A data mining approach for wheel defect detection of railway vehicles

Meisam Gordan[1], Ramin Ghiasi[1], Araliya Mosleh[2], Diogo Ribeiro[3], Vikram Pakrashi[4], and Abdollah Malekjafarian[1]

[1]Structural Dynamics and Assessment Laboratory, School of Civil Engineering, University College Dublin, Dublin, Ireland
[2]CONSTRUCT LESE, Faculty of Engineering, University of Porto, Porto, Portugal
[3]CONSTRUCT-LESE, School of Engineering, Polytechnic of Porto, Porto, Portugal
[4]UCD Centre for Mechanics, Dynamical Systems and Risk Laboratory, School of Mechanical and Materials Engineering, University College Dublin, Dublin, Ireland
Email: meisam.gordan@ucd.ie, ramin.ghiasisangani@ucd.ie, amosleh@fe.up.pt, drr@isep.ipp.pt, vikram.pakrashi@ucd.ie, abdollah.malekjafarian@ucd.ie

ABSTRACT: The combination of tight maintenance budgets and aging railway systems draw attention to the urgent need for cost-effective and efficient approaches to monitor and detect damages in railway systems. Structural Health Monitoring (SHM) has been used by many researchers in railway systems to evaluate the health condition of railway assets. However, it demands advanced data analytics due to the large datasets generated from sensor networks. Data mining is a promising computational tool that is qualified for the data extraction process. It can accurately find out the informative features, i.e., knowledge from the generated sensed data. Train wheels play a crucial role in ensuring the safe and efficient operation of trains. Defective wheels in freight railway waggons create a significant risk to railway infrastructure, often leading to considerable damage. A wheel flat occurs when a wheel becomes locked and slides along the rail, highlighting the critical importance of maintaining wheel integrity for overall railway safety. This paper presents a data-driven train wheel defect detection concept which utilises the raw acceleration responses. This data is extracted from a finite element model of Train-Track Interaction (TTI) system. A generalised form of Cross-Industry Standard Process for Data Mining (CRISP-DM) tool is used to detect wheel flat based on K-Nearest Neighbour (kNN) algorithm.

KEY WORDS: Structure condition monitoring; Data mining; Wheel flat detection; Railway vehicle; Transport infrastructure.

1 INTRODUCTION

The growing significance of railway transportation infrastructure is driving a focus on cost-effectiveness, particularly in operation and maintenance [1]. In this direction, the assessment of loads imposed by railway traffic on the infrastructure is a major concern for railway administrations in order to avoid overloaded conditions that can induce damage to railway tracks [2]. Defective wheels in railway wagons are identified as one of the most significant sources of damage to railway infrastructure and vehicles [3], [4]. Impact loads due to a defective wheel, which depend on the depth and length of the wheel flat, can be several times higher than the static load of the wheel [5]. Wheel flats are the main causes of wheel bearing damage, increase in axle temperature, axle fractures, as well as rail and concrete sleeper fractures [6]. Therefore, efficient and accurate identification of wheel flats may prevent significant damages that could lead to traffic interruptions as well as guarantee train operation safety [7], [8]. In recent years, researchers have introduced both onboard and wayside measurement techniques for detecting wheel defects. Onboard methods typically involve installing sensors directly onto the wheels to achieve comprehensive damage diagnosis and efficient wheel condition management. However, this approach is infrequently employed due to its maintenance challenges. Additionally, onboard detection techniques are predominantly utilised for monitoring track conditions rather than assessing wheel conditions. Conversely, wayside measurement systems have emerged as a superior solution for identifying wheel defects, as they allow for the assessment of all wheels' conditions during the passage of in-service trains [9].

With rapid growth of database technology, more data are collected. Obviously, there is a lot of hidden important information behind collected data. Conventional approaches like typical database processing and classical mathematical statistical methods are common tools for knowledge discovery; however, they have many drawbacks. These include difficulties in meeting real-world assumptions, time-consuming processes, and a focus solely on simplified quantitative analysis, which may not effectively address real-world problems [10]. Therefore, the analysis of information should be done in a more in-depth level in order to better make use of databases [11]. One of the powerful computing tools is data mining which is used to obtain the knowledge from raw data. In fact, data mining is useful for complex and time consuming problems that cannot be solved with traditional statistical techniques [12].

