

# **Machine Learning and Technical Analysis for Time-series price forecasting in the financial market an empirical comparison of different approaches**

Sub-title of the thesis (leave empty if not required)

**Bachelor Thesis**

*Submitted to*

**IMC University of Applied Sciences Krems**



University of  
Applied Sciences

**Bachelor Programme Informatics**

by

**David Bobek**

for the award of academic degree  
**Bachelor of Science in Engineering (BSc)**

under the supervision of  
**Dipl.-Ing. Deepak Dhungana**

Submitted on 27.11.2023



# DECLARATION OF HONOUR

I declare on my word of honour that I have written this Bachelor Thesis on my own and that I have not used any sources or resources other than stated and that I have marked those passages and/or ideas that were either verbally or textually extracted from sources. This also applies to drawings, sketches, graphic representations as well as to sources from the internet. The Bachelor Thesis has not been submitted in this or similar form for assessment at any other domestic or foreign post-secondary educational institution and has not been published elsewhere. The present Bachelor Thesis complies with the version submitted electronically.

---

David Bobek  
27.11.2023



# ABSTRACT

This study is going to be examining various different approaches to time-series analysis and price forecastin based on Machine Learning and Technical Analysis. I will be focusing on the financial market and will be trying to predict and analyse the performance of various different technical indicators and Machine learning models The purpose of this thesis is to compare the return and performance of my approaches and determine which of them is the most suitable for different scenarios. By carrying out an empirical investigation and analyzing the outcomes, I will evaluate the advantages, disadvantages and discuss potential problems.

**Keywords:** Machine Learning, Technical Analysis, Time-series price forecasting



# ACKNOWLEDGEMENTS

This is an **optional** page. Use your choice of paragraph style for text on this page. Usually, this space is for thanking your supporters and getting emotional about how grateful you are to everyone.

# Table of Contents

Declaration of honour	iii
Abstract	iv
Acknowledgements	vii
Table of Contents	viii
List of Tables	ix
List of Figures	x
1 Introduction	1
2 Example Chapter	5
2.1 Code and syntax highlighting . . . . .	5
2.2 Labels and References . . . . .	6
2.3 Mathematical Equations and Expressions . . . . .	6
2.4 Enumerations and Descriptions . . . . .	6
2.5 Adding images . . . . .	7
2.6 Colors . . . . .	7
2.7 Just a poem by Emily Dickinson . . . . .	8
2.8 Tables . . . . .	8
Bibliography	9
Appendix A Example Appendix 1	11
Appendix B Example Appendix 2	13



# List of Tables

Table 2.1   Example table . . . . . 8

# List of Figures

Figure 2.1	Old IMC Logo . . . . .	7
Figure 2.2	Including sub images! . . . . .	7

# Chapter 1

## INTRODUCTION

In today's fast-paced capitalistically driven society in which everything is based on planning for the future and trying to optimize the present decision, a topic of forecasting is well suited. In this thesis, I will be combining my passion for Data Science and Machine Learning and will try to predict the market movement based on historically available data. Currently, there are various indicators I would like to explore which machine learning algorithm is suited the best for which type of trend pattern and see how my models can be used in the industry. There are several different problems that I will be facing during this research such as Feature Selection, Overfitting or Generalization, or potential external factors that could affect the Financial Market.

### **Motivation**

The goal of this thesis is to explore and help people predict market trends and eventually make their money. Trading money without prior experience is difficult and complicated. I have decided to pursue this research in order to make this process easier. Most of the people that start trading get too scared and tend to

"FOMO" (Fear of missing out) and make incorrect decisions. I would like to therefore analyze and help people with this issue. By finding and recognizing different

patterns I could theoretically help spot trend lines and elaborate on trend patterns. These models can also be used in the energy sector by predicting electricity or fuel prices which can be crucial for energy companies and consumers. Machine learning models can analyze historical data, weather patterns, and other relevant

factors to forecast energy prices accurately. This information helps energy can potentially assist companies with the optimization of production, planning infrastruc-

ture investments, and developing pricing strategies. Machine learning models that are based on time-series analysis and forecasting have various different use cases and I see overall a large potential in using this technology in order to help third parties. Overall, establishing machine learning models that estimate pricing provides useful insights to firms and individuals, allowing them to make informed decisions, improve operations, and plan successfully in dynamic market conditions.

