

# **A Report**

**on**

## **Sentiment Analysis in Portfolio Management**

**S2-22\_MBAZG622T : Final Project Report**

**Project Work Done by: K Pravin Kumar**

**BITS ID: 2021hb59011**

**Dissertation work carried out at  
[JOYRTS Limited , Canada]**

# **Sentiment Analysis in Portfolio Management**

**S2-22\_MBAZG622T : Final Project Report**

**Project Work Done by: K Pravin Kumar  
BITS ID: 2021hb59011**

**Dissertation work carried out at  
[JOYRTS Limited , Canada]**

**Submitted in fulfilment of MBA(Finance)  
project work**

**Under the Supervision of  
Ajex Rajadurai  
[ProPlus Logics , Coimbatore]**



**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE,  
PILANI  
VIDYA VIHAR, PILANI, RAJASTHAN -  
333031.**

**April , 2023**

# Abstract

In this project, we explore the use of sentiment analysis and the Sharpe ratio to select stocks for a portfolio. Sentiment analysis is the process of analyzing text data to determine the sentiment expressed in it, such as positive, negative, or neutral. The Sharpe ratio is a measure of the risk-adjusted performance of an investment portfolio.

We first gather historical price data for a set of stocks and news articles related to these stocks. We then perform sentiment analysis on the news articles to calculate a sentiment score for each stock. We use these scores along with the Sharpe ratio to select stocks for our portfolio.

We then simulate a backtest of our portfolio over a historical time period and calculate the actual returns using the selected stocks. We compare the actual returns with the expected returns based on our sentiment analysis and Sharpe ratio approach.

Finally, we evaluate the performance of our approach and discuss its limitations and potential improvements.

Overall, this project demonstrates the use of sentiment analysis and the Sharpe ratio as a tool for stock selection and portfolio optimization, and highlights the potential benefits and challenges of using these approaches in practice.



**Signature of Student**  
**Date: 23/04/2023**



**Signature of your Supervisor**  
**Date:- 23/04/2023**

## Contents

* Modules in portfolio management	5
* Data collection	5
* Sentiment analysis	6
* Portfolio optimization	7
* BackTesting and investing	9
* Plan of work	10
* Literature references	11
* Supervisor Approval	12
* Checklist	14

## 1. Modules

There are several modules that can be used in portfolio management with sentiment analysis, each serving a specific purpose::

1. **Data Collection:** The first step in automating portfolio management with sentiment analysis is to collect data on market sentiment. This can be done by using APIs to scrape news articles, social media feeds, and other sources of market sentiment data. Alternatively, some data vendors provide sentiment data feeds that can be integrated directly into your portfolio management system.
2. **Sentiment Analysis:** Once the data has been collected, it needs to be analyzed to determine the overall sentiment of the market. Sentiment analysis algorithms can be used to automatically classify the sentiment of news articles, social media posts, and other sources of data. This analysis can be used to generate sentiment scores for different companies, sectors, or asset classes.
3. **Portfolio Optimization:** Once the portfolio has been constructed, optimization algorithms can be used to further refine the portfolio to achieve specific objectives, such as maximizing returns or minimizing risk. Optimization algorithms can take into account factors such as transaction costs, tax implications, and portfolio constraints.
4. **Integration and Testing :** The sentiment analysis outputs need to be integrated with portfolio optimization such that the weights of the stocks is calculated by taking sentiment values into consideration .Finally, the portfolio should be monitored on an ongoing basis to ensure that it remains aligned with the sentiment of the market.

## 2. Data collection

Code : [https://github.com/joyrts/portfolio\\_analysis/blob/main/2021hb59011.ipynb](https://github.com/joyrts/portfolio_analysis/blob/main/2021hb59011.ipynb)

For purpose of Sentiment analysis relieve data from newsapi so that we can control the automation of collecting news data . To collect data for sentiment analysis from NewsAPI, you can follow these steps:

1. **Sign up for a NewsAPI account:** You will need an API key to access the NewsAPI. You can sign up for a free account at <https://newsapi.org/register>.
2. **Choose your stocks:** NewsAPI offers access to thousands of news sources from around the world. You can choose the stocks that are most relevant to your analysis.

