Machine Learning Models to Predict Shareholder Returns in the Banking Industry

Le Duc Thinh [1[0009-0002-9921-6348]] and Nguyen Phuong Anh [2]

[1] International School, Vietnam National University, Hanoi
thinhld@vnu.edu.vn
[2] International School, Vietnam National University, Hanoi
phuonganhng.7512@gmail.com

Abstract. Competitive pressures have steadily driven commercial banks to strategically focus on generating returns to shareholders. This research article purpose is to analyze and provide a summary about the impacts of key financial ratios (key metrics) on the effectiveness and efficiency of commercial banking industry, which reflects on the shareholder return of these banks, using machine learning and the official data of the banking industry in the USA. In this article, we study key metrics for commercial banks, and analyze annual financial and operating data of some biggest, publicly traded commercial banks in the USA in order to find out some predictive models using machine learning algorithms, particularly for panel data, and therefore, to give investors, shareholders, or asset managers a reliable tool to evaluate and forecast the performance of commercial banks.

Keywords: Machine learning, Commercial banking, Metric, Panel data, Shareholder return.

1 Introduction

The COVID-19 pandemic had a tremendous impact on the global economy, and the banking industry has not been immune to its effects. The uncertainty made it difficult to accurately predict bank performance, as the pandemic has created unprecedented conditions that cannot be easily compared to previous economic crises. In addition, regulatory changes and government interventions to mitigate the impact of the pandemic on the economy further complicate forecasting efforts.

On the flip side, the pandemic also presented opportunities for innovation and the growth of technologies, which motivated many sectors to turn to digital solutions to adapt to the new normal. Many modern technologies, such as artificial intelligence (AI), cloud computing, or the Internet of Things (IoT), have become effective tools to store, analyze, and make predictions from the enormous amount of data.

For the decision-making process, AI and machine learning tools started to play a vital role in identifying and responding to many business problems, as well as supporting trends and demand forecasts in both short-term and long-term strategies. This is concluded in a survey done by KPMG [9] which shows that as of 2021, 83% of players in the financial services sector were using AI in some kind, hence in terms of AI penetration, financial sector came behind only the manufacturing sector, which had a 93% penetration rate. KPMG [9] also forecasts that over the period 2022-2027, the value of investments in AI in the banking and finance industry will increase at an average rate of 31.5% annually, rising from $13.9 billion in 2022 to $54.7 billion by 2027.

Among the innumerable factors impacting the stock prices and performances of banks, it is recommended to look for insights from historical data, which can be easily accessed through financial statements and other related documents. Investors or shareholders can utilize AI or machine learning tools to collect and process large amounts of data, as well as identify trends and patterns that may not be visible to human analysts. As more and more data is stored to train these models, their accuracy will continue to increase, making them reliable tools for decision-making in the future.

Total Shareholder Returns (TSR) is a key metric used to evaluate the performance of banks and assess the value delivered to shareholders over a specific period. TSR considers both capital appreciation (stock price growth) and dividends received by shareholders. It provides a comprehensive view of the overall return generated by an investment in a bank's stock.

Several research studies have been carried out on variables that influence the values of TSR. Most of these studies have focused on the relationship between TSR and several metrics that support their research objectives, or the role of TSR in indicating the performance of an organization. For example, Mburu et al. (2018) [7] studied the relationship between TSR and working capital management to establish a direct association between the two metrics. Ataünal et al. (2016) [3] studied the relationship between revenue growth and creation of shareholder value (aka TSR) using a sample of 243 non-financial Standard and Poor's 500 (S&P500) companies using 22 years of data (1993–2014). It was observed that when growth rates surpass sustainable growth rates, companies turn out ruining their value.

In the last 5 years, the number of studies about machine learning models in the finance sector has increased, using various techniques and comparing their accuracy. For instance, Al-Dmour et al. (2018) [2] tried to predict business performance by using multiple linear regression (MLR) method and neural network method. Their work compared the accuracy power of ANN (artificial neural networks) and MLR using the reliability of accounting information systems as independent variables, and business performance as a dependent variable.

