SUPERVISED MACHINE LEARNING ALGORITHMS FOR MEASURING AND PROMOTING SUSTAINABLE TRANSPORTATION AND GREEN LOGISTICS

Juliana Castaneda
Ph.D. Student, IN3 - Computer Science Dept., Universitat Oberta de Catalunya, Spain
John F. Cardona
Ph.D. Student, IN3 - Computer Science Dept., Universitat Oberta de Catalunya, Spain
Leandro do C. Martins
Ph.D. Student, IN3 - Computer Science Dept., Universitat Oberta de Catalunya, Spain
Angel A. Juan
Full Professor, IN3 - Computer Science Dept., Universitat Oberta de Catalunya, Spain

ABSTRACT

The sustainable development of freight transport has received much attention in recent years. The new regulations for sustainable transport activities established by the European Commission and the United Nations have created the need for road freight transport companies to develop methodologies to measure the social and environmental impact of their activities. This work aims to develop a model based on supervised machine learning methods with intelligent classification algorithms and key performance indicators for each dimension of sustainability as input data. This model allows establishing the level of sustainability (high, medium or low). Several classification algorithms were trained, finding that the support vector machines algorithm is the most accurate, with 98% accuracy for the data set used. The model is tested by establishing the level of sustainability of a European company in the road freight sector, thus allowing the establishment of green strategies for its sustainable development.

1. INTRODUCCIÓN

Concerning the current environmental situation regarding climate changes has impacted people and businesses, making sustainability a trend in all economic activities around the world. The constant design of strategies to mitigate the damage generated by humans' activities on the planet is more than a trend. In the business world, it is becoming a requirement. Integrating technologies to measure the impact of their activities leads to control over them and supports the strategies established to alleviate the generated impact.

Freight transport in the European Union has been growing significantly in the last decade. In 2017, it registered a total increase of 2.4 %, compared to 2016, being road freight transport (RFT) the main contributor with +4.7% (EEA, 2019). This means an increase of the demand for services in freight transport caused by the development of the global trade and influenced by the consumerism of the society (Nowakowska-Grunt and Strzelczyk, 2019; Nowicka-Skowron and Mesjasz-Lech, 2013). RFT is the main source of greenhouse gas (GHG) emissions because of the growth of its activities, which is offering important business opportunities to this sector but also challenges in the emissions reduction (Diemer and Dittrich, 2018). To achieve the proposed objectives, both governments and entrepreneurs have set out to develop sustainable strategies. Currently, there are different frameworks for sustainable freight transport (SFT) with several key performance indicators (KPIs) but with a limited agreement about the general logic and even the basic terminology to use in sustainability status of RFT providers. The common factor in assessing sustainability is that the three pillars of sustainability have to be considered and ensured that they are managed in a holistic way (Gudmundsson et al., 2016).

SFT aims to balance the economic, social and environmental dimensions of the sector in an integrated way to ensure synergy, complementarity and coherence (Zeimpekis et al., 2018; Kumar et al., 2019). From all regulations on the environmental, social and economic impacts left by RFT activities arises the need for businesses and governments to have methodologies to measure that impact. There is an exhaustive list of what SFT entails (Kumar et al., 2019). Among the characteristics, we can highlight the ability to provide safe, socially inclusive, accessible, reliable, affordable, fuel-efficient, environmentally friendly, low-carbon transport that is resistant to shocks and disruptions, including those caused by climate change and natural disasters (Youssef et al., 2017). The European Commission's 2018-2020 work programme for "the smart, green and integrated transport" called for the development and validation of new solutions that can be rapidly deployed.

These solutions should address, systematically, modes of transport, infrastructure and operating patterns, apart from integrating them into a user-friendly European transport system. This must be characterized by connectivity and intelligence, evolving according to the needs of customers and allowing the assessment of the impact of transport solutions on society and the economy, while contributing to the competitiveness of the European transport industry (European-Commission, 2017).

Currently, there is no widespread and structured way to integrate traditional and sustainable objectives of the RFT sector, creating a gap between theory and practice in the development of sustainable strategies. This leads to the question of how to integrate and evaluate the sustainability of enterprises in this sector in order to identify and mitigate negative environmental and social impacts. Recent studies have proposed machine learning techniques to analyze real-world data for decision- making problems (Kaab et al., 2019; Nilashi et al., 2018; Molina-Gómez et al., 2020; Kartal et al., 2016; Nilashi et al., 2019).

Therefore, this paper presents the development of a supervised machine learning model based on classification algorithms, for monitoring the RFT activities and determining the level of sustainability on each of its dimensions. As consequence, it allows companies to define and achieve sustainability strategies in the short, medium and long term. This paper is organized as follows: in section 2, a brief literature review on related topics is presented; section 3 details the proposed methodology; section 4 provides the experimental results in the design and development of the sustainability assessment model; section 5 contains the results of the model implementation in a RFT company; section 6 presents some managerial insights; and finally, Section 7 highlights the main conclusions of this work with future research recommendation.

2. LITERATURE REVIEW

The transport sector is essential for the productive development of any economic and social system. This indispensable sector distributes goods throughout the world and transports people to their homes, work, and schools (Crafts et al., 2005). In Europe, the transport sector represents approximately 5% of the gross domestic product, that jointly with storage, represents between 10% and 15% of the total costs of finished products (Kallas, 2011). As a result, maintaining SFT has gained growing interest within the transportation sector. According to Gatto (1995), SFT is “sustained economic development, without compromising the existing resources for future generations”. In addition, Salas-Zapata and Ortiz-Muñoz (2019) point out that sustainability itself is based on four points: (i) sustainability as a set of socio-ecological criteria that guides human action; (ii) sustainability as a vision of humanity realized through the convergence of social and ecological objectives of a given reference system; (iii) sustainability as an object, thing, or phenomenon which occurs in certain socio-ecological systems; and (iv) sustainability as an approach that involves the incorporation of social and ecological variables in the study of a human activity, process, or product. On the other hand, freight Transport “supports production, trade, and consumption activities by ensuring the efficient movement of raw materials and finished goods and their on-time delivery” (Rajabi, 2011).

