Research Journal in Business and Economics
ISSN: 3078-381X
DOI: 10.61424/rjbe
Journal Homepage: www.bluemarkpublishers.com/index.php/RJBE

RJBE
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Predictive Modeling for Business Performance Using Machine Learning Techniques

Kaniz Fatema[1] ✉ Md Sohel Mollick[2], Mohammad Yasin Hasan[3] and S. M. Arefin Shovan[4]
[1]Grand Canyon University
[2]Washington University of Science Technology (WUST)
[34]United Commercial Bank PLC
Corresponding Author: Kaniz Fatema, **E-mail**: nishikanij1@gmail.com

| ABSTRACT
In the age when operations are run with the guidance of data, predictive modeling has become an indispensable art of analyzing and predicting business. This study investigates the use of statistical analysis and machine learning methods in predicting financial performance with reference to information of the top 2000 companies across the globe. The data entails key performance indicators, including revenue, profits, assets, market value, and geographic origin. Visual analytics solutions, such as Tableau, Python, and Excel, help us identify key patterns, correlations, and financial distributions across countries and sectors. The analysis performs an extensive exploratory data analysis, and it provides information about geographic concentration, industry dominance, and financial disparities. In particular, one can speak of economic centralization, as the United States and China have the most companies and the largest market capitalization. A regression analysis further points to the importance of profits as the most important positive driver of business performance, though revenue and market value contrarily reveal negative coefficients - it is not necessarily the case that a high value or revenue is associated with high internal performance. The study shows how financial analysis should be viewed in the context of the area and financial conditions, which determine performance. This study employs a mixed-method approach, combining illustrations with statistical modeling, enabling transparency with predictive analysis. These findings demonstrate the significance of selecting features, normalizing regions, and using multivariate analysis so that predictive performance can be improved. As this paper concludes, when coupled with high-quality financial data, machine learning can provide experience to investors, business analysts, and policymakers. It suggests that external macroeconomic indicators and time-series data should be incorporated further to make more dynamic and time-aware models for future work.

| KEYWORDS

Predictive Modeling, Business Performance, Machine Learning, Financial Forecasting and Company Valuation, Exploratory Data Analysis (EDA)

| ARTICLE INFORMATION

ACCEPTED: 23 September 2025 **PUBLISHED:** 14 December 2025 **DOI:** 10.61424/rjbe.v3.i4.611

1. Introduction
1.1 Background
The modern business environment is characterized by the presence of great global competition, fast technological changes, and ever-changing customer needs (Oduro et al., 2023). In order to survive in such an environment, an organization should not only make prudent operational and financial judgments but should be able to estimate the consequences of its decisions with a significant level of accuracy. Although traditional techniques of forecasting business are helpful, it is yet not possible to overlook the intricacy of cross-dependent financial factors and the dynamism of the international economy. With the emergence of big data and computational technologies, businesses can now gather large amounts of structured and unstructured data that can allow them to draw deeper conclusions about the performance of their organizations, as well (Barchiesi & Colladon, 2019; Valdez et al., 2019). This has translated into the growing interest in data analytics and predictive modeling methods as the key corporate strategy tool (Adesina et al., 2024; Vudugula et al., 2023; Yang et al., 2025). The methods enable organizations to move to

proactive decision making and enable the executives to get to know the variable that contributes most to the performance. The estimates of future revenues, prediction of market trends, and anticipation of risks may be done on historical data and financial indicators using predictive models (Johnson et al., 2024). In addition, with the rising interdependence between countries in terms of global financial ecosystems, learning how business organizations that perform exemplary in different nations can be quite helpful in terms of information lessons and points of reference for other companies. It is against this background that the advanced perspectives of machine learning in forecasting business performance are deliberated upon in view of standard financial measures, such as revenue, profit, assets, and market value. Through practical data, the research will fill the gap between the classical financial analysis and the contemporary predictive analysis and provide scholarly knowledge along with practical usefulness to business executives and researchers.

1.2 Financial Analysis introduction to Machine Learning (ML)

Artificial intelligence - Machine learning has led to a paradigm shift in the manner in which data-driven decisions are made within different industries, such as finance and business (Akter et al., 2020; Sonkavde et al., 2023; YADAVA, 2025). In financial analysis, machine learning models, with the help of their algorithms, are trained on past data to discover patterns, identify anomalies, and make predictions (YADAVA, 2025). As opposed to classic statistical models that necessitate a user to manually define the connection among variables, ML models are able to learn and adjust autonomously; therefore, they can process vast quantities of information comprising many interrelated aspects (Abayomi et al., 2023; Li et al., 2024; Карачун et al., 2021). ML is currently being used by financial institutions and corporations to predict stock prices, credit risks, and default probability on loans, market trends, and customer behaviors (Abayomi et al., 2023; Sonkavde et al., 2023; YADAVA, 2025). Within business performance analysis, ML can be used to determine which financial indicators have the strongest impact on the success of a company, and also enables business analysts to define how business results could turn out in the future using various hypothetical scenarios. Linear Regression, decision trees, random forests, gradient boosting, and other algorithms are popular in supervised learning for predicting continuous variables, such as revenue or profit (Adekunle et al., 2021; Chun et al., 2024; Mohd et al., 2024; Ravi et al., 2024; Zolotareva, 2021). ML also increases the precision of prediction by such methods as cross-validation, feature selection, and hyperparameter tuning (Chandrasekhar & Peddakrishna, 2023; Chapman & Desai, 2023; Debal & Sitote, 2022; Marcelino et al., 2019). Use of ML in financial analysis gives companies a competitive edge since the model offers expedited and accurate findings and data-based strategies (YADAVA, 2025). Due to the business's ability to access financial data sets and the capacity to use computational power to run more and more ML algorithms, there is a high probability that ML will play an even bigger role in business forecasting. This article applies ML models to the investigation of prospective relations among the crucial financial indicators of the leading global corporations, showing how contemporary algorithms may be more efficient than traditional forecasting instruments in terms of assessing the business performance and its possible improvement.

1.3 The Value of Business Predictive Modeling

Given the importance attached to data in the current corporate world, predictive modeling is the actual method that offers future insight to businesses, depending on past and prevailing information (Adekunle et al., 2021). By predicting the drivers of crucial performances like temperatures and the market performance of the company and other performances, it will help businesses in making the right decision, minimizing uncertainty, and attaining a competitive advantage (Johnson et al., 2024). These observations facilitate strategy development, risk evaluation, resource management, and performance indicators. Predictive modeling is the actual method that offers future insight to businesses, depending on past and prevailing information. By predicting the drivers of crucial performances like temperatures and the market performance of the company and other performances, it will help businesses in making the right decision, minimizing uncertainty, and attaining a competitive advantage. These observations facilitate strategy development, risk evaluation, resource management, and performance indicators. Predictive modeling can help companies identify opportunities and threats, optimize pricing and inventory, increase customer satisfaction, and improve financial results by allowing them to forecast future situations based on the current context (Ahmed et al., 2024; Vudugula et al., 2023). Unlike descriptive analytics, which merely describe what happened in the past, predictive modeling predicts what will probably occur, hence changing decision-making to proactive (Johnson et al., 2024). Available there are organized financial data and improvements in the computational tools that have made predictive modeling precise, including scaling, and accessible to a wider audience. The predictive models implemented based on machine learning go further to augment this process by automatically detecting the patterns in big datasets and adapting to the new trends in data without involving manual control (Adekunle et al., 2021). To a business, forecasting quarterly profitability, performance in the stock market, or market value has the prospects of increasing investor confidence, operational consistency, and growth over the long term. Additionally, predictive modeling has the capacity of causing test scenarios that enable the executives to assess the consequences of decisions before actually making them (Samson-Onuorah & Bakare, 2025). Within the scope of the given research, predictive modeling is applied to analyze and predict business performance based on the top 2000 world companies, pointing out how organizations can use data science tools to enhance strategic performance and achieve maximum returns on investment (Brynjolfsson et al., 2021).

1.4 Rationale to Selecting the Dataset of the Largest Companies in the World

The largest companies in the world are based on their richness, their relevance, and their real-life applicability in business performance research (Divo et al., 2024; Oyedokun et al., 2024). The information contained in this dataset rates the most substantial 2000 businesses globally by providing major financial indicators of revenues, profits, assets, and market values, and geographical break-ups (KACAR et al., 2023). It showcases a varied number of companies in different countries and in different industries, so that it gives a global comparative analysis of the business performance (Alaminos et al., 2016; Liu et al., 2023). The introduction of high-impact finance metrics provides a strong foundation for developing predictive models that can evaluate and forecast performance metrics such as revenue and market share. Also, the dataset's structure allows applying machine learning algorithms without many preprocessing steps, as it is clean, labeled, and well-formatted (YADAVA, 2025). It is high-scoped and permits the formation of models that are robust and even broadly applicable across areas and geographies (Liu et al., 2023). The utilization of the actual company-related data increases the practical importance of the research and ensures that the results can be applied by the stakeholders who live in the corporate sector. The dataset update frequency each year is another factor that makes the analytical information up to date and valuable, as it is related to the context of the recent financial conditions. With attention given to the top organizations around the globe, the database can provide insight into what makes success at the top of the business. The applicability of this dataset to real-life corporate performance, besides the accessibility of the critically important predictor variables, indeed renders use of the dataset the ideal choice for presenting machine learning-based predictive modeling analysis across business success performance (Liashenko et al., 2023).

1.5 Research Objectives

- To design machine learning algorithms to predict business performance based on the key financial indicators of successful firms around the world.
- To examine the connection between revenue, profit, assets and market value.
- To use machine learning algorithms in predicting performance.
- To assess the performance of models according to accuracy.
- To find out the most influential determinants of business successfulness.
- To measure the trend in finances relating to various nations.
- To illustrate how effective ML can be in focusing on corporate financial analysis.

1.6 Research Questions

The key question to come up with useful machine learning models to predict business performance is, as mentioned below:

1. Which of the financial variables best represent the company revenue and profit?
2. What is the level of predicting business performance using machine learning models?
3. Are performance predictors different in countries or sectors?

1.7 Significance of Study

This study is relevant because it aimed to cover the increased interest in data-driven decision-making in the business environment, with the approach based on machine learning to forecast financial performance (Adekunle et al., 2021). Businesses are more volatile, with the digital transformation, and with more complicated financial ecosystems, it is more important than ever to be able to access accurate predictive insights (YADAVA, 2025). Findings of the given research can be used in academic literature and in the industry to help effectively enhance the accuracy of performance forecasts and formulate the most important drivers of success of the top global businesses (YADAVA, 2025). This study is an empirical study about the efficiencies of different machine learning algorithms. This helps the data scientist, analyst, and corporate strategist select the right model to use in making financial forecasts (Miller, 2023). Furthermore, the application of a sample with a worldwide representative sample strengthens the relevance of the study as it allows providing cross-country comparisons and sector-based assessment (Hellum et al., 2023). It also facilitates international benchmarking, and multinational firms can use it to gain better insights into the financial trends in other markets (Cakici et al., 2023). The results can also make a panel on investment, risk management, and performance improvement gears (Chowdhury et al., 2024; Nahar et al., 2024; Oyedokun et al., 2024). With regard to academic application, the paper bridges the median between the theoretical and real-life applications of ML and is scalable in the future of business analytics research (Adekunle et al., 2021). This study highlights the disruptive capacities of machine learning in the restructuring procedure through which businesses measure performance, build predictions, and maintain competitive advantage in the data-driven world (Adekunle et al., 2021).

