Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 3, September : 2023

ANALYSING INTENSITY OF COVID 19 IMPACTS AMONG SHIPPERS: A DECISION TREE MODEL APPROACH.

NAHLA BANU.K Research Scholar,
DR NISSAR ASSISTANT P Assistant Professor College, Tirurangadi
Nahlabanu786@gmail.com , nissarkdp@gmail.com

Abstract
The study presents an in-depth analysis of intensity of Covid -19 impacts on shippers employing a decision tree model. The research involves the collection of primary data through a structured questionnaire, focusing on understanding how the pandemic has affected the different categories of shippers. By applying decision tree model using CHAID technique, the research offers insights that can aid in adapting strategies to navigate such disruptions effectively. The study contributes to both academia and industry by shedding light on the complex interplay between virus and shipping operations, ultimately assisting stakeholders in making informed decisions in the face of uncertain and challenging circumstances.

Introduction
Cargo logistics plays a vital role in facilitating the movement of commodities across the globe, ensuring the timely delivery materials and products to meet customers demand and sustain economies. The efficient functioning of cargo logistics networks is essential for industries to thrive, supply chains to remain robust, and business to remain competitive. However, the emergence of covid 19 pandemic had profound impact on cargo logistics. Border closures travel restriction and workforce challenges disrupted supply chains, leading to increased shipping cost, delays and shortages. The shift in consumer patterns and surge in the e-commerce highlighted the need for adaptable logistics solutions. Despite these challenges, the pandemic also accelerated the adoption of digitalisation and innovative technologies, enabling the industry to enhance efficiency and resilience

Review of literature
(Larrodé et al., 2018) this paper studied the growth of opportunities of an airport in terms cargo operations .A model has been created by using the AHP methodology.it has been identified by the elements that to the lesser or greater extend ,have an impact on the potential of a cargo airport to attract air-freight activity .AHP model has one goal that is growth opportunity of air cargo logistics, four criteria namely economic factors, operative logistics, technological factors ,social legal and environmental factors and fourteen sub criteria and twenty-two attributes. The opinion of experts throughout the model construction process is crucial for the validity of the result. The findings of the study is the factor s related to airport charges and handling costs are the most influential, as well as the existence of a balance between the cargo demand between the destinations.
(Madhavika et al., 2021)**;** studied aspects of cargo handling in Chennai and Bengaluru airports .the author analyzed the functions performed by the stake holders in these two international airports. The study also examined the operations involved in imports and exports of cargo handling. The researcher identified and compared the Chennai and Bengaluru airports in terms of availability of resources, manpower, equipment, infrastructure, capacity planning, slot allocation and screen. The study includes the current cargo problems faced by all the airlines in the cargo handling process. And attempts to minimize the deficiencies in cargo handling and makes certain recommendations to maximize operational efficiency and satisfy customers need in cost effective manner.it recommends minimizes damages and pilferages and brought certain ideas to safeguard the interest the interest of all these take holders so as to give result.
(Huang et al., 2016) evaluates the service requirement of combination cargo carriers (CACCS).a gap index based fuzzy AHP was then proposed to evaluate the perceived difference toward those service

