

Business Intelligence and Machine Learning Methods for Predictive Maintenance in Greek Railways

Ilias Kalathas, Michail Papoutsidakis, Christos Drosos

Department of Industrial Design and Production Engineering, University of West Attica, Athens, Greece
Email: mipapou@uniwa.gr

How to cite this paper: Kalathas, I., Papoutsidakis, M. and Drosos, C. (2021) Business Intelligence and Machine Learning Methods for Predictive Maintenance in Greek Railways. *Open Journal of Applied Sciences*, 11, 20-35.

<https://doi.org/10.4236/ojapps.2021.111A003>

Received: March 30, 2021

Accepted: June 27, 2021

Published: June 30, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The extraction of useful information supports the process of making business decisions. In every mechanical process, application or service, the periodic maintenance of the necessary equipment is an expensive process and therefore the technicians and the supervisors have the responsibility of the proper decision making. At the railway companies, a huge amount of data is produced which, with the appropriate processing and smart business systems, can attribute quality information and knowledge. In this paper, the benefits of the business intelligence are presented with the techniques of machine learning and data mining involved of the Greek railway companies, which use obsolete procedures of maintenance. In addition, a study of the present situation is held as well as a record of the needs and requirements of the railway companies. At the same time, tools (open source of low cost) of machine learning and data mining are examined that can assist on the creation of a new strategic support of decisions for the development of the predictive maintenance of the Greek railways making a new complete system of business intelligence. Finally, the results and the motives of the railway companies are presented in order to create applications which can constitute the basic tool for the improvement of making decisions by the business's administration.

Keywords

Business Intelligence, Machine Learning, Data Mining, Decision Making, Rolling Stock, Predictive Maintenance, Weka, KNIME

1. Introduction

The correct decision-making process is one of the most important responsibilities of the executives of a company. It is very important to create data for the fast

taking of valid and precise decisions that the administration will process from selected data with the use of appropriate tools [1]. The need for the provision of improved information and the upgrading of the decision-making process was the impulse for the development of the Business Intelligence systems [2]. A complete system of Business Intelligence consists of a group of technologies, informative applications, processes, and actions in order to collect, analyze, present, and distribute the business information, which consequently will transform into knowledge [3]. The goal of business intelligence is to achieve the optimization of all processes of a business, in a short time as possible by making the best decisions. The method includes various aspects, such as analysis, predictive modeling, performance management, data mining, machine learning.

Data mining is a process that allows the extraction of useful information from the business data of a company and it is an important tool that can assist on the improvement, diagnosis, and the prediction of dangers of a business process [4]. Machine learning using software and information applications is able to learn to act based on the input data using them in a way that people would have used their previous experiences as part of their learning process setting their actions faster and more efficient [5]. Combined applications of data mining and machine learning techniques with present or new information systems are capable of transforming the company's data into strategic business decisions [6]. The accomplishment of the integration of all the systems resulting in the creation of a digital transformation of the companies consists of the business intelligence that should be adopted for the task [7]. Every company in order to accomplish its task, along with other productive factors activates and exploits the equipment available. The operative maintenance of each company's equipment is considered to be a costly process and for its implementation, fast and appropriate decisions are imposed which derive from information process and accurate information [8].

The maintenance is a business process capable of attributing profitability, productivity, and security to every business. Maintenance is defined as the process of the business for the conservation of the equipment in a desirable condition of efficient function and aims at the increase of the reliability of the equipment, the improvement of the production time and the cost decrease [9]. The types of maintenance are:

- Reactive maintenance: It is the method applied to the equipment that is constantly used until some malfunction arises. In this method, the task is interrupted, the malfunction in the equipment is restored and then it continues.
- Preventive maintenance: The basic principle of the Preventive Maintenance is the execution of maintenance tasks according to the manufacturer's standards. It is held in limited by the manufacturer deadlines and spare parts are replaced or equipment components that have been determined by the operating manual in order to minimize the malfunctions and preserve the right application of the equipment.

- Predictive maintenance: It is the method applied with regular inspections, special measurements of the equipment, application of new technologies and replacement of parts. It takes place a short while before the beginning of a malfunction, aiming at the significant decrease of the cost, in relation to the predefined parts replacement suggested by the manufacturer and sets as basic principle the predictive maintenance [10] [11].

At a constantly changing environment that every financial growth is combined with financial methods of transportation, the transportation of people and goods with the assistance of the railways is an urgent need to be achieved. The railway is one of the most inexpensive methods of transportation and its research leads to intensive financial activity [12]. The transportation companies use new applications for the optimum exploitation of data and the extraction of useful conclusions, and the railway companies are based more and more on the effective exploitation of information [13]. The rail transports perform a broad range of processes in order to achieve the main importance of their target which is the passengers' and goods' safety [14]. The railway transportation processes produce a huge amount of data with the appropriate tools of which can offer new technologies that can be analyzed and supply with the making of meaningful and precise decisions for their financial growth and provision of optimum services [15].

