Master Thesis

Predicting Repo Rollovers

Author: M.S.M. Sliepenbeek (2620324)

1st Supervisor: Prof. dr. Rob van der Mei
Company Supervisor: Natalie Kessler (De Nederlandsche Bank)
2nd Reader: Prof. dr. Mathisca de Gunst

I am thankful for the insight comments by my colleagues at De Nederlandsche Bank, especially the Data Science Hub. The views expressed here are those of the author and do not necessarily reflect the official positions of De Nederlandsche Bank. A thesis submitted in fulfillment of the requirements for the Vrije Universiteit Amsterdam Master of Science degree in Business Analytics.

August 1, 2023

Abstract

This thesis delves into the fundamental concepts and advanced techniques of predicting repurchase agreement (repo) rollover transactions. This is done with transactional level data for the European repo market which has been cleaned to work with overnight transactions to find and identify rollover transactions. This is done to implement the feature and identify possible rollover risk in the market. By analyzing the results over the time period July 2020 to July 2022 it can be said that the rollover transaction characteristics have on average a lower principal amount for the collateral, a higher agreed-upon fixed interest rate, and less clearing by a CCP. For the prediction of the rollover transaction to understand the characteristics and formulate ways that it could cause rollover risk a few supervised machine learning classifiers are used and tuned. These classifiers include Logistic Regression (as the benchmark model), Gaussian Naive Bayes, Decision Tree, Random Forest, XGBoost, and Stochastic Gradient Descent. The macro F1-scores for the Decision Tree and Random Forest models are the best with *0.79* as is the AUC-score of *0.88* and *0.89* for the Decision Tree and Random Forest model respectively. The Gaussian Naive Bayes model is overall the best in correctly identifying the rollover transactions with a True Positive normalized score of *0.816*. These models can predict whether a transaction is part of a rollover transaction and evaluate which features are important. The most important features for predicting are the indicator if the transaction has been cleared, if the transaction was set on a European trading venue and the fixed agreed-upon interest rate. Of all the transactions in the overnight European repo market around 21% is part of a rollover. Therefore, it is interesting to see that during this time period, no clear shocks can be seen over time, or in stress periods.

Acknowledgements

This thesis is written in fulfillment of the requirements for the Master of Science degree in Business Analytics at the Vrije Universiteit Amsterdam (VU). It focuses on predicting repo rollovers and the risk it brings to the European money market with the help of machine learning classifiers. The research for this thesis has been conducted during a six-month graduation internship at De Nederlandsche Bank.

First, I would like to thank my supervisor Natalie Kessler from DNB for her guidance, enthusiasm, and weekly meetings during the internship. Similarly, I thank all my other direct colleagues from the Data Science Hub for their advice and the great atmosphere during the weekly office days. It all made my time at the DNB very pleasant! Furthermore, I would like to express my gratitude towards Rob van der Mei and Matisca de Gunst from the VU for taking the time to be my first supervisor and second reader respectively. Lastly, I thank my manager Iman van Lelyveld for a spot on the team and for giving me the opportunity to write my thesis at the DNB.

Contents

1 Introduction **5**
1.1 Problem Statement . . . 7
1.2 Research Aim . . . 9
1.3 Outline . . . 11

2 Literature Review **12**
2.1 Repurchase Agreement Market . . . 12
2.2 Rollovers . . . 13
2.2.1 Rollover Risk . . . 15
2.2.2 Market freeze . . . 16
2.3 Machine Learning in Finance Markets . . . 18

3 Data **20**
3.1 Data Description . . . 20
3.2 Data Collection . . . 22
3.3 Data Cleaning . . . 23
3.3.1 SQL Cleaning . . . 23
3.3.2 Outliers . . . 27
3.4 Data Exploration . . . 29
3.5 Publicly Available Data . . . 34
3.6 Overview Summary . . . 35

4 Methodology **38**
4.1 Programming a Rollover . . . 38
4.1.1 Preparation Python . . . 39
4.2 Machine Learning Classifiers . . . 41
4.2.1 Logistic Regression . . . 42
4.2.2 Gaussian Naive Bayes . . . 43
4.2.3 Decision Trees . . . 44
4.2.4 Random Forest . . . 46
4.2.5 XGBoost . . . 46
4.2.6 Stochastic Gradient Descent . . . 48
4.3 Hyperparameter Optimization . . . 49
4.4 Performance Evaluation . . . 51
4.4.1 Confusion Matrix . . . 51
4.4.2 Accuracy . . . 52

4.4.3 Precision . . . 53
4.4.4 Recall . . . 53
4.4.5 F1-score . . . 53
4.4.6 Macro Average . . . 54
4.4.7 Weighted Average . . . 54
4.4.8 Area Under the Curve (AUC) . . . 55
4.4.9 Performance Measure Comparison . . . 56

5 Results **58**
5.1 Rollover Statistics . . . 58
5.2 Machine Learning Classifiers . . . 67
5.2.1 Logistic Regression . . . 67
5.2.2 Gaussian Naive Bayes . . . 70
5.2.3 Decision Tree . . . 71
5.2.4 Random Forest . . . 73
5.2.5 XGBoost . . . 76
5.2.6 Stochastic Gradient Descent . . . 78
5.3 Comparison . . . 79
5.4 Possible Policy Implementation . . . 82

6 Conclusion **85**

7 Discussion **87**

A Appendix **94**
A.1 Definitions . . . 94
A.2 De Nederlandsche Bank . . . 96
A.3 Business Understanding . . . 97
A.3.1 Repo Market in Numbers . . . 98
A.3.2 Liquidity . . . 100
A.4 Public Data Figures . . . 101
A.5 Additional Information on Repo Contracts . . . 105
A.6 Unoptimized Results Machine Learning classifiers . . . 106

1: Introduction

The repurchase agreement market functions mainly as a source of short-term liquidity and a market for (high-quality) collateral. In a repurchase agreement, or *repo*, a borrower (often a financial institution) sells securities (such as government bonds) to a lender (typically a money market fund or another financial institution) with an agreement to repurchase them at a later date. A properly functioning repo market is essential for the provision of liquidity and high-quality assets for many financial and non-financial institutions. The repo market is generally regarded as low risk due to its collateralized nature, where short-term loans are secured by high-quality assets, ensuring a high level of security for lenders. However, even in the eurozone's repo market, there is no such thing as a riskless financial instrument, and despite its importance much of the repo market's workings are still not well understood, including how the market reacts to changes in banking regulation. Most central banks rely on the repo market as the main channel for the transmission of monetary policy to the wider financial market and to provide emergency assistance to the monetary policy changes [1]. In this report a risk analysis is done to find repo rollovers in the European repo market as rollover transaction come with rollover risk. Therefore the main goal is to see if Machine Learning classifiers can predict repo rollover transactions and their behaviour in the market.

At the initiation of the transaction the borrower (seller) and the lender (buyer) enter into a repo agreement, where the borrower sells the securities to the lender in exchange for cash. The securities serve as collateral for the loan. The repo agreement specifies a term, which can be as short as overnight or as long as several months. The borrower agrees to repurchase the securities from the lender at the end of the term. As the initial repo term approaches expiration, both parties have a decision to make. The borrower can either repurchase the securities and close the repo or roll over the repo by entering into a new agreement with the lender to extend the term. If both parties agree to a rollover, they negotiate the terms of the new repo agreement, including the new maturity date, interest rate, and any adjustments to collateral. The borrower typically repurchases the securities and simultaneously enters into a new repo with the lender. The process of rolling over the repo can continue repeatedly, with subsequent rollovers occurring as each repo term nears expiration. While rollovers are common in the repo market, they do pose certain risks, including rollover risk. Rollover risk refers to the potential difficulties or increased costs a borrower may face when attempting to roll over their repo agreements. Some factors that contribute to rollover risk include:

1. Market conditions: Rollover risk can be influenced by prevailing market conditions, such as changes in interest rates, liquidity conditions, and overall confidence in the financial system. If market conditions deteriorate, lenders may become more reluctant to roll over repos or may demand higher interest rates or additional collateral.

2. Counterparty risk: The creditworthiness of the borrower can impact rollover risk. If the borrower's creditworthiness deteriorates, lenders may be less willing to roll over the repo or may impose stricter terms to mitigate their risk.

3. Collateral valuation: The value of the underlying collateral can fluctuate over time. If the value of the securities pledged as collateral declines significantly, the lender may require additional collateral or refuse to roll over the repo.

4. Funding availability: The ability of the borrower to secure funding in the market plays a crucial role. If there are disruptions in the funding market or the availability of financing decreases, it may become challenging for the borrower to roll over their repos.

These factors, among others, contribute to the rollover risk that borrowers face when seeking to extend their repo agreements. Understanding and managing rollover risk is essential for participants in the repo market, as disruptions or difficulties in rolling over repos can have implications for funding stability and market liquidity.

Rolling over a repo transaction can provide continued access to short-term funding, thereby enhancing liquidity for market participants. Furthermore, it allows for flexibility to the borrower, by extending the agreement, they can adjust their financing arrangements based on changing market conditions or their own requirements. In other words, the borrower can continue to use the cash obtained from the initial sale of the security (collateral), while the lender can earn additional interest on their investment. Nonetheless, rolling over a repo transaction also carries risks, such as increased counterparty risk. Extending the duration of a repo exposes both parties to an extended period of counterparty risk. If the other party defaults during the rollover transaction period, it can lead to potential losses for the parties involved. Also, if interest rates rise during the rollover period, the borrower may end up paying a higher interest rate compared to the current market rates. Lastly, if unexpected market fluctuations arise, such as a sudden need for additional liquidity or a desire to explore alternative financing options, the extended duration of the repo may limit flexibility.

1.1 Problem Statement

This project aims to find and try to predict repo rollover transactions. This research is conducted on the ground for various reasons. First of all the dataset provided to work with the European repurchase agreement market is relatively new for use at the national central banks within the eurozone. Therefore limited knowledge exists regarding the repo market or analysis of specific topics regarding the SFTR dataset. This gives the opportunity for projects within de Nederlandsche Bank (DNB) to analyze and work with this new dataset. Moreover, it is known that investors in the repo market roll over their transactions. However, no variable or relevant information exists in the data regarding the repurchase agreement market in Europe. In other words, the data set does not contain a feature that identifies when a transaction has been rolled over or is part of a rollover. Knowing when a transaction has been rolled over can provide valuable information for risk assessment, tracking rollovers helps in assessing the potential liquidity and refinancing risks associated with ongoing obligations. With this information, the European Central Bank (ECB) and DNB can evaluate the potential vulnerabilities of borrowers and lenders in managing short-term funding needs. Furthermore, it can highlight periods of increased or decreased rollover activity, which may indicate shifts in market conditions, investor preferences, or overall market sentiment. Because during periods of market turbulence or financial stress, borrowers may face challenges in rolling over their repos. An increase in failed or delayed rollovers can be a signal of liquidity stress or counterparty concerns. Monitoring rollover activity can help detect early warning signs of potential funding market disruptions. Overall, understanding when a transaction has been rolled over in a dataset provides valuable information about the maturity structure, market dynamics, funding market stress, investor behavior, and regulatory oversight in the repo market. It contributes to a comprehensive understanding of short-term funding dynamics and can be instrumental in assessing risks and market conditions. Moreover, it can be helpful to analyze when parties stop rolling over their transaction to see when rollover risk arises or as with the great financial crisis a complete repo run. A repo run refers to a situation where a significant number of participants in the repo market simultaneously attempt to stop the rollover or unwind their repo transactions, leading to a sudden and severe disruption in market liquidity. This can occur when market participants lose confidence in the ability of their counterparties to fulfill their obligations or when they anticipate difficulties in rolling over their repos. Given that there is a lack of research on this topic can be due to (1) lack of relevance, (2) impossibility, or (3) lack of data. Taking into account that such detailed repo data is new it is most probably point (3), which makes this research a contribution to this topic and market. As non-renewal of rollovers poses a substantial liquidity risk, point (1) can be ruled out. As for point (2) with this new dataset it is now possible to implement rollovers for this thesis project.

1.2 Research Aim

This research project aims to develop a Machine Learning classifier that generates rollover predictions for the European repo market. Furthermore, a simple benchmark model will be implemented to compare the other model's results. Ultimately, this report will contain an in-depth analysis of repo rollover transactions and if it is possible to predict those transactions for rollover risk mitigation. With the implementation of whether a transaction is a part of a rollover group, the DNB or ECB can have more insight into how traders behave and whether economic stress periods trigger a repo run. To my knowledge, this paper is the first to offer a comprehensive analysis of repo rollovers in the European money market. The method to find rollover transactions is structured as follows:

1. I start by defining the classification problem. The target variable to predict are repo rollovers. The repo rollovers to predict are based on the available features from the dataset and publicly available data. The binary classification label *1* means that a transaction is part of a rollover and *0* means that the transaction is not part of a rollover (the contract will not be renewed).

2. The dataset contains relevant features to program the repo rollovers. For a classification problem, the dataset should include a diverse range of data points to start solving the problem. This means that the data should be cleaned and preprocessed to handle missing values and outliers.

3. We do not know the importance of other variables for the prediction of repo rollovers. Therefore, feature selection techniques can help identify the most relevant features for the prediction problem. An indication can be made by the literature review and known important factors in the general market.

4. The dataset should be divided into a training and test set to evaluate the machine learning models. The training set will be used to train the models, while the test set is used to evaluate their performance.

5. The chosen models for the classification problem are Logistic regression (as a benchmark), Gaussian Naive Bayes, Decision Tree, Random Forest, XGBoost, and Stochastic Gradient Descent which are generally good options for a classification problem. These algorithms are then implemented in Python.

6. The chosen classification models are trained using the training dataset. Hyperparameter tuning is applied if necessary, and the performance is evaluated using evaluation metrics such as accuracy, precision, recall F1 score and area under the ROC curve (AUC-ROC). The results are then compared to find the most effective algorithm.

7. Based on the classifiers' results, the performance and feature importance of the models are analyzed. This also includes general rollover statistics found and visualized for a good overview.

8. Based on the evaluation results a conclusion can be made on the characteristics of rollover transactions and its risks. This will contribute to existing work as the dataset on the European repo market is relatively new and, to my knowledge, no data analysis has been done similar to this work.

The advantage of using Machine Learning classifiers is that they can provide enhanced predictive capabilities, adaptability, and real-time insights, making them useful tools for predicting repo rollovers and supporting risk management efforts in the repo market. Additionally, the classifiers can analyze large volumes of data and capture complex patterns and relationships that may not be easily identifiable through traditional analytical methods. Repo markets can experience shifts in dynamics and changing market conditions, while Machine learning classifiers can adapt and learn from new data, allowing them to continually update their predictive capabilities based on evolving market trends. This adaptability makes them well-suited for capturing changing patterns and improving prediction accuracy over time. By incorporating machine learning predictions into risk assessments and decision-making processes, market participants can enhance their understanding of rollover risks and make more informed decisions regarding funding strategies, collateral management, and counterparty exposures. The main research questions of this report are:

- How well do certain Machine Learning classifiers perform when predicting repo rollover transactions in the European repo market?
- Are there distinct patterns or trends in repo rollover behavior?
- What are the key characteristic features that influence repo rollover decisions?

Overall, this thesis aims to find repo rollovers in the dataset for more information and to analyze liquidity and rollover risk in the European repo market. Furthermore, with the help of machine learning classifiers, the rollover transactions can be predicted and give more interpretability, as it is essential to gain insight into the classifier's decision-making process.

1.3 Outline

The aim of this paper is to detect rollover transactions in the European repurchase agreement market and try to predict them using machine learning classifiers. The paper is divided into 7 parts, which highlight different aspects of the project. In Chapter 2 known literature is discussed and examined how this research could be compared to those or improve. In Chapter 3 the dataset is described, the processing steps explained and data analysis is presented to have a better understanding of the repo transactions as well as potential patterns. In Chapter 4 the methodology is explained, including the chosen models and performance metrics. In Chapter 5 the results are presented. First, the results regarding programming the rollover transactions and the statistics are presented. Next the results from the machine learning classifiers are presented and compared. In Chapter 6 the research is concluded and briefly summarized, describing the results and the recommended policy implementations. Concluding, in Chapter 7 the discussion and limitations.

2: Literature Review

In this Chapter known literature from the repo market and rollover risk are reviewed. In the Appendix Section A.2 and A.3 more is reported on the DNB, repo market and liquidity risk for readers that need background information. The key takeaways from the Appendix section are that: this research is conducted at DNB, the repo market is a considerably big financial market sector for safe investment, cheap borrowing, yield investment, and short-selling, most repo transactions are overnight and rollover risk is a part of liquidity risk.

2.1 Repurchase Agreement Market

The repo market has not been as extensively analyzed as other financial topics. This can be due to the fact that the repo market primarily involves short-term borrowing and lending of securities, typically between financial institutions. As such, the depth of literature may be more limited compared to more mainstream financial topics that attract broader interest. Furthermore, the technical complexity of repo transactions involves intricacies related to collateral, interest rates, maturity terms, and counterparty risks. Moreover, data on repo transactions are mostly not publicly available.

Repos are quite familiar instruments, to central banks in particular, who use them in their open market operations. The pioneering works on repos by Duffie (1996)[2], Duffie et al. (2002)[3], and more recently, Bottazzi et al. (2012)[4] provide models that explain how various constraints affect repo rate levels. Central banks have undertaken a fast-growing number of empirical studies because of their need to understand the repo market, making use of their privileged access to transaction data, most of which is their own data.

The paper of Kotidis and van Horen (2019)[5] contributes to the understanding of the repo market by examining how the market reacts to the changes in leverage ratio in capital regulation in the wake of the global financial crisis. Similarly, they use transaction-level data and show that the leverage ratio does not seem to affect collateral re-use in the repo market. The paper highlights that working with transactional-level data is important to analyze market behavior. This is also the case for this thesis where the analysis is done on transactional level data.

This paper contributes to the literature in a less complex way, as most existing articles on the repo market focus on the specific infrastructure, where it is assumed that the reader is already specialized in the repo market. Therefore, this could be a more accessible paper to start learning about the repo market.

2.2 Rollovers

A repo transaction that is rolled over to the next day means that the initial agreement has been extended or renewed for an additional day. In other words, the original parties agree to continue the repo for another day, typically at the same terms as the initial transaction. This process can continue for several days or even weeks, depending on the terms of the agreement. Rolling over a repo transaction can be beneficial for both parties, as it allows the borrower to continue to use the cash obtained from the initial sale of the security, while the lender can earn additional interest on their investment. However, rolling over a repo transaction also carries risks, such as changes in the market value of the security, counterparty default risk, and interest rate fluctuations.

From the research discussion paper of the Reserve Bank of Australia (2016) [6] they make an innovating algorithm that is able to identify rollovers in the interbank overnight cash (IBOC) market. This is one of the first papers to successfully identify rollover loans, in a different market, using stata and Matlab with the Furfine algorithm. The algorithm used works in two parts. The first part is the Furfine algorithm extended to identify rollovers. The second part identifies loans with features that cannot be picked up by the Furfine algorithm. Using transactional level data, existing Furfine-type algorithms typically identify a pair of transactions that could feasibly be considered the *first* payment and *second* payment leg of a loan. This is done by matching transactions for each pair of counterparties over two consecutive trading days, where the second transaction is equal to the repayment of the first plus a feasible amount of interest. The implied interest rate may change on a daily basis. The rollover period is set to 15 calendar days. Their algorithm is implemented in stata using over 7 million transactions in the period 2005-2015. The research paper concludes that rollover loans are set early during the day (in Australia), the number of rollover loans has significantly dropped after the Great Financial Crisis, with only 10% being rollover loans in 2015, and, the value of rollover loans are lower than non-rollover loans.

Canadian firms use substantial amounts of short-term funding. Between one-half and two-thirds of outstanding repo transactions in Canada have a term of fewer than seven days and are rolled over regularly, primarily overnight [7]. In aggregate, dealers maintain large gross positions with a tenor of fewer than seven days. Rollover risk is a vulnerability even if participants hold repo and reverse positions with matched maturities since only the borrowing leg is exposed to this risk [7]. In the United States, short-term repos are often used for longer-term financing, which is achieved by rollovers with the same counterparty. This is typically the case in transactions conducted over the phone [8]. According to the International Capital Market Association (ICMA) recommendation on SFTR reporting open repos are single transactions that run continuously until they are terminated and are not daily roll-overs into new transactions. If, however, for some reason, parties decided to contractually terminate an open repo and then transact a new open repo on the same or similar terms (for example, to reduce the size of the balance sheet on a reporting date), the new repo would be a new transaction [9]. However, ESMA's final guidelines of January 2020 imply that an open repo is periodically renewed, that is, rolled over. This is considered incorrect according to ICMA [9]. Moreover, the ICMA reports that a rollover from a maturing transaction into a new transaction should be reported only as a new transaction. As in the case of other maturing fixed-term transactions, there is no need to report the maturity date, for the transaction being rolled over [9]. This is because a rollover transaction is an express contractual agreement between the parties to an early termination of the existing repo and its replacement with a new repo. According to this information, a rollover transaction is therefore reported as a *new* transaction rather than an update, modification, or specifically rollover. This means that it is not clear when parties roll their transaction and no rollover or liquidity risk can be detected. The guidelines of ICMA is followed in this research.

In a working paper of Bank for International Settlements (BIS) [10] they mention the use of overnight repos became so prevalent before the financial crisis that Wall Street investment banks were rolling over a quarter of their balance sheet overnight. The change in the liability structure during boom times in favor of wholesale funding also leads to a shortening of funding maturities.

This paper contributes to the understanding of how rollover transactions behave and which features are important. Therefore, we can better understand the borrower and lending behavior of the parties involved. This can then help with policy changes if the market is less stable.

2.2.1 Rollover Risk

From existing literature, the specific topic on repo rollover is very slim. Even on just the rollover risk, there is not (yet) a considerable amount of literature compared to other financial topics to my knowledge. The risk in the repo market is smaller than in other markets, as the underlying collateral can be seized by the other party. Nonetheless, the great financial crisis of 2008 has shown that a market freeze can occur due to the withdrawal of rolling transactions over.

The paper of Martin et al. (2010)[11] analyzes the phenomenon of repo runs, which occur when repo lenders refuse to roll over their repo loans. It identifies factors that contribute to the occurrence of repo runs and proposes policy solutions to mitigate the risk. Features such as the unwind mechanism in the tri-party market or the difference between bilateral and tri-party repo lending and should thus contribute to a better understanding of the fragility of wholesale banking markets.

In a working paper of the BIS [10] the Bank of Italy has conducted several bottom-up liquidity stress tests focusing on potential weaknesses arising from the drying-up of wholesale funding. The bank assessed an additional stress hypothesis in the weekly liquidity risk monitoring framework. The wholesale funds and maturing term deposits enter with a 0% rollover rate. This resulted in a wholesale market freeze and the bank had to survive the shock using its available central bank-eligible securities.

Another paper on rollover risk by Bouvard et al. (2014) [12] highlights how increased transparency, such as better disclosure of information and reduced uncertainty, can mitigate rollover risk and enhance market stability. The findings emphasize the importance of transparency-enhancing policies in reducing the likelihood and severity of financial crises.

2.2.2 Market freeze

A market freeze refers to a situation in which trade does not occur despite the potential gains from trade. An example is the collapse of trading in mortgage-backed securities during the great financial crisis. Moreover, assets have become more risky, so investors are reluctant to hold them [13]. Until recently, short-term repos had always been regarded as virtually risk-free instruments and thus largely immune to the type of rollover or withdrawal risk associated with short-term unsecured obligations [14]. The intuition for the market freeze result can be explained as follows. When information arrives slowly relative to the rollover frequency, it is likely that no new information will have arrived by the next time the debt has to be refinanced. The upper bound on the amount of money that can be repaid is the debt capacity at the next rollover date. Since there is a small liquidation cost, issuing debt with a face value greater than the next period's debt capacity is unattractive. So the best the borrower can do is to issue debt with a face value equal to the next period's debt capacity assuming no new information arrives. But this locks the borrower into a situation in which he is forced to act as if his condition remains the same forever [14].

The financial crisis of 2007-09 featured large-scale losses to financial institutions from assets such as AAA-rated tranches of mortgage-backed securities. Simultaneously, markets for collateralized borrowing (repos) froze or experienced severe stress. Repos could only be rolled over with successively high levels of over-collateralization, which disrupted the financing model of broker-dealers and in fact, caused Bear Stearns to fail in March 2008 [15]. This systemic risk exposure only came to be when AAA-rated tranches of a diversified pool of mortgages fail when there is a secular decline in house prices that causes homeowners to default across the entire economy. Similarly, the only time rolling over repos becomes untenable is when the financial sector is experiencing systemic stress so that repo lenders become concerned about counterparty risk and forced selling of collateral in crowded and illiquid markets. In other words, though AAA-rated tranches and repo financing are relatively safe, their entire risk is systemic in nature [15].

The paper of Hur and Kondo (2013) [16] presents a theoretical framework to understand the dynamics of rollover risks in mitigating financial crises. It analyzes how countries' reliance on short-term debt and the possibility of sudden stops in capital inflows can lead to vulnerabilities in their financial systems. The results suggest that countries with higher levels of foreign reserves are better equipped to withstand sudden stops and rollover risk, providing a potential strategy for policymakers to enhance financial stability and reduce the likelihood of severe economic crises. Rollover risk is in their environment endogenous because the actual amount of debt that is rolled over is determined by the optimal debt arrangement. They also argue that sufficiently large shocks result in a sudden stop as all lenders refuse to roll over.

