**Predicting Company’s Financial Failure or Bankruptcy Using Financial Ratios.**

**INTRODUCTION:**

Multiple research has been done on financial stability or financial failure of an organization since 1930[1]. For the prediction of Bankruptcy, different research have been in place since late 1960’s to use financial ratios. Bankruptcy is filed when a company that has no operational source of money to operate the business and is in no position to repay the debts it owes to its creditors. Bankruptcy is a difficult situation to be as the company comes down to a stand-still state and leaves the employees and suppliers or customers in a high and financial instability state. When the competitors are increased and uncertainty in the global economy these days which is the cause to drive large organizations towards bankruptcy. The recent example of the British travel company, Thomas Cook which suddenly declared bankruptcy and left 21,000 people out of work and million dollars of investor money down the drains. Thus, it is impossible to overstate the damage caused in terms of the financial loss.[6]

The process of bankruptcy begins with a petition which is filed by the debtor or on behalf of creditors. The assets of the debtor are measured and their assets are used to repay a portion of their outstanding debt. To free itself from debt obligations, Bankruptcy filing is undertaken by the company. Debts that are not paid to creditors are forgiven for the owners. Bankruptcy filing varies in different countries. If you file for the bankruptcy in India, it will not go down well with our credit rating which means that it is tough to get a new loan if you plan to start afresh. However, it would save us from any financial trouble.

To assess and analyze the performance of firms, Financial ratios are used. Based on the various parameters, we compare and analyze a firm’s overall health. The Financial ratios are used by the Investors to get an insight to the firm’s profitability and its investment prospects. The Financial ratios are helpful in forecasting the firm’s future health by using the historical data from financial statements. Also they help to compare and contrast financial performance of firms among their competitors additionally. Inspite of many research in this area, several attempt towards exploring this relationship has made which led to limited success to the variability existence in every stock markets. The variance in volatility of these markets makes it difficult to obtain a uniform measure[5].

The Prediction of Bankruptcy is a very important task for financial institutions. Their main aim is to predict the likelihood that a firm may go bankrupt or not. Financial institutions are in need of effective prediction models in order to make appropriate lending decisions [2], stock price prediction and more. Some of the Machine learning algorithms and techniques that are used for prediction of Bankruptcy are Support Vector Machines (SVM), Naïve Bayes Classifier (NB), Classification and Regression Trees(CART), k-nearest neighbor (KNN), and MLP(Multi-Layer Perceptron). Features for the machine learning models are stepwise discriminant analysis (SDA) (Fisher, 1936), stepwise logistic regression (SLR) (Fisher & Yates, 1963), and t-testing (Zimmerman, 1997)3 and the wrapper based methods of the genetic algorithm (GA) (Holland, 1975) and recursive feature elimination (RFE) (Guyon, Weston, Barnhill, & Vapnik, 2002). We also calculate the model score, Precision, Recall, F1 score and ROC score and compare which model gives highest F1 score. We are evaluating the models using Sum of Squared Errors (SSE), Root Mean Squared Errors (RMSE), Mean Absolute Error (MAE), Relative Absolute Error (RAE).

**RELATED WORK:**

The Features that are selected such as financial ratios and techniques such as statistical or machine learning techniques are the important factors which strongly affects the performance of the prediction of Bankruptcy. While many novel prediction techniques have proposed, very few of them have analyzed the discriminatory power of the features which are related to bankruptcy prediction. The financial ratios (FRs) and the corporate governance indicators (CGIs) have been found to be important type of input variable[2]. However, the Bankruptcy prediction performance is obtained by combining CGIs and FRs has not been fully examined. Only some selected CGIs and FRs have been used in related studies and the chosen features differed from study to study. In this, we will explore all the Financial Ratio’s and Corporate Governance Indicators **through SMOTE Analysis, Bagging, Boosting and some of the Deep Learning methods**.

The prediction performance of Bankruptcy is obtained by combining seven different categories of Financial Ratio’s and five different categories of corporate governance indicators. The experimental results which is based on the real-world dataset from Taiwan proves that FR categories of solvency and profitability and the CGI categories of board structure and ownership structure are the most important features in bankruptcy prediction.

