Advances in Consumer Research
ISSN (Print): 0098-9258
ISSN(Online): 3079-1766

Vol. 2, Issue 2 (2025) https://acr-journal.com/

The Impact of Financial Data Analytics on Business Strategy

Rajesh Deb Barman[1]

[1]Assistant Professor, Department of Commerce, Bodoland University, Kokrajhar, Assam, India
Email ID: rajesh.barman19@gmail.com

Cite this paper as: Rajesh Deb Barman, (2025) The Impact of Financial Data Analytics on Business Strategy. *Advances in Consumer Research*, 2 (2), 917-922.

KEYWORDS

Financial Data Analytics, Business Strategy, Deep Learning, Hierarchy Attention Network

ABSTRACT

Financial data examination is currently a key instrument that impacts vital dynamic in the cutting-edge organization climate. This study inspects the different ways that monetary information investigation influences corporate procedure, seeing how organizations use information driven bits of knowledge to work on financial execution, smooth out processes, and get an upper hand. In this day of computerized innovation, misrepresentation is turning out to be considerably more refined and far reaching. It influences a wide range of regions, including Visa exchanges, protection cases, and then some, and it causes enormous monetary misfortunes. To distinguish false movement, monetary foundations and associations have utilized various models in view of various methods, like Deep learning, Machine Learning, and information mining. Organizations should have the right construction set up to foster their monetary show, which is additionally upheld by progressions in development that affect organizations.

.

1. INTRODUCTION

Data analytics for company is now crucial for making educated decisions and attaining long-term success in today's rapidly evolving corporate environment. The rise of digital technology and networked systems has resulted in an unprecedented amount of data for organizations, which presents both a minefield of opportunities and a labyrinth of challenges. The key to navigating this data-rich environment is the application of data analytics, a transformative tool that helps businesses to extract pertinent insights, streamline processes, and advance strategic objectives. Big Data and modern logical instruments have changed numerous parts of corporate tasks, however monetary information investigation stands apart as being particularly significant for deciding corporate system. The purposeful assessment of monetary information to determine noteworthy experiences is known as monetary information examination, and it assists organizations with settling on choices that are in accordance with their essential objectives (Zhu and Yang, 2021). Having the option to utilize and dissect monetary information has become urgent for long haul achievement and remaining serious in a period of quick mechanical forward leaps and growing business sector unpredictability. The practice of examining raw data to uncover hidden trends, correlations, and outlines is known as data analytics for business. It offers a lens through which companies can assess their past performance, forecast new trends, and suggest a smart move to gain a competitive edge. From identifying consumer preferences and market trends to streamlining procedures and boosting productivity, data analytics is being used in a variety of industries to rethink conventional business models.

.

.

2. LITERATURE REVIEW

Big Data to Enhance Financial Performance

The creation of novel technologies and architectures, by way of well as the economical extraction of value from vast amounts of data through high velocity, discovery, and/or analysis, are the definitions given in the literature Mikalef et al., (2018). Over time, BDAs have drawn the interest of policymakers as a way to guide corporate decision-making D'Angeac, (2012). A significant number of businesses have expedited their BDA initiatives in an effort to obtain a crucial understanding of how to achieve upper hand. Specialists and researchers have sorted BDAs as a potential outskirt for development, rivalry, and production (Manyika, 2011). But for others, it was a revolution that would change how people thought, work, and live. A lot of work was done to store, analyze, and show the data in light of the growing the amount, speed, and variety of data. Unfortunately, empirical evidence supporting the influence of BDAs on profitability over the long term in the banking sector and the process by which they assist banks gain an edge over their competitors is still lacking (Wamba et al., 2017). This is astonishing considering the account that corporates are increasingly entering the BDA space (Mikalef et al., 2020).

The rise of BDA changed the way that corporations produced and operated. To become competitive in the market and ensure their existence, many big businesses want to create SCs through BDAs Jun, (2019). By using advanced analytical skills to extract high-quality information from BD, BDAs assist businesses attain sustainable performance and increased operational efficiency. BDAs sincerely dedicate themselves to the organization, guaranteeing superior performance. In comparison to their less driven counterparts, companies with greater analytical drive expand three times faster, rendering to the Mckinsey and Company Global Banking Report (2018). When it comes to using BDAs, the banking sector was one of the first, followed by the pharmaceutical, insurance, energy, industrial, and agricultural sectors. BDAs, however, cannot be completely incorporated into the banking sector's decision-making procedures, culture, and business processes.