The rolling stock of the railway network is an important factor for the right function of the railways transportation as any failure detains the achievement of their strategic goals of safety, passengers' and goods' safety, reliability and velocity of the transportation task provided. The rolling stock of the railways consists of vehicles that are separated in railway wagons and engines. It is distinguished in driving material (the engines) and the traction material (the wagons). The railways are obliged to perform the transportation with safety, speed, and comfort for the passengers and with full protection for the products [16]. The mandatory requirements impose the constant monitoring of the rolling stock before the departure of the train as well as during their route. The systematic maintenance of the trains is essential since during their route many parts are subject to wear. The wear has to be dealt with on time, before it causes damage and anomalies in the circulation or affects negatively and to larger scale the train. The rolling stock maintenance is defined as the function of every railway operation for the preservation and the reliability of the trainsets for the smooth functioning, keeping the principles of the safety of the staff and the environmental protection [17]. The aim of the sector of Rolling Stock maintenance is the optimal function of the trainsets for a long period of time without problems in the minimal cost. In the Railway industry the sector of Rolling Stock maintenance ought to care for the development of the optimal and most efficient procedures of maintenance aiming at the limitation of expenses in order to improve the competitive position of the business, achieving the level of services in offer, to be the same or to the highest level [18]. With the approach of business intelligence

is feasible to avoid dysfunctions in a train which put at risk the passengers, minimizing simultaneously the operational cost, and advancing the quality of the services provided.

Business intelligence solutions supply tools with appropriate technologies in order to help collecting, integrating, storing, processing, and analyzing stored-business data. Business intelligence tools (low-cost open source) that enable changes or improvements and adapt to needs by meeting the requirements of the one who uses them are the WEKA machine learning program software and the KNIME data analysis program software. The two programs directly incorporate the latest technological developments.

In this paragraph of the paper are briefly presented relevant research with the use of the WEKA software of machine learning and the use of KNIME software of data mining and it is the starting point for the development of a smart tool that supports innovative processes for the achievement of a business intelligence system in order to make the right decisions. For the creation of a new system of complete business intelligence concerning the maintenance process in Greek railway rolling stock, the article [19] was made as a contribution. The article [19] describes the optimization of the construction process based on the reading and the prototypes as new approaches of data mining. It uses WEKA data at the Advanced Manufacturing Analytics platform in order to pinpoint hidden patterns on data related to the construction. Additionally, the study [20] refers to the data analysis that took place with machine learning techniques at WEKA, to pinpoint problems that affect the availability of the fleet of aircrafts, the impact of the spare tools availability was assessed, and an intelligent system was developed to minimize the cost of the spare tools. A systematic approach for the prediction of the cost of maintenance of the road construction equipment suggests the article [21]. In this paper the steps of collection, analysis, modeling, and validation of the data are discussed. The paper also presents the progress of different types of models analyzing seven algorithms with the use of WEKA software. The data mining approach that takes place in paper [22] can help transform maintenance processes on Greek railways. The paper [22] presents the development of a simple and easy to be used model for the prediction of an early damage of an oil pump at the oil and gas industry. The data analysis is based on real historic data from sensors. In the KNIME platform algorithms are applied where possible errors are successfully detected and classified ensuring the precision of the prediction. Additionally, the article [23] discusses the results of an industrial project oriented at the integration of data mining KNIME tools at the Enterprise Service Bus (ESB) platform. It reaches the conclusion that the machine of data mining KNIME and web services of clients can be incorporated in the present information systems and correlated with the Industry 4.0. In addition, the purpose of this paper [24] is the planning and the application of a predictive method of maintenance for farm machines. The suggested method is based on the analysis of historic data for the prediction of future failures of a machine. The promotion

of the suggested method took place with the application of an open-source tool software KNIME. The frame of improvement of the effectiveness at the editing of documents is mentioned in the paper [25]. The suggested approach uses techniques of retrieving information and identification of a passage through the techniques of machine learning. For the application of the method and the suggestion for the management of documents the KNIME software was used.

The following work presents the basic scientific knowledge relating to the field of business intelligence and describes the methods of machine learning and data mining. Immediately afterwards, the research approach to the subject takes place and examines the way, techniques, and tools (low-cost open source) of machine learning and data mining that can help reorganize maintenance processes and especially maintenance in railway rolling stock. Then, the problems faced by Greek railway companies are then recorded and follow the positive results of the use of business intelligent systems using the methods and tools of machine learning (WEKA) and data mining (KNIME). Finally, the conclusions are drawn from the implementation of a new comprehensive system of business intelligence proposed to railway companies in order to create a new decision support strategy.

2. Theoretical Background

2.1. Business Intelligence

Modern Technologies of Information and Communication are combined with historic, most of the times inactive data of the companies in order to provide with fast, secure, and reliable information to the people responsible to make decisions [26]. The combination of applications, techniques, processes, business actions and data analysis with innovative techniques for their transformation into information that will lead to more and quality knowledge the executives recommend the business intelligence [27]. The business intelligence system is a company-information system that processes data from the indoor and outdoor environment of the company and provides the administration with data in order to make fast the appropriate and meaningful decisions [28] [29]. A business intelligent system is a unique category of the system and its data are the people, the processes, and the equipment, which interact and collaborate to process data and provide the user (company) with information [30]. Every business intelligence system is constructed with constant developing levels comprising a pyramid. At the initial stage, the raw-historic data are located and at the final is the making of the final decisions as shown in **Figure 1**. For the achievement of the final stage take place processes as Data preparation, Data investigation, data visualization and data mining with optimum technique the application of algorithms of machine learning.

