INTERNATIONAL DEVELOPMENT PLANNING REVIEW
ISSN:1474-6743 | E-ISSN:1478-3401

PREDICTABILITY OF TOTAL REVENUE IN GLOBAL LARGE COMPANIES FROM FINANCIAL AND ECONOMIC PARAMETERS

Vasu Padmanabhan
Research Scholar, Department of Management, PSGR Krishnammal College for Women, Bharathiar University, Coimbatore, India

Dr. S. Nirmala
Retired College Principal and Research Supervisor PSGR Krishnammal College for Women, Bharathiar University, Coimbatore, India

Abstract

The interplay between corporate earnings and market returns has been a pivotal area of study in management research for over half a century. Central to this discourse is the exploration of various financial performance metrics, among which revenue growth, stock price, and market capitalization have recently gained prominence in terms of predictability. This paper delves into the predictability of financial performance of large global corporations, emphasizing the increasing importance of revenue as one of the primary financial parameters. The study proposes that while revenue is fundamentally influenced by the company's product or service portfolio and market reception, other factors such as past revenue performance, financial statements line items and broader economic indicators (like GDP and inflation) also play a crucial role in forecasting revenue trends.

Our research encompasses a comprehensive analysis of approximately 2000 global companies over a 12-year period (2010-2022). The findings reveal that revenues can be predicted with a Mean Absolute Percentage Error (MAPE) accuracy of 15%. To achieve this, we employed a set of regression models incorporating Ordinary Least Squares (OLS), Artificial Neural Networks (ANN), and Bayesian Neural Networks (BNN). These methodologies were specifically chosen to elucidate both non-linear relationships and to provide 99% confidence intervals for each revenue prediction. The models when applied for 10 sector-capital market groupings, showed that, the predictability of revenue for U.S. companies was markedly higher, with their accounting information proving especially pertinent in predictive models. The study also covers the impact of covid pandemic and its recovery on the revenue predictability.

Keywords: Revenue, ANN, BNN, GDP, MAPE, Earning, Ordinary Least Squares

Introduction

The nexus between a company's financial performance and its market returns has been a subject of extensive research in the field of management, accounting and economics. Historically, the primary focus of this research has been on understanding how a company's earnings and profitability impact its overall market performance. However, in recent years, there has been a paradigm shift, emphasizing other financial parameters additionally, such as revenue growth, stock prices, and market capitalization. These elements have increasingly been recognized for their predictive value in assessing a company's financial health and future performance.

In the landscape of global business, revenue generation stands as a cornerstone of corporate success. Large global organizations, in particular, emphasize the revenue as a key financial indicator, both on a quarterly and yearly basis. This emphasis on revenue is not without reason; it provides a direct reflection of a company's market reach and operational efficiency. However, revenue figures do not exist in isolation. They are the culmination of various factors, including but not limited to, the company's product or service portfolio, market reception, and the effectiveness of strategic planning and execution.

Furthermore, revenue predictability extends beyond internal company dynamics. External factors such as geopolitical events, economic conditions, and industry trends also exert significant influence. For instance, variations in Gross Domestic Product (GDP) and inflation rates can have far-reaching impacts on consumer behaviour and, consequently, on revenue streams. Additionally, the company's past performance, along with its financial position as indicated by assets, inventory levels, and other key metrics, contributes to shaping future revenue trajectories.

Given the complex and multifaceted nature of revenue predictability, this research aims to provide a comprehensive analysis of revenue trends across 2000 global companies over a 12-year period (2010-2022). Employing a diverse array of statistical methods, including Ordinary Least Squares (OLS), Artificial Neural Networks (ANN), and Bayesian Neural Networks (BNN), this study seeks to unravel the intricate patterns and relationships that define revenue predictability. The integration of these advanced analytical techniques allows for the establishment of both linear and non-linear relationships, offering deeper understanding of the factors that drive company revenues. Moreover, the inclusion of 99% confidence intervals in our predictions adds a layer of rigor and reliability, making this study a significant contribution to the existing body of knowledge on corporate financial performance.

Literature Survey

In the context of financial parameters predictability, earnings and returns stand out as most researched outcomes. Earnings forecasts have been the primary focus for financial analysts, due to their significant impact on stock prices and investor sentiment. Earnings represent a direct assessment of a company's profitability, an essential factor for investor considerations. However, revenue forecasts are gaining importance as they offer insights into a company's sales performance, market demand, and growth potential. Revenue is especially critical for evaluating companies in growth sectors or those that have not yet achieved profitability.

Initial research often relied on a profitability metric to gauge firm performance, as noted by Glick, W. H., et al. (2005). A significant majority of the 138 articles reviewed by Murugesan et al. (2016) incorporated profitability as a key performance indicator, with return on assets being particularly prevalent. Profitability was shown to be crucial for both established companies and new ventures, with a focus on sustained growth. Glick et al. (2005) highlighted that profitability metrics like earnings per share, return on equity, and return on assets, primarily driven by net income, are central in evaluating a firm's past performance. The use of Internal Cost of Capital (ICC) as a profitability measure, as discussed by Gebhardt et al. (2001) and Hou et al. (2012), reflects the complexity of predicting net income. The ICC is calculated by equating a stock's current price with the present value of expected future cash flows, offering another dimension in performance measurement.

Murugesan et al. (2016) expanded this singular perspective, by further identifying two more core variables in financial performance market value and growth. Their qualitative research underlined the importance of these dimensions in understanding a firm's overall health.

Qiu et al. (2014) examined the complexity of performance measurement from an accounting perspective, comparing accounting measures like EPS with market response measures to evaluate financial health. Market value, particularly market capitalization, has gained importance as a performance measure. This measure, reflecting external assessment and future performance expectations, is increasingly used alongside traditional profitability and growth metrics.

The concept of growth, the increase in revenue, as a performance measure, highlighted by Whetten (1987), emphasizes the significance of a firm's ability to increase in size and achieve economies of scale, thus enhancing profitability and stability.

