**MACHINE LEARNING FOR SAFETY OF RAILWAYS**

**Ankit Anand1 Sanskruti Mali2 Vidya Gaikwad3**

**Computer Engineering, Vishwakarma Institute of Information Technology, Pune**

**ABSTRACT**

Railway networks are one of the most popular and widely utilized forms of public transportation. Although safety is the most important factor in railway systems, accidents still happen because of a partial lack of safety measures. For the safety and wellbeing of those who rely on public transit, it is crucial to reduce these errors by analyzing accidents and improving safety systems utilizing cutting-edge technology like machine learning (ML).Machine learning in security regularly picks up new skills by looking for patterns in data to uncover insider risks, improve safety systems, and better detect malware[1]. So, in the context of safety in the railway industry, we must take into account the technology that is currently accessible in general and ML in particular [2]. In order to identify a pattern in earlier incidents and categorize them appropriately, this study applies the Decision Tree(DT) method for safety and accident analysis. This study's key contribution is the application of ML and a justification of how it improves automated processes, ensures safety, and yields advantages. This paper focuses on five case studies that include information on injuries and fatalities brought on by rail accidents. In order to provide a more comprehensive overview of the application of supervised ML for raising safety standards in railways, an investigation of some of these accidents as reported by INDIAN RAILWAYS [3] is conducted and published in this work.

**INTRODUCTION**

Technology has witnessed impressive evolution in the past few decades, which has in turn transformed our lives and helped us evolve with it. Right from roadways, railways, and aircraft for seamless travel to making communication effortless from any part of the world, technology has contributed more than anything to help mankind live a life of luxury and convenience[[4]](https://www.simplilearn.com/importance-of-technology-article#:~:text=Cost%20and%20Time%2DSaving,Technology%20also%20ensures%20improved%20accuracy.). It is because of technology that everyone is aware of their surroundings in a much more detailed manner. Technology has made an advancement in almost every field and it is quite evident and can be easily seen in everyday life. Efficiency gains, quicker decision-making, time and money savings, a competitive advantage, and greater innovation are all made possible by technology. Today, technology is a necessary component of our everyday existence. Every action we take, from the beginning of the day to its conclusion, involves technology.

Artificial Intelligence (AI) is one of the most revolutionary technologies of the 21st century, significantly transforming various sectors of society and the job market. Since its inception, AI has rapidly evolved, promoting paradigmatic changes that impact the way we live and work[[5]](https://www.linkedin.com/pulse/revolution-artificial-intelligence-ai-21st-century-de-oliveira/).

Artificial intelligence, which began in the 1950s with the question “Can computers think?”, has become a modern concept that means “a computing-based technology system that automates intelligent tasks generally performed by ordinary people”[[6](https://scholar.google.com/scholar_lookup?title=Artificial+Intelligence:+A+New+Synthesis&author=Nilsson,+N.J.&author=Nilsson,+N.J.&publication_year=1998),[7](http://www-formal.stanford.edu/jmc/whatisai.html),[8](https://scholar.google.com/scholar_lookup?title=Artificial+intelligence+and+natural+man&author=Boden,+M.&publication_year=1980&journal=Synthese&volume=43&pages=433%E2%80%93451)]. Artificial intelligence has been widely applied to various railway safety issues, such as railway infrastructure, trains, operations, and stations[[9]](https://www.mdpi.com/2076-3417/12/20/10572#B1-applsci-12-10572).

Numerous studies have been conducted on railway safety and related methodologies; of these, machine learning and deep learning have shown to be highly effective and yield excellent results with regard to safety-related issues.There are several reasons why railway safety defects can arise, including anomalous weather, track problems, malfunctioning systems, or human error of any kind. Regular inspection is crucial to prevent railway accidents caused by various errors.Recently, many studies have been conducted to detect defects in railways and related parts with rapidly developed artificial intelligence technology to prevent railway accidents[[9]](https://www.mdpi.com/2076-3417/12/20/10572#B1-applsci-12-10572). Diverse data sources have been used for detecting railway defects, such as images [[10](https://scholar.google.com/scholar_lookup?title=Convolutional+Neural+network+based+Online+Rail+surface+Crack+Detection&conference=Proceedings+of+the+2021+5th+International+Conference+on+Intelligent+Computing+and+Control+Systems+(ICICCS)&author=Akhila,+C.&author=Diamond,+C.A.&author=Posonia,+A.M.&publication_year=2021&pages=1602%E2%80%931606),[11](https://scholar.google.com/scholar_lookup?title=Railroad+Surface+Defect+Segmentation+Using+a+Modified+Fully+Convolutional+Network&author=Kim,+H.&author=Lee,+S.&author=Han,+S.&publication_year=2020&journal=KSII+Trans.+Internet+Inf.+Syst.+TIIS&volume=14&pages=4763%E2%80%934775)], accelerometers [[12](https://scholar.google.com/scholar_lookup?title=Condition+monitoring+of+railway+tracks+from+car-body+vibration+using+a+machine+learning+technique&author=Tsunashima,+H.&publication_year=2019&journal=Appl.+Sci.&volume=9&pages=2734&doi=10.3390/app9132734),[13](https://scholar.google.com/scholar_lookup?title=Developing+machine+learning-based+models+for+railway+inspection&author=Yang,+C.&author=Sun,+Y.&author=Ladubec,+C.&author=Liu,+Y.&publication_year=2020&journal=Appl.+Sci.&volume=11&pages=13&doi=10.3390/app11010013)], and ultrasonic sensors [[14](https://scholar.google.com/scholar_lookup?title=Fast+classification+for+rail+defect+depths+using+a+hybrid+intelligent+method&author=Jiang,+Y.&author=Wang,+H.&author=Tian,+G.&author=Yi,+Q.&author=Zhao,+J.&author=Zhen,+K.&publication_year=2019&journal=Optik&volume=180&pages=455%E2%80%93468&doi=10.1016/j.ijleo.2018.11.053),[15](https://scholar.google.com/scholar_lookup?title=Deep+learning-based+acoustic+emission+scheme+for+nondestructive+localization+of+cracks+in+train+rails+under+a+load&author=Suwansin,+W.&author=Phasukkit,+P.&publication_year=2021&journal=Sensors&volume=21&pages=272&doi=10.3390/s21010272)].

Machine learning (ML) is a technique for discovering information with self-learning techniques[16], and it has been used in every field due to its ability to obtain

useful information from large sets of data [17]. This paper applies machine learning techniques to provide a thorough review of railway safety. We have concentrated on decision trees as machine learning techniques. Decision trees are a popular machine learning model that can be applied to both regression and classification models. This paper primarily focuses on addressing track maintenance, train speed, weather, emergency breaking systems, station safety (including falls, slips, and trips from platform edges), and safety near tunnels and mountains.

