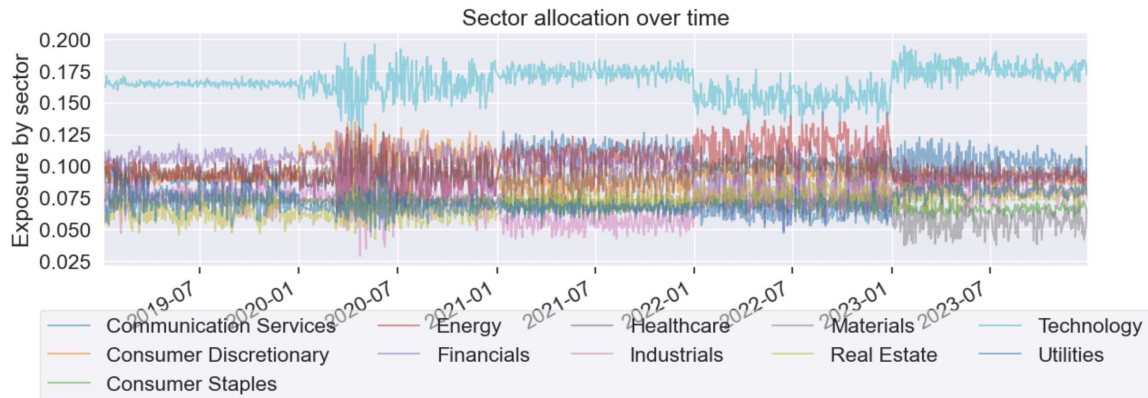
Dionica	Maksimalna vrijednost
INTC	4.93%
GOOGL	4.74%
WY	4.74%
AMZN	4.70%
HAL	4.70%
NFLX	4.58%
SHW	4.55%
PFE	4.55%
DAL	4.47%
MA	4.41%

Tablica 5.4. Prikaz deset kompanija koje su imale najveće maksimalne udjele u portfelju za podržano učenje

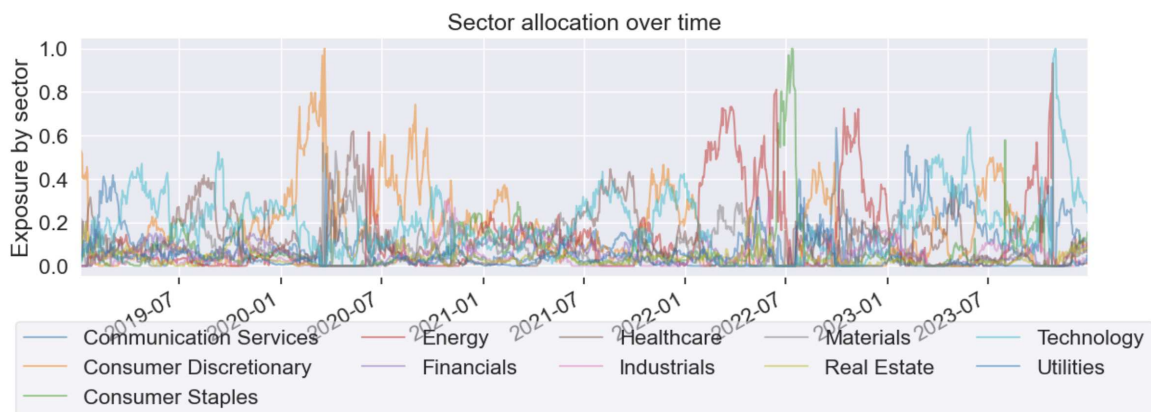
Dionica	Maksimalna vrijednost
MNST	99.97%
TSLA	96.74%
MSFT	66.37%
NFLX	63.47%
INTC	63.25%
XOM	63.10%
AMZN	62.98%
HAL	58.86%
WAL	57.91%
NEM	56.49%

Tablica 5.5. Prikaz deset kompanija koje su imale najveće maksimalne udjele u portfelju za metodu optimizacije srednje vrijednosti i varijance

Tablice 5.4. i 5.5. prikazuju najveće udjele u istih deset kompanija sa slika 5.21. i 5.22. Ovi nam podatci potvrđuju da su udjeli dionica kompanija kod podržanog učenja bili puno više diverzificirani nego što je to slučaj kod metode optimizacije srednje vrijednosti i varijance.



Slika 5.23. Graf koji prikazuje udio u sektorima industrije za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.