According to Centobelli et al. (2020), an effective sustainability program adopted by freight transport providers must include long-term environmental strategies, management execution, and information technologies (ITs) support. Its environmental strategies must focus on prior assessment of opportunities and impacted areas. In addition, SFT involves a balance between the effectiveness and efficiency of the planning and provision of transport services, and the environmental effects resulting from both economic and social circumstances. Similarly, the United Nations conference on trade and development (UNCTAD) established an ecological and socially measurable framework approach for SFT by incorporating the triple bottom line (TBL) framework (Youssef et al., 2017), which addresses the economic, environmental, and social dimensions applying indicators for defining and evaluating sustainability policies. The gathered information from the evaluation provides a broader insight for establishing sustainability guidelines, provided that these dimensions are aligned with the corresponding goals, or United Nations sustainable development goals (UN, 2015). In addition to the TBL, the global reporting initiative is put into place. According to Zhang et al. (2019), this framework that captures economic, environmental and social performance is used as an assessment of sustainability through the reliability of indicators. Additionally, UNCTAD devised a series of framework steps to achieve RFT sustainability. Furthermore, Mostert and Limbourg (2016) substantiate the growing interest in environmental sustainability research in their literature review which identifies various researchers who investigate five environmental challenges: air pollution, climate change, noise, accidents, and congestion. Correspondingly, the RFT sector's environmental sustainability program is aligned itself with measures for reducing carbon dioxide (CO_2) emissions. This alignment includes a framework of four critical points established by the evaluative and logical approach to sustainable transport indicator compilation: measurability, ease of availability, speed of availability, and interpretability.

Also, this framework is required for identifying and selecting sustainable transport indicators (Castillo and Pitfield, 2010). Altogether, research developed for assessing and measuring both logistics and transport sustainability consists of conceptual articles or empirical studies (Marchet et al., 2014).

Reaching and maintaining SFT requires more than just complying with environmental regulations and ordinances. As a result, the transportation sector must devise and incorporate green strategies into its transport operations. A strategic approach proposed by Seroka-Stolka (2014) indicates that green strategies for implementing sustainable development comprises three perspectives: the public or private (stakeholders), the operational and strategic (sustainable performance), and the local or global (geographical location). It should be noted that operational and strategic perspectives are complemented with the adoption of operational changes and the incorporation of environmental principles for strategic planning. In addition, alternative green concepts are devised for reducing the impact of road transport operations. Kadzinski et al. (2017) develop various multi objective application methods for optimizing environmentally compliant supply chains. Measuring environmental sustainability requires an extensive assessment of economic, social, and environmental principles. Although there is no definite model for measuring environmental sustainability, these three principles are fundamental for an effective and efficient sustainable project. From this perspective, additional methods supporting environmental sustainability are considered. For example, when the RFT sector adopts multi-actor and multi-criteria decision-making methodologies (Bandeira et al., 2018; Awasthi et al., 2018), and combined them with fuzzy models (Rai et al., 2017). These methodologies and models collectively allow the assessment of transport sustainability while taking into account the economic, social, and environmental principles. In addition, other factors affecting the sustainability frameworks are defined by the overall goal of the sustainability strategies, whether they be economic, social, or environmental. Moreover, measuring environmental sustainability requires aligning sustainable strategies and the three TBL dimensions mentioned in the early stages of this literature review. Today, many environmental sustainability investigations are limited to one or two TBL dimensions. As a result, not adopting the three TBL dimensions reveals that these three factors are not always attainable for measuring or evaluating environmental sustainability. With the adoption of sustainability measures, the reduction of emissions becomes an evident measurable equation. Therefore, measurable equations can lead not only to minimizing costs and GHG emissions, but also to generating green benefits (Arseculeratne and Yazdanifard, 2014). Consequently, the literature for assessing the sustainability of transportation remains limited and provides only valuable ecological methodologies and strategies and no evaluative framework that measures sustainability itself.

3. METHODOLOGY

The methodology of this research is based on supervised machine learning techniques for the assessment of sustainability through a set of sustainability KPIs. Figure 1 presents the proposed methodology in a schematic way, which consists of four main steps -the selection of the KPIs, the data preparation and training, the evaluation, and the selection of the classification algorithms- and several sub-steps.