2. Literature Review

2.1 Concept of Data-Driven Decision Making in Business Analytic

The data-driven decision-making and forecasting approach has been fueled by the increasingly difficult business environment (Rovolis & Habibipour, 2023). Companies are currently overwhelmed with huge quantities of data produced on a daily basis through operational, financial, and market activities (Elbaşı et al., 2023). The more common approaches to financial analysis, such as linear modelling models and linear assumptions, do not have the capability of identifying these patterns and associations in the data, thereby remaining inadequate to reveal them (Adekunle et al., 2021; Rao et al., 2025). This has led to the development of predictive analytics through machine learning, but it is a more successful alternative (Adekunle et al., 2021; YADAVA, 2025). The transition to predictive and prescriptive analytics models as opposed to descriptive and diagnostic systems implies the shift that is critical to the approach of organizations with regard to business strategy development, risk evaluation, and business performance enhancement (Nwaimo et al., 2020; Szukits & Móricz, 2023). Machine learning allows companies to automatically identify patterns, trends, and anomalies that they may not see using traditional methods (Adekunle et al., 2021; Elbaşı et al., 2023). This change has also led to the creation of smarter systems, which can also simulate financial situations and recommendations in real time (Aljohani, 2023). The role of sophisticated data modeling becomes critical now as business performance is shaped by new factors, including external variables such as market changes in the global business environment, economic uncertainty, and digital disruption. Firms are now trying to exploit the multivariate financial algorithms that can produce genuine forecasts. This has propelled the interest in using computational intelligence in financial models to make them more accurate and highly elastic (Chiriac, 2020; Sobrinho & Cavalcante, 2023). The use of machine learning here does not simply amount to a technological improvement but a strategic necessity of organizations that aim at maintaining a competitive edge, maximizing the bottom line, and forecasting the future with a greater level of certainty (Akindote et al., 2023; Badmus et al., 2024).

2.2 Machine Learning Models during forecasting and evaluation of financial operations

ML models have been reshaping financial projection and business performance metrics by powering the building of sound, larger, and more flexible prognostic frameworks. AI technology now revolutionizes financial forecasting through its combination of sophisticated machine learning algorithms with big data analytics and automated systems that improve both diagnosis quality and decision-making ability (YADAVA, 2025). These models are able to analyze copious quantities of prior financial information to identify patterns that can be used to have an understanding of future figures like revenue, profit margins, and market capitalization. AI helps analysts spot complex patterns which human analysts cannot detect, enhancing the accuracy of stock market trend predictions alongside credit assessment operations and market demand projections (YADAVA, 2025). Predictive models play a critical role in business optimization by enhancing decision-making, reducing operational inefficiencies, and improving resource allocation (Adekunle et al., 2021).

2.3 Usage of Common Techniques in the Predictive Financial Modeling

The financial sector has greatly impacted the monetary well-being of consumers, traders, and financial institutions. In the current era, artificial intelligence is redefining the limits of the financial markets based on state-of-the-art machine learning and deep learning algorithms (Sonkavde et al., 2023). There is extensive use of these techniques in financial instrument price prediction, market trend analysis, establishing investment opportunities, and portfolio optimization (Sonkavde et al., 2023). Predictive modeling can utilize statistical approaches, such as regression analysis and time series forecasting, or more advanced machine learning techniques, including decision trees and neural networks (Adekunle et al., 2021). A more accurate and generalizing model, Random Forest Regression, is based on constructing multiple decision trees and averaging their outputs, thereby providing better accuracy and reduction of variance. Deep learning- and machine learning-based ensemble techniques have gained popularity due to their superior performance (Sonkavde et al., 2023). For instance, a study involving ARIMA, LSTM, and Random Forest for price forecasting found that XGBoost, an ensemble learning algorithm, outperformed ARIMA and LSTM based on evaluation parameters (Sonkavde et al., 2023).

2.4 Shortage of Global Comparative Modeling in the Previous Study

Although in recent years, people have started to use machine learning in various fields of financial analysis, a certain gap has existed concerning research on the topic of a global comparative modelling of business performance. Most of the already done studies are restricted in terms of industry, geographical location, or market, and apply localized data that fail to consider the financial difference and variations in operations across countries (Fieberg et al., 2022; Laitinen et al., 2023). Since today's business is conducted in a global economy, it is critical to compare the performance measures across countries, especially for multinational companies and their investors. Factors like regulatory environments, the maturity of the market, currency stability, and trends in the region have the ability to greatly impact the financial performance of a company (Laitinen et al., 2023). But such elements are not always captured in the conventional financial modelling process. This non-cross-nation analysis prevents cross-nation benchmarking of performance and building predictive models that are really representative of different business situations (Koshiyama et al., 2021). Besides, the available studies employ small-scale data, which include the analysis of the companies of one industry and of one country, in which the results may be generalized (Lommers et al., 2021). This study fills this gap because the data used is composed of the largest 2000 contemporary companies worldwide, which provides the status of financial performance in various geographic and economic realities comprehensively. This study is a contribution to a more comprehensive idea of what makes businesses successful in the entire world by introducing a large variety of financial indicators and using machine learning methods. It also helps to develop those models, which are able to incorporate differences in terms of culture, economy, and structure, providing rich insights into the international business strategy and the comparative performance analytics (Liu et al., 2023).

2.5 Combining Non-Financial and Financial Variables in Better Forecasting

Incorporation of both non-financial and financial variables to boost performance and explanation of a model is one of the emerging debates in predictive business modeling. The conventional finance models are largely quantitative-based and only use revenue, profit, and assets as indicators. Although they are imperative when it comes to performance understanding, they fail to reflect on the expanded factors that can affect the success of businesses, e.g., organizational or environmental.

There are non-financial factors that could be decisive in determining the performance of a company, including geographic location, governance type, digital maturity, and industry and sector. The introduction of these variables in machine learning models gives a broader picture of dynamics in business by incorporating external variables and enhancing the understanding of complex patterns (Aljohani, 2023). For example, companies located in economically stable states or countries with friendly trade strategies might exhibit better performance than others with the same financial fundamentals. Similarly, companies that are well developed in digital terms or those that have a good brand equity can command better market values even when they are not making much profit. The applied machine learning models will easily handle the data diversity and complexity at that level.

The combination of the two types of variables will enhance predictive models, providing some advanced insight into the data and thus being able to improve strategic decisions. Through the "Largest Companies in the World" dataset that provides information concerning the country and some significant financial indicators, one has an opportunity to investigate these interactions empirically. By incorporating this holistic approach, one can improve forecasting as it is based on the models that are able to include the contextual factors and invisible dependencies that cannot be seen in financial data alone (Chapman & Desai, 2023). Finally, this will enhance the validity and usefulness of predictive modeling by making it more useful to investors and corporate decision-makers who need to gain more insights about what drives the performance of the company (Adekunle et al., 2021).

2.6 Empirical Study

The article titled Machine learning for predictive models in entrepreneurship research: predicting entrepreneurial action is another example in the literature that empirically demonstrates the application of the machine learning model to predictive analytics in social science. Using Global Entrepreneurship Monitor data, the study compares traditional logistic regression with machine learning approaches using XGBoost and Artificial Neural Networks. Evaluation of the models based on performance measures of accuracy, sensitivity, specificity, and AUC revealed better performance by XGBoost over other models in most of the measures. Notably, the paper singles out some of the key variables, which include self-efficacy and opportunity, as significant predictors of undertaking an entrepreneurial action. There is a particular methodological analogy with the aim of the predictive modeling of business performance, which this research highlights, the robust coverage of machine learning in nonlinear, complicated connections in behavioral information. The empirical findings favor the use of ML models at a wider business level, which makes the deployment of ML models in predicting performance legitimate.

In their article titled Crop Prediction Model Using Machine Learning Algorithms, Elbasi et al. also focus on the efficacy of machine learning in agricultural prediction through an empirical research study (Elbaşı et al., 2023). In the research, the authors benchmarked fifteen various ML algorithms, such as Bayes Net, Naïve Bayes, and Hoeffding Tree, to forecast the result of a crop based on real-time data provided by IoT sensors with farm-level variables (Elbaşı et al., 2023). The Bayes Net algorithm also showed a high classification accuracy of 97.05%, signifying the strong predictive capabilities of sophisticated ML approaches (Elbaşı et al., 2023). They also proposed a new combination scheme of features with regard to the enhancement of accuracy, emphasizing that appropriate feature selection is critical to achieving better accuracy in machine learning algorithms while analyzing agricultural data (Elbaşı et al., 2023).

In the article entitled A Review of Evaluation Metrics in Machine Learning Algorithms, Naidu, Zuva, and Sibanda perform an inclusive review of the types of performance evaluation that can be applied in machine learning algorithms (Miller et al., 2024; Plevris et al., 2022). The paper highlights the significance of using the right assessment metrics, such as Accuracy, Precision, Recall, F1-Score, ROC/AUC, and Kappa statistics, especially in the context of evaluating binary, multi-class, and multi-label classifiers (Kratsch et al., 2020). According to the authors, the lack of standardization in evaluation practices prevents any viable comparison of models, at least in a real-world scenario. Although it does not work on a particular industry, the study is especially applicable in the area of business performance forecasting because mastery of an accurate model and strategic decisions are directly correlated. In this review, there is a need to have a consistent evaluation framework used in the development of the predictive models, model selection, and performance reporting should be robust, interpretable, and in line with the business goals.

The IEEE article authored by Abdul Bujang et al., titled Multiclass Prediction Model for Student Grade Prediction Using Machine Learning, is an excellent contribution to the research, being an empirical study, as the authors use six machine learning algorithms to use a real-world student dataset (BUJANG et al., 2021). The paper solves the problem of data imbalance with the help of the Synthetic Minority Oversampling Technique and proposes a multiclass predictive model with two feature selection techniques. The study found that a combination of SMOTE and feature selection resulted in high f-measure performance, with kNN achieving up to 99.7% and J48 achieving 99.4% (BUJANG et al., 2021). This indicates its outstanding performance in the classification task. Though conducted in the field of education, the study can be quite easily transferred to the sphere of business performance prediction due to the methodology and results of the study. It shows how typical ML models can be optimized for multi-class and real-world problems and how to reduce the problem of data imbalance to help in improving the effectiveness of prediction. These experimental results are important in the design of predictable and scalable ML-based predictive solutions in business.