We focussed our review more towards the revenue prediction being the objective of this research. In the recent times, revenue predictions have ascended in prominence, now ranking as the second most prevalent forecast metric, trailing only behind earnings estimations. This trend has been documented in scholarly works such as Ertimur et al. (2011) and Lorenz & Homburg (2018). Despite this growing importance, the field of research focusing on revenue predictability at global level remains notably underexplored.

As pointed out earlier, limited research in revenue focuses on market drivers, company profile as well as the financial statement parameters. Yunika Murdayanti, et.al (2014), in their research investigates factors influencing the predictability of revenue forecast errors on initial public offers. The study finds that macroeconomic indicators significantly impact forecast accuracy, with nation's higher GDP correlating with reduced forecasting errors. Additionally, the study identifies a collective significance of firm age, size, shareholder percentage, and industry variables in influencing the accuracy of revenue forecasts and their errors.

In applying the financial parameters for revenue predictability, we see the work of Marko Kureljusic, Lucas Reisch (2022) is critical for this study. This study evaluates the efficacy of machine learning algorithms for revenue prediction in European firms under IFRS, using publicly available data. Comparing the models built on machine learning algorithms with traditional financial analysts' forecasts across 3,000 firm-year observations from 2010 to 2019, the results indicate that machine learning achieves equal or greater accuracy. This research highlights the superior performance of the random forest model among 10 analysed machine learning models, achieving a significantly lower Mean Absolute Percentage Error (MAPE) of 13.29% compared to

financial analysts' 59.48% This underscores the potential of machine learning and predictive analytics in enhancing forecasting accuracy.
The prediction of revenue from financial parameters are conducted by modelling the input parameters (independent) and performance (dependent) from financial statements. Research shows that in early times OLS regression was the main method and recently, the machine learning algorithms are also used. Typically, the MAPE is used as the base measure for prediction accuracy. Kureljusic , Lucas Reisch (2022) research demonstrates that MAPE and modified MAPE like Median absolute Percentage Error (MdAPE) are powerful accuracy measures of predictability.

The Significance of the study:

A. Imperative for Enhanced Research in Revenue Prediction

Our comprehensive review of the existing literature, combined with our analytical examination, underscores a significant need for more extensive research in the domain of revenue prediction. This necessity arises in direct response to the escalating demand and importance of revenue forecasts in financial analysis. Despite the growing reliance on revenue figures as a key indicator of corporate health and market performance, the current body of research does not adequately address this area. Consequently, there exists a substantial gap in scholarly inquiry, particularly in developing robust methodologies and models for accurate revenue forecasting.

B. The Necessity for a Global Perspective in Research

The current research landscape reveals a geographical and sectoral limitation, with most studies concentrating on specific regions or countries, and often confined to certain industry sectors. This regional and sector-specific focus significantly hampers the generalizability and applicability of the findings. Therefore, there is a pressing need for a global study that encompasses diverse economic environments and industry sectors. Such a comprehensive approach would offer a more holistic understanding of revenue forecasting dynamics and enable the development of more universally applicable models and theories.

C. Examining the Impact of the COVID-19 Pandemic on Revenue Forecasting

The COVID-19 pandemic, which struck in 2020, had a profound and unique impact on the revenue streams of companies across the globe. This unprecedented event has necessitated a revaluation of existing revenue forecasting models, which may be less effective in predicting financial outcomes in a post-pandemic economy. Understanding the specific impacts of the pandemic on revenue predictability is crucial for adapting and refining forecasting methodologies to account for such extraordinary events. This aspect of research is particularly vital for enhancing the resilience and accuracy of revenue forecasts in the face of global disruptions.

Research Objectives

In light of these identified needs, the primary aim of this research is to address these critical gaps. We intend to:

- Develop and refine algebraic and neural network models for revenue prediction that can provide high level of accuracy.
- Conduct a global analysis that transcends regional and sectoral limitations, thereby contributing to a more universal understanding of revenue forecasting.
- Investigate the specific impacts of the COVID-19 pandemic on revenue predictability and adapt forecasting models to better accommodate such global disruptions.

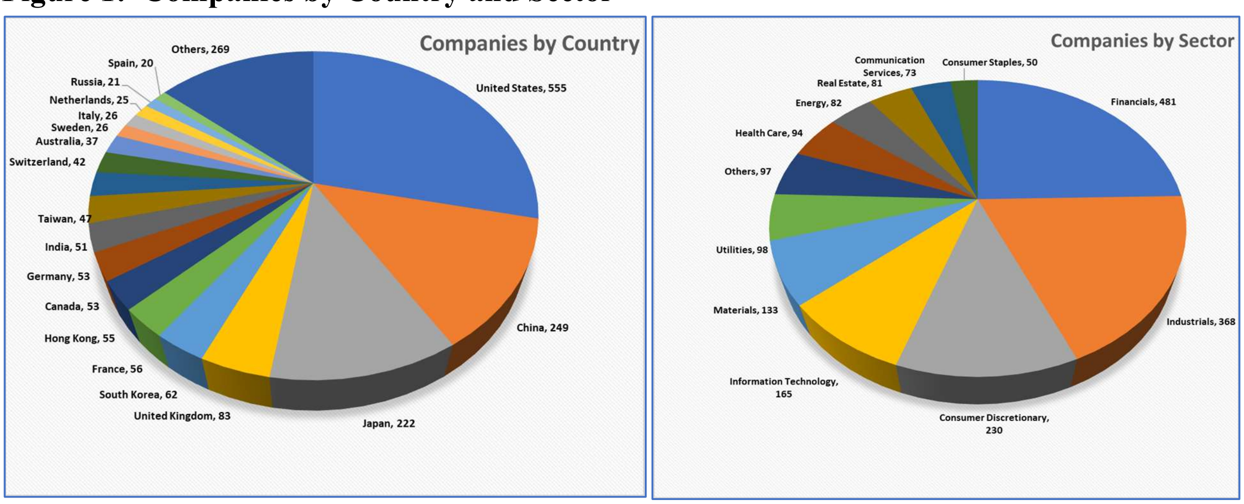
This research endeavour is committed to advancing the field of revenue forecasting by addressing these pivotal areas, thereby contributing to both academic knowledge and practical financial analysis techniques.

