Personalized Investment Recommendations Using Recommendation Systems:

Meeting the Growing Demand for Tailored Investment Solutions for Institutional Investors

Lorik Sadriu

Lund University School of Economics and Management

Master Thesis - Finance

Supervisor: Luca Margaritella

NEKN02

May 2023

Abstract

This paper presents a Deep Learning-based Hybrid Recommendation System (DLHR) designed specifically for institutional investors with public portfolio holdings on the Stockholm Stock Exchange. The objective is to provide personalized investment recommendations, complement existing portfolios, and explore untapped cross-selling opportunities. We evaluate the DLHR's effectiveness using selected metrics and compare it to widely used baseline methods. We utilize the mean-variance spanning tests to assess the potential for expanding the investment opportunity set. We find that investors seeking to maximize the Sharpe Ratio would have benefitted from adding the recommended stocks. Our results demonstrate that leveraging deep learning techniques enables effective identification and exploitation of untapped cross-selling opportunities, leading to enhanced portfolio performance and improved risk management. Thus, we conclude that implementing the DLHR approach on the Swedish stock market can provide personalized investment recommendations for institutional investors. This research highlights the practicality of employing machine learning techniques within the financial sector.

Contents

List of Tables

List of Figures

List of Abbreviations

AUM Assets Under Management

CBF Content-based Filtering

CF Collaborative Filtering

CRM Client Relationship Management

DLHR Deep Learning based Hybrid Recommendation System

HF Hybrid Filtering

kNN k-Nearest Neighbors

MAE Mean Absolute Error

MSE Mean Squared Error

1 Introduction

Learning to choose is hard. Learning to choose well is harder. And learning to choose well in a world of unlimited possibilities is harder still, perhaps too hard.

Barry Schwartz (2004)

As the financial service industry is expanding, investors now have access to an unprecedented number of investment options. For example, investors can choose from a range of stocks, bonds, mutual funds, exchange-traded funds (ETFs), and alternative investments like real estate and commodities. While this can be positive, the sheer magnitude of choices can also be overwhelming and confusing, leaving investors struggling to identify the best investment products that align with their needs and preferences. It's clear that investors require more than just generic advice - they need investment solutions tailored to their unique situations. This requirement has created a growing demand for investment advice that is both effective and personalized, and the race is on to develop solutions that can meet this need.

To meet this need, utilizing machine learning is becoming increasingly popular. Machine learning involves the development of algorithms and models capable of automatically learning and improving from data, without explicit programming. It relies on statistical techniques to analyze large datasets and identify patterns, trends, and relationships that can be used to make predictions or decisions. Recommendation systems are especially popular in this field. They use algorithms to suggest potential products based on specific inputs and criteria derived from sample data. Over the past two decades, personalized recommendation systems have become a frequent tool in several industries to offer customized product and service suggestions based on the specific preferences and requirements of the consumers.

Conventional recommendation systems comprise content-based filtering (CBF), which bases its recommendations on user/item characteristics, or collaborative filtering (CF), which bases its recommendations on past interactions. Hybrid filtering (HF) combines the strengths of both methods and delivers more effective and personalized recommendations. However, recent research has shown that deep learning techniques can further enhance the effectiveness of recommendation systems. Deep learning, a branch of machine learning, revolves around the training of artificial neural networks, which are models inspired by the intricate workings of the human brain. These networks can learn from data, discern patterns, and subsequently utilize this knowledge to make predictions or decisions. Such acquired insights find applications in a multitude of tasks, including but not limited to classification and prediction (Kiran et al., 2020).

Netflix has been a pioneer in leveraging machine learning to enhance its recommendation system. The company introduced tailored movie recommendations as early as 2000 and launched the Netflix Prize in 2006, a competition that leveraged machine learning and data mining to improve its recommendation system (Bennett et al., 2007). Today, Netflix's personalized recommendation system remains highly effective, with an impressive 80% of stream time being attributed to the system. This has resulted in significant cost savings for the company, estimated to be around $1B per year as of 2016. The effectiveness of personalized recommendation systems has been demonstrated in several industries, from e-commerce to entertainment, indicating the potential for similar success in the financial sector.

However, the majority of the prior research on recommendation systems for investment recommendation has focused on two main areas:
a)Providing investment portfolio recommendations to individual investors (Gonzales & Hargreaves, 2022) or
b)Supporting financial advisors in devising investment strategies (Musto et al., 2015). The body of research on institutional investors is relatively sparse in the area, despite their unique characteristics and potential for cross-selling opportunities. Institutional investors tend to be larger, have specific investment goals and risk preferences, and are subject to regulatory requirements. Moreover, the research in the financial area has mainly focused on conventional approaches like collaborative filtering (Sayyed et al., 2013b), content-based filtering (González et al., 2015), or hybrid filtering (Tseng, 2004) and the use of deep learning-based hybrid recommendation systems have been largely unexplored. While these approaches have shown promise in some cases, they have limitations, such as the cold start problem and data sparsity Zibriczky (2016).

In this context, this study aims to develop and test a deep learning-based hybrid recommendation system (DLHR) specifically designed for institutional investors who have public portfolio holdings on the Stockholm Stock Exchange. The objective is to provide personalized investment recommendations that complement their current portfolios and explore untapped cross-selling opportunities that have not been addressed in previous research. Through this investigation, valuable insights can be gained regarding the effectiveness of personalized recommendation systems for institutional investors and the potential advantages of employing deep learning-based techniques in the financial service industry.

To evaluate the effectiveness of the DLHR, the study will employ selected evaluation metrics to assess its precision in providing recommendations. These metrics will allow for a comprehensive assessment of the accuracy and quality of the personalized investment recommendations generated by the DLHR method. By evaluating the system's precision, the study aims to determine its efficacy in assisting institutional investors in making informed investment decisions.

Overall, this research aims to contribute to the existing literature by:
a) Providing insights into the effectiveness of personalized recommendations for institutional investors and the enhancement of cross-selling opportunities.
b) Advancing the understanding of the potential benefits of deep learning-based recommendation techniques in the financial service industry.

The thesis is organized as follows: section 2 presents the theoretical background of recommendation systems. Section 3 presents relevant literature findings underpinning this study's investigation into the employment of DLHR techniques in the financial service industry and the target group, institutional investors. Section 4 provides an overview and analysis of the data and back-testing procedure used for the empirical approach, followed by a summary of the methodology of the executed experiment in Section 5. Section 6 presents the results in conjunction with an analysis. The thesis ends with concluding remarks in section 7.

2 Recommendation Systems

To motivate our empirical analysis and shed light on the underlying model, we provide an introduction to recommendation systems and the different types of filtering. Specifically, we will begin by introducing the necessary theoretical prerequisites before delving into the existing literature on the models used.

2.1 Recommendation Systems

Recommendation systems address information overload commonly faced by users by offering personalized recommendations for exclusive content and services. The systems utilize algorithms that suggest items, products, or services to users based on different criteria. These systems typically use machine learning techniques to analyze user data and generate personalized recommendations.

Researchers have divided opinions regarding the application of recommendation systems to real-world financial problems. Several researchers mentioned below have successfully applied a recommendation system approach to solving real-world financial problems. Other researchers criticize the approach to real-world financial problems.

To understand the background and application of recommendation systems and the system deployed in this paper, one must understand the different types of recommendation systems. There are three main types of recommendation systems: collaborative filtering and content-based filtering, as well as the combination of both; hybrid filtering.

2.2 Collaborative Filtering

Collaborative filtering (CF) involves using the preferences of users with similar characteristics to the target user, also known as user-to-user filtering, to recommend products and services based on what the similar user has chosen. In addition, it also uses the preferences of the target user by tracking their likes and dislikes for product-related information to predict personal preferences and recommend products with similar characteristics, also known as user-to-product filtering. In conclusion CF generates recommendations by finding similarities between users or items.

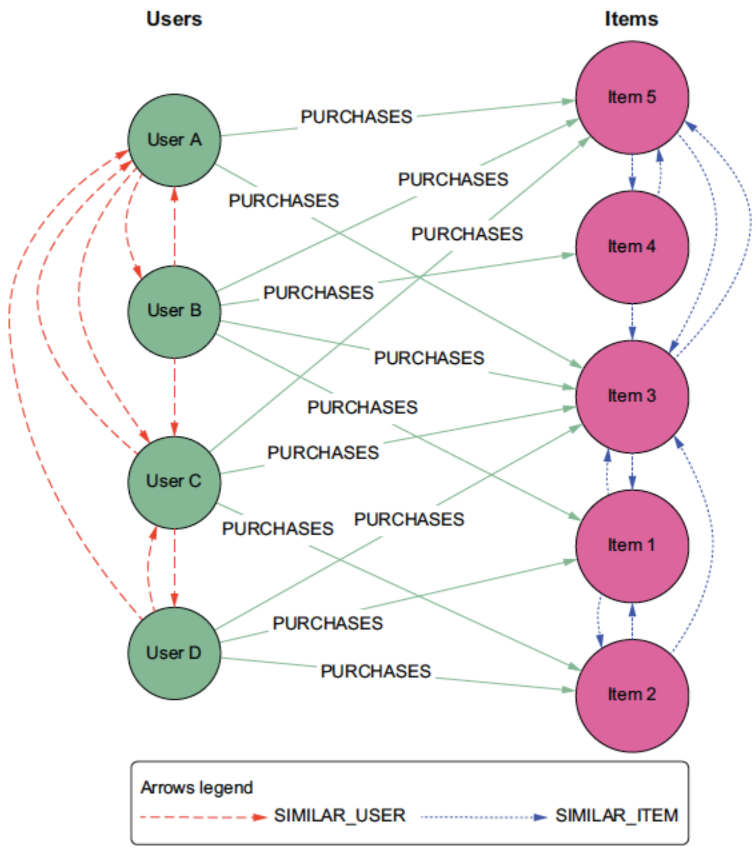

Figure 2.1: Overview of user interactions with items for CF method (Klacanova, 2022).

Figure 2.1 shows the architectural overview of user interactions with items. Both user-to-user filtering and user-to-product filtering use a similarity measure to determine which users or items are most similar. Examples of similarity measures used in CF are the Pearson correlation coefficient and the Cosine similarity measure. Once similar users or items have been identified, CF predicts a rating or preference for an item by combining the ratings or preferences of similar users or items.

CF is one of the most often used approaches when applying a recommendation system to financial settings (Zibriczky, 2016).

In their study, Sayyed et al. (2013a) conducted a preliminary investigation into the implementation of CF methods in the stock market. They use recommendation systems to suggest items to users using historical records of user purchases. These systems leverage data mining techniques to analyze historical user data, identify similarities among data items, and extract valuable hidden information or patterns. The system intends to establish relationships between new individuals and existing data items in order to determine similarities and provide recommendations. However, they did not validate these methods using stock data.

Gonzales & Hargreaves (2022) propose a CF technique to develop a stock recommender system that addresses the challenges of decision-making in the stock market. Gonzales & Hargreaves focus on understanding investors' needs and interests, and they explore three different approaches: K-Nearest Neighbour (kNN), Singular Value Decomposition (SVD), and Association Rule Mining (ARM). The three different approaches are popular CF methods used. By performing hierarchical clustering, the authors improve computational efficiency and identify similar groups of traders. They evaluate the recommended portfolios using financial metrics as well as statistical metrics. The results show that the kNN approach produces the most accurate recommendations. Overall, the study demonstrates promising recommendation results that cater to users' profiles, contribute to portfolio profitability, and minimize financial loss, thereby providing beneficial stock recommendations when considering user preferences.

In this study, we adopt the kNN approach as the CF method, similar to the one proposed by Gonzales & Hargreaves (2022).

2.3 Content-based Filtering

Content-based filtering (CBF) relies on the attributes of the items being recommended. This method suggests products comparable to what the user prefers based on the characteristics of the products and the user's previous actions or feedback. The approach creates a model using available features to explain interactions between users and items. The algorithm assesses the similarity of products and offers recommendations for items with similar attributes to previous products the user has engaged with. In conclusion CBF generates recommendations by finding similarities between the attributes of items and the user's past behavior.

CBF techniques overcome challenges associated with CF methods. They can recommend new items even when no user ratings are available, ensuring that recommendation accuracy is not affected by a lack of user preferences. These techniques can quickly adjust recommendations when user preferences change. Additionally, CBF can handle situations where different users share only identical items based on intrinsic features rather than the same items. This preserves user privacy, as recommendations can be provided without requiring the sharing of user profiles.

CBF also offers the advantage of providing explanations to users regarding how recommendations are generated. However, the technique does have its limitations. It relies on rich item metadata and well-organized user profiles for precise recommendations. Content overspecialization is another problem associated with CBF, as it restricts users from receiving recommendations similar to items already defined in their profiles (Isinkaye et al., 2015).

San Miguel González et al. (2015) propose a metadata-driven, data that describes or provides information about other data, CBF approach to peer-to-peer lending, which differs from traditional CBF. The authors present a framework that, by using vector-based and semantic models, can represent user data. While CBF algorithms are less susceptible to the cold start problem, they are usually less accurate than CF algorithms (Zibriczky, 2016).

Yoo et al. (2003) introduced the Stock Tracker, an adaptive recommendation system for trading stocks that focuses on personalized filtering based on user-deemed relevant trading information. Their approach utilizes the Moving Average Convergence Divergence (MACD) to provide buy and sell advice by analyzing the difference between two moving averages. The system incorporates an efficient algorithm that takes advantage of the fixed structure of user models and employs data-gathering techniques. The results indicate that the Stock Tracker can rapidly adapt its advice to cater to different types of users.

Similar to Yoo et al., Chalidabhongse & Kaensar (2006) proposed an adaptive user model framework for a personalized stock recommendation system. Their approach aims to provide investors with personalized and relevant information based on their individual profiles and historical interactions with the system. The system includes components such as user model initialization and updating, monitoring user interactions, and tailoring information to match user behavior and investment styles. The results demonstrate that their proposed system rapidly adapts to provide appropriate advice to users with diverse backgrounds.

2.4 The Cold Start Problem

The cold start problem is a common issue that arises in the recommendation systems field when there is not enough data available to provide accurate recommendations for new users or products. When a new user signs up for a platform or service, and there is no historical data regarding their preferences or behavior, the situation is known as a cold start for new users. Therefore, a cold start makes it difficult for the system to generate personalized recommendations for them.