A systematic and applicable tool is required to implement the data mining analysis in structural condition monitoring [13]. Currently, a number of data mining tools exist such as DMAIC (Define, Measure, Analyse, Improve, and Control), SEMMA (Sample, Explore, Modify, Model, and Assess), Cross-Industry Standard Process for DM (CRISP-DM), etc has been developed. [14]–[19]. Amongst these tools, CRISP-DM has the most extensive application [20].

Implementation of data mining in different areas of civil engineering has recently given very good results. However, the application of data mining in structural health monitoring is not used as much as expected, thus, many challenges are still ahead. Therefore, it seems a vital need is required to develop the applicability of data mining in condition monitoring [21]. To this end, in this study, a generalised form of CRISP-DM is developed to investigate the applicability of data mining for damage detection of a Train-Track Interaction (TTI) system using K-nearest neighbour (kNN).

2 DATA MINING

Data can be defined as any fact, number or text which can be proceed by a computer [22]. As the obtained pattern through data mining may be very difficult to find, it is sometimes compared to gold mining in rivers. The term "gold mining" refers to the search for gold in rocks or sand. The origination of data mining traces back to the development of artificial intelligence in 1950s [23].

2.1 Data Mining Functions and Techniques

Data mining can help in making a proactive decision based on knowledge by means of forecasting future plan. In general, data mining can be categorized into two groups, i.e., descriptive mining, and predictive mining. Each of these groups has its specific functions, (i.e., clustering, prediction, classification, exploration and association) [24] (see Figure 1). The purpose of clustering is to divide the samples into groups with related behaviour. The numerical prediction activity determines patterns, rules, or models to predict continuous or discrete target values which can also be used for other functions. Classification is used to recognize several rules which can be applied in future work to determine whether a previously unknown item belongs to a known class. Exploration is used to find out dimensionality of an input data and, eventually, the association activity is used to frequently detect occurring related objects. Based on their particular utilisations in consequence of their assumptions and drawbacks, one or combination of some of these tasks can be used to find the hidden information [25]–[27].
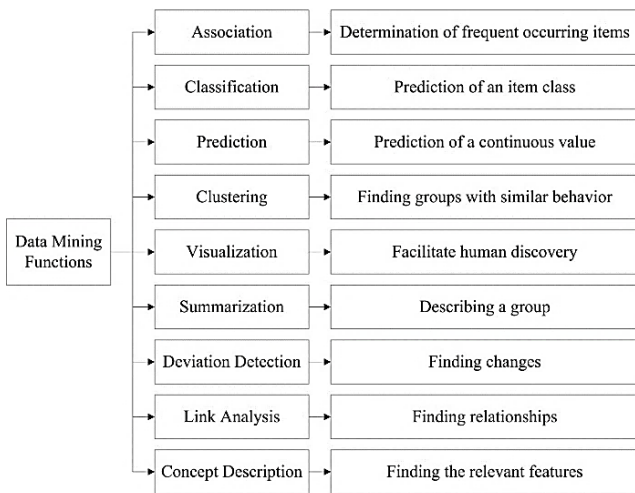

Figure 1. Data mining functions

In general, there are three types of data mining techniques which consist of statistical techniques, machine learning techniques and artificial intelligence techniques. Each of these techniques has particular algorithms for running the models to get the best solution. For example, statistical methods include regression analysis, clustering analysis, decision tree, etc., and machine learning methods such as principal component analysis, support vector machine, K-nearest neighbour, and so forth. Another category with various algorithms is artificial intelligence, e.g., fuzzy logic, genetic algorithm, artificial neural network and particle swarm algorithm [28].