Barnes and Lee (2007) [5] conducted a study implementing the feature selection method for examining the characteristics that contribute to company wealth creation in the Miscellaneous Industrials sector of the Australian Stock Exchange. They examined whether a multiple-domain model outperforms a single-domain model in forecasting company fortune by comparing conventional and artificial intelligence (AI) feature selection techniques. The study uncovered that a multiple-domain model was the most effective and that attributes such as WACC, Funds from Operation/EBITDA, and EPS were pivotal in determining the direction of change in shareholder wealth. ROA, capital turnover, and gross Debt/Cashflow were also essential characteristics for comprehending relative shareholder growth. Using neural networks and feature selection methods, the study assessed that the multi-domain model [GRNN (Classification)-Logistic into

 ISBN: 978-604-80-9774-5

GRNN (Prediction)-Logistic] could reduce the top ten features implying the health of a company from 46 to 3, while maintaining a high level of performance in both two focus points. However, not all research focuses on predicting the TSR by the influence of metrics in their financial statements, as they often focus on predicting the stock price in the short term.

There is little research focusing on the impact of many other metrics in financial statements due to the difficulties in collecting data among companies, as well as the differences in the format of the financial statements. This research aims to solve this problem, with the scope of the research focusing on large commercial banks in the USA, and information accessible and standardized by the *Federal Financial Institutions Examination Council's (FFIEC*) [11], a US government body. From then on, the research focuses on identifying the impact of key metrics in financial statements on TSR using machine learning techniques.

2 Research Questions, Methodology and Scope of Research

The aim of this research is (1) to identify the essential metrics in financial statements that influence the value of TSR, and (2) to explore different machine learning algorithms that can bring out reliable prediction models for the values of TSR, based on those metrics.

Methodology & scope of research:

Qualitative: We study various machine learning techniques used to process panel data. We also study 41 key metrics for commercial banking industry such as Net Interest Income/Average Assets, Net Loan and Lease/Assets, Growth rate of Net Loan and Leases, etc.

Quantitative: we collect and analyze data of the 30 largest banks in the US. We obtain 10 years of data (2013-2022) from Uniform Bank Performance Report (UBPR) tool, which is implemented by *Federal Financial Institutions Examination Council's (FFIEC*) [11] and their stock prices from Yahoo Finance [12]. From there we use various machine learning techniques to identify the impact of key metrics on TSR.

3 Machine Learning for Panel Data

As indicated beforehand, many financial and banking industry prediction models employ linear models and time series analysis. However, these methodologies may overlook several nonlinear relationships that have a significant impact on the performance of the bank. Despite their findings that help investors understand the trends and volatility of a particular stock or business over time, investors need to spread their investments across multiple companies and institutions rather than placing all their eggs in one basket.

Using cross-sectional analysis, investors can develop prediction models that can analyze numerous organizations over a single time frame, using a variety of attributes as predictors of business performance. In addition, the increase in factors influencing TSR and bank performance revealed the limitations of traditional linear models and time series analysis. In the meantime, the adoption of numerous non-parametric approaches that leverage the power of machine learning algorithms, such as deep neural networks, has demonstrated the potential to improve the accuracy of predictions by nonlinearly combining various factors. Masaya and Nakagawa (2020) [6] developed a cross-sectional framework for daily stock price prediction using deep learning for investment management. They discovered that deep neural network (DNN)-based stock price prediction prevails over the random forest and ridge regression.

Observations collected from multiple entities (such as individuals, businesses, and countries) over time generate panel data, a form of longitudinal data. It has both cross-sectional and time-series dimensions, making it a valuable tool for analyzing dynamic relationships and changes over time. Incorporating both within-entity (cross-sectional) and between-entity (time-series) variations, panel data analysis techniques provide insights into individual and aggregate behaviors, trends, and dependencies.

Recent research has explored the efficacy of machine learning techniques in panel data analysis across multiple disciplines. Adeboye and Alabi (2022) [1] compared the work of various deep learning models for dynamic panel data to examine the relationship between macroeconomic indicators and economic development in several African nations, thereby identifying policy implications. Ding, Wang, Zhang, and Zhong (2022) [10] utilized a machine-learning-based model to forecast house prices in several Chinese regions. The optimal model using the CatBoost algorithm allowed the authors to explain the fluctuations in housing costs and the most key variables.