The cold start problem is mainly a problem for CF algorithms due to the fact that they rely on the item's interactions to make recommendations. Without such interactions, a pure collaborative algorithm cannot suggest an item. On the other hand, CBF algorithms are less susceptible to the cold start problem in theory. As CBF rely on the features possessed by the items to make suggestions, even a new item with no interactions can still be recommended based on its features (Isinkaye et al., 2015).

2.5 Hybrid Filtering

Hybrid filtering (HF) combines the strengths of both collaborative and content-based filtering techniques. By merging these two approaches, HF can overcome several issues, e.g. the cold start problem. There are multiple ways for setting up hybrid recommendation systems, including Weighted HF and Switching hybrid filtering.

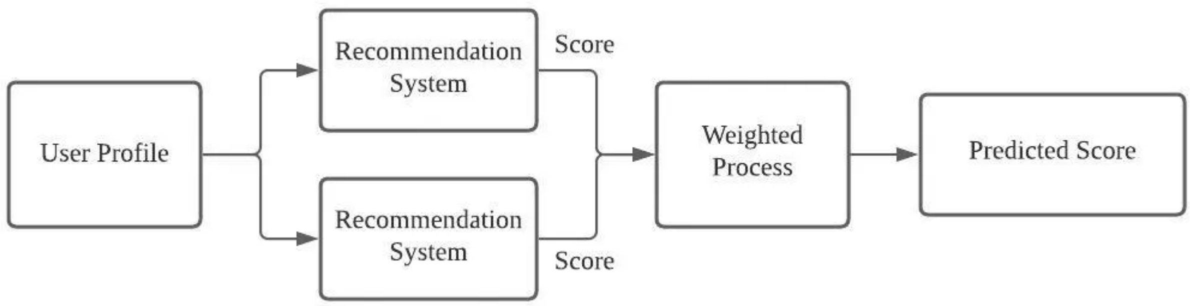

Figure 2.2: Overview of Weighted Hybrid Recommendation System (Chiang, 2021).

In Weighted HF, illustrated in figure 2.2, the system blends the outputs from different models using fixed weightings that remain constant across both the training and testing sets. In contrast, Switching HF selects a sole recommendation system based on the particular scenario. The criteria for the recommendation selector should depend on the user profile or other relevant characteristics. By adding an extra layer to the recommendation model, the Switching hybrid approach introduces a selector that determines the appropriate model to employ.

Tseng (2004) propose a novel hybrid recommendation system that integrates artificial intelligence techniques to effectively access, filter, evaluate, and incorporate information from diverse sources within the investment domain. Their system aims to provide human users with personalized decision recommendations on stock portfolio management, enabling them to optimize their investment strategies and maximize returns while considering their individual objectives and constraints.

2.6 Deep Learning Based Hybrid Recommendation System

This paper focuses on a Deep Learning Based Hybrid Recommendation System (DLHR) for our proposed recommendation system. Deep neural networks (DNNs) are a subfield of machine learning that use neural networks with three or more layers. Neural networks are inspired by the human brain and consist of layered representations of data. The input layer of the neural network model receives data, which is transformed as it passes through hidden layers. The hidden layers ultimately produce an output, also known as the last layer or model target.

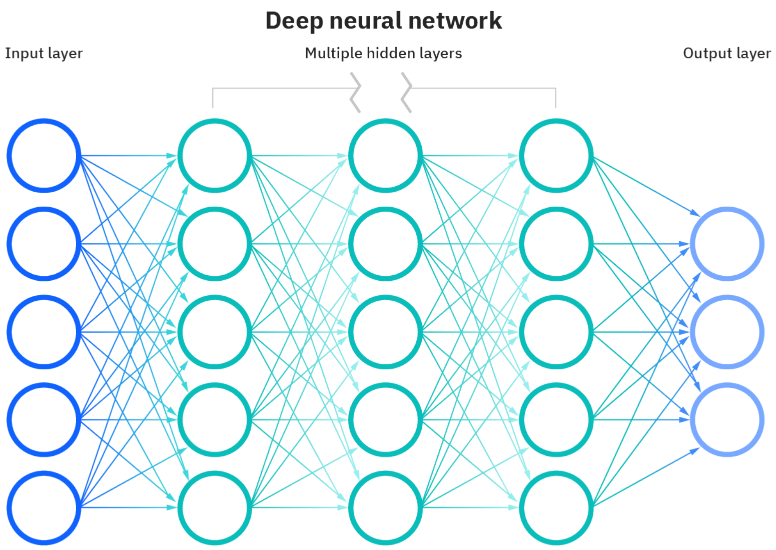

Figure 2.3: Overview of a fully connected multilayer feedforward neural network with multiple hidden layers (IBM, 2020).

Figure 2.3 illustrates the architectural layout of a multilayer feedforward neural network with three hidden layers. DNNs allow the neural network to "learn" from a large dataset, by iteratively adjusting their internal parameters presented below. While a neural network with only one layer can produce approximate predictions, incorporating hidden layers enhances accuracy by refining and optimizing the model. In a multilayered network, each neuron in a particular layer is connected to every neuron in the subsequent layer. There are no connections that skip layers or form backward loops, ensuring a strictly feedforward information flow. This sequential flow from one layer to the next facilitates efficient and effective learning.

The output $\hat{y}$ is determined by the N input values, h_i, the weights, w_i, the bias term b, and the activation function f according to the following equation.

$$\hat{y} = f\left(b + \sum_{i=1}^{N} w_i h_i\right) \tag{2.1}$$

The activation function, a critical component of neural networks, enables the model to approximate nearly any function. By introducing non-linearity into the neural network, the activation function allows for complex mappings between inputs and outputs. This non-linearity is essential for the model to learn and represent intricate relationships within the data. With the flexibility provided by the activation function, the neural network can capture and approximate a wide range of functions, making it a powerful tool for various tasks such as classification, regression, and pattern recognition.

To approximate the accuracy of the feedforward DNN, a loss function L is introduced. This function measures the error between the predicted output and the expected output of the network. As the loss function decreases, the network's robustness increases (Heaton, 2011). The most commonly used loss functions are Mean Absolute Error (MAE) and Mean Squared Error (MSE), defined as:

$$L_{MAE} = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{2.2}$$

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{2.3}$$

where $\hat{y}_i$ is the expected output and y_i the actual output, N denotes the number of cases. DNNs learn through a training process where they process various examples. Each example consists of a known input and the corresponding expected output. The model learns by determining the difference, known as the error, between the output of the model and the target output given. The model does this with the help of a backpropagation algorithm, which aims to optimize the weights by tracking the error term back to the neuron units. The model does this by computing the partial derivative of the loss function for the neuron weights and biases and adjusting them to minimize the loss function. The partial derivatives are given by:

$$[\frac{\partial L}{\partial w_{1,1}}, ..., \frac{\partial L}{\partial w_{N,N}}, \frac{\partial L}{\partial b_1}, ..., \frac{\partial L}{\partial b_N}] \tag{2.4}$$

A DLHR method combines the power of deep learning techniques with hybrid recommendation approaches to provide personalized recommendations to users.

Deep learning algorithms, specifically DNNs, are employed in the DLHR to learn patterns and representations from large volumes of data. By leveraging the hierarchical structure of neural networks with multiple hidden layers, DNNs can capture complex relationships and dependencies in the data, enabling more accurate and effective recommendation outcomes.

3 Litterature Review

In this section, we examine relevant findings from existing literature that form the basis of our study. By evaluating the application of deep learning recommendation systems, institutional investment decision-making, and cross-selling strategies, we aim to enhance the accuracy and effectiveness of our recommendation system for institutional investors.

3.1 Application of Deep Learning Recommendation Systems

Deep learning recommendation systems have gained attention as a promising alternative to conventional methods due to their inherent ability to process non-linear data effectively. One notable study by Kiran et al. (2020) addresses the limitations of collaborative filtering (CF) and content-based filtering (CBF) systems, such as the cold start problem and the linear nature of latent factors. The authors propose a deep-learning-based hybrid recommendation system (DLHR) incorporating embeddings to represent users and items. By employing a deep neural network (DNN), this approach learns non-linear factors by incorporating side information about users and items. The proposed technique outperforms the baseline methods in both cold start and non-cold start cases, as demonstrated through benchmarking on several data sets. Importantly, the proposed approach is generic and can be applied to other rating prediction data sets in recommendation systems. By incorporating side information about users and items and using a DNN, this approach can learn non-linear factors and improve recommendation performance. This aspect is crucial when working with financial data, which is often sparse in terms of explicit ratings.

In their study, Wang et al. (2017) address the issue of non-explicit selection criteria faced by editors when choosing news articles for end users. They develop a news article recommendation model that employs a dynamic attention-deep model to assist editors in selecting a subset of articles from a pool of dynamically changing news feeds. Unlike traditional approaches, this research employs deep learning techniques to understand the editor's dynamic style of article selection. By generating complex features using attention models, Wang et al. aims to represent the article style and classify whether the editor would like or reject the article.

Even though their approach is not implemented in a financial setting, the findings of this study are relevant to our research on using a deep learning-based recommendation system for financial investment recommendations. The use of deep learning attention models to capture complex features and understand user preferences can potentially enhance the accuracy and effectiveness of our recommendation system. By adapting and applying similar techniques to a pool of financial data, we may be able to create a system that learns the criteria of institutional investors and offers personalized investment recommendations based on their dynamic preferences and styles.

Jannach & Ludewig (2017) have shown that combining DNN with k-Nearest Neighbors (kNN) can improve recommendation accuracy in e-commerce applications. The kNN approach seeks to classify the k-most similar users or items by using similarity measures like Pearson correlation or cosine similarity. This approach enhances personalized recommendations by considering similarity measures between users or items. As personalized investment recommendations are the focus of our study, this work highlights the potential of combining deep learning techniques with collaborative filtering methods. This combination of deep learning and collaborative filtering methods holds immense promise. Institutional investors often face unique challenges and have distinct preferences and requirements compared to individual investors. By leveraging the fusion of DNN and kNN, we can effectively capture the intricate nuances of institutional investors' investment behaviors and preferences, resulting in more accurate and tailored recommendations.

In their research, Dong et al. (2017) present a deep learning-based hybrid recommendation system that addresses the challenge of limited information within the sparse user-item rating matrix. They propose a comprehensive deep learning framework that incorporates both item and user-side information to overcome this limitation. By incorporating both item and user-side information, this framework overcomes the limitation and improves recommendation performance. Since our study aims to develop a DLHR in a setting with sparse user-item ratings, this work provides insights into handling sparse data and incorporating additional information, which aligns with our research objectives.

In contrast to previous research, Taghavi et al. (2013) proposes a preliminary framework for a multi-agent recommendation system in computational investing. The framework utilizes a hybrid filtering technique trained with artificial neural network machine learning. It aims to provide adaptive recommendations on profitable stocks at the optimal time, considering individual investor preferences. Taghavi et al. study is valuable because it introduces a multi-agent recommendation system for computational investing, employing a hybrid filtering technique trained with artificial neural network machine learning. Moreover, their framework emphasizes adaptivity and individual investor preferences, which aligns with our goal of offering personalized investment recommendations. By integrating deep learning techniques into our hybrid recommendation method, we can dynamically adjust and update recommendations based on changing market conditions and investor preferences. This adaptability enhances the relevance and effectiveness of our recommendations, ensuring they stay aligned with the evolving needs and goals of institutional investors. While our research groups differ from theirs, we find inspiration from their work in applying our DLHR method in a financial setting.

3.2 Institutional Investment Decision-Making

Institutional investors are legal entities that manage and invest other people's money. They can take various legal forms, including profit-maximizing joint stock companies, limited liability partnerships, and statutory corporations. These investors may act independently or be part of a larger company group, such as mutual funds that are subsidiaries of banks and insurance companies.

The primary types of institutional investors include mutual funds, commercial banks, hedge funds, pension funds, and insurance companies. These institutional investors are known as intermediary investors since they handle and invest other people's money. Nonetheless, there are exceptions to this rule, as sovereign wealth funds may function as financial stabilization funds or de facto state ownership agencies. Some institutional investors, such as private equity funds, also have hybrid forms where the managing partner co-invests with limited partners.

Because of their experience and the fact that they are sometimes subject to less stringent regulations than retail investors, it is a common belief that they are better positioned to protect themselves. Overall, they are intermediary investors who manage and invest other people's money and are crucial to the financial market.

Institutional investors are known for their ability to manage large amounts of capital and diversify their investments across various asset classes. According to a 2021 report by McKinsey, the standard allocation for institutional investors is roughly 30.5% of assets to equity, 16% to real estate, 14% to infrastructure, 12.4% to private debt, and 9% to natural resources (Averstad & et al., 2023).

With the complexity and vastness of the market, institutional investors often rely on investment recommendations to inform their decision-making process. These recommendations offer valuable insights and analysis on various stocks and investment opportunities.

Chen & Cheng (2006) investigation focuses on whether institutional investors' exceptional performance can be attributed, at least partially, to their utilization of these stock recommendations. Their findings reveal compelling evidence supporting this notion. Firstly, Chen & Cheng find a positive correlation between the quarterly change in institutional ownership and consensus recommendations. This implies that institutional investors tend to increase their holdings in firms with favorable recommendations and decrease their holdings in those with unfavorable recommendations.

Moreover, after considering other factors that may influence institutional holdings, Chen & Cheng observe that firms with favorable recommendations experience a significant average increase of 0.90 percent in institutional ownership compared to those with unfavorable recommendations. This suggests that institutional investors place importance on stock recommendations when adjusting their portfolios.

To further validate their findings, Chen & Cheng employ large trades as a proxy for institutional trading activity. They note a higher frequency of buyer-initiated trades surrounding favorable recommendations, providing further evidence of institutional investors acting upon stock recommendations in their trading decisions.

Importantly, the change in institutional ownership driven by stock recommendations is found to be associated with future positive abnormal returns, indicating that the utilization of these recommendations contributes to institutional investors' overall outperformance in the market. On average, this association amounts to approximately 4.2 percent in annual returns, further highlighting the benefits gained from incorporating investment recommendations into their strategies.