2.2. Data Mining

The data mining focuses on the exploratory analysis of data for the discovery of knowledge. From huge data base using algorithms of machine learning as well as

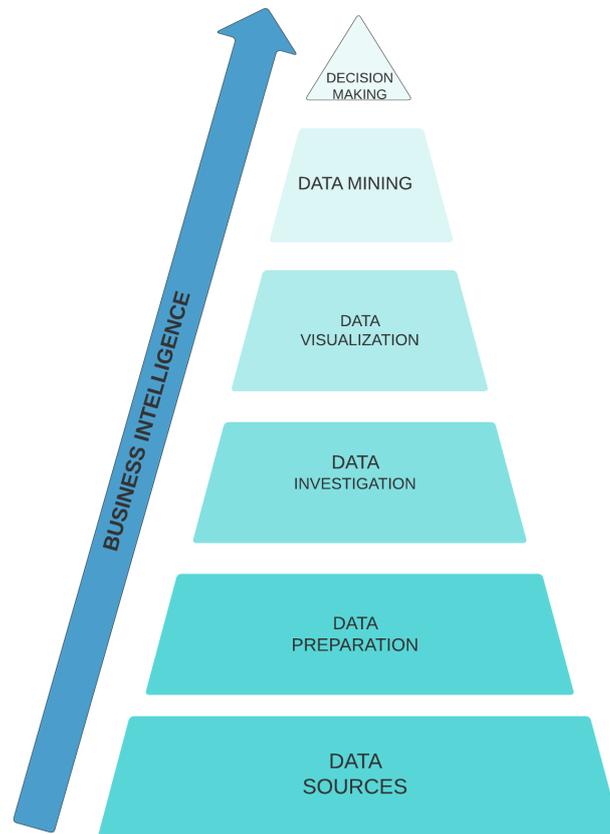


Figure 1. A traditional pyramid of business intelligence data to make better decisions.

every system that supports decisions aims at the detection analysis and classification of a huge volume of information which will be useful in order to make the right decisions [31].

Classification-Regression

The categorization and the regression are tasks of Supervised learning aiming at the discovery of connections, among a set of data. They are techniques of data mining able to predict the value of a dependent variable related to other independent variables from a total of training data that a predictive model is modified and built. The type of the dependent variable defines the difference of the two data mining techniques. For the prediction of a dependent variable that contains distinct - nominal values categorization is used, whereas for the prediction/interpretation of continuous - numerical values the technique of regression is applied [32] [33].

2.3. Machine Learning

The machine learning studies and constructs algorithms that can learn from the data and the information that are provided in any case and make predictions based on them. The problem solving using smart ways is based on the ability of the systems to learn making use of previous knowledge and experience. The ap-

plications that are created are based on algorithms of machine learning make predictions based on the data and make decisions that are expressed as the result [34].

Decision Trees

The construction of decision trees, it systemizes the data, improving the ability of management for analysis and provides the perspective of making rational decisions. The decision trees offer accessible interpretation even from non-professional users because the presentation of the knowledge is easily read, and the output of rules is comprehensible. Very important is the fact that the use of decision trees requires minimal data preparation and handles nominal and numeric input characteristics [35]. Two characteristic algorithms widely used in categorization and regression are C4.5 and o M5 respectively.

- The Iterative Dichotomiser 3 or ID3, as well as its evolution, the C4.5 and its commercial version C5.0 is the prevailing algorithm, which is used to make decision trees [36]. The criterion used for the dividing of the characteristics is called Information Gain and it expresses the decrease of entropy that will arise, if a set of observations (S) is divided into subtotals based on the values of trait A. The Information Gain is quantitatively measured with entropy. The entropy in general expresses the size of the diversity in a set of data. From the algorithm ID3 for the separation of the observations the trait with the highest information gain is chosen [37].
- The M5 algorithm is an algorithm based on decision trees and it is used when the classification receives continuous values and there are quality and quantity data [38]. The initial algorithm M5 was invented by R. Quinlan and Yong Wang made some improvements. At the algorithm M5 the data connections that are defined in their leaves, come from models of linear regression that are trained from examples located at the corresponding intervals [39]. This splitting criterion is defined as the standard deviation of the class values that reach a node as a measure of the error at the node and calculates the expected reduction in error as a result of testing each attribute at the node. Then, the attribute that maximizes the expected error reduction is selected. The result of the separation process stops when the values of the class on the node are selected that maximize the expected error reduction [40].

3. Research Approach

3.1. Machine Learning and Data Mining for Maintenance

The machine learning is a method that is used for the concept of complex models and algorithms that lead to the prediction, which is in the centre of many applications that contribute to the predictive maintenance. In the machine learning, the algorithms and the prediction systems are based on the collection of data and observations. The ability of prediction that Machine Learning offers is capable to optimize the maintenance, the improvement of the quality and the enforcement of the production [41] [42]. The data mining offers the ability of use-

ful information to be analyzed and classified from big sets of data with the use of machine learning algorithms. The supply of useful information, via the data mining process, leads to making the right decisions so as to optimize the maintenance as shown in **Figure 2** and the goals set case by case [43].

3.2. Machine Learning and Data Mining for Maintenance of Rolling Stock

Machine learning includes algorithms which allow a computing system to learn on its own based on previous knowledge and experience. The data mining consists of the application of methodologies of machine learning to large data bases. Huge mass of data is produced in the rolling stock maintenance [44]. Using data mining and machine learning information and results are produced that are useful tools which will lead to the upgrading and updating of the underlying procedures (as shown in **Figure 3**) with the creation of new innovative techniques decreasing the time of the unavailable train sets as well as the cost of the equipment and the spare parts to be serviced [45].

