Article

An Analysis of Residual Financial Contagion in Romania’s Banking Market for Mortgage Loans

Ştefan Ionescu, Nora Chirită, Ionuţ Nica and Camelia Delcea *

Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, 010552 Bucharest, Romania
* Correspondence: camelia.delcea@csie.ase.ro

Abstract: The uncertainty of the environment, the complexity of economic systems, both at the national and global economy levels, and the digital age and artificial intelligence draw attention to the existence or appearance of systemic, disruptive phenomena that can appear and propagate in different forms, producing effects that can turn into economic crises. These phenomena can be transmitted like a domino effect, and they are referred to as the contagion effect in the scientific literature. In this research, one of the four forms of financial contagion, known as residual contagion, is studied on the mortgage loan market in Romania using agent-based modeling. By considering the economic crisis of 2007–2009, also supported by the mortgage crisis, in the present paper, we aim to study the Romanian mortgage market in 2022 through the use of machine learning techniques and agent-based modeling. The purpose of this research is to capture the potential systemic risks that can outline a residual financial contagion effect. The simulation results highlight the fact that the degree of connectivity between the commercial banks in Romania and the way in which they are interconnected have a major importance in the emergence and propagation of contagion effects. The proposed approach and the obtained results can offer more insight to policymakers on how the contagion effect takes place within the banking sector.

Keywords: financial contagion; mortgage market; machine learning; NetLogo

1. Introduction

The era we live in is driven by artificial intelligence, which transforms the way the fundamental sectors of contemporary society, namely the medical, automotive, economic, financial, marketing, and retail sectors operate, but also the way each individual’s life unfolds, as we interfere with this type of technology on a daily basis [1–3]. Furthermore, the occurrence of machine learning (ML) has succeeded in significantly changing every aspect of our lives. Currently, machine learning algorithms perform tasks that, not long ago, only experts had the ability to perform [3,4].

The main motivation of this paper is to analyze and observe the impacts that artificial intelligence (AI) and its branches, machine learning (ML) and Deep Learning (DL), have on contemporary society and on important world industries such as the banking industry [5,6]. The multitude of data and information collected by a bank in Romania can have a significant contribution to help us analyze and manage artificial intelligence, as well as to build advanced risk models that predict, quantify, and appropriately estimate its associated risks. Technological innovations change the course of our lives and have a major impact on all economic systems [6,7]. The departments within the commercial banks in Romania use these AI technologies in the analysis and design of risk models related to lending products, including mortgages. For example, Romanian banks such as Banca Transilvania or ING Bank use artificial intelligence at the levels of several departments, collaborating with UiPath, which is the global company that deals with the development and automation of robotic processes, including Microsoft [8–10]. The banks also use these techniques to make
a forecast on the performance of the model based on some residual values evaluated based on the traffic light technique. Changes that exceed these values and fall on the red zone of the traffic light, meaning that the estimated value is lower than the real value, considering the residual valuation, can outline a contagion effect [11,12].

The banking network in Romania continues to adjust as we live in an era of digitization, and banks try to continuously adapt to the requirements of economic agents and of the environment. This response is a result of the complex behavior that banks exhibit as complex adaptive systems that co-evolve and organize themselves with regard to be able to offer as many banking services as possible in the new digital environment, which sometimes translates to giving up to physical agencies. This process plays a double role, as besides co-evolution and adaptation, it contributes to optimizing certain expenses of an operational nature.

The banking and financial market is influenced by major risks, including environmental issues, climate change, pollution, and other natural disasters that can impact the economy. These events can lead to financial losses in a chain reaction, affecting companies, investors, banks, and the financial market. Therefore, in order to address and mitigate these risks, banks need to adopt a sustainable approach to minimize the negative impact of potential financial contagion. Banking institutions’ integration of ESG factors (Environmental, Social, and Governance) in investment decisions and in addressing risks and opportunities related to sustainability, as mentioned earlier, enables investors to make informed decisions and reduce the risk of financial contagion. On the other hand, regulations and government policies regarding sustainability can also have a significant impact on the occurrence of financial contagion. Implementing coherent and efficient measures to promote sustainable development can reduce systemic risks as well.

From this point of view, the current paper also considers the cybernetic approach to the analyzed topic. This helps us to better understand both the transmission channels of the effects of financial contagion, as well as the impact that a faulty management of the cybernetic subsystem of credits can have. Considering this standpoint, agent-based modeling, which is specific to economic cybernetics, can be used to simulate the effect of residual banking contagion by creating a system of agents that represent individual banks and their interactions in a financial environment, which will be addressed in a practical section of the case study within this research.

The Commercial Bank, from the cybernetic perspective, is a complex adaptive system analyzed from the perspective of five subsystems [13]: credit management, deposit management, risk management, treasury management, and fund insurance management. The cybernetic approach helps us to understand the transmission channels of systemic risks in the banking sector by highlighting the eloquent properties of complex adaptive systems: co-evolution, adaptability, interconnection, and emergence or feedback mechanisms. Compared to the other sectors, the financial sector is different as it depends on people and their behaviors, to which it provides services. Banking is an area where ML procedures are applied to complex databases to combat financial fraud and cyber-attacks, save customers’ time and money, and streamline back-office activities [14]. With the development of AI, currently, banks use AI chatbots to respond to customer requests. In the financial management and advisory services sector, firms are looking at the potential of AI solutions to improve possible speculative solutions. Thus, ML has become necessary in many areas of the financial system. Among the various applications of ML in the financial system, it has been observed that one of the most important applications is fraud detection [15,16]. Another important application implemented in the financial sector is the verification of the conditions necessary to grant loans or insurance to customers [17–20].

This process can be described as ideal for ML in this sector. Calculations made by artificial intelligence can provide important predictions about the status of each customer and the conditions that they have fulfilled in order to be granted a loan.

In Romania, it has been observed that the mortgage market has registered a significant increase in recent years, according to the Stability Reports published by the NBR [21]. From
the perspective of the use of artificial intelligence models in banks in Romania, this has become more and more popular, with these algorithms being used in various fields, such as risk management, credit management, data analysis, and customer behavior analysis [22–24].

Given the importance of ML algorithms in the current era, on one hand, this paper aims to study the Romanian mortgage market in 2022 through the use of ML techniques for the purpose of better capturing the potential systemic risks that can outline a residual financial contagion effect. On the other hand, considering the confidential nature of the real data and information held by the bank, simulating a financial contagion effect is quite challenging. For this reason, this research employs agent-based modeling specific to economic cybernetics to simulate the residual contagion effect.

This paper is organized as follows: Section 2 provides a discussion related to the state of the art in the field from the perspective of the contagion phenomenon in the financial sector. Section 3 presents in detail the financial stability of Romanian banks, making a comparison between the values recorded for the prudential indicators in Romania compared to the European Union (EU) average. Section 4 presents the methodology used in this study, while Section 5 discusses the results with an accent on residual financial contagion. The paper ends with concluding remarks and future work directions.

2. The Stage of Knowledge in the Field

Until 1997, the concept of contagion meant the spread of a medical disease, as the association of the phenomenon with banking turbulence did not exist before that year [20]. According to the “Merriam-Webster” dictionary [25], contagion is associated with “(i) a contagious disease; (ii) transmission of a disease through direct or indirect contact; (iii) a disease-producing agent such as a virus; (iv) poison or contagious influence that spreads rapidly; (v) the rapid communication of an influence (such as a doctrine or an emotional state)”.

On the other hand, the Cambridge dictionary [26] associates the term contagion with the following meanings: “(i) the context in which a disease spreads by touching someone or something; (ii) the situation in which feelings, ideas or problems spread from one place to another; (iii) economic problems in a country, region, etc., they spread to another”.

The Oxford Learner’s Dictionary [27] describes the phenomenon of contagion as “(i) the spread of a disease through close contact between people; (ii) something bad that spreads rapidly by transmission from one person to another”.

From the definitions identified in the above dictionaries, it can be observed that some essential information can be extracted that connects the effect of contagion with the behavior exhibited by complex adaptive systems. For example, it can be observed that the phenomenon of contagion spreads through direct or indirect contact, which illustrates the emergent properties, interdependence, and connectivity of complex adaptive systems. Moreover, connectivity is emphasized by describing the analyzed phenomenon as “the spread of a disease through close contact between people”, illustrating the importance of the degree of connectivity.

Zheng [28] states in his research that, in the modern financial system, the division of labor and the close cooperation between different departments are determinants of the risk contagion. The authors show that the risk from stock market volatility can easily spread to other industries and departments, leading to systemic financial and economic crises [27]. Compared to the previous research, which has mostly examined financial security from the perspective of macro currency security and the security of the banking system, but has ignored its underlying problems, such as the limited rationality of investors due to a lack of financial knowledge, the work conducted by Zhang [28] analyzes micro-level investor behavior, meso-level stock market price fluctuations, and macro-level financial security in a unified framework.

Currently, ML has become necessary in many areas of the financial system, having the most important application in fraud detection. Mathur [29] predicted that the year 2027 will be one in which more and more systems will use AI, and the applications of this technology will impact the whole world, together with its important industries, including the financial
sector. This scenario features a 27-year-old entrepreneur named Godson, who works in community offices that are close to his home and is assisted by robots and Personal Drones (PA) to accomplish his work [29]. The PA drone communicates the day’s priorities to him daily. Also, at Godson’s request, the PA drone connects with the TradeBot, which is the robot that keeps track of all of Godson’s financial transactions, to present him with the monthly report on his stock market investments. The automated bots based on artificial intelligence notify Godson of potential losses and provide predictions and advice on how he can act so that losses are minimized and profit is maximized. The robots can also accept input from Godson when requested, so that processes performed by them are processed based on the manual input given by Godson [29]. These bots use algorithms that predict possible falls or rises in stock prices based on whatever information they have access to. Another innovation that artificial intelligence can introduce in the future is communication between robots and the creation of an ecosystem of robots, where they can interact. Even if in the year 2023 this scenario seems like something that is still impossible to achieve for the year 2027, we believe that it will for sure be possible in the near future.

