

# Neuronales Netz Portfolio Optimierung

*Durch die Sharpe-Quotient*

by

Patrick Spohr

09. Februar 2024

Masterstudiengang

*Finanzmathematik, Aktuarwissenschaften und Risikomanagement*

Fachbereich

*Informatik, Kommunikation und Wirtschaft*

Lehrveranstaltung

*Statistical Learning in Finance and Insurance*

Professorin

*Dr. Christina Erlwein-Sayer*

# Inhalt

<b>1</b>	<b>Motivation</b>	<b>2</b>
1.1	Portfolio Optimierung . . . . .	2
<b>2</b>	<b>Literatur</b>	<b>3</b>
2.1	Definitionen . . . . .	3
2.2	Literaturüberblick . . . . .	3
2.3	Hauptliteratur . . . . .	4
<b>3</b>	<b>Ziele</b>	<b>4</b>
3.1	Grenzen und Vermutungen . . . . .	4
3.2	Hauptziele . . . . .	5
<b>4</b>	<b>Daten</b>	<b>5</b>
4.1	Daten und Quellen . . . . .	5
4.2	Überblick der zehn beliebigen Aktien . . . . .	6
<b>5</b>	<b>Methoden</b>	<b>7</b>
5.1	Künstliches Neuronales Netz . . . . .	7
5.2	Long Short-Term Memory . . . . .	8
5.3	Softmax-Funktion . . . . .	9
5.4	Sharpe-Quotient . . . . .	10
5.5	Struktur der zwei Modelle . . . . .	10
5.6	Monatliche Portfoliowerte . . . . .	11
<b>6</b>	<b>Ergebnisse</b>	<b>12</b>
6.1	Portfoliowerte und Renditen von 1990 bis 2020 . . . . .	12
6.2	Sharpe-Quotienten . . . . .	12
6.3	Fazit und Weiterentwicklung . . . . .	12
	<b>Literaturverzeichnis</b>	<b>13</b>

# 1 Motivation

## 1.1 Portfolio Optimierung

Stellen Sie sich eine Situation vor, wobei man vielversprechende Wertpapiere fand, aber er weiß nicht, wie sie in Portfolio zugeteilt werden dürfen. Markowitz [9] zufolge habe das Verfahren, um Portfolio aufzubauen, zwei Schritte. Am Anfang würden die einzelnen Wertpapiere geprüft, indem man Erfahrungen und Beobachtungen erwäge. Danach würde man mit dieser Information ein Portfolio aufbauen. Die einfachste Lösung ist die Wertpapiere mit einem Gleichgewicht versehen und alles unter Fach und Dach zu bringen und warum nicht, die Wertpapiere sind gut, aber das ändert das Portfolio um ein Vielfaches. Die Volatilität jeder neuen Wertpapiere trägt zu dem Portfolio bei und anschließend würde das Risiko des Portfolios erhöhen oder verringern und das könnte die Wünsche der Anleger widersprechen. In Modern Portfolio Theory (MPT) sollte die steigende Volatilität eine entsprechende Rendite erwirtschaften.

Nun komme ich zu dem Hauptteil meiner Untersuchung. Mit dem zweiten Schritt von Markowitz fange ich an, nämlich, geprüfte Wertpapiere in einem Portfolio zuzuteilen. Zuerst schränke ich die Auswahl der Wertpapiere um US-Aktien ein und dann sollten sie zugeteilt werden, indem ein künstliches neuronales Netz und ein Long Short-Term Memory Layer mit der Sharpe-Quotient als Verlustfunktion eingesetzt wird. Lassen Sie mich diese Begriffe und das Modell, mit dem ich die Gewichte der Aktien gerechnet habe, in der Überschrift "Methoden" erklären.

Im Fazit, um die Leistungen der Portfolios zu verstehen, werden die optimierten Portfolios mit einem gleichgewichteten Portfolio und einem Vergleichsindex verglichen, weil die Leistung eines Portfolios ohne einen Maßstab keine konkrete Folgerung ergibt.