3.3. Machine Learning Tool-WEKA

The WEKA is Machine Learning open source software it developed at the Waikato University of New Zealand and it publicly available according to the license terms GNU General Public License. The use of WEKA offers great variety of machine learning algorithms. All the algorithms take the shape of a relatively simple table which can be read by a file or created by a simple data base question and constructed a decision tree. The reasonable default values ensure that the decision making is possible with the minimum effort. WEKA allows configuration design for data processing and in combination with graphical reproduction of algorithms can be used for predictive maintenance processes [46] [47].

3.4. Data Mining Tool-KNIME

The development of KNIME started in 2004 from a team of software engineers at the Konstanz University of Germany and it is an open-source software to analyze data, with the permission of GNU (General Public License). The KNIME

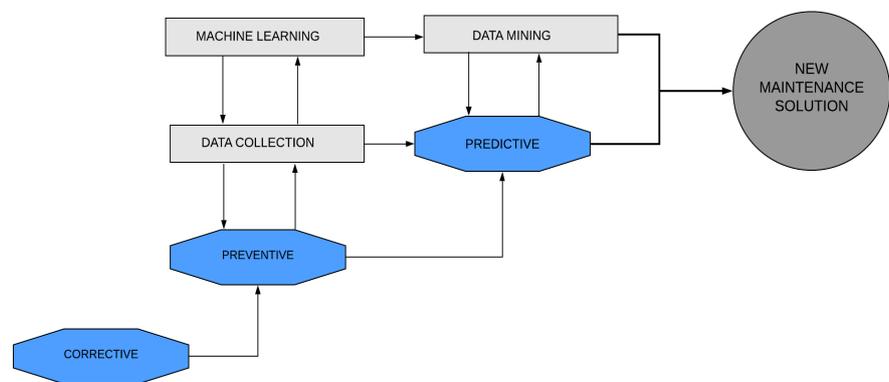


Figure 2. Integrated system for maintenance using machine learning and data mining.

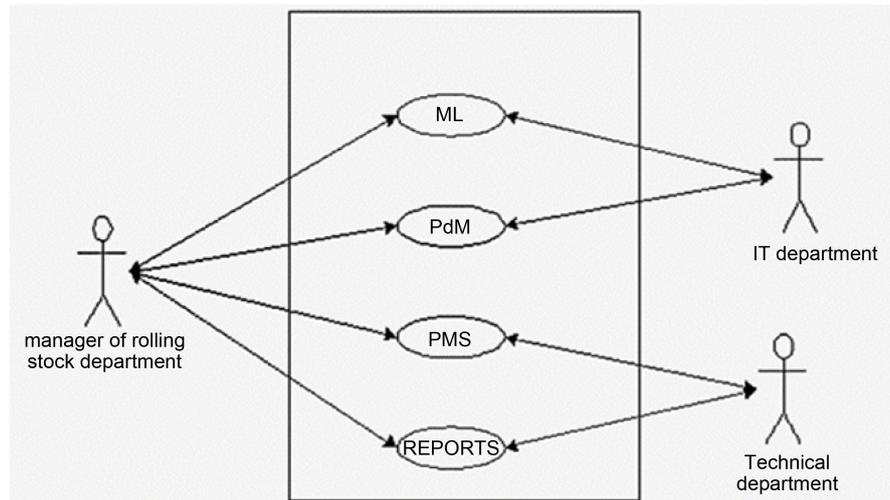


Figure 3. UML chart on new procedures maintenance rolling stock management.

is a tool of business intelligence that includes a complete environment of development and an extensible system plug-in. The open-source platform for data processing KNIME allows to the user the easy and fast analysis of many tasks as the processing, conversion, analysis, and the visual presentation of data. The user can manager huge amount of data, without being limited by the number of lines of data or the processors during the analysis of data. It is compatible with all the operative systems including Linux, Windows and Mac since its materialization has been accomplished on JAVA platform while its planning, offers to users with no programming experience the ability to use it with ease. The way of creation and development of the tool of business intelligence KNIME allows the integration of various units of loading and offers complete procedures for the appearance of the final results [48] [49].

4. Recording Problems in a Greek Company Managing Railway Rolling Stock

The railway transport in Greece is available in four railway ranges and the services are managed by different companies and organizations. Two companies are responsible for the maintenance of rolling stock on the Greek railway network, one for the urban rail network of Athens and one for the rest of the railway network outside the urban fabric. Every company is active in the maintenance for the smooth function of the Rail Rolling Stock with the ultimate goal the achievement of the safe and reliable function of the railway system. The methods of maintenance that are used from both companies are the same:

- Preventive Maintenance: Refers to the periodical inspections which take place in specific time periods or kilometers travelled or the hours of train operation. They refer to whole vehicles and their spare parts, as well.
- Current Maintenance: It is limited in small interventions which allow the rolling stock to circulate, smoothly until the next inspection.
- Unscheduled Repairs: They refer to large scale repairs or of particular me-

chanisms which have suffered damage or wear because of accidents, malfunctions, or other occasional events.

