Supply Chain Delivery Prediction by Supervised Machine Learning Classification Approach

Anurag Rajput
rajpu001@rangers.uwp.edu
University of Wisconsin Parkside
Kenosha, Wisconsin, USA

ABSTRACT

Delivery prediction plays a crucial role in supply chain management, which enables flow of goods from one place to another by delivering right item, at right place, and at right time, but there are various reasons due which delivery of goods in supply chain affected, reasons can be internal which depends on organization to organization and it can be external like rain, flood etc. In this paper I have proposed the use of classification supervised machine learning algorithms. Classification techniques are applied to predict delivery status, that is delivery will be delay or on-time. To predict delivery status, I have shown results of popular binary classification techniques such as Logistic Regression, Support vector machines, KNN, Naive Bayes method and decision tree. At last, performance of each classifier evaluated through different measures for better delivery prediction.

KEYWORDS

Supply chain delivery prediction modelling, classification techniques, Machine learning relevance

ACM Reference Format:
Anurag Rajput. 2020. Supply Chain Delivery Prediction by Supervised Machine Learning Classification Approach. In *The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20), March 11–14, 2020, Portland, OR, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10. 1145/3328778.3366876

1 INTRODUCTION

Supply Chain Management is a series of business actions that offer products, services, and information to customers, starting with suppliers and ending with customer decisions. It entails activities that efficiently integrate suppliers, manufacturers, warehouses, transporters, retailers, and customers so that the right product or service is distributed in the right quantities, to the right locations, and at the right time to reduce system-wide costs while meeting customer service levels. The primary goal of SCM is to achieve long-term competitive advantage.

Integration of information technology and increased competitiveness are required for effective supply chain management in the contemporary era of the global supply chain (SCM). Many company has implemented Supply Chain Management (SCM) as a strategy for increasing market competitiveness. The SCM philosophy proposes combining production planning, sourcing, logistics, and customer relationship management into a single, integrated set of value-adding activities that covers both internal and external operations with the firm's supply chain partners. Performance measurement is crucial in the operation of a supply chain, and various scholars have looked at the importance of performance measurement in supply chain management [1].

The purpose of this study is to provide an overview of supply chain delivery performance models that evaluate delivery performance using delivery windows. The difference between the earliest acceptable delivery time and the latest acceptable delivery time is known as a delivery window. To characterize deliveries as early, on-time, or late, timing benchmarks are utilized. Early and late deliveries are seen as delivery process flaws that waste resources in the supply chain. Early deliveries result in higher inventory holding costs, whereas late deliveries result in higher production stoppage costs and a loss of customer trust[1].

The goal of introducing supervised learning algorithms is to give a flexible platform for generating high-quality predictions and exploring previously unknown complicated challenges.

2 DATA COLLECTION

The dataset used in this paper is obtained from a company *DataCo* which is a leading international Data Information Management solutions and services company. Data consist 53 variables and 180519 observations which are both numeric and categorical in nature. For modelling relevant variables have chosen based on past study and domain based criteria, hence total variables step down from 53 to 11.

3 RELATED WORK

Alfred L. Guiffrida in his paper *"Recent Trends in Supply Chain Delivery Models"* have defined classified supply chain delivery per- formance models. The models have been classified by their key attributes such as model type (loss function versus Six-Sigma) and form of pdf/pmf (Gaussian versus non-Gaussian; continuous versus discrete) used to model the supply chain delivery time distribution.

Ramkumar Harikrishnakumar, Alok Dand, Saideep Nannapaneni, Krishna Krishnan in their research paper *"Supervised Machine Learning Approach for Effective Supplier Classification"* mentioned classification approaches for effective supply chain management. They have used supervised classication technique for model building, which is then used for decision-making.

SIGCSE '20, March 11–14, 2020, Portland, OR, USA

ACM ISBN 978-1-4503-6793-6/20/03. . . $15.00
https://doi.org/10.1145/3328778.3366876

Kamal Kumar Rajgopalan in his paper *"Global Trends in supply chain management"*, which published on 2016 have mentioned global key factors in supply chain management. Trends included in the paper are related to 1) Demand planning, 2) Globalization,
3) Increased competition and price pressures, 4) Outsourcing 5) Shortened and more complex product life cycles 6) Shortened and more complex product life cycles.