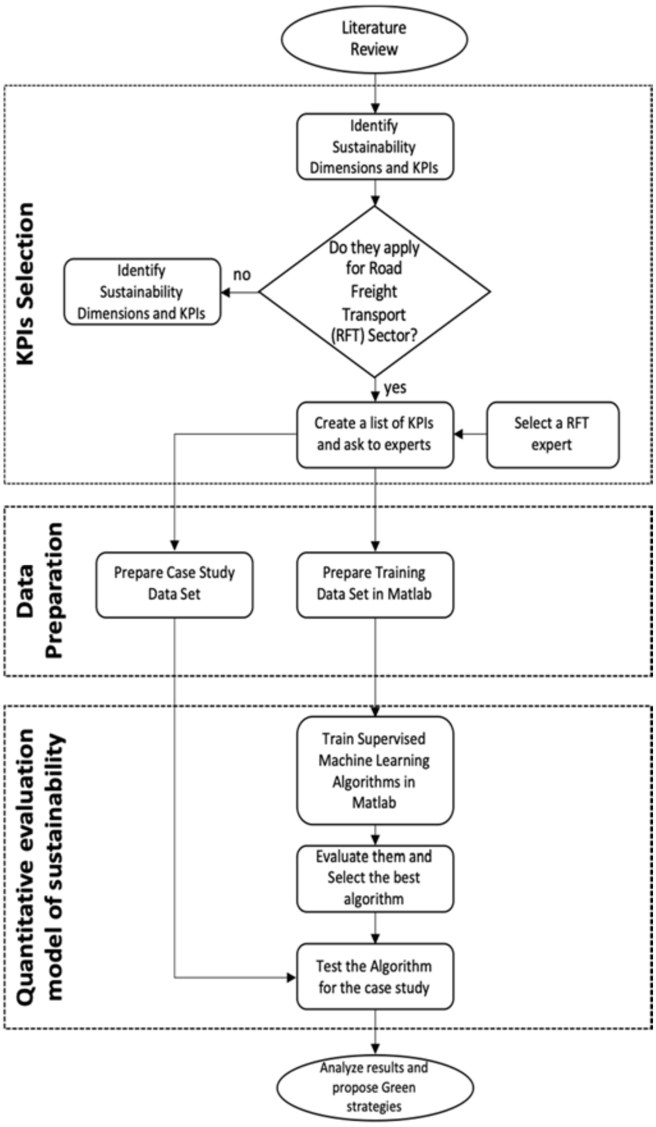

Fig. 1 – Methodology

3.1 Data selection and preparation

Sustainability comprises the TBL, economic development, environmental preservation and social development (Mihyeon Jeon and Amekudzi, 2005). Each of these is made up of a group of KPIs that allow to determine the level of sustainability in that dimension, and also serve as a reference for the quantitative evaluation of sustainability. This case study is based on the European RFT sector and the data was prepared as described in figure 2.

The methodology developed is constructed on the analysis of the KPIs included in different frameworks previously developed by governmental entities, such as the UNCTAD's framework and other scientific proposals, such as the complex performance indicators proposed by Dočekalová and Kocmanová (2016), and the assessment structures of sustainability transport networks (de Campos et al., 2019; Dobranskyte-Niskota et al., 2007; Prause and Schröder, 2015). Once the RFT expert defines the KPIs to be included in the model, the results for the evaluated company are calculated to obtain a total rate for the performance in each of the dimensions. Based on these results, its level of sustainability is measured.

In machine learning techniques, it is important to develop a correct and appropriate training data set, since the algorithms use that information to learn. Because there is no pre-defined data set for measuring sustainability for any of its dimensions, a data set is generated in Matlab with a structure similar to the well-known iris data set from Fisher and Marshall (1936). The values that represent the performance in each of the dimensions are generated as random values with a uniform distribution.

Fig. 2 – Data preparation for the case study

The methodology for the calculation is based on the weighted average. This method is also known as a weighted linear combination or scoring method. It is commonly used in multi-criteria decision-making (Chen, 2012). Generally, the weights of the relative importance may directly be assigned by decision-makers (Afshari et al., 2010). In this case, apart from selecting the most accurate KPIs to the study context, the RFT experts are responsible to assign the corresponding weights, too.

3.2 Development of the Quantitative Evaluation Model

The creation of the model to evaluate sustainability begins with the generation of the training data as aforementioned. For doing that, a series of algorithms available in “Statistics and machine learning toolboxTM” in Matlab which provides functions and apps to describe, analyze, and model data structures is employed. It includes the application called “classification learner” which allows us to train, develop, test, and evaluate several classification algorithms simultaneously. According to the results obtained in the training, the best algorithm is selected for the model development, which is determined according to the classification error (the smaller the error, the greater its accuracy in making predictions) and the metrics for performance evaluation, i.e., the predictive capability of the model (e.g., confusion matrix, cost matrix, ROC curve, etc.). The aim of training several algorithms simultaneously is to find the one that is most accurate for the type of data to be predicted. Figure 3 summarizes the workflow in Matlab for the development of the model. Within the trained algorithms, are included decision trees (Kotsiantis, 2013), discriminant analysis (DA) (Tharwat, 2016), the nearest neighbor (KNN) (Kataria and Singh, 2013; Dhanabal and Chandramathi, 2011), naive bayes (Tripathy and Rath, 2017; Al-Aidaroos et al., 2010), and support vector machines (SVM) (Kotsiantis et al., 2006; Platt, 1998).