Data collection, Methodology and Models

Due to the global scale that we set out for the research, we took top 2000 large companies as listed by Forbes 2000 -2019. Out of this 1953 companies have data over the period of 2010-2022. The financial data was collected from multiple sources that are publicly available, mainly from the respective company webpages and standardized format published by Morningstar database. The models are trained using the data from 2010 to 2016. The out of sample outcomes are predicted for 2017-2021 and analysed for each year.

While the analysis of global data provides insights at a macro level, it is equally imperative to examine the sector-specific dynamics of individual companies and the economic status of their host countries for a deeper understanding. The data represented in Figure 1 is further analysed to develop the sector-capital market groupings, simply called groupings.

Figure 1: Companies by Country and Sector



Three largest sectors were chosen as below which are part of the groupings analysis.

Financials: This sector includes banks, insurance companies, real estate investment trusts (REITs), and investment funds.

Consumer Discretionary: This sector covers companies that produce goods and services considered non-essential by consumers, like automobiles, hotels, and luxury goods.

Industrials: This sector includes aerospace and defence, machinery, construction, engineering, and transportation industries

In addition, the capital market grouping of countries is also considered for classification of the companies, based on the Kenneth French classification.

United States: United states being the number one economy in the world, it is considered as a country grouping

Developed Countries: The developed countries are based on the capital market development in those countries. The list includes Italy, Luxembourg, Sweden, United Kingdom, Ireland, Germany, Singapore, Canada, Portugal, Finland, Greece, France, Hong Kong, Japan, Denmark, Switzerland, Norway, Netherlands, Austria, Australia, Spain and Belgium

Developing Countries: The developing countries list includes Taiwan, South Korea, Saudi Arabia, Brazil, China, Thailand, India, Qatar, Malaysia, Morocco, Mexico, Israel, Turkey, Kuwait, Chile, South Africa, United Arab Emirates, Philippines, Russia, Vietnam, Argentina, Poland, Egypt, Indonesia, Hungary, Colombia, Peru, Kazakhstan, Bahrain, Nigeria and Kenya.

While the many other factors such as market potential and competitive setting, will significantly impact the revenue growth of companies, this research is mainly considering the following input variables for the predictive modelling.

Income Statement: Total Revenue, Non-Interest Expenses, Cost of Goods Sold (COGS), Net Property, Plant, and Equipment (PPE), Provision and Impairment for Loan Losses and Credit Risk, Irregular Income/Expenses, SGA (Selling, General and Administrative Expenses), Reserves/Accumulated, Comprehensive Income/Losses, Retained Earnings/Accumulated Deficit, **Operating Income,** Depreciation, Total Contractual Obligations, **Irregular income,** Provision for Income Tax, Net Income (Yes Flag), **Net Income.**

Balance sheet Variables: Total Assets, Total Liabilities, Cash and Cash Equivalents, Net Intangible Assets, Total Current Assets, Inventories, Total Current Liabilities, Total Equity

Cashflow Variables: Cash Flow from Investing Activities, Purchase/Sale of Investments, Cash Flow from Financing Activities, **Operating Cash Flow, Cash Dividend, Change in Cash, Debt Repayment, Change in Operating Capital**

In addition to incorporating direct accounting data from financial statements, our modelling also integrates financial ratios and economic parameters as input variables

Financial Ratios: Price to Earnings Ratio, Current Ratio, Net Profit Margin, Debt to Equity Ratio, Return on Equity, Return on Assets

Economic Parameters: Gross Domestic Product, Country Inflation

Models applied in this research:

Broadly the models applied in this research can be classified as Linear Models and Neural Network Models.

Linear models aim to establish the relationship between the financial outcomes and the input variables through ordinary least square method. The variables are related through a linear equation. While the primary use of OLS is to understand the functional relationship between variables, it also provides a basis for inference, such as hypothesis testing and constructing confidence intervals for the coefficients.

AR Model: Auto regressive linear model where the previous year's financial outcome is only used as input variable.

ARX Model: Auto regressive linear model with exogenous inputs, where the previous year's financial outcome along with the group of variables were used as input variables.

REG MODEL: Non-Auto regressive linear model with exogenous inputs only from the previous years' group of variables.

While OLS models fit linear relationship between the financial outcome and the input variables, the neural network model can identify the nonlinear relationships and segmented relationships and store them as weights and biases, there by potentially further improve the predictability.

AR_ANN: Auto regressive Artificial Neural Network Back Propagation (ANN-BP) model where the previous year's financial outcome is only used as input variable.

ARX_ANN: Auto regressive Artificial Neural Network Auto Regressive Back Propagation (ANN-BP) model with exogenous inputs, where the previous year's financial outcome along with the group of variables were used as input variables

REG_ANN: Non-Auto regressive Artificial Neural Network Auto Regressive Back Propagation (ANN-BP) model with exogenous inputs only from the previous years' group of variables.

The neural network Model is illustrated with the typical settings is given in the annexure.

While Neural network models use fixed weights during inference after being trained, the Bayesian models treat these weights as random variables, following a probability distribution. This allows the models to establish standard deviation and distribution of each of the models.

AR_BNN: Auto regressive Bayesian Neural Network model where the previous year's financial outcome is only used as input variable.

$$\mu_t, \sigma_t = f(W \cdot Y_{t-1} + b)$$

Equation: (1)

ARX_BNN: Auto regressive Bayesian Neural Network (ARX_BNN) model with exogenous inputs, where the previous year's financial outcome along with parameters were used as input variables

$$\mu_t, \sigma_t = f(W \cdot Y_{t-1} + V \cdot X_{t-1} + b)$$

Equation: (2)

REG_BNN: Non-Auto regressive Bayesian Neural Network (REG_BNN) model with exogenous inputs only from the previous years' group of variables

$$\mu_t, \sigma_t = f(W \cdot X_{t-1} + b)$$ Equation: (3)

Where μt is the predicted mean and σt is the predicted standard deviation at time t.