Let's understand these six trends in detail.

Demand planning :- As manufacturing sources and capacities have grown, more organizations are shifting their efforts away from plant-level production planning and toward a demand-driven focus, attempting to influence and manage demand more efficiently. Advanced demand planning systems and appropriate tactics can also aid in the discovery of data and the identification of patterns buried in a company's data systems.

Globalization :- Globalization is having a significant impact on how company is handled and transacted, even at the most local levels, thanks to improvements in communications. The supply chain is the part of a company that is most affected by the move toward a global business environment. The increased integration of a global customer and supplier base has had a significant impact on manufacturing, distribution, material sourcing, invoicing, and returns, and many companies are finding that their existing processes and technology are not flexible enough for this new business environment.

Increased competition and price pressures :- Due to the ongoing commoditization of many items, businesses must find new ways to differentiate themselves. Prices for several of a large global consumer goods manufacturer's staple products fell by as much as 60-80 percent in one occasion. They were no longer able to fetch a higher market price due to product innovation and brand equity. They made considerable cost improvements through supply chain re-design and technology in order to compete with that commoditized product.

Outsourcing :- Some people recognize that outsourcing parts or the entire supply chain can be beneficial. Companies are reaping further synergies by outsourcing all or parts of their supply chain as the marketplace improves around (1) information media and systems, (2) cost and quality of global manufacturing and distribution, and (3) product design capabilities.

Shortened and more complex product life cycles :- Many multinational corporations (MNCs), transnational corporations (TNCs), and international business corporations (IBCs) are under pressure to produce novel products and get them to market faster, while reducing cannibalization of established, high-demand products. Companies require more effective product lifecycle management systems to suit the needs of both customers and consumers. This involves a strong focus on new product launch, product discontinuance, design for manufacturability, and leveraging across all product and infrastructure features.

Collaboration between stakeholders' customers and sup- pliers :- Beyond linking information systems, the level of collab- oration extends to fully integrating business processes and orga- nizational structures among enterprises that make up the entire value chain. The ultimate goal of cooperation is to improve visi- bility across the value chain in order to make better management decisions and, as a result, lower value chain expenses. Collaboration gives essential employees along the value chain with the knowledge they need to make business-critical choices with the best available information when the correct tools, processes, and organizational structure are in place.

Paul Moynihan, Wei Dai in their paper *"Knowledge-based Service System for Supply Chain Management"*, they have explained through a research prototype that mimics supply chain management sit- uations, automated supply chain integration and optimization is possible. It will demonstrate possible cost and time savings, as well as an optimization technique that might be offered in real time, using such an approach. In addition, an empirical investigation will be conducted utilizing the prototype to demonstrate the differ- ence between combinatorial and heuristically guided optimization procedures in supply chain optimization. The supply chains will be initialized, coordinated, and monitored using Service Oriented Architecture (SOA) approaches in the prototype.

Inda Sukatia, Abu Bakar Hamida, Rohaizat Baharuna, Rosman Md Yusoffa in their research *"The Study of Supply Chain Manage- ment Strategy and Practices on Supply Chain Performance"*, they have mentioned relationship between supply chain management strat- egy and chain management practices on supply chain performance, The main data collection tool used was a questionnaire, which was administered to a total sample of 200 managers from the Malaysian manufacturing industry. Respondents were classified by their job functions, which included corporate executive, purchasing, manufacturing/production, distribution/logistic, SCM, transportation, material, and operation. The response rate was 62 percent, with 51 percent of the questionnaires being use able. The convenience sampling method was used to choose the samples. Mean, standard deviation, and correlation between independent and dependent variables were used to examine the data. The analyses involved statistical methods such as reliability and validity tests and multiple regressions. The finding showed that supply chain management practices have a significant relationship with supply chain perfor- mance statically. However, supply chain management strategy is a weak predictor of supply chain management performance.