**LITERATURE SURVEY**

| **Sr. no** | **Author name** | **DOI** | **Author Work** |
| --- | --- | --- | --- |
| **1** | Oh K | 10.3390/app122010572 | In order to improve railway safety, this paper investigates the use of artificial intelligence, particularly deep learning, with a focus on four domains: train body and bogie, railway infrastructure, operation, and station. In order to solve safety concerns, emphasis is placed on defect detection in various train components, with a special focus on high-speed trains because of their exposure to harsher environments. |
| **2** | Kaewunruen | 10.1109/ACCESS.2019.2962072 | This study promotes the use of machine learning, more especially the decision tree method, to analyze accidents and forecast passenger characteristics that may be impacted by them in order to improve safety at railway stations. The study emphasizes how crucial machine learning is to automating safety procedures and realizing its full potential for raising rail safety standards. |
| **3** | Serker I.H | 10.1007/s42979-021-00592-x | In order to analyze a wide range of data sets from cybersecurity, mobile, social media, business, IoT, and health, this study examines the critical role that machine learning plays. It seeks to give professionals and academics a fundamental framework for creating data-driven, intelligent automated systems for a range of real-world uses in the Fourth Industrial Revolution. |
| **4** | Çelik Özer | 10.31681/jetol.457046 | Among artificial intelligence, machine learning first appeared in the 1950s, though its early developments were modest. Resurrected in the 1990s, machine learning is a field that is constantly evolving and has a wide range of applications. Current research is driven by the difficulty of processing large amounts of data and asserts that the best models for new data can be found within existing data. |
| **5** | Kift Ryan | 10.1016/j.ssci.2018.07.004 | Rapid technological advancements have made the railway industry more concerned about safety, which calls for a reevaluation of system integration risks. This paper focuses on human-automation design and progressive integration efforts in the railway industry, highlighting potential health and safety hazards. |
| **6** | Vidumina Navod | - | By utilizing GPS data to identify danger zones and suggest speed reductions to drivers, a machine learning module improves road safety. The technology uses GPS data from elephants to determine a train's proximity in real-time, alerting drivers to avoid collisions and guaranteeing passenger safety. |

Railway accidents can cause significant damage because they swiftly move large numbers of people and goods across the country. But because railway infrastructure is so intricate, it is difficult and costly to maintain. Therefore, a lot of researchers are interested in applying artificial intelligence to railway safety. In this paper, deep learning approaches are primarily examined as applications of artificial intelligence for railway safety. In order to improve railway safety, deep learning techniques are first introduced in this paper. Next, we examined and categorized previous research into four example application domains: train body and bogie (door, wheel, suspension, bearing, etc.); railway infrastructure (catenary, surface, components, and geometry); operation (railway detection, railroad trespassing, wind risk, train running safety, etc.); and station (air quality control, accident prevention, etc.). Numerous studies have also been done on the use of artificial intelligence for train defect detection. The structure of a railway train is intricate and comprises many different components, including wheels, split pins, tram lines, and pantographs. Because every train accessory has a different function and environment, each one has a different degree of corrosion or durability decrease. Because even small-area or early-progress defects (cracks, cuts, aging, etc.) on trains can pose serious risks to passenger safety, fault detection and prediction are crucial. Compared to other trains, high-speed trains' components are subjected to harsher environments, so these features may be even more important.[27].

Even though station safety is a vital component of the overall design of railway systems, station accidents still happen. It's time to learn from these mistakes and advance traditional techniques by analyzing accidents and strengthening safety systems with the help of cutting-edge technologies like machine learning (ML).In order to forecast the characteristics of passengers impacted by accidents, this study investigates the use of the decision tree (DT) method in safety classification and accident analysis at railroad stations. The presentation of machine learning (ML) and an explanation of its application for ensuring safety, leveraging automated processes, and reaping the benefits of this potent technology constitute the study's crucial contribution.[28].

The digital world is rife with data in this Fourth Industrial Revolution (4IR) or Industry 4.0 era. Examples of this data include cybersecurity, mobile, social media, business, Internet of Things (IoT), and health data. The key to developing intelligent analyses of these data and correspondingly intelligent and automated applications is understanding artificial intelligence (AI), and specifically machine learning (ML). This study's primary contribution is its explanation of the principles and potential of various machine learning techniques, as well as how these techniques can be applied in the previously mentioned range of real-world application areas. Therefore, the goal of this paper is to offer a basic framework for academics and professionals in the field who wish to investigate, explore, and create data-driven automated and intelligent systems in the pertinent domains using machine learning techniques[29].

In the 1950s, machine learning emerged as a scientific field within artificial intelligence. The earliest attempts at machine learning date back to the 1950s, but during that time there were few noteworthy advances and studies in the field. Nonetheless, research in this area was revived, expanded, and continues to this day in the 1990s. This is a science that will continue to advance. This development is due to the challenge of processing and analyzing the ever-increasing amount of data. The foundation of machine learning is the idea that, because of this growing amount of data, the best model for the new data can be found among the old data. As a result, research on machine learning will continue alongside the growth of data. This study covers the background, techniques, and application domains of machine learning as well as current research in the field. The purpose of this study is to educate researchers on machine learning, a field that has gained a lot of popularity recently, and its applications[30].

The railway industry is undergoing a significant technological change that is raising new safety issues. The rate at which technological advancements and change are happening is noteworthy for an industry where mistakes can have serious repercussions. This rapid advancement in technology has led to a reevaluation of the literature on the risks involved in system integration in industries that are high-risk or safety-critical. This paper seeks to increase public awareness of potential health and safety hazards that may arise in the railway sector. The human-automation design and progressive integration efforts are two risk mechanisms that are discussed that could affect the success of an integrated railway.[31].

Using GPS data collected previously, a machine learning module is trained to identify common danger spots and recommend to the driver to reduce speed. Here is a better way to track the elephant using its GPS belt, which provides its exact location. You can use the GPS to determine the distance between the elephant's actual location and the train path, and then notify the train driver to take the appropriate action. We can guarantee passenger safety and prevent accidents to a great extent by implementing these features in real-time applications. The goal of this project is to create and evaluate a machine learning-based safety system for identifying and preventing train accidents. The system looks for patterns that might point to a higher chance of an accident or malfunction using sensor data from trains and other pertinent sources[32].

**GAPS IDENTIFIED**

Critical gaps in railway safety can be addressed with revolutionary solutions provided by machine learning and artificial intelligence (AI). Conventional maintenance methods frequently lead to unplanned equipment breakdowns and reactive responses. A solution is provided by machine learning, which makes predictive maintenance possible by analyzing sensor readings, performance metrics, and maintenance history. By being proactive, possible safety risks are avoided and downtime is reduced. Additionally, by continuously scanning real-time data for anomalies, AI algorithms can close the gap in anomaly detection. These algorithms are particularly good at picking up on minute irregularities, like strange patterns in sensor data, temperature fluctuations, or unusual vibrations, which can be used to identify possible safety concerns in advance.

By examining a variety of data sources, machine learning also tackles the problem of anticipating and averting safety incidents. Through the utilization of past incident data and present operational parameters, machine learning models are capable of predicting the probability of safety-related incidents. This makes it possible to optimize safety procedures and put preventative measures into place to stop incidents. Infrastructure monitoring automation is an additional domain where artificial intelligence (AI) shines. Image processing and computer vision systems can effectively examine tracks and infrastructure elements, detecting possible safety risks faster than with manual inspections. These developments, along with AI's capacity to improve security through video analytics and optimize traffic management, lead to a more proactive, effective, and data-driven approach to railway safety, which eventually results in a safer and more dependable railway system.