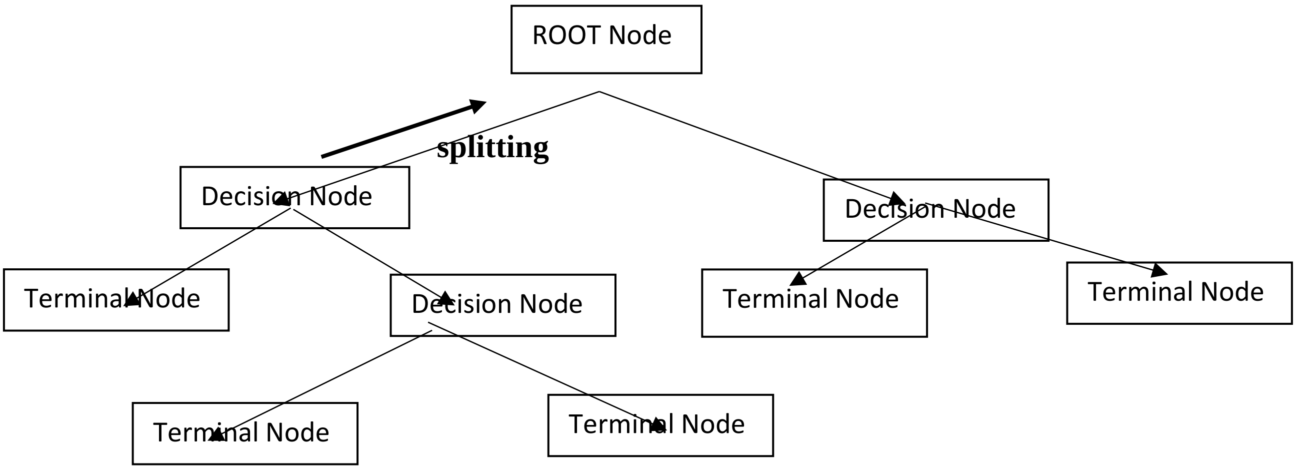
requirement attributes between CACC users and CACC operators. Finally as an empirical study, the CACC in Taiwan and their users pay much attention to service requirement attributes (SRA);the perfect cargo delivery, adequate shipping space, accurate cargo delivery, and staff professional knowledge. Finally the author suggested that improving ground operations, adopting policies of strategic alliances and enhances operating staff professional capabilities.
(Wang, 2007)the author describes service quality of air cargo sector of china airlines .this paper employs quality function deployment to integrate inside quality technology and voice of outside consumers and using "House of quality" charts, illustrates the companies performance in terms of service and offer suggestions for improvement. The conclusion shows that three main factors demanding improvement namely professionalism, physical services and correctness and positive
(Esmaeili & Pourebrahim, 2011)studied the logistics management of international airports in Kerala. The author evaluated the performance of international air cargo handled through the different airports and the services rendered by the freight forwarders were also being the preview of the study. The study concluded that Trivandrum and Kannur airports all the firms always handle perishable cargo but in Cochin it is found to be lesser. And the conclusion states that majority of the freight forwarders never performed packaging, order processing, reverse logistics functions.
(Rabinovich & Cheon, 2011)Analysed the problem towards the handling of containerized cargo and examined various facilities provided by CHA's to customers. Study concluded that CHA need to improve all kinds of their information gathering capabilities like EDI system. The study also suggested that adequate care must be taken to protect the cargo from any kind of damage or pilferage and they should provide adequate number of equipment and needs of the customers demand from time to time.
(Zailani et al., 2017) the study revolves around the overseas marketing of floriculture products and problems, prospects and strategies for the Indian floriculture industry. The study found out that among the problems of floriculture costly air freight ranked first, followed by non-availability air cargo space during peak season, high cost of capital, high cost of technology, weak domestic market, bad domestic road, non-availability of planting material.

Research methodology
The research employs a quantitative approach to investigate the impact of covid 19 among agricultural product exporters. The primary data is collected from the exporters using a structured questionnaire and the collected data is processed using regression tree analysis with CHAID Technique. The study has taken a few demographic characteristics as independent variable and the intensity of covid has taken dependent variable.

Decision Tree (CHAID Technique)
Decision tree is a non-para-metric tool which is a widely applied predictive method to develop data mining model to resolve various DM tasks. Literally this method is applied in certain situations which requires classification, graphical visualisation and dimension reduction(Milanović & Stamenković, 2016). The biggest advantage of decision tree technique is it requires less data cleansing and also not influenced by outliers or missing values to a fair degree.It also widely considered where the situations like nonlinearity arises. Moreover, it is less complex and assists in logical interpretations.
Decision tree is a tree shaped graphical representation which helps determination of possible course of action. Furthermore, its least complex rules, limited computation for classification, convenience to handle any kind of data drives the researchers wider application of decision trees. This technique segregates the population or sample data in to two or more homogeneous groups on the basis of most significant differentiator in the independent variables.
There are certain criteria for the measurement of group homogeneity and for the determination of best split of data. Suppose if the target variable is categorical, then most applied criteria is Gini coefficient, entropy measure and chi-square. These criteria is exclusively for improving the homogeneity of new groups (nodes)regarding the classification of the dependent variable, in comparison to the splitter