2.3 Machine Learning in Finance Markets

Artificial Intelligence (AI) techniques are being increasingly deployed in finance, in areas such as asset management, algorithmic trading, credit underwriting, or blockchain-based finance, enabled by the abundance of available data and by affordable computing capacity. Machine learning (ML) models use big data to learn and improve predictability and performance automatically through experience and data [17]. According to the report of the Organisation for Economic Co-operation and Development (OECD, 2021)[17] the deployment of ML in finance could amplify risks already present in financial markets given their ability to learn and dynamically adjust to evolving conditions in a fully autonomous way, and give rise to new overriding challenges and risks. Therefore, using ML to find rollover risk already present in the repo market can give more insight for policy making. Furthermore, such an existing risk as rollover risk can be associated with the inadequate use of data or the use of poor quality data that could allow for biases and discriminatory results.

Traditionally the finance sector has been using descriptive analytics to produce reports about its customers for any given services category, and this has had a big impact on the top management while making various decisions. Banking is an area where the utilization of machine learning procedures is applied to rich informational databases and can help battle financial frauds, save time and cash for clients, and computerize back-office capacities. The Bank for International Settlements (BIS) published a report on how to manage the size and complexity of data by central banks by proposing a highly automated validation workflow that outperforms traditional approaches and is suitable for a large volume of financial market time series, based on machine learning algorithms [18]. In the paper, they mention that while financial market time series data are key inputs to important policy decisions by central banks, little research focuses specifically on data validation processes for them.

The research paper of N. Garvin (2018) [19] discusses the identification of the Australian repo market microstructure using data from securities transactions. It explores how analyzing repo market activity at the transactional level can provide valuable insights into the market's underlying structure and dynamics using the Furfine algorithm. The study helps in understanding how participants engage in these short-term funding transactions. He mentions that in 2015 the market shifted toward overnight transactions

Lastly, the working paper of C. Brand, et al. (2019) [20] focuses on repo specialness premia, using ISIN-specific transaction-by-transaction data of one-day maturity repos. They document a gradual shift from cash- to securities-driven transactions in the euro area repo market. The research applies panel econometric and data mining approaches to show that by analyzing the role of financial stress and the availability of safe assets, the study provides insights into the factors shaping the repo market landscape in the region.

Overall, the use of machine learning in financial markets and central banks is becoming increasingly prominent due to its potential to analyze vast amounts of data and identify complex patterns. Nonetheless, research articles in the repo market domain still lack in comparison. The repo market's relative complexity and the need for (public) granular transactional-level data may pose challenges in obtaining comprehensive datasets for research purposes, hindering the development of machine learning applications in this specific area.

3: Data

In this chapter, the data is described to get a better understanding of the repo transactions and their attributes. In section 3.1 a thorough description of the data is given. In sections 3.2 to 3.6 data analysis is performed to understand the dataset and start working towards a clean dataset to work with for the machine learning applications.

Due to confidentiality reasons only the data up until July 2022 could be used for this thesis.

3.1 Data Description

The dataset used in this thesis is the Securities Financing Transaction Regulation (SFTR) which refers to transactions that are related to, inter alia, the build-up of leverage, procyclicality, liquidity, and maturity transformation, and interconnectedness in the financial markets [21]. Securities financing transactions (SFTs) include a repurchase transaction, securities or commodities lending, securities or commodities borrowing, buy-sell back transaction or sell-buy back transaction, and, margin lending transactions. This thesis project only uses the data on the repo transactions.

The Securities Financing Transactions Data Store (SFTDS) enables the collection of data on securities financing transactions (SFTs) as reported under the SFTR by the ECB, together with participating NCBs (national central banks), among whom DNB. The data processing steps also include the enrichment of the raw SFT data with data from relevant sources (namely SDW, CSDB, RIAD, and GLEIF)..

The SFTR data contains daily updates on every repo transaction with at least one trading party operating in the European Union (EU) via a branch, subsidiary, or parent company. The database operates on a daily updating scheme, where both the daily modifications (flow) and end-of-day status (stock) are reported for each transaction by each EU counterparty from entry until maturity. Thus, open transactions between two EU counterparties enter twice each business day. Each transaction can be traced by a Unique Transaction Identifier (UTI) over time [22]. Moreover, a repo transaction is required to be reported on a business day. Parties should assume this excludes weekends and official national holidays in the country in which they are located [9].

The SFTR dataset is relatively new (starting mid-2020)[1], and allows for new insights into money markets on a transaction level. It provides details like the counterparty network (including systemic counterparties), the types of repos traded, and characteristics such as principal amounts and repo interest rates. Furthermore, the data is refreshed daily. Transactional data relates to the transactions of the organization (in this case SFTR transaction from the ECB) and includes data that is captured, for example, when a transaction starts or at which fixed interest rate the transaction is set on. Transaction data is typically much more volatile than master data, which does not change and does not need to be created with every transaction [23]. In Figure 3.1 an example is shown of how transactional data differs from master data and that it shows all relative information regarding (in this example) the purchase of retail products.

[1] MMSR is the older dataset and contains similar details on the principal amount, interest rate, maturity, and counterparty for each repo of less than one year executed by the 48 largest euro area banks. SFTR contains all repo transactions in the EU and can therefore complement the MMSR dataset since MMSR is a subset of SFTR. The average total daily principal amount lies around EUR 645 billion reported by MMSR and EUR 1 trillion reported by SFTR.

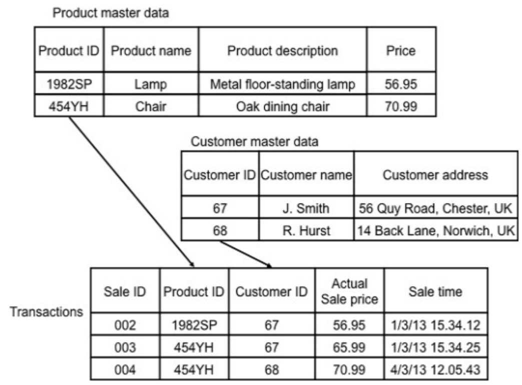


Figure 3.1: Example of transactional level data that is formed by master data. Source: Chapter 12 - Software Tools: Automated Methods for TIRM [23]

A full overview of all reported information can be found in the SFTR Reporting Guidelines published by the ESMA (2022) [2].

3.2 Data Collection

The SFTR data is collected by the ECB and regulated by the European Securities and Markets Authority (ESMA). The need to enhance the transparency of securities financing markets and thus of the financial system is responded to by ESMA reporting.

The ECB gathers data on repos through various channels, including direct reporting by market participants and data collection from central securities depositories. Market participants engaged in repo transactions are required to report their transactions to trade repositories authorized by ESMA under the European Market Infrastructure Regulation (EMIR). The ECB can access this data through ESMA and use it to monitor market developments and assess the impact of its monetary policy operations. In addition, the ECB collects data from central securities depositories (CSDs) that provide clearing and settlement services for repo transactions. CSDs are required to report information on the securities that are used as collateral in repo transactions, as well as information on the terms and conditions of the transactions. The ECB uses this data to monitor the liquidity

[2] More detailed information can be found on the website of ESMA Europe on SFTR reporting. Furthermore, an example repo can be found in ICMA SFTR sample reporting guidelines

and functioning of the euro area money markets. The ECB also conducts regular surveys of market participants to collect information on their repo transactions, including the types of collateral used, the maturity of the transactions, and the interest rates applied. This information is used to analyze market trends and inform the ECB's monetary policy decisions. Overall, the ECB employs a combination of direct reporting by market participants, data collection from CSDs, and surveys of market participants to gather data on repo transactions and monitor developments in the euro area money markets.

The SFTDS can only be accessed by staff-approved business functions of the respective member NCBs and the ECB. The daily information set was obtained from the DNB with collaboration through the ECB. After given permission to access the data the tables can be accessed through an SQL platform.

3.3 Data Cleaning

Even though the SFTR is a relatively new dataset it is immensely huge with a vast amount of information about daily repo transactions. This makes data exploration rather difficult. For each transaction, there are over 800 columns.

3.3.1 SQL Cleaning

Having over 800 features of information regarding repo transactions is overwhelming and excessive to do a clear analysis of rollovers. Furthermore, taking all these features into consideration for the analysis and machine learning applications would take too much computation power, and some features are considered duplicates of each other or unnecessary for this analysis. Therefore most of the relevant features are manually picked from the table by analyzing if they could be relevant for this research and other coworkers' projects already working from the ECB, Bundesbank, and the DNB on the SFTR dataset are analyzed for feature selection.

For a repo, it is possible to use multiple underlying securities as collateral for one given transaction. This causes the dataset to explode in rows, as for each collateral used in one transaction the dataset records them all separately. To deal with this the feature that loops over all exploded rows is set to only the first reported timestamp with all the collateral updates included. Moreover, additional information is added as features regarding the collateral such as the mean principal amount of the collateral combined and the maximum, minimum, and variance of the collateral details, etc. To confirm that there are no more duplicates or exploded rows, the data is also checked on its unique technical identifier (UTI). If there are rows with the same UTI that means it is a duplicate and therefore needs to be dropped. This also removes the transactions where both parties reported the transaction and leaves the first reported in the data. The next step is to correctly identify which date features need to be used to identify the maturity of a transaction. Agreeing upon a repo transaction involves multiple date features in the dataset, such as event date, business date, execution timestamp, reporting timestamp, start date, maturity date, and fixed maturity date. It is necessary to identify when a contract has a maturity of overnight/one business day to further predict/identify when a transaction is rolled over.

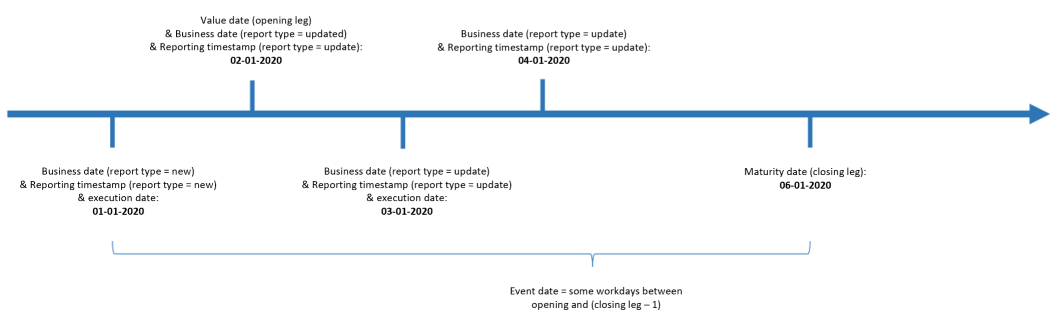


Figure 3.2: Date variables - An example SFTR data.

In Figure 3.2 an example is made to depict how one transaction has multiple date variables to analyze, note that not all transactions have to look like this example with collateralized updates. As shown in Figure 3.2 there are multiple dates to choose from when computing the maturity of a transaction, for this project, the opening leg date and closing leg date are the dates needed to compute the length of the contract. In this example that would be *06/01/2020 - 02/01/2020 = 4* days total and *2* working days (as the weekend days are excluded for the maturity of a contract). According to the financial stability department of the Deutsche Bundesbank, these are the best variables to choose when defining the maturity of a contract. With the chosen date variables the maturity of the transaction can be added as a feature for predicting repo rollovers. This also means that there are only term repos and no open repos. Where a term repo is a repo transaction that has a specified maturity date (usually the following day or week) and an open repo is a repo transaction where the dealer and the counterparty agree to the transaction without setting the maturity date [24]. The reason to only work with the term repos is that for a rollover the transaction is always renewed between the two parties, meaning that a maturity date is set every time the transaction is rolled over.

While analyzing the dataset the principal amount of the collateral (in euros or another currency), it seems that the transaction values are too large to analyze. Therefore most *Numerical features considering collateral amount* are changed into millions of euros (or another currency) as most transactions have values greater than one million euro. This also helps when using graphs, charts, or visual representations for clearer depiction and pattern recognition.

Given the size of the repo market in Europe to standardize the dataset, only the biggest European repo countries are taken, namely Germany, France, the Netherlands, and Belgium [25]. This indicates that all transactions in the cleaned dataset are from one of those countries but those transactions can be done with institutions outside of Europe, e.g. a German bank can have a transaction with an American financial institution (so those transactions are also included). This option is selected as it enhances the standardization of the data. Since this allows for a deeper understanding of the factors that influence the market in these specific countries (or just in general the big EU western market competitors). Moreover, including all countries can lead to a significant increase in complexity to interpret the results. Lastly, sometimes data availability or quality may vary across different countries. By excluding countries with limited or unreliable data, the overall quality and reliability of the analysis can be improved (this does not have to be the case for all excluded countries).

Then Central counterparty (CCP) information is joined on the involved parties to create information on whether a transaction is cleared. Repo contracts can have CCPs take on counterparty credit risk between parties to a transaction and provide clearing and settlement services for trades in foreign exchange, securities, options, and derivative contracts. It is necessary to identify whether a contract has been cleared as it can take on counterparty risk. In financial transactions, counterparty risk refers to the potential that one party in the transaction may default on its obligations. A CCP acts as a third party that interposes itself between the two original counterparties in a trade and takes on the counterparty risk by ensuring that both parties fulfill their obligations.

The data suffers from the way that some repos are reported in which way they are CCP-cleared. When party A clears a repo with party B, two trades are reported: A-CCP and CCP-B. But a meaningful measure of the repo market size should include only one trade (A-B), as the CCP is not trading so it is not really part of the market. Therefore, for cleared transactions all similar transactions are identified per UTI pattern and joined with all relevant information of the counterparty, e.g. the transaction A-CCP and B-CCP are merged with all necessary information such as country of company A and B. This also includes transactions with more CCP or other parties involved.

Next in the data cleaning steps, all the rows without a principal amount value are dropped. As that would mean no exchange of collateral happened or there was a mistake filling in the values from the transaction. Furthermore, only the transactions with an overnight maturity are selected (maturity of one business day including transactions from Friday to Monday). This already amounted to 95% of the table but it is still important to filter the maturity days as we assume that a rollover is always an overnight transaction (see more in Chapter 4.1). Lastly, the data is sorted by counterparty pair ids. Because a rollover needs to be a continuation of the original transaction that means that the counterparties stay the same for the rollover transaction. By sorting the table this way it can be easier to analyze if there are repo rollovers in the dataset and which features are the same or differ within a margin.

This results in a table of around 12 million rows and around 84 features from the original SFTR dataset. This Table is too large to download at once so it is split up into four separate files by date. Then for the analysis overall the dataset transactions between 2020-07-01 and 2022-07-01 are used, due to confidentiality reasons the data after July 2022 could not be implemented in this thesis.

3.3.2 Outliers

Since there are around 12 million rows in the data set used for analysis there are bound to be outliers. It is even more likely because the data points are all filled in manually by the reporting parties. This means that there are outliers from the miscoded variables filled in. For example, an easy mistake to make is to change the interest rate from 01% to 10%, which is obviously a big difference and not reasonable. Therefore, two options are used to remove the outliers.

The first method is to use the z-score. A z-score is defined as a measure of the divergence of different experimental observations from the most probable result, the mean [26]. In Equation 3.1 the z-score is expressed in terms of the number of standard deviations from the mean value. Where x_i is the experimental value, μ is the mean, and σ is the standard deviation.

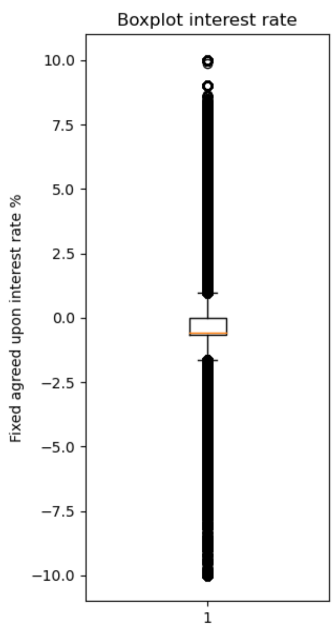
$$z = \frac{x_i - \mu}{\sigma} \tag{3.1}$$

The z-score provides a mechanism to determine the magnitude by which an observation diverges from the other observations of the dataset and if found to be large enough, the observation can be deemed an outlier [26]. Once the chosen columns for outlier detection are converted using the z-score, the center becomes zero and the z-score corresponding to each data point represents the distance from the center in terms of standard deviation. For example, a z-score of 2.5 indicates that the data point is 2.5 standard deviations away from the mean. The cutoff threshold chosen is 3, as that is the most common threshold and this threshold only takes the extreme outliers. Any z-score greater than 3 or less than -3 is considered an outlier. Using the z-score for the features *fixed agreed-upon interest rate*, the annualized rate on the principal amount of the repo transaction in accordance with the day count conventions, and *the principal amount of the underlying collateral* the outliers are filtered. These features are chosen as they are used in finding the repo rollovers in the dataset and one of the numerical features in the data.

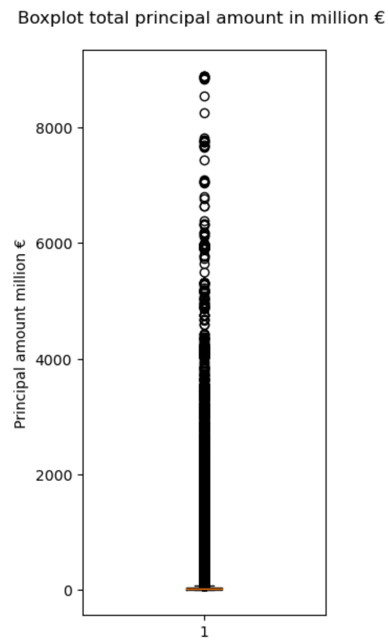
The second method is to uphold threshold values. This is when the *fixed interest rate* is above or below 10% (or 1000 basis points). Along with, the principal amount of the collateral at the value date with a threshold of 20000 million euros. Moreover, the principal amount cannot be a negative number. These thresholds are chosen as they are considered misreporting errors or abnormal transactions.

With these two methods **2.03%** of the dataset is considered an outlier of the around 12 million data points, the outlier points are dropped. It is not strange that there are many outliers in the dataset, as all transactions are manually filled in by the reporting party involved. This can cause misreporting. Moreover, a considerable amount of the transactions are so immense that there could be no possibility of that transaction being rolled over. The data quality of the SFTR table is in general not perfect. As there are still some missing values from specific dates and misreporting errors. This causes the dataset to have more outliers.

Nonetheless, since the dataset is big there could still be some outliers even after these two methods. The z-score method is applied first and then the threshold method is used. This results in the box plots Figure 3.3 below. In the box plots, there are many points outside the box for both the interest rate and principal amount features. Normally these values located outside are considered outliers. The length of the outliers provides information about the spread, it suggests a wider range of values, indicating greater variability in the dataset. Because the dataset is large, it is more likely to observe values extending beyond the box, even with outlier removal methods. As long as these values are not extreme, it is generally considered acceptable, especially when dealing with a diverse or extensive range like the interest rate and principal amount. Moreover, the principal amount of the transaction is in million of euros, and while a lot of the transactions have a value greater than one million euros the range differs greatly per transaction. This has other influencing factors such as the financial institution making the deal and the underlying securities traded.



(a) Boxplot fixed agreed-upon interest rate %



(b) Boxplot principal amount mill €

Figure 3.3: Boxplots of the interest rate (a) and principal amount (b) after outlier removal methods

3.4 Data Exploration

Over the period from 2020 to mid of 2022, the amount of total repo transactions per day is growing since the start of the dataset, this can be seen in Figure 3.4, where the moving average is the total number of transactions is the bolder line. Although there are still some dates that transactions are missing, this is most likely due to Christmas, Easter, or errors in the whole SFTR dataset. However, this also can be due to data cleaning errors, such as only using transactions that are overnight. Furthermore, the first few dates of the data start around the 17th of July instead of the beginning of July 2020. Overall, the average number of transactions grew from 20000 per day in 2020 to around 30000 transactions per day in 2022 in the biggest European repo countries with overnight transactions. This is in line with ECB stating that the secured money market is a growing financial market for central banks and other financial parties [27]. Furthermore, according to the ICMA public SFTR data [28] over the last two weeks of December 2020 new repos fell per day by 44%, which they conclude as seasonal behavior but probably also the anticipation of market disruption arising from Brexit[3]. There followed a gradual recovery in repo transactions in February 2021 and a seasonal dip at Easter (first week of April 2021), see drop 2021-04 in Figure 3.4. Then for the rest of 2021, the transactions stagnate a bit before rising again around mid-July. Again there is a collapse at the end-year, which was due to market intermediaries window-dressing their regulatory and other reports by shrinking their balance sheets [29]. The overshooting in January in reporting (as much as 35000 trades per day) suggests that parties had a lot of catching up to do after Christmas and that a lot of SFTR processing is still manual. Around February 2022 there is again a small drop immediately followed by a rising increase, this is approximately the start of the Ukraine war[4]. This could be due to a delay in reporting or market stress. At Easter 2022 the EU market took a break, which resulted in a big drop of no reports as can be seen at the big drop in Figure 3.4 April 2022.

[3] On 24 December 2020 the EU-UK Trade and Cooperation Agreement was announced, allowing goods to be sold between the two markets without tariffs or quotas. At 23:00 on 31 December 2020, the transition period ended, and the UK formally completed its separation from the EU.

[4] In February 2022, Russia recognized the Donetsk and Luhansk region as independent states and announced a special military operation in Ukraine, subsequently invading the region.

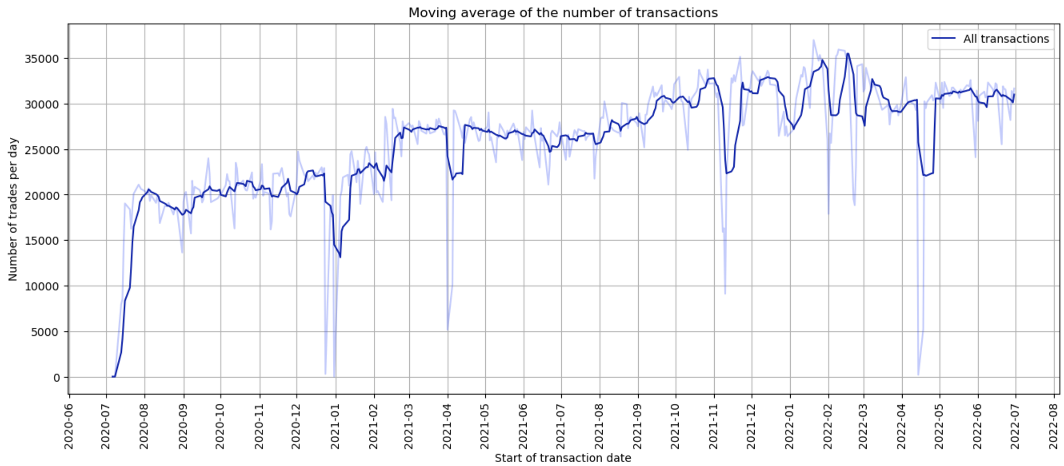


Figure 3.4: Total number of transactions in the cleaned dataset.

Next, the day on which a transaction is started seems to be equally distributed from Monday up to Thursday, see Table 3.1. However, transactions that start on Friday do not occur as much. It appears that transactions from Friday till Monday (as that is officially a maturity of one business day) do not occur as often, see Table 3.1. SFTR data guidelines require parties to report the conclusion, modification, or termination of repos no later than the working day following the conclusion, modification, or termination of the transaction [9]. If parties transact a repo during a weekend or on a public holiday in the country in which one of the parties is located, both parties should both treat that date as a working day. Trading on Friday can be more difficult since the stock market closes that day and it is also the day on which options tend to expire. Della Vigna and Pollet (2009) [30] argue that on Fridays, there is a greater probability of investors being busy with the weekend and, consequently, they do not pay the same amount of attention to investing as on other weekdays. On weekends, when the regular trading hours of financial markets are closed, there can be a decrease in activity in the secured money market. This is because financial institutions, such as banks, may have reduced operational hours or limited access to funding sources during non-business days. Additionally, some financial institutions may prefer to limit their exposure or manage their liquidity positions more conservatively over the weekend, as there may be increased uncertainty or limited market opportunities during this period. This cautious approach can lead to a decrease in transaction frequency in the secured money market from Friday to Monday. Additionally, according to Eurex (a central clearing party in the European repo market), the trading hours for overnight transactions need to be settled before 16:00 Frankfurt am Main time [31]. The American Eastern market opens at 15:30 Frankfurt am Main time, meaning there is only a small window of opportunity to trade with the United States. Lastly, in the research paper of A. Brassil et al. [6] the prediction of rollovers in the interbank overnight cash market do not involve loans on Friday.

Day of the week	**Percentage of transactions**	**Average total daily principal amount**	**Mean daily principal amount**
Monday	22.47%	€657.02 bill	€25.99 mill
Tuesday	23.43%	€695.81 bill	€26.52 mill
Wednesday	23.43%	€689.15 bill	€27.78 mill
Thursday	23.33%	€680.85 bill	€26.41 mill
Friday	7.34%	€319.82 bill	€28.27 mill

Table 3.1: Day of the week that a repo transaction takes place where the maturity is overnight (one business day). The table includes the percentage that a transaction starts on a given weekday. The average total daily principal amount per weekday is in billion EUR. The mean daily principal amount of a transaction per weekday.