**MACHINE LEARNING AND DECISION TREE BACKGROUND**

**MACHINE LEARNING**

A subfield of computer science and artificial intelligence (AI) called "machine learning" focuses on using data and algorithms to simulate human learning processes and progressively increase their accuracy.[[24]](https://www.ibm.com/topics/machine-learning).Three main categories apply to machine learning models:

Supervised machine learning

Supervised machine learning, or supervised learning, is characterized by the use of labeled datasets to train algorithms for precise outcome prediction or data classification. The model modifies its weights as input data is entered until an appropriate fit is achieved. This happens during the cross-validation process, which makes sure the model doesn't overfit or underfit. Sorting spam into a different folder from your inbox is just one example of the many real-world problems that supervised learning helps organizations solve at scale. Neural networks, naïve bayes, linear regression, logistic regression, random forest, and support vector machines (SVM) are a few techniques used in supervised learning..

Unsupervised machine learning

Unsupervised learning, sometimes referred to as unsupervised machine learning, is the process of analyzing and grouping unlabeled datasets using machine learning algorithms. These algorithms find hidden relationships or patterns in the data without requiring human assistance. This approach is perfect for customer segmentation, cross-selling tactics, exploratory data analysis, and pattern and image recognition because it can identify patterns and similarities in data. It can also be applied to dimensionality reduction, which lowers the number of features in a model. Two popular methods for this are singular value decomposition (SVD) and principal component analysis (PCA). Neural networks, probabilistic clustering techniques, and k-means clustering are some additional algorithms used in unsupervised learning.

Semi-supervised learning

A satisfying middle ground between supervised and unsupervised learning is provided by semi-supervised learning. It guides classification and feature extraction from a larger, unlabeled data set during training by using a smaller, labeled data set. The issue of insufficient labeled data for a supervised learning algorithm can be resolved through semi-supervised learning. It's also beneficial if labeling sufficient data would be too expensive.

**DECISION TREE**

A non-parametric supervised learning technique for regression and classification is called a decision tree (DT). The objective is to build a model that, by utilizing basic decision rules deduced from the data features, predicts the value of a target variable. One way to think of a tree is as a piecewise constant estimate.[[25]](https://scikit-learn.org/stable/modules/tree.html).To determine whether to divide a node into two or more sub-nodes, decision trees employ a variety of algorithms. The homogeneity of the resulting sub-nodes is increased by sub-node creation. Stated differently, we can assert that the node's purity rises in relation to the target variable. The decision tree divides the nodes based on all variables that are available, then chooses the split that produces the greatest number of homogeneous sub-nodes.[[26]](https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html).In addition to having better computational performance and simpler logical explanations, decision tree algorithms (DTs) use a set of supervised learned decision rules for predictions based on inputs of chosen predictor factors and learning from overlapping attributes. The model is derived from the CART-based DT model. Breiman's CART DT models from 1984 served as the model's inspiration for the software algorithm. DT algorithm CART generates regression trees or binary classification trees based on whether the target variable is categorical.

**WHY MACHINE LEARNING FOR RAILWAY SAFETY**

There are many benefits to applying machine learning to railway safety, and it can greatly improve the general level of efficiency and safety of railway operations. The following are strong arguments in favor of using machine learning in railroad safety:

Predictive Maintenance:

In order to forecast when components are likely to fail, machine learning models can examine sensor readings, equipment performance, and historical maintenance data. By enabling prompt maintenance, this proactive strategy lowers the possibility of unplanned malfunctions and improves general safety.

Anomaly Detection:

In real-time data, anomalies like strange vibrations, temperature swings, or other irregularities can be found by machine learning algorithms. This makes it possible to identify possible problems early on and give operators the opportunity to address them before they become serious safety risks.

Fault Detection and Diagnostics:

Railway faults can be identified and diagnosed with the help of machine learning. Through the examination of data patterns and trends, the technology is able to identify the underlying causes of problems, enabling more rapid and precise troubleshooting.

Safety Event Prediction:

Machine learning models are able to predict the probability of safety-related events by analyzing a variety of data sources, such as weather reports, train schedules, and historical incident reports. This enables operators to optimize safety protocols and take preventive action.

Automated Monitoring of Track and Infrastructure:

Tracks, signals, and other infrastructure elements can be monitored for condition using computer vision and image processing systems driven by machine learning. Automated inspections have the ability to quickly detect possible safety risks, like obstructions or track flaws.

Enhanced Security:

Advanced security system development can benefit from the application of machine learning algorithms. This enhances overall railway security by utilizing video analytics for surveillance, identifying potential security threats, and detecting anomalies in passenger behavior.

Optimized Traffic Management:

Train schedules can be made more efficient, traffic flow can be predicted, and machine learning can be used. This increases operational efficiency while also lowering the chance of collisions, which increases safety.

Human Factor Analysis:

The analysis of human factors that contribute to safety incidents can be aided by machine learning. Systems for detecting train operator fatigue or monitoring systems to make sure employees and passengers are following safety procedures are examples of this.

Continuous Improvement:

Through ongoing data analysis and incident learning, machine learning models support an ongoing culture of railway safety improvement. Targeted interventions and improved safety procedures can be put into place thanks to this iterative process.

Data-Driven Decision Making:

Railway operators can now make data-driven decisions using both historical and real-time data thanks to machine learning. As a result, safety assessments are more accurate and proactive risk mitigation strategies are made possible.

Overall, more proactive, effective, and data-driven approaches can result from the application of machine learning in railway safety, which will ultimately improve the general safety and dependability of railway systems.

