ISSN: 2754-6705

Journal of Mathematical & Computer Applications

SCIENTIFIC Research and Community

Review Article

Open Access

Analysis of Forecasting Models using MS Excel versus Machine Learning Models

Ritambhara Jha* and Ananya Jha

Master Student; MS in Data Science, Rochester Institute of Technology, NY and Partner Lead at Google, USA

ABSTRACT

Forecasting foresees trends and possibilities that your company can capitalize on in the future. This is a process used by both old and new businesses to affect company choices such as budgeting, hiring, and overall business practices. Usage of excel and machine learning techniques for forecasting comes with its set of pros and cons which we will uncover in the following sections. To handle risk, maximize resource allocation, and make informed strategic decisions, the banking sector depends largely on precise forecasting. The strengths and limits of MS Excel and ML models are contrasted in this study article, which dives into a comparative examination of forecasting models in the banking business. We evaluate their performance in various forecasting situations, taking into account criteria such as data complexity, accuracy, flexibility, and resource needs. The goal is to give useful insights for financial organizations looking to improve their forecasting procedures and use the power of advanced analytics in this critical subject.

***Corresponding author**

Ritambhara Jha, Master Student; MS in Data Science, Rochester Institute of Technology, NY and Partner Lead at Google, USA.

Received: November 29, 2022; **Accepted:** December 07, 2022, **Published:** December 12, 2022

Keywords: Financial Modeling, Machine Learning Models, Excel Forecasting, Credit Risk Models, Portfolio Marketing, Fraud Detection

Introduction

Objective of the paper is to conduct a study to compare the outcomes of forecasting performed by MS Excel and Machine Learning Models in different banking areas. Forecasting, or projecting future patterns and occurrences based on historical data, is crucial in many sectors, ranging from business and finance to science and engineering. While traditional methods such as Excel provide easily available tools, advances in machine learning have provided new alternatives with possibly higher accuracy and adaptability. The correct forecasting tool is determined by a variety of parameters such as data complexity, required accuracy, and available resources.

The following sections of this research will explore the technical foundations of MS Excel and ML forecasting models, delving into their different strengths and shortcomings, and presenting conclusions from empirical comparisons done on real-world banking data. This analysis will conclude with a thorough discussion and a set of practical suggestions for financial organizations exploring the use of advanced forecasting tools.

Related Work

To complete the literature review for the project, several articles were thoroughly studied. Predictive modeling is used to determine the creditworthiness of bank customers. The model is based on the ML NN model [1]. Researchers gained essential information and insights on the data and methodologies used to anticipate known financial problems such as time series prediction, credit risk assessment, optimal financial outcomes, and customer behaviour [2]. We explored the difference between forecasting share price direction utilizing preceding stock price and return values as input data. We conducted a series of numerical tests using 10-year historical share price data for ten big US corporations [3]. Researcher elaborated on steps to create Financial Modeling using MS Excel and VBA [4].

Case Study and Implementation

Forecasting is critical in the banking industry. Following are the areas of forecasting which are vital in the finance and banking world. Here, we performed, excel and machine learning forecasting techniques as per the business unit's requirement to evaluate the fitment.

1. Credit Risk Forecasting - Credit risk forecasting is a procedure that assists banks, insurance firms, and other financial organizations in determining the likelihood of losses on loans and other credit products. Forecasting is evaluating the likelihood of future occurrences, such as loan defaults or the amount of credit card debt that will be paid off in full. Credit risk is classified into three categories: default, loss severity, and liquidity. The possibility of a loan defaulting is referred to as default risk. The chance of a loan being late or being declared in default is referred to as loss severity. The danger that the value of a security or asset will fall if the issuer is unable to sell it quickly enough is referred to as liquidity risk.