group. On the other hand if the target variable is numerical in nature and the node carries numerical values, then the common criteria for measuring the purity and splitting of the regression tree is the sum of squared deviation from node's mean and F-Test. Decision tree structure classifies the entire set of data in to groups named "Nodes" and interconnected by tree branches .The first node in the diagram is root node which comprises of whole data of the training sets which is to be classified for the development of model. A node which again split in to another nodes is called decision nodes, also known as internal nodes. Thus, nodes having no output branches is called terminal nodes.
Therefore, in a tree structure the branch start from a root node and ends up with either decision node or terminal node. Obviously all branches and nodes explains the relationship between input variable and target variable by using *if -then* rules.in short a decision tree is a hierarchical model helps to transform the heterogeneous input data into homogeneous group and finds interrelationship between independent and dependent variables by following specified rules. The conceptual framework of decision tree as follows

Figure 1: Decision tree framework



Cross validation
The complete observed data is separated into k roughly equal-sized disjunctive divisions before being subjected to the cross-validation process, which is carried out over the course of k iterations. Tenfold cross validation, which divides the initial set of data into 10 subgroups, is the most well-known type of cross validation. The starting subset is chosen for testing in each iteration, while the remaining subsets (k-1) are used for model training. Despite the modest number of observations, cross validation provides reliable results.

CHAID Algorithm
Chi-Square Automatic Interaction, or CHAID, was developed by South African statistician Gordan V. Kass in the late 1970s. The most well-known statistical method of machine learning, CHAID, assists in determining the relationship between a categorical target variable and a number of input factors, which may be either categorical or numeric data(Milanović & Stamenković, 2016). When partitioning input data into smaller groups, the CHAID algorithm indicates that the input variable with the lowest p value is the one most strongly related with the dependent variable. As a result, the division of the predictor variables is based on the chi-square and p-value values

Table 2. Model Summery

Specifications	Growing Method	EXHAUSTIVE CHAID
	Dependent Variable	Covid Impact on exports

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, No. 3, September : 2023

	Independent Variables	Type of Shipment, Export promotion Scheme, Experience, Educational Qualification, Type of Payment, No of Branches, Type of Business
	Validation	Cross - Validation
	Maximum Tree Depth	3
	Minimum Cases in Parent Node	5
	Minimum Cases in Child Node	3
Results	Independent Variables Included	Type of Shipment, Experience, No of Branches
	Number of Nodes	8
	Number of Terminal Nodes	5
	Depth	2