4 Total Shareholder Returns and Key Financial Metrics of the Commercial Banking Industry

4.1 Total Shareholder Returns (TSR)

TSR is the total return of a stock for a given period, or the capital gain plus dividends. The annual TSR is calculated by the following formula:

$$\frac{(Price\ at\ end\ of\ year - Price\ at\ beginning\ of\ year) + Dividends}{Price\ at\ beginning\ of\ year} =$$
$$\frac{Adjusted\ closing\ price\ at\ end\ of\ year - Adjusted\ closing\ price\ at\ beginning\ of\ year}{Adjusted\ closing\ price\ at\ beginning\ of\ year}$$

The stock price used in this study is the adjusted closing price that already includes dividend. We obtain this stock price from Yahoo Finance ([19]).

In the US, all banks are required to report their data to Uniform Bank Performance Report (UBPR) under the standards set by FFIEC ([11]). Key financial metrics of the banking industry are divided by UBPR into the following 5 groups:

4.2 Earnings and Profitability Metrics

(1) Interest Income to Average Assets
(2) Interest Expense to Average Assets

 ISBN: 978-604-80-9774-5

(3) Net Interest Income to Average Assets
(4) Noninterest Income to Average Assets
(5) Noninterest Expense to Average Assets
(6) Pre-Provision Net Revenue to Average Assets
(7) Provision: Loan & Lease Losses to Average Assets
(8) Provision: Credit Loss Other Assets to Average Assets
(9) Pretax Operating Income to Average Assets
(10) Realized Gains/Losses Security to Average Assets
(11) Unrealized Gains / Losses Equity Security to Average Assets
(12) Pretax Net Operating Income to Average Assets
(13) Pretax Net Operating Income to Average Assets
(14) Net Income Attribute to Min Interest to Average Assets
(15) Net Income Adjusted for Sub-Chapter S Status to Average Assets
(16) Net Income to Average Assets

4.3 Margin Analysis

(17) Average Earning Assets to Average Asset
(18) Average Interest-Bearing Funds to Average Assets
(19) Interest Income to Average Earning Assets
(20) Interest Expense to Average Earning Assets
(21) Net Interest Income to Average Earning Assets

4.4 Loan and Lease Analysis

(22) Net Loss to Average Total LN&LS
(23) Earnings Coverage of Net Losses
(24) LN&LS Allowance to LN&LS Not Held For Sale
(25) LN&LS Allowance to Net Losses
(26) LN&LS Allowance to Nonaccrual LN&LS
(27) Total LN&LS-30-89 DAYS Past Due %
(28) Total LN&LS-90+ Days PD & Nonaccrual
(29) Non-Curr LNS+OREO to LNS+OREO

4.5 Liquidity

(30) Net Non-Core Funding Dependence $250,000
(31) Net Loans & Leases to Total Assets
(32) Net Loans & Leases to Deposits

4.6 Capitalization

(33) Leverage Ratio
(34) Total Capital Ratio
(35) Cash Dividends to Net Income
(36) Non-Curr Lns+OREO to T1 Capital+Allowance

4.7 Growth Rates

(37) Total Assets - annual change
(38) Tier One Capital 12-month growth rate
(39) Net Loans and Leases 12-month growth rate
(40) Short Term Investments 12-month growth rate
(41) Short-Term Non-Core Funding 12-month growth rate

Important Note: Due to paper length constraint, we do not explain the detail of these metrics here. But all of these can be found at the Federal Financial Institutions Examination Council's (FFIEC) Homepage ([11]). We use exact terminologies from FFIEC.