In Table 3.1 the total daily principal amount on average is also depicted for every weekday. Again from Monday to Thursday the average lies close to each other of the total principal amount of that day. Though the difference of a billion EUR is big so it could be said that on Monday trader's principal amount lies lower in general. This is also the case on Monday's mean daily principal amount is around €26 million per transaction. The average of Friday's transactions is a bit higher (€28 million per transaction). This explains that the average total daily amount is actually quite high at €320 billion, which means that the transactions on Friday are higher on the principal amount.

Figure 3.5 is the average fixed agreed-upon interest rate for all repo transactions. Repo interest rates strongly follow the ECB deposit facility rate, but are also subject to monetary policy operations, collateral demand and supply, and country-specific shocks. Repo interest rates show clear end-of-quarter seasonality effects, see Figure 3.5 in December 2020. However, during almost the whole year 2021 the interest rates are way above the average ECB interest rate (average -0.216%). An interesting observation in the Financial Stress Index (see Appendix Figure A.10 and section 3.5) is that the whole year 2021 is considered as below average. This could be the reason why the average agreed-upon interest rate is much higher in 2021. The euro repo rate traded below but closer to the deposit facility rate (DFR) until mid-2020 when secured rates began to fall back below the DFR as a result of the Eurosystem's monetary policy actions. The deposit rate is the interest rate at which banks can make overnight deposits with the central bank. This rate is always lower than the main refinancing rate. Interest rates in the repo market tend to be higher when there is increased demand for short-term funds, leading to a shortage of available cash. This can occur during periods of financial stress, market turbulence, or when market participants need to secure funds urgently. In such situations, borrowers are willing to pay higher interest rates (repo rates) to obtain the needed funds.

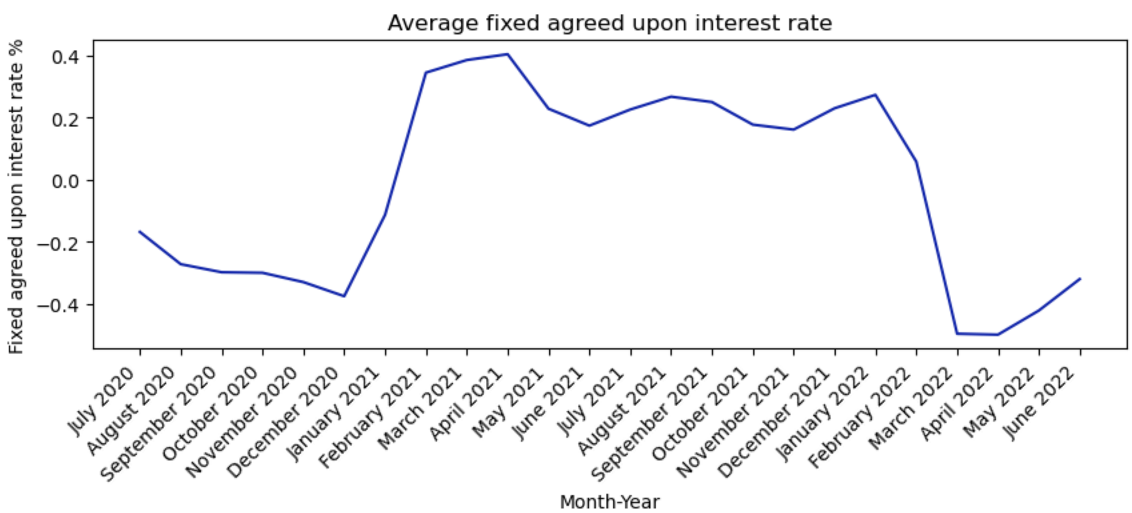


Figure 3.5: Average fixed agreed-upon interest rate for all repo transactions.

To know the distributions of the data a Shapiro-Wilk test has been conducted on the *fixed agreed-upon interest rate* and *the principal amount*. The Shapiro-Wilk test tests the null hypothesis that a sample came from a normal distribution. This resulted in a p-value of *0.0* and *0.0* for the fixed agreed-upon interest rate and principal amount respectively. This means that both p-values are less than chosen alpha level and therefore the null hypothesis is rejected that the data is normally distributed.

3.5 Publicly Available Data

Financial markets are complex systems that are influenced by a multitude of factors, including macroeconomic indicators such as GDP, inflation, and interest rates. As such, incorporating such factors into financial market analyses can provide valuable insights into market behavior and trends. Therefore some publicly available macroeconomic data was added to the SFTR dataset to explore the impact of these factors on repo rollover outcomes. This includes data from the ECB [32], the Office of Financial Research (OFR) [33] and the Federal Reserve Bank of St. Louis [34]. The included features are described in Table 3.2.

Metric	Definition
ESTR transactions	Euro short-term rate - Number of transactions Unit = not applicable
ESTR volume	Euro short-term rate - Total volume Unit = millions of euros
SOFR	USD Federal Reserve Secured Overnight Financing Rate Unit = percent per annum
Bank interest rate	Bank interest rate - loans to corporations of over EUR 1M with a floating and an IRF period of up to 3 months Unit = percent per annum
HICP	Harmonised index of consumer prices - overall index Euro area (changing composition) inflation rate Unit = percentage change
GDP	Global Economic Policy Uncertainty Index: Current Price Adjusted GDP. GDP-weighted average of national EPU indices for 20 countries Unit = index
FSI	OFR Financial Stress Index is a daily market-based snapshot of stress in global financial markets. Unit = index

Table 3.2: Descriptive analysis of the public data used besides the SFTR dataset.

The euro short-term rate (€STR) is a reference interest rate that reflects the overnight borrowing costs of banks in the euro area. It is calculated and published by the ECB based on actual transactions in the euro unsecured overnight interbank market. The €STR is used as a benchmark for a variety of financial transactions, including derivatives, loans, and securities. Both the total volume and the number of transactions are included (see Table 3.2).

The Secured Overnight Financing Rate (SOFR) is also included, it is a benchmark interest rate that reflects the cost of borrowing cash overnight using US Treasury securities as collateral. It is based on actual transactions in the US Treasury repo market.

The bank interest rate of the ECB is also included to compare with the fixed interest rate from the SFTR data. The Governing Council of the ECB sets the key interest rates for the euro area: the interest rate on the main refinancing operations, the rate on the marginal lending facility, and the rate on the deposit facility. These official interest rates form the basis for interest rates on the money market [35].

Then an important macroeconomic factor is the Harmonized Index of Consumer Prices (HICP), which is a measure of inflation in the EU. It tracks the change in prices of a basket of goods and services consumed by households across the EU. The HICP is used by the ECB to guide monetary policy and assess price stability in the euro area.

Moreover, the Gross Domestic Product (GDP) (weighted average for 20 countries around the world) is included. GDP is a measure of the economic output of a country. It represents the total market value of all goods and services produced within a country's borders. It is often used as an indicator of a country's economic health and growth.

Lastly, the Office Financial Research (OFR) Financial Stress Index (FSI) is added. The OFR FSI is constructed from 33 financial market variables, such as yield spreads, valuation measures, and interest rates. The OFR FSI is positive when stress levels are above average, and negative when stress levels are below average [33].

All these features described above and in Table 3.2 are then matched with the dates from the SFTR dataset and added as features. In Appendix A.4 are the Figures that show how these features behave during a longer time period.

3.6 Overview Summary

In Table 3.3 a short description is given about the used attributes in the available data. Each metric can have multiple different columns regarding additional information, for example, *Other parties involved* contains features such as whether a broker, CCP, or tri-party agent is involved, if the transaction has been cleared by a CCP, what the LEI of those parties are, from what country the respective party operates from, etc. The number of features used to identify the rollover transactions are limited to 84 total.

Metric	**Definition**
UTI	Unique Technical Identifier for each transaction
Date	Dates when a transaction is agreed upon, start, maturity date
Involved party id	Reporting and Other counterparty ID and country of origin
Counterparty Information	Information regarding the associated parties
Collateral	Information regarding the used collateral. - Type of collateral - Principal Amount in Euros or Dollars - Rate in Euros or Dollars - Market value and Nominal value - Haircut or Margin - Price per unit in respect of the collateral component - Risk of the security used as collateral - Rating of the collateral (SNP, Fitch) - Number of collateral updates - Moving average statistics (mean, max, min variance, etc.) - etc.
Fixed Interest Rate	Annualised interest rate on the principal amount of the repo transaction
Other parties involved	Parties involved in the transaction: CCP, tri-party agent, broker, etc.
Transaction information	Is the repo cleared? What type of transaction is agreed upon? etc.
LEI	Key reference information of legal entities
Data quality	Data quality score based on ESMA quality checks
Direction	Which direction the transaction is traded to

Table 3.3: Summary of the attributes in the dataset available

In the Table below 3.4 the descriptive statistics of some of the numerical features are depicted. It seems that the time period July 2020 to July 2022 has an average agreed fixed-upon interest rate of almost exactly zero percent. Furthermore, the rating of collateral is made numerical instead of normal credit rating such as *BB* or *A+*. The better the rating the higher the collateral rating is with a maximum of 10 (AAA) and a minimum of 1 (C)[5]. If none of the parties rated the underlying securities and the collateral was not a government bond the rating is set to zero. The average principal amount for all transactions is over €26 million, which indicates how much money is in the repo market. Most transactions have at least two collateral updates, even though all transactions are overnight. Lastly, 60% of all transactions are cleared by a CCP, from the ECB money market study of 2020 [27] around 70% of secured transactions were centrally cleared during

[5]On this page are credit ratings of the three rating agencies Moodys, S&P, and Fitch.

the period 2018-2020. Hence, this could mean that over the years clearing dropped.

Table 3.4 shows some descriptive numbers of the main features used. In Table 3.5 is a more detailed summary statistics.

	mean	std	min	median	max
Interest rate (%)	0.00	1.69	-10	-0.58	10
Collateral rating	7.86	1.14	0	9	10
Principal amount (mill €)	26.24	67.34	0	8.22	8899.82
Collateral updates	1.96	0.92	1	2	308
Cleared transactions (binary)	0.60	0.49	0	1	1

Table 3.4: Descriptive statistics for numeric features over the whole table

Table 3.5 shows a in-depth analysis of the data features per half year. It is noticeable that the average principal amount is higher in the first two quarters of 2021 and 2022 and higher during the last two quarters. Furthermore, on average 55% of the transactions are reported on a European trading venue while around 41% Over The counter trading (OTC) is. OTC is the process of trading securities via a broker-dealer network as opposed to a centralized exchange. Moreover, the transactions where both parties are located in the Eurosystem averages around 43%, meaning that the remaining 57% are transactions with non-ECB members, such as Japan or the United States.

	Mean second half 2020	**Mean first half 2021**	**Mean second half 2021**	**Mean first half 2022**
Repo interest rate	-0.30%	0.25%	0.22%	-0.22%
Average principal amount	€29.88 mill	€23.81 mill	€28.10 mill	€23.96 mill
Cleared transactions	65.29%	54.80%	59.40%	62.06%
Repo on EU trading venue	58.13%	51.68%	55.30%	56.31%
Repo on non-EU trading venue	7.16%	3.19%	3.56%	4.98%
Repo OTC trading	34.71%	44.53%	40.60%	45.72%
EU-EU repos	41.12%	43.24%	45.32%	45.12%
EU-nonEU repos	58.88%	56.76%	54.68%	54.88%
Market value	€54.69 mill	€34.77 mill	€57.17 mill	€73.27 mill
Nominal value	€23.82 mill	€21.46 mill	€26.10 mill	€26.53 mill
Collateral updates	2.20	1.97	1.80	1.95
Total number of repo trades	2.27 mill	2.84 mill	3.46 mill	3.41 mill

Table 3.5: Summarized table of the data over the time period used for analysis.

4: Methodology

The procedures for developing classification models to predict repo rollovers will be covered in this section. First, the acquisition of the dependent variable will be discussed. The second subsection describes the pre-processing methods for the data. Finally, the models that were employed will then be covered with the conducted experiments and the evaluation measures that were used to validate the results.

4.1 Programming a Rollover

The dependent variable for the predictions is the rollovers, e.g. the renewal of an existing transaction. A repo rollover transaction has several characteristics, including:

1. Continuation of the original transaction: A repo rollover transaction is a continuation of the original repo agreement between the parties, with the same securities, collateral, and terms.

2. Extension of the maturity date: The repo rollover extends the maturity date of the original repo, typically by one day or more, depending on the agreement between the parties.

3. Maintenance of the collateral: The collateral used in the original repo transaction is typically maintained throughout the rollover period, with adjustments made as necessary to reflect changes in market values or other factors.

4. Agreement on rollover terms: The terms of the rollover, including interest rates,

collateral requirements, and other conditions, are negotiated between the parties and documented in a new agreement.

5. Continuation of interest payments: Interest payments on the original repo continue during the rollover period, with adjustments made as necessary to reflect changes in interest rates or other factors.

6. Flexibility for the parties: The parties may choose to roll over the repo transaction multiple times, depending on their needs and preferences.

With the elements above in mind, conditions can be made to identify the repo rollover transaction in the dataset.

In this report, the prediction of a repo rollover is indicated by a *0* for a new transaction and *1* if the transaction is part of a rollover, from the first transaction that is part of the rollover to the last transaction. Therefore, this is a binary classification problem. To predict this classification, machine learning algorithms are used. A binary classification algorithm is a supervised learning algorithm that categorizes new observations into two classes.

4.1.1 Preparation Python

The first step to work with the data is to combine the cleaned SFTR table with the additional macroeconomic public data (see Section 3.5). This is done by merging the date of the transaction with the date of the macroeconomic factors. For example, if transaction starts on *26-01-2022* and all the macroeconomic factors on that day are implemented as a feature, e.g. the value of the ESTR rate was on that day -0.576 (see Appendix A.4 for visualizations on the public available data).

Secondly, the z-score and threshold methods are implemented to remove the outliers in the dataset. Then the dataframe columns are checked for missing values. If there are missing values for the *fixed interest rate*, *principal amount*, or *maturing date* the transaction is removed. For a lot of features, no data was filled in, as all transactions are different and can therefore differ in how many participants are included, where the transactions were reported, etc. Because of this many features have *missing values*, as they do not apply most of the time. According to Centraal Bureau Statistiek (CBS), it is recommended not to use data with more than 50% missing counts. But it depends on the feature if the data is actually missing or not applicable. For example, the feature if a transaction is cleared by a CCP has a NaN value that just means that there was no CCP involved. While the feature for collateral rating by Moody's [1] can be missing, as the reporting or counter party chose not to use that specific institution, and therefore the value is actually missing. For the other features missing values are filled with zero. According to the book of T. Chollet [36] it is generally safe to input missing values as 0, with the condition that 0 is not already a meaningful value. The model will learn from exposure to the data that the value 0 means missing data and will start ignoring the value.

Next, similarly to the research paper of the Reserve Bank of Australia [6] the dataframe is looped through on the conditions that a transaction has the same counterparties, has the same type of underlying securities, is the next day, and has a comparable interest rate and principal amount. More features are used to identify the rollover transactions, although the most important ones are mentioned. This is done with all the 84 features to identify other characteristics of a possible rollover transaction. If there are transactions that are rolled with multiple transactions the model chooses the transaction closest to Financial Stability Board (FSB) principal amount, as that feature is almost never rounded. Then the rollover transactions are matched on the whole roll period, if there are no transactions from Friday to Monday the model looks, after the Thursday to Friday transaction, for a possible renewal that starts on Monday to Tuesday, and is considered as a continuation of the rollover period. For the continuation to the prediction models a new subset of features is chosen. Considering that the data is large and some features contain the same information most features can be dropped. This results in 34 selected features to implement in the Machine Learning models.

[1] The Big Three credit rating agencies are S&P Global Ratings (S&P), Moody's, and Fitch Group.

4.2 Machine Learning Classifiers

A classifier is a function that takes the values of various *features* (independent variables or predictors, in regression) in an *example* (the set of independent variable values) and predicts the *class* that that example belongs to (the dependent variable) [37]. More formally, given an example x, the classifier is a function f that predicts the label $\hat{y} = f(x)$.

In Figure 4.1 an example of how a Machine Learning classifier works is depicted. Splitting the dataset into training and test data, the classifier learns from the training set, whose labels it can see, and uses it to predict labels for the test set, whose labels it cannot see. The predicted labels are then compared to the true labels and the evaluation metrics chosen by the classifier can be computed.

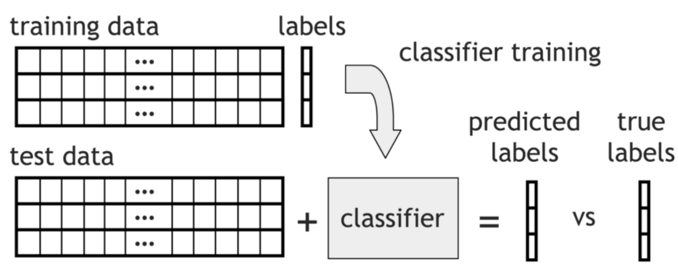


Figure 4.1: An example of how a classifier learns from the training set. Source: Pereira et al. (2009)[37]

The data is shuffled to introduce an element of randomness and split into the train and test set in order to remove bias from the data. There is no known optimal split percentage for the train and test set. But, if there is less training data, the machine learning model will show high variance in training. A common choice is the 80/20 train-test split is a common split in machine learning models. The 80/20 train-test split is generally considered good for large datasets because it strikes a balance between having enough data for training (80%) to build a robust and accurate model and having sufficient data for testing (20%) to assess the model's generalization performance effectively. For all machine learning models the 80/20 train-test split will be used. For predicting the repo rollover transaction the supervised learning algorithms are used. In supervised learning, both the input x and the output y are available, and the goal is to obtain an optimal predictive model function f. The goal is for the algorithm to learn a mapping between inputs and outputs, enabling it to make accurate predictions on new, unseen data. The main advantage of supervised learning, such as classifiers, is that it can provide precise predictions when given labeled data and is well suited for tasks like classification and regression. However, it requires labeled data, which can be costly and time-consuming to obtain, while unsupervised learning doesn't rely on labeled data, but it may produce less interpretable results and is more challenging to evaluate in the absence of ground truth.

4.2.1 Logistic Regression

The first model chosen to work with is a Logistic Regression model (LR). It is a classification technique that uses a logistic function to model the dependent variable in dichotomous nature (either a rollover or not). This classifier will be used as a benchmark model to compare the other Machine Learning classifiers. By using a benchmark model it allows for a quick assessment of performance and can provide a baseline to compare against more sophisticated models, helping to determine if the additional complexity of advanced algorithms is justified for a particular task. Logistic Regression is a simple and interpretable model that provides a good baseline for comparison. It is easy to implement and understand, making it a suitable starting point for modeling. The implementation of Logistic Regression is fitted for a binary case, where the target feature y_i takes values in the set $0, 1$ for data point i [38]. Once fitted, the probability estimates predict the probability of the positive class $P(y_i = 1|X_i)$ as in Equation 4.1.

$$\hat{p}(X_i) = expit(X_i w + w_0) = \frac{1}{1 + exp(-X_i w - w_0)} \tag{4.1}$$

As an optimization problem, binary class logistic regression with regularization with $r(w)$ minimizes the following cost functions:

$$min_w C \sum_{i=1}^{n} (-y_i log(\hat{p}(X_i)) - (1 - y_i) log(1 - \hat{p}(X_i))) + r(w) \tag{4.2}$$

In Logistic Regression, the cost function can be different, depending on the regularization method for penalization. The three main types of regularization are: *L1-norm*, *L2-norm*, and *elastic-net*. The coefficient C is another essential hyper-parameter that determines the regularization strength, smaller values specify stronger regularization and a high value tells the model to give high weight to the training data. Lastly, the *solver* type, representing the optimization algorithm type, can be set to *newton-cg*, *lbfgs*, *liblinear*, *sag* or *saga*[39]. The solvers *sag* and *saga* are both known to be suitable for large datasets. While *lbfgs* performs relatively well overall, *newton-cg* is computationally expensive because of the Hessian matrix used and *liblinear* works well with a high dimension dataset solving large-scale classification problems.

Some of the advantages of using Logistic Regression for classification are that it is easy to implement and interpret, it makes no assumptions about distributions of classes, it is usually a fast classifier and it gives a good result for many simple data sets. However, it can be prone to overfitting on high dimensional data, it assumes linearity between the dependent and independent variables and it fails to capture complex relationships.

4.2.2 Gaussian Naive Bayes

The Naive Bayes (NB) classifier is based on the Bayes theorem and is a useful classifier that is used widely in many applications such as text categorization, document judgment, and data stream classification [40]. Gaussian NB assumes that each feature has an independent capacity of predicting the output variable and follows a Gaussian (normal) distribution. According to the paper of Ali Jahromi and Mohammed Taheri (2017)[40] the input for the NB classifier, assumes the vector $(x_1, ..., x_n)$ representing the n attributes of the instance x. Let c represent the class label of the instance x. The probability of observing x given the class label c can be computed by Equation 4.3.

$$p(x_1, ..., x_n|c) = \prod_{i=1}^{n} p(x_i|c) \tag{4.3}$$

In order to predict the label of instance x, the probability of instance x in each class label is computed. The class with the maximum probability is identified as the class label of the instance x. Equation 4.4 defines the label estimation process of instance x.

$$C(x)_{NB} = \underbrace{argmax}_{c} p(c) \prod_{i=1}^{n} p(x_i|c) \tag{4.4}$$

Gaussian NB classification is a method that assumes of having a Gaussian distribution on attribute values given the class label. In Equation 4.5 the Gaussian distribution is computed, also known as the normal distribution. Where, $\mu_{c,i}$ is the mean and $\sigma^2_{c,i}$ the variance, given the class label c on the i^{th} attribute.

$$p(x_i|c) = \frac{1}{\sqrt{2\pi\sigma^2_{c,i}}} exp^{-\frac{(x_i-\mu_{c,i})^2}{2\sigma^2_{c,i}}} \tag{4.5}$$

The only parameter for the Gaussian NB classifier, in the python package Scikit-learn (Sklearn), *var smoothing* introduces a small amount of variance to the data's feature values, ensuring numerical stability and avoiding division by zero issues when computing probabilities for features with zero variance.

The main advantage of the NB algorithm is that it is a simple yet powerful algorithm and fast. However, NB assumes that all features are independent or unrelated, so it cannot learn the relationship between features. Naive Bayes can suffer from over-sensitivity to irrelevant features. If two or more features are highly correlated, they receive too much weight in the final decision as to which class a transaction belongs [41]. Furthermore, the assumption of independence among child nodes is clearly almost always wrong and for this reason, Naive Bayes classifiers are usually less accurate than other more sophisticated learning algorithms [42].

4.2.3 Decision Trees

Decision Trees (DT) are powerful, efficient, and popular approaches in data mining and knowledge discovery [43]. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piecewise constant approximation [38]. A Decision Tree model is a flowchart-like structure, where each internal node represents a test on an attribute, each branch denotes the outcome of the attribute test and each leaf node denotes the class label. The tree achieves classification by splitting the branches of a tree where each split represents a test on data attribute [43].

The function most often used to measure the quality of a split is either *Gini impurity* or *entropy*. Entropy is defined as the sum of the probability of each label times the log probability of that same label [44]. In Equation 4.6 the formula for entropy is depicted, where p_i is the probability that an arbitrary tuple belongs to class label c_i and L is the learning sample.

$$Entropy(L) = -\sum_{i=1}^{j} p_i log_2(p_i) \tag{4.6}$$

The function of the Gini index determines the purity of a specific class after splitting along a particular attribute. The best split increases the purity of the sets resulting from the split. If the learning sample L is a dataset with j different class labels. Gini is defined as in Equation 4.7, where p_i is the relative frequency in class i in L.

$$Gini(L) = 1 - \sum_{i=1}^{j} p_i^2 \tag{4.7}$$

The other important parameters in the DT model that can be tuned are the maximum depth, which specifies the maximum number of levels in the tree, limiting the number of splits and controlling the complexity of the model. Along with the parameter that specifies the minimum number of samples required to be at a leaf node, controlling the minimum size of leaf nodes and preventing the tree from creating nodes with very few instances, thus reducing overfitting. The advantages of using a Decision Tree are that it is simple to understand and interpret, it requires little data preparation and it performs well even if the assumptions are somewhat violated by the true model. The disadvantages include that learners can create over-complex trees that do not generalize the data well (overfitting) and predictions are neither smooth nor continuous.

4.2.4 Random Forest

Random Forest (RF) is an algorithm that combines the predictions from many individual randomized decision trees [45]. It uses a large number of decision trees in order to overcome the weakness of a single Decision Tree [46]. where each node the split is selected at random from among k best splits. For the k^{th} tree, a random vector Θ_k is generated, independent of the past random vectors Θ_1,...,Θ_{k-1} but with the same distribution; and a tree grown using the training set and Θ_k, resulting in a classifier $h(x, \Theta_k)$ where x is an input vector [47]. Moreover, a Random Forest is a classifier consisting of a collection of tree-structured classifiers $h(x, \Theta_k), k = 1, ...$ where Θ_k are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x. During prediction, the final class is determined by a majority vote among the individual trees, making the model less sensitive to noise and outliers in the data, and generally providing better generalization performance compared to a single decision tree.
Parameter tuning in the Python package scikitlearn for Decision Tree and Random Forest models is mostly similar as both algorithms share common hyperparameters. Such as the maximum depth, minimum samples of leaves, and the criterion.