**DATASET**

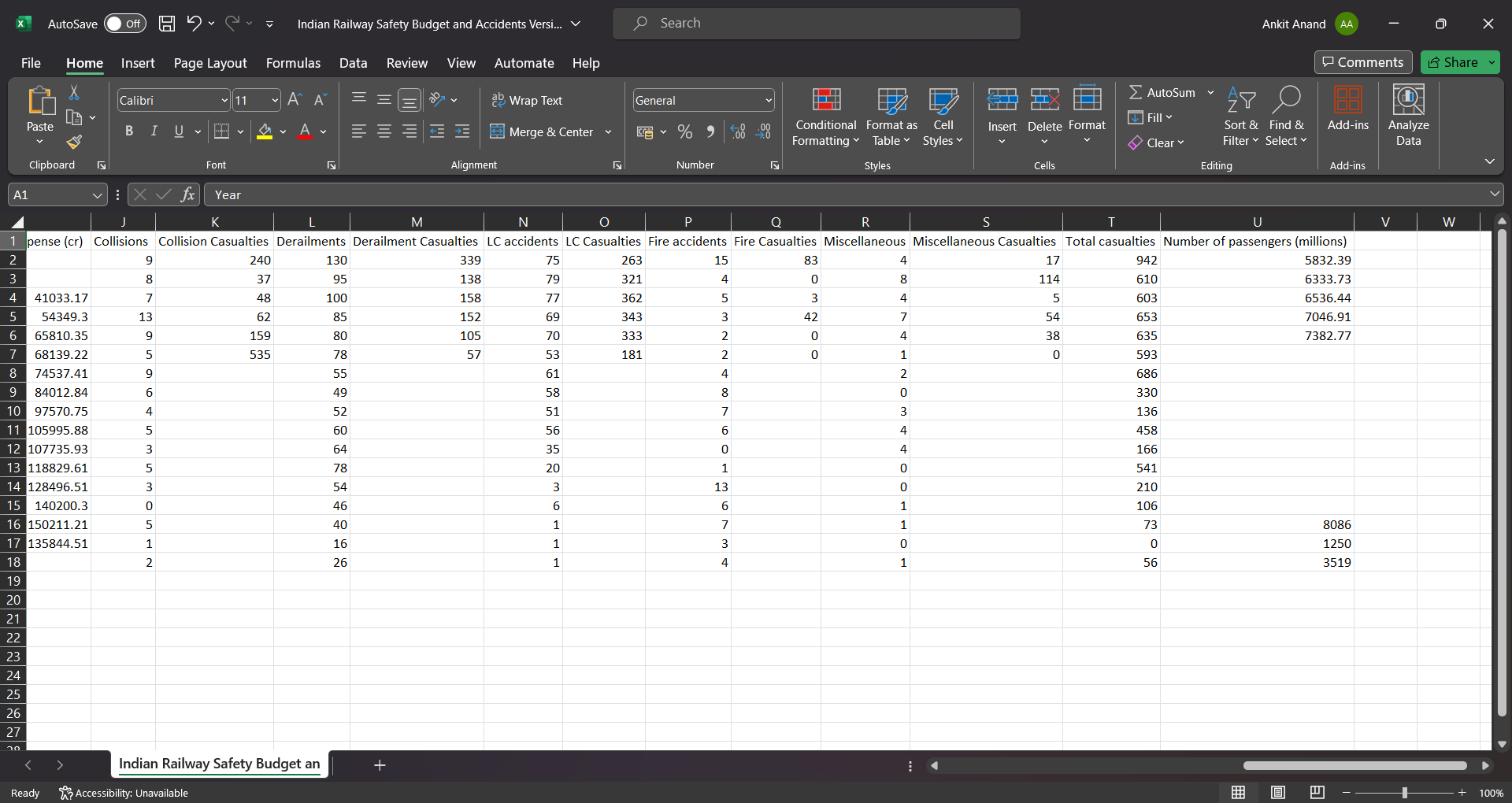
The dataset used in this study came from three different countries and contains data on more than 50 train accidents. The purpose of this dataset is to further explore the potential applications of supervised machine learning in the railroad industry. This study is important because it clarifies how management, safety, and services—particularly in station environments—can be improved through the application of machine learning (ML). Drawing from data obtained from the United States Department of Transportation[[18]](http://safetydata.fra.dot.gov/OfficeofSafety/), This model is meant to predict accidents and help decision makers. It is being used by the Rail Safety and Standards Board (RSSB)[20], the Office of Rail and Road (ORR) and London Underground[19], and the Safety Information Management System (SIMS) - Indian Railways.

We first collected the data from the various data sources, as was previously mentioned. The preparation of the data comes next. Missing values and a lot of noise are common in data. As a result, when data is prepared, it is cleaned up by removing unwanted features and adding the mean, median, or mode in place of any missing values[21]. This step takes up more than 60% of the total effort in the modeling process because of its impact on the result.

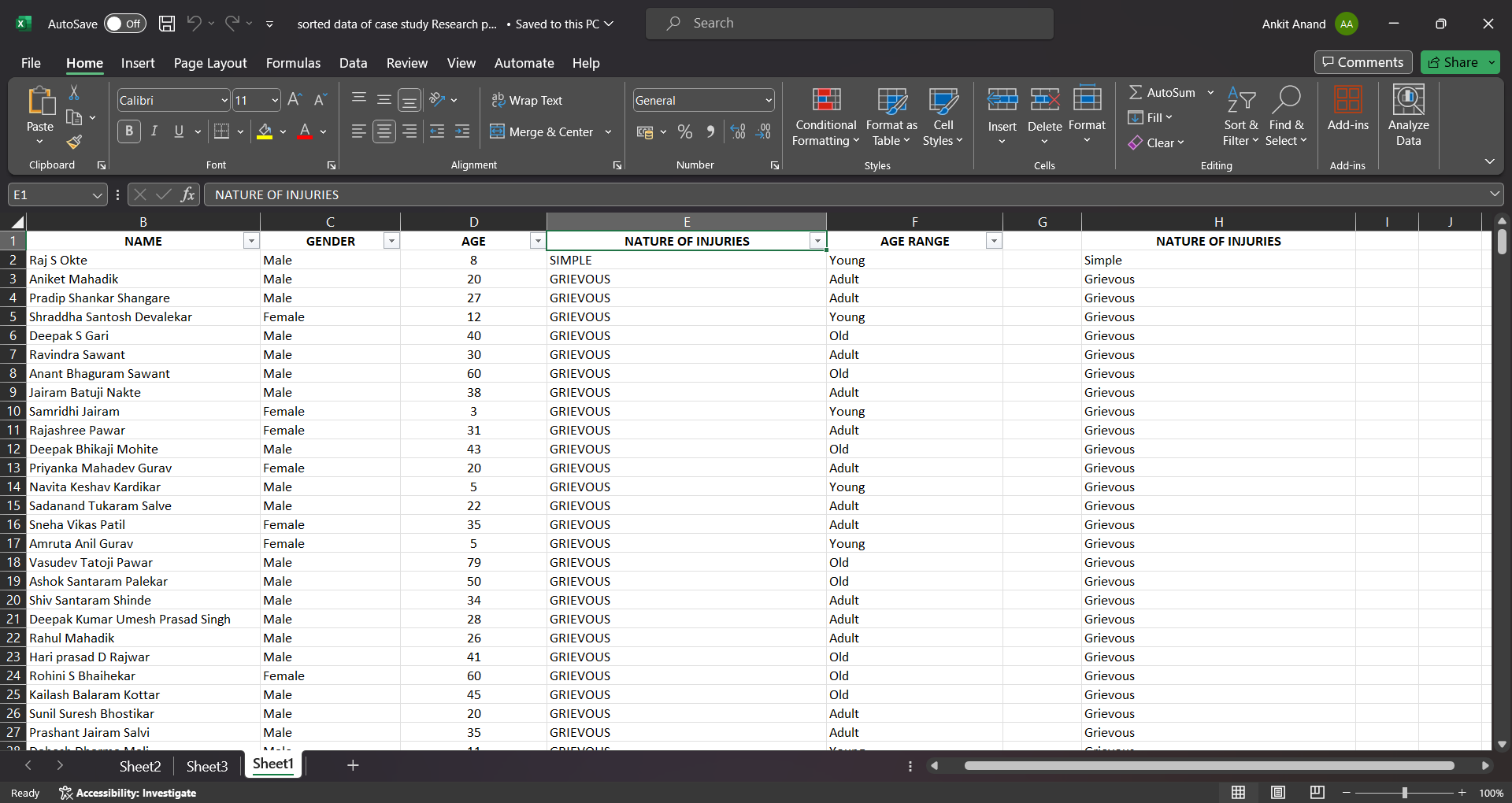
The data structure has undergone a few modifications to suit the modeling process, such as the following:

* Generalization: We generalize those attributes and make them consistent across all the data before using the dataset in the model because there are a lot of attributes that vary among the data that we have collected from different sources.
* Diminish or eliminate reductant features: Things that are not relevant to this model are removed, like accidents that happen at stations and have nothing to do with injuries sustained in train accidents.