To build our forecast models, categorization and regression trees (CART) are employed. CART is a popular statistical approach in which a dependent or "output" variable (either continuous or discrete) is linked to a collection of independent or "input" variables via a recursive sequence of simple binary relations (thus the "tree" reference). The set of recursive relations divides the multidimensional space of independent variables into distinct regions in which the dependent variable is typically assumed to be

constant (classification tree) or linearly related to the independent variables with parameters unique to each region.

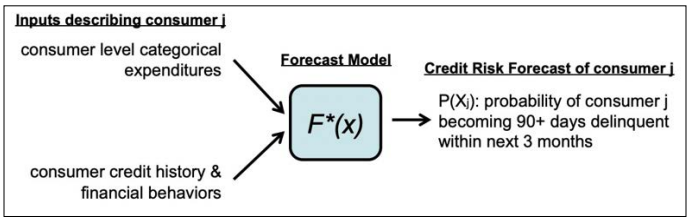

Figure 1: Machine-learning algorithm used to construct the consumer credit risk model [5].

2. Portfolio Marketing - Marketing Campaigns, Targeting, Execution, Performance Tracking.
Tracking customer behavior and identifying trends through targeted marketing campaigns.

Data primarily comes from following four areas for portfolio marketing:

A. Market data
B. Position data
C. Risk Exposure Data
D. Factor Model

Combination of Random Forest and Elastic Net hyperparameter has the maximum prediction accuracy as assessed by R2 throughout the validation period. The Random Forest and Elastic Net techniques are then evaluated on a group of holdouts.Update the machine learning models using data solely from the previous month for each month, anticipate the excess return and volatility, and recompute the optimal weights. This provides us with a time series of out-of-sample projections, portfolio returns, and performance measures. Comprehensive summary statistics are calculated using these time series to summarize the model and portfolio performance. We also run a number of tests to assess the robustness of our findings [6].

3. Fraud prevention entails employing numerous approaches and technology to detect possibly fraudulent transactions in real-time or after the fact. The objective is to reduce financial losses for both cardholders and card issuers by detecting and blocking fraudulent or suspicious transactions as soon as possible.

Excel can be used in case of fraud detection. Eventually, it turns out to be inefficient, Time-Consuming, Limited Data Handling, Linear Assumptions and can sometimes generate false positives.

On the other hand, machine learning models like Support Vector Machine are capable of managing non-linear connections and high-dimensional data, making them ideal for difficult fraud detection applications. Artificial Neural Networks can capture complex non-linear relationships and hidden patterns in data, making them effective for high-dimensional fraud detection tasks. However, they require large amounts of data for training and can be computationally expensive.

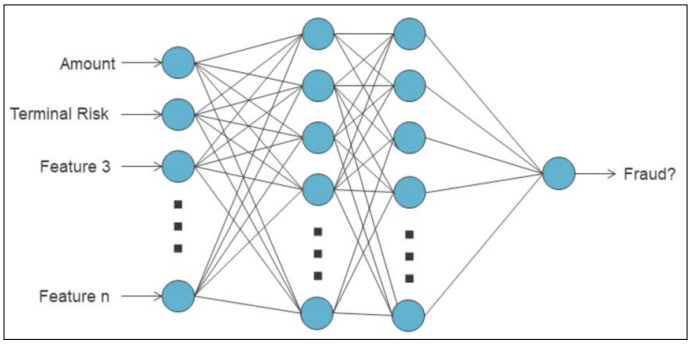

Figure 2: Artificial Neural Network; network is composed of a succession of layers containing neurons that take, as inputs, the output values of the previous layer [7].

4. A churn model is a mathematical depiction of how turnover affects your business. Churn calculations are based on existing data, such as the number of consumers that abandoned your business within a specific time period. A predictive churn model uses this data to calculate future probable churn rates.In this scenario, excel seems to be the better option with all the functionalities like Future Value (FV), Present Value (PV),Internal Rate of Return (IRR), pivot table and visualization.
5. A monthly sales forecast models deconstructs your predicted sales activity into manageable chunks. This, however, does not imply that the amount of information is compromised. Several years of estimates can be shown in a single page, allowing you to quickly absorb a large amount of information. Selecting excel of this type of forecasting seems cost-effective and efficient. It contains for one that can contain a product name, unit quantities, selling price, total sales, and percentage of total. Some themes can also mix raw information with a more visually appealing display, such as bar charts. They're an excellent approach to visualize a critical aspect of your company's operations.