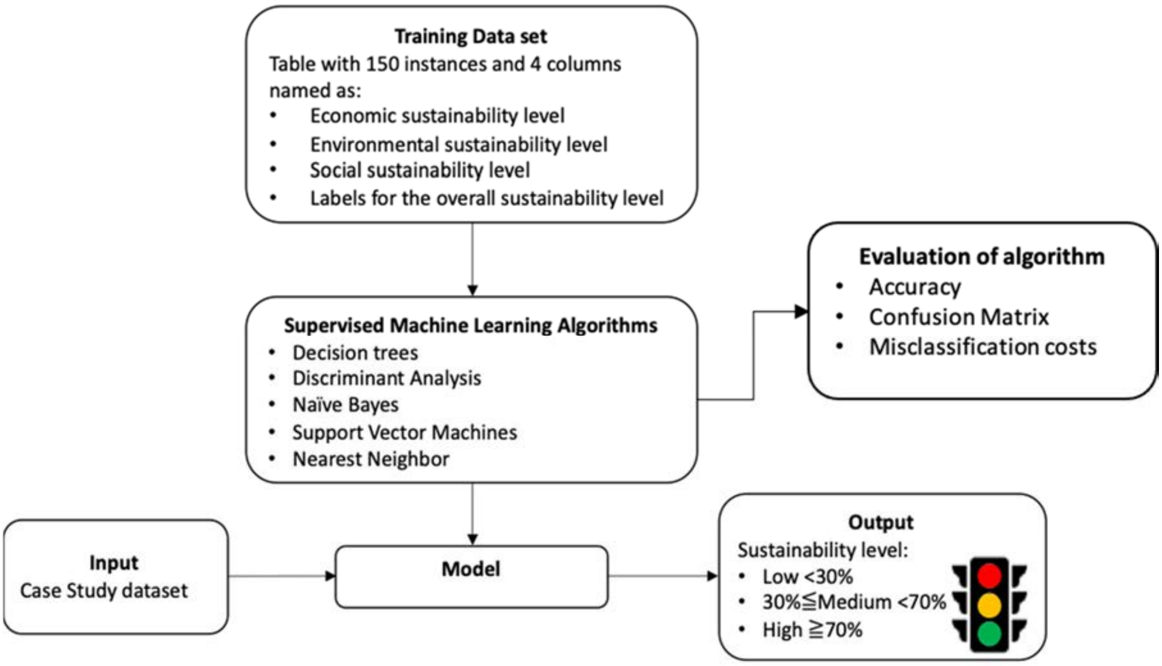

Fig. 3 – Workflow in Matlab

The validation and evaluation of the results of these algorithms are performed in terms of accuracy and classification errors. The estimation of their performance on the predictions for new data compared to the training data is determined by the cross-validation process.

The evaluation of the predictive accuracy of the fitted models is based on the performance in the automated training and the confusion matrix analysis to understand how the model has performed in each class (Amin and Ali, 2018). In addition to these metrics, the performance of the model is also evaluated through a sensitivity analysis to observe how the accuracy of the model changes according to the weights assigned to the dimensions of sustainability.

4. RESULTS AND DISCUSSION OF THE MODEL DEVELOPMENT

For testing our approach, a training data set of 150 instances was randomly generated with a uniform distribution from 0 to 1. As mentioned, these values represent the overall performance in each of the dimensions of sustainability. The data set consists of four columns, each of the first three representing a dimension of sustainability and the fourth the level of sustainability. This level measures the overall level of sustainability, being represented as one of the three following categorical values: “low”, “medium”, or “high”.

For each instance, the RFT expert has defined that its sustainability level is: *(i)* “low” when the weighted sum of the total performance in each dimension of sustainability is greater than 0% and less than or equal to 30%; *(ii)* “medium” when these results are greater than or equal to 30% and less than 70%; and *(iii)* “high” when the values are greater than or equal to 70%. The initial model was trained with the level of impact (weight) on sustainability defined by the expert which was 70% for the economic dimension, 20% for the environmental dimension, and 10% for the social dimension. For each classifier class, Table 1 presents the trained algorithms and their respective results, described by their overall accuracy, the misclassification cost, the prediction speed (in observations per second), and training time (in seconds).

Classifier class	Algorithm	Overall Accuracy	Misclassifi-cation cost	Prediction Speed (obs./s.)	Training Time (s.)
Decision trees	Fine tree	89.3%	16	1600	7.7
	Medium tree	89.3%	16	1700	7.0
	Coarse tree	86.7%	20	1500	6.4
	Boosted trees	56.7%	65	4000	11.5
	Bagged trees	88.0%	18	420	14.7
	RUSBoosted trees	89.3%	20	1500	6.4
Discriminant Analysis (DA)	Linear DA	97.3%	4	1300	9.2
	Quadratic DA	95.3%	7	2700	8.8
	Subspace DA	96.0%	6	320	14.6
Naive Bayes (NB)	Gaussian NB	88.7%	17	2700	8.2
	Kernel NB	87.3%	19	2100	9.7
SVM	Linear SVM	95.3%	7	1300	8.9
	Quadratic SVM	96.7%	5	1700	9.5
	Cubic SVM	96.7%	5	1800	9.4
	Fine Gaussian SVM	76.7%	35	1800	9.7
	Medium Gaussian SVM	95.3%	7	3100	9.1
	Coarse Gaussian SVM	79.5%	31	3100	9.5
KNN	Fine KNN	86.7%	20	2400	9.8
	Medium KNN	82.0%	27	2400	9.6
	Coarse KNN	56.7%	65	3200	10.1
	Cosine KNN	77.3%	34	3800	10.0
	Cubic KNN	82.7%	26	4300	9.9
	Weighted KNN	88.0%	18	4900	9.8
	Subspace KNN	81.3%	28	230	15.6

Table 1 – Results for all trained algorithms

According to the accuracy obtained, the best algorithm is the linear DA with an accuracy of 97.3% to define the sustainability level and the lowest misclassification costs of 4.

Although its training time of 9.2 s. is not among the lowest, it is close to the mean of all times obtained, which is equal to 9.6 s., and a prediction speed of approximately 1300 obs./s. The quadratic SVM, cubic SVM, and linear DA algorithms obtained the highest accuracy. For each of them, Figure 4 presents the obtained confusion matrix, where the number of correctly and incorrectly classified instances is observed.