In the article by Marcelino et al. titled Machine learning approach to pavement performance prediction, the authors demonstrate a comprehensive framework that addresses the formulation of predictive models that make use of machine learning in pavement management systems. As a micro-survey, the investigation implements the Random Forest algorithm to predict the International Roughness Index based on the historic information, structural conditions, climate, and traffic attributes at the period of 5- and 10-year outlooks. The two major concerns that the study touches upon are generalization performance and model reliability; the latter is of great importance to business performance forecasting. The framework provided assumes a similar approach as in civil engineering, but the longer-term predictive purposes of ML are emphasized. One of the facts noted in the paper is that researchers should rely on strong data and compare several algorithms to make effective predictions. Such empirical observations can lay the foundation of data-driven, predictive business models since they explain how ML can be used to maximize resource management and planning by making accurate predictions.

3. Description and Preprocessing of Data

3.1 Overview Of the data set

Rank	OrganizationName	Country	Revenue	Profits	Assets	MarketValue
1	JPMorgan Chase	United States	179.93 B	41.8 B	3,744.3 B	399.59 B
2	Saudi Arabian Oil Company (	Saudi Arabia	589.47 B	156.36 B	660.99 B	2,055.22 B
3	ICBC	China	216.77 B	52.47 B	6,116.82 B	203.01 B
4	China Construction Bank	China	203.08 B	48.25 B	4,977.48 B	172.99 B
5	Agricultural Bank of China	China	186.14 B	37.92 B	5,356.86 B	141.82 B
6	Bank of America	United States	133.84 B	28.62 B	3,194.66 B	220.82 B
7	Alphabet	United States	282.85 B	58.59 B	369.49 B	1,340.53 B
8	ExxonMobil	United States	393.16 B	61.69 B	369.37 B	439.39 B
9	Microsoft	United States	207.59 B	69.02 B	380.09 B	2,309.84 B
10	Apple	United States	385.1 B	94.32 B	332.16 B	2,746.21 B
11	Shell	United Kingdom	365.89 B	43.51 B	429.15 B	205.45 B
12	Bank of China	China	158.23 B	33.23 B	4,421.76 B	122.67 B
13	Toyota Motor	Japan	270.49 B	18.5 B	542.5 B	188.2 B
14	Samsung Electronics	South Korea	220.07 B	34.49 B	348.81 B	334.31 B
15	UnitedHealth Group	United States	335.94 B	20.7 B	283.68 B	460.19 B
16	Ping An Insurance Group	China	166.37 B	12.64 B	1,598.49 B	138.56 B
17	Wells Fargo	United States	108.93 B	14.5 B	1,886.4 B	142.36 B
18	Chevron	United States	232.12 B	35.78 B	255.89 B	303.54 B
18	PetroChina	China	457.4 B	22.18 B	384.58 B	122.91 B
20	HSBC Holdings	United Kingdom	73.98 B	22.19 B	2,989.7 B	151.19 B
21	TotalEnergies	France	257.59 B	21.12 B	293.03 B	151.64 B
21	Verizon Communications	United States	136.19 B	21.59 B	377.72 B	159.04 B
23	Walmart	United States	611.29 B	11.68 B	243.2 B	409.12 B
24	Citigroup	United States	120.42 B	15.02 B	2,455.11 B	90.03 B
25	China Mobile	Hong Kong	140.53 B	18.76 B	290 B	186.41 B
26	China Merchants Bank	China	72.44 B	20.56 B	1,458.48 B	129.84 B
27	Postal Savings Bank Of China	China	82.51 B	12.78 B	2,023.57 B	110.18 B
28	BP	United Kingdom	248.11 B	25.89 B	278.66 B	108.3 B
29	Volkswagen Group	Germany	293.47 B	15.63 B	633.78 B	70.16 B
30	Morgan Stanley	United States	74.16 B	10.34 B	1,199.9 B	141.76 B
31	Meta Platforms	United States	117.35 B	21.44 B	184.49 B	599.82 B
32	Sinopec	China	453.56 B	9.85 B	283.3 B	114.32 B
33	BNP Paribas	France	99.47 B	10.11 B	2,845.69 B	78.59 B
33	Goldman Sachs Group	United States	79.45 B	10.56 B	1,538.35 B	108.72 B
35	Tencent Holdings	China	82.41 B	27.26 B	227.01 B	415.36 B
36	Amazon	United States	524.9 B	4.29 B	464.38 B	1,084.06 B
37	Allianz	Germany	134.26 B	7.08 B	1,139.49 B	95.45 B

Largest companies in world | Sheet1

The dataset employed in the given research, namely, the dataset entitled "Largest Companies in the World," is highly detailed as it contains financial statistics of the 2000 biggest enterprises in the world that represent a wide array of industries and countries. It has seven important attributes, such as company rank, organization name, country, revenue, profit, assets, and market value. These indicators are part of critical financial factors used to measure the performance of the business. The data format itself comes in a tabular manner, which is appropriate in data science. It is updated on an annual basis and covers more than 1,500 distinct entries on companies in more than 100 countries. Such diversity guarantees a rich and global representation both when it comes to comparative analysis and machine learning generalization of models. All of the company entries provide quantitative information that could be utilized in regression tasks, and a categorical variable that initializes the feasibility of region division and feature encoders. The set of data has both balance sheet items and performance indicators, which makes it especially appropriate in predictive modeling. Since the data set entails the sample of companies at different geographic locations and diverse economic statuses, cross-country benchmarking and analysis of financial behavior in different sectors is possible. The fact that its usability score is 10.0 and its license type is MIT proves its quality and academic accessibility. The extensive scope of the dataset and the balanced financial metrics make it one of the ideal starting points when it comes to applying machine learning techniques in predicting the outcome of a specific business in terms of identifying the drivers of actual performance and analyzing trends in the context of the global business environment.

3.2 Cleaning and Preprocessing of Data

In advance of the application of the machine learning models, the dataset was to be ready, and several preprocessing steps were applied to it (Marcelino et al., 2019). Data cleaning is crucial as unaccounted numbers, missing values, and inconsistent data can negatively affect the performance of machine learning models (Islam et al., 2023; YADAVA, 2025). To begin with, the monetary values of the revenue, profit, assets, and the market value were provided along with the suffix B or M, which were to be translated into numerically converted float values to be used in calculation. This was transformed by dropping off the suffixes and scaling up the values as multiples of the suitable scale factors. Besides, there were columns with company names and a rank column that were not directly relevant to prediction activities; hence, they were not included as features to avoid adding extra noise to the feature set. Categorical column country containing the name of the country in which every company is located was one-hot-encoded to change it into a machine-readable form, but still retain the distinction on a country-level (Islam et al., 2023). The issue of outliers was also addressed, and missing or invalid values were handled either by imputation or by excluding that case (Zekić–Sušac et al., 2020). Normalization of data was done to make sure that financial values were on a similar scale, which is vital given the fact that some models are sensitive to input magnitudes like the linear Regression (Hasan & Ferdawsi, 2025; Ridwan et al., 2020). This process involved adjusting numeric columns to a standard scale without distorting discrepancies in the spectrum of values or losing information (Ridwan et al., 2020). Lastly, the data was divided into the training and test sets in order to validate against unseen data. This organized cleanup process made the data clean and consistent, appropriate for use in the training and evaluation of predictive machines with better reliability.

3.3 Exploratory Data Analysis (EDA)

An exploratory data analysis was utilized to get preliminary information about the dataset and to discover possible relationships among the variables (Mostafa et al., 2022). Each numerical variable was described in terms of summary statistics, including mean, median, and standard deviation, and quartile distributions. These figures showed that revenue, profit, assets, and the market value were significantly different between businesses, with the financial diversity of the global sample. Correlation analysis is useful for feature selection, especially when two features have a very high correlation, as one might be ignored without losing significant predictive power (Dev et al., 2022). Correlation was done to study linear associations between features, where a highly positive correlation was found between assets and market value, and between revenue and profit. Visual confirmation of these patterns was obtained with the help of a correlation heat map that helped identify the most suitable variables to be predicted (Islam et al., 2023). To identify scenes and the existence of outliers, boxplots and histograms of every financial metric were produced (Mostafa et al., 2022). The distributions of revenue and profit were right-skewed, with some companies that were stratospherically high; thus, a log transformation in such model specification was used to normalize the distortion in the distribution (Hasan & Ferdawsi, 2025). The visual inspection of the relationships between variables, revenue versus assets, profit versus market value, etc., was conducted on scatter plots, which confirmed that some of the relationships are linear, but some suggested nonlinear dependencies. The country-level pattern of EDA also demonstrated that some of the companies representing particular nations consistently have higher revenues and market values, which indicates possible geographical influence on company performance. Such observations were used during feature selection and model design, and decisions were made on variables to include and how to deal with skewed distributions (Dev et al., 2022). Dimensionality reduction techniques like Principal Component Analysis can also be used to understand inter-relations amongst features and optimize feature sets (Dev et al., 2022). EDA ensured important insights into the structure of the dataset and guided data-informed decisions to be made during modeling.

4. Methodology

4.1 Research Method and Forecasting Modeling Framework

The paper uses a quantitative approach to research that relies on statistical and empirical data analysis and machine learning models. The general idea is to forecast vital measures of business performance, i.e., revenue, profit, and market value, based on financial and categorical parameters of the largest companies in the world. It uses a supervised machine learning system, and learned models are trained on labeled input data (YADAVA, 2025). It is decided that regression analysis forms the essence of the task of modeling, since the dependent variables are of a continuous nature (Adekunle et al., 2021). The approach is such that data is prepared, the model training, evaluation, and validation process is followed by feature selection. The work examines the effectiveness of several machine learning algorithms based on multiple regression, and it also enables understanding which model is most advantageous for predicting financial performance and determining the most active drivers of the performance trend. To ensure model generalizability and robustness, the dataset is split into training and test sets at a 80:20 ratio. Performance is evaluated using standard evaluation metrics to assess model accuracy and interpretability. The process also involved tuning the hyperparameters and cross-validation when applicable so as to improve model settings and alleviate overfitting (Sonkavde et al., 2023). Such an approach to methodology design will allow creating sustainable and robust models with predictive capabilities at the global business level.

4.2 Forwarded Machine learning models

This research employs four machine learning algorithms in order to predict business performance measures: Linear Regression, DT Regression, Random Forest Regression, and XGBoost Regression. Every model offers distinctive features that are applicable in analyzing each level of complexity in financial datasets (Sonkavde et al., 2023). Linear Regression is used as a baseline model as it is simple and easier to interpret; it results in a linear relationship between the independent and dependent variables (Sonkavde et al., 2023). The Decision Tree Regression is a nonlinear regression that divides the dataset into branches according to the set rules of decision, on which complicated interactions of variables are captured (Adekunle et al., 2021). It is, however, prone to over-fitting unless pruned. Random Forest Regression is an ensemble learning method that creates multiple decision trees and averages their outputs to increase fidelity and reduce variance (Hewage et al., 2020; Sonkavde et al., 2023). It is especially useful when dealing with high-dimensional data that can have multicollinearity. A more sophisticated ensemble approach, XGBoost, extends the idea of boosting with an increasing number of trees whose purpose is to correct the mistakes of the trees constructed previously (Sonkavde et al., 2023). It is scaled, well-regularized, and shows good performance in predictive tasks (Sonkavde et al., 2023). The same training dataset was used to train these models and was compared with each other on a set that was reserved as a test to ensure comparisons. The reason why the tree-based models were used is to obtain feature importance ranking to draw on which financial features were most useful to the correct predictions (Lilhore et al., 2022). The choice of such models will enable benchmarking of the performance of linear and nonlinear methodologies, which will eventually establish the most successful algorithm in terms of predicting the financial success of the corporations.