### 1.2 Research Questions

"What are the comparative performance and predictive accuracy of machine learning algorithms for price forecasting in financial markets?" The purpose of this research question is to compare the performance and accuracy of several machine

learning algorithms in the context of price forecasting. By carrying out an empirical investigation and analyzing the outcomes, I will evaluate the advantages, disadvantages, and practical consequences of several machine learning algorithms for predicting financial market values, side by side to find the best for specific scenarios.

"How does a machine learning algorithm spot and classify a pattern in time-series data?" This research question seeks to elucidate the methods and strategies used

by machine learning algorithms to recognize significant patterns and derive important insights from time-series information. This question will be diving deeper

into where does a trend start and end as there are tens of different market trend patterns that tend to have similarities which could trick the algorithm

### 1.3 Research Method

In my research method, I am going to build a small application that will take data as an input. Process the data. Analyze it and based on a well-trained machine learning model will forecast the data for a set period of time. In terms of machine learning models, I am going to be using both Regression and Classification models. The

difference lies in the fact that in Regression models we are dealing with continuous output, usually represented by a floating value and in Classification we are

dealing with categorical output. Our regression output is going to be representing our time-series forecast over a certain amount of time units. In terms of categorical output, we will be classifying the trends in different categories. My progress will be following:

## 2

1. State of Art: Gathering, analyzing and extracting important information from already existing sources
2. Data Mining and cleaning : In order to get data of highest quality I will need to access it from trusted and verifiable sources. This data will most likely not be clean and I might need to do manual cleaning and feature extraction from it
3. Data Science and Analysis: This step will require a lot of Data engineering and digging down into what am I interested in data and which features are going to give us the best result based on our machine learning models
4. Training and Testing of Machine learning algorithms: By having high quality data I will be able to train the models in much fewer epochs. In this thesis I will be promoting less higher quality data over a lot of misleading and incorrect data.
- 5.

Performance Evaluation: In order to find the best machine learning model for different situations I am going to need to set a unit of measurement and will need to evaluate the performance of models. Suggested units are: (Accuracy, RMSE, MSE) • In the first step my main goal is to try to explore the thesis and information from relevant sources that are going to help me understand the complexity of this topic further. In order to stay on the correct path I will be thinking critically and selecting information from trusted and credited sources. My previous knowledge in financial markets will also be an advantage and will help me broaden the horizon of time-series analysis in financial markets. • The second step of my approach requires collecting valuable time-series data of various different markets. I will be trying to mainly focus on the American publicly traded stock market. Which will be the most beneficial for me as it is the most traded one and its movement impacts the world the most. I am expecting the collected data to not be specifically ready to use and I will have to get rid of potential issues. • Third step on my path will be Analyzing the data and performing an umbrella term called 'Data Science' which involves processes like data visualization, data exploration and statistical analysis of the data. This step will help me find

3

different trend patterns in my data and further understand what is required from me to get better results • Fourth step involves the actual creation of the machine learning models based

on training it on the train test. I am going to be experimenting with various different models and training each model until I can consider its results to be

significant enough. In this step I am also going to work with models and try to perfect them using hyper parameter tuning and focus on tweaking the settings of the model in order to improve its performance. As I am going to be dealing with both categorical and continuous data

• Fifth test is analysing the performance and evaluating them based on their performance on the performance metrics mentioned later

I will be using the following Machine learning models and analyzing their performance

• Decision Trees: Decision trees are adaptable models that can be used to forecast time series. They create a tree-like structure of decision rules by recursively partitioning the data depending on feature values. Decision trees

can handle numerical and categorical data, as well as non-linear relationships in the data.