The information systems that are used, manage the needs of companies of transport project, mainly on financial and accounting level. The accounting management includes the staff's salaries, the benefits, and the reports of the travel expenses. The financial management focuses on the fixed costs of the company, that is to say on the purchase of rolling stock spare parts and spare parts for the equipment of the stations (suppliers), on the payments of contractors (security service) and crews (cleaning of stations - trains). With reference to the spare parts repositories at every depot there is an intranet, which enables the information exchange in every plant. With the aid of the intranet takes place, the communication of the administration services, as well. This occurs firstly because they have similar activities and secondly because there are administration decisions which clearly mention that these services must collaborate with each other. It is of main importance the fact that the present information systems dispose of modular architecture in order to allow future extensions and updates. Throughout the conduct of the research, a number of problems were confirmed that the management of the rolling stock faces.

- Problems rolling stock management.
 - 1) Assignment of tasks to the technical units experimentally.
 - 2) Lack of register of trains.
 - 3) Difficulty of the calculation and the monitoring of the cost of the tasks of maintenance.
 - 4) Lack of a single treatment between the depots of the coding of the maintenance and failure tasks.
 - 5) Hand-written keeping of a huge number of records at the depots.
 - 6) Limited ability to reliably predict the consumption of spare parts and materials.
- Wording clear aims-results:
 - 1) Record keeping of the stock and the inflow and outflow of the spare parts used for the repairs and the maintenance of the trains.
 - 2) Promptly available and clear picture of the stock and information about their use.
 - 3) Analysis and illustration of all the movement of every repository and depot daily and monthly.
 - 4) Complete and detailed record of the tasks of maintenance and repair of the trains.
 - 5) Detailed condition of the trains, so as the responsible of the depot know how many trains are operative in the network any given time (real time data processing), how many trains are in the depots and for what reason (operative or not).
 - 6) Saving valuable time and space at the Rolling Stock Administration, as the approval of any process (receiving and delivering material, work orders, permits

etc.) will demand less time.

7) Decrease of time management of the reports at every depot and every depository by the current supervisor.

8) Decrease of the cost of maintenance of the trains saving working hours combined with the interruption of overtime work.

9) Decrease of the stock materials remaining in the repository and being damaged after their expiring date.

10) Decrease of cost coming from the hand-written reports on paper and ink.

11) Monitoring, documentation and comparison of the cost of operation of the trains at every railway line in order the administration of the company make the right decisions.

12) Automation of the processes that will lead to the elimination of the unnecessary bureaucracy.

5. Results of Business Intelligence at the Greek Railway Rolling Stock

A new business intelligence information system that will incorporate the techniques of machine learning and the process of data mining with the tools of open-source (WEKA - KNIME) will ensure the analysis of a huge amount of inactive data (with a low cost), that are produces through the processes of maintenance. By their interpretation, the quality information will generate for the optimum decision making by the administration every company aiming at the transformation of the process of maintenance into a predictive maintenance as shown in **Figure 4**, enabling the companies to evolve financially along with the benefit of improved services to the citizens.

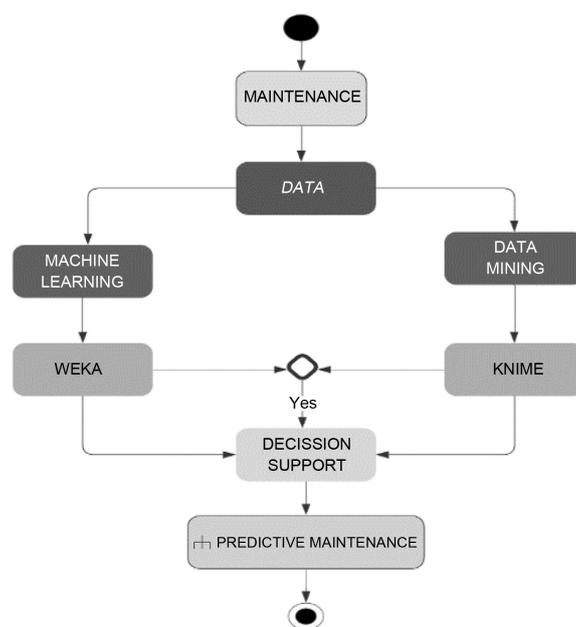


Figure 4. Flow chart diagram for predictive maintenance with business intelligence methods.

The addition of applications at the present information system that the Greek companies manage aims at the reassurance of the reliability and the availability of the Rolling stock, the decrease of the time of non-available trains and the increase of the life expectancy of the rolling stock. The applications can be the basic tool for the improvement of the making-decisions process from the administration of the company [50]. The applications suggested aim at the evolution of the management process of the rolling stock.

- Goals of the innovation recommendation:
 - 1) Right and effective management of the equipment.
 - 2) Improved planning of maintenance tasks.
 - 3) Systematic monitoring of the implementation of the maintenance tasks.
 - 4) Facilitation of the identification of the main problems of the equipment through the statistic processing of the new data.
 - 5) Decrease of the spare parts and materials stock.
 - 6) Improved management of the human resources.
 - 7) Keeping detailed records of spending man-hours and spare parts per maintenance task.
 - 8) Exportation of a single plan of maintenance tasks for the preventive and repairing maintenance.
 - 9) Improved allocation of maintenance tasks to all depots.
 - 10) Keeping a complete record of maintenance for each train.
 - 11) Implementation of a library of technical justification.
- Financial benefits:
 - 1) Pricing of monitoring of the maintenance in general.
 - 2) Planning and performing the maintenance based on the resources available (living and non-living resources).
 - 3) Decrease of the total cost of maintenance which can reach up to 40% with the elimination of overtime and the purchase of inefficient spare parts.