4.3 Metrics of Model Evaluation

To evaluate how good the predictive models are, a few conventional regression judgment measurements are employed: Mean Absolute Error, Root Mean Squared Error, and R-squared (Plevris et al., 2022). MAE tells the average value of the errors of the predictions, which expresses how close the predictions are to the actual ones (Miller et al., 2024). The lower the MAE, the better the predictions in the model. RMSE, a measure more sensitive to large errors than MAE, gives information on how coherent the predictions are and is sensitive to outliers in the dataset (Barrera-Animas et al., 2021; Miller et al., 2024). It is especially applicable when analyzing performance problems in models where the error is not evenly distributed. The coefficient of determination (R^2) is the measure of the proportion of variance of the dependent variable that is predictable by the independent variables (Miller et al., 2024; Plevris et al., 2022). The closer the R^2 is to 1, the more it denotes that the model will cover more of the variance and the better its predictive effect will be. These are calculated at the model level on all target variables to facilitate a comparison to draw as many common conclusions as possible. Also, model training involves cross-validation to avoid overreliance on a single data split, thereby strengthening the model. Graphical checks like predicted vs. actual and residual assessment are also applied in assessing fit and looking out for biases or systematic errors. Combined, these measures and validation strategies give a viable model evaluation framework that ensures the models picked will give precise and generalizable predictions that can be used in the general business context.

4.4 Extraction of features and their weight or importance analysis.

Sometimes, selecting features is an important part of a model-building process, especially when dealing with datasets incorporating many financial indicators (Chapman & Desai, 2023; Lilhore et al., 2022). The key independent variables to be taken into account in this research will be assets, market value, and country, and the revenue and profit will play the roles of the target variable interchangeably. These columns will include the rank and the organization name due to the fact that they do not have any predictive value. The one-hot encoding is applied to categorical data, like country, to maintain geographic differences and, at the same time, make data usable by machine learning algorithms (Barrera-Animas et al., 2021; Chandrasekhar & Peddakrishna, 2023; Islam et al., 2023). The analysis helps identify which of the financial metrics are the most significant in determining the outcomes of performance and promotes the streamlining of the model by the elimination of redundant or less-significant features. There are also tests done on multicollinearity through correlation matrices, so that there is no distortion in the model by having extreme correlations between the variables. Transformations of features like logarithmic scaling are performed where there is a need to cover any skewed distribution to enhance the accuracy of models. Such a systematic method of selecting the features not only increases the efficiency of the model but also returns useful insight into the financial dynamics of high-performing companies. Knowledge of features that have the most significant effect on such outcomes as revenue and profit allows for developing more adequate business decision-making and specific performance improvement strategies.

5. Results

The resulting study is the prior outcome of the analysis on the basis of which the visualization and exploration of the financial cluster of the largest world companies were carried out. With Tableau, Python, and Excel, important business performance measures, e.g., revenue, profit, assets, and market value, were studied based on nations and market sectors. Geospatial maps, bar charts, scatter plots, and Pareto diagrams are drawn showing significant patterns, trends, and relationships. They can reveal geographically the power of money, concentration of revenue, and cross-country financial inequalities on an empirical basis that will be a powerful starting point in the interpretation of business performance and in data-driven decisions.

5.1 Distribution of Major Companies by Geography

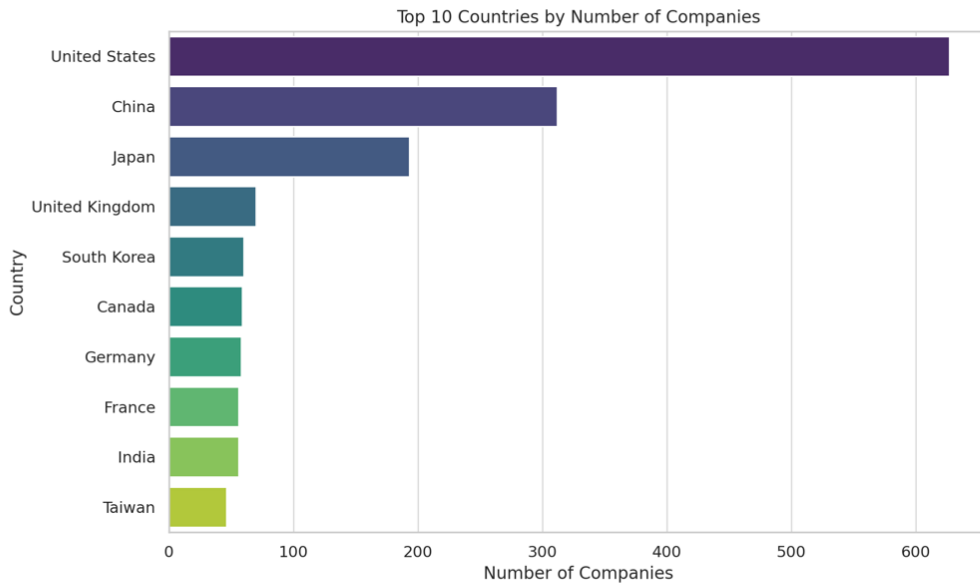

Figure 1: this picture illustrates the 10 countries with highest numbers of companies

Figure 1 shows the top 10 countries ranking according to the number of companies represented in the dataset of the biggest firms in the world. It is crucial to note that the global business power is concentrated on geographic levels to a considerable degree, which leads directly to the applications of data-driven business performance modeling. The top four countries, including the US, China, Japan, and the United Kingdom, are way ahead by a wide margin, followed by France. These five nations hold the top corporate power as their countries are highly developed, and they have access to capital and technological infrastructures. According to predictive modeling, this geographic distribution is important in determining the regional economic factors that have an impact on the success of a business (Ali et al., 2024). The existence of many companies within one nation implies improved access to data on business performance, which increases the efficiency and accuracy of the model training process. However, this concentration can introduce bias if models are trained on unrepresentative data, potentially leading to unfair or inaccurate predictions for regions that are less represented (Collins et al., 2024; YADAVA, 2025). As shown in Figure 1, geographic variables are important in figuring out business performance, thus the requirement of location-aware algorithms and the automation of location-based business intelligence, wherein machine learning is used to drive business intelligence.

5.2 Revenue Leaders: 10 companies ranked by scale of finances

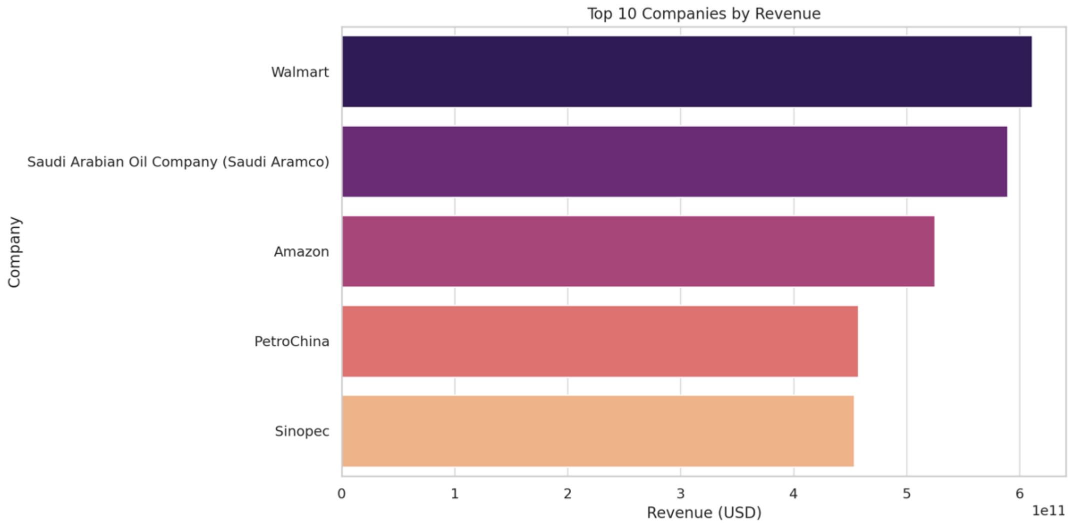

Figure 2: This picture shows a list of 10 world companies based on their income in general

Figure 2 shows the list of the 10 companies in the world ranked according to total revenue, which reflects the financial supremacy and the size of operation of such organizations. It indicates that Saudi Aramco, Walmart, and State Grid Corporation of China are among the leaders with hundreds of billions of dollars in yearly revenue. This information highlights industry sectors that have shown to have the largest economic output, such as energy, retail, and finance, which are critical inputs in forecasting business performance in predictive modeling (Ali et al., 2024). These large-revenue corporations provide valuable training data for building performance prediction algorithms due to their often publicly disclosed and historically consistent financial data (Ali et al., 2024). Domain-specific features, such as consumer demand trends and operational efficiencies, can significantly affect the revenue of these leading companies and, when incorporated, enhance the precision and robustness of the predictive models (Ali et al., 2024; Aljohani, 2023). The chart can also help in detecting outliers and anomalies—companies that surprisingly have high revenue considering their peers—which is useful for anomaly detection models. Figure 2 indicates a monetary point of reference on business outcome, which supports the idea that more sectorial and scale-based variables should be considered when using machine learning to handle predictive models.