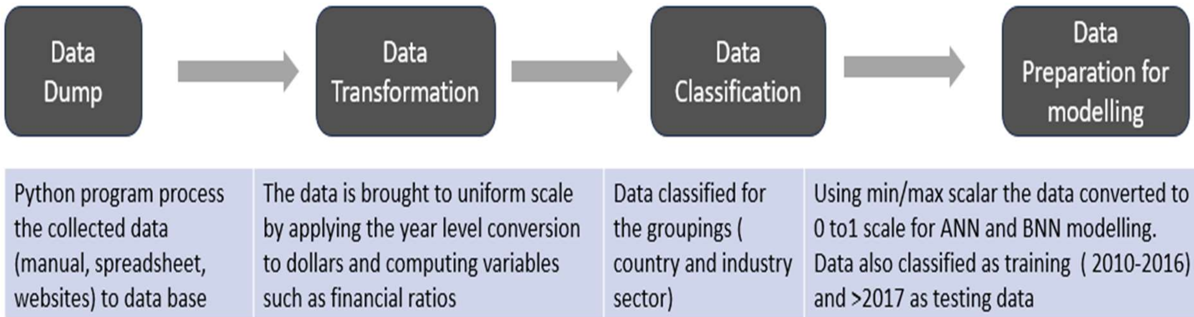
- $\mathbf{Y}_{t-1}$ represents the vector of past observations In this case, between 2010 to 2016
- $\mathbf{b}$ represents the vector of biases, also treated as random variables.
- $\mathbf{W}$ and $\mathbf{V}$ represents the matrix of weights, where each weight w_{ij} is a random variable with its own prior distribution.
- $f(\cdot)$ is a nonlinear activation function applied to the linear combination of weights and past observations.
- ϵ_t is the error term which may also be modeled with a distribution to capture the noise in the data
- $\mathbf{X}_{t-1}$ represents the vector of additional variables or predictors at time

The Bayesian neural network Model is illustrated with the typical settings is given in the annexure.

The Random walk model assumes that the current year value of the outcome to the next year outcome plus a random step. We added randomized values within 1 sigma limits of the outcome. The models consistently returned lower accuracy and not discussed at length.

Figure 2, shown as below summarizes the methodology, the input and output variables, data transformation, the data groupings, the models developed and applied for prediction, the statistical tests and the accuracy measurements.

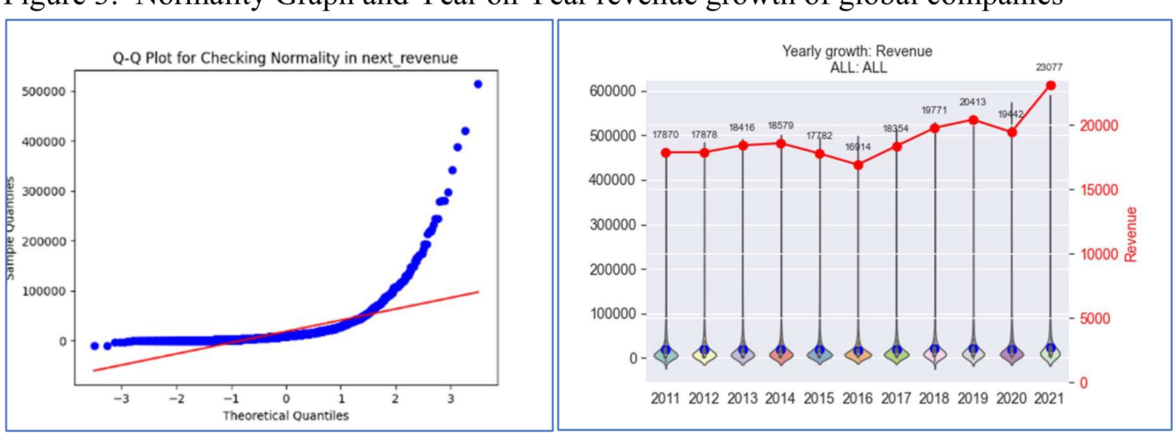
Figure 2 : Data collection and Dimensions.



Outcomes	Models	Years	Variables	Measures	Tests	Countries, Sectors
Total Revenue of the organization for the subsequent year using cross sectional analysis	• Auto Regressive (AR) • Auto Regressive with variables (ARX) • Artificial Neural Network(AR_ANN, ARX_ANN) • Bayesian Neural Network (AR_BNN, ARX_BNN) • Non Regressive (X_REG, ANN_XREG) Random Walk (RW)	2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022	Financial statement direct variables Financial Ratios Economic Data	MAPE MAPE(WIN2.5 %) ADJ RSQUARE ACCURACY MAE MSE WALD COUNT	Anderson Darling and Q-Q test for Normality Mood Median Kruskal-Wallis (KS) Kolmogorov-Smirnov (KS)	59 countries 83 countries 1951 companies from Forbes 2000-2019 list of top companies across the world

The revenue data for the companies selected across the world shows a skewed distribution with a long tail. The Q-Q plot (year 2018) as seen in Figure 3 and the Anderson-Darling test show the data is not normal. The year-on-year violin chart with mean shows that the average revenue is 3.4% of the maximum revenue. The global revenue average of the selected companies has grown from 2011 to 2021 USD 17.8 Bn to USD 23 Bn with CAGR 2.35% growth. (Figure 3) In the same period the world GDP grew from USD 73.9 Tn to 96.9 Tn with CAGR of 2.49%. showing that the companies' growth closely follows the GDP at macro level.

Figure 3: Normality Graph and Year on Year revenue growth of global companies



Model Runs, Analysis and Results

According to Marco and Lucas (2022), the projection of performance metrics is normally an integral facet of empirical accounting research. Divergent from descriptive or explanatory inquiries, forecasting analyses endeavour to ascertain the predictive capacity of specific input variables concerning a designated outcome variable. Such studies are iterative requiring several trials with various conditions, targeting the formulation of the most efficient predictive model pertinent to the research question at hand (Ding et al., 2020). This methodology, rooted in the discipline of information systems, aligns with the principles of design science research (Hevner, March, Park, & Ram, 2004), On the other hand, it is also a subject of management research from the perspective of forecasting of financial outcomes and setting organizational goals and strategies. It also helps to understand the potential peer growth and navigate successfully in the competitive settings.