Taken together, these findings suggest that institutional investors actively trade based on stock recommendations, and such trading activity plays a role in their superior performance. The utilization of investment recommendations serves as a valuable tool for institutional investors, enabling them to navigate the complexities of the market and make informed investment decisions that contribute to their overall success.

According to Basak & Pavlova (2013), the primary focus of institutional investors revolves around their performance relative to a specific index. The study presents a practical framework incorporating precise mathematical expressions, yielding the following analytical findings. The research indicates that institutional investors adjust their portfolios to favor stocks that are part of their benchmark index. This portfolio adjustment exerts upward pressure on the prices of the index stocks. As institutional investors demonstrate a stronger inclination towards riskier stocks than retail investors, they contribute to increased volatility in both the index stocks and the overall stock market. Consequently, this leads to countercyclical Sharpe ratios. Furthermore, the trading activities of institutional investors result in heightened correlations among stocks within their benchmark, generating an asset-class effect.

In portfolio optimization, institutional investors still often rely on the mean-variance model proposed by Markowitz (1952). Despite criticisms leveled against the model, it has found applicability in analyzing institutional investors' portfolios. The model's popularity and Markowitz's recognition with the Nobel Prize in Economics have contributed to its continued relevance.

However, it is crucial to note that the model has faced criticism regarding its underlying assumptions, which some argue are not fully supported by empirical evidence (Maillard et al., 2010). Critics, such as Merton (1980), highlight that the mean-variance analysis can be sensitive to input parameters, resulting in significant variations in portfolio composition even with minor changes. The sensitivity is particularly observed in expected returns. However, much of the criticism has been directed toward retail investors concerning the risk-return tradeoff. Yu & Yuan (2011) provide insights into the influence of retail investors on the risk-return tradeoff, offering two explanations. Firstly, as noise traders, retail investors often inaccurately estimate the variance of returns, thereby distorting the relationship between mean and variance. Secondly, due to restrictions on short selling, retail investors are more likely to trade when they feel optimistic rather than pessimistic.

In contrast, institutional investors are regarded as possessing greater sophistication, which makes them less susceptible to behavioral biases compared to retail investors (Bushee, 1998). They are more likely to accurately estimate the variance of returns and maintain a more balanced relationship between mean and variance in their portfolio decisions. Therefore, the mean-variance model remains applicable when analyzing institutional investors' portfolios due to their higher level of expertise and reduced susceptibility to behavioral biases.

Measuring risk-reward is of utmost importance for institutional investors as it allows them to assess the performance of their investment portfolios in a comprehensive manner. The Sharpe ratio, proposed by Sharpe (1966) and widely adopted as the de facto "gold standard" for evaluating performance in terms of risk-adjusted returns, provides a framework for comparing different investment strategies based on their ability to generate returns relative to the level of risk taken.

By integrating the measure into Markowitz's mean-variance framework, Sharpe established it as one of the most renowned metrics for performance evaluation. This ratio takes into account the trade-off between the expected return and the volatility of an investment, enabling investors to assess whether the returns generated justify the level of risk undertaken.

However, de Prado et al. (2005) shed light on the limitations of the Sharpe ratio. Their research revealed that the underlying assumptions of independence and identical distribution of returns, which are fundamental to the ratio, may not hold in certain investment strategies, such as hedge funds. These assumptions can mask significant drawdown risks, potentially exposing investors to unforeseen losses.

To address the drawdown risk limitation, the Sortino ratio was introduced as a complementary measure for evaluating and comparing the performance of fund managers with skewed return distributions Rollinger & Hoffman (2013). Unlike the traditional Sharpe ratio, which considers all deviations from the mean as risky, the Sortino ratio incorporates downside deviation. It focuses solely on returns that fall below a user-defined target or required rate of return, providing a more accurate assessment of the risk associated with an investment strategy.

In conclusion, institutional investors play a significant role in the financial market with their expertise, capital management capabilities, and utilization of investment recommendations. They contribute not only to their success but also to the overall functioning of the market.

3.3 Cross-Selling

Cross-selling is a crucial marketing strategy for businesses, particularly in the financial industry. It involves offering existing customers additional financial products or services that complement or expand their current relationship with the company. Financial institutions are increasingly utilizing cross-selling as a vital technique in customer relationship management (CRM) to deepen their existing customer relationships, increase revenue, and foster customer loyalty. The approach involves the selling of additional financial products to customers. The primary objective of cross-selling in the context of CRM is to enhance the firm's share of the customer wallet, broaden the scope of the customer relationship, and improve customer retention (Kamakura, 2008). Kamakura (2008) provides valuable insights into collaborative filtering as an analytical tool for cross-selling within a CRM setting. The study emphasizes the importance of a comprehensive and complete customer database for implementing cross-selling strategies utilizing collaborative filtering. Such a database should contain detailed information on customer activities and provide a holistic view of each customer across all contact points. Additionally, the study highlights the need for a contact-management application to manage and register customer contacts. Financial institutions can successfully implement cross-selling strategies that deepen customer relationships and increase revenue by fulfilling these key factors.

Financial advisors can also benefit from cross-selling, as it can be a way to earn additional revenue from their existing client base. However, it is crucial to approach cross-selling in a way that adheres to regulations and protects the client's best interests. The act of referring clients to additional products or services solely to receive incentives can result in negative repercussions such as customer complaints and disciplinary action.

Boustani et al. (2023) conducted a study to improve cross-selling models in retail banking by analyzing several years of account transaction data, specifically credit card transactions, using a deep-learning algorithm. They developed several models to extract features from the transaction data, which they used as input to a classifier, which showed that they had predictive power and significantly increased the accuracy of the cross-selling model. The study demonstrated that utilizing transaction data can enhance the performance of cross-selling models in retail banking.

Therefore, we will create a comprehensive and complete database to determine if the proposed method can improve cross-selling. To conduct this study, we will use publicly listed holdings data for the selected target group of institutional investors.

4 Data

4.1 Institutional Investor Data

We focus on institutional investment managers with public portfolio holdings on the Stockholm Stock Exchange. To obtain comprehensive and reliable data, we gather portfolio holdings information from the Modular Finance Holdings database. This database tracks ownership in the Swedish financial markets, providing valuable insights into the investment activities of institutional investors.

Our dataset includes portfolio holdings information from 2949 institutional investors. The data set encompasses a wide range of details that offer a comprehensive view of institutional investors and their portfolios. Firstly, we collect information about the name of each owner, allowing us to identify and analyze individual investment managers. This enables us to assess the investment strategies and preferences of different institutional investors operating on the Stockholm Stock Exchange.

Additionally, we capture portfolio value data, which gives us an indication of the financial resources managed by each institutional investor. By examining the size of their portfolios, we can identify the major players in the market and assess their potential impact on stock prices and market trends.

The number of positions held by each investor is another crucial aspect of our dataset. This information allows us to understand the diversification strategies employed by these investors. Analyzing the number of positions and the weights assigned to each position provides insights into the level of concentration or dispersion in their portfolios, highlighting their risk management approaches.

Moreover, our dataset includes country and sector overviews for each institutional investor's portfolio holdings. This enables us to understand their geographic and industry preferences. By analyzing these factors, we can identify regional or sectorial biases in their investment decisions and assess their impact on specific stocks or sectors within the Stockholm Stock Exchange.

To further investigate the investment activities, we track buying and selling activities of the institutional investors in our dataset. This data offers insights into their trading patterns, such as the frequency of transactions, the timing of trades, and the volume of shares bought or sold. These activities help us identify trends and patterns indicating investment strategies or market sentiments.

Furthermore, the performance history of institutional investors' portfolios is essential. This historical data allows us to evaluate the investment performance of each manager over time. By assessing their returns, risk-adjusted measures, and other performance metrics, we can compare and rank institutional investors based on their track records.

Lastly, our dataset includes the date of verification, ensuring that we have the most up-to-date information on institutional investors' portfolio holdings. This allows us to accurately analyze their current positions and help make assessments of their investment preferences.

4.2 Asset Overview

To investigate how recommendation systems can suggest complementary assets to institutional portfolios, this study incorporates stocks from the OMXSPI index. The OMXSPI, also known as Stockholm All-share, provides a comprehensive overview of the Stockholm Stock Exchange by encompassing all listed companies. The index follows a market-value-weighted methodology, where the weights assigned to the constituent stocks are based on their total market value. It includes all companies listed on the exchange, offering a holistic representation of market development. We source historical stock price data for this study from FinBas, a financial database provided by the Swedish House of Finance. FinBas provides daily end-of-day stock price information from the Nordic Stock Exchanges.

For this study, we select stocks from the 2021 index composition as the end point of the analyzed period. This choice is made to evaluate and compare investment recommendations for current institutional portfolios rather than constructing portfolios that mimic the index. By selecting constituents from the index at the end of the period, the study avoids potential survivorship bias (Sharpe, 1966).

Preprocessing of the data was necessary due to a few companies being listed on the exchange after 2011 and some companies being delisted, resulting in missing data. To ensure a high-quality model, stocks with missing data points were removed. Consequently, the final dataset consists of 460 stocks and their quarterly observations. The observation period ranges from 3 January 2011 to 31 December 2016.

4.3 Descriptive Statistics

To account for the quarterly reporting of institutional holdings, we utilize quarterly returns for our analyses. Adopting a larger investment interval would yield insufficient sample size observations, rendering the empirical analysis less meaningful.

To ensure the robustness of our findings, we evaluate the results of our main analyses using different investment intervals, following the approach outlined by Prakash et al. (2003). In this process, we scale the returns to a simple annual rate using the formula:

$$R_t = \frac{\text{Price}_t - \text{Price}_{t-1}}{\text{Price}_{t-1}} \times \frac{365}{\text{Days in Interval}} \tag{4.1}$$

This scaling methodology allows us to express the returns on an annual basis, facilitating easier comparison and analysis across different investment intervals.

To enhance the recommendations later on, we categorize the selected stocks into three clusters, high risk, moderate risk and low risk assets. In Table 4.1 we can see the descriptive statistics for the three clusters for the whole period, 2011-2021. When comparing average returns, the High-Risk cluster stands out with the highest mean return (0.247), followed by the Moderate-Risk cluster (0.166), and the Low-Risk cluster (0.110). This suggests that the cluster of higher-risk assets has the potential for greater returns.

In terms of volatility, the High-Risk cluster exhibits the highest standard deviation (0.191), indicating larger price fluctuations compared to the Moderate-Risk cluster (0.131) and the Low-Risk cluster (0.099). This confirms that the High-Risk cluster carries a higher level of inherent risk. Assessing risk-adjusted returns using the Sharpe ratio, the High-Risk cluster again shows a notable result, with the highest ratio (1.293). The Moderate-Risk cluster follows closely with a Sharpe ratio of 1.267, while the Low-Risk cluster has a slightly lower ratio of 1.09. A higher Sharpe ratio suggests better risk-adjusted performance.

Considering the maximum drawdown, which represents the largest quarterly decline, the High-Risk cluster experienced the largest drop (0.098), indicating higher susceptibility to significant losses. The Moderate-Risk cluster had a maximum drawdown of 0.073, while the Low-Risk cluster experienced a smaller decline (0.064).

In terms of market sensitivity, all three clusters have betas less than 1. This implies that their returns are relatively less sensitive to overall market movements, providing a certain level of stability.

Table 4.1: The table presents descriptive statistics for quarterly asset returns, of each asset cluster, scaled by a factor of four throughout the whole period. SR represents the Sharpe Ratio.

	Low-Risk	Moderate-Risk	High-Risk
N. of obs.	44	44	44
Mean	0.110	0.166	0.247
StdDev	0.099	0.131	0.191
SR	1.09	1.267	1.293
Max Drawdown	0.064	0.073	0.098
Beta	0.527	0.536	0.479

Table 4.2 displays the distribution of institutional investors from various countries investing in the Swedish stock market. The data shows that Swedish institutional investors hold a significant portion of the market, with 161 institutions representing 58.94% of the total number of institutions and accounting for 52.05% of the total market value. This indicates a significant presence and impact of local investors in the Swedish stock market. The data also demonstrates the international nature of the Swedish stock market, with institutions from various countries participating. The countries with notable representation include the USA (11.8% of institutions and 25.87% of total market value), the UK (5.46% of institutions and 4.16% of total market value), Switzerland (2.31% of institutions and 1.64% of total market value), and Norway (2.64% of institutions and 3.70% of total market value).

Table 4.2: The table presents an overview of the institutional investors in the dataset grouped by home-country

Country	Number of institutions	% of total	Total market value (MSEK)	% of total
Australia	17	0.58	29685	0.27
Austria	11	0.20	4830	0.04
Belgium	15	0.37	33287	0.31
Canada	71	1.73	104739	0.96
China	80	0.20	132301	1.21
Cyprus	66	0.51	17696	0.16
Denmark	11	2.41	131669	1.21
Finland	8	2.71	218187	2.00
France	19	2.24	130710	1.20
Germany	11	2.44	285590	2.62
Hongkong	51	0.37	8623	0.08
Iitaly	6	0.27	3096	0.03
Ireland	41	0.64	20404	0.19
Japan	34	0.37	66222	0.61
Luxemmbourg	78	1.39	70638	0.65
New Zeeland	6	0.2	5538	0.05
Norway	68	2.64	403148	3.70
Singapore	8	0.27	20236	0.19
Spain	23	0.78	13739	0.13
Sweden	161	58.94	5671032	52.05
Switzerland	1738	2.31	178545	1.64
The Netherlands	72	1.15	72428	0.66
UK	348	5.46	453555	4.16
USA	6	11.8	2818480	25.87
Total	2949	100.00	10894379	100.00

To simplify the analysis later on, we categorize the selected investors into four groups according to their portfolio value: Ultra, Large, Medium, and Small investors. The 45 largest institutional investors in our dataset can be found in Appendix A3. Table 4.3 showcases descriptive statistics for each of these investor groups. Notably, the Ultra investors possess a significantly larger total asset under management (AUM) of the combined total of the other three groups, roughly 73% of the total AUM.

When examining the number of distinct assets held in each portfolio, we observe a clear pattern: the higher the portfolio value, the greater the number of distinct assets owned by the investors.

Table 4.3: The table presents an overview of the institutional investors in the dataset grouped by AUM.

	Number of institutions	% of total AUM	Number of Positions
Ultra	737	93.6	33382
Large	739	4.3	8380
Medium	736	1.6	8334
Small	737	0.5	7305

5 Methodology

5.1 Assumptions

In constructing the investment recommendations, certain assumptions were taken into account.

