

DECISION-MAKING IN LARGE CORPORATIONS- ROLE OF BIG DATA ANALYTICS & DATA MINING

Christophar Nicholas Hendstein¹, Hiroshi Akeera Katsu^{2,*}

¹*Tel Aviv University, Tel Aviv, Israel*

²*University of Tokyo, 7 Chome-3-1 Hongo, Bunkyo City, Tokyo, Japan, hiroshik@asia.com*

Abstract

Function - The goal of this particular paper is presenting a novel framework for strategic decision making utilizing Big Data Analytics methodology.

Design/methodology/approach - In this particular research, 2 distinct machine learning algorithms, Random Forest as well as Artificial Neural Networks are used to forecast export volumes working with a considerable level of open industry information. The forecasted values are in the Boston Consulting Group Matrix to conduct strategic industry analysis.

Results - The proposed technique is validated utilizing a hypothetical case study of a Chinese business exporting freezers and refrigerators. The results indicate the proposed methodology makes exact trade forecasts and helps to conduct strategic industry evaluation properly. Furthermore, the RF performs much better compared to the ANN in terminology of forecast accuracy.

Investigate limitations/implications - This analysis provides just one case study to evaluate the proposed methodology. In future scientific studies, the validity of the suggested technique is further generalized in various item groups and nations.

Functional implications - In present day extremely competitive business environment, a good strategic industry evaluation involves exporters or importers making much better predictions along with strategic choices. To us the proposed BDA based strategy, businesses may efficiently determine business opportunities and alter their strategic choices appropriately.

Originality/value - This's the very first study to provide a holistic methodology for strategic industry evaluation using BDA. The proposed methodology effectively forecasts global trade volumes and helps with the strategic decision making practice through succeeding insights into worldwide marketplaces.

Keywords

Big data analytics, Strategic decision making, Machine learning

JEL Classification

M43, M45

DOI: <https://doi.org/10.14311/bit.2022.01.17>

Editorial information: journal Business & IT, ISSN 2570-7434, CreativeCommons license published by CTU in Prague, 2022, <http://bit.fsv.cvut.cz/>



Introduction

Today's competitive business environment calls for businesses making much better predictions for the complex business environments of theirs and respond to changing conditions immediately. To cope with this particular complexity, businesses started to incorporate a lot more data sources in the decision-making processes of theirs [2]. Choosing the best open/public information resources, making frequent posts, and integrating them along with the inner information are believed to be as several of the obstacles for companies. Firms that conquer these major information obstacles and smooth the decision-making processes of theirs will get a competitive edge in the market of theirs.

Even though discussions concerning the meaning of big data still remain, the primary attributes of big data are identified by 5 Vs: Volume, Variety, Velocity, Value and Veracity. Making use of the right data as well as the ability to derive significant results have grown to be essential for business decision-makers. The analysis of the massive volumes of information is known as Big Data Analytics. For attaining true benefit from BDA, an assortment of disciplines and techniques, which includes data, data mining, machine learning, interpersonal network analysis, signal processing, pattern recognition, techniques as well as visualization methods, could be utilized. Higher functional efficiency, much better strategic decision making, better visibility [15].

A BDA methodology for strategic decisions enhancement in customer care, and much better brand new items would be the most often agreed lots of data opportunities. Even though BDA is able to help reducing expenses, being much more nimble and attaining higher service levels, under twenty % of businesses adapted BDA within their supply chains because of individuals, culture or maybe process related obstacles [3]. With all the improved accessibility of broad data energy sources associated with global trade, importers and exporters are able to look at the improvement of the industry volumes between countries and make their strategic choices appropriately [1]. Nevertheless, they often try to arrive at some statistics about a certain market country, product and period, and consequently, obtain extremely minimal info about market behavior. With all the assistance of a holistic approach, they might constantly assess the marketplace situation from various perspectives, and that is the time when they would use this industry assessment in the decision-making process of theirs by concurrently considering the industry, product as well as time related factors. These 3 elements could be coupled with the available broad data solutions using BDA to enjoy a clearer vision concerning the particular product and country pairs that they ought to concentrate on down the road [14].