There is also a company specific angle to the revenue predictions, Gizem Topaloglu, et.al, (2023) argues, that strategic planning, organizing, directing, and control of all financial activities within a company or institution constitute financial management. Along with being crucial to fiscal management, financial management also entails applying management principles to an organization's financial assets. Predicted revenue is a key input to financial planning. Besides, predicting company's growth is one of the most important aspects of entrepreneurship research.

Coming back to the data science aspect, we conducted 5468 trials on OLS models and 3000 trails on ANN and BNN models. This helped to narrow the input variables as well as identify the best performing Neural Network models. The techniques such as factorization and principal component method were not applied as getting the coefficient of influence is critical outcome of this research. The REG model and ARX model trials show that the revenue in most of the cases is a direct function of the previous year's revenue. (illustrated as AR model). There is an overall 13.5% MAPE improvement observed over the AR model with additional input variables.

For the global set of data, with the group name called "ALL -ALL" indicating all sectors as well as all group countries, table 1 gives the details of X_REG and ARX regression coefficients, the corresponding p values and in the order of Pearson correlation. The variables which show significant influence (p values < 0.05) in both the models are highlighted.

Table 1 : Regression coefficients of selected variables in ARX , REG (OLS) models

VARIABLE	ARX	X_REG	PEARSON	ARX_P	X_REG_P
Revenue	0.9756		0.975	0	
Equity	-0.0012		0.591	0.87	
Net PPE	**0.0378**	**0.5166**	**0.566**	**0**	**0**
Depreciation	-0.6091		0.48	0	
Inventories	**0.1363**	**2.706**	**0.457**	**0**	**0**
Intangible	**0.0502**	**0.0369**	**0.375**	**0**	**0.052**

Total Assets	0.0003		0.251	0.947	
Total Liabilities	-0.0014		0.211	0.73	
Contract Obligation		-0.004	0.117		0.001
Debt Payment	-0.0054	0.3103	0.032	0.702	0
Dividend	-0.0011		0.019	0.622	
Change in Cash	0.0502		0.001	0	
Change in Op Capital	**-0.0678**	**-0.1667**	**0.001**	**0**	**0**
Cash Flow - Finance		**-0.3365**	-0.033		0
Investment Sale	-0.0124	-0.2733	-0.078	0.937	0.502
Net Investment	-0.0427	0.071	-0.078	0.786	0.862
Cash Flow Investing	-0.0002		-0.207	0.991	
Income Tax		-5.308	-0.496		0
SGA	**-0.0861**	**-1.8458**	**-0.604**	**0**	**0**

The past revenue (t-1) is a key input variable and has coefficient of 0.975. In addition, Inventories, SGA, change in operating capital are significant variables in predicting revenue in the global model. Though with poor accuracy, the REG model that does not take past revenue as part of input variable, finds Net PPE, Inventories, Debt Payment, change in operating capital, financial cash flow, investment sales, SGA and Income tax as influencing variables in predicting the overall revenue.

The hypothesis, denoted as H0, proposes that no input variable exerts a significant influence on the outcome. This hypothesis is rejected due to the presence of input variables with coefficient p-values less than 0.05, indicating significant influence. For instance, the p-value of the coefficient for past revenue is nearly zero, exemplifying this significant impact. The overall accuracy for selected models as an average of all groupings is given below in table 2. For all the regression models the mean absolute error, the root mean squared error and Adjusted R-Square shows no major variation.

The substantial improvement in the Mean Absolute Percentage Error (MAPE) for the ARX model, which is better by 13.5% compared to the AR model, underscores the influence of significant input variables on the predicted outcome. The Wald counts for the neural network models are lower than the OLS models, indicating that more data may be required for further training the model. The REG models on the other hand, shows poor accuracy as noted earlier.

Table 2 : The average accuracy comparison between key models.

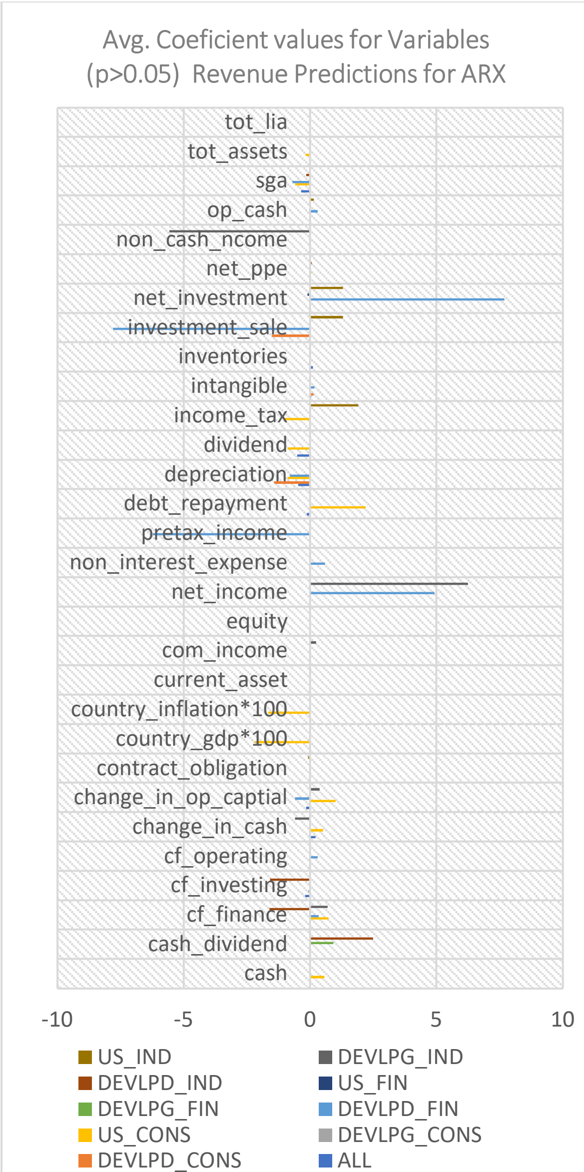
Model	**Mape**	**97.5 WIN Mape**	**MAE (BUSD)**	**RMSE (BUSD)**	**Adj R-SQ**	**Slope**	**Wald Count**
AR	0.328	0.140	2.53	5.82	0.96	0.92	7.20
ARX	0.284	0.131	2.55	5.85	0.96	0.92	7.50
ARX_ANN	0.291	0.133	2.57	5.94	0.96	0.92	6.10
ARX_BNN	0.315	0.134	2.70	6.49	0.95	0.91	6.40