To begin with, it is important to note that no short-sale constraints were applied in this study. This choice was made based on the understanding that fund managers generally have the freedom to participate in short-selling activities, making this assumption reasonable.

Secondly, due to the utilization of quarterly filings of institutional holdings, a buy-and-hold strategy for at least a quarter is assumed. However, it should be noted that this assumption poses a limitation, as the institutional investors included in this study may adopt different strategies or have shorter holding periods.

Finally, an additional assumption made in this study is that the investors included have the ability to borrow money at the risk-free rate. This assumption is based on the premise that investors can access borrowed funds at a rate that is considered risk-free in order to support their investment activities.

5.2 Constructing the Recommendation System

1. *Data pre-processing and feature engineering*: We prepare the data by cleaning, transforming, and selecting relevant features for analysis.

2. *Implicit ratings to explicit ratings*: We estimate explicit ratings from implicit ratings using appropriate assumptions or criteria.

3. *Clustering and classification*: we use K-means clustering to group similar items based on their characteristics and employ K-nearest neighbors (KNN) for classifying users with similar preferences or item characteristics.

4. *Preference generation*: we integrate the different techniques used, using hybrid filtering and deep learning, to generate accurate user preferences and recommendations.

5. Data split: Divide the data into a training set (60%), development set (20%), and test set (20%) for model training, validation, and evaluation.

6. *Evaluation*: We evaluate the DLHR model using appropriate evaluation metrics and compare its performance to baseline models.

7. *Real-world testing*: we randomly select one investor from each investor group to test and evaluate the model's performance in a real-world scenario.

5.2.1 Data Pre-Processing and Feature Engineering

Data pre-processing prepares data for machine learning by transforming it into a functionable format. The process involves handling missing or irrelevant data, cleaning inconsistencies, and standardizing data. Feature selection and hyperparameter tuning are also part of the pre-processing. The goal is to create a data set the model can work with effectively.

One accepted method of representing categorical variables like users and items is to use embeddings, which represent each categorical value as a vector in N-dimensional space. The categorical variables in this study are presented in table 4.1.

We create two embedding matrices, one for the users, E_{user}, and one for the items, E_{item}. The hyper-parameters K_U and K_I represent the number of embeddings or factors for users and items in our model. We select and tune these hyper-parameters, listed in table 4.1. These embeddings are similar to the latent factors used in collaborative filtering, but with the difference that we treat their values as parameters, like weight matrices at different hidden layers learned by the deep learning network. This means that they can capture both linear and non-linear factors of users and items. To update the weights of these embeddings, we include the matrices of embeddings for users and items as model parameters, randomly initialize them, and modify them using stochastic gradient descent.

Table 5.1: The table shows the different categorical variables deployed in the model and further specifications.

Category	**Category Specifications**	**Variables**	**Variable Specifications**
Target Variable	We focus on Implicit ratings, a type of feedback that users provide through their behavior, rather than explicitly providing a rating or feedback.	• Annual Returns • Annual Volatility • Annual Max Drawdown • Sharpe ratio	These variables measure the risk and reward of an investment.
User Features	A set of attributes or characteristics of a user that can be used to describe their properties.	• AUM • Risk profile • Investment objectives • Investment horizons • Sector allocation	The variables offer crucial insights into the behavior and preferences of investors, enabling the system to evaluate users and group them together for targeted investment opportunities or customized recommendations.
Product Feautures	A set of attributes or characteristics of an item that can be used to describe its properties.	• Adjusted closing price • Volatility • Daily change • Beta • Sector	The variables offer crucial insights into the characteristics of the products, enabling the system to evaluate products and group them together for targeted investment opportunities or customized recommendations.

We grouped the investors into small, medium, large and ultra investors based on their portfolio value and provided relevant recommendations for each group.

One of the challenges we encounter with our dataset is the absence of explicit ratings from users indicating their preferences for the stocks they invest in. These ratings are essential for building our recommendation system. Therefore, as mentioned above, we need to construct implicit ratings. To address this challenge, we follow the example of Gonzales & Hargreaves (2022) where they make the assumption that the more of the same stock an investor holds, the higher their preference for those stocks. Based on this assumption, we estimate the explicit rating of stock j for each user i by utilizing:

$$r_{ij} = 1 + 4\frac{\sum_k Q_{ij}^{(k)}}{\sum_{j,k} Q_{ij}^{(k)}} \tag{5.1}$$

The numerator represents the total volume of stock j held by user i in their kth holding. The denominator represents the total quantity of all the stocks held by user i across all holdings. We only include ratings greater than 1 in the training set or dev set respectively, so now each rating, r_{ij},in either training set or development set is a real number in the range $(1, 5]$.

5.2.2 Feauture Learning

To identify stocks with similar attributes for our recommendations, we employ a clustering algorithm. This algorithm partitions a dataset into sub-clusters, facilitating the identification of meaningful groups within the data (Isinkaye et al., 2015). An effective clustering method ensures the creation of high-quality clusters with high intra-cluster similarity and low inter-cluster similarity. In our approach, we use the clustering algorithm to categorize assets into three risk groups: High risk, Medium risk, and Low risk. This clustering process enables us to provide tailored recommendations based on the risk preferences of individual users. We cluster the items by utilizing K-means Clustering. K-means Clustering is an unsupervised machine-learning technique utilized to group similar data points based on their features. Its operation involves assigning data points to cluster centers iteratively and updating the centers to achieve convergence. A cluster represents a collection of data points grouped due to their similarities. The parameter "k" determines the desired number of centroids in the dataset. A centroid is a central point, either real or imaginary, representing the center of a cluster.

The K-means algorithm allocates each data point to one of the clusters, aiming to minimize the sum of squares within each cluster. In simpler terms, it identifies "k" centroids and assigns each data point to the nearest cluster, while ensuring that the centroids remain as compact as possible.

The term "means" in K-means refers to the process of finding the centroid, which involves averaging the data points.

We use the product features presented in table 5.1 to perform the clustering. The three clusters deployed are high risk, moderate risk and low risk assets.

5.2.3 Baseline Models

For the purpose of our comparison evaluation, we employ two baseline collaborative filtering (CF) models: user-based k-Nearest Neighbors (kNN) and item-based kNN. The kNN algorithm is employed to identify similar user preferences or items among users. It is a non-parametric and lazy learning algorithm that only trains the dataset when a recommendation is required. The kNN approach seeks to classify the k-most similar users or items by utilizing similarity measures like Pearson correlation or cosine similarity. In this study, we utilize both similarity measures for both an item-based kNN approach and user-based kNN approach, we also leverage the baseline as the CF component of our proposed system. The cosine similarity for an item-based kNN approach is calculated using the following equation:

$$\mathrm{sim}(i,j) = \frac{\sum_{u=1}^{N} r_{ui} \cdot r_{uj}}{\sqrt{\sum_{u=1}^{N} r_{ui}^2} \cdot \sqrt{\sum_{u=1}^{N} r_{uj}^2}} \tag{5.2}$$

where i and j are items, in the case of a user-based kNN approach, i and j would be two different users. r_{xy} is the rating provided by user x on item y and N is the total number of users in the dataset. After calculating the cosine similarity and the Pearson correlation we need to choose a k value, the number of nearest neighbors that are considered when making a prediction or classification. It is an important parameter that determines the influence of neighboring data points on the prediction. The choice of the optimal k value in kNN is crucial. A large k can be computationally expensive and lead to underfitting, while a small k can cause overfitting. To find the best k, we plot error rates for different k values and select the one with the lowest error rate. It is recommended to choose an odd k to avoid tie situations (Zhang, 2016). After choosing k we present the formula for the user-based algorithm to calculate the predicted rating for target user i on item j at the k nearest neighbor, using their weighted average, defined as:

$$r_{ij} = \bar{r_i} + \sum_{k} \mathrm{sim}(j,k) \frac{(r_{ik} - r_k)}{\sum_k |\mathrm{sim}(i,k)|}$$

where $\bar{r_i}$ is the average rating of user i.

5.2.4 Deep Learning Based Hybrid Recommendation System

Once we have acquired the user and item feature embeddings, we combine them with supplementary side information, including user and item features from Table 5.1, as well as the explicit ratings generated by the kNN approach. By incorporating these embeddings and content-based features, our solution functions as a hybrid recommendation system. We employ the preference generation component to forecast user preferences. This integration of models enables us to examine and comprehend data characteristics from multiple angles. Consequently, this approach overcomes the shortcomings of individual models, capitalizes on the advantages of each model, and ultimately produces more precise outcomes.

To further explore the deep nonlinear relationship between users and items, we employ deep neural networks to train the hybrid recommendation system. Deep neural networks allow us to uncover intricate patterns and capture complex interactions, providing a better understanding and prediction of user preferences.

The deep learning model comprises several neuron units connected by weights. The mathematical definition of a single neuron, also referred to as a perceptron (Rosenblatt, 1958), can be expressed as:

$$\hat{y} = f\left(b + \sum_{i=1}^{N} w_i h_i\right) \tag{5.3}$$

The output, $\hat{y}$, is determined by the N input values, h_i, the weights, w_i, the bias term, b, and the activation function, f. The activation function is vital in enabling the model to approximate a wide range of functions. Specifically, the weighted sum of the outputs from the previous layers determines the output of each neuron in the network, which, in turn, is passed through the activation function. A loss function is introduced to evaluate the accuracy of the DNN. This function measures the discrepancy between the predicted output and the expected output of the network.

Our neural network architecture consists of three layers, each performing specific computations as described in Formula 5.3. To learn and update the weights of the deep neural network (DNN), we employ the ADAM optimization algorithm proposed by Kingma & Ba (2014). The algorithm is based on a gradient descent, and efficiently updates the weights of the neural network, enabling effective learning and optimization of the model.

5.3 Evaluation Metrics for Recommendation Algorithms

Evaluating the model is crucial as the recommendations' quality can heavily impact users' experience and business value. Therefore, using established evaluation metrics that align with our objectives is critical. For this study, we evaluate the performance of our algorithm using the following evaluation metrics as sugested by Isinkaye et al. (2015). One of the initial metrics used is the mean absolute error (MAE), a statistical accuracy metric calculated as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{5.4}$$

where N is the total number of users in the dataset, y_i is the true value and $\hat{y}_k$ is the expected value. It serves as a measure of the deviation between the recommendation and the specific value from the user.

We also consider another statistical accuracy metric, the Mean Squared Error (MSE), defined as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{5.5}$$

Using both MAE and MSE in evaluating regression models is beneficial because they provide different perspectives on error magnitude, sensitivity to outliers, model comprehensibility, and decision-making trade-offs. MAE captures the average absolute difference between predicted and actual values, while MSE focuses on the average squared difference. MAE is less sensitive to outliers and is more interpretable, representing errors in the original units. MSE amplifies larger errors and may be more suitable for extreme cases.

Explicit ratings, such as the Mean Absolute Error and Root Mean Square Error, are not suitable, on their own, for evaluating recommendation systems that rely on implicit ratings. Consequently, we cannot directly evaluate the performance of implicit and explicit recommendation systems using these metrics. We use the mean average precision at K (mAP@K) to obtain a more comparable evaluation metric. This metric assesses the relevance of the recommended 'k' items while considering the item order in the list and disregards the exact rating provided by the user. This approach enables us to evaluate models with implicit or explicit data as input (Gonzales & Hargreaves, 2022).

To calculate mAP@k, the AP@k is first computed for each user in the dataset at a threshold of k items using:

$$AP@k = \frac{1}{r}\sum_{i=1}^{k} Precision(i) \cdot relevant(i) \tag{5.6}$$

$Precision(i)$, which quantifies the fraction of recommended items in the top i set that are relevant, and $relevant(i)$, an indicator function where 1 indicates that item i is relevant and 0 otherwise. The value of r represents the total number of items that are relevant to the user.

Then mAP@k is simply the average of AP@k among all users in the dataset and is given by:

$$mAP@k = \frac{1}{N}\sum_{i=1}^{N} AP@k_i \tag{5.7}$$

5.4 Mean-Variance Analysis

We use the metrics MAE, RMSE, and mAP@k to assess the quality and robustness of recommendations. However, institutional investors are more interested in understanding the potential gains and losses associated with said recommendations. Evaluating the profitability and financial risks of the investment recommendations is crucial. We aim to determine whether the assets suggested by the recommendation systems can expand investment opportunities.

We will use the mean-variance spanning test, introduced by Huberman & Kandel (1987), to perform this evaluation. This approach statistically tests the effects of adding N new assets to an existing collection of K benchmark assets. If the new frontier of the combined portfolio of N and K assets overlaps with the frontier of the portfolio of K assets, the outcome is known as "spanning." Spanning means that investors gain no benefit from adding the N assets to their portfolio. However, if the frontier of the combined portfolio is larger than the frontier of the portfolio of K assets, investors gain diversification benefits from adding the N assets.

Further, to assess the economic magnitude of the diversification benefits, we will look at the change in Sharpe Ratio of the tangency portfolio and the change in standard deviation of the global minimum-variance portfolio as suggested by Bekaert and Urias (1996).

We define $R_t = \begin{bmatrix} R_{1t} \\ R_{2t} \end{bmatrix}$ as the raw returns on $N + K$ risky assets at time t, where R_{1t} is a K-vector of the returns on the K benchmark assets and R_{2t} is an N-vector of the returns on the N test assets.

We define the expected returns on the $N + K$ assets as

$$\mu = \mathbb{E}[R_t] = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \tag{5.8}$$

and the covariance matrix of the $N + K$ risky assets as

$$V = \mathbb{V}\text{ar}[R_t] \equiv \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \tag{5.9}$$

We then estimate the following model using Ordinary Least Squares (OLS) regression.

$$R_{2t} = \alpha + \beta R_{1t} + \varepsilon_t \tag{5.10}$$

Where α represents the intercept term and ϵ represents the error term in the equation. As per Huberman & Kandel (1987), the null hypothesis of spanning is then:

$$H_0 : \alpha = 0_N, \quad \delta = 1_N - \beta 1_K = 0_N \tag{5.11}$$

Where 1_N and 1_K represent a column vector of size N and K with all elements equal to 1. When condition (5.11) holds, it implies that for every recommended asset (or portfolio of recommended assets), there exists a portfolio of the K benchmark assets with the same mean, but a lower variance compared to the recommended asset. As a result, the N recommended assets are dominated by the K benchmark assets.