2.2 Data Mining Models

Data mining process has several types of tools, e.g. Knowledge Discovery in Databases (KDD), SEMMA, DMAIC, CRISP-DM, etc. [14]–[19]. KDD process identifies novel, valid, potentially useful, and understandable forms of data. This model in presented by [16]. Generally, KDD defines the main procedures to transform the raw data into beneficial knowledge [29]. The objective of this process is to make sure that any outcome is accurate and reproducible [30]. The KDD process includes five steps which are data selection, data preparation, data transformation, data mining and interpretation. SEMMA is the abbreviation of Sample, Explore, Modify, Model, and Assess, which shows the data mining process. The SAS Institute presented a 5-stage cycle for this model [31].

DMAIC is the abbreviation of Define, Measure, Analyse, Improve and Control, which shows the procedure of the model [32]. DMAIC is a six-sigma problem-solving approach in order to remove a number of challenges, i.e., waste, defects, or quality-control problems [33]. In brief, this model can be defined as follows:

- Define the goals of the improvement activity,
- Measure the existing system,
- Analyse the system to identify ways to eliminate the gap between the current performance of the system or process and the desired goal,
- Improve the system,
- Control the new system.

CRISP-DM was introduced by a consortium of several companies such as National Cash Register (NCR) System Engineering Copenhagen from USA and Denmark, Integral Solutions Ltd. (ISL)/SPSS from USA, Daimler Chrysler AG from Germany and an insurance corporation in Netherlands, called OHRA [17], [34]. The most widespread application amongst the tools is CRISP-DM. This model has a 6-stages cycle which is described below [17], [19].

- Business understanding: Understanding of the project objectives and requirements,
- Data understanding: Getting familiar with data and initial data collection,
- Data preparation: Constructing the dataset (see Figure 2),
- Modelling: Selecting and applying modelling technique(s),
- Evaluation: Reviewing the constructed model(s),
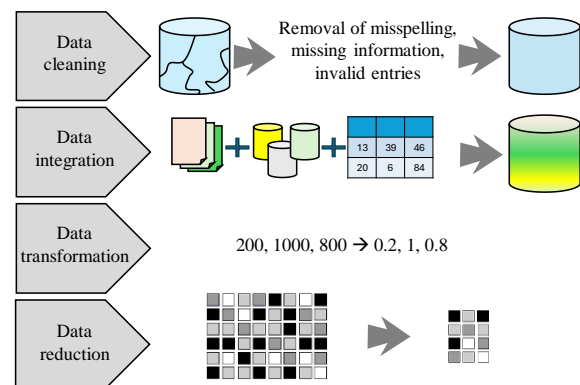- Deployment: Organizing and presenting the achieved knowledge.

Figure 2. Forms of data preparation

There are other data mining tools which can be used for knowledge discovery, as follows:

- A five-step approach was introduced by [35]. Business objectives determination, data preparation, data mining, domain knowledge elicitation, and assimilation of knowledge are the steps of this model.
- Anand & Büchner [15] proposed the another model in eight steps, i.e., human resource identification, problem specification, data prospecting, domain knowledge elicitation, methodology identification, data preprocessing, pattern discovery, and knowledge post-processing.
- A six-step discovery process was proposed by [36] for clinical diagnostic images.

3 METHODOLOGY

3.1 A Systematic Data-Driven Approach

One of the approaches in condition monitoring consists of two major components, i.e., a network of sensors to collect the response data and data mining to extract information on the structural health condition. In the same context, with development of data inverse analysis, data mining is combined with wheel defect detection of railway vehicles to improve the application of obtained data in complicated structural systems.

To achieve an effective data mining analysis, a systematic model and appropriate data mining method(s) are required. Therefore, by taking advantage of the CRISP-DM model, this study develops a generalised form of this tool that integrates the railway condition monitoring system (see Figure 3).
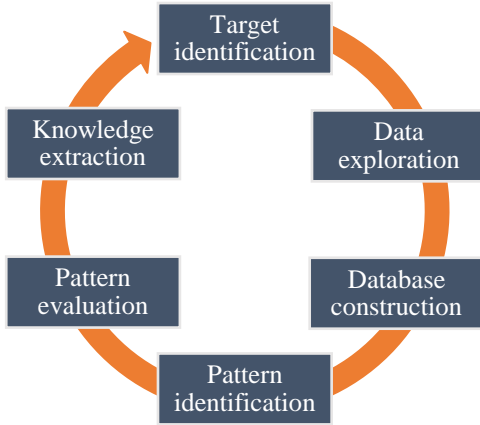