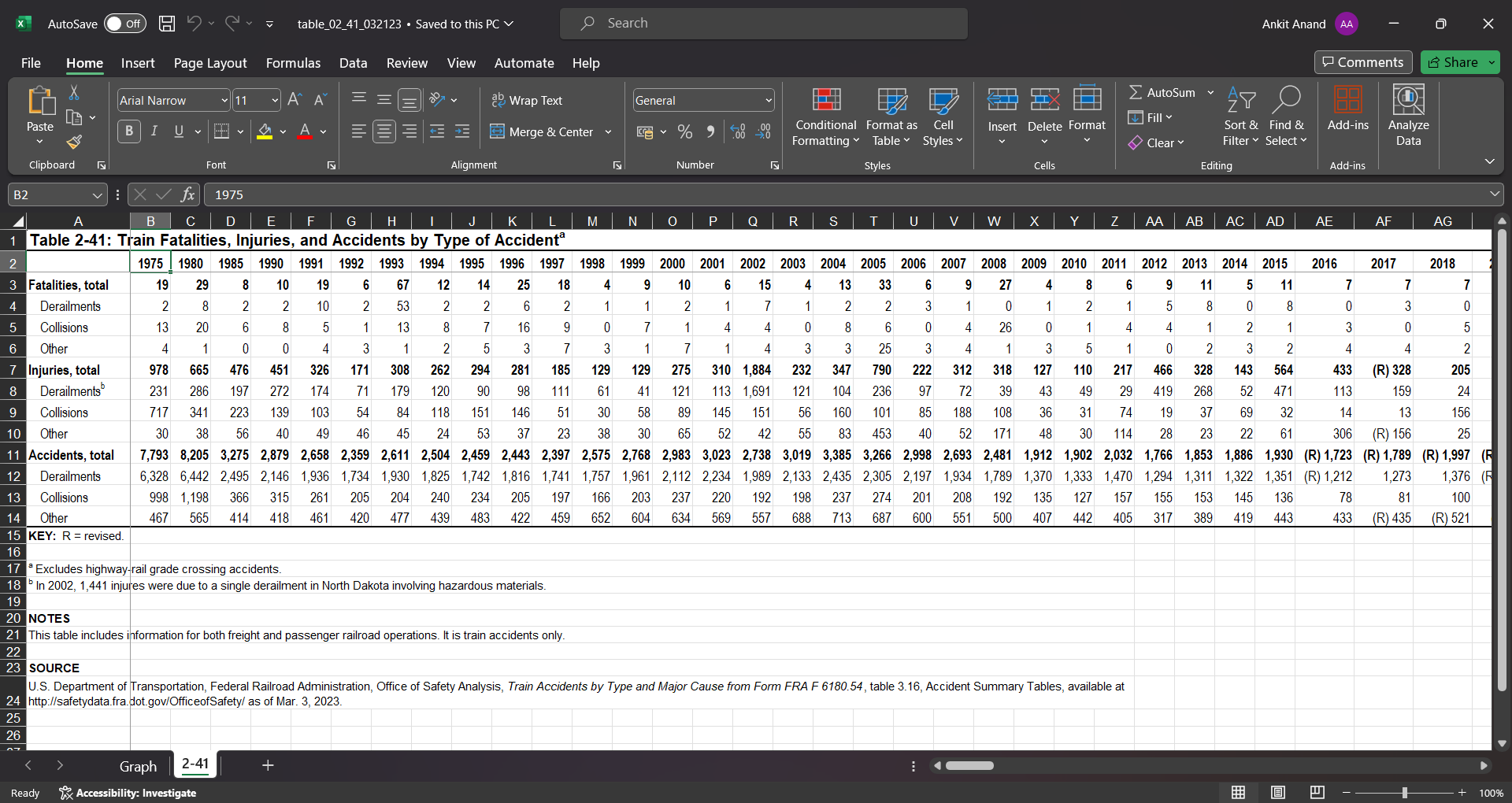
Here are some images of the raw dataset that is collected from the different sources:



img 1



img 2



img 3

The data set used in this paper is depicted in images 1, 2, and 3.

**MODEL FLOW**

We proceeded to the model preparation phase after preprocessing and eliminating all noise from the data. The necessary libraries, including pandas, matplotlib, seaborn, and scikit-learn, are first imported by the code. It loads a dataset with data on railroad accidents called "researchData.csv." After that, preprocessing is done on the dataset to get it ready for training a classification model. The 'TIME PERIOD OF ACCIDENT' and 'REASON OF THE ACCIDENT' columns contain categorical features that are one-hot encoded into a numeric format that is appropriate for machine learning. 'NATURE OF INJURIES,' the target variable, is isolated from the features. The train\_test\_split function divides the dataset into training and testing sets.