5 Data Collecting and Processing

5.1 Data Collecting

We collect and analyze the annual data of the 30 biggest commercial banks (in terms of total assets) in the U.S. Data was collected over a 10-year period, from 2013 to 2022, from the Uniform Bank Performance Report (UBPR), retrieved from the Federal Financial Institutions Examination Council (FFIEC), and the stock prices from Yahoo Finance. The authors collect data from UBPR for each bank via this link: https://cdr.ffiec.gov/public/ManageFacsimiles.aspx. To get access to the right bank, it is essential to look at the name of the bank and its headquarters location, as well as its FDIC Certificate Number (which can be found from other sources on the Internet). From then on, the authors can generate a UBPR report within a standard or custom 5-year period. All data is stored in an Excel file. All the metrics retrieved from UBPR are renamed by the formula: R + the number that the metric is assigned.

 ISBN: 978-604-80-9774-5

Bank	Year	TSR	R1	R2	R3	R4	R5	R6	R7	R9	R10	R12	R13	R14	R15	R16	R17
1. JPMorgan Chase	2013	33.71	2.04	0.27	1.76	1.99	2.64	1.11	-0.07	1.18	0.03	1.21	0.81	0.00	0.81	0.81	90.89
1. JPMorgan Chase	2014	9.67	2.02	0.22	1.80	2.00	2.62	1.18	0.05	1.13	0.00	1.13	0.76	0.00	0.76	0.76	91.55
1. JPMorgan Chase	2015	7.83	1.95	0.19	1.76	1.98	2.49	1.25	0.06	1.19	0.01	1.20	0.85	0.00	0.85	0.85	92.22
1. JPMorgan Chase	2016	38.71	2.10	0.23	1.88	2.02	2.40	1.49	0.11	1.38	0.01	1.39	0.95	0.00	0.95	0.95	92.41
1. JPMorgan Chase	2017	25.38	2.32	0.33	1.99	1.97	2.42	1.53	0.09	1.45	0.00	1.44	0.89	0.00	0.89	0.89	92.49
1. JPMorgan Chase	2018	-7.50	2.68	0.55	2.13	2.07	2.56	1.64	0.07	1.56	-0.02	1.57	1.18	0.00	1.18	1.18	91.80
1. JPMorgan Chase	2019	44.76	3.26	0.73	2.53	1.99	2.59	1.93	0.23	1.70	0.01	1.72	1.34	0.00	1.34	1.34	91.27
1. JPMorgan Chase	2020	-6.67	2.14	0.18	1.96	1.81	2.23	1.55	0.59	0.96	0.03	0.99	0.76	0.00	0.76	0.76	92.33
1. JPMorgan Chase	2021	28.96	1.67	0.06	1.61	1.62	1.98	1.24	-0.28	1.53	-0.01	1.52	1.17	0.00	1.17	1.17	92.58
1. JPMorgan Chase	2022	-14.45	2.50	0.46	2.04	1.51	2.02	1.53	0.18	1.35	-0.07	1.30	1.01	0.00	1.01	1.01	92.09
2. Bank of America	2013	29.84	2.52	0.19	2.33	2.14	2.77	1.70	0.03	1.67	0.09	1.75	1.16	0.00	1.16	1.16	88.59
2. Bank of America	2014	11.95	2.99	0.12	2.86	2.00	3.08	1.78	0.16	1.62	0.09	1.71	1.16	0.00	1.16	1.16	89.59
2. Bank of America	2015	-4.88	2.70	0.11	2.59	1.70	2.41	1.88	0.19	1.69	0.07	1.76	1.19	0.00	1.19	1.19	90.47
2. Bank of America	2016	36.55	2.72	0.12	2.60	1.57	2.20	1.97	0.22	1.76	0.03	1.79	1.19	0.00	1.19	1.19	90.87
2. Bank of America	2017	33.10	2.94	0.18	2.75	1.51	2.16	2.10	0.20	1.91	0.01	1.92	1.24	0.00	1.24	1.24	90.94
2. Bank of America	2018	-16.08	3.23	0.38	2.85	1.49	2.07	2.27	0.19	2.09	0.00	2.09	1.65	0.00	1.65	1.65	91.39
2. Bank of America	2019	44.32	3.33	0.52	2.82	1.29	2.02	2.09	0.20	1.88	0.01	1.89	1.49	0.00	1.49	1.49	91.05
2. Bank of America	2020	-12.68	2.25	0.13	2.12	1.12	1.89	1.35	0.51	0.84	0.02	0.85	0.69	0.00	0.69	0.69	92.16
2. Bank of America	2021	51.02	1.86	0.04	1.82	0.94	1.71	1.06	-0.18	1.25	0.00	1.25	1.12	0.00	1.12	1.12	92.55
2. Bank of America	2022	-26.61	2.47	0.26	2.21	0.93	1.77	1.38	0.10	1.27	0.00	1.27	1.12	0.00	1.12	1.12	92.24
3. Citigroup	2013	26.45	3.68	0.63	3.06	1.29	2.43	1.92	0.36	1.55	0.03	1.58	1.06	0.01	1.08	1.08	91.87
3. Citigroup	2014	3.61	3.49	0.53	2.96	1.20	2.66	1.50	0.29	1.21	0.00	1.21	0.76	0.01	0.75	0.75	91.72
3. Citigroup	2015	-4.35	3.44	0.48	2.97	1.27	2.35	1.88	0.39	1.49	0.02	1.51	1.00	0.00	0.99	0.99	91.60
3. Citigroup	2016	17.31	3.51	0.55	2.97	1.03	2.21	1.79	0.41	1.37	0.04	1.41	0.96	0.00	0.95	0.95	91.05
3. Citigroup	2017	24.60	3.58	0.69	2.89	1.04	2.16	1.77	0.45	1.33	0.05	1.37	0.07	0.00	0.05	0.05	91.03
3. Citigroup	2018	-28.43	3.94	0.96	2.99	1.09	2.11	1.97	0.44	1.53	0.01	1.54	1.18	0.00	1.18	1.18	92.88