Figure 3. Generalised form of CRISP-DM for structural damage detection

This applied data mining-based damage detection approach consists of six defined stages, i.e., target identification, data exploration, database construction, pattern identification, pattern evaluation and knowledge extraction. At the first stage, specimen description and experimental setup is presented. In the second stage, sensed data is collected from a healthy (baseline) and damaged train wheel. Analysis of collected data is carried out in the third stage to generate a dataset using the acceleration values of different damage cases as inputs for next stage. In the pattern identification phase, applicable machine learning, artificial intelligence, and statistical algorithms can be used to estimate the damage. In this paper, the kNN algorithm is employed to train the data and build a model for damage detection. Then, in the fifth stage, model performance is assessed using evaluation of the model. Finally, the last stage extracts valuable knowledge, i.e., damage classification.

3.2 K-Nearest Neighbour (KNN)

The kNN is an instance-based learning algorithm which can obtain a set of k objects, so called the k-nearest neighbours. The kNN utilises the mainstream vote to determine the specific class in the neighbourhood. To explain this, kNN comprises three components, i.e., (1) a group of labelled samples, (2) a distance metric among samples, and (3) k-nearest neighbour value. The distance between unlabelled and labelled query instances is calculated. At that point, the k values can be defined. Then, the class label of k-nearest neighbours is employed to finalize the class label of the sample. The following Equations present the different distance functions of the algorithm. [37]–[39].

$$Euclidean\ Distance\ Function = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \qquad (1)$$

$$Manhattan\ Distance\ Function = \sum_{i=1}^{k}|x_i - y_i| \qquad (2)$$

$$Minkowski\ Distance\ Function = \left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q} \qquad (3)$$

$$Cosine\ Distance\ Function = \frac{\sum_{i=1}^{k}x_i y_i}{\sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}} \qquad (4)$$

where, *x* and *y* represent the proportion of observations in group *k* in the nearest neighbour, *q>0*.

The kNN algorithm is robust to noisy data. Its computational complexity can be expressed by $\mathcal{O}(n * k * d)$, where *n, k,* and *d* represent the number of training examples, the hyperparameter for the kNN, and the dimension of the input space, respectively. However, its high computational cost in the calculation of distance metric can be a limitation for all databases [37]–[39].