An advantage of Random Forest is that because of the presence of multiple trees, the individual trees need not be pruned. Moreover, it runs efficiently on large datasets, can handle a great number of input variables, and gives estimates of what variables are important in the classification. A disadvantage is that by having many trees, the ability to visualize the trees is effectively lost.

4.2.5 XGBoost

XGBoost (Extreme Gradient Boosting) is an ensemble learning method that combines the predictions of multiple decision trees to create a more accurate and robust model. Tree boosting has been shown to give state-of-the-art results on many standard classification benchmarks. The XGBoost classifier has been widely recognized in a number of machine learning and data mining challenges as one of the best classifiers to use [48]. It works by iteratively adding decision trees to the ensemble, focusing on the mistakes made by the previous trees, and assigning a higher weight to misclassified instances. Additionally, XGBoost incorporates regularization techniques to prevent overfitting and uses a gradient-boosting framework to optimize a differentiable loss function, resulting in highly effective and efficient predictive models.

The model aims to minimize an objective function that consists of two parts: the loss function and a regularization term.

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{4.8}$$

Where

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda||w||^2 \tag{4.9}$$

In Equation 4.8 is the loss function that measures the prediction error of the model on the training data, where l is a differentiable convex loss function that measures the difference between the prediction $\hat{y}_i$ and the target y_i. The second term in Equation 4.8 penalizes the complexity of the model [48]. For Equation 4.9 the parameters γ and λ control the weight of each individual tree T in the model and its complexity.

The main hyperparameters for XGBoost are the number of estimators, maximum depth, and the learning rate [39]. The learning rate determines the step size at each iteration while moving towards a minimum of the loss function. The number of estimators parameter controls the number of boosting rounds or decision trees in the ensemble. The maximum depth of the individual decision trees is also implemented.

The advantage of using the XGBoost classifier is that it is known for its exceptional predictive accuracy. XGBoost is optimized for performance and computational efficiency. Furthermore, the model can handle various types of data and can be used for both classification and regression tasks. It provides valuable insights into feature importance as it ranks features based on their contribution to the model's performance, which aids in feature selection and understanding the dataset. However, achieving optimal performance with XGBoost often requires tuning a larger number of hyperparameters. Which can be time-consuming and computationally intensive. Furthermore, memory consumption can be relatively high, especially when dealing with large datasets or complex models. Lastly, it can lead to overfitting if the parameters are not tuned correctly.

4.2.6 Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) is a popular machine learning classifier widely used for training models, especially in binary classification tasks. It is a variant of the more general Gradient Descent algorithm but differs in its approach to updating model parameters. SGD is particularly beneficial for binary classification due to its efficiency in handling large datasets and its ability to adapt to changing data dynamics. SGD can be used instead of the standard Support Vector Machine (SVM), and with good parameter tuning it may yield similar or possibly even better results.

At its core, SGD optimizes the parameters of a machine learning model by iteratively updating them based on the gradients of the loss function with respect to these parameters. The loss function measures the discrepancy between the predicted outputs of the model and the actual labels. SGD computes this gradient and adjusts the model's parameters incrementally in a direction that minimizes the loss. Unlike regular Gradient Descent, which computes the gradients using the entire training dataset, SGD approximates the gradients by using a single training example or a small subset of examples (known as a mini-batch) at each iteration.

Given the training set with $(x_1, y_1), \ldots, (x_n, y_n)$ points, where $x_i \in R^m$ and $y_i \in [0, 1]$ the goal of the classifier is to learn a linear scoring function $f(x) = w^T x + b$ with $w \in R^m$ and intercept $b \in R$. To find the model parameters, the regularized training error is minimized and given in Equation 4.10, where L is the loss function that measures model fit and R is a regularization term (penalty) that penalizes model complexity; $\alpha \geq 0$ is a non-negative hyperparameter that controls the regularization strength [49].

$$E(w,b) = \frac{1}{n}\sum_{i=1}^{n} L(y_i, f(x_i)) + \alpha R(w) \tag{4.10}$$

The loss parameter specifies the loss function used to measure the model's performance during training. The different choices for L entail different classifiers: *Hinge* is equivalent to Support Vector Classification, *Squared Hinge* is like *Hinge* but then quadratically penalized, *Modified Huber* for robust regression, *log-loss* equivalent to Logistic Regression. Furthermore, the penalty parameter (regularization term R) includes the l2-norm, which is the standard regulazer for linear SVM models, and l1-norm which might bring sparsity to the model [49]. Lastly, the alpha parameter is also a main hyperparameter. The alpha is a constant that multiplies the regularization term. The higher the value, the stronger the regularization [49].

Some of the advantages of using SGD for binary classification is that SGD's is efficient to process large datasets. By updating parameters based on a subset of examples, it reduces the computational burden compared to using the entire dataset for each update. Moreover, SGD is well-suited for handling data with changing distribution or non-stationary patterns and is easy to implement. However, SGD also has some disadvantages. Namely, since SGD approximates gradients using a subset of examples, the computed gradients may be noisy, leading to less stable updates. This noise can introduce fluctuations during training and slow down convergence. Additionally, the choice of learning rate in SGD is critical. A learning rate that is too high can cause overshooting and divergence, while a learning rate that is too low can slow down convergence. Tuning the learning rate can be a non-trivial task.

4.3 Hyperparameter Optimization

Machine learning algorithms have been used widely in various applications and areas. To fit a Machine Learning model into different problems, its hyper-parameters must be tuned. Selecting the best hyperparameter configuration for Machine Learning models has a direct impact on the model's performance [39]. Tuning hyperparameters is considered a key component of building an effective ML model, especially for tree-based ML models and deep neural networks, which have many hyperparameters [39]. The generic *GridSearchCV* method from sklearn is implemented to detect the optimal hyperparameters using the grid search algorithm. Each hyperparameter value in the defined search space is evaluated by the program, with its performance evaluated using cross-validation [39]. When all the instances in the configuration space have been evaluated, the optimal hyperparameter combination in the defined search space with its performance score will be returned. Some of the advantages of using parameter tuning in ML can be that it improves the performance of the model. By finding the best combination of parameters, leading to improved accuracy, reduced overfitting, and enhanced generalization ability. Additionally, it reduces the human effort required to tune the parameters for large datasets or complex ML algorithms with a large number of hyperparameters [39].

The hyper-parameters of the machine learning models are studied in Python libraries, including sklearn and XGBoost. In the sections above of this Chapter 4 all of the main hyperparameters and their functions are already described. The specifics of the configuration space for the machine learning models are summarized in Table 4.1. It provides a summary of the machine learning classifiers, their main hyperparameters, and, suitable search space to work with. For the search space of the parameter of the Gaussian NB, *var smoothing*, a range is used from 0 to 0.001 in hundred steps (so the number of elements in the array) and an array of evenly spaced numbers on a logarithmic scale between 0 and -9, num is the number of elements in the array and set to 10.

ML algorithm	**Hyper-parameter**	**Search Space**
Logistic Regression	penalty	[l1, l2, elasticnet, None]
	C	[0.01, 1, 10, 50, 100, 150]
	solver	[lbfgs, liblinear, newton-cg, sag, saga]
Gaussian Naive Bayes	var smoothing	[np.linspace(0, 0.001, 100)] & [np.logspace(0, -9, num = 10)]
Decision Tree	max depth	[2, 3, 5, 10, 20, 30, 50]
	min samples leaf	[1, 5, 10, 20, 50, 100]
	criterion	[gini, entropy, log-loss]
Stochastic Gradient Descent	alpha	[0.0001, 0.001, 0.01, 0.1]
	loss	[hinge, log, squared hing, modified huber]
	penalty	[l1, l2, none]
Random Forest	max depth	[2, 4, 6, 10, 15]
	min samples leaf	[5, 10, 15, 20, 50, 100]
	criterion	[gini, entropy, log-loss]
Xgboost	learning rate	[0.01, 0.05, 0.1]
	max depth	[2, 4, 6]
	number of estimators	[50, 100, 200]

Table 4.1: Parameters to tune for the machine learning classifiers and the range to use.

4.4 Performance Evaluation

In this section, multiple performance evaluation methods are discussed using a diverse range to gain a deeper understanding of model performance and its relevance in predicting real-world scenarios.

4.4.1 Confusion Matrix

The confusion matrix is an evaluation method typically used in machine learning to visualize the behavior of models in supervised classification contexts. For a binary classification task, the confusion matrix is a *2 x 2* matrix that reports the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) (see Table 4.2). This evaluation method provides a more detailed analysis than simply observing the proportion of correct classification by providing a comprehensive view of the model's predictions and actual class labels.

		Actual	
		Negative	Positive
Predicted	Negative	TN	FP
	Positive	FN	TP

Table 4.2: Confusion matrix for binary classification and the corresponding array representation used. Where TN: *Normal* predicted as *normal*, FP: *normal* transaction predicted as rollover, FN: rollover transaction predicted as *normal* and, TP: rollover transaction predicted as rollover.

For this thesis, the normalized version of the confusion matrix is used. The normalized term means that each of the groupings is represented as having 1.00 samples. Thus, the sum of each row in a balanced and normalized confusion matrix is 1.00, because each row sum represents 100% of the elements in a particular topic, cluster, or class [50]. Moreover, it provides a clearer representation of the classifier's performance by showing class-wise accuracy instead of raw counts. This enables direct comparison of misclassification rates across different classes, facilitating better identification of problematic areas. Additionally, normalization accounts for class imbalances, ensuring a fair evaluation of the model's overall effectiveness.

4.4.2 Accuracy

One of the most used performance metrics is the ratio between the number of correctly classified samples and the overall number of samples, which is the accuracy. The computation method for the accuracy is given Equation 4.11, which represents the ratio between correctly predicted instances and all the instances in the dataset (where the worst value is 0 and the best value is 1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.11}$$

When the dataset is unbalanced (the number of samples in one class is larger than the number of samples in the other classes), accuracy cannot be considered a reliable measure, because it provides an over-optimistic estimation of the classifier ability on the majority class [51]. For example, if the dataset contains 100 transactions where only 15 are rollover transactions and the classifier predicts all transactions as not being a rollover that would still give an accuracy score of *0.85*, even though the prediction did not perform well.

4.4.3 Precision

Precision denotes the proportion of predicted positive cases that are correctly actual positives (see Equation 4.12). It calculates the proportion of correctly predicted true positives out of all transactions predicted as rollovers (true positives + false positives). The metric focuses on the accuracy of positive predictions, indicating how reliable the model is in identifying true positives. A higher precision value (1 is the highest, 0 lowest) indicates fewer false positives, making it useful in scenarios where false positives are costly or undesirable.

$$Precision = \frac{TP}{TP + FP} \tag{4.12}$$

4.4.4 Recall

Recall (also known as sensitivity) is used to measure the fraction of positive patterns that are correctly classified (see Equation 4.13). It calculates the proportion of correctly predicted true positives out of all actual positive transactions. The metric focuses on capturing the true positives, indicating how well the model identifies the positive instances in the dataset. A higher recall value (1 is the highest, 0 lowest) indicates fewer false negatives, making it valuable in situations where missing positive instances is a concern, such as medical diagnosis or fraud detection.

$$Recall = \frac{TP}{TP + FP} \tag{4.13}$$

4.4.5 F1-score

The F1-score is a combination of the precision and recall metric, providing a balanced measure of a classification model's performance. It is the harmonic mean of precision and recall (see Equation 4.14). The f1-score considers both the ability of the model to correctly identify positive instances (precision) and its ability to capture all positive cases (recall). It can be useful to strike a balance between precision and recall, especially in scenarios when a class imbalance exists or false positives and false negatives are equally important. A higher F1-score (1 is the highest, 0 lowest) indicates the better overall performance of the model.

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{4.14}$$

4.4.6 Macro Average

The macro average calculates the average performance metric (precision, recall, and F1-score) independently for each class. It computes the metric for each class and then takes the simple arithmetic mean across all classes. In other words, it treats each class equally, regardless of the number of instances. Macro averaging is useful when the classes need to have equal importance, regardless of their distribution or prevalence. If there are N classes the macro average can be computed as in Equation 4.15, where M_class represents the performance metric values for each class.

$$Macro\_average = \frac{(M\_class1 + M\_class2 + ... + M\_classN)}{N} \tag{4.15}$$

4.4.7 Weighted Average

The weighted average calculates the average performance metric by considering the number of instances for each class. It calculates the metric for each class and then takes the weighted mean, where the weights are determined by the support of each class. Classes with a larger number of instances contribute more to the weighted average. Weighted averaging can be useful when you want to account for class imbalance, as it gives more weight to classes with higher support, reflecting their impact on the overall performance. The weighted average can be calculated as in Equation 4.16, where the support for each class is denoted as S_class, the performance metric values as M_class, and the total number of instances as T.

$$Weighted\_average = \frac{S\_class1 \cdot M\_class1 + S\_class2 \cdot M\_class2 + ... + S\_classN \cdot M\_classN}{T} \tag{4.16}$$

4.4.8 Area Under the Curve (AUC)

The AUC (Area Under the Curve) is a performance metric commonly used in binary classification tasks, particularly in evaluating the performance of a classifier's receiver operating characteristic (ROC) curve. AUC represents the probability that a randomly chosen positive instance is ranked higher than a randomly chosen negative instance. It provides a measure of the classifier's ability to distinguish between positive and negative classes, with a higher AUC indicating better discrimination and performance. AUC ranges from 0 to 1, where an AUC of 0.5 suggests a random classifier, while an AUC of 1 represents a perfect classifier. ROC graphs are a very useful tool for visualizing and evaluating used classifiers [52]. An advantage of using the ROC is that they are able to provide a richer measure of classification performance than scalar measures such as accuracy. A basic example of a ROC graph from "An introduction to ROC analysis" [52] is given in Figure 4.2, where the calculation for the True positive and False positive Rate is given in Equation 4.17 and 4.18. From the results, the best ROC curve would be if the resulting classifier would be high to the left, as then the True positive rate is high and the False positive rate would be low. The best classifier in this example is D as seen in Figure 4.2.

$$TruePositiveRate = \frac{TP}{TP + FN} \tag{4.17}$$

$$FalsePositiveRate = \frac{FP}{FP + TP} \tag{4.18}$$

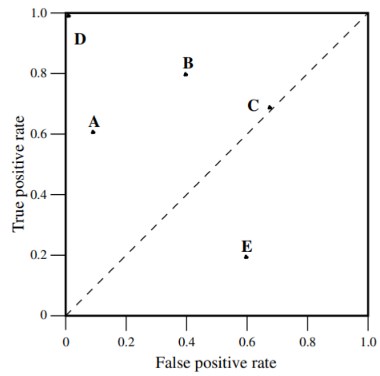


Figure 4.2: A basic ROC graph showing five discrete classifiers. Source: T. Fawcett (2005)[52]

For implementation in Python, the classifiers will be trained exactly the same, but the output method will be slightly different. For all the results the method *model.predict(x_test)* will be used as that returns the predicted class labels for the input. Because for binary classification, it returns either *0* or *1*, representing the predicted class for each sample in the training data. The method *model.predict_proba(x_test)* returns the predicted probabilities of each class for the input training data. For the binary classification, it returns a 2D array where each row represents a sample in *x_test*, and the columns represent the probabilities of belonging to the respective classes. To calculate the AUC-ROC the predicted probabilities are needed, not just the predicted class labels. This is because AUC-ROC requires information about the confidence or probability of the classifier's predictions, which is not available when using only the class labels.

4.4.9 Performance Measure Comparison

To evaluate the binary classification of repo rollover transactions and the confusion matrices, the best performance metric to use is the Macro average F1-score. Because the performance of both classes, regardless of size, is given equal importance. This then provides an overall assessment of the model's performance that is not influenced by class prevalence. Furthermore, the normalized True Positive scores of the confusion matrix will be used as it is a good measure to see how well the model predicts the rollover transactions. That means that the classification model does not miss the rollover trades so it does not miss rollover risk and liquidation risk. For this project, it is better to have a bit more transactions predicted as rollovers rather than missing them. Lastly, in the end, an AUC-ROC graph will be made with the predicted probabilities to easily visualize all the classification models and compare the results.

5: Results

In this chapter, the results of programming a rollover and the machine learning classifiers are conducted and evaluated. Finally, an analysis is provided to give an interpretation of these performances.

5.1 Rollover Statistics

In the European repo market, there is no indicator of whether a transaction has been rolled over by the two (or more) parties involved. Through the help of data cleaning, processing, and programming a first effort is done to find these transactions. First, the dataset was tidied and narrowed down to a more concise dataset to work with. Rollover transactions consist of 21% of the dataset where the average duration is 3.8 days. This can be seen as almost a whole workweek because the Friday transaction does not occur frequently. The maximum number of rolls occurring in the dataset is 35.

Figure 5.1 is the visualization of the moving average of the total number of rollovers and other transactions in the European repo market from the period of mid-2020 to mid-2022. The graph is shown with a moving average because, on some specific days, there were not a lot or no transactions at all, as is shown by the drops in Figure 5.1 by the actual transactions (the see-through lines). Besides those drops, there does not seem to be much variation through this period compared to the total number of transactions (see Figure 3.4). In Figure 5.2 are only the rollover transaction per day. Here there are no clear shocks or periods of non-rollover except for two days in November 2021. However, the non-rollover transaction also behaves in the same way. Furthermore, in the monthly meetings of the user dataset of the ECB, the dates with low spikes correspond to the missing values of the dataset due to late reporting. The average number of rollover transactions over the whole timeline lies around 5700 transactions per day. In Figure 5.2 are only the number of rollover transaction visualized over time. There are no great differences between rollover and non-rollover transactions apart from the total occurrence.

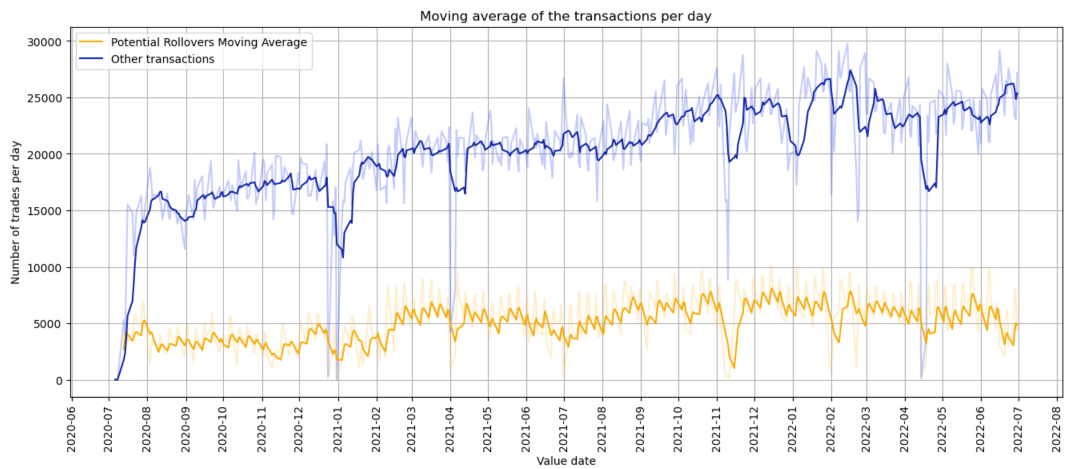


Figure 5.1: Moving average plot of the total number of repo rollover transactions (orange) and other repo transactions (blue) in Europe from 01-07-2020 till 01-07-2022.

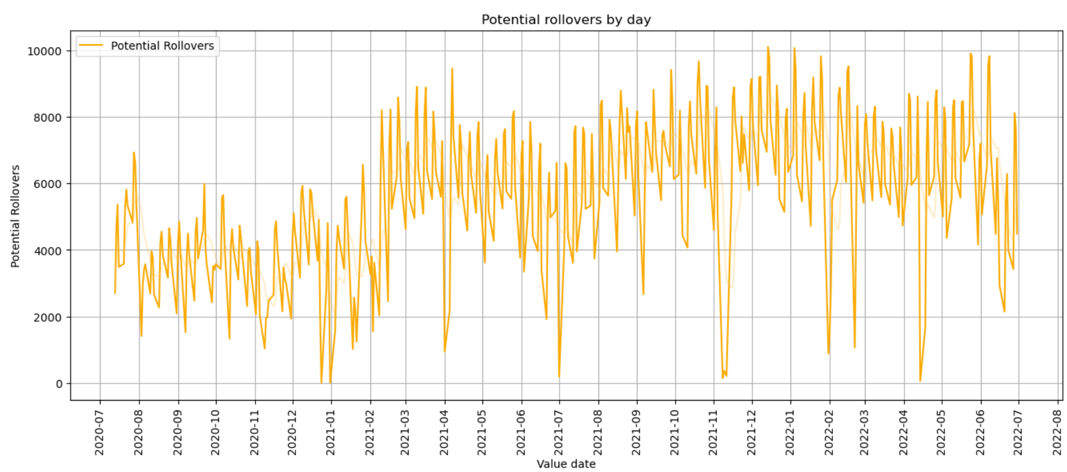


Figure 5.2: Moving average plot of the total number of only the repo rollover transactions in Europe from 01-07-2020 till 01-07-2022.

With the rollovers being programmed it is possible to compare the dataset when a transaction is a rollover and when it is not. In Table 5.1 a few of the features averages are depicted for when it is or is not a rollover transaction. A noticeable feature that differs a lot is whether a transaction has been cleared by a CCP or not. Of all the repo rollovers only 21.4% has been cleared while other transactions are cleared around 70.6%. This is interesting as according to the ECB the share of centrally cleared secured trades CCPs account for around 70% of one-day maturity trades [27]. CCPs eliminate counterparty risk by inserting themselves between the buyer and the seller of an agreed-upon trade. They do so in exchange for imposing a collateral-specific haircut to member institutions, a contribution to their 'default fund', and concentration limits [53]. As such, CCPs can help increase financial stability. However, it seems that parties that roll over their transaction do not include a CCP. This could be because if a CCP fails or faces significant financial difficulties, it can have systematic implications and impact multiple market participants. Another reason could be that as a higher proportion of trading is cleared across CCPs, more and more credit, liquidity, and operational risks will be concentrated in these institutions, which will themselves become potential sources of systemic risk [1]. Moreover, assuming parties want to roll over their transaction, having a CCP involved means that for every roll the CCP also needs to agree with the renewal. Therefore, it could be that it is considered easy to not involve a CCP or other party for a rollover transaction.

Average numerical values

	Rollover	Non-rollover
Interest rate (%)	0.125	-0.034
Principal amount (million €)	20.056	27.914
Market value (million €)	56.562	54.9134
Nominal value (million €)	19.177	26.178
Collateral updates	1.971	1.958
Cleared transactions (%)	21.395	70.645
Unit price (€)	34.865	38.853
EU trading venue (%)	17.444	65.468
nonEU trading venue (%)	3.467	4.856
OTC trading (%)	78.528	29.061
EU-EU trades (%)	20.043	50.437

Table 5.1: Rollover grouped statistics for analysis

Another interesting feature from Table 5.1 is that the rollover transactions are more likely to be OTC trading (78%) than in a trading venue in Europe (17%). This could be because rollover parties want more flexibility as with OTC trading there is no standardization of strike prices and expiration dates, so participants essentially define their own terms [54].

In Table 5.1 there is also a big difference in fixed agreed-upon interest rate, with rollover having a higher rate on average than other transactions. The agreed-upon average interest rate differs by almost 0.16% which is a big difference for these kinds of transactions with million dollar value contracts. The interest rate agreed upon can influence if the repo will be rolled over if the fixed interest rate agreed upon in the repo transaction is lower than the prevailing market rates or the cost of funds for the seller, it may be more favorable for the seller to roll over the transaction rather than seeking alternative funding sources. This is because the seller can continue to finance its securities at a lower cost. Furthermore, if market interest rates have decreased since the start of the original repo, the seller may prefer to roll over at the lower fixed interest rate instead of entering into a new transaction at the current higher rates. The agreed-upon interest rate can also influence the willingness of the counterparty to agree to a rollover. If the fixed interest rate in the repo transaction is favorable, they may be more willing to agree to the rollover. Note that overall the market interest rate changed a lot from the beginning of 2022 and continues (see Appendix Figure A.7). Therefore taking the mean can look different than the whole picture of interest rate fluctuations in the repo market, see Figure 5.3.