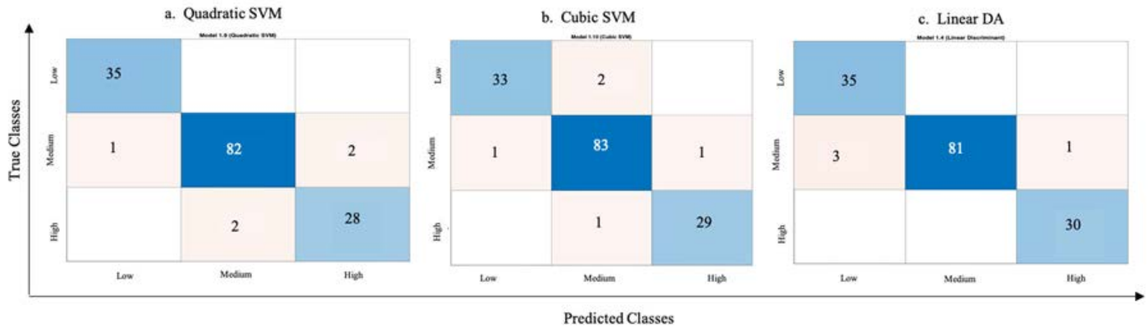

Fig. 4 – Confusion matrix [No obs.] of quadratic SVM, cubic SVM and linear DA

For the Linear DA, a total of 4 misclassifications are observed, being 2 instances as false negatives (FN) for the low and high level of sustainability and 4 false positives (FP) values for the medium level of sustainability. The results obtained show that, in general, the model can be quite accurate, with an F1-Score of 97%. The SVM algorithms have 5 misclassifications. The Cubic has, at least, one FN and one TN for each class, while the Quadratic has no TN for the low class. According to this analysis, both algorithms have a good fit to the data, even though the DA algorithm is more accurate. Comparing the number of misclassified instances, they differ only by one, being 4 for the DA and 5 for SVM. It is possible that by optimizing the hyperparameters of both algorithms, a clearer solution can be obtained as to which one of them fits better to the data used to measure the level of sustainability. When optimizing the hyperparameters of the algorithms with the Bayesian optimizer the SVM algorithm shows accuracy of 98% for measuring the level of sustainability, with 3 misclassified instances. Finally, this model is selected and exported as a code to evaluate the sustainability level of the case study which is a European RFT company.

A sensitivity analysis is performed for the model based on the SVM algorithm. 18 different scenarios were evaluated changing the weights assigned to the dimensions of sustainability. Table 2 presents the accuracy results obtained for each of them after the retraining. The results show that the accuracy of the model can change by approximately 2%, either positively or negatively from the initial 98% accuracy according to the percentage distribution given to the sustainability dimensions to define their impact on the level of sustainability. In particular, the model accuracy is more sensitive to variations where the environmental dimension has the greatest impact on the level of sustainability.

Scenario	Dimension weight			Accuracy
	Economic	Environmental	Economic	
1	0%	100%	0%	99%
2	100%	0%	0%	100%
3	0%	0%	100%	99%
4	33%	33%	33%	97%
5	50%	50%	0%	99%
6	0%	50%	50%	99%
7	50%	0%	50%	99%
8	10%	70%	20%	97%
9	20%	70%	10%	97%
10	10%	20%	70%	99%
11	20%	10%	70%	99%
12	25%	50%	25%	97%
13	50%	25%	25%	99%
14	25%	25%	50%	99%
15	70%	10%	20%	97%
16	60%	20%	20%	99%
17	20%	60%	20%	99%
18	20%	20%	60%	99%

Table 2 – Sensitivity analysis scenarios

5. MODEL IMPLEMENTATION RESULTS IN A CASE STUDY

As a case study, a European company in the freight transport sector with a global transport network is used to evaluate our methodology. This logistics service provider offers not only RFT but also other modes of transport such as rail freight, air freight, shipping, and more services. Since one of its characteristics is the decentralization in the decision-making processes, the sustainability assessment was done only for the region of southern Europe (Iberian countries). From the literature review, the UNCTAD's framework for Sustainable Freight Transport (Youssef et al., 2017) was identified as the most comprehensive framework for the freight transport sector. As KPIs are defined according to the particular circumstances of each case, table 3 presents the KPIs defined for this company with the corresponding definition and formulas according to the experts' criteria.

Dimension	KPI	Definition	Formula
Environmental	Shipments with reported CO_2 emissions	Rate of shipments with monitored CO_2 emissions in relation to total shipments in 1 year (between 0 and 1, the higher the better)	(Shipments with CO_2 emissions reported / Total shipments) * 100%
Economic	Engine Standards	The share of available Euro 6 standards-compliant vehicles (between 0 and 1, the higher the better)	% of vehicles that meet Euro 6 standards
	Transportation cost	Transportation costs as % of turnover (between 0% and 100%, the lower the better)	(Transportation costs / Total turnover) * 100%
	On-time shipments	Rate of on-time shipments in relation to total shipments (between 0 and 1, the higher the better).	[(Total shipments - Shipment delays) / Total shipments] * 100%
Social	Gender equality	Gender equality index among hired employees in the company (between 0 and 1, the higher the better. 1: very good gender equality 0: extreme gender inequality)	(Total number of women employees / Total number of men employees) * 100%
	Workforce Stability	Total workforce Stability index in the company (between 0 and 1, the higher the better. 1: very good workforce stability 0: extreme workforce instability)	(Total number of female employees / total number of employees) * 100%

Table 3 – Case Study KPI definitions and formulas

For the assessment of the level of sustainability, the company provided the data of each KPI for 2019. The weighted average methodology is applied to the company's performance values according to the impact of each of the KPIs determined by the RFT expert for each dimension. As environmental sustainability is given only by one KPI, the company presents a level of environmental sustainability of 77%. The economic dimension is defined as the one with the greatest influence on overall sustainability, and its performance is the lowest of the three with 58%. Transport costs are the most important KPI according to the weight assigned, followed by the other two. For the social dimension, both KPIs present the same level of importance, obtaining a performance of 63%. Based on these values, the input data is calculated to evaluate the sustainability level of the company.