## 2 Literatur

### 2.1 Definitionen

Als Einleitung definiere ich die Hauptbegriffe und deren Kurzwörter. Künstliches Neuronales Netz oder KNN sind nicht-lineare Algorithmen, die den Neuronen im Gehirn nachempfunden sind. Die Gestalt aller KNN hat eine Eingangsschicht, oft mehrere verborgenen Schichten und eine Ausgangsschicht [3]. Die verborgene Schicht in meinem KNN heißt Long Short-Term Memory oder LSTM. Man bezeichnet LSTM als eine Technik zur Verbesserung eines rekurrenten neuronalen Netzes, indem es um eine Art Erinnerung an frühere Erfahrungen verfügt, um das Problem des verschwindenden Gradienten zu lindern [8]. Die Verlustfunktion im KNN misst den Fehler zwischen einer Beobachtung und einer Schätzung einer Regression oder einer Klassifikation. Die Sharpe-Quotient, die ich als Verlustfunktion benutzte, ist eine Kennzahl, die die Rendite zur Volatilität eines Portfolios zeigt [11]. Zu guter Letzt ist der Risikofreier Zinssatz (Risk-Free Rate) oder RFZ, der in der Sharpe-Quotient eingesetzt wird, ein Referenzzinssatz, der theoretisch kein Ausfallrisiko enthält. Es gibt einige RFZ, die zur Verfügung stehen, und ich erläutere, den ich benutzte und warum.

### 2.2 Literaturüberblick

Ich bespreche kurz die Literatur, die mein Projekt leiteten. Die Hauptliteratur, von der ich die Idee und das Modell weiterentwickelte, heißt "Deep Learning for Portfolio Optimization" [13]. Ich erkläre diese Studie ausführlicher im nächsten Teil, denn es gibt Unterschiede, die ich mit meinem Projekt gegenüberstellen will. TensorFlow [1] und dessen Keras API [5] ermöglichte das künstliche neuronale Netz und gab mir die Gelegenheit alles schnell zu erschaffen. Andere Literatur war hilfreich alles zu verstehen, besonders mit den Begriffen künstliches neuronales Netz [3] und Long Short-Term Memory von Hochreiter und Schmidhuber, die das Verfahren initiierten.

## 2.3 Hauptliteratur

Im Anschluss daran werde ich auf die wichtigen Unterschiede zwischen meinem Projekt und der Studie von Zhang et al. [13] hinweisen. Der größte Unterschied war die Auswahl von Wertpapieren, weil sie Indexfonds oder Exchange-Traded Funds (ETFs) statt Aktien auswählten. Ein Portfolio mit ETFs von Aktien und Anleihen zeigt meines Erachtens die Kraft von diesem KNN-Modell nicht, weil ETFs schon relativ diversifiziert und nicht volatil sind. Ich wollte sehen, wie ein langfristiges Portfolio von einzelnen Aktien, rücksichtslos ihre erwarteten Renditen, optimiert wird. Indexfonds erwirtschaftet in der Regel positive Rendite, denn sie den Markt folgen und wirkt als eine sehr konservative Probe in diesem Experiment.

Als Nächstes stehe ich die Sharpe-Quotient im Fokus, die sie als Verlustfunktion anwendeten. Sie hatten keinen risikofreien Zinssatz in ihrem Modell benutzt, während ich ein Modell mit und ohne einen aufbaute. Der RFZ ist ein wichtiger Teil die Sharpe-Quotient und kann verschiedene Ergebnisse bewirken, wie in der Überschrift "Ergebnisse" dargestellt wird. In Hinsicht auf die Ergebnisse gibt es einen anderen Unterschied, den ich erwähnen muss, weil ich keine Handelskosten betrachte. Das lässt sich als eine wichtige Auslassung zeigen, denn die Aktien im ganzen Portfolio werden regelmäßig komplett wieder zugeteilt. Wegen dieser Berücksichtigung wurde zumindest ein Monat lang gewartet, bis die Gewichte des Portfolios wieder optimiert wurden.