4 RESULTS AND DISCUSSION

In this paper, a dataset was generated from modelling the dynamics of train-track interaction for research and analysis purposes [40]. As presented in Figure 4, one wheel is considered as a damaged wheel. Figure 4 also shows the details of the modelling including parameters of the model, and a set of 12 sensors mounted on the rail. The collected sensor dataset is considered as input for developing defect detection algorithm. Table 1 illustrates the undamaged (benchmark) and three damaged wheel scenarios (i.e., Low damage (LD), Medium Damage (MD), and Severe Damage (SD)) and Figure 5 shows the effect of different wheel flat severities on acceleration response of sensor 1.
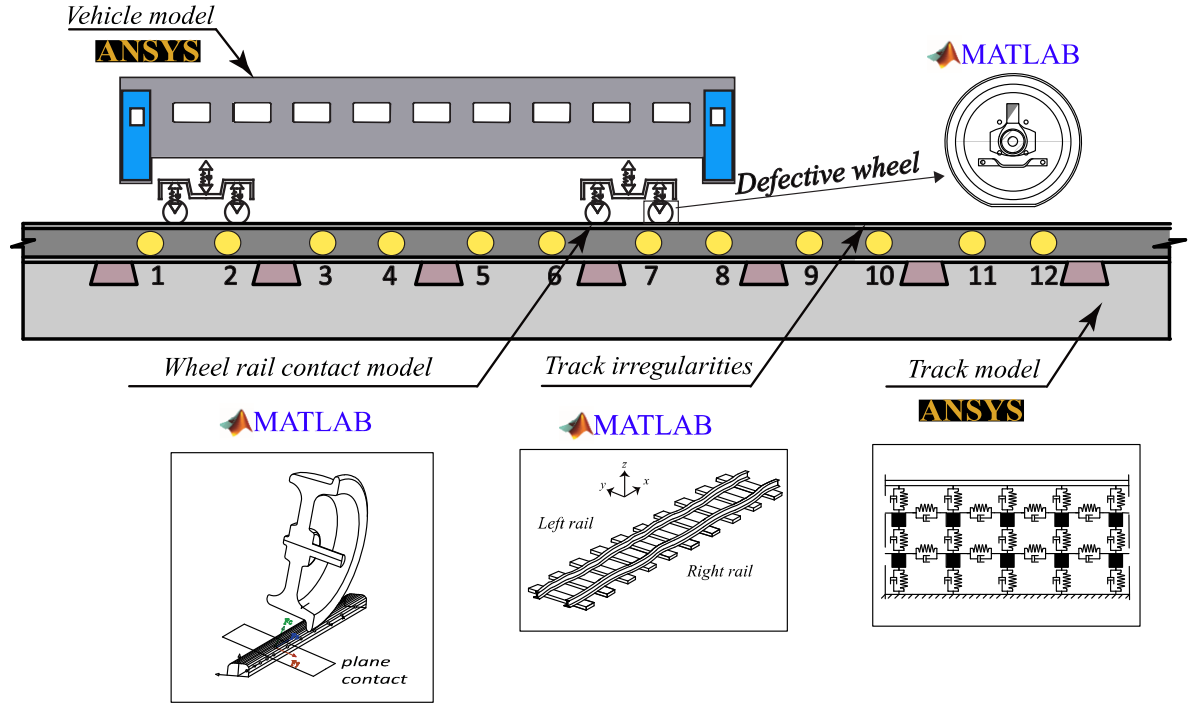
Figure 4. Schematic illustration of the train-track dynamic interaction and sensor placement

Table 1. Benchmark and damaged scenarios.

	Benchmark scenarios	Damaged scenarios
Irregularities profile	10	1 per each train
Speeds	20-60 m/s	60 m/s for Alfa Pendular
Noise ratio	5%	5%
Flat characteristics	–	Flat lengths (L): LD: 19.5 mm, MD: 36 mm, and SD: 82 mm

To apply the kNN algorithm, k-fold cross-validation was chosen due to its simplicity, ease of use, and ability to utilise all data for training and validation. The collected datasets from all 12 sensors were randomly divided into five subsets, i.e., four for training and one for testing. The healthy and defected data samples consisted of 100 and 150, respectively. After several trials, the k value of the kNN model, which represents the number of neighbours, was selected as 10. The kNN model was built using three chosen predictors, and the 'cosine' metric was employed for distance computation. The mean of the nearest neighbour values was used to predict the target range. The classification accuracy (CA) based on the confusion matrix was used to assess the detection capability of method, which is the percentage of samples correctly classified and evaluated as:

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \qquad (5)$$

where, TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively. Table 2 compares the performance across the 12 measurement positions

in two schemes (i.e., binary, and multi-class classification). In binary classification, the aim is to detect the presence of the damage. However, in multi-class classification, severity of defect can also be detected. As can be seen from this table, the average binary defect detection CA is 95.2% and the fourth sensor achieved the highest accuracy. This is likely because it is positioned closest to the wheel crack. Nevertheless, when employing multi-class classification, the maximum accuracy dropped to 80%, indicating that while the kNN algorithm is highly accurate in detecting damage, it presents lower accuracy in classifying the severity of the damage. One possible explanation for this accuracy reduction is the nature of the input dataset. Since we utilised raw acceleration data in the paper, this data representation might not be ideal for classifying severity. Therefore, it is necessary to consider using a higher level of data representation, such as the time-frequency domain [41], as input for the kNN algorithm.
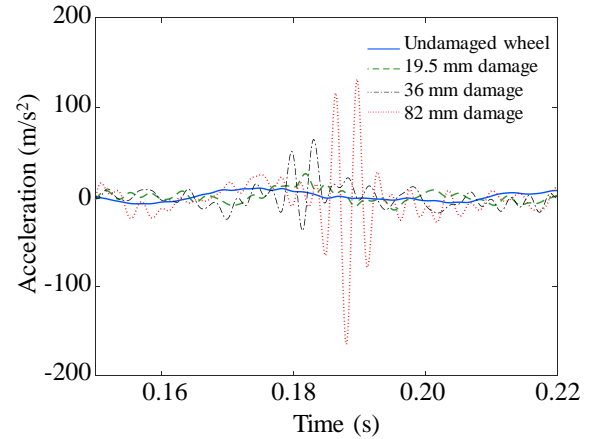