3. Retrieve news articles: Once you have identified the sources you want to use, you can use the NewsAPI to retrieve the news articles. You can specify a variety of parameters, including the source, the date range, and the keywords you are interested in.

```
newsapi = NewsApiClient(api_key= NEWS_API_KEY)
keywrdr = <ticker name>
my_date = date.today() - timedelta(days=1)
# datetime.strptime('10-Apr-2021', '%d-%b-%Y')

articles = newsapi.get_everything(q = keywrdr,
                                  from_param = my_date.isoformat(),
                                  to = (my_date + timedelta(days = 1)).isoformat(),
                                  language="en",
                                  #sources = ",".join(sources_list),
                                  sort_by="relevancy",
                                  page_size = 100)
```

For stock analysis for portfolio optimization the historical trading data can be retrieved from yahoo finance . To retrieve stock data from Yahoo Finance, you can use their API or web scraping methods

```
import yfinance as yf
stock_data = yf.download(stock_symbol, start=start_date, end=end_date)
```

### 3. Sentiment analysis

Code : [https://github.com/joyrts/portfolio\\_analysis/blob/main/2021hb59011.ipynb](https://github.com/joyrts/portfolio_analysis/blob/main/2021hb59011.ipynb)

The main objectives of sentiment analysis in portfolio management are to use information about public opinion and market sentiment to inform investment decisions and improve the performance of a portfolio.

- Identify trends and patterns in public opinion: Sentiment analysis can be used to identify trends and patterns in public opinion about a particular stock or market trend, which can help portfolio managers to make more informed investment decisions.
- Measure and anticipate market reactions: Sentiment analysis can be used to measure and anticipate market reactions to news events, announcements, and other factors that may affect a stock's performance.

- Quantify and manage risk: Sentiment analysis can be used to quantify and manage risk by identifying negative sentiment that may indicate a stock is overvalued, or positive sentiment that may indicate a stock is undervalued.
- Monitor brand reputation: Sentiment analysis can also be used to monitor brand reputation by identifying and addressing negative sentiment before it becomes a major problem.
- Identify potential opportunities: Sentiment analysis can be used to identify potential investment opportunities by identifying positive sentiment about a particular stock or market trend that may indicate future growth.

Once you have the text from the articles, you can perform sentiment analysis using a natural language processing (NLP) library like NLTK. These libraries provide functions to analyze the sentiment of a piece of text and return a score indicating whether the text is positive, negative, or neutral.

```
from nltk.sentiment.vader import SentimentIntensityAnalyzer  
SentimentIntensityAnalyzer().polarity_scores()
```

Once you have performed sentiment analysis on the news articles, you can analyze the results to identify patterns or trends. You may want to compare the sentiment of articles from different sources or look for changes in sentiment over time.

#### 4. Portfolio optimization

Code : [https://github.com/joyrts/portfolio\\_analysis/blob/main/2021hb59011.ipynb](https://github.com/joyrts/portfolio_analysis/blob/main/2021hb59011.ipynb)

Portfolio optimization using the Sharpe ratio involves selecting a combination of assets that provides the highest possible return for a given level of risk. The Sharpe ratio is a measure of risk-adjusted return and is calculated as the excess return of the portfolio over the risk-free rate divided by the standard deviation of the portfolio's return.

The steps for portfolio optimization using the Sharpe ratio are as follows:

- Define the investment universe: Identify the set of assets that are available for investment. These assets can be stocks, bonds, mutual funds, ETFs, etc.

- Determine the expected returns: Estimate the expected returns for each asset in the investment universe. This can be done by analyzing historical data, market trends, financial statements, and other relevant information.
- Calculate the standard deviation: Calculate the standard deviation of the returns for each asset in the investment universe. This measures the level of risk associated with each asset.
- Determine the correlation matrix: Calculate the correlation coefficient between each pair of assets in the investment universe. This helps to identify the diversification benefits of adding a particular asset to the portfolio.
- Define the constraints: Define any constraints that must be taken into account when constructing the portfolio, such as maximum and minimum weights, sector diversification requirements, and other investment guidelines.
- Optimize the portfolio: Use an optimization model or software to find the combination of assets that maximizes the Sharpe ratio while satisfying the constraints. This involves solving a complex mathematical problem that takes into account the expected returns, standard deviation, and correlation matrix for each asset.