## 3 Ziele

### 3.1 Grenzen und Vermutungen

Wie ich gerade erwähnt habe, wird das Portfolio nur mit Aktien aufgebaut. Angesichts dieser Bedingung schränke ich den Umfang des Portfolios weiter, um zu verständlichen Schlüssen zu führen. Zuerst kommen alle Aktien aus demselben Börse, NASDAQ, also nur US-Aktien. Zweitens werden die Aktien beliebig ausgewählt, die im Zeitraum 1990-2019 stetig gehandelt wurden. Von diesen Grenzen folgen die Vermutungen, dass alle Aktien schon geprüft wurden und zumindest zahlungsfähig waren, weil sie für 30 Jahre auf dem Markt gehandelt wurden. Konkret bedeutet, dass die Analysten in der Lage waren, Unternehmen zu vermeiden, die

bankrottgegangen waren. Das ist wichtig, weil das Modell nur funktioniert, wenn die Form die Daten gleich sind oder anders gesagt, die Anzahl der Tage auf dem Markt gleich sind. Außerdem wählte ich US-Treasury Staatsanleihen als einen RFZ, weil ich nur US-Aktien haben.

## 3.2 Hauptziele

Im Folgenden gehe ich auf die Hauptziele meines Projektes. Was hatte ich am Ende vor, zu erzielen? Am wichtigsten werden Portfolios von beliebigen Aktien durch die Sharpe-Quotient mit und ohne einen RFZ optimiert. Das heißt genau, dass ich die Gewichte jeder Aktie für einen Zeitraum erzeuge. Dadurch, dass ich die Portfoliowerte bekomme, multipliziere ich die Gewichte und Aktienkurs. Danach werden die Leistung von den optimierten Portfolios mit einem gleichgewichtigen Portfolio und einem Vergleichsindex verglichen. 30 Jahre über kann ich die monatliche Rendite jedes Portfolio rechnen, um die Sharpe-Quotient jedes Portfolio auszurechnen. Ich hatte keine Annahme, weil mein Projekt weit weg genug von der Hauptliteratur [13], dass ich nicht wusste, ob die optimierten Portfolios besser als ein gleichgewichtiges Portfolio oder ein Vergleichsindex würden. Im besten Fall würden die optimierten Portfolios dem Vergleichsindex überlegen sein.

# 4 Daten

## 4.1 Daten und Quellen

Ich gehe nun über zu den Daten und beschreibe, woher ich die Daten bekomme und was ich aus den Daten auszog. Zusammenfassend hatte ich drei Datenquellen: Kaggle, Yahoo! Finance, und FRED, der für Federal Reserve Economic Data steht. Ich erhielt den täglichen Aktienpreis und Datum aus Stock Market Dataset auf Kaggle [10]. Der Vergleichsindex ist der S&P 500 und kommt aus Yahoo! Finance [6]. Es ist zum Bemerken, dass der Ticker GSPC heißt. Aus dem GSPC erhielt ich auch den täglichen Indexpreis und Datum. Zuletzt bekam ich den risikofreien Zinssatz, 10-Year US Treasury Yield, von FRED [7]. Der RFZ hat ein Datum und ein tägliches Prozent. Seltsamerweise gab es paar Lücken im Datensatz

von FRED und ich musste deswegen die füllen, indem ich der Durchschnitt von dem frühen und nächsten Prozent rechnete. Ich denke, das beeinflusst die Ergebnisse nicht viel, weil es sich über 30 Jahren erstrecken, aber ich mitteile es trotzdem.

## 4.2 Überblick der zehn beliebigen Aktien

Bei den Aktien steht die Frage im Mittelpunkt, ob sie wirklich ein typisches oder realistisches Portfolio ähneln. Obwohl sie beliebig ausgewählt wurden, fand ich es wichtig, die relevant Information über sie auszurechnen, weil es passieren könnte, dass das Portfolio, egal von den Gewichten, den S&P 500 komplett übertrifft, wenn es Ausreißer wie Monster Beverage Corp (MNST) enthält, der über 150.000% Gesamtrendite von 1990 bis 2019 einbrachte.