Source: Authors calculation

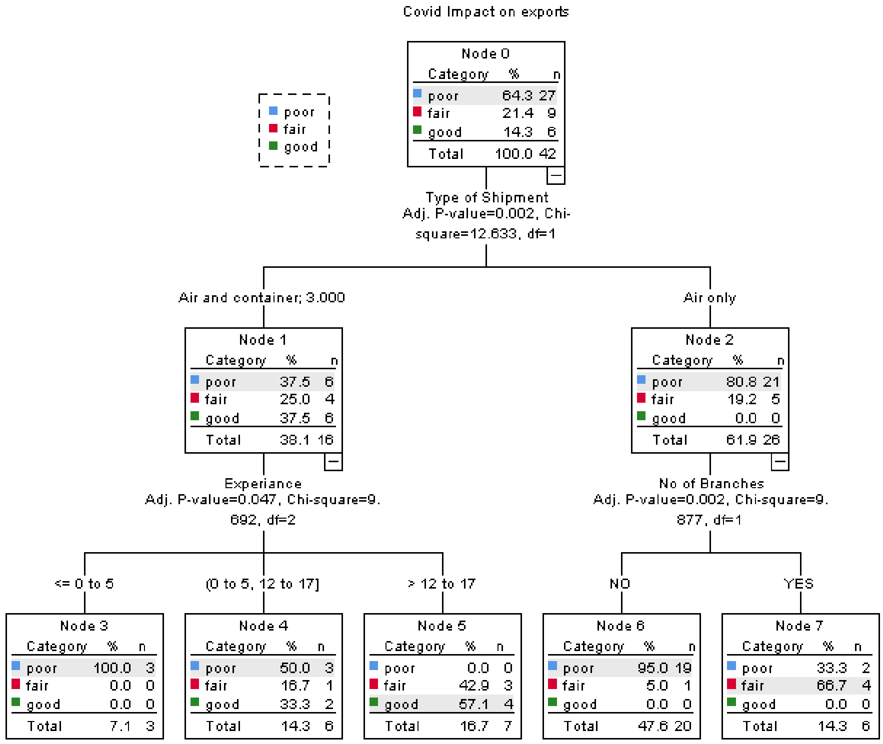
According to the CHAID procedure's results, the constructed model has a total of 8 nodes, 5 of which are terminal nodes, within 3 levels of tree depth. There are a total of six independent variables that were initially stated; however, only three of those variables are included in the final model because the other three do not have any statistically significant associations with the covid effect. The best subsets of predictors were chosen, in other words, depending on the values of the Chi-square criterion and corresponding p values. Thus, a significant reduction in the model's dimensionality is accomplished by applying the CHAID Algorithm. Along with fewer predictors in the final model, it also had fewer categories for the predictors.

Modelling Results

The objective of this technique is to develop a classification model for identification of intensity of covid impact among shippers based on the interdependence of selected personnel characteristics. figure 1 reveals that *type of shipment* is considered as the best predictor in determining or classifying whether the impact is good, fair and bad. alternatively type of shipment is the most significant input variable (independent variable) having substantial association with the target variable (dependent variable). The observed data of this independent variable has strong ability to distinguish and categorize the covid effect .Statistical significance of *type of shipment* was determined, with α=.05 using following values: ($\chi 2$ =12.633,df=1,p-value =0.002). The independent variable namely *type of shipment* serves as the first predictor ,splits the root node,i.e the sample of 42 respondents in to two division containing the categorisation of type of shipment as node1 and node2.In the tree 16 respondents is belonging to node 1(air and container) while the rest of respondent lies in node 2(air only). However, both classifications are dominated by poor covid impact.

Figure 1

CHAID DECISION TREE



In Node 1, approximately 38% respondents of using both air and container shipment is marked as *poor* impact of covid outbreak, while 25%,38% responses are belongs to *fair* and *good* respectively. While in the second node shippers using only air transportation for shipment showing 80%,19.2% as poor and fair impact respectively. Within the second level of the decision tree a couple of statistically significant variables were identified which is *experience* and *number of branches*.experiance is significant for splitting node1 (χ^2 =9.692,*df=2,p-value =0.047)*.according to this classification following three groups of respondents were obtained(Node 3;<=0 to 5,Node 4; 0 to 5,12 to 17,Node 5;>12 to 17).Node 3 explains the shippers belonging to air and container category having the experiences of<= 0 to 5 ,100% responses are marked as *poor*. When it comes to the category 0 to 5,12 to 17, tree demonstrates poor as 50% fair as 17% and good as 33%, where poor is the dominant one. Meanwhile node 5(experience >12 to 17) explains 57 % responses in good category and the remaining 43% remaining 43% are classified as fair

Further number of branches is significant for splitting node 2($\chi 2$ =9.877, df=1, p-value =0.002). while considering node 6(shippers having no branches) majority of responses flows towards poor category (95%),at the same time node 7(shippers having branches)fair takes the dominance approximately 67%,the rest 33 % falling into the poor category.

Assessment of accuracy of CHAID model

Tables 3 and 4 provide information on the established CHAID Model's accuracy and predictive potentiality because these metrics are crucial for understanding the model's performances.