• AdaBoost: AdaBoost is a technique for combining numerous weak learners (in this case, decision trees) to build a strong learner. Each weak learner is trained on a subset of the data, with a greater emphasis on previously misclassified occurrences. The predictions of the weak learners are then

weighted based on their performance. By combining the strengths of numerous decision trees, AdaBoost can increase overall predicting accuracy.

- Support Vector Regression (SVR): SVR is a machine learning model that performs regression tasks using support vector machines. By including lagged variables as input features, it can be modified for time series forecasting.
- Linear Regression: Linear Regression is the simplest machine learning model on this list, but in certain cases, its sheer speed is unrivaled. ARIMA (Autoregressive Integrated Moving Average) models are commonly used for time series analysis and forecasting. They can handle both stationary and non-stationary time series and capture autocorrelation and trend in the data.

4

- Seasonal ARIMA (SARIMA) models: SARIMA models are ARIMA extensions that incorporate seasonality in the data. They are appropriate for time series data that exhibit regular patterns over defined time intervals.
- K-means: K-means is a clustering algorithm that works on the principle of

grouping points that are close together into different clusters. This is a classification algorithm and could very well be used to find different parts of trends

in data.

- Neural Networks: Neural Network is the most complex approach in the list however the modern adaption of this model provides a lot of insights and blasting accuracy in some cases. By comparing these models in various different situations we can conclude which of these models is the most suitable in each scenario. All of these Algorithms will be tested evaluated and needed to make outcomes from

1.4 Structure This thesis will be composed out of 2 parts. The first part I will be focused on theory and explanation of my results and how each of the algorithms work with the trend pattern. This part will be mostly composed from our findings and already existing knowledge as a state of art on this problem. The second part will be a full working application based on already listed machine learning models classifying a certain set of data. This application will present half of my bachelor's thesis and therefore I really need to focus on correctness of my approach and mitigation of any potential failures that could lead to unclear results about which I will talk about later. The Bachelor's thesis will be stored in multiple locations in order to prevent any accidents of deletion. The technology stack I will use in this project will be explained in the further stages of this document

# Chapter 2

## EXAMPLE CHAPTER

This is only an example of a chapter! Anyways, all thesis should have a problem statement – not necessarily as a separate chapter though. Only after you know the problem, it will be possible for you to evaluate the results of what you did. If you want to see examples of evaluations, have a look at how graph visualizations are evaluated here [\[1\]](#).

### 2.1 Code and syntax highlighting

You may sometimes want to add code snippets to your thesis. You can do so by using `lstlisting`. Use this with care, as code should not be extensively presented in the thesis. Here is an example.

```
def addition ():  
    print("I_am_adding_numbers_here!")  
    n = float(input("Enter_the_number:_"))  
    t = 0 // Total number enter  
    ans = 0  
    while n != 0:  
        ans = ans + n  
        t+=1  
        n = float(input("Enter_another_number_(0_to_end):_"))  
    return [ans, t]
```

## 2.2 Labels and References

See [chapter 1](#) for interesting stuff and see a cool logo in [Figure 2.1](#). If you are still not convinced, try adding a footnote<sup>1</sup>. Its easy to add citations, just use a bibtex file to list your references and cite them here like this [\[3\]](#). If you want to read a cool paper [\[2\]](#), just contact the author of the paper. Haha, that was funny!