**DATA:**

Company bankruptcy was defined based on the business regulations of the Taiwan Stock Exchange. The Taiwan Stock Exchange is a financial institution located in Taipei, Taiwan. It has over 900 listed companies. It was established in 1961 and began operating as a stock exchange on 9 February 1962. The data were collected from the Taiwan Economic Journal for the years 1999 to 2009. The Taiwanese Bankruptcy Prediction data was obtained from UCI Machine Learning Repository.

The Taiwanese Bankruptcy Dataset contains:

* 95 features (X1-X95, business regulations of Taiwan Stock Exchange)
* 1 Vector of labels

Our aim is to use these features to understand their impact on the selected models and how they can help us recognizing the companies that are close to bankrupcty. All the variables that are part of this dataset are numeric and complete in nature. The data includes a majority of numerical attributes that help understand the possibility of bankruptcy. We will use various predictive models to see how accurately we predict which companies will face bankruptcy in the future. This is the same data set that was used by Liang D and Co[2] for research and donated to UCI community.

|  |  |
| --- | --- |
| X1 Cost of Interest-bearing Debt | X49 Fixed Assets Per Employee |
| X2 Cash Reinvestment Ratio | X50 total assets to GNP price |
| X3 Current Ratio | X51 Return On Total Assets(C) |
| X4 Acid Test | X52 Return On Total Assets(A) |
| X5 Interest Expenses/Total Revenue | X53 Return On Total Assets(B) |
| X6 Total Liability/Equity Ratio | X54 Gross Profit /Net Sales |
| X7 Liability/Total Assets | X55 Realized Gross Profit/Net Sales |
| X8 Interest-bearing Debt/Equity | X56 Operating Income /Net Sales |
| X9 Contingent Liability/Equity | X57 Pre-Tax Income/Net Sales |
| X10 Operating Income/Capital | X58 Net Income/Net Sales |
| X11 Pretax Income/Capital | X59 Net Non-operating Income Ratio |
| X12 Working Capital to Total Assets | X60 Net Income-Exclude Disposal Gain or Loss/Net Sales |
| X13 Quick Assets/Total assets | X61 EPS-Net Income |
| X14 Current Assets/Total Assets | X62 Pretax Income Per Share |
| X15 Cash/Total Assets | X63 Retained Earnings to Total Assets |
| X16 Quick Assets/Current Liability | X64 Total Income to Total Expenses |
| X17 Cash/Current Liability | X65 Total Expenses to Assets |
| X18 Current Liability to Assets | X66 Net Income to Total Assets |
| X19 Operating Funds to Liability | X67 Gross Profit to Sales |
| X20 Inventory/Working Capital | X68 Net Income to Stockholder's Equity |
| X21 Inventory/Current Liability | X69 One if Net Income is Negative for the Last Two Years; Zero Otherwise |
| X22 Current Liabilities/Liability | X70 (Inventory +Accounts Receivables) /Equity |
| X23 Working Capital/Equity | X71 Total Asset Turnover |
| X24 Current Liabilities/Equity | X72 Accounts Receivable Turnover |
| X25 Long-term Liability to Current Assets | X73 Days Receivable Outstanding |
| X26 Current Liability to Current Assets | X74 Inventory Turnover |
| X27 One if Total Liability exceeds Total Assets; | X75 Fixed Asset Turnover |
| X28 Equity to Liability | X76 Equity Turnover |
| X29 Equity/Total Assets | X77 Current Assets to Sales |
| X30 (Long-term Liability+Equity)/Fixed Assets | X78 Quick Assets to Sales |
| X31 Fixed Assets to Assets | X79 Working Capital to Sales |
| X32 Current Liability to Liability | X80 Cash to Sales |
| X33 Current Liability to Equity | X81 Cash Flow to Sales |
| X34 Equity to Long-term Liability | X82 No-credit Interval |
| X35 Liability to Equity | X83 Cash Flow from Operating/Current Liabilities |
| X36 Degree of Financial Leverage | X84 Cash Flow to Total Assets |
| X37 Interest Coverage Ratio | X85 Cash Flow to Liability |
| X38 Operating Expenses/Net Sales | X86 CFO to Assets |
| X39 (Research and Development Expenses)/Net Sales | X87 Cash Flow to Equity |
| X40 Effective Tax Rate | X88 Realized Gross Profit Growth Rate |
| X41 Book Value Per Share(B) | X89 Operating Income Growth |
| X42 Book Value Per Share(A) | X90 Net Income Growth |
| X43 Book Value Per Share(C) | X91 Continuing Operating Income after Tax Growth |
| X44 Cash Flow Per Share | X92 Net Income-Excluding Disposal Gain or Loss Growth |
| X45 Sales Per Share | X93 Total Asset Growth |
| X46 Operating Income Per Share | X94 Total Equity Growth |
| X47 Sales Per Employee | X95 Return on Total Asset Growth |
| X48 Operation Income Per Employee |  |

**IMPLEMENTATION:**

Comparing to all the feature selection methods Step wise Linear Regression was giving good percentage [2].