The primary goal of this particular analysis is developing a holistic methodology for strategic industry analysis. From our proposed methodology, a variety of international trade data are believed to be by using managerial implications and BDA are devised for a business. The Boston Consulting Group Matrix is used as the strategic management application in this research. The study of ours additionally offers a nonparametric forecasting technique using machine learning algorithms with a considerable length of information [4]. Note that ML algorithms have grown to be important parts of real life BDA applications. In this particular research, 2 distinct ML algorithms, Random Forest as well as Artificial Neural Networks are already put on to forecast the export volumes [6]. The prior scientific studies in the literature just created forecasting designs, though we used forecasting as an input to our recommended strategic industry style and then created a holistic Big Data Analytics framework for strategic industry analysis. Additionally, our proposed BDA strategy employs much more variety of information as well as machine learning options to forecast export volumes compared to the other types in the literature. In order to exhibit the usefulness of the methodology of ours, a hypothetical case study of a Chinese business exporting freezers and refrigerators to the United Kingdom is presented. This particular item class is among the key export solutions of China, and the United Kingdom is 1 of the key importers. In order to validate the effectiveness of the methodology of ours more thoroughly, three distinct sub-products are tested, as well as the results are discussed [5].

The majority of this particular paper is structured as follows. Area two summarizes the literature on the forecasting of related studies and trade volumes. In Section three, the research goals as well as the problem definition are provided in detail. The recommended alternative BDA framework, which often combines the strategic industry evaluation style and forecasting global trade, is offered with Section four. Area five shows the application of the suggested methodology and reports the outcomes of the case study of detail. Future work and conclusion are summarized in Section six.

Literature Review

Demand management is among the primary key tasks of Supply Chain Management and BDA allows businesses to comprehend their market requires higher. A good knowledge of the market demands aids designing supply chains to react quicker and better to changing customer as well as supplier must have. Based on the most recent literature review scientific studies, the quantity of publications on BDA has grown considerably for the final 5 years in parallel with the increased supply chain management studies. Demand management by utilizing BDA blankets twenty five % of all BDA in the SCM studies. These studies mostly focus on forecasting, predicting and synchronizing customer demands.

Based on Hu as well as Akter and Sazu et al. (2022), BDA could be put on to different interdisciplinary fields. For instance, BDA is successfully utilized in healthcare data analysis, economic industry prediction, and customer churn prediction. Besides the accessibility of information, the development of computational power as well as ease of use of BDA solutions have helped businesses to automate the decision-making process of theirs using Artificial Intelligence strategies. Thus, the investments running a business Intelligence technologies greatly enhance the online business functionality of businesses through predictive and prescriptive analysis using BDA [7]. Nevertheless, present scientific studies revealed that these investments might boost business efficiency just in case they're integrated into the decision making processes. Consequently, in this particular study, we concentrate on utilizing BDA for forecasting in global trade and conducting strategic industry analysis as forecasting is among the primary key areas that BDA is able to assist companies to make greater decisions [9].

Within the previous years, different forecasting designs have been used for trade forecasting that utilizes extrapolation, economic models and time-series, agent based computational economics versions, and machine learning algorithms. 2 mainstream research approaches, nonparametric and parametric approaches, have been utilized in trade forecasting versions [8]. AutoRegressive Integrated Moving their, Holt-Winter, and Average variants would be the most often used parametric time series forecasting strategies in trade amount forecasting. Bailey and Dale utilized the Box Jenkins way to forecast US merchandise exports. Sazu & Jahan (2022) believed export demand elasticities in developing and manufacturing places by time series methods [6]. Sazu and Jahan (2022) forecasted seaborne trade passes of 4 various commodity solutions by utilizing Vector Auto Regression [7]. Sazu investigated seasonality in Pakistan's merchandise exports as well as imports by making use of a univariate ARIMA design [8]. Jahan & Sazu (2022) used ARIMA to calculate spice import as well as export volumes as well as output actions for China and India [9]. Khan used ARIMA, Var and Holt-Winter strategies to forecast Bangladesh's entire imports. Likewise, Sazu (2022) used ARIMA, VAR, Engle Granger Single Equation and Vector Error Correction models to forecast South Africa's farming exports imports. Jahan et al. (2022) constructed univariate time series models to forecast the export need for molding as well as chipboard volume for Peninsular Malaysia [13].

The study of ours plays a role in the literature by presenting a holistic Big Data Analytics framework for strategic industry analysis. The prior scientific studies in the literature just centered on creating forecasting strategies to foresee global trade volume; however, the study of ours will be the very first to utilize BDA for strategic industry analysis. Additionally, our BDA based forecasting unit varies from the prior nonparametric forecasting research since it engages additional quantity and variety of

information sources. Additionally, our proposed strategy extends the BCG matrix by including the predictions of future market and trends scenarios to assist businesses to generate much better strategic industry choices [10].

Problem definition

For evaluation of the future and current business situations in global trade, 3 unique aspects are considered. The very first part is definitely the representation of the marketplace. An individual country, a geographical region or maybe a team of nations with typical perspectives, could be viewed as being a marketplace. The next one will be the prediction horizon, which can differ from monthly to a few years. The 3rd part will be the product granularity. Item granularity is handled based on the Harmonized Coding System and commodity Description levels

The study of ours addresses 3 important study questions: How could we utilize available information energy sources in strategic industry analysis for exporters/ importers?