Business analytics and the performance of companies

"A process of converting data into actions via evaluation and conclusions in the context of managerial decision-making and problem-solving" is how the definition of business analytics (Liberatore & Luo, 2008). Utilizing data, statistical analysis, quantitative approaches, technical advancements, and computer- or mathematical-based models to help managers make more informed decisions and gain a deeper understanding of their business operations is what Evans (2012) characterizes as business analytics.
For others, the field of study known as "business analytics" focuses on creating new insights, systems viewpoints, and comprehensive understandings of a company's business environment (Ram & Delen, 2018). According to Delen and Ram (2018), it is a unique area of analytics that uses its tools, methods, and concepts to create answers for challenging business issues. In order to help people make timely, wise decisions that will help them survive, flourish, innovate, and expand, the business analytics academic stream aims to provide information on best practices. Descriptive, diagnosis, discovery, forecasting, and prescriptive analytics are all areas of focus for business analytics to encourage decision-making based on data and produce better intermediate results and financial performance.

3. METHODOLOGY

With an accentuation on what information driven methods mean for vital preparation, execution the board, monetary extortion location, and advancement, this study investigates the connection between monetary information examination and company methodology. Through the joining of hypothetical structures and genuine models, the study tries to offer an exhaustive understanding of the essential advantages that monetary information investigation offers.

To preserve pertinent information and facilitate algorithmic analysis, the unstructured financial data which was in text must be transformed into a numeric format, whereas the financial variables represent organized tabular data and don't need any significant preprocessing. Consequently, the DL elements of the suggested HAN model are covered in the ensuing subsections.

Deep Learning (DL)

Predictive modelling can be done with unstructured text data once it has been numerically represented. Traditional text classification tactics offer sparse lexical characteristics, such as TF-IDF, and then apply linear modelling or kernel techniques to this representation. New methods for learning textual representations have recently been added to DL, such as Convolutional Neural Networks (CNN) Kalchbrenner et al., (2014) and Recurrent Neural Networks (RNN) Yin et al., (2017). Because the RNN architecture preserves the input sequence, it is frequently utilized in video processing, language production, and natural language understanding Kalchbrenner and Blunsom, (2013). Long-term addictions can be acquired by the model thanks to an LSTM, a unique kind of RNN made up of different gates that decide whether information is remembered, updated, or retained. Important information is stored in a distinct cell that serves as a memory, and an LSTM selectively maintains or alters prior knowledge Tixier, (2018). Accordingly, significant information might continue for a drawn-out period without being supplanted by the new data sources.

Hierarchy Attention Network (HAN) More refined Semantic ordered progressions, like the associations between words,

sentences, and records, are additionally adopted into thought by DL strategies. Approaches to hierarchical document construction have been discussed. Rao and associates (2018). HAN was created to handle the relatively new idea of word and phrase contexts, where a word or sentence's meaning may vary based on the document Yang et al., (2016). HAN initially identifies the important words in an expression before assessing whether sentences, while considering context, are significant in a document (see Figure 1) in order to compute the document encoding. The concept acknowledges that a word's occurrence in one sentence may be meaningful, whereas another word's presence in another sentence may not be. The HAN creates a document representation by first building sentence vectors from words, then using the attention mechanism to aggregate these sentence grids into a record portrayal. An attention mechanism that determines important weights and an encoder that produces pertinent contexts make up the model. The same methods are used, first at the word level and afterward at the sentence level.
Word Count. Structured tokens W_{it} representing the word "i" in a phrase $t \in [1, T]$ are created from the input. In addition, each token is allocated to a pre-trained anchoring matrix W_e multivariate vectors $X_{it} = W_e W_{it}$. Words are therefore represented numerically byX_{it}, which is a representation of the term in an uninterrupted vector interstellar. Encoder for words. The vectorized tokens address the contributions of the ensuing layer. In spite of the fact that Yang et al., (2016) involved GRU for encoded information, we pick LSTM since it performed better on the enormous text groupings that are given. To get the word comments in the ongoing methodology, a bidirectional LSTM is presented. The word embedded matrix is the only parameter that differs between the two unidirectional LSTMs that make up the model. Sentences undergo processing from right in the initial forward LSTM, they undergo processing from left to right, while in the retrograde LSTM, they are treated from left to left. At each time step t, the pair of phrase embedded data is concatenated to produce an inner version of the bi-directional LSTM h_{it}.