Figure 5. Acceleration time series signals for different wheel flat lengths

Table 2. Classification accuracy of sensors

Sensor Number	Binary CA	Multi-class CA
Sensor 1	0.9160	0.7640
Sensor 2	0.9240	0.7840
Sensor 3	0.9840	0.7880
Sensor 4	**0.9920**	**0.8000**
Sensor 5	0.9800	0.7880
Sensor 6	0.9640	0.7800
Sensor 7	0.9440	0.7480
Sensor 8	0.9440	0.7720
Sensor 9	0.9360	0.7720
Sensor 10	0.9440	0.7960
Sensor 11	0.9480	0.7720
Sensor 12	0.9520	0.7640

Figure 6 shows the output of the kNN algorithm applied to data from the fourth sensor. The results indicate that the kNN model achieved a high overall CA of 99% in defect detection of the train wheel.
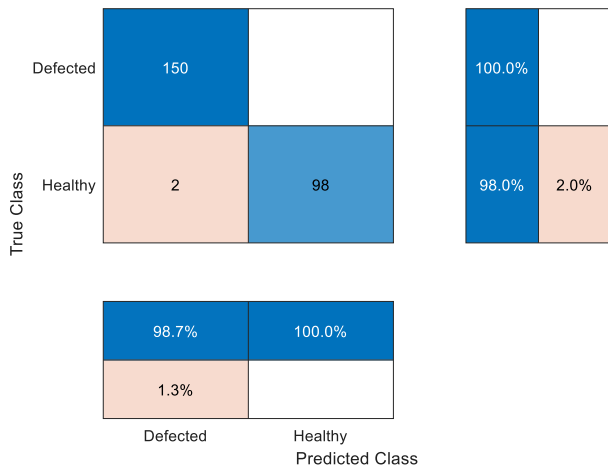

Figure 6. Confusion matrix of defect detection using kNN for sensor 4

5 CONCLUSION

This paper presented a data mining-based approach for detecting wheel flats in railway vehicles using the kNN algorithm. The approach utilised sensor data from a finite element model of the TTI system and a generalised form of the CRISP-DM framework for data mining analysis. The results demonstrate the effectiveness of the developed approach in binary classification for wheel flat detection. The kNN model achieved a high classification accuracy (CA) of 99% using data from the sensor closest to the wheel. However, multi-class classification for severity levels resulted in a lower CA (i.e., 80%). This shows that raw acceleration data might not be an ideal input for detecting the severity of the wheel flat.

REFERENCES

[1] M. Mohammadi, A. Mosleh, C. Vale, D. Ribeiro, P. Montenegro, and A. Meixedo, "An Unsupervised Learning Approach for Wayside Train Wheel Flat Detection," *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23041910.

[2] A. Mosleh, P. A. Costa, and R. Calçada, "A new strategy to estimate static loads for the dynamic weighing in motion of railway vehicles," *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit*, vol. 234, no. 2, pp. 183–200, 2020, doi: 10.1177/0954409719838115.

[3] A. Mosleh, A. Meixedo, D. Ribeiro, P. Montenegro, and R. Calçada, "Machine learning approach for wheel flat detection of railway train wheels," *Transp. Res. Procedia*, vol. 72, no. 2022, pp. 4199–4206, 2023, doi: 10.1016/j.trpro.2023.11.354.

[4] A. Mosleh, A. Meixedo, D. Ribeiro, P. Montenegro, and R. Calçada, "Automatic clustering-based approach for train wheels condition monitoring," *Int. J. Rail Transp.*, vol. 11, no. 5, pp. 639–664, 2023, doi: 10.1080/23248378.2022.2096132.