Numerically, the company scored 62% for overall sustainability. Categorically, a high level of sustainability is achieved from a performance of 70%, the company is 8% away from reaching a high level of sustainability, so it has a medium level of sustainability. The greatest weight of the economic dimension on the overall sustainability, and transport costs representing more than 70% of the total turnover, influence negatively the overall performance of this dimension. The results for the other two KPIs of this dimension are good, as on-time deliveries are at 88% and engine standards (Euro VI) are at 90%. These results only represent 40% of overall sustainability. As the environmental dimension is only 20% relevant, its performance only contributes to the overall sustainability by 15%.

The social dimension only represents 10% of the total, contributing 6.3% to the total. With an equitable distribution of the weights, an overall return of 68% is obtained, which only represents a difference of 6% concerning the real value obtained, being also an average level of sustainability. This result means that the company must improve the performance of its sustainability indicators, especially transport costs.

The sustainable strategies are proposed based on the previous results obtained for overall sustainability and each of its dimensions. The selected KPIs reveal the strategies currently proposed by the company for its sustainable development. It can be seen that the company has as its strategy to implement concepts such as the use of IT systems to monitor and control CO_2 emissions, to use environmentally friendly vehicles, to monitor and control the costs and efficiency of transport, and to ensure the equality and well-being of both its employees and the society in general. Understanding the current strategies and performance of the company leads to a medium level of sustainability that allows for the identification of which strategies and which dimension of sustainability should be focused on in the future. Within the company's results, it is noted that all its KPIs are defined on the basis of European regulations for the RFT. Currently, the company does not have any environmental sustainability indicators that actually show the impact of its activities. The integration of a system for measuring and monitoring GHG and CO_2 emissions as well as fuel consumption is a starting point and a valid strategy for the near future.

The evaluated company needs a more solid long-term strategy to continue its sustainable development. Promoting sustainable transport and involving all stakeholders in the development of the strategy is the best way to promote sustainability among customers and employees and to increase business. The proposed methodology for a sustainable strategy consists of establishing KPIs with a clear objective for each of the dimensions. In this case, for each of the dimensions, different KPIs are proposed based on the available frameworks for sustainable RFT. It is also proposed to include as many externalities caused by activities such as accidents, air pollution, climate change, noise, and management as possible. In addition, maintain the commitment to the continuous improvement of its performance for the KPIs that have been initially established for each of the three dimensions of sustainability. This allows for the evaluation of their level of sustainability and to maintain a historical record of the evolution of their sustainable development.

6. MANAGERIAL INSIGHTS

The growing awareness of sustainability in society is putting pressure on companies to integrate the principles of sustainable responsibility into their strategies and policies. Beyond the development of quantitative criteria for evaluating the sustainability of companies based on automatic learning techniques, such as the methodology developed in this work, companies in the RFT sector need to define and adopt sustainable strategies that integrate their three pillars. In the methodology developed, it can be observed that in order to apply these methods, a whole subsequent administrative process at the strategic level is also necessary, which initiates with the definition of sustainability objectives that integrate the three dimensions. Within the objectives, the key performance indicators for each dimension must be integrated and the performance in each dimension, and the general sustainability must be evaluated, as it has been done for the case study. As a final and starting point of a new strategic sustainable cycle, it is required the commitment of the stakeholders supported by ongoing monitoring, reporting, and communications among stakeholders that, at the same time, promote awareness and engagement. This becomes a cycle that must be constantly updated to continue the sustainable development of the company.

Today’s customers are concerned about sustainable development (León et al., 2014). The development and integration of these quantitative models that integrate the three dimensions of sustainability support the decision-making process that integrates sustainability criteria. These methodologies teach companies that they can establish guidelines for their sustainable development that guide them in setting objectives and at the same time evaluate the company’s performance in relation to them. Besides, they are adapted to the particular situation of each company or context of the study. This can be seen in that the input data can vary, i.e., the KPIs, and yet these tools fulfill their purpose. In general, the adoption of this type of strategy shows the social and environmental responsibility that companies in the RFT sector have and how they contribute to sustainable development.

7. CONCLUSIONS AND FUTURE RESEARCH

Nowadays, people and many businesses around the world are trying to develop strategies to mitigate the damage generated by humans' activities on the planet, and therefore, reducing the environmental impacts caused by climate changes. With the increase of road freight transport in Europe, the demand for related services in freight transport has been increased and, consequently, greenhouse gas emissions have been potentialized. To overcome this problem, in this paper, we developed a model based on supervised machine learning methods based on classification algorithms to integrate and evaluate the sustainability of enterprises in the road freight transport sector. This methodology aims to monitor the RFT activities and determining the level of sustainability on each of its sustainability dimensions.

For testing our methodology, a data set was generated in Matlab to represent the overall performance in each of the dimensions of sustainability. Each algorithm has been trained through this data, and that one which presented the best performance was selected to evaluate the sustainability dimensions of a European company in the freight transport sector with a global transport network. According to the results, the optimized SVM classifier obtained using Bayesian optimization has presented the best adaptation to the data and predicted with greater accuracy the level of sustainability. For environmental sustainability, the company presented a level of 77%. For the economic sustainability dimension, the company got 58%, which is mainly represented by transport costs (the most important KPI). Finally, for the social dimension, a performance of 63% was concerned. Numerically, the company got a 62% of sustainability out of the 100% possible, being the company 8% away from reaching a high level of sustainability. Therefore, it implies that the company needs a more solid long-term strategy to continue its sustainable development, where promoting sustainable transport and involving all stakeholders in the development of the strategy is the best way to promote sustainability among customers and employees and to increase business.