6. Conclusion

With the application of business intelligence systems at the transport section and especially at the Greek railway companies that use outdated processes, the improvement of whole process of making decisions is achieved. Business intelligent systems with the use of machine learning techniques and tools of data mining lead the most effective management of the Rolling Stock, making at the most use of the abilities of the living and non-living sources of every company achieving the goals of the company for the provision of high-quality services. Simultaneously, it reassures the desideratum of the companies for low-cost operation of the trains, reduction of bureaucracy due to automation of processes and the minimizing of errors and overlooking, while there is better management of the constantly increased amount of information and data. The combination of machine learning and data mining with the use of open-source programs (gratis available) reassures the basic politics of the companies for security, protection,

reliability, speed and passengers' satisfaction, without the charge of the budget of the company. Additionally, with the use of business intelligence, a new plan of monitoring can be implemented for the reliability and availability of the equipment under maintenance. This could lead to an efficient control of the consumption of materials for every task for the on-time prediction for the replacement of stock. The digital transformation of companies with the use of business intelligence provides with the appropriate information for fast, valid and right decisions which set new rules for the maintenance of the trains aiming at the implementation of the predictive maintenance to reassure the main target of the railway transportations which is the safety of passengers and goods.

Acknowledgements

All authors would like to thank the University of West Attica and specifically the Post Graduate Program of Studies (MSc) "New technologies in Shipping and Transport", for the financial support provided to them to undertake this research project.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Schultheis, M.P. (2016) The Impact of Business Intelligence Systems on the Perceived Quality of Strategic Decision Making. Ph.D. Thesis, Curtin University, Bentley. <http://hdl.handle.net/20.500.11937/2527>
- [2] Kitsios, F. and Kamariotou, M. (2021) Artificial Intelligence and Business Strategy towards Digital Transformation: A Research Agenda. *Sustainability*, **13**, 2025. <https://doi.org/10.3390/su13042025>
- [3] Azeroual, O. and Theel, H. (2018) The Effects of Using Business Intelligence Systems on an Excellence Management and Decision-Making Process by Start-Up Companies: A Case Study. *International Journal of Management Science and Business Administration*, **4**, 30-40. <https://doi.org/10.18775/ijmsba.1849-5664-5419.2014.43.1004>
- [4] Rao, A.V., Kumari, G.L.A. and Rao, M.V.B. (2008) Blended Attacks Exploits, Data Mining Analytics for Business Intelligence and Decision Support. *Oriental Journal of Computer Science and Technology*, **1**. <http://www.computerscijournal.org/?p=2073>
- [5] Pirani, Z., Marewar, A., Bhavnagarwala, Z. and Kamble, M. (2017) Application of Business Intelligence Using Machine Learning Approach. *International Journal of Computer Applications*, **165**, 28-31. <https://doi.org/10.5120/ijca2017914131>
- [6] Kabakchieva, D. (2009) Business Intelligence Applications and Data Mining Methods in Telecommunications: A Literature Review. *Proceedings of International Conference on Software, Services & Semantic Technologies*, Sofia, 28-29 October 2009. <http://hdl.handle.net/10506/241>
- [7] Yafooz, W.M.S., Abu Bakar, Z.B., Fahad, S.K.A. and Mithon, A.M. (2019) Business Intelligence through Big Data Analytics, Data Mining and Machine Learning. In:

