**CHAPTER TWO**

**LITERATURE REVIEW**

**2.1 Historical Development of Employee Appraisal And Assessment**

The history of performance appraisal is quite brief. (www.whatishumanresource.com, 2000) Its roots in the early 20th century can be traced to Taylor's pioneering Time and Motion studies. But this is not very helpful, for the same may be said about almost everything in the field of modern human resources management.

As a distinct and formal management procedure used in the evaluation of work performance, appraisal really dates from the time of the Second World War - not more than 60 years ago. Yet in a broader sense, the practice of appraisal is a very ancient art. In the scale of things historical, it might well lay claim to being the world's second oldest profession! There is, says Dulewicz (Dulewicz, 1989)"... a basic human tendency to make judgements about those one is working with, as well as about oneself."

Appraisal, it seems, is both inevitable and universal. In the absence of a carefully structured system of appraisal, people will tend to judge the work performance of others, including subordinates, naturally, informally and arbitrarily. The human inclination to judge can create serious motivational, ethical and legal problems in the workplace. Without a structured appraisal system, there is little chance of ensuring that the judgements made will be lawful, fair, defensible and accurate. Performance appraisal systems began as simple methods of income justification. That is, an appraisal was used to decide whether or not the salary or wage of an individual employee was justified.

The process was firmly linked to material outcomes. If an employee's performance was found to be less than ideal, a cut in pay would follow.

On the other hand, if their performance was better than the supervisor expected, a pay rise was in order.

Little consideration, if any, was given to the developmental possibilities of the appraisal. It was felt that a cut in pay, or rise, should provide the only required impetus for an employee to either improve or continue to perform well. Sometimes this basic system succeeded in getting the results that were intended; but more often than not, it failed.

Institutionalization of performance appraisal started as far back as the industrial revolution when it was used as a means of measuring organizational efficiency (Fandray, 2001). Wren (1994) affirmed that Performance appraisal was incepted when Robert Owen used a wooden coloured block to measure the achievement of employees working in the cotton Mills in Scotland at the close of work hours. During that era, it was utilized as a disciplinary mechanism for punishing poor performance (Kennedy & Dresser 2001). This resulted in the negative notation of the appraisal system which turned out to be despised by both the appraiser and the appraisee. As confirmed by Robert and Pregitzer (2007) “performance appraisal is a yearly rite of passage in organizations that triggers dread and apprehension in the most experienced, battle-hardened manager”. The above quote summarizes the extent to which the appraisal process is disliked by the evaluators. Subsequently, organizations tried to refine the methods linking it to other administrative matters including reward, promotion, training and so forth, arguing that employees achievements should not only be measured but evaluated and managed (Kennedy & Dresser, 2001). Despite the historical perspective, an appraisal is both inevitable and universal. There has been much analysis and wide criticisms of the effectiveness and use of performance appraisal within the organizational context but up to recent times, the issue is still being debated among scholars, academicians and professionals.

**2.2 Traditional methods of Performance Appraisal**

**2.2.1 Rating Scales Method**

Rating Scales Method is a commonly used method for assessing the performance of the employees and well-known traditional method of performance appraisal of employees (https://sites.google.com/site/whatishumanresource/traditional-methods-of-performance-appraisal, 2000). Many corporations and companies example in the country India, telecommunications company likely airtel and US IT companies like Dell Corporation are using this method for evaluating the employees and subsequently take decisions on the concerned employee.

Depending upon the job of an employee under this method of appraisal traits like attitude, performance, regularity, accountability, and sincerity, etc, are rated with the scale from 1 to 10. 1 indicates negative feedback and 10 indicates positive feedback as shown below.

The attitude of the employee towards his superiors, colleagues, and customers

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Extremely poor Excellent

Regularity in the job

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Extremely                                                                                    Excellent

poor

Under this method of performance appraisal, an employee may be assessed by his superiors, colleagues, subordinates or sometimes by his customers which all depends on nature of the company or job which is added where the employee. An appraiser is a person who appraises employee will give a rating for every trait given by marking or choosing a number based on his observation and satisfaction. ultimately all numbers chosen or marked will be added to determine the highest score gained by an employee. An employee who scored more points will be treated as a top performer following  descending scored employees will be treated as a low performer and the least scored employee will be treated as non-performers

**2.2.2 Essey Appraisal Method**

This traditional form of appraisal, also known as “Free Form method” involves a description of the performance of an employee by his superior. The description is an evaluation of the performance of any individual based on the facts and often includes examples and shreds of evidence to support the information. A major drawback of the method is the inseparability of the bias of the evaluator.   
  