[5] A. Mosleh, P. Montenegro, P. Alves Costa, and R. Calçada, "An approach for wheel flat detection of railway train wheels using envelope spectrum analysis," *Struct. Infrastruct. Eng.*, vol. 17, no. 12, pp. 1710–1729, 2021, doi: 10.1080/15732479.2020.1832536.

[6] Z. Zhang, S. Wei, B. Andrawes, D. A. Kuchma, and J. R. Edwards, "Numerical and experimental study on dynamic behaviour of concrete sleeper track caused by wheel flat," *Int. J. Rail Transp.*, vol. 4, no. 1, pp. 1–19, 2016, doi: 10.1080/23248378.2015.1123657.

[7] A. Mosleh, P. A. Costa, and R. Caçada, "Development of a Low-Cost Trackside System for Weighing in Motion and Wheel Defects Detection," *Int. J. Railw. Res.*, vol. 7, no. 1, pp. 1–9, 2020, [Online]. Available: http://ijrare.iust.ac.ir/article-1-268-en.html

[8] C. Vale, "Wheel flats in the dynamic behavior of ballasted and slab railway tracks," *Appl. Sci.*, vol. 11, no. 15, 2021, doi: 10.3390/app11157127.

[9] B. Pintão, A. Mosleh, C. Vale, P. Montenegro, and P. Costa, "Development and Validation of aWeigh-in-Motion Methodology for Railway Tracks," *Sensors*, vol. 22, no. 5, 2022, doi: 10.3390/s22051976.

[10] W.-W. Wu, Y.-T. Lee, M.-L. Tseng, and Y.-H. Chiang, "Data mining for exploring hidden patterns between KM and its performance," *Knowledge-Based Syst.*, vol. 23, no. 5, pp. 397–401, 2010.

[11] J. Lingxia, "Research on Distributed Data Mining System and Algorithm Based on Multi-Agent," University of Quebec, 2009.

[12] D. W. M. Hofmann and J. Apostolakis, "Crystal structure prediction by data mining," *J. Mol. Struct.*, vol. 647, no. 1–3, pp. 17–39, 2003.

[13] M. Gordan *et al.*, "Data mining-based damage identification of a slab-on-girder bridge using inverse analysis," *Measurement*, vol. 151, p. 107175, 2020, doi: 10.1016/j.measurement.2019.107175.

[14] A. Azevedo and M. F. Santos, "KDD, SEMMA AND CRISP-DM:A Parallel Overview," in *IADIS European Conference Data Mining*, IADIS, 2008, pp. 182–185.

[15] S. S. Anand and A. G. Büchner, *Decision Support Using Data Mining*. London: Financial Times Management, 1998.

[16] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "Knowledge Discovery and Data Mining : Towards a Unifying Framework," *KDD*, vol. 96, pp. 82–88, 1996.

[17] P. Chapman *et al.*, "CRISP-DM 1.0 Step-by-step data mininng guide," 2000.

[18] L. A. Kurgan and P. Musilek, "A survey of Knowledge Discovery and Data Mining process models," *Knowl. Eng. Rev.*, vol. 21, no. 1, pp. 1–24, 2006.

[19] S. C. Chen and M. Y. Huang, "Constructing credit auditing and control & management model with data mining technique," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5359–5365, May 2011.

[20] M. Saltan, S. Terzi, and E. Ug, "Backcalculation of pavement layer moduli and Poisson's ratio using data mining," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2600–2608, 2011.

[21] M. Gordan, H. A. Razak, Z. Ismail, K. Ghaedi, Z. X. Tan, and H. H. Ghayeb, "A hybrid ANN-based imperial competitive algorithm methodology for structural damage identification of slab-on-girder bridge using data mining," *Appl. Soft Comput. J.*, vol. 88, p. 106013, 2020, doi: 10.1016/j.asoc.2019.106013.

[22] T. Yang, Ping-Bai, and Y.-S. Gong, "Spatial Data Mining Features between General Data Mining," *2008 Int. Work. Educ. Technol. Train. 2008 Int. Work. Geosci. Remote Sens.*, pp. 541–544, 2008.