Future work could be derived on the basis of this paper. This model could be implemented for other companies and in other economic sectors by modifying the KPIs and adapting them according to the studied context. This would make it possible to verify that the model is not only limited to the RFT sector, but it serves to determine the level of sustainability regardless of the sector being evaluated. This therefore provides an opportunity to explore how accuracy may be affected by the results of the context. On the other hand, the developed SML model is subject to a certain level of subjectivity or bias since the parameters were defined by an expert in the sector. Therefore, the subjectivity could be mitigated by integrating this SML methodology with optimization methods based on heuristics and metaheuristics associated to sustainability criteria such as fuel consumption, external costs, CO_2 emissions, among others. These methodologies are characterized by the use of algorithms that allows for the optimal selection of KPIs that maximizes sustainability based on their impact level. A hybrid model such as this would not only allow a more objective and standardized evaluation of the level of sustainability but would also automatically establish the sustainability strategies.

ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Ministry of Science (PID2019-111100RB-C21 / AEI /10.13039/501100011033, RED2018-102642-T), and the Erasmus+ Program (2019-I-ES01-KA103-062602).

REFERENCES

AFSHARI, A., MOJAHED, M., YUSUFF, R.M., (2010). Simple additive weighting approach to personnel selection problem. International Journal of Innovation, Management and Technology 1, pp. 511.

AL-AIDAROOS, K.M., BAKAR, A.A., OTHMAN, Z., (2010). Naive bayes variants in classification learning, in: 2010 International Conference on Information Retrieval & Knowledge Management (CAMP), IEEE. pp. 276-281.

AMIN, M., ALI, A., (2018). Performance evaluation of supervised machine learning classifiers for predicting healthcare operational decisions. Wavy AI Research Foundation: Lahore, Pakistan.

ARSECULERATNE, D., YAZDANIFARD, R., (2014). How green marketing can create a sustainable competitive advantage for a business. International business research 7, pp. 130.

AWASTHI, A., OMRANI, H., GERBER, P., (2018). Investigating ideal-solution based multicriteria decision making techniques for sustainability evaluation of urban mobility projects. Transportation Research Part A: Policy and Practice 116, pp. 247-259.

BANDEIRA, R.A., D'AGOSTO, M.A., RIBEIRO, S.K., BANDEIRA, A.P., GOES, G.V., (2018). A fuzzy multi-criteria model for evaluating sustainable urban freight transportation operations. Journal of cleaner production 184, pp. 727-739.

DE CAMPOS, R.S., SIMON, A.T., DE CAMPOS MARTINS, F., (2019). Assessing the impacts of road freight transport on sustainability: A case study in the sugar-energy sector. Journal of Cleaner Production 220, pp. 995-1004.

CASTILLO, H., PITFIELD, D.E., (2010). Elastic-a methodological framework for identifying and selecting sustainable transport indicators. Transportation Research Part D: Transport and Environment 15, pp. 179-188.

CENTOBELLI, P., CERCHIONE, R., ESPOSITO, E., (2020). Evaluating environmental sustainability strategies in freight transport and logistics industry. Business Strategy and the Environment 29, pp. 1563-1574.

CHEN, T.Y., (2012). Comparative analysis of saw and topsis based on interval-valued fuzzy sets: Discussions on score functions and weight constraints. Expert Systems with Applications 39, pp. 1848-1861.

CRAFTS, N., ET AL., (2005). The ‘Death of Distance’: What does it mean for economic development? World Economics 6, pp. 1-14.

DHANABAL, S., CHANDRAMATHI, S., (2011). A review of various k-nearest neighbor query processing techniques. International Journal of Computer Applications 31, pp. 14-22.

DIEMER, R., DITTRICH, F., (2018). Transport in the european union-current trends and issues. European Commission, Directorate-General Mobility and Transport, Brussels.

DOBRANSKYTE-NISKOTA, A., PERUJO, A., PREGL, M., 2007. Indicators to assess sustainability of transport activities. European Comission, Joint Research Centre.

DOČEKALOVÁ, M.P., KOCMANOVÁ, A., (2016). Composite indicator for measuring corporate sustainability. Ecological Indicators 61, pp. 612-623.

EEA, (2019). Transport: increasing oil consumption and greenhouse gas emissions hamper eu progress towards environment and climate objectives. URL: https://www.eea.europa.eu/themes/transport/term/increasing-oil-consumption-and-ghg.

EUROPEAN-COMMISSION, (2017). C(2020)6320 of 17 september 2020 horizon 2020 - work programme 2018-2020smart, 11. green and integrated transport. (revised).

FISHER, R.A., MARSHALL, M., (1936). Iris data set. RA Fisher, UC Irvine Machine Learning Repository 440, pp. 87.

GATTO, M., (1995). Sustainability: is it a well defined concept?

GUDMUNDSSON, H., HALL, R.P., MARSDEN, G., ZIETSMAN, J., (2016). Frameworks, in: Sustainable Transportation. Springer, pp. 171-202.

KAAB, A., SHARIFI, M., MOBLI, H., NABAVI-PELESARAEI, A., CHAU, K.W., (2019). Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. Science of the Total Environment 664, pp. 1005-1019.