In Figure 5.3 the annualized interest rate on the principal amount of the repurchase transaction in accordance with the day count conventions is plotted for rollover transactions (in orange) and other transactions (in blue). In the secured money market, rates on repos experienced in part a delayed or even impaired transmission, mainly owing to an increase in collateral scarcity in 2022 [55]. Moreover, the limited supply of short-term securities and increased demand for such instruments contributed to only a partial pass-through of the policy rate increase to government bill rates. There is again a clear difference at the end and start of the year as in January 2021 the rates are significantly increasing for both types of transactions. To make it better understandable the weighted average interest rate with principal amount is plotted below in Figure 5.4. Again the rollover transaction (orange) has a higher mean fixed agreed-upon interest rate than the non-rollover transactions. A few spikes seem to be the same (Christmas and Easter) however now it is clear that the interest rate in the second half of 2021 has a big spike in both transaction types. ICMA stated that concerns about the 2021 year-end and the potential for a collateral shortage, particularly in the euro market, were being raised as early as October [56]. Participants report that many investors began positioning for the turnaround this time. Furthermore, participants cite a shortage of collateral, perhaps largely as a result of the ECB Targeted Longer-term Refinancing Operations, and a lack of access, whether directly or through local intermediaries, to NCB holdings [56]. This shortage could be an explanation for why the average and weighted average is higher during the year-end of 2021.

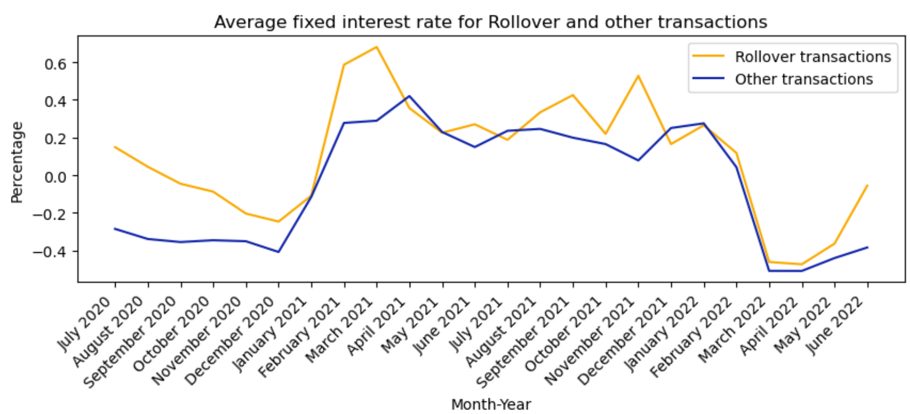


Figure 5.3: Fixed interest rate in Europe from 01-07-2020 till 01-07-2022.

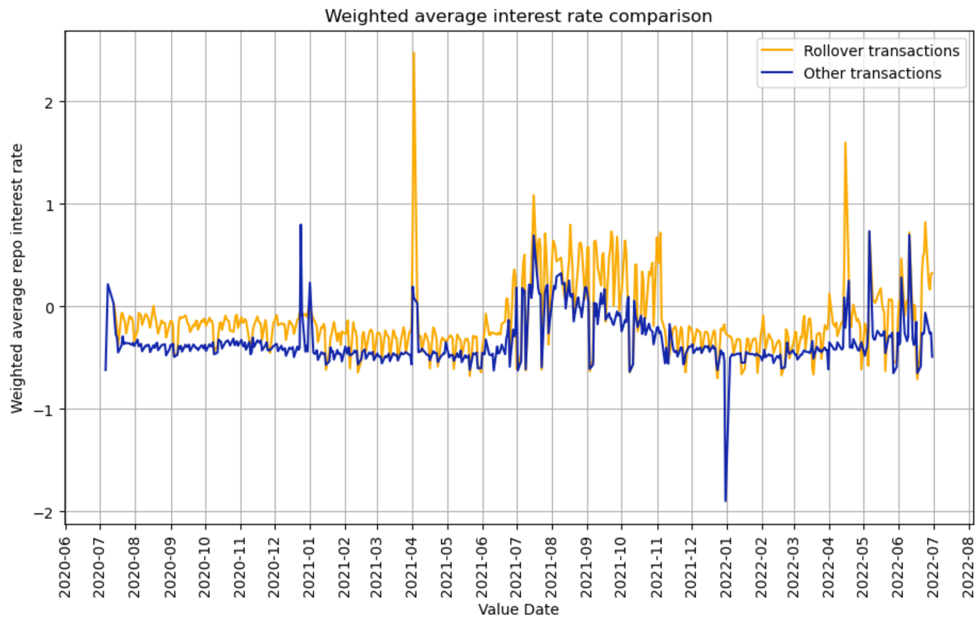


Figure 5.4: Weighted average of the fixed interest rate in Europe from 01-07-2020 till 01-07-2022.

The difference in other and rollover transactions for the cleared indicator is also visualized in Figure 5.5. Here there is a clear difference in the number of rollover transactions (orange line) that are being cleared compared to the other transaction (blue line). According to ICMA analysis [29], fluctuations and trends in the EU repo market were mainly driven by CCP-cleared repo transactions. An interesting pattern in CCP-cleared repo in the EU market is its tendency to rise gradually over the year and then step down at the start of the next year. This is especially noticeable in the rollover transactions in around January 2021. The other transactions are on average very stable around 70% cleared with some peaks, e.g. Christmas, Easter. A drawback of using a CCP could be that more credit, liquidity, and operational risk will be concentrated in these institutions, which will themselves become potential sources of systemic risk [1]. Parties that want to roll their trades most likely do not want to have more liquidity risk involved and choose to not have a CCP.

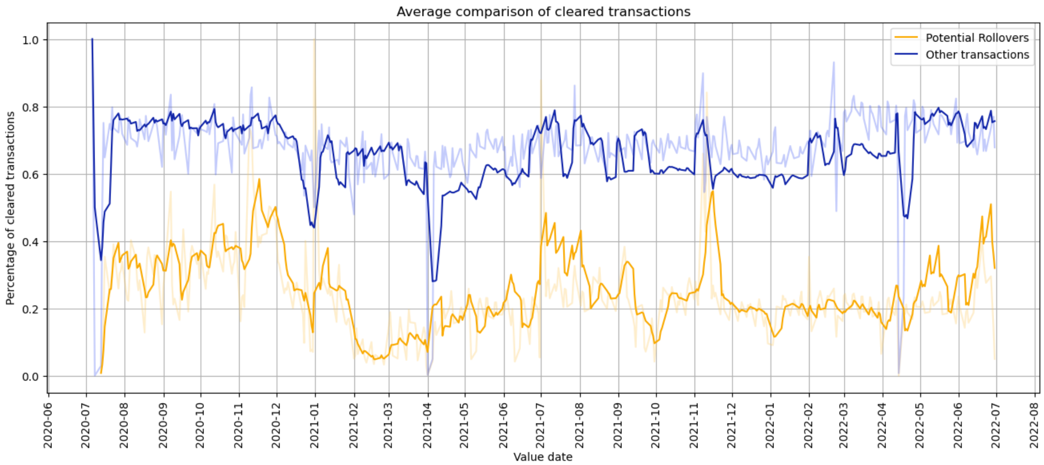


Figure 5.5: Percentage of transactions that have been cleared or not from 01-07-2020 till 01-07-2022. The bold lines are the moving average while the see-through lines are the actual values.

In Figure 5.6 is the sum of the principal amount of the *other* repo transactions and *rollover* transactions per day. The graph shows that the total principal amount over time of the rollover transactions lies lower as they only consist of 21% of the market and the average principal amount of a rollover trade is €20 million while that of a non-rollover transaction is €28 million. The lowest spikes are in January 2021, April 2021, November 2021, and April 2022.

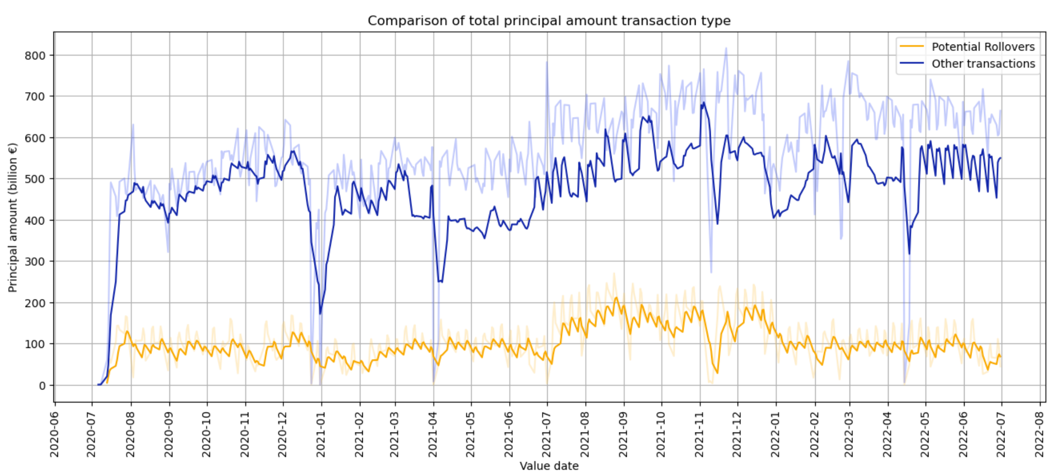


Figure 5.6: Sum of the principal amount for each day. The bold lines are the moving average while the see-through lines are the actual values.

Figure 5.7 and 5.8 show how the average price of the principal amount moves over time and how its rollover transactions compare. On average have the rollover transactions a lower principal amount (around €20 million per day) while the normal transactions are valued higher (around €28 million per day). It seems that rollover transactions move more with seasonality up and down on average (see Figure 5.7). This includes the rise of the average principal amount in December 2020 and 2021 followed by a big decrease in January 22021 and 2022. During 2021 the transaction types are both closer to each average principal amount. This could be due to the potential collateral shortage, talked about. No clear shocks can be seen in the plots.

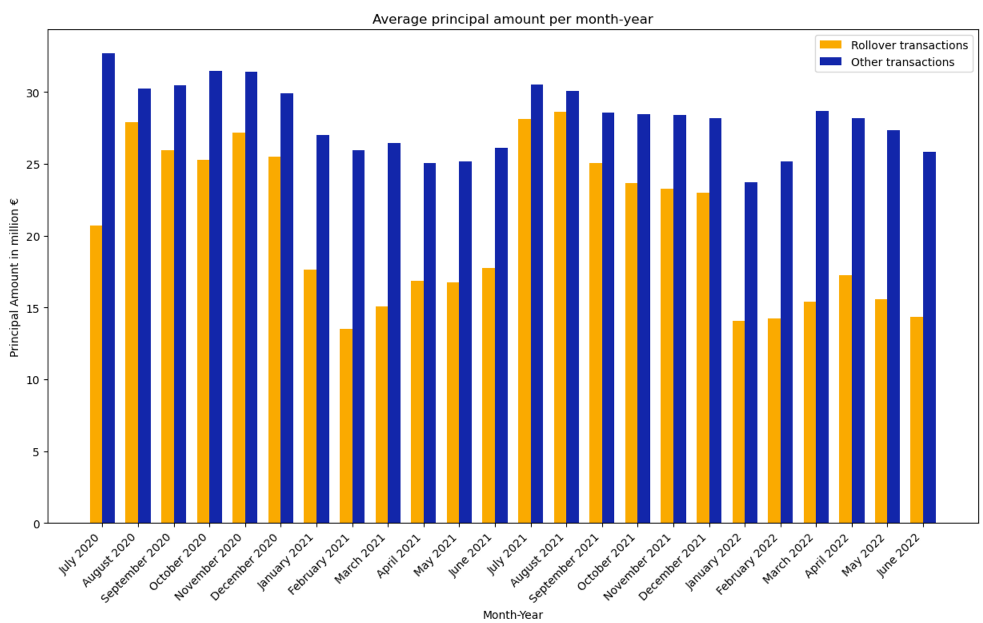


Figure 5.7: Average barplot of the principal amount per month. The bold lines are the moving average while the see-through lines are the actual values.

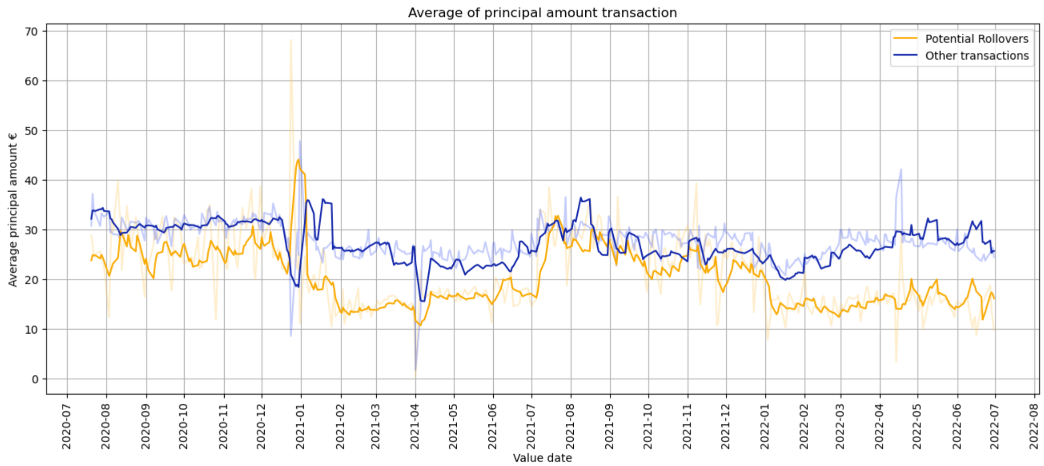


Figure 5.8: Average plot of the principal amount per day. The bold lines are the moving average while the see-through lines are the actual values.

The financial market and the whole economy are divided into institutional sector codes. In Table 5.2 the majority of all transactions are from sector S122[1], namely the *Deposit taking corporations, except the central bank*. There are more sectors in the data but only the most frequent ones are depicted in Table 5.2. Most rollover transactions are located in Sector S122, as are most repo transactions.

Sector	**Overall percentage**	**Percentage is a rollover**
S122	92.79%	23.87%
S125	3.17%	25.34%
S124	2.62%	25.52%
S123	1.12%	33.01%%
S127	0.13%%	48.20%
S129	0.11%	26.64%
S126	0.04%	43.31%

Table 5.2: Analysis percentage per sector

In Table 5.3 the description of the types of collateral used in a transaction are depicted in total overall percentage and percentage of when the transaction is a rollover. It is interesting to see that for around 75% of the total transactions from that percentage, only 13% of those government securities are rollover transactions, which is less than the 21% in total rollover transactions.

[1]The other institutional sector codes can be found on: ECB securities holding statistics

Collateral type	Overall percentage	Percentage is a rollover
Equities	2.39%	53.90%
Govt. securities	74.47%	12.95%
other	23.13%	56.94%

Table 5.3: Analysis of collateral type percentages on the dataset.

The percentage of rollover transactions in the datasets amounts to around 21% of the total. The average duration of a rollover is around 4 business days. This number can be explained due to the number of transactions on Friday (to Monday) being low. A transaction is defined as a rollover if the transaction contract is repeated on the next business day. However, most transactions take place from Monday to Thursday, see Table 3.1. With these results a new discussion point can arise: is a transaction also a rollover when the first contract is from Thursday to Friday and continues from Monday to Tuesday? From the Canadian repo market report (2016), Table A.1, repo transactions have a term of fewer than seven days and are rolled over regularly, primarily overnight (i.e., maturity transformation).

5.2 Machine Learning Classifiers

All the Machine Learning classifiers are first trained with all 34 features and default parameters and then optimized with feature importance and hyperparameter tuning.

5.2.1 Logistic Regression

The best features for the unoptimized Logistic Regression model are shown in Figure 5.9. Where the best predictors for rollover transactions are the *cleared indicator* and *if the trading venue is in Europe*. With recursive feature elimination a subset of the features are chosen to try and optimize the model. This starts with all features and recursively eliminates the least important features based on their importance scores (see Figure 5.9). After removing the least important features, the classifier is retrained, and the process is repeated until the desired number of features or performance metrics is reached. This approach allows to iteratively discard features that contribute the least to the classification task. The optimized model has 8 features.

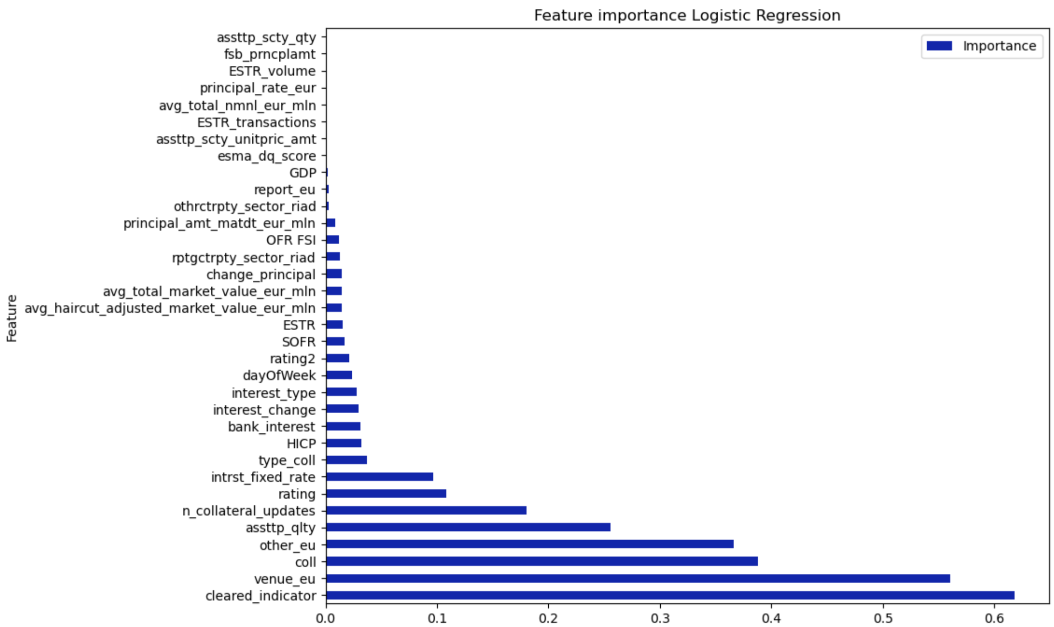


Figure 5.9: Feature importance with the Logistic Regression model.

The results of hyperparameter tuning for the Logistic Regression model, depicted in Table 5.4, demonstrate the impact of different regularization strengths on the model's performance. The optimal parameters for the Logistic Regression give the best results regarding the macro F1-score. The model does imply a penalty (regularization) during training. This could mean that the model does impose any additional constraint on the coefficients during training, allowing it to take any value to minimize the loss function. It is interesting to see that the best solver is the *newton-cg* solver as it is particularly suitable for problems when the data is not too large.

The results for the Logistic Regression classifier are depicted in Table 5.5 and 5.6. The model optimized classifies pretty accurately the predicted rollover transactions. It almost perfectly predicts the non-rollover transactions and makes a few mistakes in predicting the rollovers as the TP normalized rate is *0.411*. The resulting macro F1-score is *0.71* overall. Logistic Regression is sensitive to class imbalances since the majority class (*Other* *transactions*) significantly dominates the minority class (*Rollover transactions*) in terms of the number of instances, the classifier might struggle to find patterns and relevant features associated with rollovers. Therefore, it was essential to apply recursive feature elimination and parameter tuning. As Logistic Regression relies on informative features to make predictions. If the used features do not capture the patterns or characteristics of the rollover transaction, the model might struggle to find relationships in the features and ends up predicting the majority class for all instances.

Logistic Regression hyperparameters	Range	Best
penalty	l1, l2, elasticnet, None	l2
C	0.01, 1, 10, 50, 100, 150	100
solver	lbfgs, liblinear, newton-cg, sag, saga	newton-cg

Table 5.4: Overview of hyperparameter tuning of the Logistic Regression classifier and the best results for the optimized model.

Logistic Regression			
	Precision	Recall	F1-score
No	0.85	0.95	0.90
Yes	0.72	0.41	0.52
Accuracy			0.84
Macro avg	0.78	0.68	**0.71**
Weighted avg	0.82	0.84	0.82

Table 5.5: Classification result for the Logistic Regression classifier.

Logistic Regression		Actual	
		Negative	Positive
Predicted	Negative	0.955	0.045
	Positive	0.589	**0.411**

Table 5.6: LR normalized confusion matrix

The logistic regression classifier used as a benchmark does produce good results, especially after tuning the hyperparameters. In general, the LR model performs acceptable in predicting rollover transactions. With a macro F1-score of *0.71* and a TP normalized rate of *0.411*. The best features are *cleared indicator*, *if the trading venue is located in Europe* and, *if the collateral is a government security.*

5.2.2 Gaussian Naive Bayes

In Figure 5.10 the feature importance is visualized using the decision tree classifier, as the Gaussian Naive Bayes classifier from sklearn does not have a feature importance function. Therefore, the feature importance of the three tree-based models is implemented where the best results come back when using the best 12 features from the Random Forest model. For the hyperparameter selection only the var smoothing parameter is tuned with the GridSearchCV algorithm in Python, the parameters to tune are shown in Table 5.7.

Gaussian Naive Bayes hyperparameters	Range	Best
var smoothing	[np.linspace(0, 0.001, 100)] & [np.logspace(0,-9,num = 10)]	0

Table 5.7: Overview of hyperparameter tuning of the Gaussian Naive Bayes classifier and the best results for the optimized model.

The best result from the grid search is $var\_smoothing = 0.0$. Setting the var smoothing parameter to zero means that no smoothing is applied, and it assumes that there are no instances in the training data where a feature has a variance of zero for any given class. The classifier is therefore handling zero variances by not adding any smoothing or adjustments to the probability distribution calculation. This result could suggest that either the training data does not have instances with zero variances, or the smoothing technique used in Gaussian NB does not improve the classifier's performance on this specific dataset. The optimal *var_smoothing* parameter can vary depending on the specific characteristics of the data.

Using the selected features and setting *var_smoothing* to zero the result for the Gaussian NB are depicted in Table 5.8 and the confusion matrix in Table 5.9. The optimized model has a macro F1-score of *0.68*, which is a bit lower than the LR model. However, the TP normalized rate is *0.816*, which is better at predicting than the LR model. This could be because the Gaussian NB model assumes that the features are normally distributed within each class. The normality assumption of most of the features is not normal as stated in Section 3.4.

Gaussian Naive Bayes			
	Precision	Recall	F1-score
No	0.93	0.70	0.80
Yes	0.43	0.82	0.56
Accuracy			0.73
Macro avg	0.68	0.76	**0.68**
Weighted avg	0.82	0.73	0.75

Table 5.8: Classification result for the optimized Gaussian Naive Bayes.

Gaussian Naive Bayes		Actual	
		Negative	Positive
Predicted	Negative	0.703	0.297
	Positive	0.184	**0.816**

Table 5.9: Gaussian Naive Bayes normalized confusion matrix

Overall, the Gaussian NB model performs a bit better than the Logistic Regression model in correctly predicting rollover transactions. With a macro F1-score of *0.68* and a TP normalized rate of *0.816*. The most important features from the RF model were implemented as the feature selection method.

5.2.3 Decision Tree

In Figure 5.10 are the most important features of the Decision Tree model that were not optimized with parameter tuning. Again the recursive feature elimination method is used and reduced to 13 features. What is interesting to see in Figure 5.10 is that the Decision Tree model clearly sets feature *cleared indicator* as the most important factor in the prediction of repo rollovers. This is in line with the results from Table 5.1, where the mean of rollover transaction do not seem to be cleared by a central counterparty or other parties.

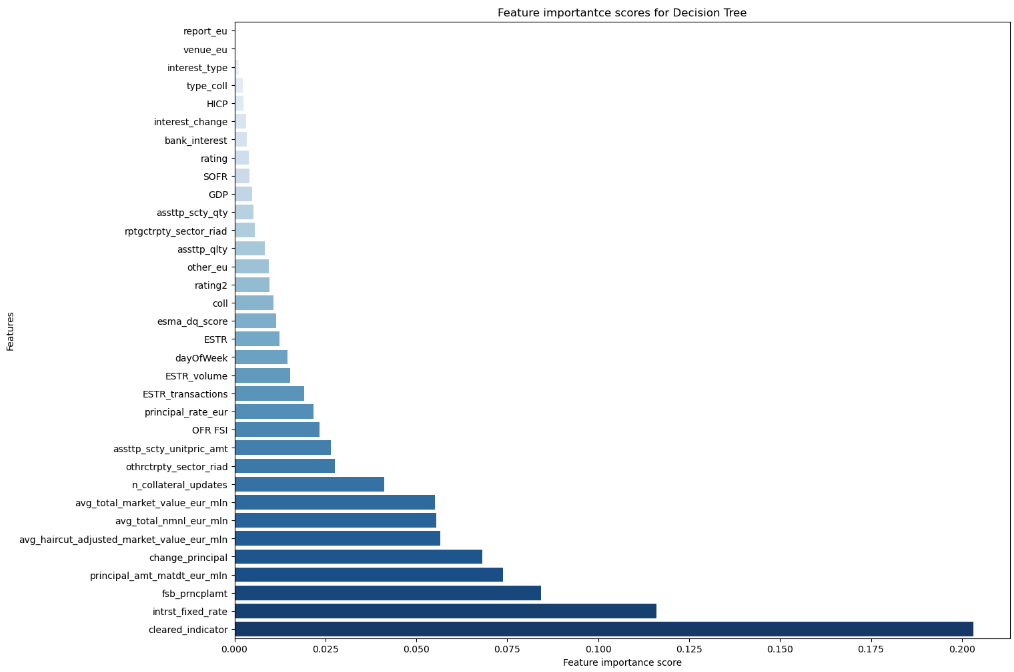


Figure 5.10: Feature importance with the unoptimized decision tree classifier

Next, the hyperparameters used to try and optimize the decision tree model results are in Table 5.10. The best hyperparameters and features are then used for the optimized model. The overall results are then in Table 5.11 where the most important measure are in bold. A comparison with the other models is written in Section 5.3. Overall, the classifier predicts quite nicely the rollover transactions from the other transactions.