<b>Ticker</b>	<b>Name</b>	<b>Geschäft</b>
COHR	Coherent, Inc.	Laser
CTB	Cooper Tire & Rubber Company	Reifen
EQT	EQT Corporation	Energie
GOLD	Barrick Gold Corporation	Bergarbeit
NSEC	National Security Group, Inc.	Versicherung
OTTR	Otter Tail Corporation	Energie
PEP	PepsiCo, Inc.	Getränke
SKYW	SkyWest, Inc.	Fluglinie
SNFCA	Security National Financial Corporation	Lebensversicherung
WY	Weyerhaeuser Company	Abholzung

Aus aller beliebig ausgewählten Aktien ist SkyWest, Inc. am besten, der die Summe aller monatlichen Renditen 635,7% erwirtschaftete und eine durchschnittliche monatliche Sharpe-Quotient 0,825 hatte. Alle Aktien im Portfolio hatten eine durchschnittliche monatliche Sharpe-Quotient zwischen null und eins, also es gibt keine außergewöhnliche Aktie. Die Aktien gehörten zu verschiedenen Industrien und deswegen ist das Portfolio diversifiziert in dieser Hinsicht. Die Aktien in der Tabelle oben baute das Portfolio auf, das ich im Modell eingesetzt hatte.

## 5 Methoden

### 5.1 Künstliches Neuronales Netz

Dieser Abschnitt behandelt die wichtigen Teile eines künstlichen neuronalen Netzes, besonders ein Feed-Forward KNN. Das KNN hat drei Teile: eine Eingangsschicht  $x_1, \dots, x_d$  mit  $d$  Inputs, eine oder mehrere verborgenen Schichten,  $z_1, \dots, z_M$  mit  $M$  Hidden-Units, und eine Ausgangsschicht  $y_1, \dots, y_c$  mit  $c$  Output-Units.

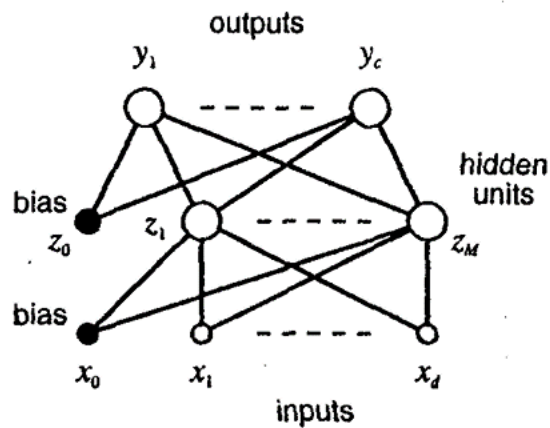


Figure 4.1. An example of a feed-forward network having two layers of adaptive weights. The bias parameters in the first layer are shown as weights from an extra input having a fixed value of  $x_0 = 1$ . Similarly, the bias parameters in the second layer are shown as weights from an extra hidden unit, with activation again fixed at  $z_0 = 1$ .

Figure 1: Feed-Forward Neuronales Netz [3]

Daten fließen in der Eingangsschicht, wo die Eingabe und die Gewichte jeder Neuronenverbindung multipliziert werden, um eine Ausgabewert zu ergeben. Für jedes Neuron wird der Aktivierungszustand durch eine Aktivierungsfunktion mit dessen Eingabewert bestimmt, sodass ein Neuron feuert oder ruht. Die Ausgabe dieser Aktivierungsfunktion geht weiter in den nächsten Schichten und wie vor, dieselben Verfahren gebrauchen, bis die Ausgangsschicht, wobei der Fehler der Schätzung durch eine Verlustfunktion gerechnet wird. Dann fängt Back-Propagation an, wo, von hinten nach vorne, die Gewichte jeder Neuronenverbindung mit



Hilfe des Fehlers durch ein Gradientenverfahren aktualisiert werden. Diese Feed-Forward und Back-Propagation Verfahren wiederholen sich, bis der Fehler minimisiert ist. Die Training-Genauigkeit sollte in der Regel nicht viel größer als die Validation-Genauigkeit, weil das Over-Fitting bedeuten könnte und weist darauf, dass das Modell dem Geräusch oder der Varianz folgt.

## 5.2 Long Short-Term Memory

Die verborgene Schicht meines Modelles ist LSTM und ich erkläre die Wichtigkeit dieser Weiterentwicklung des KNN-Verfahrens. Lassen Sie mich zuerst kurz die Schwäche vom traditionellen Feed-Forward Verfahren hervorheben. Feed-Forward KNN hat sogenannt kurze Erinnerung oder "Short-Term Memory," weil die Neuronen keine Informationen über frühere Neuronen verfügen. Normalerweise ist das kein Problem, weil die frühere Information keinen großen Einfluss auf die nächste Information haben. Das ändert sich, wenn Zeitreihen als Eingangsdaten einfließen, weil die früheren Daten wichtige Information enthalten, die die Ergebnisse verbessern können. Zum Beispiel sind Aktienkurse Zeitreihen und die früheren Preise beeinflussen die nächsten.