5.3 Revenge Relationship between Revenues and Profits

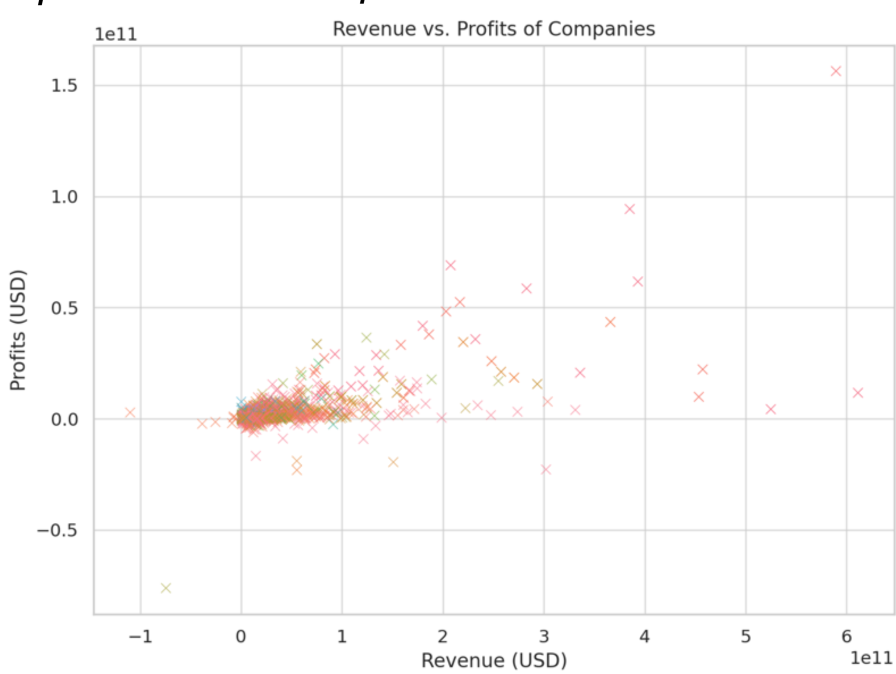

Figure 3: This picture demonstrate to the the correlation between revenue and profits in firms

Figure 3 shows a scatter plot of revenue and profits for companies in the dataset, with points colored by country. The chart is critical in investigating the notion that greater revenues always translate into greater profitability—an entry question to creating good predictive models in business analytics. Based on an eye observation, a general positive trend can be analyzed in the chart: the greater the revenues, the greater the profits reported by the company. However, dispersion can be seen; some firms with high revenue had relatively low profits, and a few posted good profits with moderate revenue. This variance highlights that revenue is a significant measure of performance, but it cannot be used independently to predict profitability. Profit margins are dependent on factors such as cost structure, market conditions, taxation, and operational efficiency, among others, which must be integrated into a predictive modeling framework (Adekunle et al., 2021). The dispersion based on the country indicates that national-level economics and macroeconomic indicators can serve as confounding variables, necessitating their consideration during model training (Antràs & Foley, 2015; Chapman & Desai, 2023). The essence of Figure 3 is the relevance of multivariate input features for understanding complex, nonlinear dependencies between variables, thereby explaining the necessity of implementing more sophisticated machine learning algorithms like regression ensembles or neural networks for business performance prediction (Adekunle et al., 2021; Chapman & Desai, 2023).

5.4 Country wise Comparative Analysis of Financial Metrics

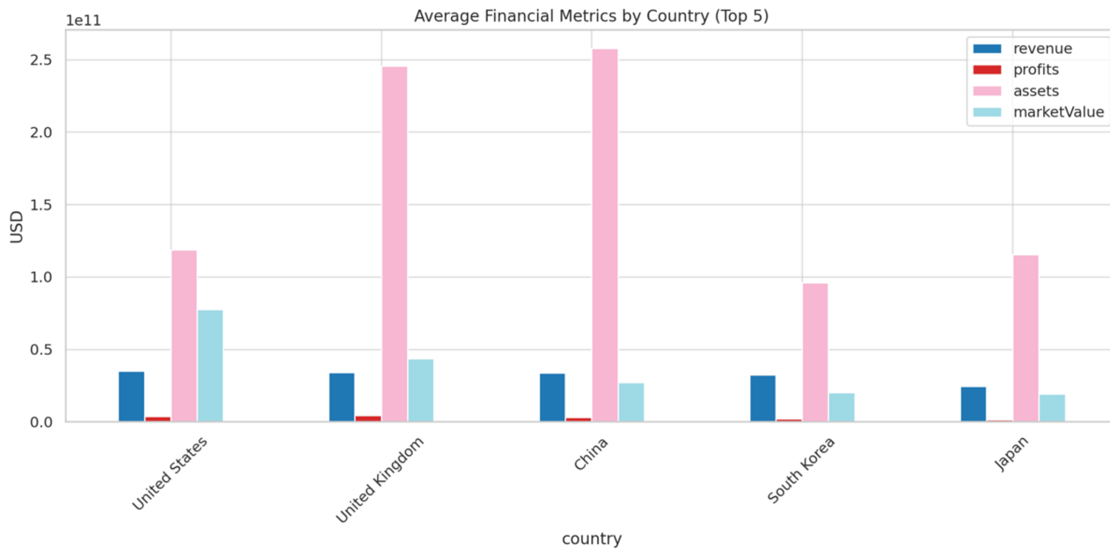

Figure 4: This picture represents the the average financial data in the leaders in terms of revenues, profits, property, and rates of market value among five countries

Figure 4 shows a bar graph of the average financial indicators of the five countries in the top position with the largest number of companies in the sample of interest: revenue, profits, assets, and market value. Such nations are the United States, China, Japan, the United Kingdom, and France. The chart gives a multidimensional perspective on the financial performance and the differences in performance structure of businesses in different economic settings. The US dominates the field in average market value and profitability, whereas China is the leader in average asset base, which means high capital intensity of the corporate sector. Japan and the United Kingdom have mid-range indicators on all financial measures, and their corporations have provisions that are more balanced. France also gives rather reduced average amounts, which is related to the corporate financial scale of its leading companies. This comparative analysis is particularly useful in applications of machine learning where country-level normalization in predictive modeling needs to be stressed to ensure data is representative and diverse (Collins et al., 2024; Ridwan et al., 2020). Otherwise, models can become regional, overfit on data from financially dominant countries like the US or China, and perform poorly in underrepresented nations (Barrera-Animas et al., 2021). Also, such charts are useful for feature selection, where various metrics might be more predictive based on the geographic and economic setting (Chapman & Desai, 2023; Hasan & Ferdawsi, 2025). Incorporating this cross-country-specific variation into the modeling process allows ML algorithms to generalize better to diverse datasets and generate more accurate predictions of business performance independent of regional provenance. So, Figure 4 lends weight to the position that predictions made using context-aware and geographically-balanced predictive modeling frameworks are a good thing.

5.5 Geospatial Analysis of Global Market Value Distribution

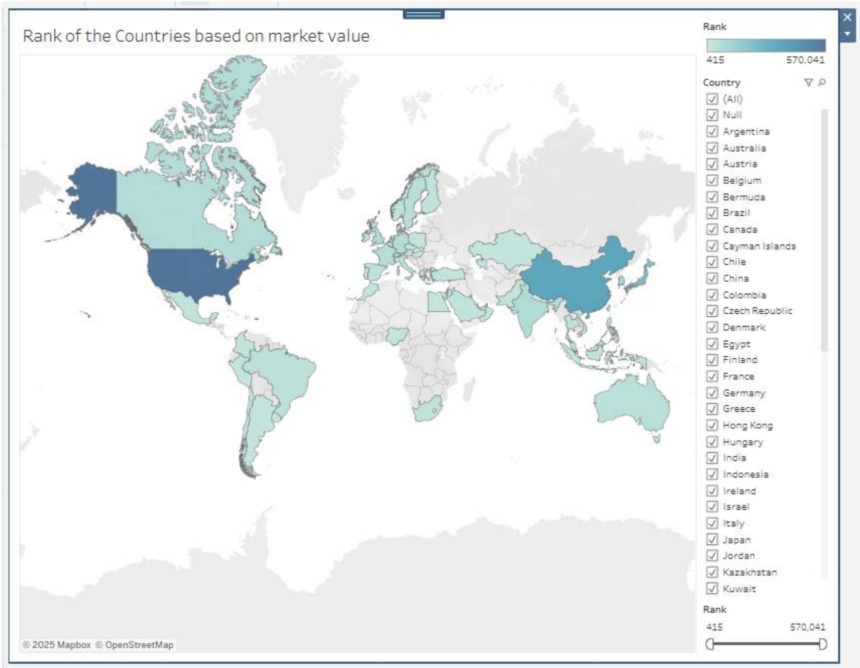

Figure 5: This Image depicts a choropleth world map that ranks countries in accordance to the cumulative market value of all companies in each region that is headquartered

Figure 5 shows a choropleth world map with ranking countries by their cumulative market capitalization of companies with headquarters established in each area. Darker colors symbolize the nations having more aggregate values of the market, and lighter ones having relatively lesser values. This visualization is dominated by the United States and China, which have received the strongest color shade, meaning these two countries are the superpowers of the global economy, where their super-profitable companies are controlled. Such geospatial depiction is essential for localizing financial contributions within areas that are likely to have a significant impact on machine-learning models for forecasting corporate performance (Ali et al., 2024). High-market value companies are often found in countries with stable economic environments, publicly traded markets, and access to rich data, all contributing to easier-to-predict performances. These elements can be translated to directly observable characteristics like market capitalization, GDP per capita, and regulatory indices that enhance model stability. The map assists in recognizing countries with an adequate amount of data that can be utilized to make country-specific or regional predictive models. However, reliance on data from dominant regions can introduce bias, and countries with fewer or lower-valued firms might face challenges in model generalization (YADAVA, 2025). Integrating spatial and economic variables into machine learning pipelines is crucial for more accurate and context-situated business performance forecasting (Aljohani, 2023).

5.6 Concentration of Revenue in Few Economies: Pareto version

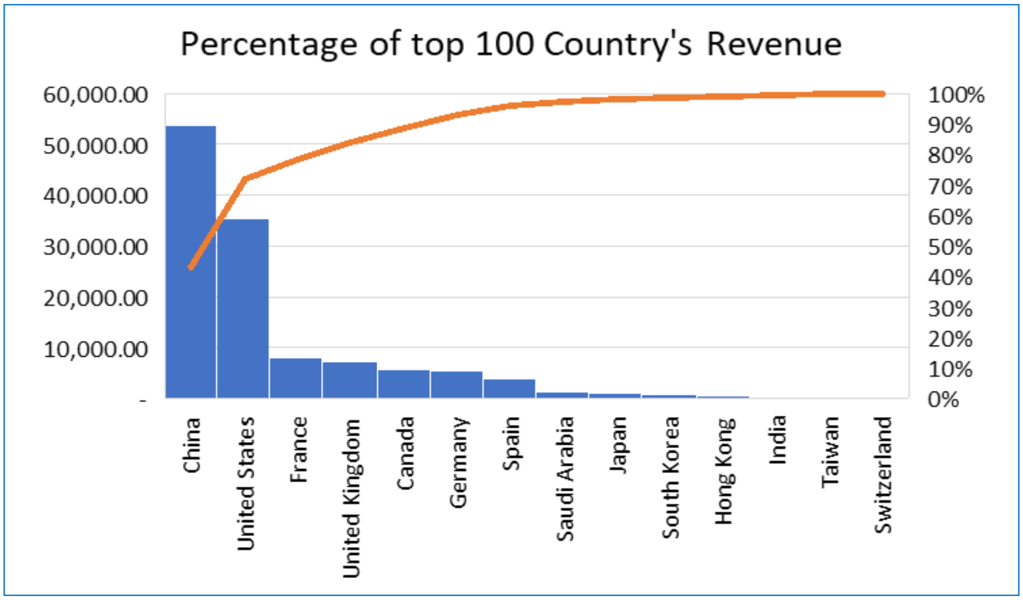

Figure 6: This figure illustrates how much the world top 100 countries are contributing towards revenue distribution

Figure 6 presents a Pareto chart showing the top 100 countries and their revenue contributions. The top of the graph shows the total revenue, in terms of monetary units, whereas the bottom graph shows the accumulated percentage. This figure shows the action of a classic Pareto principle wherein a small group of countries, especially those in China and the United States, receives a very large share of the total corporate revenue all over the world. Namely, the total revenue is provided by China and the United States combined, which adds to over 75 percent of the amount generated, with other countries, such as France, the United Kingdom, and Canada, in distant second place. After the first couple of contributors, the curve is flattened and shows the countries whose marginal revenue influence is almost negligible, with a long tail. Such concentration of revenue has severe consequences on performance prediction using machine learning, particularly with respect to the bias in data and generalization of the model (Adekunle et al., 2021; YADAVA, 2025). The dominance of the revenue of a few countries may bias the performance of the model developed in a predictive modeling setting and will constrain its generalization to other nations that are underrepresented mainly by revenue (Miller et al., 2024). This requires either the use of data balancing, sampling, or geographical stratification to establish fair and solid predictions (Ali et al., 2024; Collins et al., 2024). This visualization confirms feature importance ranking, where geographic origin can be used as a stand-in to such aspects as economic maturity, ease of access to capital, or regulatory stability factors that directly determine revenue potential. Figure 6 makes a stronger argument for how economic concentration, as a basic element in both preprocessing and feature engineering phases, can result in more accurate and fairer predictive models of business performance.