Fig. 1. A part of the final dataset used for building prediction models.

5.2 Handling Missing Values

In financial data analysis, missing values are a common challenge that requires careful consideration to ensure accurate and reliable results. Especially for time series data, absent values will make it challenging to construct a reliable prediction model. However, in reality, managing missing values in time series data is unavoidable and crucial, as a variety of factors and issues disrupt the data collection process daily, such as holidays, lost documents, recording errors, etc.

The authors of this investigation selected the exclusion and zero-filling strategies. This includes excluding columns with numerous missing values that cannot be reliably supplied. This study retains variables with a small number of missing values in the analysis and fills in the missing rows with zeros. This method allows us to retain most of the available data while mitigating potential distortions caused by a small number of missing values.

After this step, the data has 296 rows (not including the first row) and 39 columns, which has the column for Name of Bank, Year, TSR, and 36 metrics.

5.3 Data Reduction - Collinearity

Collinearity, also known as multicollinearity, identifies and evaluates the presence of high correlations or linear relationships between predictor variables in a statistical model. Collinearity occurs when two or more predictor variables in a regression or predictive modeling context are highly correlated, making it difficult to distinguish their individual effects on the dependent variable.

In this step, the authors ran a heatmap to detect the multicollinearity among variables. Some variables will be excluded before working on the feature selection step.

- **Dropped R1 and R2**. Reason: R3 is calculated by subtracting R1 from R2.
- **Dropped R15**. Reason: The R15 has similarities with R16 with the difference in the adjusted net income. Their high correlation shows that the adjusted part is insignificant. Hence the authors decided to remove the R15 - Net Income Adjusted Sub S/ Average Assets (15).

- **Dropped R19 and R20**. Reason: R21 is calculated by subtracting R19 from R20.

After this step, the number of metrics besides TSR is decreased to 31 metrics.

5.4 Data Preparation for Machine Learning Models

This study will explore different approaches to build a reliable prediction model for TSR. For models taking TSR as the dependent variable, all 31 metrics extracted from UBPR will be used as the independent ones.

Several ways to split the data into training and testing datasets have been explored. The authors decide to split the data into training and testing datasets with a ratio of 8:2, where the training dataset will contain all records in the first 8 years (2013-2020) and the testing dataset will contain all records in the latest 2 years (2021 and 2022). Note that the data is treated as panel data.