Fares et al. [30] developed research in which they analyzed, from a holistic perspective, the specialist literature on how artificial intelligence algorithms are used in the banking sector. The analyzed period was 2005–2020. The authors examined 44 research works using a systematic review and a content analysis of the topic as their methods. The research conclusions were focused on how the banking sector expands on three key research areas in which AI algorithms are used: customer, strategy, and processes. Ris et al. [31] present in their research the fact that with today’s technological advances, with the automation of processes, banks must adapt to the evolution in order not to lose the opportunity to transform their business models and to prevent certain risks such as bank fraud. They concluded that the automation process in the banking sector, along with the use of AI algorithms, can improve the performance of the banking business process.

More and more, research is emphasizing the importance of using AI models in the analysis process regarding the granting and monitoring of loans. However, according to the research carried out by Sadok et al. [32], it is emphasized that certain limitations regarding ethical, legislative, and legal issues must also be considered.

The phenomenon of financial contagion is gaining the attention of more and more supervisory institutions in Romania and is analyzed in increasing numbers of scientific articles. Although there is a low probability of the manifestation of the risk of contagion, it can appear at any time, considering that a minor shock, untreated, can spread systemically in the banking network and have major effects. Trenca and Dezsi [33] examine in their paper the behavior of the Romanian stock market from the perspective of the formation of the phenomenon of financial contagion in the global market. They used the MS-VAR model on three countries, between 1997 and 2012, and their results conclude that during the financial crisis, no financial contagion effects were observed on the Romanian market due to the easy integration of the Romanian stock market with the world market.

In another research, Carausu [34] used a wavelet analysis in order to analyze whether there was a financial contagion effect between the capital market in the US and Romania at the beginning of the economic crisis of 2007–2009. They discovered that the capital market in Romania and the US were in perfect synchronization during the crisis for large transaction periods, meaning that the Romanian market was sensitive to short-term systemic shocks from the US market. This aspect could have outlined a fundamental long-term contagion.

On the other hand, Tilica [35] studies the presence of the day-of-the-week (DOW) effect in the process of forming the financial contagion effect that can be observed in individual economic sectors on the post-communist markets of Eastern Europe. The author noticed that the markets that offer specific national sector indices from the financial crisis period of 2007–2009 are Romania, Russia, and Poland. He used the GJR-GARCH model on a regression model that considers both the crisis period and working days.

Bookstaber [36] carried out research in which he described how the agent-based approach can be used in modeling financial crises. It clearly shocks the agents’ interactions
and the way they respond to changes in the environment. Practically, Bookstaber believes that this explains the phenomenon of contagion and the cascades that can appear due to financial leverage, the concentration of the market formed by agents, and because of the liquidity factor. In addition, Bookstaber [36] highlights how agent-based modeling overcomes the limitations of traditional economic models when analyzing financial crises.

Systemic risk and financial contagion are closely related concepts that refer to how problems in a financial institution or financial sector can impact and spread throughout the entire financial system [37]. Financial contagion is the phenomenon in which issues or disruptions in a financial institution or financial market can transmit and affect other financial institutions or economic sectors. It can have a negative impact in the form of a chain reaction, spreading from one institution to another and potentially leading to the collapse of the entire financial system. Systemic risk is associated with the possibility that an event or a series of events can cause significant disruptions throughout the financial system. This risk arises when there are interconnectedness and interdependence among financial institutions and financial markets, so that problems in one specific location can rapidly propagate and affect the entire system [37–39].

Residual financial contagion refers to the secondary and lingering effects after an initial financial crisis has been resolved or when a financial institution experiences major problems. Even if the initial source of contagion has been addressed, the consequences can persist and spread throughout the financial system. For example, in the case of a bank collapse, there is a possibility that other banks or financial sectors may be negatively affected even after the initial crisis has been addressed.

Thus, systemic risk can lead to the effect of financial contagion, and the latter can result in residual financial contagion, amplifying and propagating the negative effects within the financial system. It is important to understand these connections and develop appropriate risk management mechanisms to prevent or mitigate the impact of financial crises.

Given the close connection between systemic risk and financial contagion, this study can make significant contributions to green practices and sustainability as well. Commercial banks in Romania provide loans to companies for green investments and financing, highlighting the importance of ESG (Environmental, Social, and Governance) factors. As this is a relatively new niche, the risks involved can be quantified differently, and the emergence of residual financial contagion risk can rapidly form and propagate if not addressed proactively. Therefore, through simulations conducted using agent-based modeling, the performance and resilience of financial institutions and green financing can be analyzed to identify the synergy between them. Currently, various studies address green practices and sustainability across different fields. For instance, Tsagkanos et al. [40] conducted a study that examines the relationship between corporate green bonds and commodities, employing methodologies such as Value-at-Risk-based copulas. On the other hand, Mahmood et al. studied how the perception and disclosure of corporate social responsibility (CSR) by firms determine their financial performance. According to Naeem et al. [41], financial markets are also exposed to extreme and uncertain circumstances that amplify tail risk, including sustainable, religious, and conventional markets. Therefore, they apply a neural network to identify religious and conventional investments with a high exposure to tail risk.

The Chinese government is actively promoting the issuance of green credit by commercial banks as a key aspect of China’s green finance system. Lian et al. [42], in their research, focus on the impact of green credit on the financial performance of banks, considering the diverse regional context of green development. Using panel data from 34 Chinese commercial banks between 2007 and 2018, the study empirically analyzes the effects of green credit on financial performance using a fixed effects model. The findings demonstrate that green credit has a positive impact on commercial banks’ financial performance, particularly through its effect on the rate of return on interest-bearing assets. Moreover, the level of green development enhances the economic benefits of green credit, and higher green economic growth and government support through environmental policies amplify the positive influence of green credit on banks’ financial performance. In conclusion, commercial banks...
are encouraged to actively engage in green credit activities, while the government should strengthen policies that incentivize green credit to foster the coordinated development of green finance and a sustainable economy.

3. The Level of Financial Stability Banking in Romania

The degree of uncertainty on macroeconomic developments, both at the national and global levels, has increased substantially in the recent period of time [43]. This increase was supported by the outbreak of the armed conflict between Russia and Ukraine and by the major implications of the energy crisis [43]. In this context, the analysis of systemic risks is an important action for the supervisory authorities, which, in the case of Romania, is made by the National Bank of Romania, which publishes periodic financial stability reports [43].

Thus, according to the report on financial stability published in June 2022 [43], the National Bank of Romania (NBR) identified two severe systemic risks and two high systemic risks. These are presented in Table 1.

Table 1. Map of systemic risks regarding financial stability in Romania.

<table>
<thead>
<tr>
<th>Systemic Risk Event</th>
<th>Risk Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global uncertainties in the context of the energy crisis, the war in Ukraine, and the COVID-19 pandemic.</td>
<td>Severe Systemic Risk</td>
</tr>
<tr>
<td>The deterioration of internal macroeconomic balances, including deterioration as a result of geopolitical and international developments.</td>
<td>Severe Systemic Risk</td>
</tr>
<tr>
<td>The delay in reforms and the absorption of European funds, especially through the National Recovery and Resilience Plan.</td>
<td>High Systemic Risk</td>
</tr>
<tr>
<td>The risk of non-payment of loans contracted by the non-governmental sector</td>
<td>High Systemic Risk</td>
</tr>
</tbody>
</table>


For each risk presented in Table 1, the colors used represent the intensity of the risk and the arrows indicate its perspective for the next period.

One of the important risks for the banking sector remains the credit risk. Regarding the state of non-performing loans at the population level, the non-performing rate decreased insignificantly in June 2022 compared to March 2022, according to the NBR report [6]. This decrease may be a False Positive result as many debtors have appealed to GEO 37/2020 for the suspension of payment installments due to financial difficulties [44]. However, an improvement in the payment capacity was observed in the case of mortgage loans [45].

Over time, there have been several financial crises such as the Asian crisis, the South Sea Bubble, or the recent global crisis of 2007–2009. Many of these had the financial instabilities of the banks as the main starting point. Central banks play an important role in maintaining the financial stability of commercial banks [46–49].

Thus, the NBR ensures the design of a macroprudential supervision framework with the aim of limiting the frequency and impact of potential future crises, as well as the contribution made to the safeguarding of the financial system. Financial stability in Romania has become the main objective of the NBR since 1999 when the Financial Banking Policy Directorate was established and later, in 2004, with the establishment of the Financial Stability Directorate [50].

According to the Eurostat source, Romania ranks eighth in the European Union from the perspective of price stability. Even though inflation has grown strongly in the last three
years, Romania has gradually improved its position compared to other EU countries, as can be seen in Figure 1.