The HAN is applied subsequent to Kränkel and Lee's application Kränkel and Lee, (2019). Both linguistic and quantitative variables are used in the data used for training set to train the DL model. The financial ratios are thus concatenated with the textual information gathered in the part before this one. Using stratified random sampling, the model predicts the probability of fraud in the relevant validation and assessment partitions. The HAN-based fraud detection model's framework, which includes the output parameters for each layer are displayed in Figure 2. 150 neurons make up the LSTM layer, which also has a final dense layer dimensions of 6 and a HAN dense dimension of 200. 300 dimensions are available for the sentence and word annotation in this instance because to include a mix of both backward and forward LSTMs. In the last layer of the HAN, dropout regularization is used to prevent over-fitting. Before passing through the soft max function that generates the fraud probability, the final step involves concatenating the generated 200-size document representation with 47 ratios of finances and putting it into a dense layer. A number of batches of 32 and 17 epochs were employed for training, following hyperparameter adjustment on the train validating set.

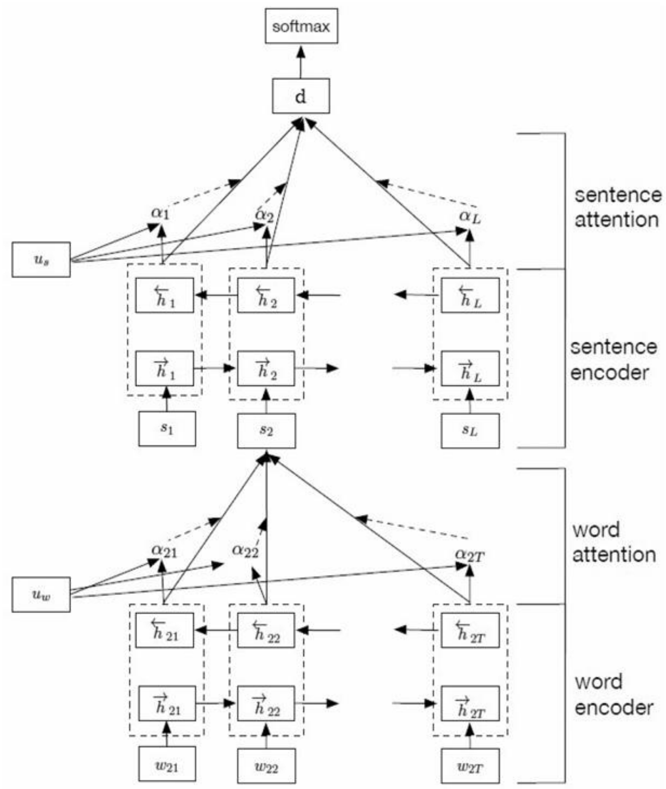

Figure 1 Construction of HAN; adapted from Yang et al., (2016)

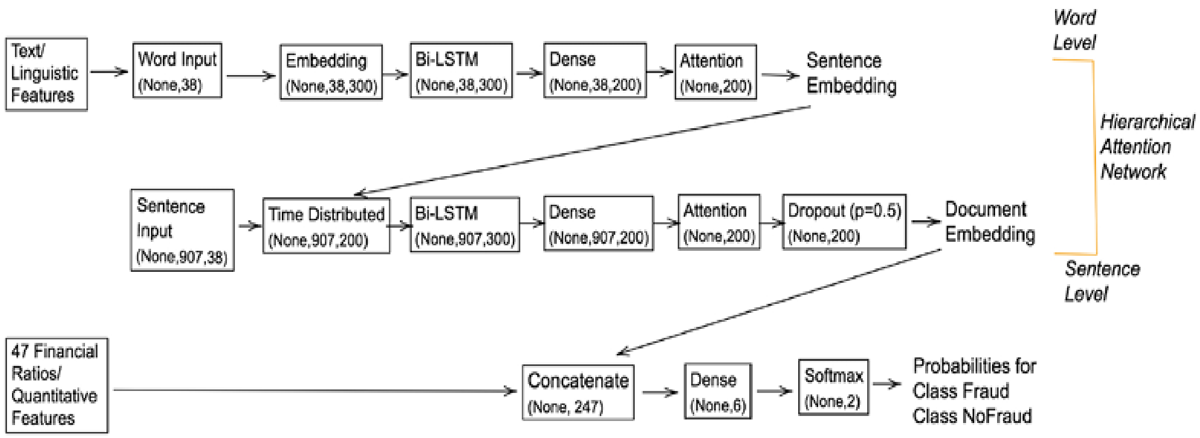

Figure 2 Structure of the Fraud Detection Model Based on HAN; adapted from Craja et al., (2020)