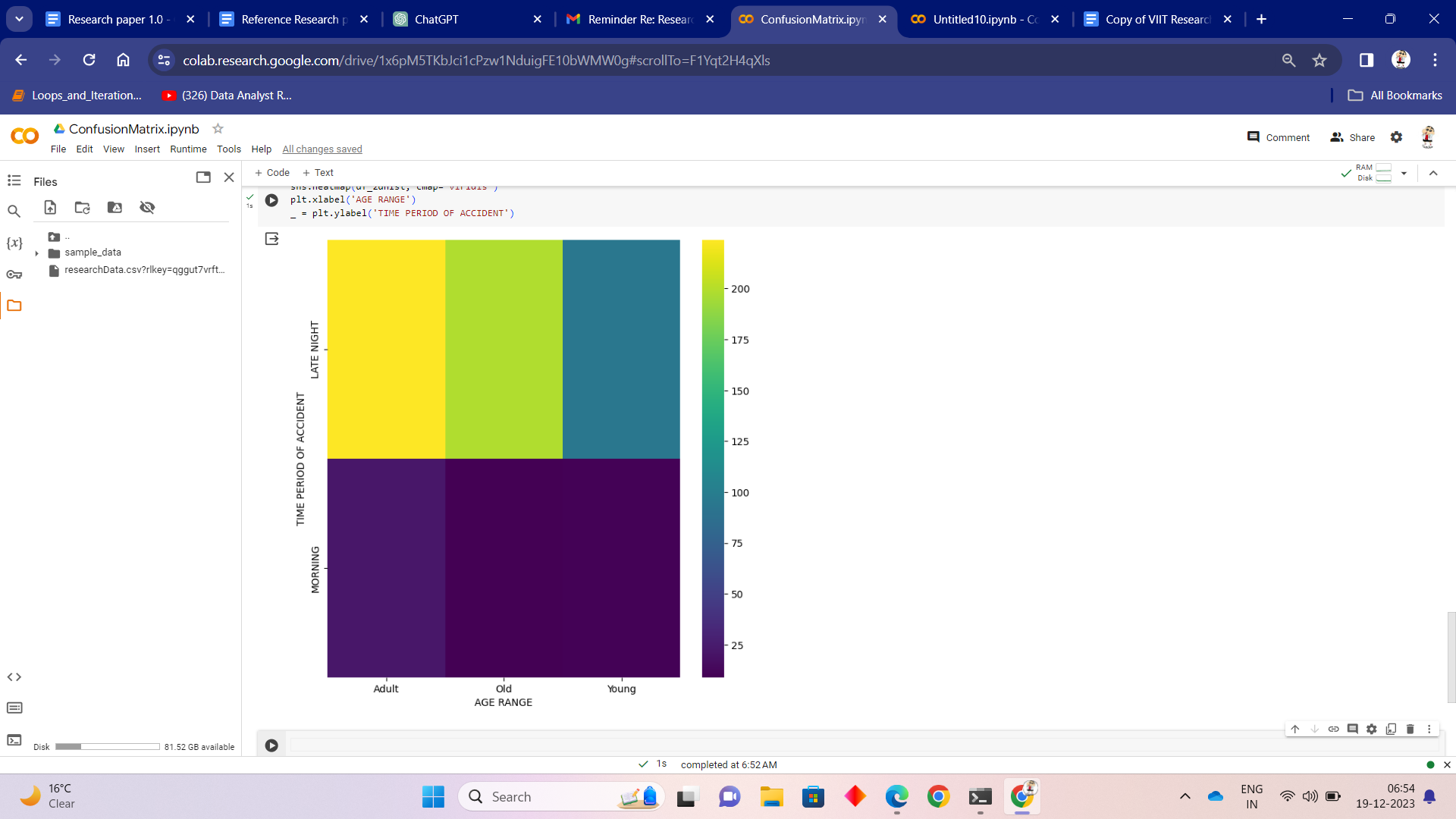
The model entails instantiating a classifier model after importing all necessary libraries. For this task, the RandomForestClassifier is chosen by the code. Next, it uses the training data (X\_cls\_train, y\_cls\_train) to train the model. The target variable for the test dataset (X\_cls\_test) is predicted by the model after training, and the predictions are kept in y\_cls\_pred. With the aid of seaborn and matplotlib, the code creates a confusion matrix heatmap to illustrate the effectiveness of the classification model. This heatmap offers information about how well the model can predict various classes. Together with a comprehensive classification report that includes metrics like precision, recall, and F1-score for every class, the classification accuracy is computed and printed.

Ultimately, the model highlights how critical it is to assess the performance of the classification model using metrics and visualization. A more thorough evaluation of the model's performance can be obtained by consulting the confusion matrix heatmap, accuracy, precision, recall, and F1-score. The code can be expanded upon or altered for additional enhancements. It demonstrates a thorough method for developing and assessing a classification model for predicting the type of injuries sustained in railroad accidents based on predetermined features.