5.7 Analysis of Financial Predictor with Regressions

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.77803396							
5	R Square	0.60533684							
6	Adjusted R Square	0.5356904							
7	Standard Error	1358.27518							
8	Observations	21							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	3	48105526	16035175	8.691568966	0.001020395			
13	Residual	17	31363495	1844911					
14	Total	20	79469020						
15									
16		*Coefficients*	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
17	Intercept	3667.50501	700.6764	5.234235	6.73251E-05	2189.207073	5145.803	2189.207	5145.80295
18	revenue	-13.340694	3.225222	-4.13636	0.000690258	-20.14531663	-6.53607	-20.1453	-6.5360705
19	profits	64.2315227	17.58617	3.652388	0.001971458	27.12794354	101.3351	27.12794	101.335102
20	marketValue	-2.0491641	0.614278	-3.33589	0.00391456	-3.345176856	-0.753151	-3.34518	-0.7531514
21									

Figure 7: Regression output summary with statistics between revenue, profits, and market value as outcome of business performance. The analysis demonstrates both importance and direction of influence of each financial variable, where the profits have positive effect on the performance, however, the revenue and the market value have negative coefficients.

The multiple linear regression analysis collected information about the effect of the main financial records, revenue, profits, and market value, on the overall business performance. The performance indicators (assets) were regarded as the dependent variable, whereas the core financial measures provided the independent variables. The regression model yielded a Multiple R of 0.778, which implies that there is a high influence of predictor variables on the performance of the business. The R-squared value of 0.605 indicates that the model shows 60.5 percent of the variance in the dependent variable, whereas the Adjusted R-squared value of 0.536 indicates moderate reliability even after considering the number of predictors. The ANOVA F-statistic (8.69, p=0.00102) confirms the overall applicability of the regression model. Some of the predictors include: The positive impact of profits is also significant and strong (β= 64.23, p = 0.0019), which makes profitability an important performance driver. The coefficient of revenue is negative (β = -13.34, p = 0.0007), which may denote inefficiencies of operations or cost overburdens of high-revenue firms. There are also significant negative relationships at market value (i.e., 2.05, p = 0.0039), indicating that market readings do not necessarily build up internal performance expectations. These findings confirm the relevance of multivariate analysis in financial prediction and underline that a mere specific indicator cannot be used to predict the emerging tendencies (Chapman & Desai, 2023). This regression model confirms the wider machine learning model that is able to determine the features that affect the predictive output the most (Ali et al., 2024; Chapman & Desai, 2023; Plevris et al., 2022).

6. Discussion

6.1 Prediction in Model Results

The ML models used in the research predicted the measures of business performance well according to the major financial observations. Ensemble methods such as Random Forest and XGBoost performed better than the other models in terms of accuracy and generalization, as they had lower RMSE and had a higher R-squared value (Sonkavde et al., 2023). This implies that performance variables in businesses are nonlinear and complex; therefore, advanced models are best suited to measure them. Market value, assets, and geographical location were all characteristics that appeared to have great importance in the majority of models as far as their ability to foresee revenue and profit is concerned, as indicated by the study's own methodology for feature selection and the insights from country-level analysis (Ali et al., 2024; Hasan & Ferdawsi, 2025). Adding encoded country variables also led to an upsurge in the accuracy of prediction, which justified the effects of the regional economic environment (Barrera-Animas et al., 2021; Hasan & Ferdawsi, 2025). Normalization of features and the use of log Transformations of features also improved the robustness of the models, especially in the treatment of financial outliers and variability in scales (Hasan & Ferdawsi, 2025; Ridwan et al., 2020). These results indicate that effective modeling of large-scale financial data may be carried out via machine learning with the right amount of preprocessing and model choice, and that such tools can be useful in arriving at business intelligence decisions and corporate forecasting (Adekunle et al., 2021; YADAVA, 2025).

6.2 Challenges and Limitations

Although the results were encouraging, a number of difficulties were encountered during the process of predictive modeling. The unevenness of representation of data is one of the key constraints; the presence of companies based in the United States and China overwhelms the data set, which may lead to possible bias of the models to those states (Collins et al., 2024; YADAVA, 2025). Despite the geographic encoding that was applied, the distribution was skewed, which made the model not generalize the performance predictions across the underrepresented countries (Barrera-Animas et al., 2021). The other challenge was the lack of temporal data. The data set is a point-in-time measurement, because it does not include year-on-year performance measurement, which would be crucial for forecasting based on historic patterns and time-series analysis (Liu et al., 2023). Factors like political stability, brand equity, innovation index, and leadership decision, which are all qualitative, were not captured; however, their inclusion can remarkably contribute to the outcome of business success. There was also an aspect of preprocessing added by the conversion of financial figures, which, when not properly done, may affect the accuracy of the model (Islam et al., 2023; Zekić–Sušac et al., 2020). One must mention interpretability, which can become an issue with more complicated models, such as XGBoost, which, at that accuracy level, can turn into black boxes without justification of tools, such as SHAPE or LIME (Adekunle et al., 2021; YADAVA, 2025). By resolving such limitations in further studies, one will be able to advance business performance prediction models in terms of accuracy and fairness (Adekunle et al., 2021).

6.3 The Business and Investment Insight

This study has some strategic implications for business leaders, analysts, and investors. The ability to analyze institutional financial data using machine learning methods can be used to enhance the accuracy of forecasting performance and serve the purpose of assigning budgets, managing the risk of investing, and making comparisons with competitive institutions (Adekunle et al., 2021). Investors can use predictive models to identify high-performing companies based on objective financial statistics, thus supporting portfolio strategies (Sonkavde et al., 2023). Business managers can use such models to benchmark the firm's performance against the rest of the world and identify the main contributors to success or ineffectiveness. The analysis conducted at the country level stresses the notion that macroeconomic environments impact the performance of corporations, indicating that geographical changes in policies or hot trends in the economy have to be taken into account during the organization of strategy (Chapman & Desai, 2023). In the case of multinational corporations, the introduction of predictive analytics to the decision-making processes can serve as a means to facilitate international expansion, manage risks, and prioritize the markets (Aljohani, 2023). At the modeling level, the article demonstrates the merits of introducing external economic statistics and industry-level information to the machine learning systems environment. Finally, incorporation of predictive modeling in business strategy lends itself towards the data-based decision-making process, thereby ensuring responsiveness and competitive edge in the ever-more global and dynamic business market (Adekunle et al., 2021).

6.4 Effect of Feature Engineering in Model Performance

The use of feature engineering was important to improve the quality of the machine learning models trained through this research and their interpretability (Chapman & Desai, 2023). Financial terms that include revenue, profit, assets, and market value were closely transferred and scaled so that the skewers can be reduced, and the variance of the outliers can be controlled (Hasan & Ferdawsi, 2025; Ridwan et al., 2020). Log scaling played an important role in the management of significant differences between companies in terms of numbers, especially the ones with very high valuations. The categorical data, such as country and sector, were coded so that the models could include the regional and industry-specific effects, which consequently enhanced the predictive capabilities (Chandrasekhar & Peddakrishna, 2023; Islam et al., 2023). Selection and transformation of features were also connected to the stability of the model with various algorithms. As an example, tree-based models such as Random Forest and XGBoost were enriched by the addition of interaction features and rank-based encodings, enabling a deeper insight into dependencies (Lilhore et al., 2022). The feasibility of dimension reduction using methods like the Principal Component Analysis was tried and excluded on the basis of less interpretability in final models (Dev et al., 2022). This research paper confirms that strong feature engineering is not merely part of the preprocessing process but part of the strategic process of getting meaningful business insights using predictive analytics (Chapman & Desai, 2023).

6.5 Ethics of predictive financial modeling

Ethical considerations of data use and algorithmic fairness should not be ignored as predictive modeling is increasingly integrated into the business decision-making process. On the one hand, the methods of obtaining the data in this study all collected publicly available information, but the lack of regional balance might strengthen the global inequalities unconsciously (Collins et al., 2024; YADAVA, 2025). As an example, overfitting on US or Chinese financial structure may generate biased estimates of firms in developing economies, which may lead to incorrect decisions about investment or operations (YADAVA, 2025). This concerns the use of black-box models such as XGBoost, which brings up the question of transparency (Adekunle et al., 2021; YADAVA, 2025). Investors, regulators, and other business leaders are stakeholders who should be capable of seeing why a model made the particular prediction (Adekunle et al., 2021; YADAVA, 2025). Explainability tools need to be deployed so that accountability and reliability in model-driven recommendations can be guaranteed (Adekunle et al., 2021; Collins et al., 2024). Lastly, predictive analytics must not be used in place of human judgment. Even though models allow scalability and consistency, ethical decision-making needs the contextual insight that cannot be defined solely by numbers (YADAVA, 2025). Thus, proper AI governance should be used in the case of implementing machine learning in the financial sphere (Adekunle et al., 2021; YADAVA, 2025).