![Table: Risk indicators of the banking sector monitoring by European Banking Authority](image)

<table>
<thead>
<tr>
<th>Risk indicators of the banking sector monitoring by European Banking Authority</th>
<th>Romania</th>
<th>EU (average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1 equity ratio</td>
<td>18.8</td>
<td>16.6</td>
</tr>
<tr>
<td>The rate of non-performing loans</td>
<td>2.8</td>
<td>3.8</td>
</tr>
<tr>
<td>The degree of coverage with provisions a non-performing loans</td>
<td>65.9</td>
<td>43.8</td>
</tr>
<tr>
<td>Return on Equity (ROE)</td>
<td>15.5</td>
<td>7.9</td>
</tr>
<tr>
<td>Cost/Income</td>
<td>53.6</td>
<td>53.4</td>
</tr>
<tr>
<td>Loans/Deposits (population and non-financial companies)</td>
<td>69</td>
<td>110</td>
</tr>
</tbody>
</table>

Figure 1. The situation of prudential indicators in Romania. Source: Financial Stability Report [45].

Regarding the banking sector in Romania, according to the information presented in the Financial Stability Report [45] by the NBR, the financial and prudential position of the banking system in Romania is at an adequate level compared to the average recorded at the level of the European Union. However, the risks are increasing. As can be seen in Figure 1, the risk of non-payment of loans is a vulnerable subject, considering the deterioration of macroeconomic conditions, as well as the uncertainties of their future developments. The increase in inflation associated with the increase in the interest rate amplifies the vulnerability of this increase in non-performing loans. Also, even if the average is very high at the EU level regarding banks’ debts to the government sector and assets, falling within the worst prudential interval established by the EBA, the situation is not favorable in Romania either.

Even if the situation regarding financial stability in Romania is relatively stable at present, the uncertainties of the macroeconomic environment, the energy crisis, Russia’s aggression against Ukraine, and the low level of financial education in Romania still represent challenges that the Romanian economy faces. Although Romanian banks are considered stable and well capitalized, there are several concerns regarding the high level of non-performing loans, which has increased significantly in recent years. The supervisory authorities must maintain their increased interest in identifying solutions for the problems they face.

Considering these aspects, the analysis of contagion effects becomes a very important topic to analyze, including the high degree of uncertainty of the political, economic, and social contexts, as well as the existing vulnerabilities of the banking sector in Romania, emphasizing the current research.

The analysis of residual financial contagion contributes to economic growth through stabilizing and consolidating the financial system, restoring confidence in it, and developing appropriate policies and regulations to better prevent and manage financial crises. By creating a safer and more predictable environment for investments and economic activity, it can stimulate innovation, productivity, and sustainable growth. For example, by conducting this analysis, the vulnerabilities and remaining risks in the financial system can be identified. This enables authorities and institutions to take preventive measures to reduce the risk of recurrence of the contagion effect, thereby creating a more stable and secure environment for economic growth. Furthermore, regulations and policies can be improved by providing valuable insights learned from past effects. Thus, by strengthening regulations, the financial system can be stabilized, creating a favorable climate for investments and economic growth [38,51].
4. Methodology


ML is a field of computer science that is based on the use of statistical tests, techniques, and methodologies to ensure the ability to learn from past experiences and to improve the behavior of models, as well as the way they perform certain tasks, over time [52,53].

Considering the economic uncertainties faced by banking institutions, the standardized traditional methods used in the development of bank credit risk models are insufficient to face the new challenges. Thus, specific machine learning algorithms are useful, beneficial tools for banking institutions in Romania. By using them, more accurate and efficient models can be developed to help us evaluate the credit risk and to reduce the high risk of default that Romania has been facing in recent years. Even if these algorithms have proven to be more accurate than traditional credit risk assessment models, we cannot say that they represent the perfect solution because there are also certain risks related to the input data used and the historical data on the basis of which the models were developed, which, in time, may no longer have a high discrimination power [53,54].

In the following, the most used ML algorithms in the development of credit risk models will be briefly presented, which will be employed in analyzing and evaluating the predictive ability of some debtors to repay the second mortgage loan based on indicators that are influenced by customer behavior, such as the debt held for the first mortgage loan and the value of the loan. The algorithms are presented through a brief description and through the needed implementation in Python for ensuring the reproducibility of the results obtained through the use of these algorithms.

Random Forest—a supervised learning algorithm that helps us build a set of decision trees. This algorithm is used most of the time for classification problems and can provide a relative importance of the variables used in the prediction [55,56]. To implement this model in Python, the following lines of code were used, according to Figure 2:

```
# Define model
rf_clf = RandomForestClassifier(random_state = 0)
```

Figure 2. Developing the Random Forest model.

Basically, the estimator is instantiated in a control object variable. The random_state argument receives the value 0 so that with each run of the model, the results remain consistent.

In the second step, we defined the set of hyperparameters, as can be seen in Figure 3.

```
# Define hyperparameters
param_grid = [{'max_depth': [5, 10, 50, 100], 'max_features': ['auto', 'sqrt'],
               'n_estimators': [100, 200, 500], 'min_samples_split': [5, 10, 20, 100],
               'min_samples_leaf': [5, 10]]
```

Figure 3. Defining the hyperparameters for Random Forest model.

The third step, which can be seen in Figure 4, consisted of adjusting the hyperparameters and training the model on the training set. Within this stage, the k fold cross-validation process was also performed. At the end of the training stage, the best tuned hyperparameters, the best accuracy score, and the best estimator, resulting from the hyperparameter tuning and model training, were displayed. We also added a method that calculates the running time of the model on the training set, which will be displayed in a later output.

In the last step, we evaluated the accuracy and performance of the model by calculating the indicators for performance measurement, to which the confusion matrix and the classification ratio were added. Figure 5 illustrates the last described step necessary to evaluate the performance of the Random Forest algorithm.
K-Nearest Neighbors (KNN)—a supervised learning algorithm based on the idea that new data instances are classified based on the classes of their neighbors in the data set [57].

The purpose of this ML algorithm is to search for the nearest neighbors of a known query point in order to be able to assign a class label to that point. In this sense, the KNN algorithm considers the following steps [57].

The distance values are determined. The distance between the interrogation point and the other points is calculated. Based on these calculations, we determine the decision boundaries in different regions.

To determine the distance metrics, several measures [58] can be applied such as Euclidean distance, Manhattan distance, Minkowski distance, and Hamming distance.

The formula to calculate the Euclidean distance is

\[ (a, b) = \sqrt{\sum_{i=1}^{n} (b_i - a_i)^2} \]  \hspace{1cm} (1)

Another well-known distance metric is the Manhattan distance (MD). This measures the absolute value between two points and has the following formula:

\[ d(a, b) = \left( \sum_{i=1}^{m} |a_i - b_i| \right) \]  \hspace{1cm} (2)

The Minkowski distance \( (M_p) \) is based on the Euclidean and Manhattan distance metrics, being their generalized form. This can be confirmed by replacing the parameters \( p \) in the formula below with 2 to obtain the Euclidean distance, and \( p = 1 \) is the Manhattan distance.

\[ M_p = \left( \sum_{i=1}^{n} |a_i - b_i|^{\frac{1}{p}} \right)^{\frac{p}{p}} \]  \hspace{1cm} (3)

When working with Boolean vectors or strings, the Hamming distance is used. This can be represented by the following formula:

\[ D_H = \left( \sum_{i=1}^{k} |a_i - b_i| \right) \]  \hspace{1cm} (4)

\[ a = b; D = 0 \text{ or } a \neq b; D \neq 1 \]  \hspace{1cm} (5)
Logistic Regression—a supervised learning algorithm that uses the sigmoid function to make a binary classification. Basically, the sigmoid function is a mathematical function that transforms continuous values into values between 0 and 1 [59,60]. This can be represented by the following formula [61,62]:

\[
F(x) = \frac{1}{1 + e^{-x}} = \frac{e^x}{e^x + 1}
\]

Logistic regression is based on the probability that a sample belongs to a class. The main feature of this probability must be continuous and bounded between (0, 1). The graphical representation is shown in Figure 6.

![Figure 6. Representation of the logistic/sigmoid function.](image)

Extra Trees Classifier—an algorithm similar to Random Forest. The difference is that it uses the entire data set and builds the decision tree by randomly selecting features and cut points.

This algorithm can be an efficient method for reducing the variant and improving the generalization [63].

Naïve Bayes—a supervised learning algorithm that is based on Bayes theorem. Naïve Bayes can be used for classification problems and is effective when the independent variables are few and highly correlated [64,65].

AdaBoost—a supervised learning algorithm that tries to improve the performance of a weak classifier by successively building stronger classifiers. In each round, a higher weight is given to the wrong examples so that the next strong classifier focuses on these examples. AdaBoost can be used for classification problems [64].

XGBoost—an algorithm that is an extended implementation of Gradient Boosting Machine (GBM), which is another supervised learning algorithm used for classification and regression problems. XGBoost uses an additive model where new classifiers are added to correct the mistakes of previous classifiers [65,66].

LightGBM—an open-source algorithm package, patented in 2016 by Microsoft and, like the original Gradient Boosting, uses decision trees to solve classification problems as well as other specific machine learning problems [67].
Hyperparameters are used to control the behavior of the automatic learning model and to optimize its performance in order to reduce the risk of residual model [68]. These are the configuration parameters of the algorithm that are not learned by the model, but are specified by the user before training the model. If the hyperparameters are set incorrectly, then the performance of the model may be sub-optimal. Choosing the right hyperparameters can improve model performance and reduce the risk of overfitting or underfitting the model. Thus, for each model analyzed in this research, hyperparameter tuning was applied.

To test the performance and accuracy of the models on the testing set, a series of performance indicators can be applied that measure the accuracy of models and provide results about their performances on the predictions on the testing set.

The confusion matrix [69,70] contains the total number of predicted values and actual values. It is a very popular metric used to solve classification problems (Table 2). This one can be applied to binary classifications, but also to multi-class classification problems. Classification models have multiple categorical outputs. Most indicators that measure prediction errors calculate the total error in the model but cannot find the individual cases of error in the model. The model may misclassify more categories than others, but we cannot see this using a standard indicator that measures accuracy.