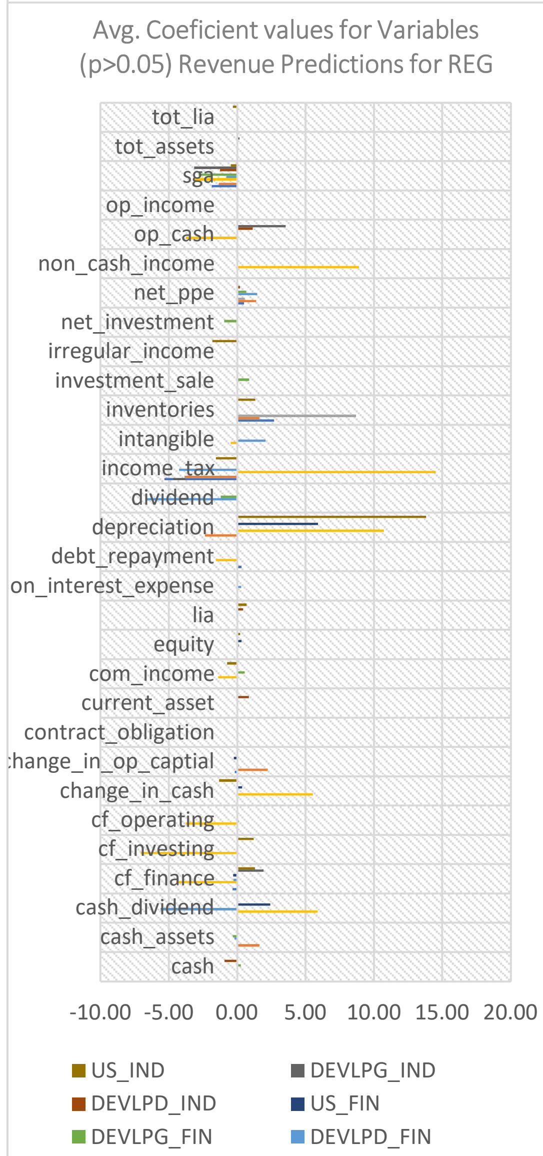
X_REG	0.794	0.646	8.46	16.34	0.67	0.75	6.40

We generated distinct models for each grouping and observed different set of influencing variables form the optimized regression equations. This means that the industrial sector and the capital market grouping impacts the level of influence of each variable over the revenue prediction. For example, Cash Flow – Investing, Cash Flow Finance, Change in Operating Capital, Depreciation and Investment sale significantly influence 3 or more groupings. It also means that Cashflow and Income statement has more bearing on future income. US – Consumer Discretionary and Developed Countries – Financials have the maximum input variables that appear in the prediction equation of the ARX model. This also reflects in the coefficient of the total revenue which is 87% as against the 97% in the global model. Interestingly, for the less accurate REG models, we find US – Consumer Discretionary and Developed Countries – Financials show highest number of variables (13) influencing the predictability. The Adjusted R Square and slope of trend chart between the predicted outcome and actual outcome is much higher in both cases, showing that Non regressive financial parameters have high influence in the revenue predictability. Another important observation is that the total predictors in the US and Developed countries is greater than 20 while in developing countries it is 9, indicating potentially the market/ economic / business environment variation between the developed countries or the maturity of financial reporting, which may require further investigation.

Following diagram shows (Figure 4) that variables like Dividend, Income Tax are strongly influencing the next year income covering all the 10 groupings

Figure 4 : Coefficients of significant variables across the groupings.