Risk

Table 3

Method	Estimate
Re substitution	.214
Cross-Validation	.429

Source: Authors Calculation

Table 4

Classification

Observed	Predicted			
	poor	fair	good	Percent Correct
poor	25	2	0	92.6%
fair	2	4	3	44.4%
good	2	0	4	66.7%
Overall Percentage	69.0%	14.3%	16.7%	78.6%
Growing Method: EXHAUSTIVE CHAID Dependent Variable: Covid Impact on exports				

Table 3 presents prediction risk as a percentage of inaccurately classified observations.to be more specific, results demonstrates that, if the characteristics of shippers of three independent variables are known, the risk that the shippers are incorrectly classified in terms of covid impact is 21.4% and 42.9% when a test sample is used in model cross validation.in simple words roughly 21%
Probability out of 100% is likely to be misclassified, which means 79% classified in the correct group. Then table 4 showing the classification matrix also known as confusion matrix. The matrix shows the three categories of dependent variable (good, fair, bad), and also displays modelled and empirical values. it can be stated that the overall accuracy of the model is 78.6%. in other words, the model has accurately classified 33(main diagonal of the matrix)out of 42 shippers in the observed data. The percentage structure of predicted (modelled values) according to the categories of the dependent variable (69%:14.3%:16.7%) is not significantly different from that of the actual data

Conclusion and Discussions

The study analyses the relationship of shippers profile and covid impact by taking 42 responses, extracted a set of predictors which are found to be statistically significant (in the hierarchical structure it has shown that *type of shipment* has the strongest interaction with the target variable (covid impact). however, it is extremely important to have a continuous observation by taking large number of responses to ensure the stability of the presented model. Besides increasing the number of observations, increasing the number of input variable also helps to improve the model performance.

References:

Esmaeili, A., & Pourebrahim, F. (2011). Assessing trade potential in agricultural sector of Iran: Application of gravity model. *Journal of Food Products Marketing*, *17*(5), 459–469. https://doi.org/10.1080/10454446.2011.583534

Huang, S. H. S., Tseng, W. J., & Hsu, W. K. K. (2016). An assessment of knowledge gap in service quality for air freight carriers. *Transport Policy, 50,* 87–94. https://doi.org/10.1016/j.tranpol.2016.06.006

Larrodé, E., Muerza, V., & Villagrasa, V. (2018). Analysis model to quantify potential factors in the growth of air cargo logistics in airports. *Transportation Research Procedia*, *33,* 339–346. https://doi.org/10.1016/j.trpro.2018.10.111

Madhavika, W. D. N., S.J.A.N.S., J., Ehalapitiya, K. H. S. M., Wickramage, T. U., Fernando, W. M. D., & Jayasinghe, A. V. T. A. (2021). An Exploration on Supply Chain Resilience Capabilities of Tea Exporting Companies: Special Reference to Pre COVID-19 and During COVID-19 Pandemic. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.4116157

Milanović, M., & Stamenković, M. (2016). CHAID Decision Tree: Methodological Frame and Application. *Economic Themes*, *54*(4), 563–586. https://doi.org/10.1515/ethemes-2016-0029

Rabinovich, E., & Cheon, S. H. (2011). Expanding horizons and deepening understanding via the use of secondary data sources. *Journal of Business Logistics*, *32*(4), 303–316. https://doi.org/10.1111/j.0000-0000.2011.01026.x

Wang, R. T. (2007). Improving service quality using quality function deployment: The air cargo sector of China airlines. *Journal of Air Transport Management*, *13*(4), 221–228. https://doi.org/10.1016/j.jairtraman.2007.03.005

Zailani, S., Shaharudin, M. R., Razmi, K., & Iranmanesh, M. (2017). Influential factors and performance of logistics outsourcing practices: an evidence of malaysian companies. *Review of Managerial Science*, *11*(1), 53–93. https://doi.org/10.1007/s11846-015-0180-x