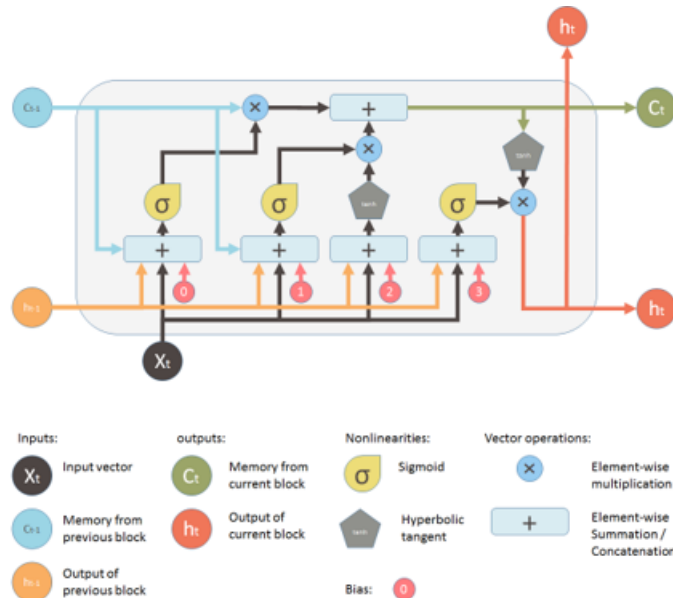


Figure 2: Long Short-Term Memory Unit [12]

Kurz gesagt, LSTM wurde gebildet, um Back-Flow-Problemen von Feed-Forward KNN "Short-Term Memory" zu lindern. Dadurch, dass ich das beste Verfahren für meine Daten benutzen will, wendete ich, wie meine Hauptliteratur [13], LSTM an. Ich bemerke auch, dass LSTM eine Art rekurrentes neuronales Netz oder RNN ist, weil die Ausgaben der früheren Neuronen erinnert werden. Im RNN heißt Back-Propagation unter einen anderen Namen: "Back-Propagation Through Time" oder BPTT. Der Grundidee bleibt wie vor, aber das Verfahren nutzt eine Reihe von Zeitschritten mit Eingaben und Ausgaben Paare und rechnet die kumulierten Fehler über jeden Zeitschritt [4].

### 5.3 Softmax-Funktion

Wie früher erwähnt hat KNN eine Aktivierungsfunktion. Mein Modell benutzte die Softmax-Funktion, die eine normalisierte Exponentialfunktion ist.

$$\sigma(x)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{für } j = 1, \dots, K \text{ und } x = \text{Input Vektor}$$

Die Softmax-Funktion wird eingesetzt, um eine Wahrscheinlichkeitsverteilung über alle unterschiedlichen möglichen Ereignisse zu rechnen. Die Funktion versichert für Klassifikation mit mehr als zwei Klassen, dass die Wahrscheinlichkeit aller Klassen eins ergibt und das ist wichtig für die Gewichte der Aktien im Portfolio, weil die Summe aller Gewichte auch eins ist. Es ist sinnvoll zu erwähnen, dass die Softmax-Funktion der Sigmoid-Funktion ähnelt aber die Summe aller Werte in der Sigmoid-Funktion nicht eins ist.

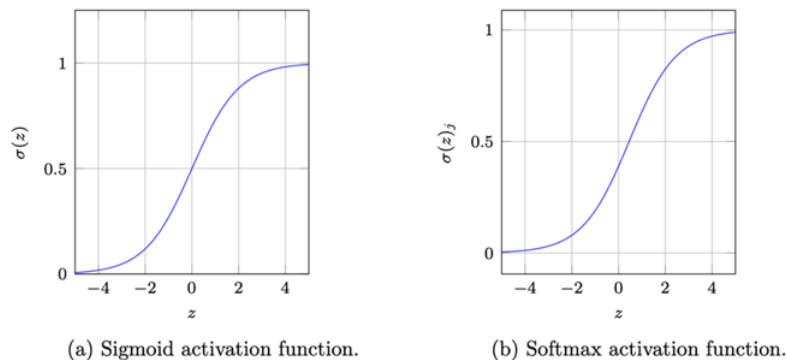


Figure 3: Long Short-Term Memory Unit [2]