To calculate the maximum Sharpe ratio using Yahoo Finance code, you can use Python and its libraries like pandas, numpy, yfinance, and scipy. Here's an example code that fetches historical price data from Yahoo Finance, calculates the optimal portfolio weights, and displays the results

After the data is downloaded the following steps need to be performed

- \* Calculate the expected returns and covariance matrix of the assets

```
returns = np.log(prices / prices.shift(1)).mean()
```

- \* Build covariance matrix

```
np.log(prices / prices.shift(1)).cov()
```

- \* Calculate portfolio returns using the following formula

**Weightage of the stock in the portfolio \* returns**  
**Calculate portfolio volatility**

```
portfolio_volatility = np.sqrt(np.dot(weights.T, np.dot(cov_matrix, weights)))
```



\* Calculate Sharpe ratio

$$\text{Sharpe ratio} = (\text{portfolio\_return} - \text{risk\_free\_rate}) / \text{portfolio\_volatility}$$

Calculation of optimal weights for the portfolio using recursive programming . The program downloads the historical price data of the assets specified in the tickers list, calculates the expected returns and covariance matrix, defines an objective function that maximizes the Sharpe ratio, defines the constraints and bounds of the optimization problem, solves the problem using the minimize function from scipy.optimize, and prints the optimal portfolio weights and the maximum Sharpe ratio.solve the optimization problem to find the optimal portfolio weights

```
initial_weights = [1 / len(tickers) for _ in range(len(tickers))]  
opt_result = minimize(neg_sharpe_ratio, initial_weights, args=(returns, cov_matrix),  
method='SLSQP', bounds=bounds, constraints=constraints)  
opt_weights = opt_result.x
```

Now lets add the sentiment sensor to adjust the weights so that we can use the sharpie ratio to calculate the the optimal portfolio weights

```
portfolio_return = np.dot(weights, returns.values) + np.dot(weights,  
list(sentiment_scores.values()))  
portfolio_volatility = np.sqrt(np.dot(weights.T, np.dot(cov_matrix, weights)))  
sharpe_ratio = portfolio_return / portfolio_volatility
```

## 5. Backtesting and investing

Code : [https://github.com/joyrts/portfolio\\_analysis/blob/main/testing.ipynb](https://github.com/joyrts/portfolio_analysis/blob/main/testing.ipynb)

We first define the tickers to backtest, the date range, and the weights of the portfolio. We then fetch the historical prices of the assets using the yf.download() function from the yfinance library.

```
tickers = ['AC.TO', 'RNW.TO', 'FTS.TO', 'NVEI.TO', 'SU.TO']  
start_date = '2023-04-15'  
end_date = '2023-04-22'
```

Next, we calculate the daily returns of the assets using the pct\_change() method of the prices DataFrame. We then use the dot product (dot()) of the returns DataFrame and the weights list to calculate the daily returns of the portfolio.

```
returns = prices.pct_change()  
  
# Calculate the daily returns of the portfolio  
portfolio_returns = returns.dot(weights)
```

Finally, we calculate the actual portfolio return over the past week by taking the product of the daily returns of the portfolio plus one, and subtracting one from the result. We then print the actual portfolio return over the past week using the `print()` function with a formatted string to display the result as a percentage. Note that this is just an example code and you may need to customize it to suit your needs.

## 6. Plan of Work

Phases	Start Date-End Date	Work to be done	Status
Abstract	11 Jan 2023 – 22 Jan 2023	Literature Review and prepare Dissertation abstract/outline	Completed
Design & Development	23 Jan 2023 – 28 Feb 2023	Design & Development Activity	Completed
Mid Semester Report/ Review	15 Mar 2023 - 22 Mar 2023	Submit Dissertation to Supervisor & Additional Examiner for review and feedback	Completed
Integration	22 Mar 2023 - 31 Mar 2023	Integrating sentiment analysis and portfolio optimization modules	Completed
Testing	01 Apr 2023 - 15 Apr 2023	Software Testing, User Evaluation & Conclusion	Completed
Final Submission	16 Apr 2023 - 23 Apr 2023	Final Review and submission of Dissertation	Completed

## 6. Literature References

The state art is the base for any successful project. In current project, the literature inclined towards the new domain of conversational information retrieval is considered. The following are referred journals from the preliminary literature review.