We then calculate the Wald, Likelihood Ratio and Lagrange Multiplier tests as described by Kan & Zhou (2008)

$$W = T(\lambda_1 + \lambda_2) \sim \chi_2^2 \tag{5.12}$$

$$LR = T\left(\ln(1 + \lambda_1) + \ln(1 + \lambda_2)\right) \sim \chi_2^2 \tag{5.13}$$

$$LM = T\left(\frac{\lambda_1}{1 + \lambda_1} + \frac{\lambda_2}{1 + \lambda_2}\right) \sim \chi_2^2 \tag{5.14}$$

They follow an asymptotic chi-squared distribution with two degrees of freedom, and where λ_1 and λ_2 are the eigenvalues of the $\hat{H}\hat{G}^-1$ matrix where

$$\hat{H} = \begin{bmatrix} \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} & \hat{\alpha}'\hat{\Sigma}^{-}1\hat{\delta} \\ \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\delta} & \hat{\delta}'\hat{\Sigma}^{-1}\hat{\delta} \end{bmatrix} \tag{5.15}$$

$$\hat{G} = \begin{bmatrix} 1 + \hat{\mu}_1'\hat{V}_{11}^{-1}\hat{\mu}_1 & \hat{\mu}_1'\hat{V}_{11}^{-1}1_K \\ \hat{\mu}_1'\hat{V}_{11}^{-1}1_K & 1_K'\hat{V}_{11}^{-1}1_K \end{bmatrix} \tag{5.16}$$

Where $\hat{\Sigma}$ represents the residual variance.

To gain a better understanding of the importance of the null hypothesis and to aid in our later discussion regarding the allocation of weights to the recommended assets, we analyze two portfolios located on the minimum-variance frontier of the N + K assets, the global minimum variance (GMV) portfolio and the tangency (T) portfolio. The weights assigned to these portfolios are presented below, as outlined in Kan & Zhou (2008).

$$W_T = \frac{V^{-1}\mu}{1_{N+K}'V^{-1}\mu} \tag{5.17}$$

$$W_{GMV} = \frac{V^{-1}\mu}{1_{N+K}'V^{-1}1_{N+K}} \tag{5.18}$$

6 Empirical Analysis

In this section, we present the outcomes of our empirical methodology, which was introduced in Section 5. Furthermore, we critically analyze and discuss these results.

6.1 Performance Evaluation of the System

We start by presenting the performance of the DLHR and comparing it to four baseline methods. The evaluation metrics provide insights into the performance of the recommendation system models and are presented in Table 6.1.

For the Mean Absolute Error (MAE), DLHR achieves the lowest value of 0.68, indicating the smallest average difference between predicted and actual ratings. The low MAE suggests that DLHR performs better in accurately predicting user preferences than the other models. On the other hand, User KNN – Pearson has the highest MAE value of 0.85, implying that it has a higher average prediction error and may face challenges in accurately predicting user preferences.

Moving on to the Mean Squared Error (MSE), DLHR again performs the best with the lowest value of 0.95 and has the lowest average squared difference between predicted and actual ratings, showcasing better overall prediction accuracy. Conversely, User KNN – Pearson has the highest MSE value of 1.25, suggesting that it may struggle in accurately predicting user preferences and could result in higher prediction errors.

Considering the Mean Average Precision at 10 (mAP@10), DLHR achieves the highest value of 0.73, indicating better precision in the top 10 recommended items. Users are more likely to find relevant recommendations among the top 10 suggestions from DLHR. On the other hand, User KNN – Pearson has the lowest mAP@10 value of 0.61, implying lower precision in the top 10 recommendations and potentially fewer relevant suggestions for users.

Lastly, the adjusted R^2 (Coefficient of Determination), a modified version of R-squared that has been adjusted for the number of predictors in the model, metric reflects how well the models fit the data. DLHR achieves the highest adjusted R^2 value of 0.82, suggesting that it is the best-fitting method for the data compared to the other models. This indicates that DLHR captures a higher portion of the variance in the actual ratings, resulting in more accurate predictions. User KNN – Pearson has the lowest adjusted R^2 value of 0.68, implying a weaker fit to the data and potentially more difficulty in capturing the underlying patterns in the user-item ratings.

Table 6.1: The table presents MAE, MSE, mAP@10 and R-squared for the different systems deployed.

Method	MAE	MSE	mAP@10 (Higher is better)	R^2_{Adj}
User KNN – Cosine	0.82	1.15	0.63	0.72
User KNN – Pearson	0.85	1.25	0.61	0.68
Item KNN – Cosine	0.78	1.10	0.67	0.75
Item KNN – Pearson	0.80	1.12	0.66	0.74
DLHR	**0.68**	**0.95**	**0.73**	**0.82**

6.2 Mean-Variance Analysis

Table 6.2 reports the results of the mean-variance spanning test for the three different periods. The asymptotic tests reject the null hypothesis of spanning at a 1% significance level for the Likelihood Ratio and the Lagrange Multiplier test and a 5% significance level for the Wald test in the training period. For the development period, the Wald test and the Lagrange Multiplier are statistically significant at a 1% level, while the Likelihood Ratio test is significant at a 5% level. In the test period all three tests reject the null hypothesis at a 1% significance level. Hence, we can conclude that the recommended assets proposed by the system expand the investors' mean-variance efficient frontier and, therefore, their investment opportunity set.

Table 6.2: The table shows the results for the mean-variance spanning test of N-recommended assets with K-benchmark assets for all three periods.

	Training Period		Development Period		Test Period	
	Value	p-Value	Value	p-Value	Value	p-Value
Wald	8.409	0.015**	10.494	0.005***	22.987	0.000***
Likelihood Ratio	10.013	0.007***	8.002	0.018**	24.708	0.000***
Lagrange Multiplier	23.674	0.000***	28.437	0.000***	29.120	0.000***

***, **, * denote statistical significance at the 1%, 5%, 10% level.

Table 6.3 represents the tangency (T) and global minimum-variance (GMV) portfolios for the mean-variance spanning test involving N-assets and K-benchmark-assets for the Ultra investor across the three asset clusters.

We can see that the T portfolios have an allocated weight of 40.5% to 43.9% across all asset clusters to the N-recommended assets, while the GMV portfolios have an allocated weight of 8.7% to 37.8%. This suggests that the T portfolios are more skewed towards the recommended assets than the GMV portfolio.

Moving on to the metrics, the mean (expected quarterly return) for the N+K portfolio is 0.397, 0.403, and 0.525 for low, moderate, and high-risk assets, respectively. The standard deviation (StdDev), which measures the portfolio's volatility, is 0.220, 0.218, and 0.237 for low, moderate, and high-risk assets, respectively.

The Sharpe Ratio (SR) represents the risk-adjusted return, calculated by dividing the mean return by the standard deviation. For low-risk assets, the SR is 1.57 for the T portfolio and 1.32 for the GMV portfolio. Similarly, for moderate-risk stocks, the SR is 1.76 for the T portfolio and 1.10 for the GMV portfolio. For high-risk assets, the SR is 2.13 for the T portfolio and 1.21 for the GMV portfolio. Higher SR values indicate better risk-adjusted performance.

Table 6.3: The table presents the tangency (T) and global minimum-variance (GMV) portfolios of the Ultra investor for the mean-variance spanning test of N-Recommended assets with K-benchmark-assets for all three asset clusters.

	Low Risk Assets		Moderate Risk Assets		High Risk Assets	
Ultra	T	GMV	T	GMV	T	GMV
N	41.0%	37.8%	43.9%	31.5%	40.5%	8.7%
K	59.0%	62.2%	56.1%	68.5%	59.5%	92.3%
Mean	0.397	0.259	0.403	0.233	0.525	0.266
StdDev	0.220	0.182	0.218	0.193	0.237	0.203
SR	1.57	1.32	1.76	1.10	2.13	1.21
Δ StdDev		-23.5%		-18.9%		-14.6%
Δ SR	28.6%		44.3%		74.6%	

SR depicts the Sharpe Ratio. Δ denotes the change of the respective metric of the N+K portfolio with regards to the respective metric of the K benchmark portfolio.

The Large investor exhibits comparable findings to those observed in the Ultra investor case. The T portfolios exhibit a higher bias towards the suggested assets (27.2%-74.1%) than the GMV portfolios (-6.4%-17.5%). However, it is worth noting that the GMV portfolio assigns negative weights to the N-recommended assets for high-risk assets, implying a short position on those assets. In addition, the T-high-risk asset portfolio demonstrates the highest Sharpe ratio (1.57), while the GMV-low-risk portfolio exhibits the lowest standard deviation (0.182).

Table 6.4: The table presents the tangency (T) and global minimum-variance (GMV) portfolios of the Large investor for the mean-variance spanning test of N-Recommended assets with K-benchmark-assets for all three asset clusters.

	Low Risk Assets		Moderate Risk Assets		High Risk Assets	
Large	T	GMV	T	GMV	T	GMV
N	27.2%	17.5%	30.6%	14.7%	74.1%	-6.4%
K	72.8%	82.5%	69.4%	85.3%	25.9%	106.4%
Mean	0.218	0.153	0.267	0.162	0.292	0.109
StdDev	0.178	0.140	0.190	0.151	0.199	0.141
SR	1.11	0.95	1.30	0.94	1.37	0.63
Δ StdDev		-30.4%		-25.2%		-30.3%
Δ SR	23.3%		44.4%		52.2%	

SR depicts the Sharpe Ratio. Δ denotes the change of the respective metric of the N+K portfolio with regards to the respective metric of the K benchmark portfolio.

Similar findings are observed for the Medium investor, as seen with the Ultra investor. The T portfolios exhibit a higher bias towards the suggested assets (40.0%-40.6%) than the GMV portfolios, except for the GMV-Low risk portfolio that assigns the highest weight value of 42% to the N recommended assets. In addition, the T-high-risk asset portfolio demonstrates the highest Sharpe ratio (1.30), while the GMV-low-risk portfolio exhibits the lowest standard deviation (0.153), while the GMV-moderate-risk portfolio has a higher standard deviation (0.164) than the GMV-high-risk portfolio (0.159).

Table 6.5: The table presents the tangency (T) and global minimum-variance (GMV) portfolios of the Medium investor for the mean-variance spanning test of N-Recommended assets with K-benchmark-assets for all three asset clusters.

	Low Risk Assets		Moderate Risk Assets		High Risk Assets	
Medium	T	GMV	T	GMV	T	GMV
N	40.0%	42%	40.6%	17.4%	40.5%	4.0%
K	60.0%	58.0%	59.4%	82.6%	59.5%	96.0%
Mean	0.286	0.231	0.315	0.205	0.441	0.235
StdDev	0.205	0.153	0.270	0.164	0.223	0.159
SR	1.30	1.05	1.42	0.94	1.89	1.28
Δ StdDev		-27%		-22.6%		-24.8%
Δ SR	36.8%		49.5%		98.0%	

SR depicts the Sharpe Ratio. Δ denotes the change of the respective metric of the N+K portfolio with regards to the respective metric of the K benchmark portfolio.

We observe distinct outcomes for the Small investor concerning the allocations of the T and GMV portfolios. Concerning the low and high-risk asset clusterings, the GMV portfolios display a higher skewness towards the recommended assets (78.2% and 82.9%) than the T portfolios (13.1% and 25.8%). Furthermore, we observe that the GMV-Low and T-High portfolios assign greater weights to the N-recommended assets than to the K-benchmark assets (21.8% and 17.1%), a trend not observed among the three other investor groups, except for the T-high risk portfolio for the large investor with a 74.1% weight allocated to the N-recommended assets.

Table 6.6: The table presents the tangency (T) and global minimum-variance (GMV) portfolios of the Small investor for the mean-variance spanning test of N-Recommended assets with K-benchmark-assets for all three asset clusters.

	Low Risk Assets		Moderate Risk Assets		High Risk Assets	
Small	T	GMV	T	GMV	T	GMV
N	13.1%	78.2%	25.8%	20.0%	82.9%	1.4%
K	86.9%	21.8%	74.2%	80.0%	17.1%	98.6%
Mean	0.279	0.190	0.317	0.171	0.444	0.215
StdDev	0.220	0.162	0.228	0.161	0.260	0.168
SR	1.18	1.05	1.30	0.94	1.63	1.28
Δ StdDev		-28.7%		-29.1%		-26.2%
Δ SR	19.2%		31.3%		64.6%	

SR depicts the Sharpe Ratio. Δ denotes the change of the respective metric of the N+K portfolio with regards to the respective metric of the K benchmark portfolio.

Overall, the N-recommended assets had a net positive impact on the T and GMV portfolios. For the T portfolios, they increased the Sharpe ratio, For the GMV portfolios, the recommended assets helped lower the volatility. The T-High risk portfolio for the medium investor had the highest increase, 98%, in terms of Sharpe ratio. The GMV-Low risk portfolio for the large investor had the most significant decrease, -30.4%, in terms of standard deviation.

6.3 Financial Performance

This section focuses on showcasing the financial performance of the tangency and GMV portfolios containing the N+K assets, compared to the benchmark K asset portfolio. The financial performance metrics used in this section are presented and defined in Appendix A1. We provide an overview of the cumulative returns as well as risk and return statistics for the testing period spanning from start of 2020 to end of 2021 for all four types of institutional investors.

Starting with the Ultra investor, figure 6.1 shows the cumulative returns for the 6 different N+K portfolios and the Ultra benchmark portfolio. The T-High portfolio demonstrated strong outperformance compared to the benchmark Ultra, delivering higher cumulative returns, as illustrated in figure 6.1. The inclusion of the N-recommended assets in the T portfolios plays a significant role in boosting the portfolio's performance.