Table 2. Confusion matrix.

<table>
<thead>
<tr>
<th>Actual Value</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Value</td>
<td>True Positive (TP)</td>
<td>False Positive (FP)</td>
</tr>
</tbody>
</table>

As can be seen, the confusion matrix is a quadratic matrix, with its size being given by the number of classes predicted by the model. It contains four important quantities: True Positive, True Negative, False Positive, and False Negative.

The True Positive size shows the number of examples correctly classified in the positive class and the True Negative size shows the number of examples correctly classified in the negative class. The other two quantities show the number of examples wrongly classified in the positive and negative classes, respectively.

We will use the confusion matrix to calculate a series of performance metrics, such as accuracy, recall, precision, F1 score, and ROC curve, of the selected classification models.

The calculation formulas for these metrics are as follows [70]:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{8}
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{9}
\]

\[
F1 \text{ Score} = \frac{2 \frac{\text{Recall}}{\text{Recall} + \text{Precision}}}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}} \tag{10}
\]

The ROC curve illustrates the relationship between the True Positive rate (TP) and the False Positive rate (FP) at different decision threshold cut points.
4.2. Agent-Based Modeling

Commercial Bank is an economic and financial agent that offers specific services and products, such as granting loans, deposits, customer management, and intermediary payments, both for individuals and legal entities [13,70,71].

Cybernetics, as defined by Norbert Wiener, is an interdisciplinary science that deals with the study of communication and control of complex systems in complex environments [72].

The economy, analyzed from a cybernetic perspective [70,72], is a complex system made up of complex agents that interact both with each other and with the environment, and these interactions can be understood based on the feedback mechanisms that are formed. This is made up of agents such as all companies, banks, the population, households, the government, and all the agents that form and define the national economy or the global economy, depending on how it is analyzed at the system level.

The Commercial Bank, analyzed as a cybernetic system, must be understood from the perspective of the function and interaction between its subsystems. Considering the central topic of the work, we will focus on the cybernetic subsystem of credit management, which represents the interface between the Commercial Bank and the Credit Market, but also realizes the interaction between the Bank and the environment through the services and products offered to clients. This interaction between autonomous agents, the bank’s customers, will be analyzed through agent-based modeling (ABM) with the aim of understanding both their behavior and the behavior of the banking system.

Agent-based modeling [73,74] is an efficient method that offers tools necessary to understand, simulate, and solve real-life situations so that we can understand the complex behavior of economic agents and the factors that influence their behavior, unlike the model based on equations. ABM is based on computational modeling that can incorporate different traits and characteristics of the agents, helping us to understand the interaction between them, which is an aspect that can be difficult to analyze or model if modeling based on equations is used.

Traditional mathematical models, such as those that use logistic models, the Lotka-Volterra model, or models based on differential equations cannot explain the behavior of agents in decision making or behavioral modification [75].

An agent-based modeling cannot work without considering the particularities of individual entities or agents. These particularities [76,77] refer to the properties of the agents, the behavior during the movement, the behavior during the change in direction, or the behavior in the interactions with the other agents.

There are many definitions in the literature formed by different researchers for agent-based modeling that are more or less similar, but their meanings are partly the same. Therefore, the meaning of agent-based models is given by models programmed in computers, using various mathematical or statistical formulas, with the aim of capturing the evolution and behavior of one or more individuals, but within a given environment, it is well established. There are different conditions depending on what a person wants to simulate or study [77,78]. The benefit provided by agent-based modeling is the ability to present the results and the process of reaching them in an interactive form rather than an abstract one, even if the initial step in solving the problem involves an abstract, theoretical structure, such as a formula or an indicator [79].

Regarding the use of agent-based modeling in the analysis of the financial contagion effect, the properties of agents that can be found in such models are heterogeneous, emergence, complexity, social interaction, adoption, risk and reward, learning, informational asymmetry, interdependence, feedback, adaptability, realistic behavior, and exploring the space of possibilities. The properties are further described in Table 3.

Agent-based modeling involves computer systems where agents interact with one another within an environment according to well-established rules [72]. However, the rules and conditions established by researchers can lead to unpredictable outcomes at the macro level, as interactions between agents and their environment can result in unexpected results during experiments. These rules and conditionings can nevertheless offer insight into predicting behavior at the micro level [85].
Table 3. Agent characteristics.

<table>
<thead>
<tr>
<th>Agent Characteristics</th>
<th>Description of the Property of ABM Agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneity</td>
<td>Agent-based models provide the ability to simulate each specific entity of interest, such as ants, customers, households, small businesses, large corporations, governments, and more. Additionally, these models allow for the incorporation of realistic or irrational behaviors. By utilizing a “bottom-up” approach, these models can capture the diversity of the real world, making them highly intriguing. Basically, agents are different from each other in terms of behavior, degree of connectivity, and portfolio size.</td>
</tr>
<tr>
<td>Emergence</td>
<td>One of the most important emergent properties of agents refers to the appearance of behaviors or characteristics at the system level that cannot be predicted or explained by the individual analysis of agents. This feature makes agent-based models very useful in analyzing and simulating complex systems such as financial markets. According to [81], one of the most notable instances of emergent behavior in the financial market and economics is Adam Smith’s concept of the invisible hand, which illustrates how the self-interested behaviors of individual agents in the economy can converge to generate optimal outcomes for society as a whole.</td>
</tr>
<tr>
<td>Complexity</td>
<td>By using agent-based models, we can acknowledge the interconnected and non-linear nature of financial markets. This involves modeling the bank and the market as a complex adaptive system. Rather than disregarding complexity, agent-based models actively embrace it. Although agents in agent-based models are considered autonomous entities, the complexity property helps us to program agents so that they have complex behaviors and decisions based on an input of information and variables.</td>
</tr>
<tr>
<td>Social Interaction</td>
<td>In the case of the analysis of banking systems using agent-based modeling, the social interaction property shows us that agents interact with each other through the banking network. For example, you can simulate a network of customers who have mortgages offered by the Romanian banking system. This network can be an eloquent example that highlights the described property.</td>
</tr>
<tr>
<td>Adoption</td>
<td>Agents can adopt different strategies, being autonomous agents, but these are also influenced to some extent by market circumstances and by the level of financial education that the agent has.</td>
</tr>
<tr>
<td>Risk and reward</td>
<td>Through this feature, agents evaluate their potential risk and reward to make decisions about their portfolio. For example, in the case of accessing a second mortgage, the agents can assess the risk of defaulting on the loan, but also the reward of purchasing another property.</td>
</tr>
<tr>
<td>Learning</td>
<td>In terms of learning, agents can learn from past experiences and adapt their strategies.</td>
</tr>
<tr>
<td>Informational asymmetry</td>
<td>Agents may have different or incomplete information about markets and other agents, which can influence their decisions and lead to risk propagation.</td>
</tr>
<tr>
<td>Interdependence</td>
<td>Decisions made by one agent can directly or indirectly impact other agents and the market as a whole.</td>
</tr>
<tr>
<td>Feedback</td>
<td>Agents can receive feedback from the market and adjust their strategies accordingly, which can lead to changes in market behavior and performance.</td>
</tr>
<tr>
<td>Adaptability</td>
<td>Agents may be able to adapt to changes in the economic environment and other market conditions.</td>
</tr>
<tr>
<td>Realistic behaviors</td>
<td>One of the strengths of agent-based models is their ability to generate realistic behavior by observing actual human behavior. Research in behavioral economics has shown that people frequently use heuristics rather than making fully rational decisions. Consequently, there are several models that investigate the consequences of situations where purely rational options are not viable or too costly, or when agents’ environments undergo changes over time [81].</td>
</tr>
<tr>
<td>Exploring the space of possibilities</td>
<td>A major advantage of agent-based modeling is that it allows for the simulation for each individual agent of numerous scenarios to observe the entire population of agents.</td>
</tr>
</tbody>
</table>

Source: References [80–84].

Regarding the use of agent-based modeling in the analysis of the residual contagion effect, we will use the NetLogo IT solution, and we will have the model developed by Gai and Kapadia as a reference. In their study [86], they created a model for two important contagion channels in financial systems. The primary focus is on how losses can spread through the complex network of direct exposures to counterparties after an initial default. However, the secondary effects of distress on asset prices in some financial institutions can cause other financial entities to decrease the value of their assets, potentially triggering additional default rounds. Contagion stemming from direct interconnections between interbank claims and liabilities can thus be intensified through indirect contagion from the asset side of the balance sheet, particularly in illiquid markets for key assets of the financial system.

Gai and Kapadia [86] assumed a financial network in which a number of financial intermediaries are randomly interconnected through interbank exposures that are directed
and weighted by assets and liabilities owned by the banks in the network. In addition, they also considered the size of these exposures important in the analysis of the contagion effect.

From the perspective of graph theory, in the case of the analyzed network, the banks represent the nodes of the network, and the interbank exposures represent the arcs. One of the important properties of the network is its degree distribution. Thus, Gai and Kapadia [86] considered, for example, that in a directed graph, each node has 2 degrees. An internal degree indicates a bank’s interbank assets or exposures, and an external degree represents interbank liabilities.

Thus, the model assumes the following assumptions:

- The total assets of each bank in the network are made up of interbank assets and illiquid external assets such as mortgages;
- The total positions of the interbank assets of each bank in the network are uniformly distributed, following a normal Gaussian distribution, and are independent of the number of links that a bank has in the network;
- Interbank liabilities are determined endogenously;
- The only component of a bank’s liabilities are exogenous data of customer deposits.