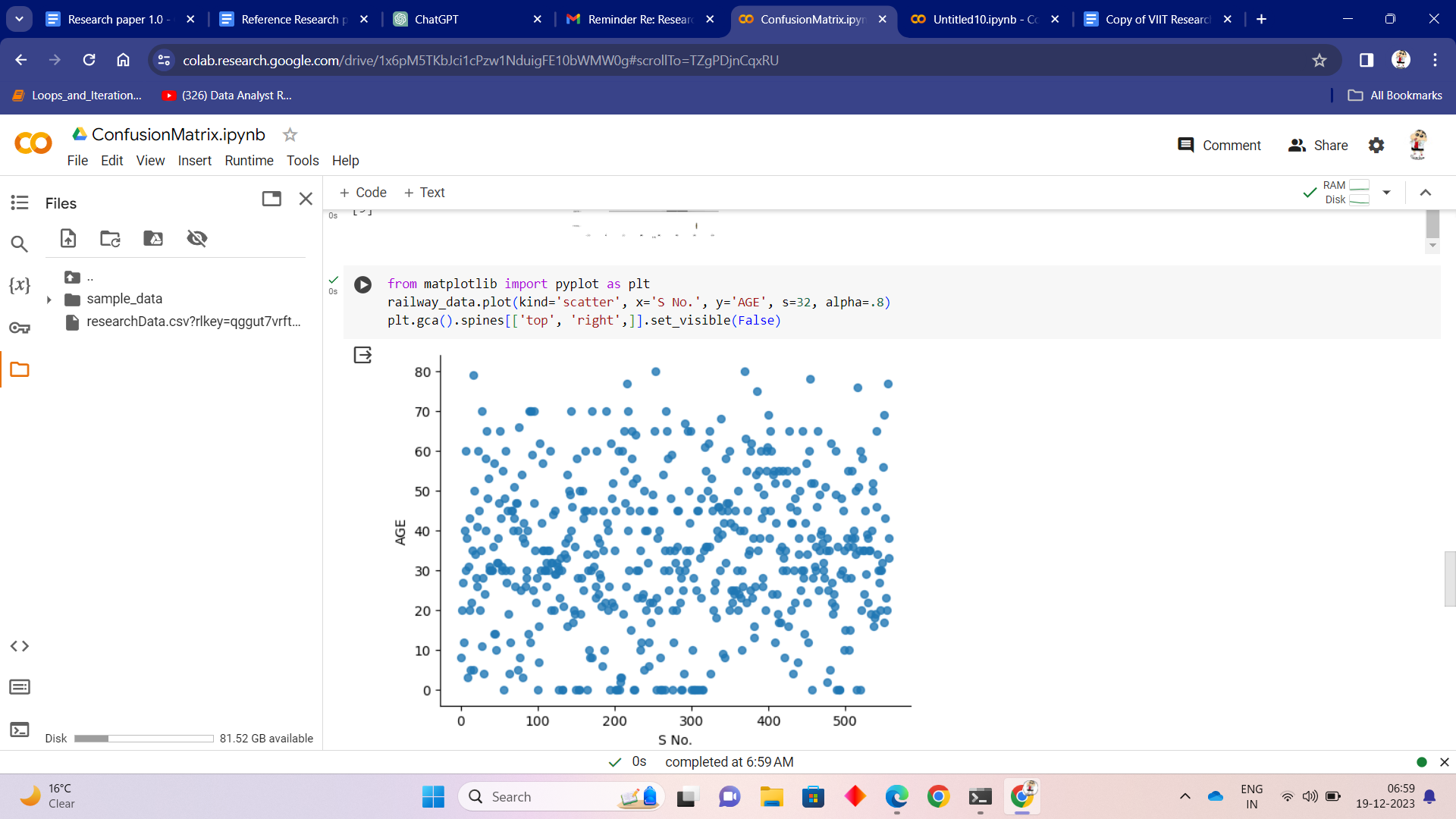
**GRAPHS**



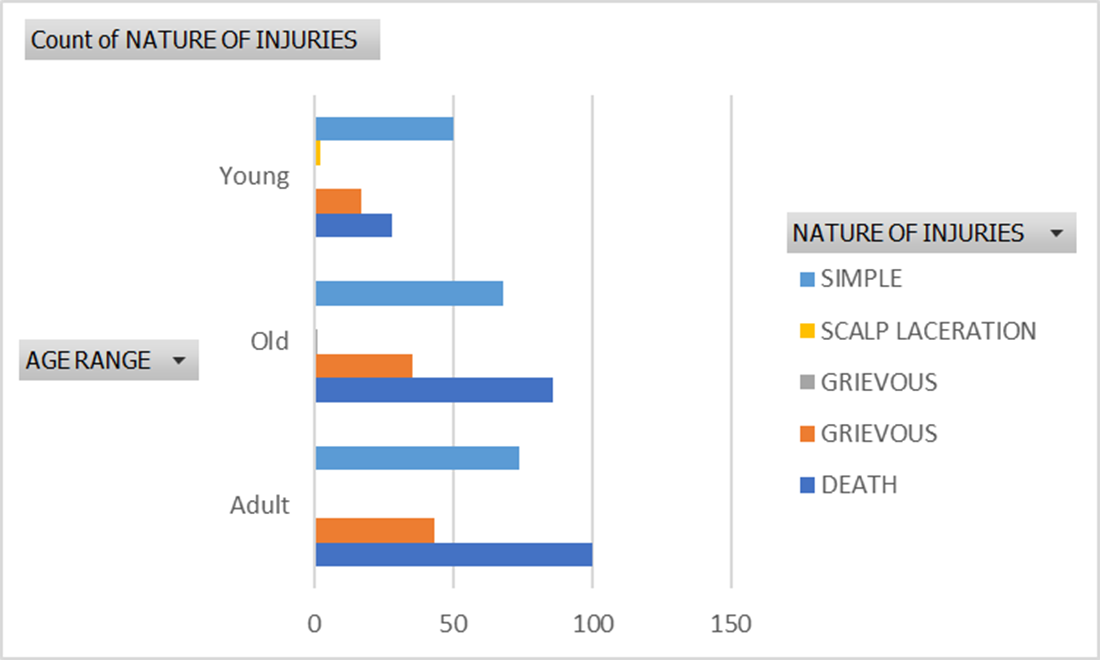
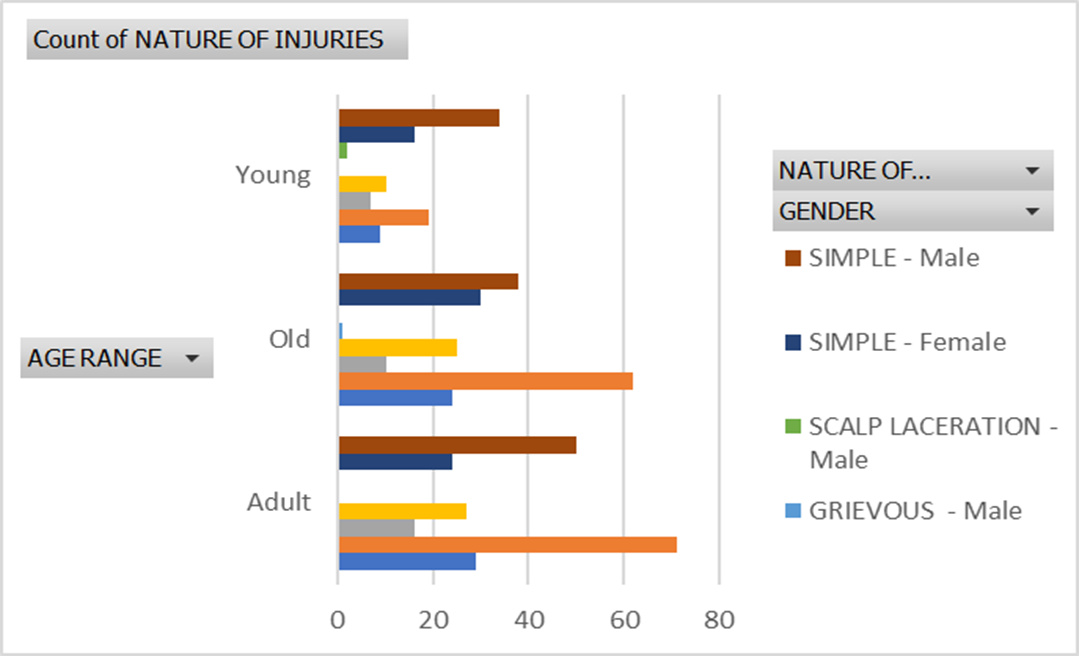
img 4: Confusion matrix graph for the model



img 5: relation between time period of accident and age range of passenger



img 6: cluster graph for the model which shows the age at which more accidents occur.



img 7 & img 8: Both the graph gives the count of nature of injuries in a particular age range and in a particular gender.

**CONCLUSION**

To sum up, the incorporation of machine learning into railway safety is a significant development that will have a profound impact on the sector. Machine learning algorithms have proven to be able to proactively close long-standing gaps in conventional safety measures when applied to predictive maintenance, anomaly detection, safety event prediction, and overall operational optimization. Railway operators can anticipate and prevent potential hazards before they escalate by using machine learning, which also improves the accuracy and efficiency of safety protocols, by utilizing historical data and real-time monitoring. Decision trees and other interpretable machine learning models make it easier to comprehend the variables affecting safety and provide stakeholders the knowledge and tools they need to make well-informed decisions and carry out focused interventions.

Additionally, the use of machine learning fosters a culture of ongoing advancement in railroad safety. These technologies make it possible to improve safety procedures and put creative solutions into practice by iteratively analyzing data and learning from previous incidents. The role of machine learning in fostering a more secure and dependable railway environment is further reinforced by its capacity to automate infrastructure monitoring, optimize traffic management, and improve security. With a trajectory towards a proactive, data-driven, and highly secure operational landscape, the outlook for railway safety is positive as the industry continues to embrace the potential of AI and ML. The partnership between technological innovation and railway safety points to a dynamic future in which innovations will be made to guarantee the safety of personnel and passengers as well as the general effectiveness of railway systems.

**FUTURE SCOPE**

The present research has provided valuable insights regarding the application of machine learning to railway safety. There are several avenues that need to be investigated and developed further, just like in any research project. This section explores the work's potential future expansions and focuses on key areas where further research and innovation could improve our comprehension of railway safety via the use of machine learning (ML) and related applications.

This work uses machine learning algorithms to predict future incidents in railway safety based only on numerical data from previous train accidents. But if we dig a little deeper into this issue, we might be able to apply deep learning to this model and try to use some image recognition software, which would simplify and improve the accuracy of accident prediction for this model.

This work is mainly about machine learning, so it does not discuss deep learning or neural networks. This may open the door to a range of neural network algorithms that could be applied in the future to generate an improved model or the same one.

The prediction model used in this paper is solely based on numerical data. In contrast, a CNN model can quickly detect those cracks in image data about the railway track, forecast the future before any trains pass by, and send out a signal to the nearest railway station to initiate track-changing work[22, 23]. Not only can this image data be used to identify issues with the trains' doors, wheels, bearings, suspensions, and other parts.