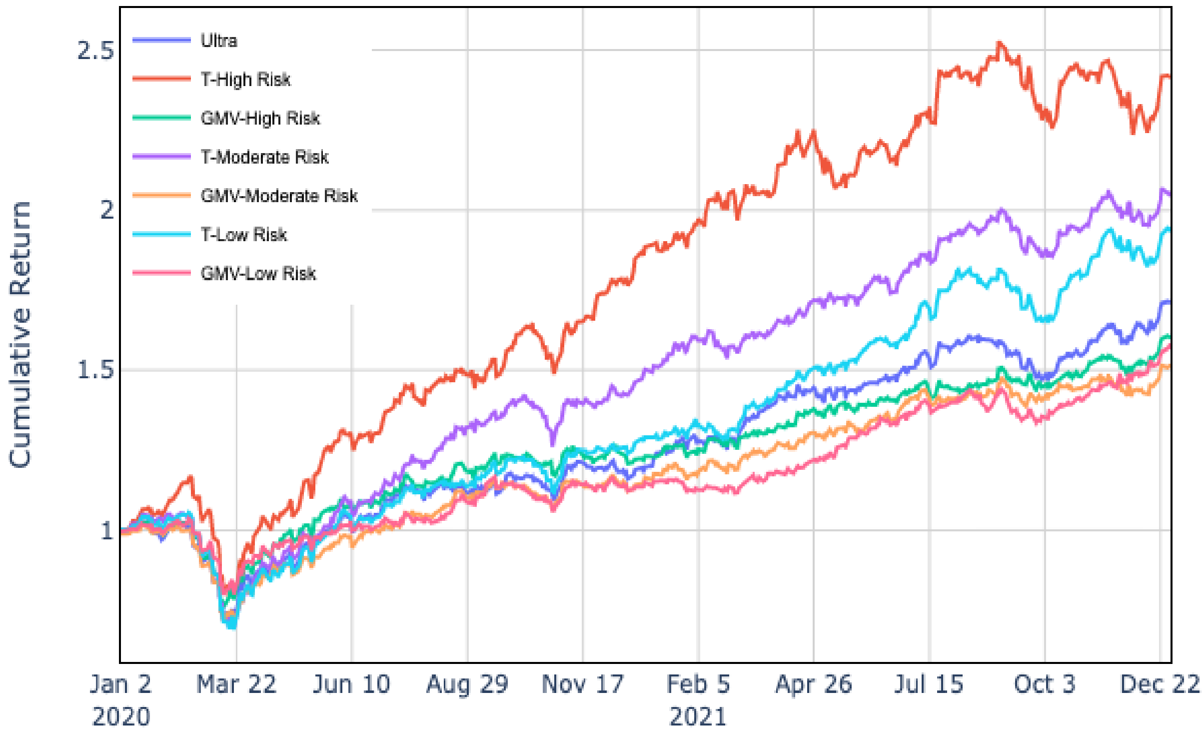

Figure 6.1: Cumulative returns for the K benchmark (Ultra) and N+K portfolios during 2020-2021.

In table 6.1, we present the return-risk statistics for all of the portfolios. The recommended assets increase the portfolio's Sharpe ratio, indicating improved risk-adjusted returns compared to the benchmark. The T-High portfolio showcases the highest Sharpe ratio (2.13) among the portfolios compared to the benchmark (1.22), suggesting that the recommended assets aid in enhancing its risk-adjusted performance.

Similarly, the T-High portfolio exhibits the highest Sortino ratio (2.62) compared to the benchmark (1.73), indicating superior risk-adjusted returns with a focus on downside volatility. The N-recommended assets likely contribute to this outcome by improving the portfolio's downside protection.

On the other hand, the GMV portfolios display lower volatility compared to the benchmark. The inclusion of the N-recommended assets in the GMV portfolios has the effect of decreasing their volatility. This suggests that the recommended assets aid in constructing diversified portfolios with reduced risk, resulting in lower overall volatility. While the GMV portfolios have lower returns compared to the T-High portfolio and the benchmark, it's important to note that their primary objective is to achieve the global minimum variance, prioritizing risk reduction rather than maximizing returns. Including the N-recommended assets in the GMV portfolios contributes to constructing portfolios that minimize volatility while still generating positive returns. The GMV Low portfolio records the lowest annual volatility (18.12%) compared to the benchmark (23.80%). It also records the lowest maximum drawdown (-24.52%) compared to the benchmark (-30.31%).

Overall, the N-recommended assets have a net positive impact on the T and GMV portfolios. For the T portfolios, they increase the Sharpe ratio, indicating improved risk-adjusted returns and higher performance compared to the benchmark. For the GMV portfolios, the recommended assets help lower the volatility, resulting in more stable performance.

Table 6.7: Risk and return statistics for the K benchmark (Ultra) and N+K portfolios during 2020-2021.

Portfolio	Annual Return (%)	Cumulative Return (%)	Annual Volatility (%)	Sharpe Ratio	Sortino Ratio	Maximum Drawdown(%)
Ultra	31.05	71.36	23.80	1.22	1.73	-30.31
T-High	52.50	141.35	23.70	2.13	2.62	-32.35
T-Moderate	40.30	105.79	21.80	1.76	2.24	-34.14
T-Low	39.7	94.49	22.00	1.57	2.05	-34.58
GMV-High	26.62	60.02	20.33	1.21	1.75	-26.43
GMV-Moderate	23.27	51.70	19.33	1.10	1.52	-30.25
GMV-Low	25.93	58.27	18.12	1.32	1.85	-24.52

Similar to the case of the Ultra investor, the findings for the Large portfolio reveal compelling insights. The T-High portfolio consistently outperforms the benchmark portfolio, delivering higher cumulative returns, as seen in figure 6.2, superior risk-adjusted metrics, and enhanced downside protection. However, it is worth noting that the outperformance of the T-High portfolio, while significant, may not be as extreme as observed in the case of the Ultra investor. Including the N-recommended assets within the T-High portfolio significantly contributes to its overall performance. Likewise, the GMV portfolios, like their counterparts in the Ultra investor case, exhibit lower volatility compared to the benchmark Large portfolio. However, this time the GMV-High portfolio performs the worst in terms of returns.

All N+K portfolios record a lower maximum drawdown than the Large portfolio, unlike with the Ultra investor where only the GMV portfolios record a lower maximum drawdown.

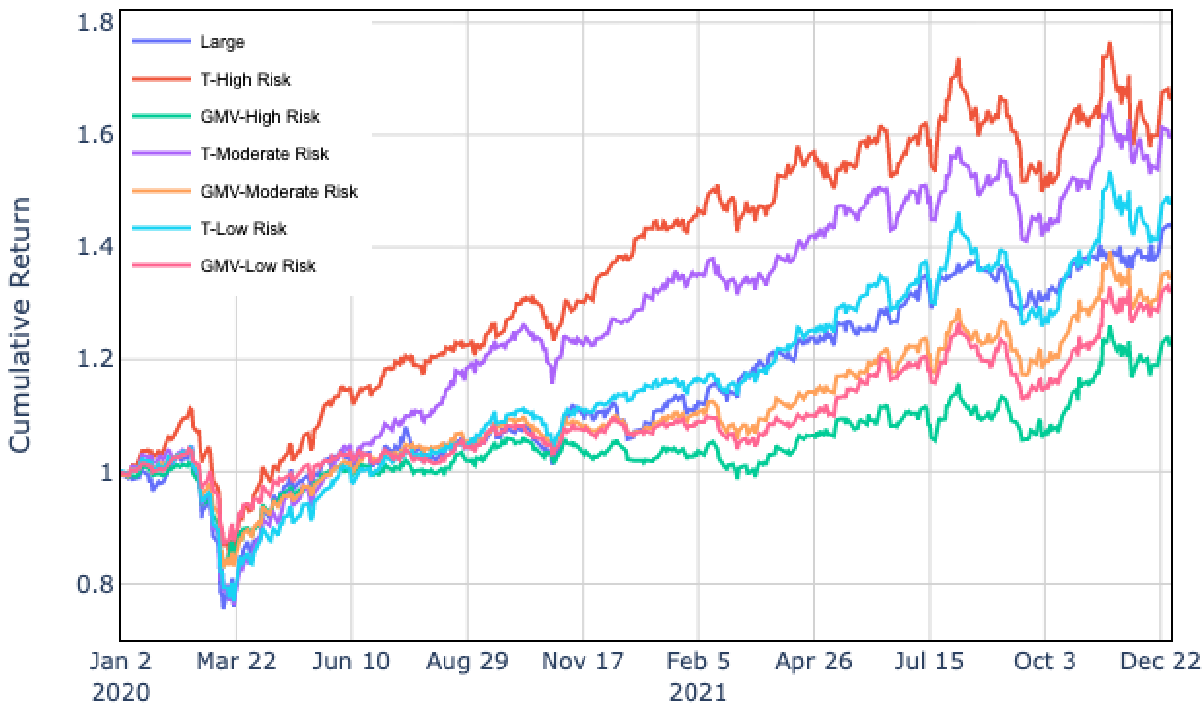

Figure 6.2: Cumulative returns for the K benchmark (Large) and N+K portfolios during 2020-2021.

As seen in table 6.8, the T-High portfolio outperforms the benchmark portfolio and the other portfolios in terms of annual return (29.2%) and with a Sharpe ratio of 1.63 compared to the benchmark of 0.99. However, the T-Moderate comes close in performance with a return of 26.65% and a Sharpe ratio of 1.3. The GMV-Low records the lowest annual volatility (14.03%) compared to the benchmark (20.16%).

Table 6.8: Risk and return statistics for the K benchmark (Large) and N+K portfolios during 2020-2021.

Portfolio	Annual Return (%)	Cumulative Return (%)	Annual Volatility (%)	Sharpe Ratio	Sortino Ratio	Maximum Drawdown(%)
Large	20.15	44.15	20.16	0.90	1.26	-27.47
T-High	29.20	66.57	19.85	1.37	1.87	-23.77
T-Moderate	26.65	60.09	18.96	1.30	1.76	-26.43
T-Low	21.79	48.08	17.82	1.11	1.50	-26.08
GMV-High	10.86	22.8	14.06	0.63	0.90	-18.29
GMV-Moderate	16.17	34.79	15.07	0.94	1.30	-20.14
GMV-Low	15.33	32.84	14.03	0.95	1.35	-16.16

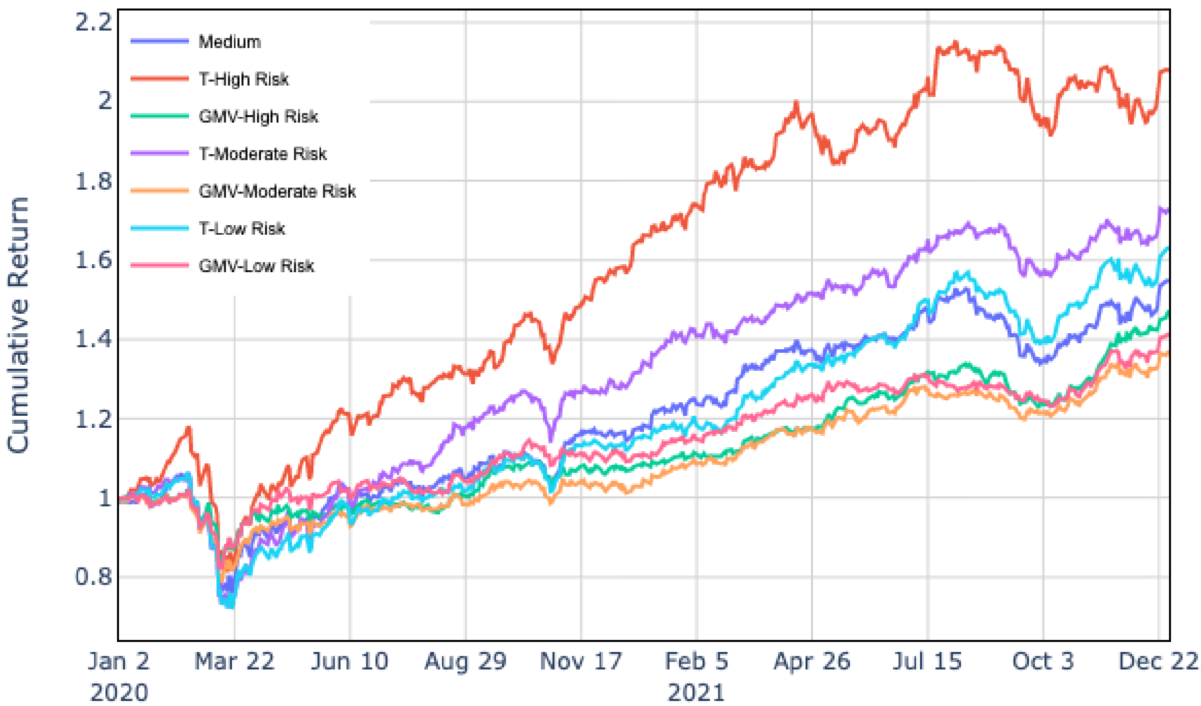

Figure 6.3: Cumulative returns for the K benchmark (Medium) and N+K portfolios during 2020-2021.

In table 6.9, we see similar results as earlier, the T-High portfolio performs better in terms of annual return, Sharpe ratio, and Sortino ratio. The GMV-Low (-19.67%) has once again the lowest drawdown, and the T-Moderate portfolio (-31.42%) has the highest drawdown.

Table 6.9: Risk and return statistics for the K benchmark (Medium) and N+K portfolios during 2020-2021.