Decision tree hyperparameters	Range	Best
max depth	[2, 3, 5, 10, 20, 30, 50]	30
min samples leaf	[1, 5, 10, 20, 50, 100]	50
criterion	[gini, entropy, log-loss]	entropy

Table 5.10: Overview of hyperparameter tuning of the Decision Tree classifier and the best results for the optimized model.

Decision Tree			
	Precision	Recall	F1-score
No	0.90	0.94	0.92
Yes	0.75	0.61	0.67
Accuracy			0.87
Macro avg	0.82	0.78	**0.79**
Weighted avg	0.86	0.87	0.87

Table 5.11: Classification result for the optimized Decision Tree.

Decision Tree		Actual	
		Negative	Positive
Predicted	Negative	0.943	0.057
	Positive	0.393	**0.607**

Table 5.12: Decision Tree normalized confusion matrix

The Decision Tree model gives good results and predicts all transactions the most accurately. Moreover, the DT model performs better than the LR model in predicting rollover transactions. With a macro F1-score of *0.79* and a TP normalized rate of *0.607*. The most important features to predict the model are *cleared indicator*, *the fixed agreed-upon interest rate* and, *the principal amount of the FSB at value date.*

5.2.4 Random Forest

For the implementation of the RF model, a new feature importance function is used to see if there are any differences with the DT model and if those features differ by how much, why, and does the model perform better. Random Forest feature importance is limited to decision trees and ensembles of decision trees. Per split in the decision tree, the information gain is assigned to the splitting feature as its importance measure [57]. This importance measure can be accumulated per feature over all trees. Other feature importance measures for tree-based methods are the feature's depth in the tree or the total number of instances that are used for classification [57]. In Figure 5.11 the most important features are depicted. Most features are similar for the RF and DT model, only the order in how much importance weight is given differs slightly. The *cleared indicator* is for both models the best feature. Then most features have a different order of importance. For the RF model the features *trading venue is located in Europe*, *number of collateral updates*, *if the collateral is a government bond* and *if the counterparty is located in Europe* are rated higher than for the DT model. For example, the feature *coll*, whether the collateral is government security, its weight is greater for the Random Forest model. This could be because government securities are considered of higher quality and therefore have a lower chance of defaulting or losing in value. The recursive feature elimination results for the Random Forest in 12 remaining features.

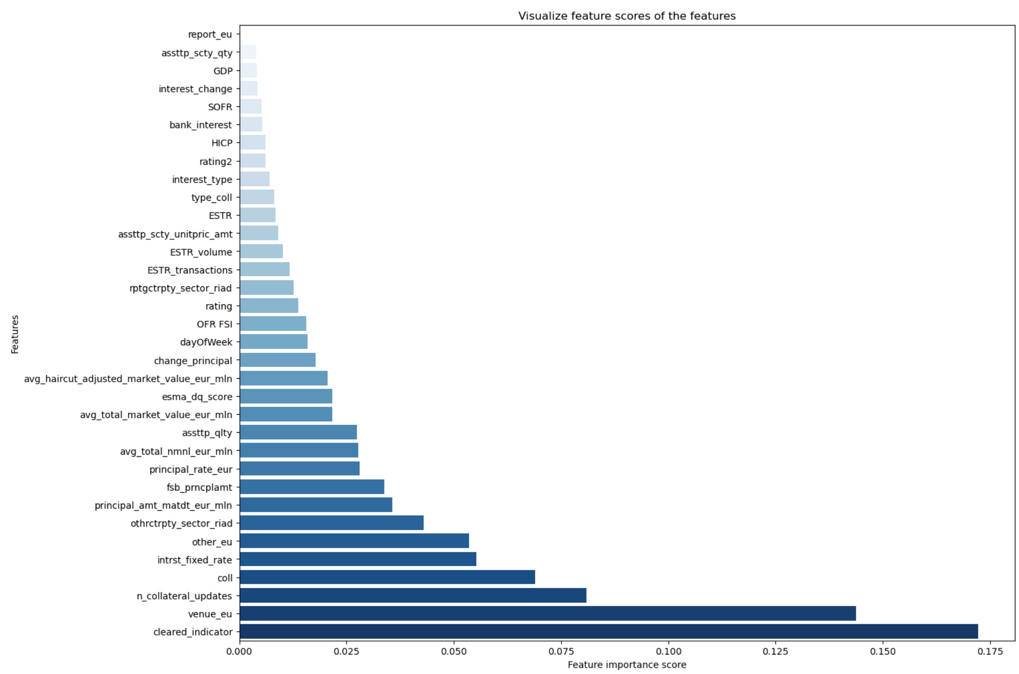


Figure 5.11: Feature importance with the Random Forest model.

Moreover, the hyperparameters for the RF model are tuned, and the range and best parameters are shown in Table 5.13. The *max depth* and *number of estimators* are lower for the RF than the DT model, and the *criterion* parameter is for both *entropy*.

RF hyperparameters	Range	Best
max depth	[2, 4, 6, 10, 15]	15
min samples leaf	[5, 10, 15, 20, 50, 100]	10
criterion	[gini, entropy, log-loss]	entropy

Table 5.13: Overview of hyperparameter tuning of the Random Forest classifier and the best results for the optimized model.

Furthermore, the results from the RF classifier are depicted in Table 5.14 and 5.15. Both the RF and DT model have the same macro F1-score of *0.79*, which means that both have similar levels of prediction performance. A reason for this could be that this is not uncommon or that the data is quite noisy or still has many outliers. The DT model predicts the TP (*0.607*) cases slightly better than the RF model (*0.591*). Thus, the DT model is the preferred choice for rollover risk analysis, as it is more likely to predict the rollover transactions correctly.

RF

	Precision	Recall	F1-score
No	0.89	0.95	0.92
Yes	0.77	0.59	0.67
Accuracy			0.87
Macro avg	0.83	0.77	**0.79**
Weighted avg	0.86	0.87	0.86

Table 5.14: Classification result for the optimized RF classifier.

RF		Actual	
		Negative	Positive
Predicted	Negative	0.949	0.051
	Positive	0.409	**0.591**

Table 5.15: Random Forest normalized confusion matrix

The Random Forest model gives good and similar results to the Decision Tree model. Yet, the Decision Tree model performs slightly better in predicting the TP cases and therefore is more useful for correctly identifying rollover transactions. With a macro F1-score of *0.79* and a TP normalized rate of *0.591* the Random Forest model performs well.

The most important features to predict the model are *cleared indicator*, *if the trading venue is in Europe* and, *the number of collateral updates*.

5.2.5 XGBoost

For the XGBoost model feature importance SHAP (SHapley Additive exPlanation) is employed to interpret the results and analyze the importance of individual features, see Figure 5.12. SHAP offers an insightful means to interpret the results from a complex algorithm such as XGBoost [58]. The technique is not only capable of evaluating the importance and direction of the impacts of a feature on the output of the model, but it can also extract complex and nonlinear joint impacts of features on the output of a model [58]. The feature importance is the best for the *cleared indicator* again, yet for the other features rankings are different than for the Decision Tree and Random Forest model. Here, the feature that if a *trading venue* is located in Europe is scored higher than in the DT model. The same goes for the features *day of the week* and *the counterparty is located in Europe*. The best 12 features are chosen with recursive feature elimination and used for the optimized model of XGBoost.

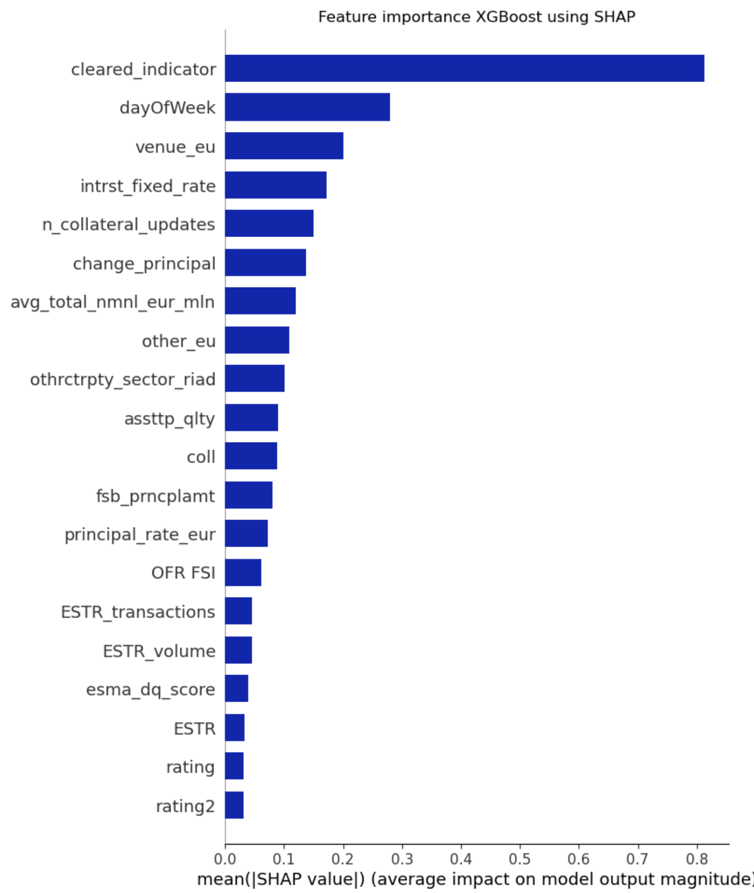


Figure 5.12: Feature importance with the XGBoost model.

The best-tuned parameters are in Table 5.16, where the *learning rate* is *0.1* which determines the step size at each iteration while moving towards a minimum of the loss function. The number of estimators is *200* and the maximum depth is *6*.

XGBoost hyperparameters	Range	Best
learning rate	[0.01, 0.05, 0.1]	0.1
max depth	[2, 4, 6]	6
number of estimators	[50, 100, 200]	200

Table 5.16: Overview of hyperparameter tuning of the XGBoost classifier and the best results for the optimized model.

With the feature selection and the parameter tuning the results are in Table 5.17 and 5.18. The highest macro F1-score is *0.78*. This indicates that the model has exhibited a strong ability to generalize and make accurate predictions on both the positive and negative classes. Although the ability to correctly predict TP cases is *0.553*, which is lower than the Decision Tree model (*0.607*) but not significantly.

XGBoost			
	Precision	Recall	F1-score
No	0.88	0.95	0.92
Yes	0.75	0.55	0.64
Accuracy			0.86
Macro avg	0.82	0.75	**0.78**
Weighted avg	0.86	0.86	0.86

Table 5.17: Classification result for the optimized XGBoost classifier.

XGB		Actual	
		Negative	Positive
Predicted	Negative	0.949	0.051
	Positive	0.447	**0.553**

Table 5.18: XGBoost normalized confusion matrix

Overall, the XGBoost model performs better than the Logistic Regression model in predicting rollover transactions. With a macro F1-score of *0.78* and a TP normalized rate of *0.553*. Along with different results for feature importance where the features *cleared indicator*, *day of the week* and *trading venue located in EU* are the most informative to the model.

5.2.6 Stochastic Gradient Descent

For the SGD classifier used from the sklearn package in Python, there is no feature importance function. Therefore, the Decision Tree parameters are used. The features from the Random Forest and XGBoost model gave the same results. Additionally, hyperparameter tuning is done, see Table 5.19 for the optimal parameters obtained. It is interesting to see that the best *loss* parameter is the *modified_huber*, which is another smooth loss that brings tolerance to outliers as well as probability estimates.

In Table 5.20 and 5.21 are the results from the classifier model. The model is biased towards not predicting the transactions as the recall is *0.07* for predicting rollover transactions. This classifier does not work for prediction repo rollover transactions to mitigate rollover risk.

SGD hyperparameters	Range	Best
alpha	[0.0001, 0.001, 0.01, 0.1]	0.001
loss	[hinge, log, squared hing, modified huber]	modified huber
penalty	[l2, l1, none]	none

Table 5.19: Overview of hyperparameter tuning of the Stochastic Gradient Descent classifier and the best results for the optimized model.

Stochastic Gradient Descent			
	Precision	Recall	F1-score
No	0.79	0.92	0.85
Yes	0.34	0.14	0.20
Accuracy			0.75
Macro avg	0.57	0.53	**0.53**
Weighted avg	0.70	0.75	0.71

Table 5.20: Classification result for the optimized Stochastic Gradient Descent classifier.

SGD		Actual	
		Negative	Positive
Predicted	Negative	0.924	0.076
	Positive	0.860	**0.148**

Table 5.21: Stochastic gradient descent normalized confusion matrix

The performance of the SGD model is not good and performs worse the Logistic Regression model. The macro F1-score is *0.53* with the TP cases normalized at *0.148*. The features are optimized using the Decision Tree feature importance.

5.3 Comparison

The results in Table 5.22 suggests that the Decision Tree model has a slight advantage in correctly identifying positive instances compared to the Random Forest and XGBoost classifiers, which are similar tree models. Overall, the best predictive models are the

Gaussian Naive Bayes and Decision Tree model.

While the benchmark model, Logistic Regression, and Stochastic Gradient Descent model perform the worst at predicting rollover transactions. It is pleasant to see that the more complicated models outperform the benchmark model. This could mean that for future implementation more complicated models predict when a repo transaction is rolled over and those models could prevent the change of rollover risk.

There can be several reasons why the Gaussian Naive Bayes classifier might outperform other models like Decision Trees, XGBoost, Random Forest, Stochastic Gradient Descent, and Logistic Regression in certain scenarios. A possible reason could be that Gaussian NB assumes that all features are conditionally independent given the class label. This could mean that the assumption holds and allows the model to effectively capture the relationships between the features and the target variable. Furthermore, the model requires fewer parameters to tune than the other models and is therefore less likely to overfit. Additionally, Gaussian NB performs well if there is class imbalance and the model requires fewer resources to compute compared to more complex models like XGBoost of Random Forest.

Comparison ML results		
Classifier	Normalized TP	F1-macro
LR	0.411	0.71
Gaussian NB	**0.816**	0.68
DT	0.607	**0.79**
SGD	0.148	0.53
RF	0.591	**0.79**
XGB	0.553	0.78

Table 5.22: Best results of the machine learning classifiers test set performance. With feature selection and hyperparameter optimization.

In Figure 5.13 all the classifiers results are visualized. In the Figure are the scores for the recall, F1-score, and accuracy. The classifier results show that the F1-score is always in between recall and accuracy (as it is the harmonic mean), except for Gaussian NB.

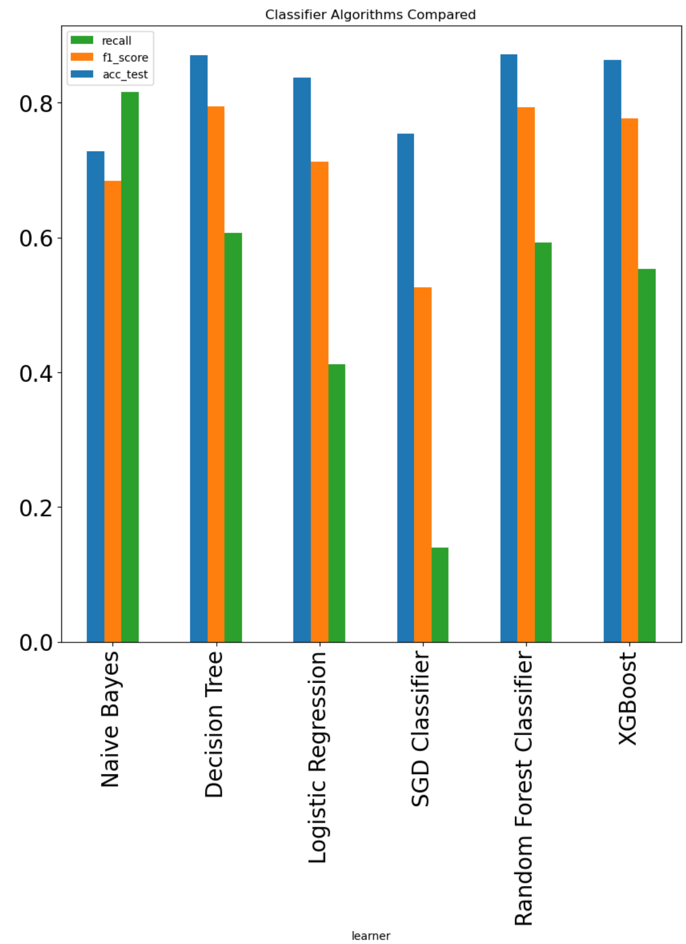


Figure 5.13: Comparing the results from classifiers - this graph will be run again for better results.

In Figure 5.14 the results from the predicted probabilities are visualized in an AUC-ROC graph. In the legend of the Graph, the AUC score is also illustrated, where the Decision Tree, Random Forest, and XGBoost model all have good scores of *0.88*, *0.89* and *0.87* respectively. These most fitting models are all in the left upper part of the ROC curve, the higher and more left the curve the better the prediction. There is a small (insignificant) difference between the three models as the Random Forest and Decision Tree have a somewhat higher True Positive Rate than the XGBoost model. This means that the predicted probabilities slightly differ from the prediction in the binary case, as shown in Table 5.22. However, these differences are limited and still show-case that the tree-based models perform well in the prediction of repo rollover transactions. The Gaussian

Naive Bayes model is clearly not better at predicting than the tree-based model but still has a relatively good score of *0.80* prediction probability. For the Logistic Regression model, performance is better than in the binary case for predicting rollovers with a *0.82* AUC-score. Lastly, the SGD model (see purple line Figure 5.14) has the worst AUC-score *0.52*, which means that the model predicts almost randomly all of the transactions. This could mean that the model is not suitable.

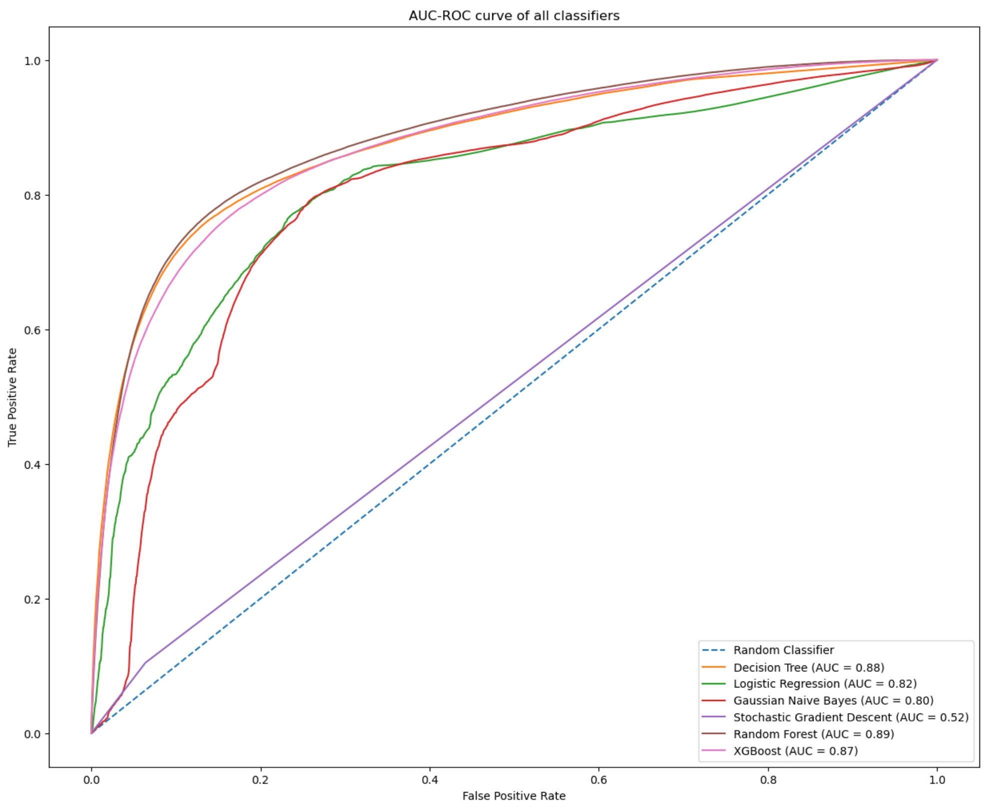


Figure 5.14: Results of the AUC-ROC predicted probabilities for all classifiers.

5.4 Possible Policy Implementation

To mitigate the risk of repo runs, market participants, regulators, and central banks employ various measures. These can include enhancing transparency and disclosure practices, implementing robust risk management frameworks, establishing collateral eligibility criteria, and providing liquidity support through central bank facilities during periods of stress.

Additionally, regulatory oversight and supervision of the repo market aim to ensure the resilience and stability of the financial system.

Interest rate changes are out of the control of individuals, so it could be difficult to minimalize rollover risk. The parties involved in the transactions could hedge on the interest rate and as the ECB has been doing for the last year is to control the three key interest rates accordingly.

Liquidity stress testing can be implemented with the knowledge that certain features are more important to identify rollover transactions and therefore the test can focus on rollover risk mitigation. Stress testing is an important tool in developing a complete picture of an institution's liquidity risk profile. Adequately designed and properly implemented liquidity stress tests can generate valuable information on a bank's liquidity profile that cannot be generated from a limited set of standardized liquidity metrics [10]. A liquidity stress test implementation of sudden market downturns or counterpart defaults with the data on repo rollovers and focus on its key characteristics would evaluate how financial institutions manage their liquidity positions.

With this thesis, a start has been made to work with the SFTR data to find and predict repo rollover transactions. With more knowledge from economics about the repo market a more complicated model can be built to see if the rollover transaction can be forecasted, given that some economic factors move a certain way. For example, suppose collateral credit rating drops significantly, what does the model predict will happen to the transactions? It would be interesting to see more research on such models.

Although this data is confidential it could be interesting for the ECB/DNB to know which unique parties are more likely to roll over their trades in recent history. If they would suddenly stop during a *stress* period or an economic change. It could indicate some form of a market freeze.

It can be useful to look at other markets' data to see if investors roll over their trades with the same characteristics (such as clearing by a CCP) and overall mean. Then to monitor whether the markets are losing confidence in a particular institution or has identified risks at an institution, it is useful to collect information on equity prices, credit default swap (CDS) spreads, money-market trading prices, the situation of roll-overs and prices for various lengths of funding, the price/yield of bank debenture or subordinated debt in the secondary market [59].

Another policy measure could be to report more regularly to provide details on funding sources, liquidity positions, and risk management practices, which could help the ECB/DNB monitor their rollover risk exposure. This could include the Liquidity Coverage Ratio, which is designed to ensure that banks hold a sufficient reserve of high-quality liquid assets to allow them to survive a period of significant liquidity stress for 30 calendar days. Or the Net Stable Funding Ratio (NSFR), which requires financial institutions to maintain a stable funding profile in relation to their assets and off-balance sheet exposures over a one-year horizon.

6: Conclusion

The goal of this paper was to analyze and predict repurchase agreement rollover transactions with machine learning classifiers. By providing an in-depth analysis of the transactional level data in the European repo market and defining what rollover transaction characteristics are the dependent variable is constructed in the dataset. During the data preparation, a noticeable result was that the overnight repo transaction do not occur much from Friday to Monday, while still having a similar average principal amount per transaction. From the graphs and tables created, we can see how the rollover transactions behave compared to the non-rollover transactions. Although, there are some missing dates due to holidays or misreporting issues. Overall, no clear shocks can be seen and therefore there is no indication of rollover risk during the analyzed time period. The results in finding the rollover transactions is that around 21% of all transactions are rolled over. This means that those parties continue to use the cash obtained from the initial sale of the security for a longer period of time than one business day. Furthermore, no clear shocks can be seen over time, or in stress periods. This means that events that happened during those years are not *stressful* enough that traders stop rolling over their repos. The results also showcase the important features that characterize a rollover transaction from a non-rollover transaction. The most apparent result is the cleared indicator. If a transaction has not been cleared by any other involved party such as a CCP the transaction is much more likely to be a rollover. The other meaningful features are the agreed-upon fixed interest rate, the indicator of whether the trading venue is located in Europe, and the type of collateral. Moreover, on average the fixed agreed-upon interest rate of a rollover is higher and the average principal amount is lower compared to a non-rollover transaction. Additionally, all the Machine Learning classifiers were able to predict the rollover transaction and the Decision Tree model was able to do it best with a macro F1-score of *0.79*. While the Gaussian Naive Bayes classifier was overall better able to identify the rollover transactions with a True Positive normalized rate of *0.784*. For future research, a liquidity stress test can be implemented focusing on the key rollover characteristics or forecasting models.

7: Discussion

A key challenge in providing the rollover transactions in the repo market is the sheer size of the database, due to the (double) reporting of daily status and transaction reports and the over 800 contract fields per entry. This makes it particularly challenging to distinguish between misreporting, data glitches, and interesting patterns worthwhile exploring. With that in mind, having put significant effort into developing the cleaning codes in a way that it can be reused and extended to the other subsets of the SFTR data. Another, consideration for the indicator if a repo transaction is a rollover could be that humans tend to agree with *simple* numbers. When two parties agree upon a principal value of the underlying collateral people tend to use rounded numbers, e.g. 10 million euro government bond. This makes it a little bit harder to say for sure if the transaction that occurs is rolled over or if the parties use similar/constant numbers.