The condition for the bank’s solvency is expressed by the following equation:

$$(1 - \theta)A_{IB} + qA_{IM} - L_{IB} - D_i > 0$$

where:

- $A_{IB}$ represents the interbank assets;
- $A_{IM}$ represents the illiquid assets;
- $L_{IB}$ represents the interbank liabilities;
- $D_i$ represents deposits;
- $\theta$ represents the proportion of banks that have defaulted on their obligations to bank;
- $q$ denotes the price at which the illiquid asset can be resold.

The model will calculate the following variables using the relationships below:

$$A_{IB} = b_i * c$$

$$A_{IM} = b_i * (1 - c)$$

$$L_{IB} = \frac{0.2}{E_{out}}$$

$$D_i = A_{IB} + A_{IM} - L_{IB}$$

where:

- $b_i$ represents bank size;
- $c$ represents the constant that is set with a value between 0 and 1;
- $E_{out}$ represents the number of outgoing links.

In this research, the NetLogo computer solution was chosen because it allows for the programming and simulation of interactions among individual agents in a virtual environment [87]. The NetLogo platform provides an intuitive graphical interface and facilitates the visualization of simulation results [87,88]. Users can program the behavior of agents and adjust parameters to explore different scenarios and effects. NetLogo is freely available and supported by an active community that provides resources, examples, and pre-defined models that users can utilize and modify according to their specific needs [77,89]. Recently, NetLogo has been used in various applications in the scientific literature, including but not limited to modeling the human behavior involved in an evacuation process [90,91], reducing the risk associated with airplane boarding in pandemic conditions [92], understanding the issues related to fairness and efficiency in an economic context [93], economic
recovery after an endemic situation [94], optimization in flexible job-shop scheduling [95], and modeling dynamical processes on complex networks [96].

Regarding the simulation of residual contagion effect, the NetLogo environment allows for the definition of agents, interactions among agents within the financial network, modeling of risk and contagion, simulation and evaluation of scenarios, and analysis of the obtained results aimed at assessing the effects of contagion on the overall financial and economic system. The results can highlight specific vulnerabilities, evaluate the effectiveness of policies and risk reduction measures, and provide a framework for understanding the mechanisms of financial contagion propagation [97,98].

5. Case Study: The Analysis of Banking Financial Contagion from a Cybernetic Approach
5.1. Application of Machine Learning

In the following, the banking market of mortgage loans in Romania at the beginning of 2022 is analyzed for the first seven banks in terms of size from the perspective of both assets and the volume of loans granted. The data were taken from the official online platforms of the banks included in the analysis. These banks are the Romanian Commercial Bank (RCB), the CEC Bank (CECB), the ING Bank (INGB), the Transilvania Bank (TB), the Raiffeisen Bank (RB), the Romanian Development Bank (RDB), and the Unicredit Bank (UB). The analysis presented in Figure 7 was carried out in order to observe the evolution of interest rates on mortgage loans for several banks in Romania and to compare the values. The interest rates used are all variable, and the EAI represents the effective annual interest and is an indicator that expresses the total cost of the loan in percentage form.

From the perspective of the analysis of existing mortgages on the banking market in Romania, as can be seen in Figure 7, the bank that offers the lowest interest rate is the CEC Bank. However, the EAI does not have the lowest value, as the bank has several commissions such as credit administration fees that exceed the EAI of the mortgage loans offered by the Romanian Commercial Bank. At the opposite pole, Unicredit Bank has both the highest interest rate and EAI. From the perspective of the client who is applying for a loan, it is important for them to consider the EAI, not just the interest rate, because they must estimate their correct payment capacity to be able to repay the loan.

ROBOR is the reference index used for loans in RON. It is determined by the National Bank of Romania (NBR), based on a specific methodology. It represents an average of the interest rates at which Romanian banks are willing to lend to each other and is published every working day at 11:00 AM. Starting from May 2019, according to NBR 19/2019, the ROBOR index from the variable interest component was replaced by the reference index for consumer loans, IRCC. So, all banks have variable interest expressed according to the IRCC.

The quarterly reference index for consumer loans (IRCC) of the interbank money market is calculated as the arithmetic mean of the daily values recorded in all working days of the respective quarter. After the analysis, we noticed that the evolution trend of the consumer reference index follows a polynomial series.

We observe that the lowest values of IRCC are in the last trimester of 2014 and the second trimester of 2021, respectively, and the highest value is recorded in the third trimester of 2022.

The increase in the inflation rate and the interest rate can constitute a favorable environment for the propagation of financial contagion effects. In Figure 9, we can observe the evolution of the inflation rate from the second trimester of 2019 to the third trimester in 2022. The highest inflation rate recorded during the inflation period is in the second quarter of 2022, and the lowest inflation rate was observed in 2019T3.
The Romanian banking market of mortgage loans (2022)

<table>
<thead>
<tr>
<th>Year</th>
<th>RCB</th>
<th>CECB</th>
<th>INGB</th>
<th>TB</th>
<th>RB</th>
<th>RDB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022</td>
<td>4.36%</td>
<td>5.13%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td>4.28%</td>
<td>5.02%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td>4.06%</td>
<td>4.98%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>3.91%</td>
<td>4.86%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>3.82%</td>
<td>4.79%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7.** Data on interest on mortgage loans from top 7 banks in Romania. Source: Authors’ processing according to the data published on the banks’ official websites.

In Figure 8, we denoted 2019T2 as the values recorded for the second trimester in 2019.

The values of the quarterly reference index - IRCC

\[ y = 0.0007x^2 - 0.0094x + 0.0405 \]

\[ R^2 = 0.8739 \]

**Figure 8.** Evolution of the quarterly reference index—IRCC. Source: Authors’ processing according to the data published on the National Bank of Romania website [99].

The significant increase in inflation can propagate a financial contagion effect through the following channels [100–106]:

- An increase in the production costs of companies, which will have the effect of decreasing profit margins and, possibly, increasing consumer prices. Thus, other effects can be generalized with an impact on the occurrence of potential systemic events of residual contagion. For example, this increase in costs makes it difficult for companies to repay loans. This can lead to the imbalance of the credit risk model that is the basis of the granting of loans by the bank. It is enough to have calculation residues at a sufficiently large bank and with a high degree of connectivity in the financial economic network, due to which the rate of non-performing loans will increase.
- A decrease in purchasing power can negatively affect the sales and profits of the companies.
- An increase in the interest rate is a normal reaction of the central banks, as a rule, to combat the increase in inflation. This can have the effect of decreasing investments and consumption because loans become more difficult for consumers to access.
- An increase in the interest rate can also lead to a decrease in the value of financial assets, leading to significant losses for investors, which favors triggering a financial contagion effect by affecting other financial institutions or even the financial market in which it operates.

![Inflation rate graph](image)

**Figure 9.** Evolution of the inflation rate in Romania. Source: Authors’ processing according to the data published on the National Institute of Statistics website [100].

To prevent such situations, financial supervisory authorities can take measures to limit the rise in inflation and interest rates through appropriate macroeconomic policies. It is also important that banks and other financial institutions have adequate risk management and diversification of their portfolios to reduce the risk of financial contagion in the event of adverse events.

For these reasons, the current research also aims to test machine learning algorithms for mortgage credit models.

Next, we will analyze and predict the ability of some debtors to repay the second mortgage loan and their final status, considering other variables that are important in the analysis of customer behavior, such as the amount of the loan, the remaining debt from the first mortgage loan, etc.

The data that we model below have a sensitive nature; they were taken from the official website of the Central Bank of England [107]. The data set contains information on second mortgages purchased or home equity loans and contains a set of 5960 observations and 13 variables. Two of these variables are non-numeric and eleven are numeric. The analyzed time horizon is 2018–2021.

Based on this data set, we model the relationship between the dependent variable BAD and the predictor variables LOAN, MORTDUE, VALUE, DELINQ, and DEBTINC. The description of the variables is provided in Table 4. The initial variables and loading of the database before the process of processing, cleaning, and preparing it can be seen in Figure 10.
Table 4. Description of variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad</td>
<td>Quantitative</td>
<td>A bad customer is defined, in this case, as a customer who has payment delays or has not repaid the loan by the due date and has penalties such as garnishment of salary, court executions, etc.</td>
</tr>
<tr>
<td>Loan</td>
<td>Quantitative</td>
<td>The value (amount) of the requested loan.</td>
</tr>
<tr>
<td>Mortdue</td>
<td>Quantitative</td>
<td>The amount owed for the mortgage loan.</td>
</tr>
<tr>
<td>Value</td>
<td>Quantitative</td>
<td>The value of the mortgaged property.</td>
</tr>
<tr>
<td>Reason</td>
<td>Qualitative</td>
<td>The purpose of the loan requested. We have 2 possible options: loan consolidation (DeptCon) or home renovation (HomeImp).</td>
</tr>
<tr>
<td>Job</td>
<td>Qualitative</td>
<td>Scorecard for credit applicants’ occupations.</td>
</tr>
<tr>
<td>Yoj</td>
<td>Quantitative</td>
<td>Seniority at work.</td>
</tr>
<tr>
<td>Derog</td>
<td>Quantitative</td>
<td>The number of exemptions.</td>
</tr>
<tr>
<td>Delinq</td>
<td>Quantitative</td>
<td>The number of loans that have not been paid.</td>
</tr>
<tr>
<td>Clage</td>
<td>Quantitative</td>
<td>Duration of the oldest line of credit (months).</td>
</tr>
<tr>
<td>Ning</td>
<td>Quantitative</td>
<td>Number of recent credit requests.</td>
</tr>
<tr>
<td>Clno</td>
<td>Quantitative</td>
<td>Number of lines of credit.</td>
</tr>
<tr>
<td>Debtinc</td>
<td>Quantitative</td>
<td>Debt-to-income ratio.</td>
</tr>
</tbody>
</table>

Figure 10. Defining the k-fold cross validation methodology for the testing set. Source: Authors’ processing in Python.