Under this method, the rater is asked to express the strong as well as weak points of the employee’s behaviour. This technique is normally used with a combination of the graphics rating scale because the rater can elaborately present the scale by substantiating an explanation for his rating. While preparing the essay on the employee, the rater considers the following factors: 

* Job knowledge and potential of the employee;
* Employee’s understanding of the company’s programmes, policies, objectives, etc.;
* The employee’s relations with co-workers and superiors;

How is technology helping transform performance appraisal and management?  
  
The most common cliché is “performance appraisal is dead“. What it means is the annual performance appraisal is dead and most organizations today are adopting regular, real-time, anecdotal and casual feedback. Just as young people do not write e-mails anymore they WhatsApp similarly we don't have to write a long performance appraisal at the end of the year. Companies are moving to real-time performance appraisal and that is why in our system everything is shareable and real-time. And technology plays the most crucial role in providing the platform to have fast and quick and informal communication.

* The employee’s general planning, organizing and controlling ability;
* The attitudes and perceptions of the employee, in general.

Essay evaluation is a non-quantitative technique. This method is advantageous in at least one sense, i.e., the essay provides a good deal of information about the employee and also reveals more about the evaluator. The essay evaluation method, however, suffers from the following limitations:

* It is highly subjective; the supervisor may write a biased essay. The employees who are sycophants will be evaluated more favourably than other employees.
* Some evaluators may be poor in writing essays on employee performance. Others may be superficial in explanation and use flowery language which may not reflect the actual performance of the employee. It is very difficult to find effective writers nowadays.

The appraiser is required to find time to prepare the essay. A busy appraiser may write the essay hurriedly without properly assessing the actual performance of the worker. On the other hand, the appraiser takes a long time, this becomes uneconomical from the viewpoint of the firm, because the time of the evaluator (supervisor) is costly.

**2.2.3 Ranking Method**

Under the ranking method, the manager compares an employee to other similar employees, rather than to a standard measurement. An offshoot of ranking is the forced distribution method, which is similar to grading on a curve. Predetermined percentages of employees are placed in various performance categories, for example, excellent,above average, average, below average, and poor,. The employees ranked in the top group usually get the rewards (raise, bonus, promotion), those not at the top tend to have the reward withheld, and those at the bottom sometimes get punished.

**2.2.3.1 Why and when do we use the ranking method?** Managers have to make evaluative decisions, such as who is the employee of the month, who gets a raise or promotion, and who gets laid off. So when we have to make evaluative decisions, we generally have to use ranking. However, our ranking can, and when possible should be based on other methods and forms. Ranking can also be used for developmental purposes by letting employees know where they stand in comparison to their peers—they can be motivated to improve performance (https://sites.google.com/site/whatishumanresource/traditional-methods-of-performance-appraisal, 2000).

**2.2.4 Paired Comparison**

A better technique of comparison than the straight ranking method, this method compares each employee with all others in the group, one at a time. After all the comparisons on the basis of the overall comparisons, the employees are given the final rankings.

**2.2.5 Critical Incidents Location**

This technique of performance appraisal was developed by Flanagan and Burns (Flanagan, 1954). The manager prepares lists of statements of very effective and ineffective behaviour of an employee. These critical incidents or events represent the outstanding or poor behaviour of employees on the job. The manager maintains logs on each employee, whereby he periodically records critical incidents of the workers' behaviour. At the end of the rating period, these recorded critical incidents are used in the evaluation of the workers’ performance.

2.2.6 CHECKLIST METHOD  
The rater is given a checklist of the descriptions of the behaviour of the employees on the job. The checklist contains a list of statements on the basis of which the rater describes the on the job performance of the employees.

Another simple type of individual evaluation method is the checklist. A checklist represents, in its simplest form, a set of objectives or descriptive statements about the employee and his behaviour. If the rater believes strongly that the employee possesses a particular listed trait, he checks the item; otherwise, he leaves the item blank. A more recent variation of the checklist method is the weighted list. Under this, the value of each question may be weighted equally or certain questions may be weighted more heavily than others. The following are some of the sample questions in the checklist.