Which strategic management models as well as open details can be utilized together for strategic industry analysis?

How could we produce extra value by utilizing Big Data Analytics in the selected strategic industry evaluation technique?

These're the primary investigation objectives that we try to resolve using our proposed BDA framework. For testing the methodology of ours, we show a hypothetical case study where a Chinese business exports refrigerators as well as freezers all around the planet, in which one of the main markets of its will be the United Kingdom. We just have to believe a yearly export worth for the hypothetical situation provider to calculate the relative market share of its in the BCG matrix for demonstration purposes. Without loss of generality, a arbitrary export worth could be selected, as we demonstrate in Section five. Thus, the proposed technique could be quickly put on to a genuine business so long as the company 's annual export information might be gotten.

Methodology

Our suggested Big Data Analytics strategy is grounded on CRISP DM. By utilizing the actions of CRISP DM as phases of the framework of ours, we included the necessary steps to produce a holistic methodology for industry analysis using BDA. The proposed methodology of ours is done in 2 levels that are different in parallel: Strategic Market Analysis, and Forecasting International Trade. These two levels and the steps of theirs are describe in detail in the next sections.

Strategic industry evaluation model

Exporters, importers as well as logistics service providers are believed to be as the likely users of the methodology of ours. These businesses are exporting, facilitating or importing motions of items from one country to the next. Their portfolios include product as well as country pairs. Profitable profile management is crucial to be one step forward in modern competitive business environment. International bilateral trade information between lands are publicly accessible readily available by the companies [11]. Nevertheless, the primary key concern for the companies is finding a correct way of utilizing these wide open international trade data in the decision-making processes of theirs. International trade data are needed for the case study of ours in 2 ways. For starters, trade information provide us the import sector size of a nation for a certain product in a particular period. Thus, the industry growth rate could be estimated using the information. Next, by mixing the market size with all the income of the business, a business is able to derive the relative market share of its.

Forecasting global trade by utilizing big data analytics

To effectively forecast international trade, potential factors and the dynamics affecting bilateral trade between nations must be identified initially. This's the business understanding phase of the industry forecasting. These factors could be grouped into two primary groups: Product specific trade info, and Country or maybe global conditions connected factors [13]. The primary elements impacting bilateral trade would be the supply and demand elements of the similar nations.

Random forest

Random forest is a machine learning technique, and that is popular for each category as well as regression. The RF programs are reported to yield effective outcomes for little datasets with a somewhat big number of characteristics. Breiman developed the ensemble learning strategy by using various decision trees [16]. Prediction by making use of a single choice tree is able to result in an overfitting issue in the coaching set, that produces low quality success in the evaluation set. Thus, the RF algorithm can be used not only in forecasting but additionally in the feature selection procedures. In the arbitrary forest algorithm, the very best descriptive mixture is selected out of a lot of trees produced throughout the algorithm. The basic practice of RF algorithms is defined [12].

Artificial neural networks

Synthetic Neural Networks are based upon the concept to mimic biological neural methods by connected and simple processing nodes. ANN has been used in different regions for the final 2 years, financial analysis, credit scoring, including forecasting, and fraud evaluation. ANN incorporates computational components which are created to imitate the natural central nervous system. ANN is dependent on the accumulation of knowledge throughout training sessions. Both ANN and RF algorithms rely on a variety of parameters [14].

Results

As describe in detail in Section three, our hypothetical case study organization is assumed to export 3 kinds of items from China to a number of nations. In order to show the methodology of ours, just the export volumes to the United Kingdom are forecasted. The information set applied to the forecasting design is put together from various open data energy sources which are provided. The resulting information set is from 2006 April to 2018 March with 144 month data points.

Just before using the tuning procedure, 3 preliminary analyses are performed testing the consequences of the input month combinations, include choice and reliant adjustable transformation choices on the prediction quality. The very first examination is deciding on which month or month mixtures of the past trade information must be taken as input elements in the forecasting version. The past trade information of 3 months, 3 as well as 4 months, and 3 with 5 months are analyzed. Next, the portion of selected characteristics is tested with fifty %, 100 %, 75 %, in which fifty % implies that just half of the characteristics are in the model. The 3rd you're determining the use of the reliant variable. In this particular analysis, 3 choices are searched: exact same, logarithm transformation as well as square root transformation. Based on the outcomes of the preliminary analyses, the very best combinations are discovered [15].