Slika 5.24. Graf koji prikazuje udio u sektorima industrije za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Slike 5.23. i 5.24. prikazuju udjele u portfeljima po sektorima kojima dionice mogu pripadati. Kao što se spomenulo u poglavlju 2., dionice su raspoređene u 11 sektora industrije, s tim da najviše njih pripada sektoru tehnologije - njih sedam. Sve ostale sektore čini po tri ili četiri kompanije. Kao i u prethodnim grafovima, možemo primijetiti veliku razliku među metodama podržanog učenja te optimizacije srednje vrijednosti i varijance. Metoda podržanog učenja u pravilu jednoliko ulaže u sve sektore industrije s malom iznimkom sektora tehnologije. Međutim, to ima smisla s obzirom na to da najveći broj dionica pripada tom sektoru. Također, možemo primijetiti da se iz godine u godinu minimalno mijenjanju trendovi ulaganja po sektorima što ima smisla ako se sjetimo da je za svaku godinu testiranja zadužen jedan nezavisni agent treniran na posebnom skupu za treniranje. S druge strane, metoda optimizacije srednje vrijednosti i varijance je puno

volatilnija te nema jasne trendove ulaganja u određene sektore. Možemo primijetiti da je čak u tri navrata imala 100% udjela u jednom sektoru i to svaki put u drugom sektoru. U 2020. godini to je bio sektor diskrecijske potrošnje, u 2022. sektor nediskrecijske potrošnje dok je krajem 2023. godine to bio sektor tehnologije.



Slika 5.25. Prikaz broja dionica za koje je postojao barem nekakav ulog u svakom trenutku za portfelj generiran tehnikom podržanog učenja u periodu od 2019. do 2023. godine.



Slika 5.26. Prikaz broja dionica za koje je postojao barem nekakav ulog u svakom trenutku za portfelj generiran metodom optimizacije srednje vrijednosti i varijance u periodu od 2019. do 2023. godine.

Na slikama 5.25. i 5.26. prikazani su grafovi koje pružaju uvid u broj dionica u kojima je svaka od metoda imala uložene udjele. Drugim riječima, ako je $w_i > 0$ u trenutku t , onda smatramo da portfelj ima udio u dionici i . Ovaj graf također prikazuje uvid u volatilnost pojedine metode. Portfelji izgenerirani podržanim učenjem u većini perioda

testiranja imaju udjela u svih 42 dionice. Jedino se u 2020. godini dogodilo da je u par navrata portfelj imao udjela u 41 kompaniji. S druge strane, kod metode optimizacije srednje vrijednosti i varijance stvar je drugačija. Prosječan broj dionica u portfelju tu iznosi samo 24. U niti jednom trenutku udio nije prešao brojku od 40 dionica. Također, možemo primijetiti da je u doba krize u 2020., 2022. te krajem 2023. godine portfelj težio tome da ima udjela u što je moguće manje dionica.

Na temelju svih metrika i grafova koji su prikazani u ovome poglavlju možemo zaključiti da je tehnika dubokog podržanog učenja imala bolje rezultate od metode optimizacije srednje vrijednosti i varijance na skupu podataka od ukupno 5 godina - od 2019. do 2023 godine. Ukupni i prosječni prinosi su veći kod podržanog učenja, ali je puno veća razlika u volatilnosti. Mogli smo primijetiti da je metoda optimizacije srednje vrijednosti i varijanci sklonija većoj volatilnosti, pogotovo u periodu krize kao što je ona iz 2020. godine uzrokovana pandemijom COVID-19 i kriza iz 2022. godine. Također, ta metoda je sklona ulaganju u manji broj dionica u odnosu na podržano učenje koje je u gotovo svako danu testiranja imalo udjele u svih 42 dionice.

6. Zaključak

U ovome radu generirali smo portfelj koji se sastoji isključivo od dionica 42 kompanije iz najpoznatijeg svjetskog indeksa S&P500 podijeljenih u 11 sektora industrije. Portfelji su generirani s dvije metode - tradicionalna metoda optimizacije srednje vrijednosti i varijance te metoda dubokog podržanog učenja koja je posljednjih nekoliko godina znatno dobila na popularnosti. Na temelju rezultata prezentiranih u ovome radu možemo zaključiti kako je tehnika dubokog podržanog učenja ostvarila bolje rezultate. Portfelj generiran dubokim podržanim učenjem imao je veće prinose, ali posebno velika razlika bila je u volatilnosti. Kod podržanog učenja, algoritam je nastojao u svakom trenutku imati diverzificiran portfelj te se tako bolje mogao nositi s kriznim vremenima. Posebno uočljiva je razlika u performansama tijekom krize iz 2020. godine tijekom pandemije COVID-19 te tijekom 2022. godine kad je globalno tržište dionica bilo u velikom padu. U nekoliko navrata portfelj generiran metodom optimizacije srednje vrijednosti i varijance imao je skoro 100% udjela u samo jednoj dionici, što nije željeno ponašanje. Također, duboko podržano učenje imalo je bolji povrat po jedinici rizika što je vidljivo kroz Sharpeov omjer.