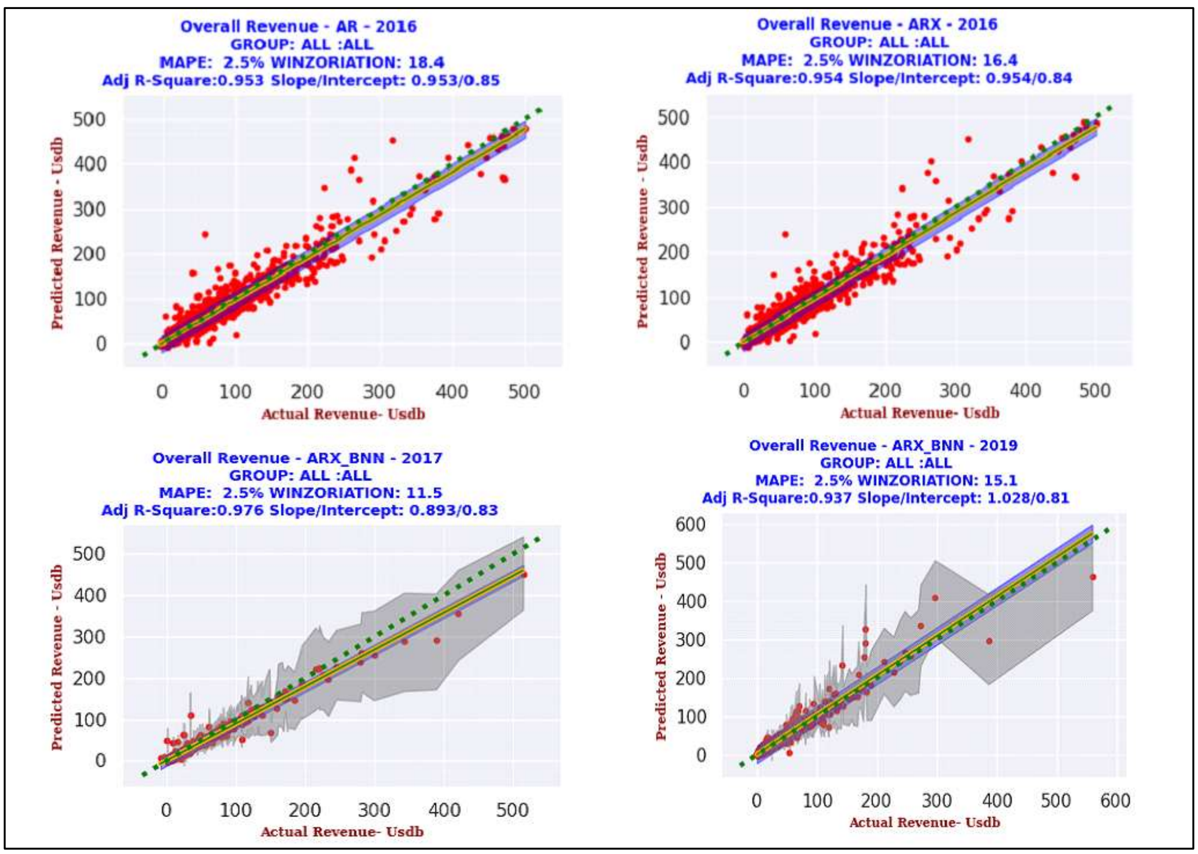


In ALL-ALL grouping in ARX modelling, Inventories, Intangibles, Change in Operating Capital, Net PPE have significant coefficients in the regression equation.

Atleast one variable in all 10 groupings is observed to have a coefficient in the regression equation whose p-value < 0.05 and hence the Ho is rejected. The details of coefficients and p-value is given in the annexure in table 5,6 and 7.

Objective 2 & 3: Predictive Modelling Technique Evaluation and the influence of covid pandemic on predictability of revenue: The OLS models are best suited for quantifying the influence of each parameter over the predicted outcome. However due to potential non linearity and ability to have categorical variables along with continuous variables in the input side, the neural network models are also promising. The neural network model with non linear activation functions is opt for handling non linearity. The common type of neural network model (ANN models) predicts deterministic outcome for a given set of input variables. However, the Bayesian Neural network models (BNN models) as explained in equation 1, 2 and 3 have been developed with two heads, viz., one for the average of the predictions and the other head being for the standard deviations, thus capable of giving confidence level for each set of input variables for any given alpha level, thus moving to a stochastic outcome. Hence these models are added into this research. The predicted data set need a visual representation to evaluate any major anomaly. The following Figure 5, shows the global (ALL-ALL) model diagrams for visual interpretation. The chart allows to develop 3 sigma confidence intervals based on the linear relationship. The slope less than one for the test data (marked as 2016 in the graph title), shows that predicted values are less than the actuals, indicating further potential to improve the prediction accuracy. The BNN charts shows individual data point confidence interval limits in addition to the population level confidence intervals. The confidence intervals become much wider at higher revenue, potentially due to the higher variability of the input variables. In 2019, the chart shows that the slope is greater than one, indicating the revenue dip due to the covid pandemic and the inability of the model to completely absorb this variation from the input variables and hence higher MAPE.

Figure 5 : X-Y chart view of predicted vs actual revenue for different models



While the visual observations give insights into the mechanics of the prediction models, the statistical tests of mood median test, Kruskal-Wallis (KS) and Kolmogorov-Smirnov (KS) show in table 3 that, the ARX models, have less test results with p-value less than the level of significance (alpha = 0.05). It means that the H0 is rejected in 13 tests in ARX Models and in 18 tests in AR model. In all the cases together, the ARX_BNN model best performs (0nly in 8 tests the null hypothesis is rejected). This is visible even in the REG models where, REG_BNN shows more acceptance of Ho Model.

Table 3: Bt Year and Model, Hypothesis rejection for statistic to be the same.(H0 rejected)

Year	**AR**	**AR ANN**	**AR BNN**	**ARX**	**ARX ANN**	**ARX BNN**	**REG BNN**	**REG**	**REG ANN**	**Yearly Total**
2016	4	8	3	2	1		15	16	12	**61**
2017	1	5	1				6	9	10	**32**
2018	1	3					6	9	9	**28**
2019	5	8	5	5	4	5	11	13	14	**70**
2020	3	3	3	3	3	3	5	3	7	**33**
2021		2					8	11	10	**31**
ALL	4	9	3	3	2		14	14	13	**62**
Model Level Total	**18**	**38**	**15**	**13**	**10**	**8**	**65**	**75**	**75**	**317**

The yearwise comparison clearly indicating that in 2019 more number of rejections of the hypothesis than other years in the out of sample testing. Due to the larger data set in ALL years and test of 2010-2016, the number of rejections are higher. The lower and comparable rejections for the year 2020, 2021 shows the robustness of the models. The higher AR-ANN model Ho rejections show that the auto regressive neural network models may not be effective, even the accuracy parameters show less deviation.

For individual capital market grouping and sector the ARX model was defined with variables selected to either match the accuracy measurements or fair better. (Refer table 4). The higher degree of reduction of 97.5% MAPE comparing to the over all MAPE shows the long tail of high revenue companies in each of the groupings. United States 97.5% MAPE is 34% better than developed as well developing countries, again showing the business and economic maturity. The developed countries financial MAPE is significantly higher which requires further study.