* Is the employee really interested in the task assigned?    Yes/No
* Is he/she respected by his colleagues (co-workers)              Yes/No
* Does he/she give respect to his superiors?                          Yes/No
* Does he/she follow instructions properly?                            Yes/No
* Does he/she make mistakes frequently?                              Yes/No

**2.3 Modern Method Of Appraisal**

**2.3.1 Assessment Centres**

An assessment centre is a place to evaluate an individual potentiality and performance, so as to position he/she in the core functional areas in an organisation. normally, organisations outsource assessment centres instead of making them on their own.

This method of performance appraisal is being opted by the RBI (Reserve Bank of India) for assessment of its officers (Performance appraisals: What India Inc. plans to offer to its employees, 2019).

An assessment centre typically involves the use of methods like social/informal events, tests and exercises, assignments being given to a group of employees to assess their competencies to take higher responsibilities in the future. Generally, employees are given an assignment similar to the job they would be expected to perform if promoted. The trained evaluators observe and evaluate employees as they perform the assigned jobs and are evaluated on job-related characteristics.   
  
The major competencies that are judged in assessment centres are interpersonal skills, intellectual capability, planning and organizing capabilities, motivation, career orientation etc. assessment centres are also an effective way to determine the training and development needs of the targeted employees.   
  
Nearly 30% of companies seek assessment centre services while moving an employee from an executive position to a managerial position. 20% of companies said they use the centre's services when seeking a position on senior management level (modern methods of performance appraisal, 2000).

**2.3.2 BEHAVIORALLY ANCHORED RATING SCALES**  
  
Behaviorally Anchored Rating Scales (BARS) is a relatively new technique which combines the graphic rating scale and critical incidents method. It consists of predetermined critical areas of job performance or sets of behavioural statements describing important job performance qualities as good or bad (for example the qualities like inter-personal relationships, adaptability and reliability, job knowledge etc). These statements are developed from critical incidents.   
  
In this method, an employee’s actual job behaviour is judged against the desired behaviour by recording and comparing the behaviour with BARS. Developing and practising BARS requires expert knowledge.  
  
A behaviorally anchored rating scale is an employee appraisal system where raters distinguish between successful and unsuccessful job performance by collecting and listing critical job factors. These critical behaviours are categorized and appointed a numerical value which is used as the basis for rating performance.

**2.3.3 HUMAN RESOURCE ACCOUNTING METHOD**  
  
Human Resource Accounting is a method to measure the effectiveness of personnel management activities and the use of people in an organization.HRA is the process of Assigning, budgeting, and reporting the cost of human resources incurred in an organization, including wages and salaries and training expenses.  
  
Human resources are valuable assets for every organization. Human resource accounting method tries to find the relative worth of these assets in terms of money. In this method, the Performance appraisal of the employees is judged in terms of the cost and contribution of the employees. The cost of employees include all the expenses incurred on them like their compensation, recruitment and selection costs, induction and training costs etc whereas their contribution includes the total value added (in monetary terms). The difference between the cost and the contribution will be the performance of the employees. Ideally, the contribution of the employees should be greater than the cost incurred on them.

**2.3.4 CUSTOMER FEEDBACK METHOD**

This method is truly one of the modern methods of the performance appraisal system. Customer feedback method is used, especially for sales staff who deal with sales activity in the organisation. Under this method of appraisal system, customer feedback is directly linked with employee performance. This method of assessment could be unbiased and reliable since customers who are outsiders may give correct judgement about employee performance than the insiders who are superiors.  
  
Companies like Cognizant and Wipro software solutions (modern methods of performance appraisal, 2000) are using customer feedback method to assess the performance of their sales staff in order to hike salaries.

Some industry experts say placing a higher weightage on customer feedback may fail to motivate employees because customer feedback may vary from customer to customer. Depending on the current state of mind, psychology and perception of the customers may give different feedback for the same degree of sales service.

### 2.4 PERFORMANCE APPRAISAL CRITICISM

An issue with performance appraisals is that differentiating individual and organizational performance can be difficult. And if the evaluation's construction doesn't reflect the culture of a company or organization, it can be detrimental. Employees report general dissatisfaction with their performance appraisal processes (Gary Roberts, 2007). Other potential issues include:

* Distrust of the appraisal can lead to issues between subordinates and supervisors or a situation in which employees merely tailor their input to please their employer.
* Performance appraisals can lead to the adoption of unreasonable goals that demoralize workers or incentivize them to engage in unethical practices.
* Some labour experts believe that the use of performance appraisals has led to lower use of merit- and performance-based compensation (Chen, 2018).
* Performance appraisals may lead to unfair evaluations in which employees are judged not by their accomplishments but by their likeability. They can also lead to managers giving underperforming staff a good evaluation to avoid souring their relationship.
* Unreliable raters can introduce a number of biases that skew appraisal results toward preferred characteristics or ones that reflect the rater's preferences.
* Performance appraisals that work well in one culture or job function may not be useful in another.