Unatoč vrlo dobrim rezultatima dubokog podržanog učenja, tako generiran portfelj nije spreman za ulaganje u stvarnom svijetu. Prije nego što bi se mogao stvarno koristiti, trebalo bi bolje modelirati okolinu u kojoj djeluje agent podržanog učenja. U ovome radu nismo uzimali u obzir transakcijske troškove te smo pretpostavili da su promjene u portfelju trenutne što može imati veliki utjecaj u stvarnom svijetu ulaganja. Također, treba napomenuti da se za treniranje modela podržanog učenja može koristiti znatno više resursa nego što se koristilo za metodu u opisanu u ovom radu. Tako možemo potencijalno imati znatno robusniji model koji bi bolje generalizirao i koji bi mogao raditi i s više paralelnih agenata. Isto tako, s jačim modelom mogli bismo trenirati agente podržanog

učenja na većem skupu podataka s više dionica i duljim vremenskim periodom što bi također dovelo do bolje generalizacije modela.

Literatura

- [1] H. Markowitz, “Portfolio selection”, *The Journal of Finance*, sv. 7, br. 1, str. 77–91, 1952. [Mrežno]. Adresa: <http://www.jstor.org/stable/2975974>
- [2] Wikipedia, “Burza”, siječanj 2022., page Version ID: 6268806. [Mrežno]. Adresa: <https://hr.wikipedia.org/w/index.php?title=Burza&oldid=6268806>
- [3] —, “Dividenda”, lipanj 2021., page Version ID: 5895524. [Mrežno]. Adresa: <https://hr.wikipedia.org/w/index.php?title=Dividenda&oldid=5895524>
- [4] I. Webster, “S&P 500 Returns since 1957”. [Mrežno]. Adresa: <https://www.officialdata.org/us/stocks/s-p-500/1957>
- [5] Y. finance, “S&P 500 (^GSPC) Stock Historical Prices & Data - Yahoo Finance”. [Mrežno]. Adresa: <https://finance.yahoo.com/quote/%5EGSPC/history/>
- [6] S. D. J. Indices, “S&P 500®”. [Mrežno]. Adresa: <https://www.spglobal.com/spdji/enhttps://www.spglobal.com/spdji/en/indices/equity/sp-500/>
- [7] Y. finance, “Yahoo Finance - Stock Market Live, Quotes, Business & Finance News”. [Mrežno]. Adresa: <https://finance.yahoo.com/>
- [8] —, “yfinance”, srpanj 2024. [Mrežno]. Adresa: <https://pypi.org/project/yfinance/>
- [9] S. Sood, K. Papatotiriou, M. Vaiciulis, i T. Balch, “Deep Reinforcement Learning for Optimal Portfolio Allocation: A Comparative Study with Mean-Variance Optimization”, *ICAPS 2023*, 2023.
- [10] W. F. Sharpe, “The Sharpe Ratio (Fall 1994)”, u *Streetwise*, P. L. Bernstein i F. J. Fabozzi, Ur. Princeton University Press, prosinac 1998., str. 169–185.