There are 5 machine learning methods used in this study to compare and evaluate their performance to predict the value of TSR, listed in the full name and abbreviation below:

- Vector Autoregression (VAR)
- Random Forest Regression (RF)
- XGBoost Regression (XGB)
- AdaBoost (ADA)
- Deep Neural Network (DNN)

5.5 Performance Evaluation Metrics

This study employs four metrics to assess model performance. Two primary metrics are used to evaluate machine learning models: R-squared (measuring model fit) and MAPE (measuring model prediction error).

- Root Mean Square Error (RMSE)
- Mean Absolute Error (MAE)
- R-squared (R^2)
- Mean Absolute Percentage Error (MAPE)

 ISBN: 978-604-80-9774-5

6 Machine Learning Models

6.1 Feature Selection

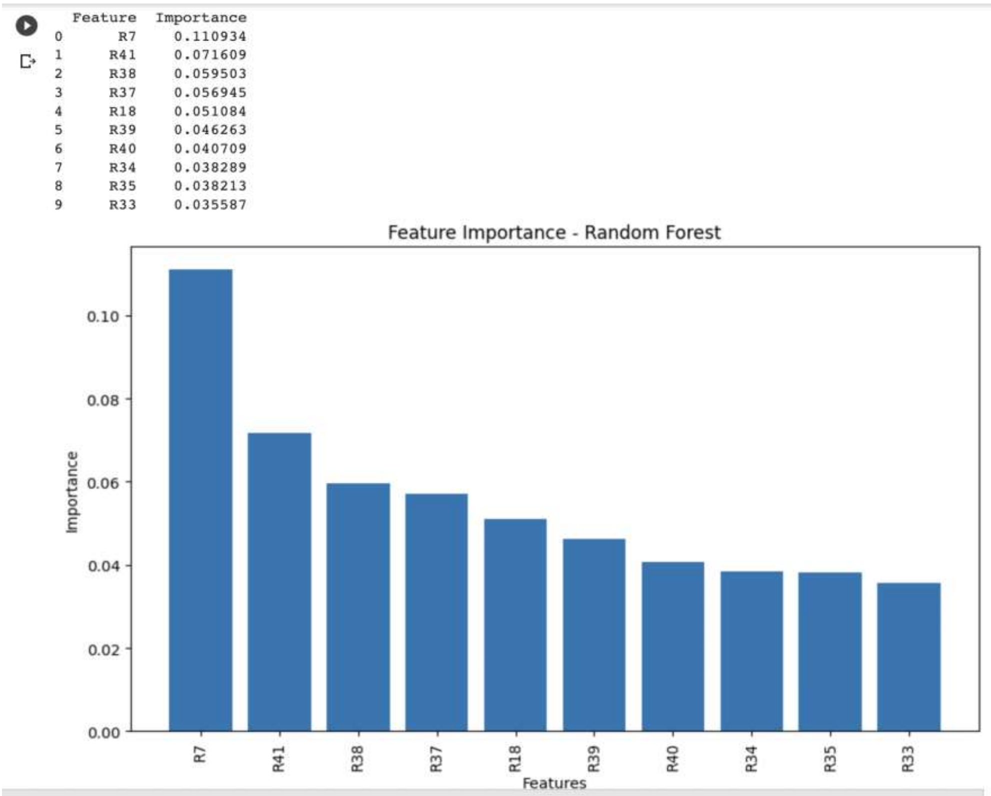

Fig. 2. Feature selection of TSR, retrieved from Google Colab.

As shown in the graph, R7 (Provision: Loan & Lease Losses to Average Assets) is the variable having the strongest prediction power for all models for TSR. All variables in group “Growth Rates" (R37-R41) appeared in the top 10 features, considered they are important influential factors in determining the performance and returns of the banks.

From the correlation heatmap below, it can be observed that TSR has little to no correlation with 31 metrics. R7 continues to be a variable that has a significant relationship with TSR, having the highest correlation value (0.11). The top 10 variables having a high correlation with TSR are in different groups, with the absolute value ranging from 0.06 to 0.09. Based on these results, the authors decide to select all variables for the prediction model.