- Sharma, N., Chakrabarti, A. and Balas, V., Eds., *Data Management, Analytics and Innovation. Advances in Intelligent Systems and Computing*, Vol. 1016, Springer, Singapore, 217-230. https://doi.org/10.1007/978-981-13-9364-8_17
- [8] Franciosi, C., Iung, B., Miranda, S. and Riemma, S. (2018) Maintenance for Sustainability in the Industry 4.0 Context: A Scoping Literature Review. *IFAC-PapersOnLine*, **51**, 903-908. <https://doi.org/10.1016/j.ifacol.2018.08.459>
- [9] Agustiady, T.K. and Cudney, E.A. (2018) Total Productive Maintenance. *Total Quality Management & Business Excellence*. <https://doi.org/10.1080/14783363.2018.1438843>
- [10] Ran, Y., Zhou, X., Lin, P., Wen, Y. and Deng, R. (2019) A Survey of Predictive Maintenance: Systems, Purposes and Approaches. *IEEE Communications Surveys & Tutorials*, **20**, 1-36. <https://arxiv.org/pdf/1912.07383.pdf>
- [11] Carvalho, T.P., Soares, F.A.A.M.N., Vita, R., Francisco, R.P., Basto, J.P. and Alcala, S.G.S. (2019) A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance. *Computers & Industrial Engineering*, **137**, Article ID: 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- [12] Petrović, M., Jenić, V. and Kaužljjar, D. (2010) The Role of Railway Stations in Increasing Transport Demand. *WIT Transactions on the Built Environment*, **111**, 249-260. <https://doi.org/10.2495/UT100231>
- [13] Ovchynnikova, V., Kuzmenko, A., Yusupova, T., Toropova, V. and Gontar, N. (2019) Digital Transformation of Innovative Business Processes on Railway Transport. *SHS Web of Conferences*, **67**, Article No. 01009. <https://doi.org/10.1051/shsconf/20196701009>
- [14] Rama, D. and Andrews, J.D. (2016) Railway Infrastructure Asset Management: The Whole-System Life Cost Analysis. *IET Intelligent Transport Systems*, **10**, 58-64. <https://doi.org/10.1049/iet-its.2015.0030>
- [15] Consilvio, A., Solís-Hernández, J., Jiménez-Redondo, N., Sanetti, P., Papa, F. and Mingolarra-Garaizar, I. (2020) On Applying Machine Learning and Simulative Approaches to Railway Asset Management: The Earthworks and Track Circuits Case Studies. *Sustainability*, **12**, 2544. <https://doi.org/10.3390/su12062544>
- [16] Nabian, M.A., Alemazkoo, N. and Meidani, H. (2019) Predicting Near-Term Train Schedule Performance and Delay Using Bi-Level Random Forests. *Transportation Research Record: Journal of the Transportation Research Board*, **2673**, 564-573. <https://doi.org/10.1177/0361198119840339>
- [17] Alkali, B. (2019) Railway Rolling Stock Fleet Predictive Maintenance Data Analytics. *International Journal of Railway Technology*, in Press.
- [18] Dinmohammadi, F., Alkali, B., Shafiee, M., Bérenguer, C. and Labib, A. (2016) Risk Evaluation of Railway Rolling Stock Failures Using FMECA Technique: A Case Study of Passenger Door System. *Urban Rail Transit*, **2**, 128-145. <https://doi.org/10.1007/s40864-016-0043-z>
- [19] Gröger, C., Niedermann, F. and Mitschang, B. (2012) Data Mining-driven Manufacturing Process Optimization. *Proceedings of the World Congress on Engineering*, London, 4-6 July 2012, Corpus ID: 14461305.
- [20] Rahim, A., Hamizah, F., Noor, N.M., Abdullah, H. and Annanurov, B. (2020) Cost Minimization of Aircraft Critical Components for Planning and Maintenance Requirements. *International Journal of Integrated Engineering*, **12**, 40-53. <https://doi.org/10.30880/ijie.2020.12.07.005>
- [21] Bayzid, S.M., Mohamed, Y. and Al-Hussein, M. (2016) Prediction of Maintenance

- Cost for Road Construction Equipment: A Case Study. *Canadian Journal of Civil Engineering*, **43**, 480-492. <https://doi.org/10.1139/cjce-2014-0500>
- [22] Orrù, P.F., Zoccheddu, A., Sassu, L., Mattia, C., Cozza, R. and Arena, S. (2020) Machine Learning Approach Using MLP and SVM Algorithms for the Fault Prediction of a Centrifugal Pump in the Oil and Gas Industry. *Sustainability*, **12**, 4776. <https://doi.org/10.3390/su12114776>
- [23] Massaro, A., Maritati, V., Galiano, A., Birardi, V. and Pellicani, L. (2018) ESB Platform Integrating Knime Data Mining Tool Oriented on Industry 4.0 Based on Artificial Neural Network Predictive Maintenance. *International Journal of Artificial Intelligence & Applications*, **9**, 1-17. <https://doi.org/10.5121/ijaa.2018.9301>
- [24] Lüttenberg, H., Bartelheimer, C. and Beverungen, D. (2018) Designing Predictive Maintenance for Agricultural Machines. *26th European Conference on Information Systems (ECIS 2018)*, Portsmouth, Research Papers 153. https://aisel.aisnet.org/ecis2018_rp/153
- [25] Rahmaoui, O., Souali, K. and Ouzzif, M. (2020) Towards a Documents Processing Tool using Traceability Information Retrieval and Content Recognition through Machine Learning in a Big Data Context. *Advances in Science, Technology and Engineering Systems Journal*, **5**, 1267-1277. <https://doi.org/10.25046/aj0506151>
- [26] Wieder, B. and Ossimitz, M.-L. (2015) The Impact of Business Intelligence on the Quality of Decision Making—A Mediation Model. *Procedia Computer Science*, **64**, 1163-1171. <https://doi.org/10.1016/j.procs.2015.08.599>
- [27] Grabińska, A. and Ziora, L. (2019) The Application of Business Intelligence Systems in Logistics. Review of Selected Practical Examples. *Sciendoc: System Safety: Human-Technical Facility—Environment*, **1**, 1028-1035. <https://doi.org/10.2478/czoto-2019-0130>
- [28] Mishra, B.K., Hazra, D., Tarannum, K. and Kumar, M. (2016) Business Intelligence using Data Mining Techniques and Business Analytics. *International Conference System Modeling & Advancement in Research Trends (SMART)*, Moradabad, 25-27 November 2016, 84-89. <https://doi.org/10.1109/SYSMART.2016.7894496>
- [29] Yusof, E.M.B.M. and Yusof, A.R.M. (2013) The Study on the Application of Business Intelligence in Manufacturing: A Review. *International Journal of Business Intelligence Research*, **4**, 43-51. <https://doi.org/10.4018/jbir.2013010104>
- [30] Bordeleau, F.-È., Mosconi, E. and Santa-Eulalia, L.A. (2018) Business Intelligence in Industry 4.0: State of the Art and Research Opportunities. HICSS. <https://doi.org/10.24251/HICSS.2018.495>
- [31] Bhaskaran, S.S. (2017) An Investigation into the Knowledge Discovery and Data Mining (KDDM) Process to Generate Course Taking Pattern Characterized by Contextual Factors of Students in Higher Education Institution (HEI). Ph.D. Thesis, Brunel University, London. <http://bura.brunel.ac.uk/handle/2438/15880>
- [32] Çiğşar, B. and Ünal, D. (2019) Comparison of Data Mining Classification Algorithms Determining the Default Risk. *Scientific Programming*, **2019**, Article ID: 8706505. <https://doi.org/10.1155/2019/8706505>
- [33] Alam, I., Farid, D.M. and Rossetti, R.J.F. (2019) The Prediction of Traffic Flow with Regression Analysis. In: Abraham, A., Dutta, P., Mandal, J., Bhattacharya, A. and Dutta, S., Eds., *Emerging Technologies in Data Mining and Information Security. Advances in Intelligent Systems and Computing*, Vol. 813, Springer, Singapore, 661-671. https://doi.org/10.1007/978-981-13-1498-8_58
- [34] Chopra, R. (2018) Machine Learning. Khanna Publishing House, New Delhi.
- [35] Bukhsh, Z.A., Saeed, A., Stipanovic, I. and Doree, A.G. (2019) Predictive Maintenance