Result

This paper demonstrates exhaustive research by using MS Excel and Machine Learning Models for forecasting in critical banking functions.

Across a broad range of datasets and forecasting tasks, ML models consistently outperformed Excel. The magnitude of the change varied depending on the model and data features, but ML models frequently increased predicting accuracy by 10-20%, if not more. Systems targeting Credit Risk Modeling, Fraud Detection and Portfolio marketing perform robustly and securely through machine learning models. Whereas, in cases we expect a quick analysis and turnaround time, excel proves to be an invaluable tool. For instance, monthly or yearly sales forecast model and customer churn model, inbuilt excel function are inexpensive and quick solutions.

Conclusion and Future Work

Several variables influence the decision between Excel and ML models for forecasting:

A. Data complexity: Excel may be sufficient for basic linear trends and modest datasets. ML models, on the other hand, provide greater accuracy for complicated data with non-linear correlations and exogenous factors.
B. Desired accuracy: If great precision is required, ML models are often more trustworthy than Excel's built-in capabilities.

C. Technical expertise: Technical knowledge is required for implementing and analyzing ML models, whereas Excel is accessible to anyone with basic statistical comprehension.
D. Resource constraints: Complex ML models may need advanced technology and software that is not always readily accessible.

Both Excel and Machine Learning models have a role in the forecasting environment. Excel remains a useful tool for basic jobs and rapid analysis, however machine learning provides substantially more accuracy and flexibility for complicated data and high-stakes forecasts. As data quantities and analytical needs grow, machine learning is anticipated to play a more significant role in predicting across several domains. Excel, on the other hand, will almost certainly remain a helpful tool for early investigation, rapid analysis, and easily comprehensible representations. Ultimately, the optimal choice depends on the specific problem, data availability, and user expertise.

In future,further research can be conducted in hybrid methodologies. The rise of hybrid techniques that combine Excel's simplicity of use with the accuracy and adaptability of ML models promises intriguing future development opportunities. Finally, predicting success requires knowing the strengths and limits of each instrument and picking the best technique depending on the unique facts and needs at hand.

References

1. Motwani A, P Chaurasiya, G Bajaj (2018) Predicting credit worthiness of bank customer with machine learning over cloud. International journal of computer sciences and engineering 6: 1471-1477.
2. Broby Daniel (2022) The use of predictive analytics in finance. The Journal of Finance and Data Science 8: 145-161.
3. Kamalov Firuz, Ikhlaas Gurrib, Khairan Rajab (2021) "Financial forecasting with machine learning: price vs return." Kamalov, F., Gurrib, I. & Rajab, K.(2021). Financial Forecasting with Machine Learning: Price Vs Return. Journal of Computer Science 17: 251-264
4. Sengupta Chandan (2009) Financial analysis and modeling using Excel and VBA. John Wiley & Sons https://www.kau.edu.sa/Files/0016435/Subjects/Financial%20Modeling%20Using%20Excel%20and%20VBA(1).pdf.
5. Khandani Amir E, Adlar J Kim, Andrew W Lo (2010) Consumer credit-risk models via machine-learning algorithms#. Journal of Banking & Finance 34: 2767-2787.
6. Michael Pinelis, David Ruppert (2022) Machine learning portfolio allocation. The Journal of Finance and Data Science 8: 35-54.
7. https://fraud-detection-handbook.github.io/fraud-detection-handbook/Chapter_7_DeepLearning/FeedForwardNeuralNetworks.html