4 MACHINE LEARNING ALGORITHM SELECTION

Machine learning models have gotten a lot of attention and recogni- tion among the methodologies mentioned above, both in academia and in industry. Machine learning algorithms (MLA) are divided into three categories in Figure 1: supervised learning, unsupervised learning, and reinforcement learning. Labeled input and output data are accessible for supervised learning. The inputs and outputs are not labeled in an unsupervised learning process, and the algorithms identify unknown patterns in the data. In reinforcement learning, software agents determine the system's optimum behavior and take measures to maximize its performance automatically. Classification, regression, clustering, and association analysis are all solved using the first two types of algorithms. In this paper I have shown results of some popular classification techniques. 1) Logistic Regression 2) Naive Bayes 3) K nearest neighbor algorithm 4) Decision Tree 5) Support Vector Machines.

ATTRIBUTES SELECTION FOR DELIVERY PREDICTION

Days for shipment :- It is turn around time provided by dataco from source to destination.

Shipping Mode :- This is the class defined according to an organization business policy, for this dataset shiping mode have classes 1) Standard 2) First Class 3) Second Class 4) Same day.

Order Item Quantity :- Number of Items in a particular shipment.

Latitude, Longitude :- Latitude and longitude of delivery station.

Type :- Payment type 1) Cash, 2) Debit, 3) Payment, 4) Debit.

Order Status :- It is a order status at some specific moment like pending, On-hold, closed etc.

MODEL DEVELOPMENT

Data Preprocessing :- The data is cleaned and preprocessed at first to ensure that there are no flaws in the data, such as noise, redundancy, or missing data, that could impair the model's overall performance.

Data splitting :- 80 percent data were chosen for training the model using a random sampling method, and 20 percent were used for modeling analysis. A bigger fraction of the dataset is used to train the classification model in order to improve its accuracy.

Algorithm selection :- The algorithm selection is the second step in the model generation. In this paper, I have selected supervised classification algorithms like KNN, Logistic regression, Naïve Bayes, Decision Tree, and SVM methods for classifying the data. The selected model is then trained and the performance of the models are evaluated.

MODEL VALIDATION

Accuracy Comparison:- All of the individual models' accuracy is measured, and the models with the highest accuracy are compared to those of other models.

Recall:- Recall is the number of true positives divided by the number of true positives plus the number of false negatives.

F1 Score:- The F1-score takes the harmonic mean of a classifier's precision and recall to create a single statistic. It's mostly used to compare the results of two different classifiers.

RESULT

Classifier/Performance Parameters	Accuracy	Recall	F1 Score
Logistic Regression	68	61	71
Naïve Bayes	69	61	70
Support Vector Machine	54	47	0.3
KNN	71	68	69
Decision Tree	82	82	81

CONCLUSION

To classify delivery prediction into two groups, on-time and delay, I have presented a supervised learning-based approach. In this paper, I have presented a multi-class classification model that provides a suitable framework for identifying the appropriate class as well as the ap- propriate algorithm identification. The model gives an accurate, efficient, and effective technique for predicting delivery. A larger dataset will be used in future work to train and assess the model's performance when it is subjected to a high level of complexity. Fu- ture research should also look into various optimization algorithms that can improve the efficiency of all classifiers, particularly when modeling complex scenarios.

REFERENCES

[1] [1] Alfred L. Guiffrid, Recents Trends In Supply chain Delivery Models: 2014, International Journal of Social, Management, Economics and Business Engineering Vol:8 No:6, 2014.

[2] [2] Ramkumar Harikrishnakumar, Alok Dand, Saideep Nannapaneni, Krishna Krishnan, 2019, Supervised Machine Learning Approach for Effective Supplier Classification

[3] 18th IEEE International Conference on Machine Learning and Applications (ICMLA), 2019.

[4] [3] Kamal Kumar Rajagopalan, Global trends in Supply Chain Management, ZIJBEMR, Vol.6 (1), JANUARY (2016), pp. 99-112, 2016.

[5] [4] Inda Sukatia *, Abu Bakar Hamida , Rohaizat Baharuna , Rosman Md Yusoffa, The Study of Supply Chain Management Strategy and Practices on Supply Chain Performance, The 2012 International Conference on Asia Pacific Business Innovation & Technology Management, 2012.