We identified and estimated several possible models using machine learning algorithms, and later checked the accuracy of the models and compared the results between them to decide which model is better, i.e., the most suitable to make predictions based on the data set regarding the repayment of mortgage loans of some applicants.

The application was made in Python.

To carry out the analyses presented above, the following steps were followed:

1. Database processing: Cleaning, processing, and preparing the data set used in the analysis and subsequent modeling, as well as splitting it into a training set and testing set, respectively.
2. Dimensioning the variables: Splitting the data set into a training set (80%) and testing set (20%), respectively.
3. Standardizing the predictor variables by eliminating the maximum differences between the variables.
4. Defining the k-fold cross-validation methodology to test the performance of the model for the testing set.
5. Estimating the models on the training set and testing the accuracy and performance of the model on the testing set.

For step 1, we imported the following specific libraries in Python: <<pandas>>, <<numpy>>, <<matplotlib.pyplot>>, <<math>>, and <<time>> (Appendix E).

The first step was to identify the missing values and 0 values and replace them with the column mean for each value. This step was performed so that the model could be estimated as accurately as possible. Later after processing, there were no more missing values for the numeric columns (Appendix A).

In the next step, we divided the data set into a training set (train) and a testing set (test), keeping only the significant variables for the analysis (Appendix B).

In the training set, we included about 80% of the observations, and in the testing set, we included about 20% of the observations.
The dependent variable BAD was assigned to the Y variable in the application, and the predictor variables LOAN, MORTDUE, VALUE, DELINQ, and DEBTINC were assigned to the X variable.

The next step was to standardize the predictor variables (variable X) so that all their values had a mean equal to 0 and a standard deviation equal to 1 (Appendix C). Thus, the values of these predictor variables were found in the interval \([-1, 1]\). This standardization was performed to make the analysis as accurate as possible, eliminating large differences in the values between variables, thus bringing them into a common range of values with limits, so that there were no discrepancies in the analysis of the models. The normalization method, min–max scaling (MinMaxScaler()), could also be used, but the normalization method of standard scaling (StandardScaler()) was used, because this technique forms a Gaussian normal distribution of the data, so that the model learns the parameters in a more efficient way. Also, the standard scaling standardization method preserves the useful information of outliers and determines that the algorithm is not easily influenced by outliers, compared to the min–max scaling normalization method.

In practice, the following standardization formula is used:

\[
z = \frac{x - \mu}{\sigma}
\]

where:
- \(\mu\) represents the mean of the data distribution;
- \(\sigma\) represents the standard deviation of the data distribution;
- \(z\) represents the standard score.

It is not necessary to use dummy variables, because the dependent variable BAD is a binary discrete variable, which has only two numerical classes (1/0). Dummy variables are variables that take discrete values, usually 0 or 1, to represent the presence or absence of a particular attribute or characteristic. They are used in statistical analysis and econometric modeling to allow for the measurement of the causal effects of a set of independent variables on a dependent variable. If we had factor variables with more than two classes (n classes), we would have had to create \(n - 1\) dummy variables in order to perform the analysis. The creation of dummy variables could have been performed using various methods including LabelEncoder(), OneHotEncoder(), LabelBinarizer(), and OrdinalEncoder(), which assigns binary numeric values of 1 and 0 to categorical, multi-class variables depending on the variables’ classes.

In the next step, the k-fold cross-validation methodology was defined to test the model performance on the testing set.

The k-fold cross-validation method was used to split the data set into a training set and a testing set, so that after each split, each part of the training set became the testing set, and each part of the testing set became the training set, respectively, so at each stage, the accuracy of the model was tested on new data. This aspect can be seen in Figure 10. This cross-validation method is specifically used to reduce overfitting.

Next, we applied several models to the processed data set. We estimated the models on the training set and tested their accuracy and performance on the testing set. Afterwards, we analyzed and compared the results, choosing the optimal model for forecasting.

For each model, hyperparameter tuning was performed. To test the performance and accuracy of the models on the testing set, we chose a number of performance indicators that measure the accuracy of the models and provide the results of their performance on the predictions on the testing set.

The confusion matrix contains the total number of predicted values and actual values. It is a very popular metric used to solve classification problems. This can be applied to binary classifications, but also to multi-class classification problems. Classification models have multiple categorical outputs. Most indicators that measure prediction errors calculate the total error in the model but cannot find the individual cases of error in the model. The model may misclassify more categories than others, but we cannot see this using
a standard indicator that measures accuracy. Several specific machine learning algorithms were tested, including Random Forest, K-Nearest Neighbors, Logistic Regression, Extra Trees Classifier, Naïve Bayes, AdaBoost, XGBoost, LightGBM, and the test results are summarized in Table 5.

Table 5. Test results for machine learning algorithms applied to the data set.

<table>
<thead>
<tr>
<th>Models</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>86.66</td>
<td>78.71</td>
<td>49.19</td>
<td>60.55</td>
<td>72.85</td>
</tr>
<tr>
<td>K-Nearest Neighbors</td>
<td>85.99</td>
<td>76.47</td>
<td>47.18</td>
<td>58.35</td>
<td>71.67</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>80.36</td>
<td>88.89</td>
<td>6.45</td>
<td>12.03</td>
<td>53.12</td>
</tr>
<tr>
<td>Extra Trees Classifier</td>
<td>81.28</td>
<td>79.13</td>
<td>10.48</td>
<td>18.91</td>
<td>55.19</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>80.87</td>
<td>73.81</td>
<td>12.5</td>
<td>21.38</td>
<td>55.66</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>88.26</td>
<td>80.00</td>
<td>58.06</td>
<td>67.29</td>
<td>77.13</td>
</tr>
<tr>
<td>XGBoost</td>
<td>88</td>
<td>78.07</td>
<td>58.87</td>
<td>67.13</td>
<td>77.25</td>
</tr>
<tr>
<td>LightGBM</td>
<td>87</td>
<td>75.41</td>
<td>55.65</td>
<td>64.03</td>
<td>75.44</td>
</tr>
</tbody>
</table>

The first measure of model performance is accuracy testing. This is calculated as the ratio of correctly predicted observations to the total number of observations. To evaluate the performance of each proposed model, the accuracy was calculated. The higher its value, the better the model and the better it is at predicting the observations. In terms of accuracy, the best value was obtained for the AdaBoost model, followed by XGBoost, LightGBM, Random Forest, and K-Nearest Neighbors. Another indicator for performance testing is the precision score, which is calculated as the ratio of correctly anticipated positive observations to the total number of predicted positive observations. A high accuracy score value indicates a low False Positive rate, that is, when the predicted situation is close to or similar to the current situation. For our models, the highest value was obtained by the Extra Trees Classifier, followed by Logistic Regression and AdaBoost.

The next indicator for evaluating the performance of the models is the recall score, which refers to the sensitivity between the correctly predicted positive observations and all actual observations.

A value greater than 0.5 is considered to indicate good model performance from the perspective of sensitivity to observations. From this point of view, the best models are XGBoost and AdaBoost. In order to be able to draw the most pertinent conclusion about the identification of the optimal model, we also use the F1 score, which is calculated as a weighted average between the precision score and the recall score. The best value was obtained by AdaBoost, followed by XGBoost and LightGBM.

The last indicator to measure the performance of the models used is the AUC Score, which depends on the ROC curve. The ROC curve is a graphical representation that shows the relationship between the True Positive rate and the False Positive rate. The AUC can take values between 0 and 1. If the chosen model has 100% wrong predictions, then the AUC for this model will have a value of 0. Conversely, if all the predictions of the model are correct, then the AUC will have a value of 1. From the perspective of the AUC score, the best model was XGBoost, followed by AdaBoost, LightGBM, Random Forest, and K-Nearest Neighbors.

Finally, we plotted the AUC-ROC curve for XGBoost (Appendix D), as it was the model with the best value for the AUC score and performed very well in most of the accuracy and performance evaluation tests, being a possible optimal model (please see Figure 11).
Figure 11. Graphical representation of the AUC-ROC curve.

The graphical representation of the ROC curve above illustrates the False Positive rate on the OX axis, and the True Positive rate on the OY axis. The coordinate points (0,0) represent the wrong classification of all examples, and coordinates (1,1) represent the correct classification of all examples. Considering the AUC value of 77.33%, this indicates a good performance of the classification model.

Regarding the emergence of the financial contagion effect, in this case, we can investigate the correlations of the prediction errors. This can become problematic because it can cause incorrect estimates of standard errors, thus distorting the conclusions of the model.

Residual financial contagion refers to the spread of negative financial effects of an economic or financial event to other institutions or markets through their interconnections and relationships. This type of contagion can be caused by a variety of factors, such as the insolvency of a financial institution, liquidity crises, market volatility, or rapid changes in interest rates.

Residual financial contagion can have a significant impact on the global economy by creating a domino effect in the chain of financial institutions and markets. This can lead to a significant decline in financial assets, a decrease in liquidity, and an increase in credit risk for other financial institutions.

To prevent or minimize residual financial contagion, financial regulators and supervisors can take precautionary measures, such as introducing strict risk supervision and control rules, continuously monitoring financial markets, applying stress tests, and maintaining adequate capital and liquidity.