4. RESULTS AND DISCUSSION

Four potential outcomes are considered while tackling the binary classification problem of financial statement fraud detection: True negative (FN) indicates that a fraud case was incorrectly classified as non-fraud, true positive (TP) indicates that a False negatives (FN) show that a fraud case was mistakenly categorized as non-fraud, false positives (FP) show that a non-fraud case was mistakenly labeled as fraud, and fraud cases were correctly classified as non-fraud. F1-score, accuracy, sensitivity (sometimes called recall or TP rate), specificity (also called TN rate), and precision were all taken into consideration in numerous earlier research to quantify predictive performance, West and Bhattacharya, (2016). Accuracy, F1-score, F2-score, sensitivity, specificity, and AUC are among the metrics used in this study to assess model performance.
A combination of accuracy (the proportion of cases correctly categorized as fake) and sensitivity (the number of fraudulent occurrences the classifier failures) yields the F-score. It gauges the accuracy and resilience of the models' classification of fraudulent cases:

$$F_\beta - score = (1 - \beta^2) \times \frac{precision \times sensitivity}{(\beta^2 \times precision) + sensitivity}$$

In light of this, The F2-score (harmonic average of precision by sensitivity) is used in this study as a supplement to the F1-score, which is better suited for fraud detection because it gives sensitivity a larger weight than precision. The ability of a model to correctly rank fraud vs non-fraud cases is gauged by the AUC. The model's capacity to distinguish between situations involving fraud and those that are not improves with an AUC. The AUC is employed in this investigation and is recommended over accuracy in fraud detection since it is resilient to unequal class distributions, Purda and Skillicorn, (2015). In order to calculate F1- and F2-scores, model-based fraud probability must have a cutoff. The threshold that optimizes the difference between FP rate and sensitivity is chosen, and the classification results are assessed using it. The ideal threshold for the HAN model is 0.05, meaning that an assertion is deemed dishonest if its fraud likelihood exceeds 5%.

The effectiveness of a number of classification methods in detecting fraud was compared. In order to classify fraud, the algorithms employ financial indications of feature groups. Comparable outcomes from the out-of-sample set of tests are shown in Table 1. The test set's baseline accuracy in identifying every case as non-fraudulent (the vast mainstream class) is provided below in Fig 3.

Table 1 Comparable outcomes from Financial Data Analytics

Model	F1 Score	F2 Score	Sensitivity	Specificity	Accuracy	AUC
LR	0.4878	0.759	0.6944	0.7654	0.8363	0.773
SVM	0.4736	0.7606	0.6277	0.793	0.839	0.7672
ANN	0.4673	0.6946	0.7944	0.6685	0.6891	0.7675
RF	0.5609	0.7903	0.7777	0.799	0.8764	0.87
XGB	0.594	0.8402	0.677	0.882	0.8592	0.858