<https://doi.org/10.1515/9781400829408-022>

- [11] Investopedia, “Understanding the Sharpe Ratio”. [Mrežno]. Adresa: https://www.investopedia.com/articles/07/sharpe_ratio.asp
- [12] G. Cornuéjols i R. Tutuncu, *Optimization Methods in Finance*. Cambridge University Press, 01 2007. <https://doi.org/10.1017/CBO9780511753886>
- [13] PyPortfolioOpt, “PyPortfolioOpt”. [Mrežno]. Adresa: <https://pyportfolioopt.readthedocs.io/en/latest/>
- [14] Python, “Welcome to Python.org”, kolovoz 2024. [Mrežno]. Adresa: <https://www.python.org/>
- [15] O. Ledoit i M. Wolf, “Honey, I Shrunk the Sample Covariance Matrix”, *The Journal of Portfolio Management*, sv. 30, br. 4, str. 110–119, srpanj 2004. <https://doi.org/10.3905/jpm.2004.110>
- [16] Wikipedia, “Reinforcement learning”, kolovoz 2024., page Version ID: 1239979142. [Mrežno]. Adresa: https://en.wikipedia.org/w/index.php?title=Reinforcement_learning&oldid=1239979142
- [17] L. P. Kaelbling, M. L. Littman, i A. W. Moore, “Reinforcement Learning: A Survey”, travanj 1996., arXiv:cs/9605103. <https://doi.org/10.48550/arXiv.cs/9605103>
- [18] Megajuice, “English: Diagram showing the components in a typical Reinforcement Learning (RL) system. An agent takes actions in an environment which is interpreted into a reward and a representation of the state which is fed back into the agent. Incorporates other CC0 work: <https://openclipart.org/detail/202735/eye-side-view> <https://openclipart.org/detail/191072/blue-robot> and <https://openclipart.org/detail/246662/simple-maze-puzzle>”, travanj 2017. [Mrežno]. Adresa: https://commons.wikimedia.org/wiki/File:Reinforcement_learning_diagram.svg
- [19] waldoalvarez, “English: Markov Decision Process”, svibanj 2017. [Mrežno]. Adresa: https://commons.wikimedia.org/wiki/File:Markov_Decision_Process.svg

- [20] Wikipedia, “Markov decision process”, kolovoz 2024., page Version ID: 1241302840. [Mrežno]. Adresa: https://en.wikipedia.org/w/index.php?title=Markov_decision_process&oldid=1241302840
- [21] J. Moody i M. Saffell, “Learning to trade via direct reinforcement”, *IEEE Transactions on Neural Networks*, sv. 12, br. 4, str. 875–889, srpanj 2001., conference Name: IEEE Transactions on Neural Networks. <https://doi.org/10.1109/72.935097>
- [22] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, i O. Klimov, “Proximal Policy Optimization Algorithms”, kolovoz 2017., arXiv:1707.06347 [cs]. <https://doi.org/10.48550/arXiv.1707.06347>
- [23] DhanushKumar, “PPO Algorithm”, veljača 2024. [Mrežno]. Adresa: <https://medium.com/@danushidk507/ppo-algorithm-3b33195de14a>
- [24] Ray, “RLLib: Industry-Grade Reinforcement Learning — Ray 2.34.0”. [Mrežno]. Adresa: <https://docs.ray.io/en/latest/rllib/index.html>
- [25] Quantopian, “pyfolio — pyfolio 0.9.3+28.g3dcf744 documentation”. [Mrežno]. Adresa: <https://pyfolio.ml4trading.io/index.html>

Sažetak

Ovaj rad uspoređuje dvije različite metode za generiranje portfelja klasa imovine koristeći povijesne cijene američkih tvrtki. Metoda optimizacije srednje vrijednosti i varijance je tradicionalna metoda koja već dugo služi za svrhu generiranja portfelja, dok je duboko podržano učenje relativno nova metoda koja je znatno dobila na popularnosti u posljednjih nekoliko godina zbog njene široke primjene i velikih mogućnosti u raznim područjima. Pokazano je da na istom skupu dionica tehnika dubokog podržanog učenja ostvaruje znatno bolje rezultate od metode optimizacije srednje vrijednosti i varijance, posebno u smislu volatilnosti. Time je pokazan veliki potencijal koji sadrži podržano učenje u financijskoj domeni. U radu se detaljno opisuje arhitektura dubokog podržanog učenja koja se koristila u ovu svrhu.

Ključne riječi: duboko podržano učenje; metoda optimizacije srednje vrijednosti i varijance; generiranje portfelja

Abstract

This paper compares two different methods for generating portfolios of assets using historical prices of US companies. The mean-variance optimization method is a traditional method that has long served the purpose of portfolio generation, while deep reinforcement learning is a relatively new method that has gained considerable popularity in the last few years due to its wide application and great possibilities in various fields. It is shown that on the same set of stocks, the deep reinforcement learning technique achieves significantly better results than the mean and variance optimization method, especially in terms of volatility. This shows the great potential of reinforcement learning in the financial domain. The paper describes in detail the deep reinforcement learning architecture used for this purpose.

Keywords: deep reinforcement learning; mean-variance optimization; portfolio generation