There is no clear definition of when a transaction is officially rolled over. In this report, a rollover transaction is considered to have an overnight maturity of one business day. Is a transaction also a rollover when the first contract is from Thursday to Friday and continues from Monday to Tuesday? The SFTR dataset from the ECB is very confidential and can not be published publicly. Therefore, only people with access to the data could be consulted for questions. Furthermore, this thesis topic is very specific in a distinct market section. This could mean that, as the writer of this thesis is not an expert on this topic, mistakes were made in assumptions. This is also the case when looking for a definition for rollover trades in the repo market as ICMA and ESMA have different definitions for those transactions in the SFTR dataset. The implemented extra features from public data eventually did not perform well during the Machine Learning classifiers. The macroeconomic factors did not affect whether a transaction would be rolled over or not. One of the main issues could be that the number of rollover transactions over the years is quite steady (see Figure 5.1), therefore taking public data that is matched per day would not help. For future work an algorithm could be created to update a nowcasting model, each time new data is published, based on the SFTR dataset.

Bibliography

[1] *Frequently Asked Questions on Repo*. International Capital Market Association. 2019. URL: https://www.icmagroup.org/market-practice-and-regulatory-policy/repo-and-collateral-markets/icma-ercc-publications/frequently-asked-questions-on-repo/. (accessed: 23.05.2023).

[2] D. Duffie. "Special repo rates". In: *The Journal of Finance* 51.2 (1996), pp. 493–526. DOI: https://doi.org/10.1111/j.1540-6261.1996.tb02692.x.

[3] D. Duffie, N. Gârleanu, and L. H. Pedersen. "Securities lending, shorting, and pricing". In: *Journal of Financial Economics* 66.2-3 (2002), pp. 307–339. DOI: https://doi.org/10.1016/S0304-405X(02)00226-X.

[4] J.-M. Bottazzi, J. Luque, and M. R. Páscoa. "Securities market theory: Possession, repo and rehypothecation". In: *Journal of Economic Theory* 147.2 (2012), pp. 477–500. DOI: https://doi.org/10.1016/j.jet.2010.11.004.

[5] A. Kotidis and N. Van Horen. "Repo market functioning: The role of capital regulation". In: (July 2018).

[6] A. Brassil, H. Hughson, and M. McManus. *Identifying Interbank Loans from Payments Data*. Tech. rep. Reserve Bank of Australia, Dec. 2016.

[7] C. Garriott and K. Gray. *Canadian repo market ecology*. Tech. rep. Bank of Canada Staff Discussion Paper, 2016. DOI: https://doi.org/10.34989/sdp-2016-8.

[8] Bank for International Settlements Committee on Payment and Settlement Systems (CPSS). *Strengthening repo clearing and settlement arrangements*. 2010.

[9] R. Comotto. *ICMA Recommendations for Reporting under SFTR*. Apr. 2023.

[10] *Liquidity stress testing: a survey of theory empirics and current industry and supervisory practices*. Tech. rep. Bank for International Settlements, Oct. 2013.

[11] A. Martin, D. R. Skeie, and E.-L. Von Thadden. "Repo runs". In: *Review of Financial Studies* 27.4 (2010), p. 42.

[12] M. Bouvard, P. Chaigneau, and A. De Motta. "Transparency in the financial system: Rollover risk and crises". In: *The Journal of Finance* 70.4 (2015), pp. 1805–1837. DOI: https://doi.org/10.1111/jofi.12270.

[13] Y. Leitner. "Why Do Markets Freeze?" In: (2011).

[14] V. V. Acharya, D. Gale, and T. Yorulmazer. "Rollover Risk and Market Freezes". In: *The Journal of Finance* 66.4 (Jan. 2010), pp. 1177–1209. DOI: https://doi.org/10.1111/j.1540-6261.2011.01669.x.

[15] V. Acharya and A. Krishnamurthy. "Why bankers must bear the risk of 'too safe to fail' assets". In: (Mar. 2010).

[16] S. Hur and I. O. Kondo. "A theory of rollover risk, sudden stops, and foreign reserves". In: *Journal of International Economics* 103 (2013), pp. 44–63. DOI: https://doi.org/10.1016/j.jinteco.2016.08.006.

[17] M. Belhaj and Y. Hachaıchi. "Artificial Intelligence, Machine Learning and Big Data in Finance Opportunities, Challenges, and Implications for Policy Makers". In: (Aug. 2021). URL: https://www.oecd.org/finance/artificial-intelligence-machine-learning-big-data-in-finance.htm.

[18] M. Erdem and T. Park. *A novel machine learning-based validation workflow for financial market time series*. Tech. rep. Bank for International Settlements, 2022.

[19] N. Garvin. "Identifying repo market microstructure from securities transactions data". In: *RBA Research Discussion Papers rdp2018-09, Reserve Bank of Australia* (2018).

[20] C. Brand, L. Ferrante, and A. Hubert. *From cash-to securties-driven euro area repo markets: the role of financial stress and safe asset scarcity*. Tech. rep. European Central Bank, 2019.

[21] European Securities and Markets Authority. *SFTR Reporting*. URL: https://www.esma.europa.eu/data-reporting/sftr-reporting. (accessed: 02.03.2023).

[22] N. Kessler, I. van Lelyveld, and E. van der Woerd. *The European Equity Lending Market: Exclusive Security Lending Agreements*. Apr. 2023.

[23] Alexander Borek, Ajith K. Parlikad, Jela Webb, and Philip Woodall. "Software Tools: Automated Methods for TIRM". In: *Total Information Risk Management*. Ed. by A. Borek, A. K. Parlikad, J. Webb, and P. Woodall. Boston: Morgan Kaufmann, 2014, pp. 237–269. ISBN: 978-0-12-405547-6. DOI: https://doi.org/10.1016/B978-0-12-405547-6.00012-2. URL: https://www.sciencedirect.com/science/article/pii/B9780124055476000122.

[24] Nathan Reiff, Gordon Scott, and Yarilet Perez. *Repurchase Agreement (Repo): Definition, Examples, and Risks*. Mar. 2023. URL: https://www.investopedia.com/terms/r/repurchaseagreement.asp. (accessed: 12.06.2023).

[25] S. Corradin, J. Eisenschmidt, M. Hoerova, T. Linzert, G. Schepens, and J.-D. Sigaux. *Money markets, central bank balance sheet and regulation*. Discussion Paper. 2020.

[26] V. Aggarwal, V. Gupta, P. Singh, K. Sharma, and N. Sharma. "Detection of spatial outlier by using improved Z-score test". In: *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE. 2019, pp. 788–790. DOI: 10.1109/ICOEI.2019.8862582.

[27] *Euro money market study 2020*. Apr. 2020. URL: https://www.ecb.europa.eu/pub/euromoneymarket/html/ecb.euromoneymarket202104_study.en.html#toc10.

[28] R. Comotto. *The first year of SFTR public data on repo*. Tech. rep. Sept. 2021.

[29] *ICMA analysis: SFTR public data for repo in 2022*. Tech. rep. ICMA group, Apr. 2023.

[30] S. DellaVigna and J. M. Pollet. "Investor inattention and Friday earnings announcements". In: *The journal of finance* 64.2 (2009), pp. 709–749. DOI: https://doi.org/10.1111/j.1540-6261.2009.01447.x.

[31] *Repo Market*. [Online; accessed 24-July-2023]. URL: https://www.eurex.com/ex-en/markets/eurex-repo/euro-repo-market.

[32] *Statistical Data Warehouse*. European Central Bank. URL: https://sdw.ecb.europa.eu/. (accessed: 03.05.2023).

[33] *OFR Financial Stress Index*. Office of Financial Research. URL: https://www.financialresearch.gov/financial-stress-index/. (accessed: 03.05.2023).

[34] S. R. Baker, N. Bloom, and S. J. Davis. *Global Economic Policy Uncertainty Index: Current price adjusted GDP*. FRED, Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/GEPUCURREN. (accessed: 03.05.2023).

[35] *ECB interest rates*. De Nederlandsche Bank. URL: https://www.dnb.nl/en/the-euro-and-europe/the-ecb-s-monetary-policy/ecb-interest-rates/. (accessed: 03.05.2023).

[36] F. Chollet. *Deep learning with Python*. Simon and Schuster, 2021.

[37] F. Pereira, T. Mitchell, and M. Botvinick. "Machine learning classifiers and fMRI: a tutorial overview". In: *Neuroimage* 45.1 (2009), S199–S209. DOI: 10.1016/j.neuroimage.2008.11.007.

[38] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830. (accessed on: 14.06.2023).

[39] L. Yang and A. Shami. "On hyperparameter optimization of machine learning algorithms: Theory and practice". In: *Neurocomputing* 415 (2020), pp. 295–316. DOI: https://doi.org/10.1016/j.neucom.2020.07.061.

[40] A. Haghpanah Jahromi and M. Taheri. "A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features". In: *2017 Artificial intelligence and signal processing conference (AISP)*. IEEE. 2017, pp. 209–212. DOI: 10.1109/AISP.2017.8324083.

[41] C. A. Ratanamahatana and D. Gunopulos. "Feature selection for the naive bayesian classifier using decision trees". In: *Applied artificial intelligence* 17.5-6 (2003), pp. 475–487. DOI: 10.1080/713827175.

[42] S. B. Kotsiantis. "Supervised machine learning: A review of classification techniques". In: *Emerging artificial intelligence applications in computer engineering* 160.1 (2007), pp. 3–24.

[43] Priyanka and D. Kumar. "Decision tree classifier: a detailed survey". In: *International Journal of Information and Decision Sciences* 12.3 (2020), pp. 246–269. DOI: https://doi.org/10.1504/IJIDS.2020.108141.

[44] T. Suryakanthi. “Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm”. In: *International Journal of Advanced Computer Science and Applications* 11.2 (2020), pp. 612–619. DOI: 10.14569/IJACSA.2020.0110277.

[45] B. Hacibedel and R. Qu. “Understanding and Predicting Systemic Corporate Distress: A Machine-Learning Approach”. In: (July 2022).

[46] A. E. Maxwell, T. A. Warner, and F. Fang. “Implementation of machine-learning classification in remote sensing: An applied review”. In: *International Journal of Remote Sensing* 39.9 (2018), pp. 2784–2817. DOI: https://doi.org/10.1080/01431161.2018.1433343.

[47] L. Breiman. “Random forests”. In: *Machine learning* 45 (2001), pp. 5–32. DOI: https://doi.org/10.1023/A:1010933404324.

[48] T. Chen and C. Guestrin. “Xgboost: A scalable tree boosting system”. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794.

[49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. *Stochastic gradient descent*. [Online; accessed 24-July-2023]. 2011. URL: https://scikit-learn.org/stable/about.html#citing-scikit-learn.

[50] S. Simske. “Chapter 4 - Meta-analytic design patterns”. In: *Meta-Analytics*. Morgan Kaufmann, 2019, pp. 147–185. ISBN: 978-0-12-814623-1. DOI: https://doi.org/10.1016/B978-0-12-814623-1.00004-6.

[51] D. Chicco and G. Jurman. “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation”. In: *BMC genomics* 21 (2020), pp. 1–13. DOI: https://doi.org/10.1186/s12864-019-6413-7.

[52] T. Fawcett. “An introduction to ROC analysis”. In: *Pattern recognition letters* 27.8 (2006), pp. 861–874. DOI: https://doi.org/10.1016/j.patrec.2005.10.010.

[53] C. Boissel, F. Derrien, E. Örs, and C. Thesmar. “Systemic risk in clearing houses: Evidence from the European repo market”. In: *Journal of Financial Economics* 125.3 (2017), pp. 511–536.

[54] A. Ganti. *OTC Options: Difference From Standard Options, Risks*. 2022. URL: https://www.investopedia.com/terms/o/otcoptions.asp.

[55] *Annual Report*. Tech. rep. 2022. URL: https://www.ecb.europa.eu/pub/pdf/annrep/ecb.ar2022~8ae51d163b.en.pdf.

[56] A. Hill. *The European repo market at 2021 year-end. An ICMA European Repo & Collateral Council (ERCC) briefing note*. Tech. rep. International Capital Market Association, 2022.

[57] N. Burkart and M. F. Huber. “A survey on the explainability of supervised machine learning”. In: *Journal of Artificial Intelligence Research* 70 (2021), pp. 245–317.

[58] A. B. Parsa, A. Movahedi, H. Taghipour, S. Derrible, and A. K. Mohammadian. "Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis". In: *Accident Analysis & Prevention* 136 (2020), p. 105405.

[59] *Supervisory review process Liquidity monitoring metrics*. Tech. rep. Bank for International Settlements, 2019.

[60] *What is Collateral? Definition, Meaning, and Example*. Yieldstreet. May 2023. (accessed: 18.04.2023).

[61] *Principal*. bdc. URL: `https://www.bdc.ca/en/articles-tools/entrepreneur-toolkit/templates-business-guides/glossary/principal#:~:text=Principal%5C%20is%5C%20the%5C%20amount%5C%20of,fee%5C%20for%5C%20making%5C%20money%5C%20available`. (accessed: 18.04.2023).

[62] J. Chen. *What Is Window Dressing in Finance?* URL: `https://www.investopedia.com/terms/w/windowdressing.asp`.

[63] T. Segal, E. Estevez, and Y. Perez. *Rollover Risk*. July 2022. URL: `https://www.investopedia.com/terms/r/rollover-risk.asp`. (accessed: 19.04.2023).

[64] ICMA European Repo Council. *A Guide to Best Practice in the European Repo Market*. July 2015.

[65] R. Heijmans and S. G. Yun. "Analysis of risk factors in the korean repo market based on us and european repo market experiences during the global financial crisis". In: *Journal of Financial Market Infrastructures* 4.1 (2015), pp. 25–58. DOI: `http://dx.doi.org/10.2139/ssrn.2580594`.

[66] *Understanding repo: A cash building block*. BlackRock, 2022. (accessed: 18.04.2023).

[67] R. Comotto. *International Capital Market Association European Repo Market Survey*. 44. Zurich, Mar. 2023.

[68] A. Hayes, G. Scott, and Y. Perez. *Understanding Liquidity and How to Measure It*. Mar. 2023. URL: `https://www.investopedia.com/terms/l/liquidity.asp#:~:text=Tara%20Anand%20%2F%20Investopedia-,What%20Is%20Liquidity%3F,turn%20it%20back%20into%20cash.`. (accessed: 04.07.2023).

[69] *Repo market functioning*. Apr. 2017.

[70] S. Kördel N. de Vette B. Klaus and A. Sowiński. *Financial Stability Review. Gauging the interplay between market liquidity and funding liquidity*. Tech. rep. May 2023.

Appendix A: Appendix

A.1 Definitions

Collateral Collateral refers to marketable financial securities securities, such as bonds, or other types of assets, such as non-marketable assets or cash [60]. If a borrower defaults on the loan, the lender can seize the collateral and sell it to recoup its losses.

CCP Central Clearing Counterparty (CCP) is a financial institution that takes on counterparty credit risk between parties to a transaction and provides clearing and settlement services for trades in foreign exchange, securities, options, and derivative contracts. CCPs are highly regulated institutions that specialize in managing counterparty credit risk.

Principal amount Principal is the amount of money a company borrows when it takes a loan. This amount is recorded on a promissory note as proof of the debt owed. In all but the rarest of situations, the borrower must pay interest, which is the lender's fee for making money available. The interest is calculated on the principal and almost always paid monthly [61].

Repurchase agreement transaction a repo is a short-term (usually overnight) collateralized loan, in which the borrower (of cash) sells security (typically government bonds as collateral) to the lender, with a commitment to buy it back later at the same price plus interest.

Liquidity is a key measure of how well financial markets are working. It refers to how easily assets can be bought or sold—and when it dries up, it can be disruptive. This may pose risks to financial stability.

Repo rate is, in effect, an interest rate on loans collateralized by a specific instrument. A *special* repo rate is a rate significantly below prevailing market riskless interest rates [2].

Window dressing refers to the practice of altering financial data to appear more attractive to investors. It can be identified by carefully evaluating a fund's holdings and looking for suspicious trades coinciding with the end of a quarter or fiscal year [62].

Credit a contract agreement in which a borrower receives a sum of money or something of value and repays the lender at a later date, generally with interest.

Repo rollover transaction being renewed. A rollover occurs when the account owner has an existing repo deal with a counterparty and then executes a new repo deal with the same counterparty where no new money is exchanged, however, leverages the collateral already in place.

Rollover risk is the possibility that a borrower cannot refinance by borrowing to repay existing debt. Generally, the shorter-term the maturing debt, the greater the borrower's rollover risk. Rollover risk reflects economic conditions (e.g. liquidity and credit markets) versus a borrower's financial condition [63].

Leverage ratio is any kind of financial ratio that indicates the level of debt incurred by a business entity against several other accounts in its balance sheet, income statement, or cash flow statement.

Over the Counter Over the Counter (OTC) is the process of trading securities via a broker-dealer network as opposed to on a centralized exchange like MTFs.

Trading Venue Trading venue is defined as a regulated market, an MTF or an OTC. MTF (Multilateral Trading Facility) is a European term for trading system that facilitates the exchange of financial instruments between multiple parties.

DNB De Nederlandsche Bank (DNB) is an organization that supervises the government of the Netherlands and other important financial institutions. They report back to the European Central Bank (ECB) and are the central bank of the Netherlands.

ECB The European Central Bank is the central bank for the euro area and is responsible for monetary policy and financial stability within the Eurozone countries. Its main objective is to maintain price stability and control inflation.

MMF Money Market Fund (MMF) is a kind of mutual fund that invests in highly liquid, near-term instruments. Though not quite as safe as cash, money market funds are considered extremely low-risk on the investment sprectrum.

Monetary policy Monetary policy is a set of actions to control a nation's overall money supply and achieve economic growth.

Liquidity stress testing Liquidity stress test is a financial analysis technique to assess the ability of a firm or market to withstand liquidity shocks or adverse conditions. It involves simulating scenarios to measure potential liquidity shortfalls and the impact of such events on a company's ability to meet its financial obligations [10].

FSB The Financial Stability Board (FSB) is an international body that monitors and makes recommendations about the global financial system.

SFTR Securities Financing Transaction Regulation (SFTR) is the dataset used in this thesis that includes all securities financing transactions (SFTs) with at least one trading party operating in the European Union (EU).

SFTs Securities financing transactions (SFTs) include a repurchase transaction, securities or commodities lending, securities or commodities borrowing, buy-sell back transaction or sell-buy back transaction, and, margin lending transactions.

ESMA European Securities and Markets Authority (ESMA) is the EU's financial markets regulator and supervisor. ESMA's mission is to enhance investor protection, promote orderly financial markets and safeguard financial stability.

ICMA International Capital Market Association (ICMA) is a self-regulatory organization and trade association for participants in the capital markets.

BIS The Bank of International Settlements (BIS) is an international financial institution which is owned by member central banks. Its primary goal is to foster international monetary and financial cooperation while serving as a bank for central banks.

Repo run In a repo run, a leveraged financial institution may not be able to roll over its short-term borrowing, despite it being collateralized [11].

UTI Unique Transaction Identifier (UTI) is a unique reference assigned to the SFT in order to identify the trade.

Market value Market value is the price an asset would fetch in the marketplace, or the value that the investment community gives to a particular equity or business.

Nominal value Nominal value of a security, often referred to as face or par value, is its redemption price and is normally stated on the front of that security.

NCB National Central Bank (NCB) refer to a central bank of an EU member state.

Securities A security is a tradeable financial asset.

A.2 De Nederlandsche Bank

De Nederlandsche Bank (DNB) is an organization that supervises the government of the Netherlands and other important financial institutions. They report back to the European Central Bank (ECB) and are the central bank of the Netherlands. The DNB is committed to safeguarding financial stability and thereby contributing to sustainable prosperity in the Netherlands. They ensure price stability and balanced macroeconomic development in Europe, a shock-resilient financial system, and a secure, reliable, and efficient payment system. Furthermore, they are a sound and ethical financial institution that fulfills its obligations and commitments.

The DNB is the central bank of the Netherlands, responsible for implementing monetary policy, promoting financial stability, and supervising financial institutions. It was established in 1814 and is one of the oldest central banks in the world. DNB's main objective is to maintain price stability and support the economic policies of the Dutch government. To achieve this goal, DNB sets the interest rate and conducts open market operations to influence the supply of money and credit in the economy. DNB also supervises banks, insurance companies, and other financial institutions to ensure that they operate safely and soundly and that they comply with applicable laws and regulations. In addition to its domestic responsibilities, DNB is also involved in international cooperation and participates in various global organizations such as the European Central Bank (ECB) and the Bank for International Settlements (BIS). DNB is an independent institution and operates under the supervision of the Dutch parliament. It is accountable to the public through regular reporting on its activities and policies. Overall, DNB plays a critical role in the Dutch economy and financial system, promoting stability and growth while safeguarding the interests of consumers and the broader public.

A.3 Business Understanding

The repo market is one of the largest and most actively traded sectors in the short-term credit markets and is an important source of liquidity for many money market funds (MMFs). The tri-party repo is the most widely used form of repo across MMFs (mostly in the United States), where a third party – a custodian bank or clearing organization – serves as an intermediary between the counterparties to the deal. According to the International Capital Market Association (ICMA) [1], a repo performs four basic functions:

1. One party can invest cash secured against the asset provided as collateral – **safe investment**.

2. The counterparty can borrow cash in order to finance a long position in an asset, in an amount and at a repo rate that reflects, among other things, the collateral provided to the lender – **cheap borrowing**.

3. One party can earn a return by lending out an asset that is in demand in the market, in exchange for cheap cash, which can be used for funding or reinvested for profit – **yield investment** for securities investors.

4. The counterparty can borrow an asset in order to sell and establish a short position or to deliver in order to settle a sale that has already been agreed – **short-selling**.

The repo market functions mainly as a source of short-term liquidity and a market for (high-quality) collateral. In a repo, one party is selling securities in order to buy these back at a slightly different price. Effectively, this party is borrowing cash, collateral-backed, with a typical maturity of overnight up to 3 days, at a certain repo interest rate. However, longer-dated transactions are also observed. The use of repo transactions is also linked to the use of derivatives, as both the required and the received cash collateral (initial margin, variation margin) for derivatives is often borrowed (lent) in the repo market on a daily basis. Examples of underlying securities are government securities, (corporate) debt securities, and (main index) equities. As usually high-quality collateral is used, EU repo rates are close to the ECB's deposit facility, reflecting the relatively low risk of the transactions.

Money markets include several markets and instruments which share the characteristic of providing short-term funding or collateral, with the maturity of transactions of up to and including one year. This includes short-term unsecured loans, secured short-term loans, sovereign bills, commercial papers, certificates of deposit, and money market mutual funds. Participants in money markets include banks, non-bank financial institutions such as investment funds and money market funds, as well as non-financial corporations. In an unsecured transaction, liquidity (cash) is exchanged for a promise of repayment at a future date (most commonly overnight). In a secured transaction, the trade is collateralized. Secured transactions have been on an increasing path, and this trend has been accentuated by the Global Financial Crisis. The share of the secured transactions in Europe in total (secured plus unsecured) turnover increased from around 60% in 2005 to more than 95% in 2019 [25].

According to the ICMA survey [1], repo is a more stable source of short-term wholesale funding than unsecured deposits because collateral in the form of high-quality liquid assets hedges the credit and liquidity risk of lenders. This means that lenders are more willing to offer longer-term funding and are less likely to refuse to roll over lending, even in a stressed market.

The typical terms to be affirmed during a repo transaction are transaction date, purchase date, repurchase date (if applicable), collateral (ISIN), the nominal value of collateral, the market value of collateral, purchase price, repo rate, or interest rate index and spread, the currency of purchase price, counterparty, buy or sell and the settlement account [64]. The terms of a repo are specified in Table A.1, at a minimum, the counterparties, the security bought and sold, the date of purchase and the term, the repo rate, and the haircut should be included in an agreement or contract.

Counterparties	A repo has two counterparties: the seller and the buyer. The seller is so-called because it begins the repo agreement by selling a security. If the repo agreement is interpreted as a collateralized loan, the seller is the cash borrower and the buyer is the cash lender.
Purchase date	The date on which the purchase of securities occurs. This date is also called the settlement date and the first leg.
Repurchase date	The date on which the repurchase of securities occurs. This transaction is also called the maturity date and the second leg. The repurchase date can be left open and determined later.
Collateral	The security sold on the purchase date. This security is called collateral.
Equivalent securities	The security repurchased on the repurchase date. Securities are equivalent if they are (i) of the same issuer; (ii) part of the same issue; and (iii) of an identical type, nominal value, description, and amount as the collateral. If the collateral matures before the repurchase date, the equivalent security is the sum of money equivalent to the proceeds of the redemption.
The purchase price (of value)	The cash paid for the collateral. The purchase price is also called the value of the repo. This number is the notional value that appears on a firm's balance sheet.
The repuchase price	The cash paid for the collateral repurchase.
The repo rate	The interest rate paid on the cash to compensate the cash lender for the collateralized loan.
Haircut	The difference between the purchase price and the market value of the collateral. A haircut protects repo buyers against some market risk should they have to liquidate the collateral.