**SMOTE:**

SMOTE stands for **Synthetic Minority Oversampling Technique.** SMOTE is a machine learning technique that solves problems that occur when using an **imbalanced data set**. **Imbalanced data** is data in which observed frequencies are very different across the different possible values of a categorical variable. Basically, there are many observations of some type and very few of another type. Example : The classification model that allows us to use customer data to make a prediction of whether the visitor will buy the new product.

Most e-commerce shoppers do not buy: often, many come for looking at products and only a small percentage of visitors actually buy something. **Our data set will be imbalanced, because we have a huge number of non-buyers and a very small number of buyers.** Accuracy is a bad machine learning metric when working with imbalanced data.

The most straightforward method to counteract class imbalance is undersampling. Undersampling means that **you discard a number of data points of the class that is present too often**.The disadvantage of undersampling is that you lose a lot of valuable data.

Another simple solution to imbalanced data is oversampling. Oversampling means making duplicates of the data that is the least present in your data set. You then add those duplicates to your data set.

SMOTE is an algorithm that performs data augmentation by creating **synthetic data points** based on the original data points. SMOTE can be seen as an advanced version of oversampling, or as a specific algorithm for data augmentation. The advantage of SMOTE is that you are **not generating duplicates**, but rather creating synthetic data points that are **slightly different** from the original data points.

SMOTE is an improved alternative for oversampling

The **SMOTE algorithm** works as follows:

* You draw a random sample from the minority class.
* For the observations in this sample, you will identify the k nearest neighbors.
* You will then take one of those neighbors and identify the vector between the current data point and the selected neighbor.
* You multiply the vector by a random number between 0 and 1.
* To obtain the synthetic data point, you add this to the current data point.

This operation is actually very much like **slightly moving the data point in the direction of its neighbor**. This way, you make sure that your synthetic data point is **not an exact copy**of an existing data point while making sure that it is **also not too different** from the known observations in your minority class.

The challenge of working with imbalanced datasets is that most machine learning techniques will ignore, and in turn have poor performance on, the minority class, although typically it is performance on the minority class that is most important.

One approach to addressing imbalanced datasets is to oversample the minority class. The simplest approach involves duplicating examples in the minority class, although these examples don’t add any new information to the model. Instead, new examples can be synthesized from the existing examples. This is a type of [data augmentation](https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/) for the minority class and is referred to as the **Synthetic Minority Oversampling Technique**, or **SMOTE** for short.

This procedure can be used to create as many synthetic examples for the minority class as are required.

**LOGISTIC REGRESSION:**

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, **it gives the probabilistic values which lie between 0 and 1**.

Linear Regression is used for solving Regression problems, whereas **Logistic regression is used for solving the classification problems**.

In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).

The curve from the logistic function indicates the likelihood of something such as whether the cells are cancerous or not, a mouse is obese or not based on its weight, etc.

Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.

Logistic Regression can be used to classify the observations using different types of data and can easily determine the most effective variables used for the classification.

Logistic regression is used to calculate the probability of a binary event occurring, and to deal with issues of classification. For example, predicting if an incoming email is spam or not spam, or predicting if a credit card transaction is fraudulent or not fraudulent. In a medical context, logistic regression may be used to predict whether a tumor is benign or malignant. In marketing, it may be used to predict if a given user (or group of users) will buy a certain product or not. An online education company might use logistic regression to predict whether a student will complete their course on time or not.

logistic regression is used to predict the likelihood of all kinds of “yes” or “no” outcomes. By predicting such outcomes, logistic regression helps [**data analysts**](https://careerfoundry.com/en/blog/data-analytics/difference-between-data-scientist-and-data-analyst/) (and the companies they work for) to make informed decisions. In the grand scheme of things, this helps to both minimize the risk of loss and to optimize spending in order to maximize profits.

It is the go-to method for binary classification problems (problems with two class values).

