**LITERATURE SURVEY**

# 1) Comparative Analysis of Classification Techniques using WEKA on Different Datasets

# AUTHORS: Mahesh Parmar

Data mining is the analytic process designed to explore large amounts of data in search for consistent patterns and systematic relationships between variables and then to validate the findings by applying the detected patterns to new subsets of data. Data mining software are analytical tools for analyzing data. Weka is a data mining tools, contains many machine leaning algorithms and provides the facility to classify our data through various algorithms. Classification techniques a model is built based on training data and applied to test data in broad applications. In this paper, two classification algorithms are used for analyzing datasets. The main aims to show the comparative Analysis of decision tree (J48) and Backpropagation classification algorithm using WEKA tool and find out which technique is most suitable for user working on different datasets. The best algorithm based on the Bank datasets and Vote dataset is MLP classifier with accuracy respectively of 73.75. % and 96.32%.

# 2) Prediction of occupational accidents using decision tree approach

# AUTHORS: [Sobhan Sarkar](https://ieeexplore.ieee.org/author/37085804722); [Atul Patel](https://ieeexplore.ieee.org/author/37086065974); [Sarthak Madaan](https://ieeexplore.ieee.org/author/37086065064); [Jhareswar Maiti](https://ieeexplore.ieee.org/author/37085803056)

# The focus of the present study is to build a predictive model which not only could predict the occupational incidents but also provide rules for explaining accident scenarios like near-miss, property damage, or injury cases. Classification and regression tree (CART) is used for prediction purpose. Furthermore, the parameters of CART have been tuned by grid based tuning and genetic algorithm (GA). The experimental results show that the GA optimized CART provides better accuracy than others. Additionally, the best rules extracted from GA optimized CART are discussed in order to adopt better safety precautionary measures at work.

# 3) A Novel Paradigm of Melanoma Diagnosis Using Machine Learning and Information Theory

**AUTHORS:** [**Kailash Chandra Giri**](https://ieeexplore.ieee.org/author/37088391451)**;**[**Mayank Patel**](https://ieeexplore.ieee.org/author/37088389295)**;**[**Amit Sinhal**](https://ieeexplore.ieee.org/author/38489854600)**;**[**Diwakar Gautam**](https://ieeexplore.ieee.org/author/37085434337)

# Harmful Melanoma, basically the most extremely dangerous sort of epidermis malignancy, has a phenomenal conclusion whenever taken care of inside the reparable early ranges. Early determination and careful extraction is presumably the most vigorous cure of melanoma. This work utilizes a record set of 184 clinical dermatoscopic pictures of skin injuries, in which 144 pictures are of dangerous sores and 40 photos are of the amiable sore, picture pre-handling, and division techniques are utilized to separate melanoma from considerate pigmented sores. Otsu and Entropy fundamentally based picture division rules are cultivated which improves the execution. The appropriate outcomes demonstrate that Havrda Entropy and Harris Corner Detector based melanoma analysis approach accomplish greater affectability concerning Otsu and Harris based joined methodology. The separated geometrical, fringe and shading highlight set is conveyed to characterize an outlining limit among considerate and dangerous classes of melanoma, and it is seen that entropy-based neural learning approach outflanks to Otsu based neural learning approach individually.

# 4) Survey paper on crime prediction using ensemble approach

**AUTHORS :** [**Ayishu Almaw**](https://www.researchgate.net/profile/Ayishu-Almaw)

# Crime is a foremost problem where the top priority has been concerned by individual, the community and government. This paper investigates a number of data mining algorithms and ensemble learning which are applied on crime data mining. This survey paper describes a summary of the methods and techniques which are implemented in crime data analysis and prediction. Crime forecasting is a way of trying to mining out and decreasing the upcoming crimes by forecasting the future crime that will occur. Crime prediction practices historical data and after examining data, predict the upcoming crime with respect to location, time, day, season and year. In present crime cases rapidly increases so it is an inspiring task to foresee upcoming crimes closely with better accuracy. Data mining methods are too important to resolving crime problem with investigating hidden crime patterns.so the objective of this study could be analyzing and discussing various methods which are applied on crime prediction and analysis. This paper delivers reasonable investigation of Data mining Techniques and ensemble classification techniques for discovery and prediction of upcoming crime.

# 

# 5) AN ALGORITHM FOR PREDICTIVE DATA MINING APPROACH IN MEDICAL DIAGNOSIS

# AUTHORS: Shakuntala Jatav1 and Vivek Sharma2

# The Healthcare industry contains big and complex data that may be required in order to discover fascinating pattern of diseases & makes effective decisions with the help of different machine learning techniques. Advanced data mining techniques are used to discover knowledge in database and for medical research. This paper has analyzed prediction systems for Diabetes, Kidney and Liver disease using more number of input attributes. The data mining classification techniques, namely Support Vector Machine(SVM) and Random Forest (RF) are analyzed on Diabetes, Kidney and Liver disease database. The performance of these techniques is compared, based on precision, recall, accuracy, f\_measure as well as time. As a result of study the proposed algorithm is designed using SVM and RF algorithm and the experimental result shows the accuracy of 99.35%, 99.37 and 99.14 on diabetes, kidney and liver disease respectively.