Table A.1: Terms of the repo contract. Source C. Garriott & K. Gray (2016)[7]

A.3.1 Repo Market in Numbers

The repo market falls under the secured money market segment, which is the largest segment of the euro money markets, accounting for 60% of total daily transaction volumes and 51% of total outstanding amounts [25]. In a secured loan transaction, the borrower has to provide collateral to the lending counterparty to secure the loan. In an unsecured money market transaction, the lending counterparty does not require collateral for the liquidity provided [65]. Government bonds are the dominant type of collateral in the euro-denominated repo market, accounting for 85% of all transactions, and most-used government bonds are issued in the biggest euro-area economies. Government bonds are more secure in the repo market due to their lower default risk, the creditworthiness of the issuer, and high liquidity in financial markets. Additionally, their stable interest rates and, acceptance as collateral by central banks contribute to the appeal in repo transactions of the use of government bonds. Given the size, breadth, and depth of the respective sovereign debt markets, bonds issued by governments in Germany, France, Italy, Spain, the Netherlands, and Belgium account for 95% of the total public collateral used in secured transaction [25].

The daily volume of the tri-party repo market according to the Federal Reserve Bank of NY (as of March 9, 2022) is $3.6 trillion dollars [66]. From a survey of the ICMA group the European Repo market in 2021 was around 9 billion in the month of December [67]. This represents that the repo market around the world is a significant portion of the overall financial market and highlights the importance of the repo market in providing liquidity to the financial system. Most transactions in the secured money market are concentrated in the one-day maturity bucket [27]. Where the overnight and maturity of one day (spot/next and tomorrow/next) are the majority of secured transactions, under which the repo market falls, see left side Figure A.1. The repo market is mostly overnight because it allows participants to borrow and lend with a short-term horizon, typically lasting one business day. This short-term nature provides flexibility and ensures that both borrowers and lenders can manage their liquidity needs and adjust their position daily based on changing market conditions and financial requirements.

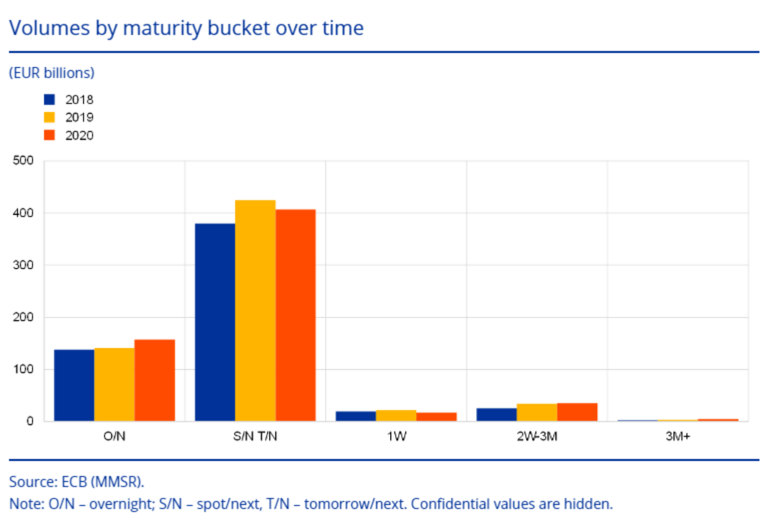


Figure A.1: Overview of the European secured money market maturities over the time period 2018-2020. Source: Euro money market study 2002 (ECB) [27]

According to the ECB, the repo market turnover tends to drop at quarter-end and year-end, for the purpose of window dressing [27]. Close to regulatory reporting dates banks reduce the size of their balance sheets in order to comply with prudential regulation and to optimise bank levies. During these periods a general reduction of borrowing activity is therefore observed when banks' leverage ratio is close to the regulatory minimum and/or they prefer to increase netting. This can also be seen in Figure A.2. Where the closer investors are toward the end of a year, the higher the funding is. Market participants have increased significantly since 2017 (see Figure A.2) and has been cited by market participants as one of the reasons behind the less stressed year-ends in recent years (up to year-end of 2020).

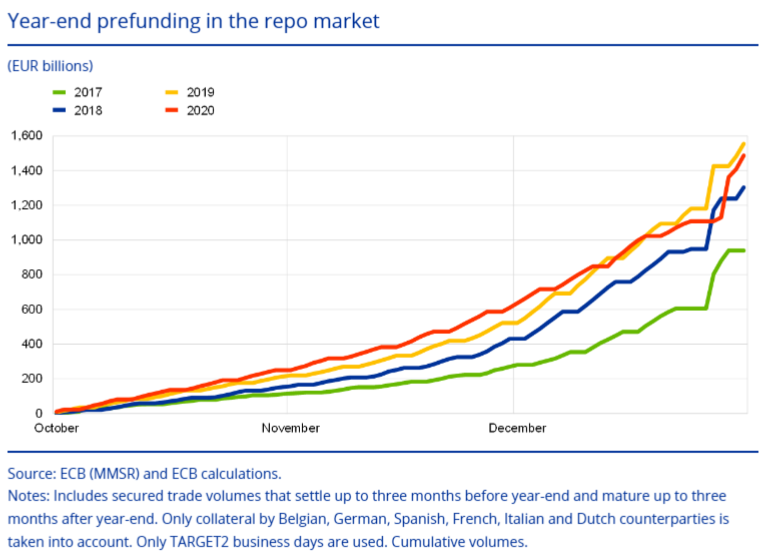


Figure A.2: Visualisations of the year-end funding activity in the European repo market. Prefunding consists of conducting trades that span the year-end reporting date but are agreed upon well ahead of the settlement date. Source: Euro money market study 2002 (ECB) [27]

A.3.2 Liquidity

To understand rollover risk first liquidity needs to be explained as it is a part of liquidity risk. Liquidity refers to the ease and speed with which an asset or security can be bought or sold in the market without significantly impacting its price. It represents the degree of market activity and the ability to convert an asset into cash quickly and at a fair price. High liquidity is desirable as it provides efficiency, facilitates smooth transactions, and reduces the risk of price volatility [68].

A well-functioning repo market also supports liquidity in other markets, thus contributing to the efficient allocation of capital in the real economy. However, excessive use of repos can also facilitate the build-up of leverage and encourage reliance on short-term funding [69]. When liquid, financial markets provide efficient, reliable pricing of financial assets, and funding can be easily obtained at a fair price [70]. Liquidity consists of two main pillars: market liquidity and funding liquidity. Market liquidity may be defined as the ability to rapidly execute large transactions at a low cost and with a limited price impact. Funding liquidity indicates the ease of borrowing conditions and capital flows in global financial markets. Market liquidity has several dimensions which are closely, but not solely, related to trading conditions in financial markets (see Figure A.3). Market liquidity has several dimensions which are closely, but not solely, related to trading conditions in financial markets (see Figure A.3). Here the ECB defines rollover risk as: "If a bank relies to a large extent on short-term funding (e.g. via repos), it runs the risk that it might be unable to replace or roll over maturing short-term funding or that it might only be able to do so at significantly higher costs"[70]. Moreover, funding liquidity also includes haircut/margin risk (see Figure A.3) which surfaces if a financial intermediary is funded via a secured loan (e.g. a repo transaction) and the haircut or margin applied to the collateral increases. In any of these cases, the financial firm might be forced to exit positions or liquidate assets. Since early 2022, market liquidity and funding liquidity have declined simultaneously as volatility has risen. Lower funding liquidity largely reflects the environment of higher interest rates. At the same time, market functioning has remained orderly, keeping rollover risk contained [70].

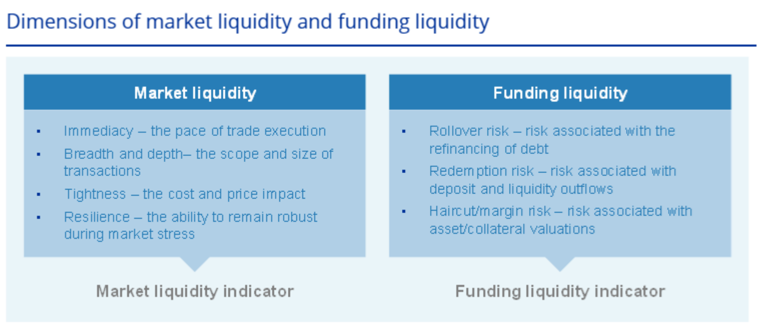


Figure A.3: Market and Funding liquidity risk. Source: ECB [70]

Funding liquidity risk is a risk where the firm is unable to roll over its maturing debt obligations with immediacy at some point in the future.

A.4 Public Data Figures

In this section are the extra features from section 3.5 visualized over time. All the data was downloaded on 03-05-2023. Both ESTR Figures A.4 and A.5 show a very deep peak around the end of 2021 and the beginning of 2022, indicating hints of seasonal adjustment or window dressing. The same peak can be seen at the end of 2022/beginning of 2023. In Figure A.6 the Secured Overnight Financing Rate of the USA is shown over the years 2015 up to May 2023. Around the end of February 2020, the rate deeply declines to almost zero until February 2022. This declined period is around the same period as the COVID-19 lockdowns in the United States. After that period the SOFR increases substantially. Figure A.7 of the bank rates is similar to Figure A.6 of SOFR in that around mid-2022 the rate starts to increase significantly.

In the Figure for inflation A.8 a different curve can be seen. The inflation rates rise from the beginning of 2021 up to the end of 2022, where it seems to steadily decline in the year 2023 so far.

For the Figure of the GDP Index A.9 the visualization is less clear than the ones above. The GDP Index is based on newspaper coverage frequency. Several types of evidence indicate that the index proxies for movements in policy-related economic uncertainty [34].

Lastly, Figure A.10 shows the Financial Stress Index. The OFR FSI is positive when stress levels are above average, and negative when stress levels are below average. From half of the year 2020 up until around February 2022 the stress level is negative and thus seems to indicate no disruptions in the normal functioning of financial markets. From that point onward are the financial markets in a more stress level, this is also around the time that Russia invaded and occupied parts of Ukraine (24 February 2022), indicating that the war could be a contributing factor to the FSI level.

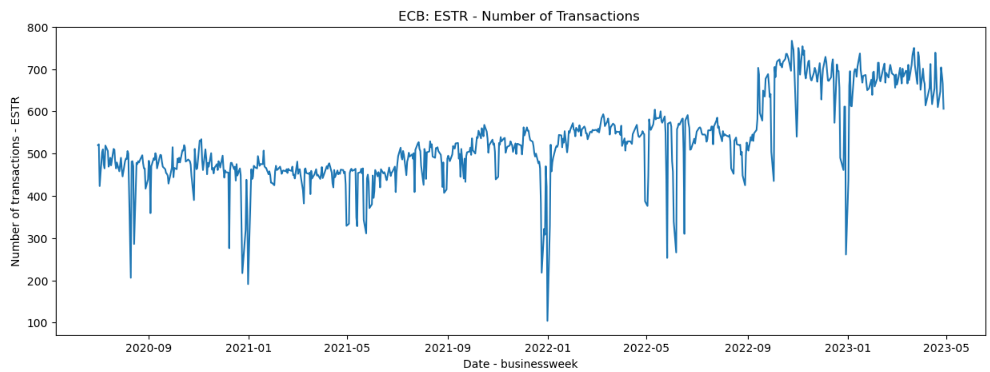


Figure A.4: Visualisations of the EST rate - number of transactions. Source ECB [32]

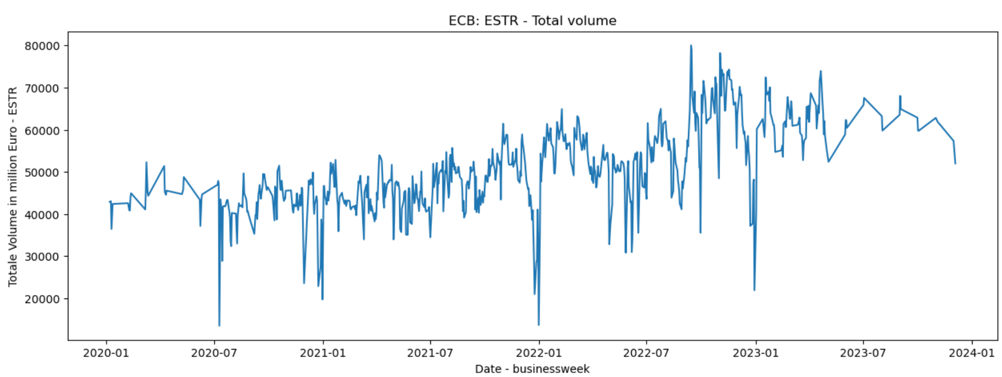


Figure A.5: Visualisations of the EST rate - Total volume. Source ECB [32]

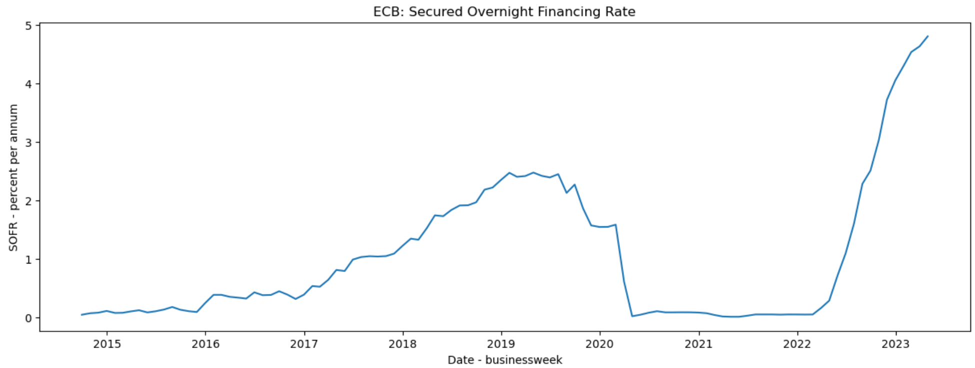


Figure A.6: Visualisations of the SOFR. Source ECB [32]

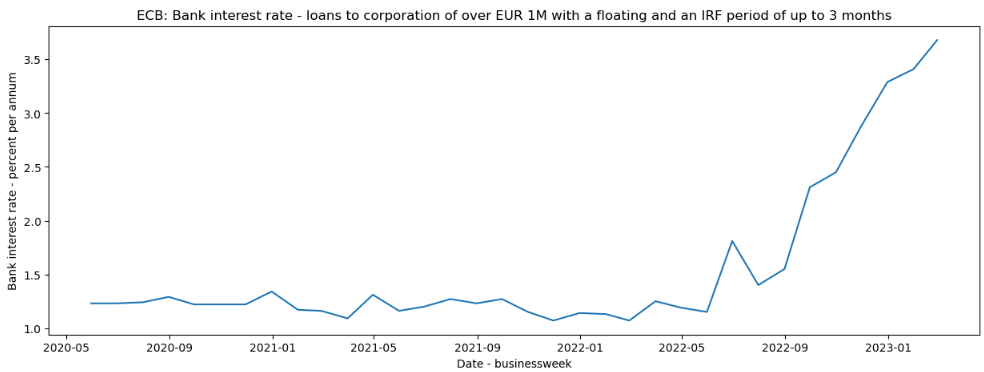


Figure A.7: Visualisations of the bank interest rate - loans to corporations of over 1 million euro with a floating and IRF period of up to 3 months. Source ECB [32]

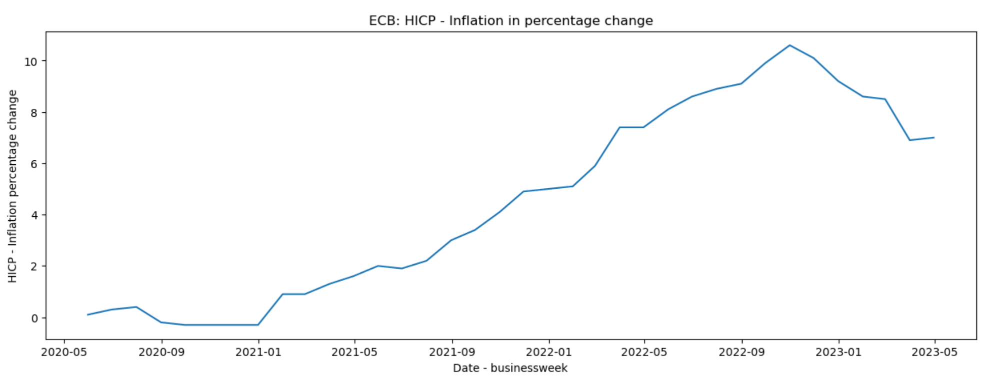


Figure A.8: Visualisations of HICP - overall index in the Euro area. Source ECB [32]

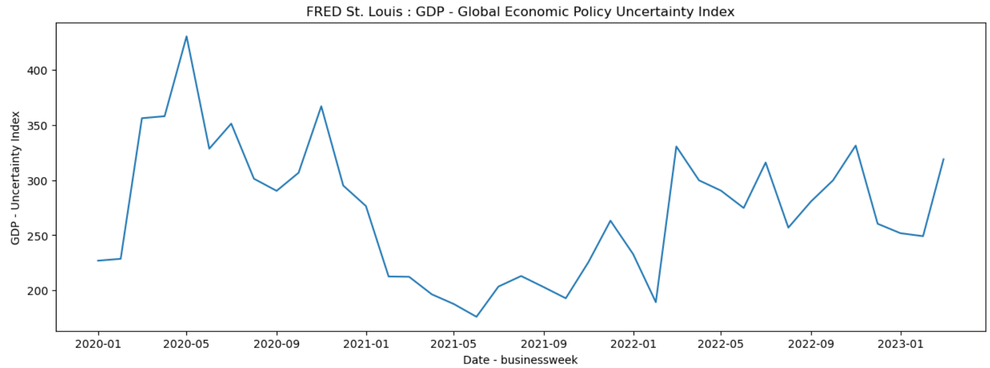


Figure A.9: Visualisations of the GDP weighted average index. Source FRED St. Louis[34]

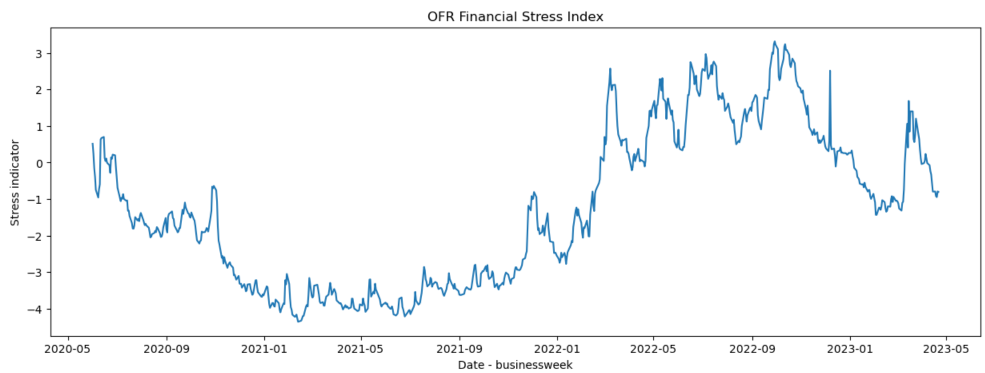


Figure A.10: Visualisations of the Financial Stress Index. Source OFR [33]

A.5 Additional Information on Repo Contracts

Types of repo contracts	
General collateral (GC) repo	A repo quoted for any collateral on a menu of securities. A common kind of GC repo is the GC repo, for which the collateral can be any bond on a particular list of issued bonds. Another common kind of GC repo is the mortgage-backed securities (MBS) GC repo, for which the collateral can be any security on a list of MBS.
Specific/special repo	A repo for which the desired collateral is a single, named security. If the repo rate for a specific repo is substantially less than the rate for a GC repo, the collateral of a specific repo is said to be "on special". A rate substantially less than the GC repo rate incentivizes the securities lender to send since it can obtain cash at a cheaper rate than the market.
Open-term repo	A repo for which the repurchase date is not identified and will be set later through the right of an early return.
Closed-term repo	A repo for which the repurchase date is identified. Most repos, even overnight repos, are closed in a term.
Evergreen repo	A term repo that automatically renews its term until either counterparty gives notice, at which point it becomes a term repo with the set term. Seen differently, it is an open-term repo with an extended notice period.
Floating-rate repo	A repo for which the repo rates are linked to a benchmark plus or minus a spread.
Sell/buybacks or buy/sellbacks	Sell/buybacks are an older-style repurchase agreement. They differ from repo in two main ways: • The interest payments are all embedded in the repurchase price. • The manufactured payments are not paid until the repurchase date.

Table A.2: Types of repo contracts. Source C. Garriott & K. Gray (2016)[7]

Participants in the repo market	
Dealers	Repo originated as a means for dealers to finance long or short positions they acquired during the normal course of their market-making activities. Without repo, dealers had to borrow either via loans in the unsecured market or by issuing commercial paper. Unsecured credit is expensive, and repo gave dealers a flexible means to access secured financing.
Leveraged investors	Investors who wish to develop levered positions can use repo to fund a set of asset purchases by using the purchased assets as repo collateral. Since repo is collateralized, it is a cheaper way to achieve leverage than unsecured sources. In Canada, pension funds and insurance companies may use repo to achieve leverage.
Prime brokers	Prime brokers offer securities lending and repo services to their clients as part of their value-added services. They may repo clients' investments (held by the broker) to provide clients with cheaper funding.
Repo conduits	A repo conduit is a special-purpose vehicle whose purpose is to lend to other investment firms via repo. The conduit finances itself by issuing asset-backed commercial paper secured by the collateral the conduits receiver as it lends. Repo conduits often make money by accepting illiquid or unusual collateral that is difficult to repo otherwise.
Risk-averse cash investors	Repo offers investors who have cash an opportunity to earn interest on the cash without taking on much risk, as repo is collateralized. Money-market mutual funds are an example of risk-averse cash investors.
Central banks	Central banks can use the repo market as a way to manage the cost of quantity of credit in an economy. Repo is a preferred tool in many central banks because of the size of the repo market and because it reduces the credit risk being taken with public funds.

Table A.3: Participants in the repo market. Source C. Garriott & K. Gray (2016)[7]

A.6 Unoptimized Results Machine Learning classifiers

All the unoptimzied results of the Machine Learning classifiers, run with all 34 features and default parameters.

GNB

	Precision	Recall	F1-score
No	0.78	1.00	0.88
Yes	0.00	0.00	0.00
Accuracy			0.78
Macro avg	0.39	0.50	**0.44**
Weighted avg	0.61	0.78	0.69

Table A.4: Classification result for the unoptimized GNB classifier.

GNB		Actual	
		Negative	Positive
Predicted	Negative	1.000	0.000
	Positive	1.000	**0.000**

Table A.5: Unoptimzed GNB normalized confusion matrix

LR			
	Precision	Recall	F1-score
No	0.78	1.00	0.88
Yes	0.00	0.00	0.00
Accuracy			0.78
Macro avg	0.39	0.50	**0.44**
Weighted avg	0.61	0.78	0.69

Table A.6: Classification result for the unoptimized LR classifier.

LR		Actual	
		Negative	Positive
Predicted	Negative	1.000	0.000
	Positive	1.000	**0.000**

Table A.7: Unoptimzed LR normalized confusion matrix

DT			
	Precision	Recall	F1-score
No	0.89	0.89	0.89
Yes	0.60	0.62	0.61
Accuracy			0.83
Macro avg	0.75	0.75	**0.75**
Weighted avg	0.83	0.83	0.83

Table A.8: Classification result for the unoptimized DT classifier.

DT		Actual	
		Negative	Positive
Predicted	Negative	0.987	0.013
	Positive	0.379	**0.621**

Table A.9: Unoptimzed DT normalized confusion matrix

SGD			
	Precision	Recall	F1-score
No	0.79	0.94	0.86
Yes	0.30	0.09	0.14
Accuracy			0.76
Macro avg	0.55	0.52	**0.50**
Weighted avg	0.68	0.76	0.70

Table A.10: Classification result for the unoptimized SGD classifier.

SGD		Actual	
		Negative	Positive
Predicted	Negative	0.925	0.075
	Positive	0.854	**0.146**

Table A.11: Unoptimzed SGD normalized confusion matrix

RF			
	Precision	Recall	F1-score
No	0.92	0.97	0.95
Yes	0.64	0.37	0.46
Accuracy			0.91
Macro avg	0.78	0.67	**0.71**
Weighted avg	0.89	0.91	0.90

Table A.12: Classification result for the unoptimized RF classifier.

RF		Actual	
		Negative	Positive
Predicted	Negative	0.974	0.026
	Positive	0.634	**0.366**

Table A.13: Unoptimzed RF normalized confusion matrix

XGB			
	Precision	Recall	F1-score
No	0.89	0.95	0.92
Yes	0.75	0.56	0.64
Accuracy			0.86
Macro avg	0.82	0.76	**0.78**
Weighted avg	0.86	0.86	0.86

Table A.14: Classification result for the unoptimized XGB classifier.

XGB		Actual	
		Negative	Positive
Predicted	Negative	0.948	0.052
	Positive	0.438	**0.562**

Table A.15: Unoptimzed XGB normalized confusion matrix