- ance Using Tree-Based Classification Techniques: A Case of Railway Switches. *Transportation Research Part C: Emerging Technologies*, **101**, 35-54. <https://doi.org/10.1016/j.trc.2019.02.001>
- [36] Cherfi, A., Nouira, K. and Ferchichi, A. (2018) Very Fast C4.5 Decision Tree Algorithm. *Applied Artificial Intelligence*, **32**, 119-137. <https://doi.org/10.1080/08839514.2018.1447479>
- [37] Kyrkos, E. (2015) Business Intelligence and Data Mining [eBook]. Hellenic Academic Libraries Link, Athens, Chapter 4. <http://hdl.handle.net/11419/1231>
- [38] Kaveh, A., Hamze-Ziabari, S.M. and Bakhshpoori, T. (2017) M5' Algorithm for Shear Strength Prediction of HSC Slender Beams without Web Reinforcement. *International Journal of Modeling and Optimization*, **7**, 48-53.
- [39] Lin, L., Wang, Q. and Sadek, A.W. (2016) A Combined M5P Tree and Hazard-Based Duration Model for Predicting Urban Freeway Traffic Accident Durations. *Accident Analysis and Prevention*, **91**, 114-126. <https://doi.org/10.1016/j.aap.2016.03.001>
- [40] Quilan, J.R. (1992) Learning with Continuous Classes. Word Scientific in Proceeding AI'92, Singapore, 343-348.
- [41] Accorsi, R., Manzini, R., Pascarella, P., Patella, M. and Sassi, S. (2017) Data Mining and Machine Learning for Condition-Based Maintenance. *Procedia Manufacturing*, **11**, 1153-1161. <https://doi.org/10.1016/j.promfg.2017.07.239>
- [42] Jimenez-Cortadi, A., Irigoien, I., Boto, F., Sierra, B. and Rodriguez, G. (2020) Predictive Maintenance on the Machining Process and Machine Tool. *Applied Sciences*, **10**, 224. <https://doi.org/10.3390/app10010224>
- [43] Rezig, S., Achour, Z. and Rezg, N. (2018) Using Data Mining Methods for Predicting Sequential Maintenance Activities. *Applied Sciences*, **8**, 2184. <https://doi.org/10.4018/978-1-5225-3176-0>
- [44] Kohli, S., Senthil Kumar, A.V., Easton, J.M. and Roberts, C. (2018) Innovative Applications of Big Data in the Railway Industry. IGI Global, 1-395. <https://doi.org/10.4018/978-1-5225-3176-0>
- [45] Oliffa, H. and Liu, Y. (2017) Towards Industry 4.0 Utilizing Data-Mining Techniques: A Case Study on Quality Improvement. *Procedia CIRP*, **63**, 167-172. <https://doi.org/10.1016/j.procir.2017.03.311>
- [46] Asha Kiranmai, S. and Jaya Laxmi, A. (2018) Data Mining for Classification of Power Quality Problems Using WEKA and the Effect of Attributes on Classification Accuracy. *Protection and Control of Modern Power Systems*, **3**, Article No. 29. <https://doi.org/10.1186/s41601-018-0103-3>
- [47] Kulkarni, E.G. and Kulkarni, R.B. (2016) Article: Weka Powerful Tool in Data Mining. *IJCA Proceedings on National Seminar on Recent Trends in Data RTDM*, **2**, 10-15.
- [48] Feltrin, L. (2015) KNIME an Open Source Solution for Predictive Analytics in the Geosciences [Software and Data Sets]. *IEEE Geoscience and Remote Sensing Magazine*, **3**, 28-38. <https://doi.org/10.1109/MGRS.2015.2496160>
- [49] Jagla, B., Wiswedel, B. and Coppée, J.-Y. (2011) Extending KNIME for Next-Generation Sequencing Data Analysis. *Bioinformatics*, **27**, 2907-2909. <https://doi.org/10.1093/bioinformatics/btr478>
- [50] Hadj-Mabrouk, H. (2019) Contribution of Artificial Intelligence and Machine Learning to the Assessment of the Safety of Critical Software Used in Railway Transport. *AIMS Electronics and Electrical Engineering*, **3**, 33-70. <https://doi.org/10.3934/ElectrEng.2019.1.33>