**2.5 DECISION TREE**

A **decision tree** is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility (Brid, 2018). It is one way to display an algorithm that only contains conditional control statements.

A decision tree is a flowchart-like structure in which each internal node represents a “test” on an attribute (e.g. whether a coin flip comes up heads or tails), each branch represents the outcome of the test, and each leaf node represents a class label (decision taken after computing all attributes). The paths from root to leaf represent classification rules.

Tree based learning algorithms are considered to be one of the best and mostly used supervised learning methods. Tree based methods empower predictive models with high accuracy, stability and ease of interpretation. Unlike linear models, they map non-linear relationships quite well. They are adaptable at solving any kind of problem at hand (classification or regression). Decision Tree algorithms are referred to as **CART** **(Classification and Regression Trees)**.

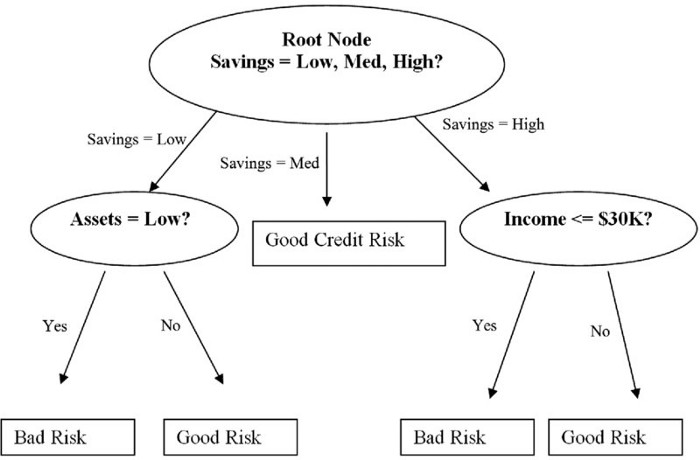


Figure 2.0 Decision tree example

**2.6 APPLICATIONS FOR DECISION TREE** :

Decision trees have a natural “if … then … else …” construction that makes it fit easily into a programmatic structure. They also are well suited to categorization problems where attributes or features are systematically checked to determine a final category. For example, a decision tree could be used effectively to determine the species of an animal.

A decision tree can be applied in any of the below sceneriors

* When the user has an objective he is trying to achieve, like to maximise profit or optimise cost
* When there are several courses of action
* There is a calculable measure of benefits of several alternatives
* When there are uncertainties concerning which outcome will actually happen.
* When they are events beyond the control of the decision maker, like environmental factors.

As a result, the decision making tree is one of the more popular classification algorithms being used in Data Mining and Machine Learning. Example applications include:

* + Evaluation of brand expansion opportunities for a business using historical sales data
  + Determination of likely buyers of a product using demographic data to enable targeting of limited advertisement budget
  + Prediction of likelihood of default for applicant borrowers using predictive models generated from historical data
  + Help with prioritization of emergency room patient treatment using a predictive model based on factors such as age, blood pressure, gender, location and severity of pain, and other measurements
  + Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal.

Because of their simplicity, tree diagrams have been used in a broad range of industries and disciplines including civil planning, energy, financial, engineering, healthcare, pharmaceutical, education, law, and business.

**2.7 TYPES OF DECISION TREES**

Decision trees used in data mining are mainly of two types (Bhumika Gupta, 2017):

• **Classification** tree in which analysis is when the predicted outcome is the class to which the data belongs. For example outcome of loan application as safe or risky.

• **Regression** tree in which analysis is when the predicted outcome can be considered a real number. For example population of a state.

Both the classification and regression trees have similarities as well as differences, such as procedure used to determine where to split.

There are various decision trees algorithms namely ID3(Iterative Dichotomiser 3), C4.5, CART(Classification and Regression Tree), CHAID(CHi- squared Automatic Interaction Detector), MARS. Out of these, we will be discussing the more popular ones which are ID3, C4.5, CART.

**2.7.1 ID3(Iterative Dichotomiser)**

ID3 is an algorithm developed by Ross Quinlan (https://en.wikipedia.org/wiki/Ross\_Quinlan, 2019) used to generate a decision tree from a dataset .