Also, effective communication and increased transparency between financial institutions and regulators can help to prevent residual financial contagion by raising awareness of the risks and vulnerabilities in the financial system.

5.2. Application of Agent-Based Modeling

There were cases of the residual financial contagion effect in Romania. During the financial crisis of 2008, but also in 2012, the problems in the banking system in Romania had a domino effect on other financial institutions.

In 2008, the banking system in Romania and beyond was exposed to significant credit risks generated by the rapid increase in non-performing loans and the exposure to currency risk. This context led to liquidity problems and the need to resort to emergency financing
from the National Bank of Romania. In 2012, the same causes led to the emergence of the residual financial contagion effect.

The measures taken by the financial supervision authorities in Romania regarding the safety and stability of the financial system were as follows [50]:

- The establishment of the Financial Supervision Authority in 2013 with the aim of improving the supervision of financial institutions in Romania.
- The elaboration and introduction of stricter rules and regulations for financial institutions, such as increasing capital requirements and liquidity requirements for banks.
- Improving transparency by disseminating relevant information for the financial market and introducing periodic reports communicated by the press.
- Developing stress tests and increasing the frequency of audits or inspections that financial institutions are subject to in order to identify risks and take preventive measures.

According to public information on the website of the National Bank of Romania, there are currently 25 commercial banks and seven branches of foreign credit institutions. That is, we have a number of 32 banks [108] that form the banking network in Romania.

Regarding the simulation of the transmission of the residual financial contagion effect, we used agent-based modeling, starting from the research carried out by Gai and Kapadia [86]. The developed model was used to explore how shocks in one part of the financial system can lead to contagion and even the collapse of the entire system.

Based on this information, we will illustrate and simulate the residual financial contagion effect using agent-based modeling and the NetLogo IT solution. We assume that the banking network formed by the 32 banks in Romania are connected to each other through financial transactions (loans, deposits, financial transfers, etc.). Each bank has a capital level and certain banks grant loans to other banks.

The variables used in the model are summarized in Table 6.

**Table 6.** The description of variables used in NetLogo.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of agents (banks)</td>
<td>The agents modeled in NetLogo are the 32 banks that form the banking network in Romania.</td>
</tr>
<tr>
<td>Interbank assets</td>
<td>Represents the loans the bank has received from other banks on the interbank market.</td>
</tr>
<tr>
<td>Illiquid assets</td>
<td>Represents some form of illiquid assets on the balance sheet.</td>
</tr>
<tr>
<td>Interbank liabilities</td>
<td>Represents the obligations to other banks on the interbank market.</td>
</tr>
<tr>
<td>Deposits</td>
<td>Bank deposits.</td>
</tr>
<tr>
<td>Bank size</td>
<td>Banks have different sizes based on random normal distribution.</td>
</tr>
<tr>
<td>Ticks</td>
<td>One tick represents 3 days.</td>
</tr>
</tbody>
</table>

Source: Authors’ processing according to [86].

Regarding the simulation made in NetLogo, starting from the model developed by Gai and Kapadia, we added procedures to the model to simulate the state of default starting from the bank that has the largest size, and simulated a scenario in which the state of default was installed at the bank with the smallest size. The interface can be viewed in Appendix F. Regarding the description of the modules that can be seen in the interface, they are detailed in Table 7.

Figure 12 shows an example of a simulation that was made in NetLogo, which considers the following input in the initial stage:

- The agents (banks) were initialized;
- The interdependencies between the agencies are built randomly, with each bank having a certain size;
- The simulation is based on the scenario that one of the banks with the smallest size goes into default.
Table 7. The description of parameters used in NetLogo.

<table>
<thead>
<tr>
<th>Setup Parameters</th>
<th>Meaning</th>
<th>Possible Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Banks</strong></td>
<td>The total number of banks that will be considered for the proposed analysis.</td>
<td>Between 1 and 80 Since we are analyzing the behavior that may occur in Romania, 32 banks were selected, since this is the total number of the banks present in the country.</td>
</tr>
<tr>
<td><strong>Setup</strong></td>
<td>Initializes the model.</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>mu</strong></td>
<td>The average value of the interbank assets.</td>
<td>Between 0 and INF In order to have a Gaussian distribution of the interbank assets for our analysis, mu should be 0.</td>
</tr>
<tr>
<td><strong>sigma</strong></td>
<td>The standard deviation of the interbank assets.</td>
<td>Between 0 and INF In order to have a Gaussian distribution of the interbank assets for our analysis, sigma should be 1.5.</td>
</tr>
<tr>
<td><strong>Default random bank</strong></td>
<td>One random bank from their total amount is going to be in a default state.</td>
<td>1</td>
</tr>
<tr>
<td><strong>Default smallest bank</strong></td>
<td>One of the smallest banks is going to be in a default state.</td>
<td>1</td>
</tr>
<tr>
<td><strong>Default biggest bank</strong></td>
<td>One of the biggest banks is going to be in a default state.</td>
<td>1</td>
</tr>
<tr>
<td><strong>go</strong></td>
<td>Runs the model.</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Count links</strong></td>
<td>The total amount of links that are between all of the banks.</td>
<td>Between 1 and 32²</td>
</tr>
<tr>
<td><strong>Defaulted banks</strong></td>
<td>The number of banks in default-state at a specific moment in time.</td>
<td>Between 1 and 32</td>
</tr>
<tr>
<td><strong>Non-defaulted banks</strong></td>
<td>The number of banks in a non-default-state at a specific moment in time.</td>
<td>Between 0 and 31</td>
</tr>
</tbody>
</table>

Source: Authors’ processing according to [86].

Figure 12. Simulation of the financial contagion effect in NetLogo.

In the Final State image, it can be seen that the entire network goes into default, with the financial contagion effect being transmitted systemically throughout the entire network. The network was completely contaminated, outlining an effect that can be understood as an economic crisis. The conclusions of this simulation are that although the size of the bank in default is the smallest, connectivity in the network is very important as a property of complex adaptive systems. The selected bank of size 1, as can be seen, interacts through lending relationships or by holding assets with 24 banks out of the remaining 31, which is an aspect that favored the transmission of the contagion effect. Moreover, the bank of size 1 is connected with the bank with the largest size (19), which is an aspect that also influences it.

However, in order to have a holistic picture of the above conclusion, we propose the analysis of three scenarios: the bank that defaults is randomly selected, the defaulting bank with the largest size is selected, and the bank with the smallest size is chosen.
For each scenario, 1000 simulations were carried out. Ticks was considered to be the default entry time, and one tick was considered to be 3 days. For the three scenarios, the results are centralized in Table 8.

**Table 8.** The results of the financial contagion simulation for the 3 scenarios.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Statistics</th>
<th>Interbank Assets</th>
<th>Illiquid Assets</th>
<th>Interbank Liabilities</th>
<th>Deposits</th>
<th>Bank Size</th>
<th>Tick</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default Random Bank</td>
<td>Average Non-Contagion</td>
<td>0.530662021</td>
<td>2.455400697</td>
<td>0.001776756</td>
<td>2.984286</td>
<td>2.986062718</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Max. Value Non-Contagion</td>
<td>19.6</td>
<td>78.4</td>
<td>0.135232349</td>
<td>98</td>
<td>98</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Min. Value Non-Contagion</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
<td>0.878655</td>
<td>1</td>
<td>360</td>
</tr>
<tr>
<td>Default Biggest Bank</td>
<td>Average Financial Contagion</td>
<td>0.66056338</td>
<td>2.642253521</td>
<td>0.620101426</td>
<td>2.682715</td>
<td>3.302816901</td>
<td>2.443662</td>
</tr>
<tr>
<td></td>
<td>Max. Value Financial Contagion</td>
<td>15</td>
<td>60</td>
<td>4.078809524</td>
<td>73.89019</td>
<td>75</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Min. Value Financial Contagion</td>
<td>0.2</td>
<td>0.8</td>
<td>0</td>
<td>−2.99836</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Default Smallest Bank</td>
<td>Average Non-Contagion</td>
<td>5.48487</td>
<td>25.336</td>
<td>0.003419</td>
<td>30.81745</td>
<td>30.82087</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Max. Value Non-Contagion</td>
<td>64.2</td>
<td>256.8</td>
<td>0.337621</td>
<td>321</td>
<td>321</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Min. Value Non-Contagion</td>
<td>0</td>
<td>2.4</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Average Financial Contagion</td>
<td>5.837297</td>
<td>23.40054</td>
<td>0.54245</td>
<td>28.69539</td>
<td>29.23784</td>
<td>2.462162</td>
</tr>
<tr>
<td></td>
<td>Max. Value Financial Contagion</td>
<td>45.8</td>
<td>183.2</td>
<td>5.9</td>
<td>228.96</td>
<td>229</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Min. Value Financial Contagion</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4.207451</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Average Non-Contagion</td>
<td>0.181226054</td>
<td>0.818773946</td>
<td>0.002115523</td>
<td>0.99788</td>
<td>1</td>
<td>358.6303</td>
</tr>
<tr>
<td></td>
<td>Max. Value Non-Contagion</td>
<td>0.2</td>
<td>1</td>
<td>0.316666667</td>
<td>1</td>
<td>1</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Min. Value Non-Contagion</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
<td>0.68333</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Average Financial Contagion</td>
<td>0.199524941</td>
<td>0.800475059</td>
<td>0.70634049</td>
<td>0.29137</td>
<td>1</td>
<td>2.418052</td>
</tr>
<tr>
<td></td>
<td>Max. Value Financial Contagion</td>
<td>0.2</td>
<td>1</td>
<td>12.8</td>
<td>1</td>
<td>1</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Min. Value Financial Contagion</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
<td>−6.2114</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Following the three realized scenarios, the following observations were concluded regarding the transmission of the financial contagion effect in the banking network in Romania:

- The size of the bank is an important element that influences the transmission of the contagion effect.
- The degree of connectivity given by banking transactions carried out in the network has a significant impact on the transmission of the systemic default shock and favors the propagation of the contagion effect.
- The level of deposits is again an important element compared to the level of liabilities (interbank loans). In general, according to the simulations, if the level of deposits is higher than the interbank loans, the contagion effect is blurred.
- We noticed that in about 1 week (2 ticks), the contagion effect can spread and be installed in the entire banking network.