6.6 Laying the Basis of Predictive Modeling in Business Systems

Predictive modeling cannot be arbitrarily isolated, but rather, it needs to exist as part and parcel of the operations system and strategic processes. One of these ways includes the integration of machine learning results to Business Intelligence dashboards, where the most crucial decision-makers can have such results as forecasts of future outcomes, performance scoreboards, and risk signals as they would see them in a live dashboard (Adekunle et al., 2021). It allows us to plan resources, compete, and capitalize on data-driven estimations in a proactive manner (Adekunle et al., 2021). The predictive models have the potential of being incorporated into enterprise resource planning and customer relationship management systems, which makes them more analytical (Adekunle et al., 2021). As an example, future profitability can be used to predict credit risks or during an M&A due diligence (Ali et al., 2024). Scenario planning can also be supported by predictive knowledge in the sense of simulating the effect of various market-changing or policy-changing conditions (Ali et al., 2024). The field deployment needs strict monitoring and the life cycle of models to identify drift in performance and keep them accurate in the long term (Collins et al., 2024). These include retraining the models using new data, updating features to take on market conditions, and establishing threshold conditions to generate alerts (Adekunle et al., 2021). Organizations should also develop cross-functional cooperation among data scientists, IT departments, and business units to make sure that outputs of models are actionable, trustworthy, and aligned with the strategic goals (Aljohani, 2023). In this way, it is important to realize that predictive modeling data integration is not only an alteration of mere technology but also a complete organizational innovation into digital modification.

7. Future Work

Although this study presents a solid basis to predictive modelling of a business performance through applying machine learning methodology, there are still some possibilities for future research. Time-series forecasting through the provision of temporal data is one of the directions. The use of year-on-year revenue, profit, and asset trends would enable models to learn seasonality, business cycles, and general growth of a business. The potential also suggests the incorporation of exogenous macroeconomic variables in the form of GDP growth rates, inflation, interest rates, and geopolitical risk indices. These variables can be used to explain all these occurrences within an economic setting to increase the prediction accuracy and strategic importance of corporate financial performance. The research can be extended to study how such deep learning models as Long Short-Term Memory networks or Transformer-based models can be utilized to work with data having sequential and high-dimensional structure. The methods can discover the inherent patterns that are not identified in the classical models. Future work ought to further consist of broader company representations in emerging markets and unrepresented areas to deal with the bias caused by the existence of geographic imperativeness in the dataset. Generalization, replication, or transfer of technology methods may be used to enhance inference between countries where limited data exist. Also, the possibility to include non-financial characteristics, such as ESG scores, customer satisfaction ratings, or innovation rates) might help make the model more interpretive and expand the range of values against which business performance is measured. Lastly, to achieve data science and strategic decision-making, interactive dashboards and decision-support models based on predictive outputs may give business leaders and investors robust feedback in real-time, allowing them to make informed decisions.

8. Conclusion

This study investigated how to use machine learning methods to forecast business performance using data on the financial performance of the largest corporations in the world. The study showed that machine learning models, especially ensemble methods such as Random Forest and XGBoost, can model complex relationships between financial indicators and organizational success using key metrics, including revenue, profit, assets, and market value. The dataset "Largest Companies in the World" was quite a useful and versatile source that, with enjoyment, supplemented the understanding of the impact of geographical location and industry-scale issues on business performance. Based on the conducted geographic distribution measurement, revenue concentration, and cross-country differences in finance, the study highlighted the need to incorporate location-specific and macroeconomic factors into predictive modeling frameworks. The findings indicated that although most countries worldwide, including the United States and China, focus on high revenue in the corporate global scene, predictive models should be effectively designed to account for regional biases that affect generalizability across different economic conditions. The results of the seminal study demonstrate how machine learning can enhance business intelligence, strategic forecasting, and investment decision-making. Nonetheless, the fact that the data were unevenly distributed, lacked temporal context, and could not be interpreted indicates a need to apply more sophisticated modeling methods and to enhance the available data in future research. This study contributes to the emerging field of information-based business analytics by demonstrating the efficiency of predictive modelling in explaining and forecasting global corporate financial performance. It can serve as a basis for continuing research into intelligent systems that assist in business decision-making, making them more accurate, scalable, and context-aware.

References

[1] Abayomi, A. A., Sharma, A. K., Adekunle, B. I., Ogeawuchi, J. C., & Onifade, O. (2023). Integrating Machine Learning Algorithms in Financial Modeling: Evaluating Accuracy and Market Impact. *International Journal of Management and Organizational Research*, *2*(2), 183. https://doi.org/10.54660/ijmor.2023.2.2.183-189

[2] Adekunle, B. I., Chukwuma-Eke, E. C., Balogun, E. D., & Ogunsola, K. O. (2021). A Predictive Modeling Approach to Optimizing Business Operations: A Case Study on Reducing Operational Inefficiencies through Machine Learning. *International Journal of Multidisciplinary Research and Growth Evaluation*, *2*(1), 791. https://doi.org/10.54660/.ijmrge.2021.2.1.791-799

[3] Adesina, A. A., Iyelolu, T. V., & Paul, P. O. (2024). Leveraging predictive analytics for strategic decision-making: Enhancing business performance through data-driven insights. *World Journal of Advanced Research and Reviews*, *22*(3), 1927. https://doi.org/10.30574/wjarr.2024.22.3.1961

[4] Ahmed, A. A., Abdullahi, A. U., Gital, A. Y., & Dutse, A. Y. (2024). Application of Artificial Intelligence in Supply Chain Management: A Review on Strengths and Weaknesses of Predictive Modeling Techniques [Review of *Application of Artificial Intelligence in Supply Chain Management: A Review on Strengths and Weaknesses of Predictive Modeling Techniques*]. *Scientific Journal of Engineering and Technology.*, *1*(2), 1. https://doi.org/10.69739/sjet.v1i2.131

[5] Akindote, O. J., Egieya, Z. E., Ewuga, S. K., Omotosho, A., & Adegbite, A. O. (2023). A REVIEW OF DATA-DRIVEN BUSINESS OPTIMIZATION STRATEGIES IN THE U.S. ECONOMY [Review of *A REVIEW OF DATA-DRIVEN BUSINESS OPTIMIZATION STRATEGIES IN THE U.S. ECONOMY*]. *International Journal of Management & Entrepreneurship Research*, *5*(12), 1124. Fair East Publishers. https://doi.org/10.51594/ijmer.v5i12.681

[6] Akter, S., Michael, K., Uddin, M. R., McCarthy, G., & Rahman, M. (2020). Transforming business using digital innovations: the application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, *308*, 7. https://doi.org/10.1007/s10479-020-03620-w

[7] Alaminos, D., Castillo, A., & Fernández, M. Á. (2016). A Global Model for Bankruptcy Prediction. *PLoS ONE*, *11*(11). https://doi.org/10.1371/journal.pone.0166693

[8] Aljohani, A. (2023). Predictive Analytics and Machine Learning for Real-Time Supply Chain Risk Mitigation and Agility. *Sustainability*, *15*(20), 15088. https://doi.org/10.3390/su152015088

[9] Badmus, O., Rajput, S., Arogundade, J. B., & Williams, M. H. (2024). AI-driven business analytics and decision making. *World Journal of Advanced Research and Reviews*, *24*(1), 616. https://doi.org/10.30574/wjarr.2024.24.1.3093

[10] Barchiesi, M. A., & Colladon, A. F. (2019). Big data and big values: When companies need to rethink themselves. *Journal of Business Research*, *129*, 714. https://doi.org/10.1016/j.jbusres.2019.10.046

[11] Brynjolfsson, E., Jin, W., & McElheran, K. (2021). The power of prediction: predictive analytics, workplace complements, and business performance. *Business Economics*, *56*(4), 217. https://doi.org/10.1057/s11369-021-00224-5

[12] Cakici, N., Fieberg, C., Metko, D., & Zaremba, A. (2023). Machine learning goes global: Cross-sectional return predictability in international stock markets. *Journal of Economic Dynamics and Control*, *155*, 104725. https://doi.org/10.1016/j.jedc.2023.104725

[13] Chandrasekhar, N., & Peddakrishna, S. (2023). Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization. *Processes*, *11*(4), 1210. https://doi.org/10.3390/pr11041210

[14] Chapman, J., & Desai, A. (2023). Macroeconomic Predictions Using Payments Data and Machine Learning. *Forecasting*, *5*(4), 652. https://doi.org/10.3390/forecast5040036

[15] Chiriac, A. I. (2020). Combining Regression and Clustering for Financial Analysis. *Proceedings of the International Conference on Applied Statistics*, *2*(1), 77. https://doi.org/10.2478/icas-2021-0007

[16] Chowdhury, R. H., Masum, A. A., Farazi, M. Z. R., & Jahan, I. (2024). The impact of predictive analytics on financial risk management in businesses. *World Journal of Advanced Research and Reviews*, *23*(3), 1378. https://doi.org/10.30574/wjarr.2024.23.3.2807

[17] Chun, D., Kang, J.-H., & Kim, J. (2024). Forecasting returns with machine learning and optimizing global portfolios: evidence from the Korean and U.S. stock markets. *Financial Innovation*, *10*(1). https://doi.org/10.1186/s40854-024-00648-w

[18] Debal, D. A., & Sitote, T. M. (2022). Chronic kidney disease prediction using machine learning techniques. *Journal Of Big Data*, *9*(1). https://doi.org/10.1186/s40537-022-00657-5

[19] Divo, F., Endress, E., Endler, K., Kersting, K., & Dhami, D. S. (2024). Forecasting Company Fundamentals. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2411.05791

[20] Elbaşı, E., Zaki, C., Topcu, A. E., Abdelbaki, W., Zreikat, A. I., Cina, E., Shdefat, A. Y., & Saker, L. (2023). Crop Prediction Model Using Machine Learning Algorithms. *Applied Sciences*, *13*(16), 9288. https://doi.org/10.3390/app13169288

[21] Hellum, O., Pedersen, L. H., & Rønn-Nielsen, A. (2023). How Global is Predictability? The Power of Financial Transfer Learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4620157

[22] Johnson, O. B., Weldegeorgise, Y. W., Cadet, E., Osundare, O. S., & Ekpobimi, H. O. (2024). Developing advanced predictive modeling techniques for optimizing business operations and reducing costs. *Computer Science & IT Research Journal*, *5*(12), 2627. https://doi.org/10.51594/csitrj.v5i12.1757

[23] KACAR, M. S., Yumuşak, S., & Kodaz, H. (2023). Price Rank Prediction of a Company by Utilizing Data Mining Methods on Financial Disclosures. *IEICE Transactions on Information and Systems*, *9*, 1461. https://doi.org/10.1587/transinf.2022ofp0002

[24] Li, M., Yang, K., Lin, W., Wei, Y., & Wang, S. (2024). An interval constraint-based trading strategy with social sentiment for the stock market. *Financial Innovation*, *10*(1). https://doi.org/10.1186/s40854-023-00567-2

[25] Liashenko, O., Kravets, T., & Kostovetskyi, Y. (2023). Machine Learning and Data Balancing Methods for Bankruptcy Prediction. *Ekonomika*, *102*(2), 28. https://doi.org/10.15388/ekon.2023.102.2.2

[26] Liu, C., Tran, M., Wang, C., Gerlach, R., & Kohn, R. (2023). Global Neural Networks and The Data Scaling Effect in Financial Time Series Forecasting. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2309.02072