1. *"Python Machine Learning" by Sebastian Raschka and Vahid Mirjalili - This book provides an introduction to machine learning using the Python programming language and covers a wide range of machine learning libraries, such as scikit-learn, TensorFlow, and Keras.*
2. *"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron - This book is an practical guide to machine learning that covers the use of the most popular machine learning libraries, such as scikit-learn, Keras, and TensorFlow.*
3. *"The Intelligent Investor" by Benjamin Graham - This book is considered a classic in the field of value investing and provides a comprehensive guide to making sound investment decisions. It covers the principles of investment, risk management and valuation, and how to apply them in the stock market.*
4. *"Portfolio Theory and Capital Markets" by William Sharpe - This book is considered a classic in the field of portfolio management and is widely used as a textbook in finance courses. It covers the Capital Asset Pricing Model (CAPM) and Modern Portfolio Theory (MPT) in depth.*
5. *"Natural Language Processing with Python" by Steven Bird, Ewan Klein, and Edward Loper - This book provides a comprehensive introduction to natural language processing and covers a wide range of techniques and tools, including the use of the NLTK library in Python.*
6. *"Sentiment Analysis and Opinion Mining" by Bing Liu - This book provides a comprehensive introduction to sentiment analysis and opinion mining, including the use of natural language processing and machine learning techniques.*

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**  
**WORK INTEGRATED LEARNING PROGRAMMES (WILP) DIVISION**  
**First Semester 2022-2023**

**S2-22\_MBAZG622T - MBA(Finance) - Pre-Final Evaluation Sheet**

**BITS ID No. : 2021hb59011**

**NAME OF THE STUDENT : K Pravin Kumar**

**EMAIL ADDRESS : kpravinraj@gmail.com**

**NAME OF THE SUPERVISOR: Ajex Rajadurai**

**PROJECT TITLE : Sentiment Analysis in Portfolio Management**

***Project Work Final Evaluation (Please put a tick (☑) mark in the appropriate box)***

S. No.	Evaluation Component	Excellent	Good	Fair	Poor
1.	Final Project Report	Excellent			
2.	Final Seminar and Viva-Voce	Excellent			

S. No.	Evaluation Criteria	Excellent	Good	Fair	Poor
1	Technical/Professional Competence	Excellent			
2	Work Progress and Achievements	Excellent			
3	Documentation and expression	Excellent			
4	Initiative and Originality	Excellent			
5	Research & Innovation	Excellent			
6	Relevance to the work environment	Excellent			

Please ENCIRCLE the Recommended Final Grade: **Excellent** / Good / Fair / Poor

**Remarks of the Supervisor:**

Pravin , Great job on your project on portfolio optimization and management with sentiment analysis! Your work demonstrated a thorough understanding of the subject matter and showcased your ability to apply various analytical techniques to the problem at hand. I was particularly impressed with your use of sentiment analysis to inform your investment decisions. This innovative approach shows a deep understanding of the importance of considering qualitative factors alongside quantitative data. Your analysis was comprehensive and your recommendations were well-supported by your findings. You also showed strong attention to detail in your implementation of the optimization algorithms, which resulted in a well-balanced and diversified portfolio. Overall, your project was well-organized, well-executed, and demonstrated a high level of technical proficiency. I have no doubt that your work will have a significant impact on the field of portfolio optimization and management. Congratulations on a job well done!

	<b>Supervisor</b>
Name	G . Ajax Rajadurai
Qualification	BE , MBA
Designation	CEO , ProPlus Logics Solutions Pvt. Ltd
Employing Organization & Location	ProPlus Logics Solutions Pvt. Ltd
Phone Number	
Mobile Number	+91 9994434213
Email Address	<a href="mailto:ajex@propluslogics.com">ajex@propluslogics.com</a>
Signature	
Place & Date	23/04/2023 - Coimbatore , India

## **Checklist of Items for the Final Dissertation / Project / Project Work Report**

This checklist is to be attached as the last page of the final report.