[23] M. Gordan *et al.*, "State-of-the-art review on advancements of data mining in structural health monitoring," *Measurement*, vol. 193, p. 110939, 2022, doi: 10.1016/j.measurement.2022.110939.

[24] T. Pang-Ning, M. Steinbach, and V. Kumar, *Introduction to data mining*. Boston: Pearson Addison-Wesley, 2006. doi: 10.1016/0022-4405(81)90007-8.

[25] T. Y. Chen and J. H. Huang, "Application of data mining in a global optimization algorithm," *Adv. Eng. Softw.*, vol. 66, pp. 24–33, 2013.

[26] M. K. Obenshain, "Application of data mining techniques to healthcare data.," *Infect. Control Hosp. Epidemiol.*, vol. 25, no. 8, pp. 690–695, Aug. 2004.

[27] S.-H. Liao, P.-H. Chu, and P.-Y. Hsiao, "Data mining techniques and applications – A decade review from 2000 to 2011," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 11303–11311, Sep. 2012.

[28] S. Saitta, B. Raphael, and I. F. C. Smith, *Data Mining : applications in civil engineering*. Saarbrücken: VDM, 2009.

[29] T. Miranda, A. G. Correia, M. Santos, L. Ribeiro, and P. Cortez, "New Models for Strength and Deformability Parameter Calculation in Rock Masses Using Data-Mining Techniques," *Int. J. Geomech.*, vol. 11, no. 1, pp. 44–58, 2011.

[30] R. B. Buchheit, J. H. Garrett, Jr., S. R. Lee, and R. Brahme, "A Knowledge Discovery Case Study for the Intelligent Workplace," *Comput. Civ. Build. Eng.*, pp. 914–921, Aug. 2000.

[31] X. Zhu, "Agile mining : a novel data mining process for industry practice based on Agile Methods and visualization," Master dissertation, University of Technology Sydney, 2017. [Online]. Available: https://opus.lib.uts.edu.au/handle/10453/123178

[32] D. Stevens, "The Leveraging Effects of Knowledge Management Concepts In the Deployment of Six Sigma in a Health Care Company," PhD Thesis, Walden University, 2006.

[33] T. Pyzdek and P. Keller, *The Six Sigma Handbook*, Third Edit., vol. 35, no. 9. New York, 2002.

[34] I. B. Fernandez, S. H. Zanakis, and S. Walczak, "Knowledge discovery techniques for predicting country investment risk," *Comput. Ind. Eng.*, vol. 43, no. 4, pp. 787–800, 2002.

[35] P. Cabena, P. Hadjinian, R. Stadler, J. Verhees, and A. Zanasi, *Discovering data mining: from concept to implementation*. Prentice-Hall, Inc. Upper Saddle River, NJ, 1998.

[36] K. J. Cios, A. Teresinska, S. Konieczna, J. Potocka, and S. Sharma, "A knowledge discovery approach to diagnosing myocardial perfusion," *IEEE Eng. Med. Biol. Mag.*, vol. 19, no. 4, pp. 17–25, 2000.

[37] H. Son, C. Kim, N. Hwang, C. Kim, and Y. Kang, "Advanced Engineering Informatics Classification of major construction materials in construction environments using ensemble classifiers," *Adv. Eng. Informatics*, vol. 28, no. 1, pp. 1–10, 2014, doi: 10.1016/j.aei.2013.10.001.

[38] I. L. Martínez, "Structural Damage Assessment under Uncertainty," PhD Dissertation, University of California, 2010.

[39] X. Wu *et al.*, *Top 10 algorithms in data mining*, vol. 14, no. 1. 2007. doi: 10.1007/s10115-007-0114-2.

[40] A. Guedes *et al.*, "Detection of Wheel Polygonization Based on Wayside Monitoring and Artificial Intelligence," *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23042188.

[41] T. Buckley, B. Ghosh, and V. Pakrashi, "A Feature Extraction & Selection Benchmark for Structural Health Monitoring," *Struct. Heal. Monit.*, vol. 22, no. 3, pp. 2082–2127, 2023, doi: 10.1177/14759217221111141.