Logistic regression is named for the function used at the core of the method, the logistic function.

The [logistic function](https://en.wikipedia.org/wiki/Logistic_function), also called the sigmoid function was developed by statisticians to describe properties of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment. It’s an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1, but never exactly at those limits.

1 / (1 + e^-value)

Where e is the [base of the natural logarithms](https://en.wikipedia.org/wiki/E_(mathematical_constant)) (Euler’s number or the EXP() function in your spreadsheet) and value is the actual numerical value that you want to transform.

The coefficients (Beta values b) of the logistic regression algorithm must be estimated from your training data. This is done using maximum-likelihood estimation.

[Maximum-likelihood estimation](https://en.wikipedia.org/wiki/Maximum_likelihood) is a common learning algorithm used by a variety of machine learning algorithms, although it does make assumptions about the distribution of your data (more on this when we talk about preparing your data).

The best coefficients would result in a model that would predict a value very close to 1 (e.g. male) for the default class and a value very close to 0 (e.g. female) for the other class. The intuition for maximum-likelihood for logistic regression is that a search procedure seeks values for the coefficients (Beta values) that minimize the error in the probabilities predicted by the model to those in the data (e.g. probability of 1 if the data is the primary class).

We are not going to go into the math of maximum likelihood. It is enough to say that a minimization algorithm is used to optimize the best values for the coefficients for your training data. This is often implemented in practice using efficient numerical optimization algorithm

**RANDOM FOREST CLASSIFIER:**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of *combining multiple classifiers to solve a complex problem and to improve the performance of the model.*

As the name suggests, ***"Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."*** Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

**The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.**

Since the random forest combines multiple trees to predict the class of the dataset, it is possible that some decision trees may predict the correct output, while others may not. But together, all the trees predict the correct output. Therefore, below are two assumptions for a better Random forest classifier:

There should be some actual values in the feature variable of the dataset so that the classifier can predict accurate results rather than a guessed result.

The predictions from each tree must have very low correlations.

**XGBOOST:**

XGBoost is an [ensemble learning](https://courses.analyticsvidhya.com/courses/ensemble-learning-and-ensemble-learning-techniques?utm_source=blog&utm_medium=an-end-to-end-guide-to-understand-the-math-behind-xgboost) method. Sometimes, it may not be sufficient to rely upon the results of just one machine learning model. Ensemble learning offers a systematic solution to combine the predictive power of multiple learners. The resultant is a single model which gives the aggregated output from several models.

**CATBOOST:**

CatBoost is a recently open-sourced machine learning algorithm from Yandex. It can easily integrate with deep learning frameworks like Google’s TensorFlow and Apple’s Core ML. It can work with diverse data types to help solve a wide range of problems that businesses face today. To top it up, it provides best-in-class accuracy.

It is especially powerful in two ways:

* It yields state-of-the-art results without extensive data training typically required by other machine learning methods, and
* Provides powerful out-of-the-box support for the more descriptive data formats that accompany many business problems.

“CatBoost” name comes from two words “**Cat**egory” and “**Boost**ing”.

As discussed, the library works well with multiple **Cat**egories of data, such as audio, text, image including historical data.

“**Boost**” comes from gradient boosting machine learning algorithm as this library is based on gradient boosting library. Gradient boosting is a powerful machine learning algorithm that is widely applied to multiple types of business challenges like fraud detection, recommendation items, forecasting and it performs well also. It can also return very good result with relatively less data, unlike DL models that need to learn from a massive amount of data.

**Experiment:**

Backward Elimination feature selection technique is used to select the features for the machine learning models and entire lot of features have been sent to the deep learning models. But SMOTE method is used to oversample the examples in the minority class and train and test dataset has been generated.

We are comparing the result generated by Liang D & Co [2] with our results as the underlying data set remains the same. FC is combination of Financial Ratio’s and Corporate Governance Indicators and he compared performance of the model with FC and Financial Ratio’s. It proved that FC are little better than Financial Ratio’s alone. For this experiment we are considering FC’s along with SMOTE sampling method and going to evaluate some of the Boosting, Bagging and ANN algorithms. Results of our experiment with comparison with results of Ling D[2] is below, by the way step wise linear regression filter for the features performed well than others for the author.