Table 4: Accuracy measures across the sector-capital market groupings

Group Countries	Sector	**Model**	**Mape**	**97.5 WIN Mape**	**MAE (BUSD)**	**RMSE (BUSD)**	**Adj R-SQ**	**Slope**	**% Wald**
ALL	ALL	AR	0.34	0.17	3.04	8.58	0.96	0.904	0.50

		ARX	0.30	0.15	3.05	8.59	0.96	0.904	0.50
DEVELOPED COUNTRIES	CONSUMER DISCRETIONARY	AR	0.19	0.11	3.33	7.40	0.98	0.914	0.58
		ARX	0.14	0.11	3.46	7.99	0.98	0.902	0.58
	FINANCIALS	AR	1.05	0.23	3.68	11.27	0.83	0.845	0.67
		ARX	0.93	0.20	3.71	10.93	0.84	0.854	0.67
	INDUSTRIALS	AR	0.17	0.13	2.66	5.54	0.94	0.877	0.50
		ARX	0.17	0.12	2.68	5.57	0.94	0.876	0.58
DEVELOPING COUNTRIES	CONSUMER DISCRETIONARY	AR	0.17	0.14	2.71	4.88	0.98	0.91	0.67
		ARX	0.16	0.13	2.74	4.89	0.98	0.912	0.75
	FINANCIALS	AR	0.16	0.13	1.02	2.20	0.99	0.97	0.67
		ARX	0.16	0.13	1.02	2.19	0.99	0.971	0.67
	INDUSTRIALS	AR	0.77	0.20	2.85	5.11	0.98	0.927	0.42
		ARX	0.58	0.18	2.81	5.08	0.98	0.935	0.50
UNITED STATES	CONSUMER DISCRETIONARY	AR	0.16	0.10	2.31	4.77	0.98	0.963	0.67
		ARX	0.15	0.10	2.29	4.67	0.98	0.969	0.67
	FINANCIALS	AR	0.14	0.09	1.37	3.07	0.98	0.972	0.75
		ARX	0.13	0.10	1.52	3.30	0.98	0.969	0.83
	INDUSTRIALS	AR	0.14	0.11	2.33	5.35	0.95	0.924	0.58
		ARX	0.12	0.09	2.28	5.30	0.95	0.93	0.50

Conclusions and further study

The revenue prediction is the second critical financial outcome of the companies. This research conducted at global scale using OLS and Neural networks demonstrate high level of accuracy for

revenue predictability, specifically in Auto Regression models. when applied in conjunction with the of financial and economic parameters as inputs. This high accuracy surpasses the analyst predictions and particularly significant given that analysts typically possess detailed insights into the specific factors influencing the financial performance of the companies they monitor. In their 2022 study, Marko Kureljusic and Lucas Reisch observed that analyst expectations on revenue of European companies displayed 50% variability. Contrastingly, our global model exhibited a 17% variation, which is substantially higher. Further, the autoregressive exogenous (ARX) models outperformed standard autoregressive models by 14%.
In terms of geographical analysis, the predictability of revenue for U.S. companies was markedly higher, with their accounting information proving especially pertinent in revenue predictability. Following the U.S. companies, the companies from developed countries showed moderate level of predictability, while those in developing countries presented greater opportunities for refining these models.
On a global scale, our research demonstrates that the Auto Regressive Exogenous-Bayesian Neural Network (ARX_BNN) model achieves superior performance in predictability tests. This finding opens up further avenues for research expansion. Our study covered 3 different sectors, suggesting the potential to extend this analysis to additional sectors on a global scale. Moreover, each sector can be further dissected into specific industries, where industry-relevant variables can enhance predictability. For instance, in the banking sector, factors such as deposit growth in host countries, mortgage rates, and consumer spending patterns could be incorporated as additional variables to refine the predictive accuracy. The research work done in net income can further boost the research in revenue also.

References

1. Ding, K., Lev, B., Peng, X., Sun, T., & Vasarhelyi, M. A. (2020). Machine learning improves accounting estimates: Evidence from insurance payments. Review of Accounting Studies, 25(3), 1098-1134. https://doi.org/10.1007 /s11142-020-09546-9
2. French, Kenneth R. Description of Fama/French Factors. (dartmouth.edu)
3. Glick, W.H., Washburn, N.T. and Miller, C.C. (2005). The Myth of Firm Performance. Proceedings of the Annual Meeting of American Academy of Management, Honolulu, 5-10 August 2005.
4. Gizem Topaloglu, Tolga Ahmet KalaycÄ±1, Kaan Pekel1, Mehmet Fatih Akay2. International Journal of Mathematics and Computer in Engineering 2(1):117-122, DOI:10.2478/ijmce-2024-0009, October 2023.
5. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75-105. https://doi.org/10.2307/25148625
6. Hou, Kewei & van Dijk, Mathijs & Zhang, Yinglei. (2010). The Implied Cost of Capital: A New Approach. Ohio State University, Charles A. Dice Center for Research in Financial Economics, Working Paper Series, 53. 10.2139/ssrn.1229964.

7. Kureljusic, Marko and Reisch, Lucas. (February 28, 2022). Revenue forecasting for European capital market-oriented firms: A comparative prediction study between financial analysts and machine learning models. Corporate Ownership & Control, 19(2), 159–178. https://doi.org/10.22495/cocv19i2art13. Available at SSRN: https://ssrn.com/abstract=4057309
8. Lorenz, Tanja & Homburg, Carsten. (2018). Determinants of analysts' revenue forecast accuracy. Review of Quantitative Finance and Accounting. 51. 10.1007/s11156-017-0675-4.
9. MurugesanSelvam, JayapalGayathri, VinayagamoorthiVasanth, Kasilingam Lingaraja, Sigo Marxiaoli. (July 2016). Determinants of Firm Performance: A Subjective Model. International Journal of Social Science Studies, Vol. 4, No.7.
10. Qiu, Xin Ying & Srinivasan, Padmini & Hu, Yong. (2014). Supervised Learning Models to Predict Firm Performance With Annual Reports: An Empirical Study. Journal of the American Society for Information Science and Technology. 65. 10.1002/asi.22983.
11. Whetten, David A. (1987). "Organizational Growth and Decline Processes." Annual Review of Sociology, vol. 13, pp. 335–58. JSTOR, http://www.jstor.org/stable/2083252.
12. Yunika Murdayanti. (2014). Factors Affecting Prediction of Revenue Forecast Error Study on IPO Firms at Indonesians Stock Exchange Period 2007-2012. Society of Interdisciplinary Business Research, vol 4(1).