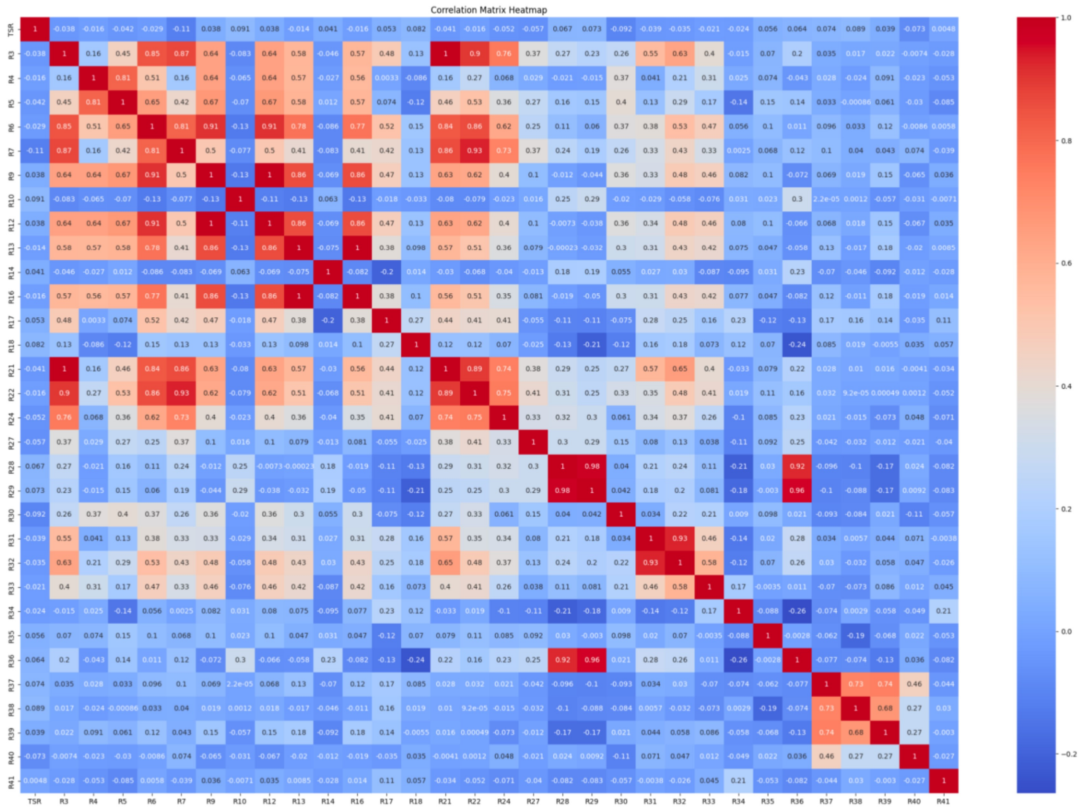

Fig. 3. Correlation heatmap for TSR, retrieved from Google Colab.

6.2 Machine Learning Models

Table 1. Results from machine learning models for TSR.

Method	RMSE	MAE	R-squared	MAPE
RF	31.5834	27.7770	0.1018	306.3049
ADA	30.5013	26.4095	0.1623	286.9792
XGB	33.0571	27.8628	0.0160	337.3246
VAR	41.3222	13.0423	0.1986	347.2531
DNN	28.3718	23.5897	0.2751	325.9167

 ISBN: 978-604-80-9774-5

As the table above presents, DNN is the model with the highest fitting degree (R-squared = 0.27, MAPE = 325.91), while ADA has the best prediction effect (R-squared = 0.16, MAPE = 286.97).

7 Conclusion and Recommendation

7.1 Conclusion

This study explores different machine learning approaches to predict TSR throughout a particular period (2013-2022) for the top 30 biggest commercial banks in the U.S., using metrics from financial reports retrieved from an open-access database of the U.S. government (FFIEC). The results of the best models are evaluated by R-squared and MAPE, and they show that DNN is the model with the highest fitting degree, while ADA has the best prediction effect. A feature selection using random forests and a correlation matrix are conducted to show the top variables impacting the prediction model for TSR.