**CONCLUSION:**

Reference:

1. J. L. Bellovary, D. E. Giacomino, and M. D. Akers, ‘‘A review of bankruptcy prediction studies: 1930 to present,’’ J. Financial Educ., vol. 33, pp. 1–42, Dec. 2007.
2. [Liang, D., Lu, C.-C., Tsai, C.-F., and Shih, G.-A. (2016) Financial Ratios and Corporate Governance Indicators in Bankruptcy Prediction: A Comprehensive Study. European Journal of Operational Research, vol. 252, no. 2, pp. 561-572.](https://isslab.csie.ncu.edu.tw/download/publications/1.pdf)

# 3) USE OF FINANCIAL RATIOS TO MEASURE THE QUALITY OF EARNINGS

[Gullett, Nell S](https://www.proquest.com/indexinglinkhandler/sng/au/Gullett,+Nell+S/$N;jsessionid=33D95995239DD71173E40EA37D5426BC.i-03bc6ea303836550a); [Kilgore, Ronald W](https://www.proquest.com/indexinglinkhandler/sng/au/Kilgore,+Ronald+W/$N;jsessionid=33D95995239DD71173E40EA37D5426BC.i-03bc6ea303836550a); [Geddie, Mary F](https://www.proquest.com/indexinglinkhandler/sng/au/Geddie,+Mary+F/$N;jsessionid=33D95995239DD71173E40EA37D5426BC.i-03bc6ea303836550a" \o "Click to search for more items by this author). [**Academy of Accounting and Financial Studies Journal**](https://www.proquest.com/pubidlinkhandler/sng/pubtitle/Academy+of+Accounting+and+Financial+Studies+Journal/$N/29414/OpenView/2124417228/$B/75040B6CC80246EDPQ/1;jsessionid=33D95995239DD71173E40EA37D5426BC.i-03bc6ea303836550a)**; Arden**[Vol. 22, Iss. 2,](https://www.proquest.com/indexingvolumeissuelinkhandler/29414/Academy+of+Accounting+and+Financial+Studies+Journal/02018Y04Y01$232018$3b++Vol.+22+$282$29/22/2;jsessionid=33D95995239DD71173E40EA37D5426BC.i-03bc6ea303836550a) (2018): 1-12

5) Ferris, Erin. (2018). Dividend Taxes and Stock Volatility. International Tax and Public Finance, 25(2), 377–403. <https://doi.org/10.1007/s10797-017-9455-2>

4) E. I. Altman, ‘‘Financial ratios, discriminant analysis and the prediction of corporate bankruptcy,’’ J. Finance, vol. 23, no. 4, pp. 589–609, Sep. 1968

J. A. Ohlson, ‘‘Financial ratios and the probabilistic prediction of bankruptcy,’’ J. Accounting Res., vol. 18, no. 1, pp. 109–131, Apr. 1980. [32] J. Brozyna, G. Mentel, and T. Pisula, ‘‘Statistical methods of the bankruptcy prediction in the logistics sector in Poland and Slovakia,’’ Transformations Bus. Econ., vol. 15, no. 1, pp. 80–96, 2016.

P. P. M. Pompe and A. J. Feelders, ‘‘Using machine learning, neural networks, and statistics to predict corporate bankruptcy,’’ Comput.-Aided Civil Infrastruct. Eng., vol. 12, no. 4, pp. 267–276, Jul. 1997. [35]

J. Min and Y. Lee, ‘‘Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters,’’ Expert Syst. Appl., vol. 28, no. 4, pp. 603–614, May 2005.

N. Ghatasheh, H. Faris, R. Abukhurma, P. A. Castillo, N. Al-Madi, A. M. Mora, A. M. Al-Zoubi, and A. Hassanat, ‘‘Cost-sensitive ensemble methods for bankruptcy prediction in a highly imbalanced data distribution: A real case from the Spanish market,’’ Prog. Artif. Intell., vol. 9, no. 4, pp. 361–375, Dec. 2020.

Comparing the Performance of Deep Learning Methods to Predict Companies’ Financial Failure H. ALJAWAZNEH 1 , A. M. MORA 2, P. GARCÍA-SÁNCHEZ 1 AND P. A. CASTILLO-VALDIVIESO 1

6) Corporate Bankruptcy Prediction using Machine Learning Techniques, Shantanu Deshpande, Dr. Vladimir Milosavljevic, [National College of Ireland](http://norma.ncirl.ie/4312/1/shantanudeshpande.pdf)