Additionally, there are numerous opportunities for safety at train stations. We have failed to consider a number of potential hazards to station safety in our attempts to predict incidents and identify the reasons behind incidents that happen on railroad tracks. Future researchers might be able to create a system that takes a picture of every person entering the station and compares it to the criminal database kept by the government thanks to technological innovation. A siren that only sounds when someone tries to cross the platform from the tracks—rather than by the escalator or foot bridge, which are intended for that purpose—may also be installed by the railway administration.

To sum up, the forthcoming advancements delineated in this section have the potential to augment our comprehension of Machine Learning for Railway Safety and convert these perspectives into pragmatic implementations. The field of railway safety is expected to witness a substantial contribution from image recognition systems as long as researchers pursue the avenues outlined here.

**REFERENCES**

1. <https://www.cisco.com/c/en/us/products/security/machine-learning-security.html#:~:text=In%20security%2C%20machine%20learning%20continuously,by%20uncovering%20suspicious%20user%20behavior>.
2. HAMAD ALAWAD 1 , SAKDIRAT KAEWUNRUEN 1 , AND MIN AN 2

Digital Object Identifier 10.1109/ACCESS.2019.2962072

1. https://www.indianrail.gov.in/enquiry/StaticPages/StaticEnquiry.jsp?StaticPage=index.html
2. <https://www.simplilearn.com/importance-of-technology-article#:~:text=Cost%20and%20Time%2DSaving,Technology%20also%20ensures%20improved%20accuracy>.
3. <https://www.linkedin.com/pulse/revolution-artificial-intelligence-ai-21st-century-de-oliveira/>
4. <https://scholar.google.com/scholar_lookup?title=Artificial+Intelligence:+A+New+Synthesis&author=Nilsson,+N.J.&author=Nilsson,+N.J.&publication_year=1998>
5. <http://www-formal.stanford.edu/jmc/whatisai.html>
6. <https://scholar.google.com/scholar_lookup?title=Artificial+intelligence+and+natural+man&author=Boden,+M.&publication_year=1980&journal=Synthese&volume=43&pages=433%E2%80%93451>
7. <https://www.mdpi.com/2076-3417/12/20/10572#B1-applsci-12-10572>
8. <https://scholar.google.com/scholar_lookup?title=Convolutional+Neural+network+based+Online+Rail+surface+Crack+Detection&conference=Proceedings+of+the+2021+5th+International+Conference+on+Intelligent+Computing+and+Control+Systems+(ICICCS)&author=Akhila,+C.&author=Diamond,+C.A.&author=Posonia,+A.M.&publication_year=2021&pages=1602%E2%80%931606>
9. <https://scholar.google.com/scholar_lookup?title=Railroad+Surface+Defect+Segmentation+Using+a+Modified+Fully+Convolutional+Network&author=Kim,+H.&author=Lee,+S.&author=Han,+S.&publication_year=2020&journal=KSII+Trans.+Internet+Inf.+Syst.+TIIS&volume=14&pages=4763%E2%80%934775>
10. <https://scholar.google.com/scholar_lookup?title=Condition+monitoring+of+railway+tracks+from+car-body+vibration+using+a+machine+learning+technique&author=Tsunashima,+H.&publication_year=2019&journal=Appl.+Sci.&volume=9&pages=2734&doi=10.3390/app9132734>
11. <https://scholar.google.com/scholar_lookup?title=Developing+machine+learning-based+models+for+railway+inspection&author=Yang,+C.&author=Sun,+Y.&author=Ladubec,+C.&author=Liu,+Y.&publication_year=2020&journal=Appl.+Sci.&volume=11&pages=13&doi=10.3390/app11010013>
12. <https://scholar.google.com/scholar_lookup?title=Fast+classification+for+rail+defect+depths+using+a+hybrid+intelligent+method&author=Jiang,+Y.&author=Wang,+H.&author=Tian,+G.&author=Yi,+Q.&author=Zhao,+J.&author=Zhen,+K.&publication_year=2019&journal=Optik&volume=180&pages=455%E2%80%93468&doi=10.1016/j.ijleo.2018.11.053>
13. <https://scholar.google.com/scholar_lookup?title=Deep+learning-based+acoustic+emission+scheme+for+nondestructive+localization+of+cracks+in+train+rails+under+a+load&author=Suwansin,+W.&author=Phasukkit,+P.&publication_year=2021&journal=Sensors&volume=21&pages=272&doi=10.3390/s21010272>
14. R. Xu, J. Han, W. Qi, J. Meng, and H. Zhang, ‘‘Railway fastener image recognition method based on multi feature fusion,’’ IOP Conf. Mater. Sci. Eng., vol. 397, Aug. 2018, Art. no. 012119, doi: 10.1088/1757- 899x/397/1/012119.
15. . Dai, ‘‘A machine learning approach for optimisation in railway planning,’’ Ph.D. dissertation, Dept. Math. Comput. Sci., Delft Univ. Technol., Delft, The Netherlands, 2018.
16. <http://safetydata.fra.dot.gov/OfficeofSafety/>
17. <https://dataportal.orr.gov.uk/statistics/health-and-safety/rail-safety/>
18. <https://safety.indianrail.gov.in/sims/>
19. <https://www.javatpoint.com/data-preparation-in-machine-learning>
20. <https://scholar.google.com/scholar_lookup?title=Random+forest+based+diagnosis+approach+for+rail+fault+inspection+in+railways&conference=Proceedings+of+the+2016+National+Conference+on+Electrical,+Electronics+and+Biomedical+Engineering+(ELECO)&author=Santur,+Y.&author=Karak%C3%B6se,+M.&author=Akin,+E.&publication_year=2016&pages=745%E2%80%93750>
21. <https://scholar.google.com/scholar_lookup?title=Deep+convolutional+neural+networks+for+detection+of+rail+surface+defects&conference=Proceedings+of+the+2016+International+Joint+Conference+on+Neural+Networks+(IJCNN)&author=Faghih-Roohi,+S.&author=Hajizadeh,+S.&author=N%C3%BA%C3%B1ez,+A.&author=Babuska,+R.&author=De+Schutter,+B.&publication_year=2016&pages=2584%E2%80%932589>
22. <https://www.ibm.com/topics/machine-learning>
23. <https://scikit-learn.org/stable/modules/tree.html>
24. <https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html>
25. Oh K, Yoo M, Jin N, Ko J, Seo J, Joo H, Ko M. A Review of Deep Learning Applications for Railway Safety. Applied Sciences. 2022; 12(20):10572. https://doi.org/10.3390/app122010572
26. Alawad, Hamad & Kaewunruen, Sakdirat & An, Min. (2019). Learning From Accidents: Machine Learning for Safety at Railway Stations. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2962072.
27. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592-x
28. Çelik, Özer. (2018). A Research on Machine Learning Methods and Its Applications. 10.31681/jetol.457046.
29. Crawford, Elise & Kift, Ryan. (2018). Keeping track of railway safety and the mechanisms for risk. Safety Science. 110. 10.1016/j.ssci.2018.07.004.
30. Vidumina, Navod. (2023). Train Accident Prevention & Breakdown Detection using Machine Learning based Safty System.