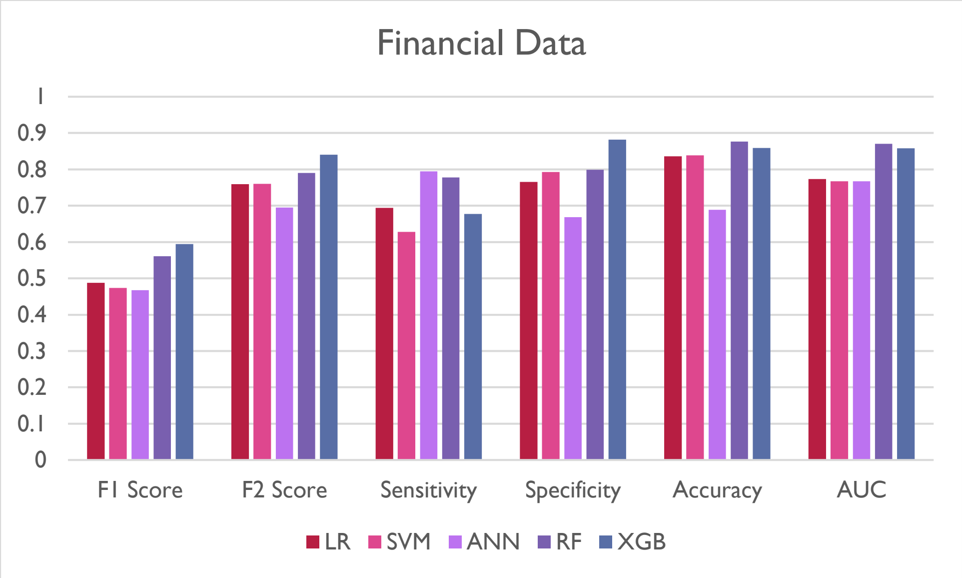

Figure 3 Comparable outcomes from Financial Data Analytics

Using financial information (FIN) for modelling has been the greatest frequently used method. The method acts as a standard for this investigation. There appears to be a non-linear relationship between financial metrics and a report's fraud state, as seen by the high accuracy and AUC of the models based on trees RF and XGB in forecasting fraud on FIN. This outcome is consistent with the idea that RF did particularly well when dealing with high-dimensional fraudulent financial data. Given the error cost imbalance, additional performance metrics, including F2-score, were used to address the classification models' practical usefulness. Together with the model's sensitivity, the higher predictive performance should be taken into consideration to version for the consequences of failing to identify the deceitful case.

5. CONCLUSION

Investigation of financial information impacts business procedure since it gives organizations the information and assets, they need to arrange multifaceted and changing commercial centres effectively. Financial examination assists firms with making long haul progress and keep an upper hand by supporting monetary execution, advancing development, upgrading risk the executives, and working on essential preparation. A critical part of hierarchical execution will keep on being the strategic use of monetary data investigation as the business biological system changes. That's what the discoveries demonstrate, in contrast with the benchmark models, the DL model altogether worked on its AUC. The outcomes show that while most ML models perform better at precisely recognizing genuine affirmations, they are less powerful at distinguishing fake cases. Conversely, the DL model is especially fit to recognize false circumstances accurately

REFERENCES

[1] Craja, P., Kim, A., & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. Decision Support Systems, 139, 113421.

[2] D'Angeac, G. D. (2012). Big data: The management revolution-Harvard business review. Harv. Bus. Rev., 90, 60-66.

[3] Evans, J. R. (2012). Business analytics. Pearson Higher Ed.

[4] Jun, S. (2019). Bayesian structural time series and regression modeling for sustainable technology management. Sustainability, 11(18), 4945.

[5] Kalchbrenner, N., & Blunsom, P. (2013, October). Recurrent continuous translation models. In Proceedings of the 2013 conference on empirical methods in natural language processing (pp. 1700-1709).

[6] Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A convolutional neural network for modelling sentences. arXiv preprint arXiv:1404.2188.

[7] Kränkel, M., & Lee, H. E. L. (2019). Text Classification with Hierarchical Attention Networks. In Seminar Information Systems, HU-Berlin, February (Vol. 8).

[8] Liberatore, M. J., & Luo, W. (2010). The analytics movement: Implications for operations research. Interfaces, 40(4), 313-324.

[9] Manyika, J. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 1.

[10] Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. Information & Management, 57(2), 103169.

[11] Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. Information systems and e-business management, 16, 547-578.

[12] Purda, L., & Skillicorn, D. (2015). Accounting variables, deception, and a bag of words: Assessing the tools of fraud detection. Contemporary Accounting Research, 32(3), 1193-1223.

[13] Ram, S., & Delen, D. (2018). Introduction to the inaugural issue of journal of business analytics. Journal of Business Analytics, 1(1), 1-1.

[14] Rao, G., Huang, W., Feng, Z., & Cong, Q. (2018). LSTM with sentence representations for document-level sentiment classification. Neurocomputing, 308, 49-57.

[15] Tixier, A. J. P. (2018). Notes on deep learning for nlp. arXiv preprint arXiv:1808.09772.

[16] Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of business research, 70, 356-365.

[17] West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: a comprehensive review. Computers & security, 57, 47-66.

[18] Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016, June). Hierarchical attention networks for document classification. In Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies (pp. 1480-1489).

[19] Yin, W., Kann, K., Yu, M., & Schütze, H. (2017). Comparative study of CNN and RNN for natural language processing. arXiv preprint arXiv:1702.01923.

[20] Zhu, X., & Yang, Y. (2021). Big data analytics for improving financial performance and sustainability. Journal of Systems Science and Information, 9(2), 175-191.

❖❖❖❖❖