Portfolio	Annual Return (%)	Cumulative Return (%)	Annual Volatility (%)	Sharpe Ratio	Sortino Ratio	Maximum Drawdown(%)
Medium	22.12	48.89	21.18	0.95	1.35	-24.53
T-High	44.10	107.98	22.30	1.89	2.23	-31.19
T-Moderate	31.48	73.0	20.70	1.42	1.74	-31.42
T-Low	28.64	65.15	20.50	1.30	1.59	-32.00
GMV-High	23.50	51.30	15.92	1.28	1.78	-19.67
GMV-Moderate	20.50	37.05	16.40	0.94	1.28	-22.39
GMV-Low	23.11	51.29	15.29	1.05	1.53	-17.92

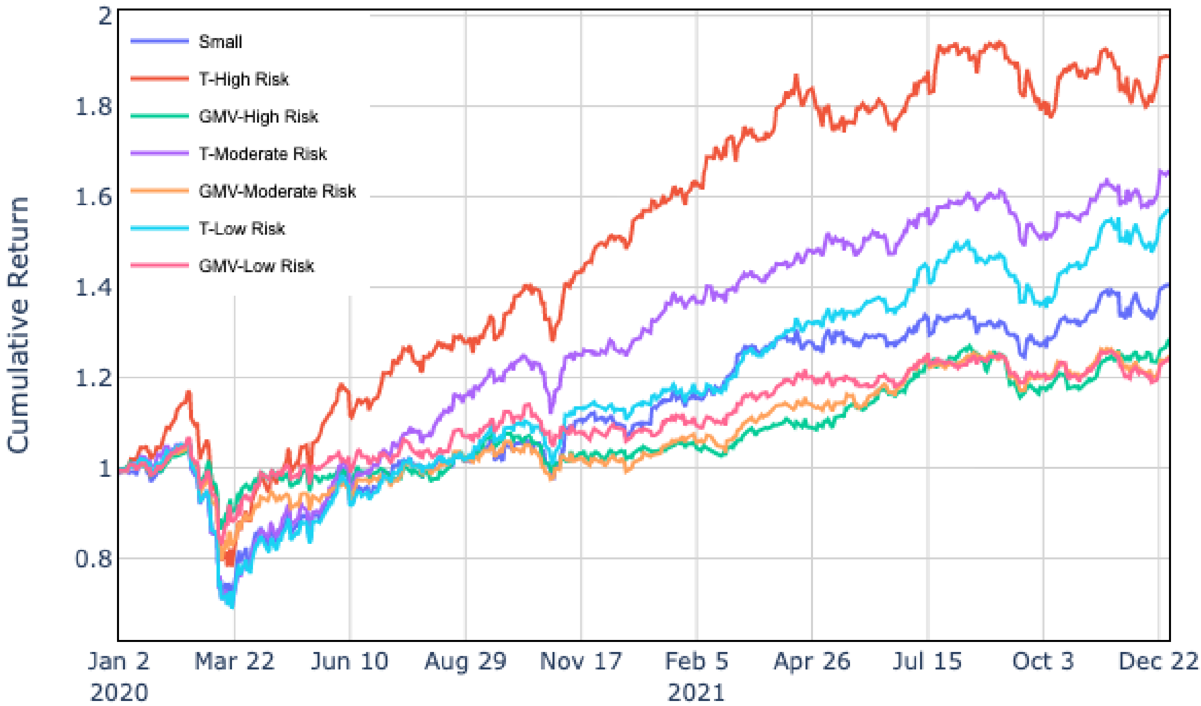

Figure 6.4: Cumulative returns for the K benchmark (Small) and N+K portfolios during 2020-2021.

In the case of the small investor we see once again that the T-High portfolio significantly outperforms the benchmark portfolio and the other portfolios in terms of annual return (44.4%) and with a Sharpe ratio of 1.63 compared to the benchmark of 0.99. The GMV-Moderate records the lowest annual volatility (16.1%) compared to the benchmark (22.73%).

Table 6.10: Risk and return statistics for the K benchmark (Small) and N+K portfolios during 2020-2021.

Portfolio	Annual Return (%)	Cumulative Return (%)	Annual Volatility (%)	Sharpe Ratio	Sortino Ratio	Maximum Drawdown (%)
Small	24.58	54.91	22.73	0.99	1.39	-28.54
T-High	44.43	107.98	26.03	1.63	2.23	-31.19
T-Moderate	31.68	73.00	22.83	1.30	1.74	-31.42
T-Low	27.90	63.26	21.95	1.18	1.56	-20.48
GMV-High	21.48	47.34	16.78	1.28	1.78	-17.88
GMV-Moderate	17.14	37.05	16.11	0.94	1.28	-22.39
GMV-Low	19.02	41.45	16.21	1.05	1.53	-19.67

6.3.1 Correlation

In Figure 6.5, we present the correlation coefficients, based on the Pearson correlation coefficient defined in Appendix A2, between the original Ultra investor portfolio and the tangency portfolios, providing insights into their interrelationships. These correlation coefficients indicate the extent to which portfolios are correlated with one another.

Among the tangency portfolios, all of them demonstrate a strong correlation with the benchmark portfolio. Notably, the T-High portfolio exhibits the highest correlation coefficient of 0.96, indicating a close association with the original portfolio. This suggests that the T-High portfolio closely tracks the performance of the original portfolio, reflecting a high degree of similarity.

Conversely, the T-Low portfolio displays the lowest correlation coefficient of 0.91 with the benchmark portfolio. This suggests a relatively weaker correlation with the original portfolio, implying some divergence in performance. The T-Low portfolio's lower correlation indicates that its returns may not align as closely with the original portfolio compared to the other tangency portfolios.

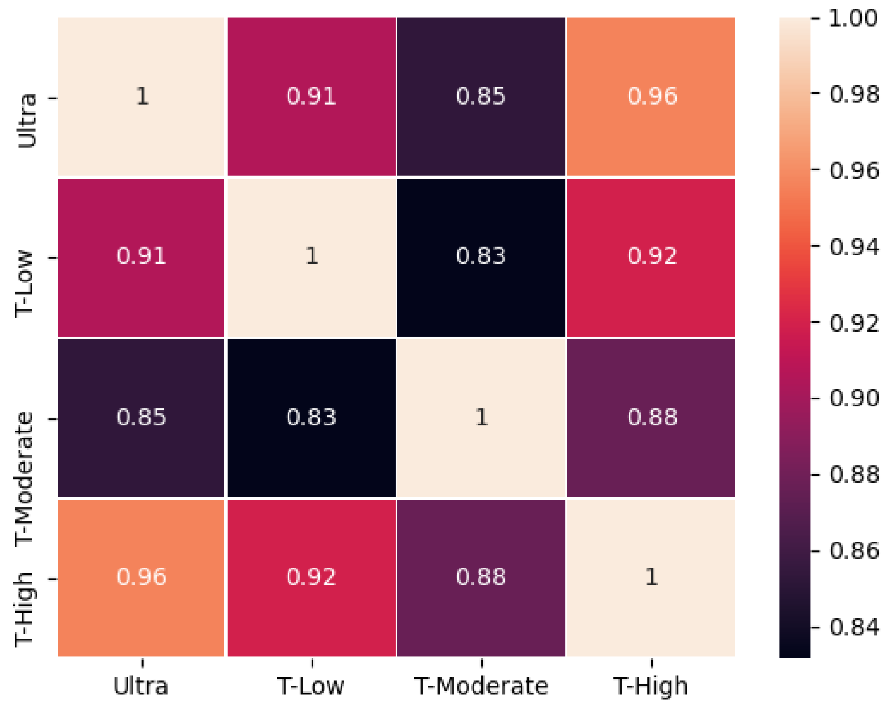

Figure 6.5: Heat map of the correlation coefficients between the original Ultra investor portfolio and the T portfolios

In Figure 6.6, we present the correlation coefficients between the original Ultra investor portfolio and the GMV portfolios, providing insights into their interrelationships. Notably, the GMV portfolios demonstrate a weaker correlation with the benchmark portfolio than the tangency portfolios.

The GMV-L portfolio exhibits the weakest correlation coefficient of 0.68, indicating a divergence in association with the original portfolio. On the other hand, the GMV-H portfolio exhibits the highest correlation coefficient of 0.78, reflecting a moderate level of correlation with the original portfolio. This implies that the GMV-H portfolio's performance aligns relatively more closely with the benchmark, displaying a greater similarity.

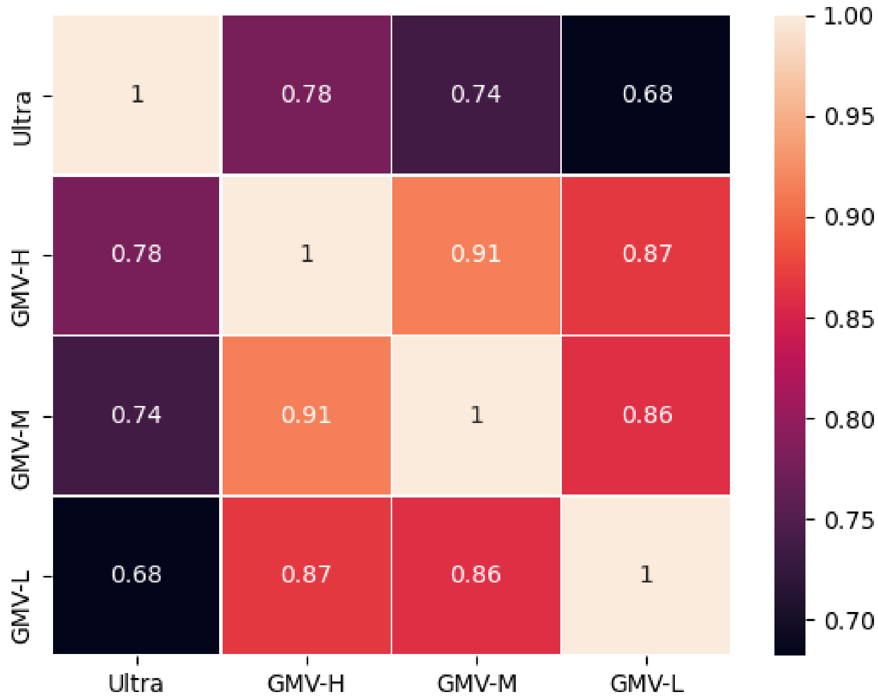

Figure 6.6: Heat map of the correlation coefficients between the original Ultra investor portfolio and the T portfolios

6.4 Discussion

The main findings presented in the previous sections provide valuable insights into the potential of personalized investment recommendation systems to uncover untapped cross-selling opportunities for institutional investors. By tailoring investment recommendations to complement the investors' current portfolios, the DLHR system demonstrated its ability to identify previously overlooked assets.

The performance evaluation of the DLHR system revealed its superiority over the baseline methods in accurately predicting user preferences. This suggests that the system has a higher capability to identify investment opportunities that align with the specific needs and preferences of institutional investors. Through the utilization of deep learning-based techniques, the DLHR system effectively captured underlying patterns in user-item ratings, leading to more accurate predictions and personalized recommendations.

The mean-variance spanning test show that the recommended stocks expand the opportunity set of the unconstrained institutional investor. The inclusion of the recommended assets improves the Sharpe Ratio of the tangency portfolio and reduces the standard deviation of the global minimum-variance portfolio. The results indicated that the recommended assets proposed by the DLHR system expanded the investors' mean-variance efficient frontier and their investment opportunity set. This implies that the personalized investment recommendations not only complemented the investors' existing portfolios but also introduced new assets that improved the overall risk-return profile.

Additionally, the analysis of the tangency (T) and global minimum-variance (GMV) portfolios revealed interesting insights. The recommended assets showed a higher bias in the T portfolios compared to the GMV portfolios, indicating that they had a greater influence on the asset allocation decisions in the T portfolios. This suggests that the recommended assets potentially uncovered new cross-selling opportunities for institutional investors.

Furthermore, the financial performance evaluation of the tangency and GMV portfolios provided evidence of the positive impact of the recommended assets on risk-adjusted returns and volatility reduction. The T portfolios exhibited higher risk-adjusted returns compared to the benchmark portfolios, while the GMV portfolios displayed lower volatility. These findings indicate that the recommended assets had the potential to enhance the performance and risk management strategies of institutional investors.

By incorporating the recommended assets into their portfolios, institutional investors can potentially diversify their holdings, access new market opportunities, and optimize their risk-return trade-offs. This highlights the untapped cross-selling opportunities that personalized investment recommendation systems can uncover by providing tailored investment suggestions that complement investors' current portfolios. In summary, the findings of this study suggest that personalized investment recommendation systems, such as the DLHR system developed and tested, have the potential to identify and exploit untapped cross-selling opportunities for institutional investors. By leveraging deep learning techniques and providing personalized recommendations, these systems can enhance portfolio performance and risk management strategies and expand the investment opportunity set for institutional investors on the Stockholm Stock Exchange.

6.4.1 Caveats

The DLHR approach utilized in this study has certain caveats that should be taken into consideration as they may impact the real-world performance and efficiency of the models. One important consideration is the testing period of two years, which may not fully capture the long-term value and performance of the portfolio constructed by the DLHR algorithms. In real-world scenarios, market capitalization-weighted tracking portfolios have a lengthy performance history and low transaction costs. However, the presented models do not incorporate transaction costs, which are a real consideration when implementing these strategies.

Managing the DLHR algorithms in a real-world environment introduces additional challenges. Timely and accurate data delivery becomes crucial as the models rely on up-to-date and correct data for making informed decisions. While the approach suggests quarterly portfolio recommendations, portfolio managers in the real world often engage in high-frequency trading, making decisions at a much faster pace. This means the models do not account for price changes that occur when the market is closed or situations where a company goes bankrupt and ceases trading.

Another important caveat relates to the holding period assumption. While the study assumes a buy-and-hold strategy for at least a quarter, institutional investors may adopt different investment strategies or have shorter holding periods. Hedge funds or mutual funds, for example, often have varying investment goals and trading strategies that result in shorter time horizons for holding stocks. Therefore, the recommendations generated by the study's method might not accurately reflect the needs and preferences of individual or retail investors with different investment horizons.

The method used in the study constructs implicit ratings based on the assumption that the more of the same stock an investor holds, the higher their preference for those stocks. However, this assumption might not capture the full complexity of investor preferences. Investors have diverse strategies, risk tolerances, and preferences going beyond the number of holdings. Factors such as company fundamentals, industry trends, and market conditions significantly affect investment decisions. Therefore, it's crucial to interpret the recommendations generated by the method with this limitation in mind and consider additional factors beyond the number of holdings.

Additionally, the quality and effectiveness of the recommendation system heavily rely on data pre-processing and feature engineering. These steps involve cleaning and transforming the data and selecting relevant features for analysis. It's important to acknowledge that limitations and biases are present during these processes, which can impact the performance of the recommendation system. For instance, incomplete or inaccurate data cleaning can distort information, and inadequate feature selection may lead to recommendations lacking accuracy or relevance. Conducting these steps with care and thoroughness is essential to mitigate any potential limitations or biases that could affect the reliability of the recommendations.

7 Conclusion

The staggering amounts of investment options led us to explore the employment of recommendation systems for personalized investment recommendations. In conclusion, the research conducted in this thesis has successfully demonstrated the effectiveness of personalized investment recommendation systems, specifically the use of a Deep learning-based hybrid recommendation system (DLHR), in uncovering untapped cross-selling opportunities for institutional investors on the Stockholm Stock Exchange. The tailored recommendations provided by the DLHR system complemented existing portfolios, identified overlooked assets, and offered personalized suggestions that diversified holdings, access new market opportunities, and optimized risk-return trade-offs.