**This checklist is to be duly completed, verified and signed by the student.**

	<b>Is the final report neatly formatted with all the elements required for a technical Report?</b>	<b>Yes</b>
	Is the Cover page in proper format as given in Annexure A?	<b>Yes</b>
	Is the Title page (Inner cover page) in proper format?	<b>Yes</b>
	(a) Is the Certificate from the Supervisor in proper format? (b) Has it been signed by the Supervisor?	<b>Yes</b> <b>Yes</b>
	Is the Abstract included in the report properly written within one page? Have the technical keywords been specified properly?	<b>Yes</b> <b>Yes</b>
	Is the title of your report appropriate? <b>The title should be adequately descriptive, precise and must reflect scope of the actual work done.</b> Uncommon abbreviations / Acronyms should not be used in the title	<b>Yes</b>
	Have you included the List of abbreviations / Acronyms?	<b>Yes</b>
	Does the Report contain a summary of the literature survey?	<b>Yes</b>
	Does the Table of Contents include page numbers? i. Are the Pages numbered properly? (Ch. 1 should start on Page # 1) ii. Are the Figures numbered properly? (Figure Numbers and Figure Titles should be at the bottom of the figures) iii. Are the Tables numbered properly? (Table Numbers and Table Titles should be at the top of the tables) iv. Are the Captions for the Figures and Tables proper? v. Are the Appendices numbered properly? Are their titles appropriate	<b>Yes</b> <b>Yes</b> <b>Yes</b> <b>Yes</b> <b>Yes</b> <b>Yes</b>
	Is the conclusion of the Report based on discussion of the work?	<b>Yes</b>
	Are References or Bibliography given at the end of the Report? Have the References been cited properly inside the text of the Report? Are all the references cited in the body of the report	<b>Yes /</b> <b>Yes</b> <b>Yes</b>
	Is the report format and content according to the guidelines? The report should not be a mere printout of a PowerPoint Presentation, or a user manual. Source code of software need not be included in the report.	<b>Yes</b>

### **Declaration by Student:**

I certify that I have properly verified all the items in this checklist and ensure that the report is in proper format as specified in the course handout.

**Place:** Toronto, Canada  
**Date:** 23/04/2023

**Signature of the Student**  
**Name:** K Pravin Kumar

*K. Pravin Kumar*

**ID No.: 2021hb59011**

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI**  
**WORK INTEGRATED LEARNING PROGRAMMES (WILP) DIVISION**  
**SECOND SEMESTER OF ACADEMIC YEAR 2022-2023**

**S2-22\_MBAZG622T : PROJECT OUTLINE**

<b>STUDENT ID No.</b>	2021hb59011
<b>NAME OF THE STUDENT</b>	K Pravin Kumar
<b>STUDENT'S EMAIL ADDRESS</b>	<a href="mailto:021hb59011@wilp.bits-pilani.ac.in">021hb59011@wilp.bits-pilani.ac.in</a>
<b>STUDENT'S EMPLOYING ORGANIZATION &amp; LOCATION</b>	JOYRTS Limited , Canada
<b>SUPERVISOR'S NAME</b>	G . Ajex Rajadurai
<b>SUPERVISOR'S EMPLOYING ORGANIZATION &amp; LOCATION</b>	ProPlus Logics Solutions Pvt. Ltd
<b>SUPERVISOR'S EMAIL ADDRESS</b>	ajex@propluslogics.com
<b>ADDITIONAL EXAMINER'S NAME</b>	Pradhi Suresh Mathialagan
<b>ADDITIONAL EXAMINER'S EMPLOYING ORGANIZATION &amp; LOCATION</b>	Cognizant Technology Solutions
<b>ADDITIONAL EXAMINER'S EMAIL ADDRESS</b>	pradhisuresh.mathialagan@cognizant.com
<b>DISSERTATION / PROJECT / PROJECT WORK TITLE</b>	Sentiment Analysis in Portfolio Management