**CLASSIFICATIO|N OF DECISION TREE**

**Abstract**

Decision tree classifiers are regarded to be a standout of the most well-known methods to data classification representation of classifiers. Different researchers from various fields and backgrounds have considered the problem of extending a decision tree from available data, such as machine study, pattern recognition, and statistics. In various fields such as medical disease analysis, text classification, and many more the employment of Decision tree classifiers has been proposed in many ways. This paper provides a detailed approach to the decision trees.. As a result, the uses of different types of datasets are discussed and their findings are analyzed.

**INTRODUCTION**

Nowadays, technology has developed a lot, especially in the field of Machine Learning , which is useful for reducing human work. In the field of artificial intelligence, ML integrates statistics and computer science to build algorithms that get more efficient when they are subject to relevant data rather than being given specific instructions. Besides speech recognition, image detection, text localization, etc. ML is the study of computational algorithms that are enhanced from experience automatically. It is considered as an artificial intelligence subset [1, 2]. Orderly to produce foretelling or decision without being specifically programmed to do so, ML algorithms create a model population based on a sample, defined as 'training data' [3, 4]. In a broad area of applications, like email filtering and computer vision, ML algorithms are utilized where it is difficult or impractical to create traditional algorithms to implement functions required [5]. Two major mechanisms can be broken into ML training data is machine learning. It also analyzes the testing data and creates a derived task that can be used for new examples to map [12, 13]. Each data input object, however, has a class label pre- assigned. The primary function of supervised algorithms is to learn a model that creates the same labeling preferably for the data offered and popularizes well on unseen data. This is the major aim of algorithms for classification [14]

A decision tree is a tree-based technique in which any path beginning from the root is described by a data separating sequence until a Boolean outcome at the leaf node is achieved [24 - 27]..

In this paper, a comprehensive review is performed for the latest and most efficient approaches that have been performed by researchers in the past three years about decision trees in different areas of machine learning. Also, the details of this method, such as using datasets, and the findings achieved are summarized. In addition, this study highlighted the most commonly used approaches and the highest accuracy methods achieved.

The organization of the remaining paper is as follows: Section II contains the decision tree algorithm mentioning its types, benefits, and drawbacks; Section III gives a Literature Review on decision tree Algorithm; Section IV comparison and discussion on the decision tree, and the last section contains the conclusion.

**DECISION TREE ALGORITHM**

One of the widely used techniques in data mining is systems that create classifiers [32]. In data mining, classification algorithms are capable of handling a vast volume of information. It can be used to make assumptions regarding categorical class names, to classify knowledge on the basis of training sets and class labels, and to classify newly obtainable data [33]. Classification algorithms in machine learning contain several algorithms, and in this work, the paper focused on the decision tree algorithm in general.

The conceptual rules are much easier to construct than the numerical weights in the neural network of connections between nodes [37, 38]. Mainly for grouping purposes, DT is used. Moreover, DT is a usually utilized classification model in Data Mining [39]. The nodes and branches are composed of each tree. Each node represents features in a category to be classified and each subset defines a value that can be taken by the node [40, 41]. Because of their simple analysis and their precision on multiple data forms, decision trees have found many implementation fields [42]

**Decision Tree Algorithms**

There are several Types of DT algorithms such as: Iterative Dichotomies 3 (ID3), Successor of ID3 (C4.5), Classification And Regression Tree(CART) [44], CHi-squared Automatic Interaction Detector(CHAID) [45], Multivariate Adaptive Regression Splines (MARS) [46], Generalized, Unbiased, Interaction Detection and Estimation (GUIDE), Conditional Inference Trees (CTREE) [47],[48], Classification Rule with Unbiased Interaction Selection and Estimation (CRUISE), Quick, Unbiased and Efficient Statistical Tree (QUEST) [49], [50]. Table I shown comparison between the frequently used algorithms for the decision tree [51].

**Entropy and Information Gain**

Entropy is employed to measure a dataset's impurity or randomness [52], [53]. The value of entropy always lies between 0 and 1. Its value is better when it is equal to 0 while it is worse when it is equal to 0, i.e. the closer its value to 0 the better

Information gain is one metric used for segmentation and is often called mutual information. This intuitively informs how much knowledge of a random variable's value [57, 58]. It’s the opposite of entropy, the higher its value is the better. The data gain 𝐺𝑎𝑖( 𝑆 , 𝐴 ) is defined as the following on the definition of entropy [59, 60], as shown in “equation (2)”.

**Benefits and Drawbacks of decision tree**

The DT algorithm is part of the supervised learning algorithm family, and its main objective is to construct a training model that can be used to predict the class or value of target variables through learning decision rules inferred from the training data. The DT algorithm can be used to solve regression and classification problems, but it has benefits and drawbacks [61 - 63].

De Felice et al. [66] suggested a decision tree algorithm to recognize known and novel clinical indications before treatment for survival in Locally Advanced Rectal Cancer (LARC). The analytics showed that even non-experts in the field, in particular classification trees, can easily interpret the tree-based machine learning process. Validation errors need to be managed to even achieve their statistical capacity. Around 2007 and 2014, patients with histologically confirmed LARC

1. R. L. De Mántaras, “A distance-based attribute selection measure for decision tree induction,” *Machine learning*, vol. 6, no. 1, pp. 81–92, 1991.
2. S. Taneja, C. Gupta, K. Goyal, and D. Gureja, “An enhanced k-nearest neighbor algorithm using information gain and clustering,” in *2014 Fourth International Conference on Advanced Computing & Communication Technologies*, 2014, pp. 325–329.