The superiority of the DLHR system over baseline methods in accurately predicting user preferences highlights the value of employing deep learning techniques to capture underlying patterns in user-item ratings. The effectiveness of the DLHR is in line with similar previous research conducted in the area (Kiran et al. (2020), Huang et al. (2019)), suggesting that personalized recommendation systems can identify investment opportunities that align accurately with the specific needs and preferences of institutional investors.

Furthermore, the study found that investment recommendations are crucial for institutional investors and significantly contribute to their ability to make profits, in line with Chen & Cheng (2006). The tailored recommendations provided by the DLHR system were instrumental in guiding investment decisions and maximizing returns.

The mean-variance analysis conducted in this research further supported the potential for untapped cross-selling opportunities by expanding the mean-variance efficient frontier and investment opportunity set for the different types of institutional investors. Furthermore, the examination of tangency (T) and global minimum-variance (GMV) portfolios revealed crucial insights into the influence of recommended assets on asset allocation decisions. This finding signifies that the suggested assets have a substantial impact on shaping the optimal distribution of investments within institutional portfolios. By suggesting stocks that modify the initial asset allocation choices, this analysis uncovers new avenues for cross-selling opportunities among institutional investors. These opportunities arise from the ability to offer complementary products that align with the preferred portfolio, thus enhancing the overall portfolio performance and meeting the diverse needs and preferences of institutional investors.

Moreover, the financial performance evaluation of the portfolios with the added recommended assets demonstrated a positive impact of the recommended stocks on risk-adjusted returns and volatility reduction. The tangency portfolios exhibited higher risk-adjusted returns than the original benchmark portfolios, while the GMV portfolios displayed lower volatility. These findings indicate that the recommended assets have the potential to enhance performance and risk management strategies for institutional investors. We also show that institutional investors demonstrate a stronger inclination towards riskier stocks, in line with Basak & Pavlova (2013), as seen in the close association of the high-risk tangency portfolio with the original portfolios.

It is crucial to acknowledge the limitations of this study, which include the restricted testing period, the absence of transaction costs, and the assumptions made regarding holding periods and investor preferences. Future implementations of the DLHR system should address these limitations and biases to ensure the reliability of the recommendations.

Additionally, further research can explore the employment of personalized recommendation systems in different financial contexts and address the identified limitations. Incorporating richer data sources, such as an extended testing period, a broader index covering more countries, or including additional financial factors like financial news, could lead to more robust and comprehensive results. Exploring personalized recommendation systems beyond stocks, and considering other asset classes, such as private equity and private debt, would also contribute to a more comprehensive understanding of their effectiveness in supporting institutional investors.

In summary, this research has provided valuable insights into the potential of personalized recommendation systems for institutional investors, particularly the DLHR system. By leveraging deep learning techniques, these systems can effectively identify and exploit untapped cross-selling opportunities, enhance portfolio performance, and improve risk management strategies. Future advancements in personalized recommendation systems hold the promise of unlocking new opportunities and improving investment decision-making processes for institutional investors in the financial service industry.

References

Averstad, P., & et al. (2023, Mar). *Mckinsey global private markets review 2023: Private markets turn down the volume.* McKinsey & Company. Retrieved from https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/mckinseys-private-markets-annual-review

Basak, S., & Pavlova, A. (2013). Asset prices and institutional investors. *American Economic Review*, *103*(5), 1728–1758.

Bennett, J., Lanning, S., et al. (2007). The netflix prize. In *Proceedings of kdd cup and workshop* (Vol. 2007, p. 35).

Boustani, N., Emrouznejad, A., Gholami, R., Despic, O., & Ioannou, A. (2023). Improving the predictive accuracy of the cross-selling of consumer loans using deep learning networks. *Annals of Operations Research*, 1–18.

Bushee, B. J. (1998). The influence of institutional investors on myopic r&d investment behavior. *Accounting review*, 305–333.

Chalidabhongse, T. H., & Kaensar, C. (2006). A personalized stock recommendation system using adaptive user modeling. In *2006 international symposium on communications and information technologies* (pp. 463–468).

Chen, X., & Cheng, Q. (2006). Institutional holdings and analysts' stock recommendations. *Journal of Accounting, Auditing & Finance*, *21*(4), 399–440.

Chiang, J. (2021, Jun). *7 types of hybrid recommendation system.* Analytics Vidhya. Retrieved from https://medium.com/analytics-vidhya/7-types-of-hybrid-recommendation-system-3e4f78266ad8

de Prado, M. M. L., Peijan, A., et al. (2005). *Measuring loss potential of hedge fund strategies* (Tech. Rep.). University Library of Munich, Germany.

Dong, X., Yu, L., Wu, Z., Sun, Y., Yuan, L., & Zhang, F. (2017). A hybrid collaborative filtering model with deep structure for recommender systems. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 31).

Gonzales, R. M. D., & Hargreaves, C. A. (2022). How can we use artificial intelligence for stock recommendation and risk management? a proposed decision support system. *International Journal of Information Management Data Insights*, *2*(2), 100130.

González, B. S. M., del Alamo, J. M., & Yelmo, J. C. (2015). A personal data framework for exchanging knowledge about users in new financial services. In *Finrec* (pp. 19–26).

Heaton, J. (2011). Introduction to the math of neural networks (beta-1). *Heaton Research Inc.*

Huang, Z., Yu, C., Ni, J., Liu, H., Zeng, C., & Tang, Y. (2019). An efficient hybrid recommendation model with deep neural networks. *IEEE Access*, *7*, 137900–137912.

Huberman, G., & Kandel, S. (1987). Mean-variance spanning. *The Journal of Finance*, *42*(4), 873–888.

IBM. (2020). *"What are Neural Networks?"*. (Available online: https://www.ibm.com/se-en/cloud/learn/neural-networks [Accessed 20 December 2021])

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian informatics journal*, *16*(3), 261–273.

Jannach, D., & Ludewig, M. (2017). When recurrent neural networks meet the neighborhood for session-based recommendation. In *Proceedings of the eleventh acm conference on recommender systems* (pp. 306–310).

Kamakura, W. A. (2008). Cross-selling: Offering the right product to the right customer at the right time. *Journal of Relationship Marketing*, *6*(3-4), 41–58.

Kan, R., & Zhou, G. (2008). Tests of mean-variance spanning. In *Afa 2001 new orleans meetings, olin working paper.*

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kiran, R., Kumar, P., & Bhasker, B. (2020). Dnnrec: A novel deep learning based hybrid recommender system. *Expert Systems with Applications*, *144*, 113054.

Klacanova, A. (2022, May). *Intro to collaborative filtering.* GraphAware. Retrieved from https://graphaware.com/blog/business/intro-to-collaborative-filtering.html

Maillard, S., Roncalli, T., & Teïletche, J. (2010). "The Properties of Equally Weighted Risk Contribution Portfolios". *The Journal of Portfolio Management*, *vol.36*(no.4), pp.60–70.

Markowitz, H. (1952, March). "Portfolio Selection". *The Journal of Finance*, *vol.7*(no.1), pp.77-91.

Merton, R. C. (1980). "On Estimating the Expected Return on the Market: An exploratory investigation". *Journal of financial economics*, *vol.8*(no.4), pp.323–361.

Musto, C., Semeraro, G., Lops, P., De Gemmis, M., & Lekkas, G. (2015). Personalized finance advisory through case-based recommender systems and diversification strategies. *Decision Support Systems*, *77*, 100–111.

Prakash, A. J., Chang, C.-H., & Pactwa, T. E. (2003). Selecting a portfolio with skewness: Recent evidence from us, european, and latin american equity markets. *Journal of Banking & Finance*, *27*(7), 1375–1390.

Rollinger, T. N., & Hoffman, S. T. (2013). Sortino: a 'sharper'ratio. *Chicago, Illinois: Red Rock Capital*.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.

San Miguel González, B., Álamo Ramiro, J. M. d., & Yelmo Garcia, J. C. (2015). A personal data framework for exchanging knowledge about users in new financial services.

Sayyed, F., Argiddi, R., & Apte, S. (2013a). Collaborative filtering recommender system for financial market. *Int. J. Eng. Adv. Technol.(IJEAT) ISSN*, 2249–8958.

Sayyed, F., Argiddi, R., & Apte, S. (2013b). Generating recommendations for stock market using collaborative filtering. *Int. J. Comput. Eng. Sci*, *3*, 46–49.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, *39*(1), 119–138.

Taghavi, M., Bakhtiyari, K., & Scavino, E. (2013). Agent-based computational investing recommender system. In *Proceedings of the 7th acm conference on recommender systems* (pp. 455–458).

Tseng, C.-C. (2004). Portfolio management using hybrid recommendation system. In *Ieee international conference on e-technology, e-commerce and e-service, 2004. eee'04. 2004* (pp. 202–206).

Wang, X., Yu, L., Ren, K., Tao, G., Zhang, W., Yu, Y., & Wang, J. (2017). Dynamic attention deep model for article recommendation by learning human editors' demonstration. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining* (pp. 2051–2059).

Yoo, J., Gervasio, M., & Langley, P. (2003). An adaptive stock tracker for personalized trading advice. In *Proceedings of the 8th international conference on intelligent user interfaces* (pp. 197–203).

Yu, J., & Yuan, Y. (2011). Investor sentiment and the mean–variance relation. *Journal of Financial Economics*, *100*(2), 367–381.

Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbors. *Annals of translational medicine*, *4*(11).

Zibriczky, D. (2016). Recommender systems meet finance: a literature review. In *Proc. 2nd int. workshop personalization recommender syst* (pp. 1–10).

Appendix A

A.1 Financial Performance Measures

In this study, we implement eight different performance measures to evaluate the five different portfolio construction strategies. Performance statistics, discussed below, correlation, concentration and turnover. The chosen measures are in line with previous research within the area of portfolio evaluation (Maillard et al., 2010). The mentioned measures should provide an in-depth discussion of the differences between the compared portfolio strategies. They should also tell whether DRL is a competitive and valuable strategy in the area.

The Swedish risk-free interest rate for the whole period 2011-2021, obtained from the Swedish national bank, is used whenever a risk-free rate is needed for the performance measures.

Annual return

Annual return is the annualised geometric average return. The Annual return for each portfolio was calculated using the bellow formula (CAGR=compound annual growth rate):

$$CAGR = \left(\left(\frac{Final\ value}{Initial\ value} \right)^{\frac{1}{years}} \right) - 1 \tag{A.1}$$

Cumulative Return

Cumulative returns are cumulative sum of the daily returns. It can also be calculated as a single number, based on the final and initial value:

$$CR = \frac{Final\ value - Initial\ value}{Initial\ value} \tag{A.2}$$

Annual volatility

The annual volatility (standard deviation) is the annualised volatility. The annual volatility of the portfolio was calculated by multiplying the daily volatility by $\sqrt{252}$, since there are usually 252 trading days in a year.

Sharpe ratio

The Sharpe ratio is a financial measure introduced by Nobel laureate Sharpe (1966). The Sharpe ratio measures the performance of an investment´s, one single security or portfolio of securities, return compared to its risk. Sharpe (1966) defines the ratio as:

$$S_P = \frac{E[r_P - r_f]}{\sigma_P} \tag{A.3}$$

Maximum drawdown

Maximum drawdown is the largest peak-to-through downturn of the portfolio value. The maximum drawdown of each portfolio was calculated as:

$$MDD(T) = \max\{0, \max_{t \in (0,t)} P(t) - P(T)\} \tag{A.4}$$

Where T is the time at the end of the period and $P(t)$ is the stock price at time t.

Sortino ratio

The Sortino ratio is similar to the Sharpe ratio, it aims to provide risk-adjusted return. The Sortino ratio of the portfolios was calculated using the following formula:

$$\text{Sortino Ratio} = \frac{E(r)}{\sigma_d} \tag{A.5}$$

Where σ_d is the standard deviation of the negative results (downside).

A.2 Pearson correlation coefficient

The correlation of the portfolios and the items is computed using the Pearson correlation coefficient, calculated as the covariance of the two portfolios divided by the product of their standard deviations.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{A.6}$$

$$= \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)\left(\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right)}} \tag{A.7}$$

A.3 Top 40 Institutional Investors

Investor	Portfolio Value (MSEK)	Country
Investor	516653,6	Sweden
BlackRock	514150,9	USA
Vanguard	385124,3	USA
Capital Group	321645,8	USA
Swedbank Robur Fonder	320489,6	Sweden
Norges Bank	249767,2	Norway
AMF Pension & Fonder	227656,0	Sweden
Alecta Tjänstepension	219939,4	Sweden
Handelsbanken Fonder	214027,8	Sweden
SEB Fonder	176733,4	Sweden
Fidelity Investments (FMR)	170953,0	USA
Stefan Persson & Familj	141803,3	Sweden
Industrivärden	137764,0	Sweden
Knut och Alice Wallenbergs Stiftelse	135704,5	Sweden
Geely Holding	130141,7	China
Volkswagen AG	125848,6	Germany
T. Rowe Price	124723,2	USA
Gustaf Douglas	111896,1	Sweden
Nordea Fonder	109787,6	Finland
Melker Schörling AB	109787,1	Sweden
Wellington Management	106091,9	USA
Lundbergföretagen AB	105313,0	Sweden
Fjärde AP-fonden	90547,3	Sweden
Första AP-fonden	84756,7	Sweden
Folksam	77542,9	Sweden
Carl Bennet	75523,7	Sweden
Länsförsäkringar Fonder	73444,3	Sweden
State Street Global Advisors	68812,9	USA
EQT	68621,5	Sweden
Cevian Capital	68193,9	Sweden
Investment AB Latour	64616,1	Sweden
PRIMECAP	64537,0	USA
Avanza Pension	63742,6	Sweden
Invesco	61697,4	USA
Tredje AP-fonden	60021,9	Sweden
Livförsäkringsbolaget Skandia	59125,8	Sweden
Fredrik Lundberg	58545,6	Sweden
Storebrand Fonder	56512,4	Sweden
Government of Japan Pension Investment Fund	56097,8	Japan
Spiltan Fonder	55373,0	Sweden