[6] [5] Paul Moynihan, Wei Dai, Knowledge-based Service System for Supply Chain Management, 2010 7th International Conference on Service Systems and Service Management,JAPAN, 2010.

[7] [6]P. R. C. Gopal, and J. Thakkar, "A review of supply chain performance measures and metrics: 2000-2011,"International Journal of Productivity and Performance Management, vol. 61, no. 5, pp. 518-547, 2012.

[8] [2] P. R. Martin, and J. W. Patterson, "On measuring company performance within a supply chain," International Journal of Production Research, vol. 47, no. 9, pp. 2449-2460, 2009.

[9] [3] G. P. Kurien, and M. N. Qureshi, "Study of

performance measurement practices in supply chain management," International Journal of Business, Management and Social Sciences, vol. 2, no. 4, pp. 19-34, 2011.

[10] [4] M. A. Bushuev, and A. L. Guiffrida, "Optimal position of supply chain delivery window: Concepts and general conditions," International Journal of Production Economics, vol. 137, no. 2, pp. 226-234, 2012.

[11] [5] A. Gunasekaran, C. Patel, and E. Tirtiroglu, "Performance measures and metrics in a supply environment,"International Journal of Production and Operations Management, vol. 21, no. 1/2, pp. 71-87, 2001.

[12] [6] A. L. Guiffrida,and R. Nagi, "Cost characterizations of supply chain delivery performance,"International Journal of Production Economics, vol. 102, no. 1, pp. 22-36, 2006.

[13] [7] D. Garg, Y. Naraharai, and N. Viswanadham, "Achieving sharp deliveries in supply chains through variance reduction,"European Journal of Operational Research, vol. 171, no. 1, pp. 227-254, 2006.

[14] [8] M. Roy, R. K. Gupta, and T. Dasgupta, "A technique for determining the optimal mix of logistics service providers of a make-to-order supply chain by formulating and solving a constrained nonlinear cost optimization problem,"Decision Sciences Letters, vol. 2, pp. 1-14, 2013.

[15] [9] M. Safaei, S. Issa, M. Seifert, K. D. Thoben, and W. Lang, "A method to estimate the accumulated delivery time uncertainty in supply networks," In: Dynamics in Logistics, Lecture Notes in Logistics, Kreowski, H.-J., Scholz-Reiter, B. and Thoben, K.-D. (eds.), Springer-Verlag, Berlin, Heidelburg, 2013, pp. 337-347. [10] M. Safaei, M. Seifert, and K. D. Thoben, "Proposed model for increasing the delivery performance in dynamic supply network,"World Academy of Sciences, Engineering and Technology, vol. 39, pp. 620- 623, 2010.

[16] [11] A. L. Guiffrida, "A probability based model for evaluating delivery performance,"Journal of the Academy of Business and Economics, vol. 9, no. 3, pp. 95-104, 2009.

[17] [12] A. L. Guiffrida, and M. Y. Jaber,"Managerial and economic impacts of reducing delivery variance in the supply chain,"Applied Mathematical Modeling, vol.32, no. 10, pp. 2149-2161, 2008.

[18] [13] A. L. Guiffrida, M. Y. Jaber, and R. A. Rzepka, "An economic model for justifying the reduction of delivery variance in an integrated supply chain," Information Systems and Operational Research, vol. 46, no. 2, pp. 147-153, 2008.

[19] [14] A. L. Guiffrida, R. A. Rzepka, and M. Y. Jaber, "A convolution algorithm for evaluating supply chain delivery performance," Proceedings of the 40th Hawaii International Conference on System Sciences (HICSS-40), 9 pages, CD-ROM, IEEE Computer Society, January 3-6, 2007.

[20] [15] A. L. Guiffrida, and R. Nagi, "Economics of managerial neglect in supply chain delivery performance,"The Engineering Economist, vol. 51, no.1, pp. 1-17, 2006.

[21] [16] T. Ngniatedema, and A. L. Guiffrida, "A generalized discrete model for evaluating supply chain delivery performance," Proceedings of the International Conference on Management and Information Systems (ICMIS-2013),Bangkok, Thailand, September 22-24, 2013, pp. 426-430.