7.2 Recommendations for Investors

At an investor's glance, profitability is one of the first priorities when considering potential investments. The banking industry often presents a compelling proposition for investors seeking long-term growth and stability. According to the Global Banking Annual Review by McKinsey (2022) [8], global revenue in the industry reached a 14-year high in 2022, growing by around 345 billion USD after a decade of plateau returns.

As a result, to measure directly what investors will get when allocating their money in the banking sector, TSR should be well-understood. After examining some US commercial and investment banks' financial reports, provisions and funding activity are believed to have the most influence on the shareholder return metric. On the side of income statements, the provision for loan and lease losses is often treated as an expense item or a type of 'contra-asset' account (Balla et al., 2012, [4]). Although it could be accepted that the higher the provision, the lower the potential value that shareholders might have, the provision is still crucial for risk management. On the side of funding activity, the short-term non-core describes how federal funds as well as commercial paper were purchased and managed. This activity often generates regular returns, which contribute positive value to investors with such short-term assets.

7.3 Limitation of Study & Recommendations for Further Research

Due to the low number of banks and the short period of time, plus the fact that the total amount of training and testing dataset is too small, overfitting occurs, despite some efforts to conduct hyperparameter tuning and modifications to improve the models of the author. For machine learning methods, it is recommended to collect more data from other commercial banks to do further research, thus giving the model a better generalization effect and making more accurate predictions.

References

1. N. O. Adeboye & N. O. Alabi: Deep-Learning Modelling of Dynamic Panel Data for African Economic Growth. Journal of Econometrics and Statistics, 2(1), 47-60 (2021).
2. Al-Dmour, A. H. & Al-Dmour, R. H.: Applying Multiple Linear Regression and Neural Network to Predict Business Performance Using the Reliability of Accounting Information System". International Journal of Corporate Finance and Accounting (IJCFA), 5(2), 12-26 (2018).
3. Ataünal, Levent; Gürbüz, Ali Osman; Aybars, Asli: Does high growth create value for shareholders? Evidence from S&P500 firms. European Financial and Accounting Journal, 11(3), 25-38 (2016).
4. E. Balla, J. R. Morgan, and J. Romero: Loan Loss Reserve Accounting and Bank Behavior. Federal Reserve Bank of Richmond or the Federal Reserve System (2012).
5. M. B. Barnes, V. C. S. Lee: Feature Selection Techniques, Company Wealth Assessment and Intra-sectoral Firm Behaviours. In: Huang, DS., Heutte, L., Loog, M. (eds) Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues. ICIC 2007. Lecture Notes in Computer Science, vol 4681. Springer, Berlin (2007).
6. A. Masaya & K. Nakagawa: Cross-sectional Stock Price Prediction using Deep Learning for Actual Investment Management. In: 2020 International Artificial Intelligence and Blockchain Conference (AIBC 2020), 9-15, Association for Computing Machinery, Nagoya (2020).
7. L. W. Mburu: The Relationship Between Working Capital Management and Total Shareholder Return of Manufacturing Firms Listed at the Nairobi Securities Exchange. Master's degree research project, School of Business, University of Nairobi, Nairobi (2018).
8. McKinsey: McKinsey's Global Banking Annual Review, retrieved from: https://www.mckinsey.com/industries/financial-services/our-insights/global-banking-annual-review (2022), last accessed 2024/01/01.
9. P. Mingsakul: Tech Trends in the Banking Sector in 2023. Retrieved from https://www.krungsri.com/en/research/research-intelligence/tech-trend-2023 (2023), last accessed 2024/01/01.
10. Ding, Xiafei & Wang, Weiya & Zhang, Yiqian & Zhong, Xiaoyuqian: Machine Learning-based Models for House Price Prediction in Provincial Administrative Regions of China. 10.2991/aebmr.k.220307.035 (2022).
11. Federal Financial Institutions Examination Council's (FFIEC) Homepage, https://cdr.ffiec.gov/public/Default.aspx, last accessed 2024/01/01.
12. Yahoo Finance Homepage, https://finance.yahoo.com/, last accessed 2024/01/01.

 ISBN: 978-604-80-9774-5