## 5.4 Sharpe-Quotient

Anschließend stelle ich die Sharpe-Quotient vor, die als Verlustfunktion verwendet wird. Die Sharpe-Quotient enthält drei Teile: die Rendite des Portfolios, ein risikofrier Zinssatz und die Volatilität oder die Standardabweichung der Renditen. Was sagt die Sharpe-Quotient genau? Eine Sharpe-Quotient über eins bedeutet, dass das Portfolio eine Rendite mehr als der risikofreie Zinssatz erwirtschaftet und deren Ergebnis größer als die Volatilität der Rendite ist. Einfacherweise wird er Anleger für das Risiko hervorragend entschädigt. Eine Sharpe-Quotient gleich eins zeigt, dass Renditen und Risiken in einem ausgewogenen Verhältnis stehen. Unter eins zeigt eine unterdurchschnittliche Entschädigung für das entsprechende Risiko [11].

$$Sharpe = \frac{R_p - R_f}{\sigma_p}$$

$R_p$  = Portfolio Renditen  
 $R_f$  = risikofreier Zinssatz  
 $\sigma_p$  = Standardabweichung der Renditen

In meinem Model ist A Aktien mit x Close Preis die Features und das Y Target ist das y Gewicht für jede N Aktie. Typischerweise würde die Verlustfunktion minimiert aber die Sharpe-Quotient in meinem Modell wird negiert, weil das Minimum von einer negierten Funktion das Maximum ist. Das Maximum von der Sharpe-Quotient bedeutet, die größte Rendite für die entsprechende Volatilität des Portfolios.

## 5.5 Struktur der zwei Modelle

Ich bespreche kurz die überall Struktur der zwei Modelle, die ich benutzte, um das monatliche Gewicht für jede Aktie zu bestimmen. Das erste Modell bleibt ungeändert von dem Modell in der Hauptliteratur [13]. Das bedeutet, dass das Modell ein RNN mit einer LSTM-Schicht mit 64 Einheiten ist. Dieses erste Modell setzt keinen risikofreien Zinssatz in die Sharpe-Quotient-Verlustfunktion ein. Das zweite Modell beachtet den risikofreien Zinssatz, um der optimalen Gewichte zu generieren. Die Eingangsschicht enthält die täglichen Preise und tägliche prozentuelle Änderung jede Aktie. Am Ende ergibt das Modell das monatliche Gewicht für jede Aktie, die die Sharpe-Quotient optimiert.

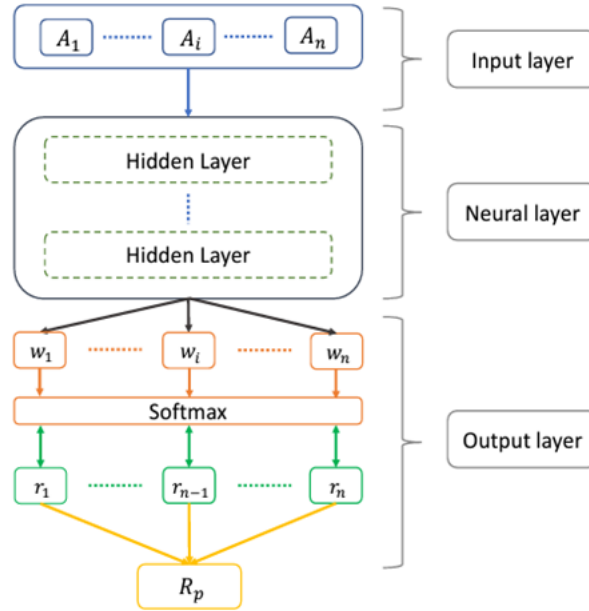


Figure 2: Model architecture schematic. Overall, our model contains three main building blocks: input layer, neural layer and output layer.