You will find twenty eight features in both ANN and RF models. Since each characteristic in the data set of ours has various ranges, the information of each characteristic are normalized based on min max normalization. Following normalization, information are split into training as well as evaluation sets, the place that the instruction set is eighty % of the whole information set with 115 observations, as well as the test set will be the remaining twenty % with twenty eight observations. Of the split procedure, information stratification is used to make certain the information of every month to be

incorporated in the train as well as test sets. Thus, the accumulation of particular weeks in the test or exercise sets is avoided. Next, the hyperparameters are tuned for equally algorithms based on the evolved process summarized [16].

Both ML techniques have been applied through the use of "scikit learn" open source libraries of Python on a Windows PC. Additionally, Python libraries of "MLPRegressor" as well as "RandomForestRegressor" had been employed for RF and ANN implementations, respectively.

The RF algorithm yields higher R2 values compared to the ANN algorithm for most product types. Among 3 merchandise types, the largest R2 great is noticed in Product 84.18.50, and the lowest R2 value is observed in Product 84.18.40. Nevertheless, for Product 84.18.40, the median R2 great is elevated to 0.531 when RF is used. Like Product 84.18.40, the ANN design results in a scaled-down R2 importance for 84.18.10. Plus, probably the highest median R2 worth of the ANN design is 0.802 for merchandise 84.18.50, but the RF design accomplishes 0.921 median R2 worth for that service. Thus, based on these results, the RF design accomplishes higher R2 values and works much better compared to the ANN design [18].

Just how can we utilize open data energy sources in strategic industry analysis for exporters/importers? Which strategic management models as well as open details can be utilized together for strategic industry analysis? Just how can we produce extra value by utilizing Big Data Analytics in the selected strategic industry evaluation technique?

Discussion

The model of ours uses the popular BCG matrix, that is among the most widely used portfolio management methods in training as well as the literature. The proposed methodology effectively forecasts industry volumes between countries and helps to determine future market and trends scenarios. Therefore, the case study demonstrates the proposed design helps you to conduct strategic industry evaluation properly.

From the study of ours, the primary elements affecting trade volume are labeled as each global and product specific elements. Additionally, in accordance with the outcomes of the nonparametric forecasting literature, both RF and ANN yielded forecasts that are accurate in the case study of ours with 3 products [19]. Nevertheless, RF performed much better compared to ANN in terminology of forecast accuracy in addition to robustness of the outcomes. It summarizes the outcomes of the research and discusses the findings relating to the research questions of ours.

Conclusion and investigation implications

Today's increased worldwide competition forces businesses to make much better predictions as well as strategic decisions considering the business environments of theirs. Boston Consulting Group Matrix is among the most known management aids that revolutionized strategic control. Nevertheless, it's some issues in training, like the difficulty of its in calculating market share as well as market growth rate. Besides, it shows just the present business environment and doesn't create some insight into the world. By correctly forecasting the long term business environment, several future insight info could be put into the BCG Matrix [20]. Consequently, businesses are able to determine brand new business and trade opportunities from the forecasts of the industry volumes between nations. Nevertheless, correct forecasting is now drastically more challenging as a result of the improved intricacy of the globalization as well as competition between nations and supply chains. Using BDA, forecasts that are accurate may be achieved and these forecasts could be utilized in strategic industry analysis. In this particular research, we suggest a holistic methodology for strategic decision making using BDA.

The case study of ours demonstrates the likely customers can readily adjust the proposed methodology to arrange the strategic market decisions of theirs based on the future and current industry trends. The results show the RF design yields much more precise forecasts than the ANN design. Nevertheless, each ANN and RF models offer effective forecasts for trade volumes. Identifying trivial characteristics, transforming the dependent adjustable with logarithm and including previous trade volume info contribute significantly to the correct forecasts. Rather than utilizing all features obtained from the information sources, using fifty % or maybe seventy five % of the characteristics improves the forecasting accuracy for those products in both ANN and RF versions. The results likewise show the tuning process helps you to find better plus more robust results. Thus, feature selection as well as tuning procedures enhance the forecast accuracy for those products.

Practical contributions

In order to show our proposed strategy, we offered a case study of a hypothetical Chinese business exporting freezers and refrigerators to the United Kingdom. At the very first stage of the methodology of ours, the BCG Matrix is used with the international trade information as the present market situation. Next, we include the forecasted succeeding values to the BCG Matrix with our suggested Big Data Analytics technique. In order to evaluate the effectiveness of the suggested methodology, 3 distinct sub product organizations within the key product class of refrigerators plus freezers are selected in the case study.