The simulations carried out are based on the analysis of the residual contagion effect, which represents a situation in which a problem that a bank has can affect other banks, even after the initial problem has been solved. The residual financial contagion is transmitted in the network through direct exposures, as we observed in the simulations, when a bank has debts or loans to other banks, especially when they are non-performing. In dealing with the emergence of the risk of financial contagion, regulatory authorities must have an active
role in monitoring and managing systemic events in the banking system. Considering the simulations above, we can conclude that regardless of the form of financial contagion, its effects can be devastating when these effects cannot be predicted, or when they are not detected in time to be avoided. The simulation carried out in NetLogo offers an important perspective on systemic risks and can help us to identify effective solutions to manage these risks and to reduce the effects of financial contagion.

6. Conclusions

Considering the uncertainty of the current economic environment, financial contagion can be caused by a variety of factors such as fluctuations in the interest rate or exchange rate, interconnections between banks, the increase in the inflation rate, the increase in production costs, or the lack of diversification.

Currently, the specialized literature presents the factors that contribute to the creation of financial contagion, and most of the time, spillover contagion is the most frequently studied form.

Our research highlights the importance of studying financial contagion and its impact on the banking market within the broader context of economic and business sustainability. The analysis sheds light on the potential risks and vulnerabilities within the mortgage loan sector, which are crucial factors to consider for sustainable economic growth and stability. By exploring the link between financial contagion and sustainability, this study contributes to a comprehensive understanding of the economic and business aspects of sustainability in the specific context of the banking market in Romania.

The novelty of this research is the focus on residual contagion. This form of contagion is illustrated by the fact that when a bank is negatively affected or goes into default, it can transmit negative effects to other banks through the banking network and through the phenomenon of financial interconnection. This can be described as a residual effect because it can form following other phenomena of financial contagion, as can be seen in the simulations made in NetLogo. When a bank faces a solvency problem or cannot pay its debts to other banks in its interconnected network, these defaulted banks thus face liquidity problems and cannot face payment requests from their customers, which can lead to the collapse of other banks in their network.

It is very important for a bank to have a healthy and diverse loan portfolio. From this perspective, the bank develops, implements, and periodically validates credit risk models that are adapted to the economic context and the profile of the loaned client. Machine learning algorithms are recommended by the latest research in the field for the construction of these credit risk models and have a significant impact on all types of bank loans.

Based on the results obtained by the ML algorithms, the best models from the perspective of performance metrics (accuracy, precision, recall score, F1 score, and AUC curve) are the boosting algorithms XGBoost, AdaBoost, and LightGBM.

After processing the data and applying the model to them, the accuracy tests as well as the other performance measurement indicators performed very well, indicating that the probability of predicting a customer’s loan repayment is very high. This ability to predict a loan repayment by customers has a major impact and brings value to banking systems. Through these predictions, the bank can manage and resolve cases more effectively and plan its processes in a more efficient manner. These types of predictions can be valuable assets and improve processes not only in banking systems, but in many other industries as well.

Based on the agent-based modeling made in NetLogo, the model developed by Kai and Kapadia [86] helped us build some simulations to highlight how the residual financial contagion effect propagates in the banking network in Romania. The simulations highlighted the fact that the degree of connectivity between the commercial banks in Romania and the way in which they are interconnected have a major importance in the emergence and propagation of contagion effects.

The conclusions of this research can be considered as both general conclusions that can be applied to banking systems worldwide, as well as conclusions that are specific to
Romania. However, from the perspective of regulations and policies, banking institutions may be subject to different regulatory frameworks. In the case of Romania, the National Bank of Romania issues and enforces certain regulations and procedural frameworks, but banks may also be subject to specific criteria set by the European Central Bank, guidelines issued by the European Banking Authority, or other supervisory institutions. The results of applying this research in any of the banking institutions in Romania can contribute to identifying scenarios that highlight certain residual risks that have not been observed until now. These aspects can be formalized through the improvement of internal and external regulatory frameworks.

Future research directions in the area of exploring the residual financial contagion effect should consider the implications of cybernetic systems on the banking network by illustrating the properties of complex adaptive systems. For example, in the research carried out in this paper, the following properties can be easily identified: connectivity, which is illustrated by the relationships and interconnections between the banks that form the banking network in Romania, and co-evolution and adaptability, which are highlighted by the way in which certain banks stabilize to the default effects of other banks and adapt to the current economic context by using machine learning algorithms. On the other hand, besides the analysis of the banking network as a complex, cybernetic, and dynamic adaptive system, it is interesting to apply econometric tests to investigate the effects of banking residuals and the emergence of the residual contagion effect. Another important research direction can be to analyze the impact of green credit on financial contagion. By doing so, we can identify aspects that explain how measures of green credit issuance or environmental performance indicators are incorporated into models. Researchers can outline assumptions and hypotheses regarding the relationship between green practices and the propagation of residual contagion in the banking sector.

Author Contributions: Conceptualization, Ş.I.; data curation, I.N.; formal analysis, Ş.I.; investigation, Ş.I., N.C. and I.N.; methodology, Ş.I.; project administration, C.D.; software, Ş.I.; supervision, C.D.; validation, Ş.I., N.C. and I.N.; visualization, N.C. and C.D.; writing—original draft, Ş.I.; writing—review and editing, N.C., I.N. and C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: This paper was co-financed by the Bucharest University of Economic Studies during the PhD program.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

![Figure A1. Identifying missing values.](image1)

![Figure A2. Replacing missing values and 0 values with the mean.](image2)
Figure A2. Replacing missing values and 0 values with the mean.

Appendix B

```python
# Drop non-numeric columns
df = df.drop(['REASON', 'JOB'], axis=1)
print(df.head())
```

Figure A3. The first 5 observations of the data set with the columns REASON and JOB removed.

Appendix B

```python
# Split dataset into train and test
y = df.iloc[:, 0].values
X = df.iloc[:, [1, 2, 3, 8, 12]].values
# print(y.head())
# print(X.head())
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=1, test_size=0.2)
```

Figure A4. Splitting the data set into training and testing sets.

Appendix C

```python
# Feature scaling
sc_X = StandardScaler()
X_train = sc_X.fit_transform(X_train)
X_test = sc_X.transform(X_test)
print(X_train)
```

Figure A5. Observing the structure/dimensionality of the variables on training and testing sets.

Appendix C

```python
from sklearn import metrics
auc = metrics.roc_auc_score(y_test, y_pred_xgb)
false_positive_rate, true_positive_rate, thresholds = metrics.roc_curve(y_test, y_pred_xgb)
```

Figure A6. Data standardization.

Appendix D

```python
plt.figure(figsize=(10, 8), dpi=100)
plt.axis('scaled')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.title('AUC & ROC curve')
plt.plot(false_positive_rate, true_positive_rate, 'g')
plt.fill_between(false_positive_rate, true_positive_rate, facecolor='lightblue', alpha=0.7)
plt.text(0.95, 0.05, 'AUC = 0.84', ha='right', fontsize=12, weight='bold', color='blue')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.show()
```

Figure A7. Designing the AUC-ROC curve.
Appendix E

Figure A8. Library’s import.

Appendix F

Figure A9. NetLogo interface.

References


13. Sargents, H. Algorithmic decision-making in financial services: Economic and normative outcomes in consumer credit. *AI Ethics* **2022**. [CrossRef]


15. Herrmann, H.; Masawi, B. Three and a half decades of artificial intelligence in banking, financial services, and insurance: A systematic evolutionary review. *Cybern. Syst. J.* **2022**. [CrossRef]


19. Herrmann, H.; Masawi, B. Three and a half decades of artificial intelligence in banking, financial services, and insurance: A systematic evolutionary review. *Cybern. Syst. J.* **2022**. [CrossRef]


33. Trenca, I.; Dezsi, E. Financial contagion on the Romanian stock market. In Finante—Provanicile Viitorului (Finance—Challenges of the Future); University of Craiova, Faculty of Economics and Business Administration: Craiova, Romania, 2012; Volume 1, pp. 27–36.
40. Tsagkanos, A.; Sharma, A.; Ghosh, B. Green Bonds and Commodities: A New Asymmetric Sustainable Relationship. Sustainability 2022, 14, 6852. [CrossRef]
42. Lian, Y.; Gao, J.; Ye, T. How does green credit affect the financial performance of commercial banks?—Evidence from China. J. Clean. Prod. 2022, 344, 131069. [CrossRef]
47. Hoppit, J. The myths of the South Sea Bubble. Trans. R. Hist. Soc. 2002, 12, 141–165. [CrossRef]
65. Aggarwal, C.C. Neural Networks and Deep Learning; Springer International Publishing AG: Cham, Switzerland, 2018.


83. Chakraborty, S. Developing Agent-Based Models to Study Financial Markets. Digital Commons @ University of South Florida. Available online: https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=10119&context=etd (accessed on 16 February 2023).


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.