## 2.3 Mathematical Equations and Expressions

Basic equations in  $\text{\LaTeX}$  can be easily "programmed". Fermat's Last Theorem (sometimes called Fermat's conjecture, especially in older texts) states that no three positive integers  $a$ ,  $b$ , and  $c$  satisfy the equation

$$a^n + b^n = c^n$$

for any integer value of  $n$  greater than 2. The cases  $n = 1$  and  $n = 1$  have been known since antiquity to have infinitely many solutions. And because its so much fun, here is an integral for you - thank me later!

$$\int_0^1 x^2 + y^2 dx$$

Do you want a more complex formula, I have no idea what it means, but it looks pretty.

$$\prod_{i=1}^n \sum_{i=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - p^{-s}}$$

## 2.4 Enumerations and Descriptions

Here is a simple list:

1. The labels consists of sequential numbers.
2. The numbers starts at 1 with every call to the enumerate environment.

Here is another list:

1. The labels consists of sequential numbers.
  - The individual entries are indicated with a black dot, a so-called bullet.
  - The text in the entries may be of any length.

<sup>1</sup>did you like it?



2. The numbers starts at 1 with every call to the enumerate environment.

Maybe such descriptions are also useful. These look neat to me. What do you think?  
Oh, I forgot, this document is not a tutorial.

**Short** This is a shorter item label, and some text that talks about it. The text is wrapped into a paragraph, with successive lines indented.

**Rather longer label** This is a longer item label. As you can see, the text is not started a specified distance in – unlike with other lists – but is spaced a fixed distance from the end of the label.

## 2.5 Adding images

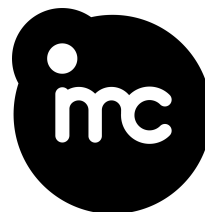
Adding a simple image is easy. Adding complex images is also easy. What is a complex image anyway?



Figure 2.1: Old IMC Logo



(a) New Logo since 2023



(b) New IMC Logo since 2023

Figure 2.2: Including sub images!

## 2.6 Colors

1. IMC Blue

```
\textcolor{imcblue}
```

## 2. IMC Color for Science and Technology

```
\textcolor{imctech}
```

## 3. IMC Corporate Color

```
\textcolor{imcgray}
```

## 4. IMC Corporate Color 2

```
\textcolor{imcorange}
```

## 2.7 Just a poem by Emily Dickinson

I'm nobody! Who are you?  
Are you nobody, too?  
Then there's a pair of us — don't tell!  
They'd banish us, you know.  
How dreary to be somebody!  
How public, like a frog  
To tell your name the livelong day  
To an admiring bog!

## 2.8 Tables

Country List			
Country Name or Area Name	ISO ALPHA 2 Code	ISO ALPHA 3 Code	ISO numeric Code
Afghanistan	AF	AFG	004
Aland Islands	AX	ALA	248
Albania	AL	ALB	008
Algeria	DZ	DZA	012
American Samoa	AS	ASM	016
Andorra	AD	AND	020
Angola	AO	AGO	024

Table 2.1: Example table

# BIBLIOGRAPHY

- [1] M. Burch, W. Huang, M. Wakefield, H. C. Purchase, D. Weiskopf, and J. Hua, "The state of the art in empirical user evaluation of graph visualizations," *IEEE Access*, vol. 9, pp. 4173–4198, 2021. [Online]. Available: <https://doi.org/10.1109/ACCESS.2020.3047616>
- [2] D. Dhungana, A. Haselböck, and S. Wallner, "Generation of multi-factory production plans: Enabling collaborative lot-size-one production," in *46th Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2020, Portoroz, Slovenia, August 26-28, 2020*. IEEE, 2020, pp. 529–536. [Online]. Available: <https://doi.org/10.1109/SEAA51224.2020.00088>
- [3] B. Huettner, "The elements of technical writing (2nd ed.) - book review," *IEEE Transactions on Professional Communication*, vol. 45, no. 1, pp. 59–60, 2002.



# Appendix A

## EXAMPLE APPENDIX 1

Appendices should be used for supplemental information that does not form part of the main research. Remember that figures and tables in appendices should not be listed in the List of Figures or List of Tables.



# Appendix B

## EXAMPLE APPENDIX 2

Appendices should be used for supplemental information that does not form part of the main research. Remember that figures and tables in appendices should not be listed in the List of Figures or List of Tables.