Figure 4: Long Short-Term Memory Unit [13]

## 5.6 Monatliche Portfoliowerte

Um die Leistung dieses Modell zu schätzen, wurde eine Simulation durchgeführt, wo man diese erzeugten Gewichte benutzen, um die Sharpe-Quotient des Portfolios zu optimieren und hohe Renditen zu erwirtschaften. Die monatliche Portfoliowerte wurde gerechnet, indem die optimalen Gewichte für jede Aktie im Portfolio für Zeit  $t$ , wurde mit den Aktienpreisen für Zeit  $t+1$  multipliziert. Das simulierte den Aktienhandel mit Hilfe von den erzeugten Gewichten des Modelles. Mit den monatlichen Portfoliowerten zur Hand konnten die monatlichen Renditen gerechnet werden. Das simulierte monatliches Trading und damit konnten die monatlichen Renditen summiert, um die Leistung des Portfolios auszuwerten.

$$V_p = \sum_{j=1}^A \sum_{i=1}^M P_{i+1}^T \cdot w_i$$

$M = 360$  Monate

$A = 10$  Aktien

$P$  = tägliche Preise von Aktie A im Monat  $M$

$w$  = monatliche Gewicht von Aktie A im Monat  $M$

$P_{1+1}$  = Preise im Februar 1990 von Aktie A

$w_1$  = Gewicht im Januar 1990 von Aktie A

Insgesamt wurden vier Portfolios miteinander verglichen, um die Leistungsfähigkeit des Modelles zu schätzen: S&P 500 Vergleichsindex, das gleichgewichtige Portfolio, das optimierte Portfolio ohne den risikofreien Zinssatz und das optimierte Portfolio mit dem risikofreien Zinssatz. Nachdem die Renditen der Portfolios gerechnet wurden, wurden die Sharpe-Quotienten jedes Portfolio gerechnet, um zu sehen, ob die Sharpe-Quotient jedes Portfolio maximiert wurde.

## 6 Ergebnisse

### 6.1 Portfoliowerte und Renditen von 1990 bis 2020

Hi

### 6.2 Sharpe-Quotienten

Hi

### 6.3 Fazit und Weiterentwicklung

Hi

# Literaturverzeichnis

- [1] Martin Abadi et al. “TensorFlow: A system for large-scale machine learning”. In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. 2016, pp. 265–283. URL: <https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf>.
- [2] Data Basecamp. May 2023. URL: <https://databasecamp.de/en/ml/softmax-function>.
- [3] Christopher M. Bishop. “Neural Networks for Pattern Recognition”. In: (1995). URL: <https://people.sabanciuniv.edu/berrin/cs512/lectures/Book-Bishop-Neural%20Networks%20for%20Pattern%20Recognition.pdf>.
- [4] Jason Brownlee. *A Gentle Introduction to Backpropagation Through Time*. Aug. 2020. URL: <https://machinelearningmastery.com/gentle-introduction-backpropagation-time/>.
- [5] François Chollet et al. *Keras*. 2015. URL: <https://keras.io>.
- [6] Yahoo! Finance. *S&P 500 (^GSPC)*. URL: <https://finance.yahoo.com/quote/%5EGSPC/>.
- [7] FRED. *Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis*. URL: <https://fred.stlouisfed.org/series/DGS10>.
- [8] Sepp Hochreiter and Jürgen Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735. eprint: <https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf>. URL: <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [9] Harry Markowitz. “Portfolio Selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. ISSN: 00221082, 15406261. URL: <http://www.jstor.org/stable/2975974>.
- [10] Oleh Onyshchak. *Stock Market Dataset*. URL: <https://www.kaggle.com/datasets/jacksoncrow/stock-market-dataset>.

- [11] William F. Sharpe. “The Sharpe Ratio”. In: *The Journal of Portfolio management* (1994). URL: <https://web.stanford.edu/~wfisharpe/art/sr/sr.htm>.
- [12] Shi Yan. *Understanding LSTM and its diagrams*. Mar. 2016. URL: <https://blog.mlreview.com/understanding-lstm-and-its-diagrams-37e2f46f1714>.
- [13] Zihao Zhang, Stefan Zohren, and Stephen Roberts. “Deep Learning for Portfolio Optimization”. In: *The Journal of Financial Data Science* 2.4 (Aug. 2020), pp. 8–20. ISSN: 2640-3943. DOI: 10.3905/jfds.2020.1.042. URL: <http://dx.doi.org/10.3905/jfds.2020.1.042>.