KADZINSKI, M., TERVONEN, T., TOMCZYK, M.K., DEKKER, R., (2017). Evaluation of multi-objective optimization approaches for solving green supply chain design problems. Omega 68, pp. 168-184.

KALLAS, S., (2011). Transport 2050: Comission Outlines Ambitious Plan to Increase Mobility and Reduce Emmisions. Technical Report. Technical Report March.

KARTAL, H., OZTEKIN, A., GUNASEKARAN, A., CEBI, F., (2016). An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification. Computers & Industrial Engineering 101, pp. 599-613.

KATARIA, A., SINGH, M., (2013). A review of data classification using k-nearest neighbour algorithm. International Journal of Emerging Technology and Advanced Engineering 3, pp. 354-360.

KOTSIANTIS, S.B., (2013). Decision trees: a recent overview. Artificial Intelligence Review 39, pp. 261-283.

KOTSIANTIS, S.B., ZAHARAKIS, I.D., PINTELAS, P.E., (2006). Machine learning: a review of classification and combining techniques. Artificial Intelligence Review 26, pp. 159-190.

KUMAR, A., CALZAVARA, M., VELAGA, N.R., CHOUDHARY, A., SHANKAR, R., (2019). Modelling and analysis of sustainable freight transportation. International Journal of Production Research 57(19), pp. 6086-6089.

LEÓN, R., JUAN, A.A., ET AL., (2014). Promoting corporate social responsibility in logistics throughout horizontal cooperation. Managing Global Transitions 12, pp. 79-93.

MARCHET, G., MELACINI, M., PEROTTI, S., (2014). Environmental sustainability in logistics and freight transportation. Journal of Manufacturing Technology Management. 25(6), pp. 775-811.

MIHYEON JEON, C., AMEKUDZI, A., (2005). Addressing sustainability in transportation systems: definitions, indicators, and metrics. Journal of infrastructure systems 11, pp. 31-50.

MOLINA-GÓMEZ, N.I., RODRÍGUEZ-ROJAS, K., CALDERÓN-RIVERA, D., DÍAZ-ARÉVALO, J.L., LÓPEZ-JIMÉNEZ, P.A., (2020). Using machine learning tools to classify sustainability levels in the development of urban ecosystems. Sustainability 12, pp. 3326.

MOSTERT, M., LIMBOURG, S., (2016). External costs as competitiveness factors for freight transport—a state of the art. Transport Reviews 36, pp. 692-712.

Nilashi, M., Cavallaro, F., Mardani, A., Zavadskas, E.K., Samad, S., Ibrahim, O., (2018). Measuring country sustainability performance using ensembles of neuro-fuzzy technique. Sustainability 10, pp. 2707.

NILASHI, M., RUPANI, P.F., RUPANI, M.M., KAMYAB, H., SHAO, W., AHMADI, H., RASHID, T.A., ALJOJO, N., (2019). Measuring sustainability through ecological sustainability and human sustainability: A machine learning approach. Journal of Cleaner Production 240, pp. 118162.

NOWAKOWSKA-GRUNT, J., STRZELCZYK, M., (2019). The current situation and the directions of changes in road freight transport in the european union. Transportation Research Procedia 39, pp. 350-359.

NOWICKA-SKOWRON, M., MESJASZ-LECH, A., (2013). Globalization and the development of logistics infrastructure of the freight transport by road. Regional Integration: Europe, the Mediterranean and the World Economy. 53rd.

PLATT, J., (1998). Sequential minimal optimization: A fast algorithm for training support vector machines. Microsoft.

PRAUSE, G., SCHRÖDER, M., (2015). Kpi building blocks for successful green transport corridor implementation. Transport and telecommunication journal 16, pp. 277-287.

RAI, H.B., VAN LIER, T., MEERS, D., MACHARIS, C., (2017). Improving urban freight transport sustainability: Policy assessment framework and case study. Research in Transportation Economics 64, pp. 26-35.

RAJABI, M., (2011). Modeling the energy freight-transportation network. Logistics Operations and Management: Concepts and Models, 441.

SALAS-ZAPATA, W.A., ORTIZ-MUÑOZ, S.M., (2019). Analysis of meanings of the concept of sustainability. Sustainable Development 27, pp. 153-161.

SEROKA-STOLKA, O., 2014. The development of green logistics for implementation sustainable development strategy in companies. Procedia-Social and Behavioral Sciences 151, pp. 302-309.

THARWAT, A., (2016). Linear vs. quadratic discriminant analysis classifier: a tutorial. International Journal of Applied Pattern Recognition 3, pp. 145-180.

TRIPATHY, A., RATH, S.K., (2017). Classification of sentiment of reviews using supervised machine learning techniques. International Journal of Rough Sets and Data Analysis (IJRSDA) 4, pp. 56-74.

UN, G.A., (2015). Transforming our world: the 2030 Agenda for Sustainable Development: Draft resolution referred to the United Nations summit for the adoption of the post-2015 development agenda by the General Assembly at its sixty-ninth session. Technical Report. A/70/L. 1, 18 September 2015. http://undocs.org/A/70/L.1.

YOUSSEF, F., BENAMARA, H., WELLER, M., (2017). UNCTAD framework for sustainable freight transport, in: United Nations Conference on Trade and Development Geneva, Switzerland.

ZEIMPEKIS, V., AKTAS, E., BOURLAKIS, M., MINIS, I., (2018). Sustainable Freight Transport: theory, models, and case studies. Springer 63.

ZHANG, X., KANELLOS, N.V., PLANT, E., (2019). Environmental sustainability of logistics service providers: a systematic literature review on indicators for city logistics, in: Symposium on Logistics, pp. 405.