As evidenced in the case study, foreseeing market situations offers crucial managerial insights for businesses. Businesses are able to create much better strategic choices based on upcoming industry trends and grab business opportunities. On the flip side, if industry development is expected for a product, airers4you is able to plan to boost manufacturing capability and make brand new contracts with vendors to fulfill raw material demands. Based on the objective and the readily available information sets, the actions of the proposed framework could be easily used in various other real life uses inside the supply chain, for example need management, supplier risk management, product quality management, along with predictive upkeep.

Limitations as well as potential research

For demonstrating the usefulness of the suggested technique, this particular analysis provides just one case of a hypothetical Chinese business. For future work, the methodology of ours could be tested on various other country pairs that have the same product types. As an additional extension, the suggested technique could be put on to different country and products pairs to determine the substantial factors impacting the bilateral trade volumes between nations. This particular study is centered on foreseeing 3 weeks ahead; however, some other forecast horizons could be examined to check the proposed methodology. Lastly, some other machine learning methods could be used to compare with ANN and RF.

References

- [1] AMIR E. Khandani, Adlar J. Kim, Andrew W. Lo, Consumer credit-risk models via machine-learning algorithms, ELSEVIER 2010.
- [2] AKTER Jahan, S., & Sazu, M. H. (2022). Rise of mobile banking: a phoenix moment for the financial industry. *Management & Datascience*, 6(2).
- [3] CHENG-Lung Huang, Mu-Chen Chen, Chieh-Jen Wang, "Credit scoring with a data mining approach based on support vector machines", ELSEVIER Expert Systems with Applications 33 (2007) 847–856.
- [4] CHITRA, B. Subashini, "Data Mining Techniques and its Applications in Banking Sector", IJETAE 2013.
- [5] PRASAD, D., Associate Professor Hyderabad Business School, GITAM University, Hyderabad "Prediction of Churn Behavior Of Bank Customers Using Data Mining Tools" 2012

-
- [6] SAZU, M. H. (2022). Does Big Data Drive Innovation In E-Commerce: A Global Perspective?. *SEISENSE Business Review*, 2(1), 55-66.
- [7] ISENBERG, D. T., Sazu, M. H., & Jahan, S. A. (2022). How Banks Can Leverage Credit Risk Evaluation to Improve Financial Performance. *CECCAR Business Review*, 3(9), 62-72.
- [8] SAZU, M. H., & Jahan, S. A. (2022). How Big Data Analytics is transforming the finance industry. *Bankarstvo*, 51(2), 147-172.
- [9] JAHAN, S. A., & Sazu, M. H. (2022). Factors Affecting The Adoption Of Financial Technology Among The Banking Customers In Emerging Economies. *Financial Studies*, 39.
- [10] CHITRA, K., B.Subashini, "Customer Retention in Banking Sector using Predictive Data Mining Technique", ICIT 2011.
- [11] KUCHIPUDI Sravanthi, Tatireddy Subba Reddy, "Applications of Big data in Various Fields", IJCSIT 2015.
- [12] MOHAMMAD Ridwan Ismail, Besut, Terengganu, Malaysia, "A Multi-Layer Perceptron Approach for Customer Churn Prediction", *International Journal of Multimedia and Ubiquitous Engineering* Vol.10, No.7 (2015).
- [13] SAZU, M. H. (2022). How Machine Learning Can Drive High Frequency Algorithmic Trading for Technology Stocks. *International Journal of Data Science and Advanced Analytics (ISSN 2563-4429)*, 4(4), 84-93.
- [14] SAZU, M. H., & Jahan, S. A. (2022). Can big data analytics improve the quality of decision-making in businesses?. *Iberoamerican Business Journal*, 6(1), 04-27.
- [15] SAZU, M. H., & Jahan, S. A. (2022). Impact of big data analytics on business performance. *International Research Journal of Modernization in Engineering Technology and Science*, 4(03), 367-378.
- [16] SAZU, M. H., & Jahan, S. A. (2022). Impact of blockchain-enabled analytics as a tool to revolutionize the banking industry. *Data Science in Finance and Economics*, 2(3), 275-293.
- [17] N. SUN; J. G. Morris, J. Xu, X. Zhu, "iCARE: A framework for big data-based banking customer analytics", IEEE 2014.
- [18] OYENIYI A.B. Adeyemo, "Customer Churn Analysis in Banking Sector Using Data Mining Techniques", IEEE 2015.
- [19] UTKARSH S., Santosh Gopalkrishnan, "Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks", ELSEVIER 2015.
- [20] XINDONG W., Vipin, K., Quinlan. J. R., Joydeep Ghosh, Qiang Yang, "Top 10 algorithms in data mining", Springer 2008.