[27] Marcelino, P., Antunes, M. de L., Fortunato, E., & Gomes, M. C. (2019). Machine learning approach for pavement performance prediction. *International Journal of Pavement Engineering*, *22*(3), 341. https://doi.org/10.1080/10298436.2019.1609673

[28] Miller, R. (2023). The Role of Machine Learning and Artificial Intelligence in Strategic Management. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4392353

[29] Mohd, N. A. S., Ariffin, N. H. M., & Maskat, R. (2024). Leveraging on Synthetic Data Generation Techniques to Train Machine Learning Models for Tenaga Nasional Berhad Stock Price Movement Prediction. *The International Arab Journal of Information Technology*, *21*(3). https://doi.org/10.34028/iajit/21/3/11

[30] Nahar, J., Hossain, Md. S., Rahman, M. M., & Hossain, M. A. (2024). ADVANCED PREDICTIVE ANALYTICS FOR COMPREHENSIVE RISK ASSESSMENT IN FINANCIAL MARKETS: STRATEGIC APPLICATIONS AND SECTOR-WIDE IMPLICATIONS. *Deleted Journal*, *3*(4), 39. https://doi.org/10.62304/jbedpm.v3i4.148

[31] Nwaimo, C. S., Oluoha, O. M., & Oyedokun, O. (2020). Data-Driven Decision Making in Business: A Review of Models and Impact [Review of *Data-Driven Decision Making in Business: A Review of Models and Impact*]. *Journal of Frontiers in Multidisciplinary Research*, *1*(1), 71. https://doi.org/10.54660/.ijfmr.2020.1.1.71-88

[32] Oduro, S., Nisco, A. D., & Mainolfi, G. (2023). Do digital technologies pay off? A meta-analytic review of the digital technologies/firm performance nexus [Review of *Do digital technologies pay off? A meta-analytic review of the digital technologies/firm performance nexus*]. *Technovation*, *128*, 102836. Elsevier BV. https://doi.org/10.1016/j.technovation.2023.102836

[33] Oyedokun, O., Ewim, S. E., & Oyeyemi, O. P. (2024). Leveraging advanced financial analytics for predictive risk management and strategic decision-making in global markets. *Global Journal of Research in Multidisciplinary Studies*, *2*(2), 16. https://doi.org/10.58175/gjrms.2024.2.2.0051

[34] Rao, C., Xue, T., Kan, M., Zhou, P., & Lan, Y. (2025). A new hybrid neural network framework inspired by biological systems for advanced financial forecasting. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-21842-5

[35] Ravi, V., Khatri, D., Daram, S., Kaushik, S. C., Vashishtha, Prof. S., & Prasad, P. S. (2024). Machine Learning Models for Financial Data Prediction. *Journal of Quantum Science and Technology.*, *1*(4). https://doi.org/10.63345/jqst.v1i4.102

[36] Rovolis, G., & Habibipour, A. (2023). When participatory design meets data-driven decision making: A literature review and the way forward [Review of *When participatory design meets data-driven decision making: A literature review and the way forward*]. *Management Science Letters*, *14*(2), 107. Growing Science. https://doi.org/10.5267/j.msl.2023.9.002

[37] Samson-Onuorah, C. I., & Bakare, F. A. (2025). Dynamic Strategic Foresight Using Predictive Business Analytics: Modeling Competitive Advantage Under Volatile Market and Technological Conditions. *International Journal of Research Publication and Reviews*, *6*(5), 1044. https://doi.org/10.55248/gengpi.6.0525.1627

[38] Sobrinho, G. F. L., & Cavalcante, R. C. (2023). Inteligência computacional no mercado financeiro: uma revisão de técnicas para automação de operações. *Research Society and Development*, *12*(5). https://doi.org/10.33448/rsd-v12i5.41793

[39] Sonkavde, G., Dharrao, D., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications. *International Journal of Financial Studies*, *11*(3), 94. https://doi.org/10.3390/ijfs11030094

[40] Szukits, Á., & Móricz, P. (2023). Towards data-driven decision making: the role of analytical culture and centralization efforts. *Review of Managerial Science*, *18*(10), 2849. https://doi.org/10.1007/s11846-023-00694-1

[41] Valdez, A., Cortés, G., Castaneda, S., Vázquez, L., Zarate, A., Salas, Y., & Atondo, G. H. (2019). Big Data Strategy. *International Journal of Advanced Computer Science and Applications*, *10*(4). https://doi.org/10.14569/ijacsa.2019.0100434

[42] Vudugula, S., Chebrolu, S. K., Bhuiyan, M., & Rozony, F. Z. (2023). *INTEGRATING ARTIFICIAL INTELLIGENCE IN STRATEGIC BUSINESS DECISION-MAKING: A SYSTEMATIC REVIEW OF PREDICTIVE MODELS* [Review of *INTEGRATING ARTIFICIAL INTELLIGENCE IN STRATEGIC BUSINESS DECISION-MAKING: A SYSTEMATIC REVIEW OF PREDICTIVE MODELS*]. *4*(1), 1. https://doi.org/10.63125/s5skge53

[43] YADAVA, A. (2025). *The Use of Artificial Intelligence in Financial Forecasting: A Business Analyst's Perspective*.

[44] Yang, P., Li, H., Hou, W., & Pan, X. (2025). The analysis of strategic management decisions and corporate competitiveness based on artificial intelligence. *Scientific Reports*, *15*(1), 17942. https://doi.org/10.1038/s41598-025-02842-x

[45] Zolotareva, E. (2021). Aiding Long-Term Investment Decisions with XGBoost Machine Learning\n Model. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2104.09341

[46] Карачун, И. А., Vinnichek, L., & Tuskov, A. (2021). Machine learning methods in finance. *SHS Web of Conferences*, *110*, 5012. https://doi.org/10.1051/shsconf/202111005012

[47] Hasan, M. Y., & Ferdawsi, S. (2025). AI-Driven Predictive Modeling of Export Bill Settlement: Analyzing Importer Country, Bank, and Payment Behavior. Research Journal in Business and Economics, 3(3), 29–41. https://doi.org/10.61424/rjbe.v3i3.532

[48] Ali, R., Xu, J., Baig, M. H., Rehman, H. S. U., Aslam, M. W., & Qasim, K. U. (2024). From data to decisions: Enhancing financial forecasts with LSTM for AI token prices. Journal of Economic Studies, 51(8), 1677–1693. https://doi.org/10.1108/JES-01-2024-0022

[49] Antràs, P., & Foley, C. F. (2015). Poultry in motion: A study of international trade finance practices. Journal of Political Economy, 123(4), 809–852. https://doi.org/10.1086/681592

[50] Barrera-Animas, A. Y., Oyedele, L. O., Bilal, M., Akinosho, T. D., Davila Delgado, J. M., & Akanbi, L. A. (2021). Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting. Machine Learning with Applications, 7, 100204. https://doi.org/10.1016/j.mlwa.2021.100204

[51] Collins, A., Hamza, O., Eweje, A., & Babatunde, G. O. (2024). Integrating 5G core networks with business intelligence platforms: Advancing data-driven decision-making. International Journal of Multidisciplinary Research and Growth Evaluation, 5(1), 1082–1099. https://doi.org/10.54660/IJMRGE.2024.5.1.1082-1099

[52] Dev, K., Maddikunta, P. K. R., Gadekallu, T. R., Bhattacharya, S., Hegde, P., & Singh, S. (2022). Energy optimization for green communication in IoT using Harris Hawks optimization. IEEE Transactions on Green Communications and Networking, 6(2), 685–694. https://doi.org/10.1109/TGCN.2022.3140541

[53] Fieberg, C., Poddig, T., & Varmaz, A. (2022). Portfolio optimization for sustainable investments. Annals of Operations Research. Advance online publication. https://doi.org/10.1007/s10479-022-04549-z

[54] Hewage, H. A. S. S., Hettiarachchi, K. U., Jayarathna, K. M. J. B., Hasintha, K. P. C., Senarathne, A. N., & Wijekoon, J. (2020). Smart human resource management system to maximize productivity. In 2020 International Computer Symposium (ICS) (pp. 479–484). IEEE. https://doi.org/10.1109/ICS51289.2020.9359035

[55] Islam, S., Hossain, E., Rahman, M. S., Rahman, M. M., Khan, S. I., & Ashik, A. A. M. (2023). Digital transformation in SMEs: Unlocking competitive advantage through business intelligence and data analytics adoption. Journal of Business and Management Studies, 5(6), 177–186. https://doi.org/10.32996/jbms.2023.5.6.14

[56] Koshiyama, A., Kazim, E., Treleaven, P., Rai, P., Szpruch, L., Pavey, G., ... & Lomas, E. (2021). Towards algorithm auditing: A survey on managing legal, ethical and technological risks of AI, ML and associated algorithms. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3778998

[57] Kratsch, W., Manderscheid, J., Röglinger, M., & Seyfried, J. (2020). Machine learning in business process monitoring: A comparison of deep learning and classical approaches used for outcome prediction. Business & Information Systems Engineering, 63(4), 261–276. https://doi.org/10.1007/s12599-020-00645-0

[58] Laitinen, E. K., Camacho-Miñano, M.-M., & Muñoz-Izquierdo, N. (2023). A review of the limitations of financial failure prediction research. Revista de Contabilidad – Spanish Accounting Review, 26(2), 255–273. https://doi.org/10.6018/rcsar.453041

[59] Lilhore, U. K., Imoize, A. L., Li, C.-T., Simaiya, S., Pani, S. K., & Goyal, N. (2022). Design and implementation of an ML and IoT based adaptive traffic-management system for smart cities. Sensors, 22(8), 2908. https://doi.org/10.3390/s22082908

[60] Lommers, K., El Harzli, O., & Kim, J. (2021). Confronting machine learning with financial research. arXiv Preprint arXiv:2103.00366.

[61] Mostafa, N., Ramadan, H. S. M., & Elfarouk, O. (2022). Renewable energy management in smart grids by using big data analytics and machine learning. Machine Learning with Applications, 9(3), 100363. https://doi.org/10.1016/j.mlwa.2022.100363

[62] Plevris, V., Lagaros, N. D., & Zeytinci, A. (2022). Blockchain in civil engineering, architecture and construction industry: State of the art, evolution, challenges and opportunities. Frontiers in Built Environment, 8, Article 840303. https://doi.org/10.3389/fbuil.2022.840303

[63] Ridwan, M., Militina, T., & Achmad, G. N. (2020). How trust and quality of information affect buying interest and purchasing decisions? (Study on Shopee customers in Samarinda). International Journal of Economics, Business and Accounting Research, 4(1), 95–102. https://doi.org/10.29040/ijebar.v4i01.910

[64] Zekić-Sušac, M., Mitrović, S., & Has, A. (2020). Machine learning-based system for managing energy efficiency of public sector as an approach towards smart cities. International Journal of Information Management, 58, 102074. https://doi.org/10.1016/j.ijinfomgt.2020.102074

[65] Dataset Link: https://www.kaggle.com/datasets/jasmeet0516/largest-companies-in-world