To construct a decision tree, ID3 uses a top-down, greedy search through the given sets, where each attribute at every tree node is tested to select the attribute that is best for classification of a given set . Therefore, the attribute with the highest information gain can be selected as the test attribute of the current node. ID3 is based on Occam’s razor. In this algorithm, small decision trees are preferred over the larger ones. However, it does not always construct the smallest tree and is, therefore, a heuristic algorithm (Hemlata, 2013). For building a decision tree model, ID3 only accepts categorical attributes. Accurate results are not given by ID3 when there is noise and when it is serially implemented. Therefore data is preprocessed before constructing a decision tree (Anuj & Robin, 2013). For constructing a decision tree information gain is calculated for each and every attribute and attribute with the highest information gain becomes the root node. The rest possible values are denoted by arcs. After that, all the outcome instances that are possible are examined whether they belong to the same class or not. For the instances of the same class, a single name class is used to denote otherwise the instances are classified on the basis of splitting attribute.

**2.7.1.1 Advantages of ID3**

• The training data is used to create understandable prediction rules.

• It builds the fastest as well as a short tree.

• ID3 searches the whole dataset to create the whole tree.

• It finds the leaf nodes thus enabling the test data to be pruned and reducing the number of tests.

• The calculation time of ID3 is the linear function of the product of the characteristic number and node number (Qing-Yun, Chun-ping, & Hao, 2016).

**2.7.1.2 Disadvantages of ID3**

• For a small sample, data may be over-fitted or overclassified.

• For making a decision, only one attribute is tested at an instant thus consuming a lot of time.

• Classifying the continuous data may prove to be expensive in terms of computation, as many trees have to be generated to see where to break the continuum.

• One disadvantage of ID3 is that when given a large number of input values, it is overly sensitive to features with a large number of values (Badr & Abdelkarim, 2011).

**2.7.2 Classification And Regression Trees (CART) .**

Classification And Regression Tree was introduced by Breiman in 1984 (Leo, 1984) . CART algorithm builds both classification and regression trees. The classification tree is constructed by CART by the binary splitting of the attribute. Gini Index is used as selecting the splitting attribute. The CART is also used for regression analysis with the help of regression tree. The regression feature of CART can be used in forecasting a dependent variable given a set of predictor variable over a given period of time. CART have an average speed of processing and supports both continuous and nominal attribute data.

**2.7.2.1 Advantages of CART**

• CART can handle missing values automatically using surrogate splits.

• Uses any combination of continuous/discrete variables.

• CART automatically performs variable selection.

• CART can establish interactions among variables.

**2.7.2.2 Disadvantages of CART**

• CART may have unstable decision trees.

• CART splits only by one variable.

• Non-parametric.

**2.7.3 C4.5**

C4.5 is an algorithm used to generate a decision tree which was also developed by Ross Quinlan. It is an extension of Quinlan’s ID3 algorithm. C4.5 generates decision trees which can be used for classification and therefore C4.5 is often referred to as statistical classifier (wiki/C4.5\_algorithm, 2017). It is better than the ID3 algorithm because it deals with both continuous and discrete attributes and also with the missing values and pruning trees after construction. C5.0 is the commercial successor of C4.5 because it is a lot faster, more memory efficient and used for building smaller decision trees.C4.5 performs by default a tree pruning process(wikipedia.org/wiki/Decision\_tree\_pruning, 2019). This leads to the formation of smaller trees, more simple rules and produces more intuitive interpretations.

C4.5 follows three steps in tree growth (Devinder, Rajiv, & Sunil, 2015):

• For splitting of categorical attributes, C4.5 follows the similar approach to ID3 algorithms. Continuous attributes always generate binary splits.

• Selecting attribute with the highest gain ratio.

• These steps are repeatedly applied to new tree branches and growth of the tree is stopped after checking of stop criterion. Information gain bias the attribute with more number of values. Thus, C4.5 uses Gain Ratio which is a less biased selection criterion.

**2.7.3.1 Advantages of C4.5**

• C4.5 is easy to implement.

• C4.5 builds models that can be easily

2.7.3.2 Disadvantages of C4.5

• A small variation in data can lead to different decision trees when using C4.5. • For a small training set, C4.5 does not work very well.

**2.7.4 Advantages of Decision Tree:**

* **Easy to Understand**: Decision tree output is very easy to understand even for people from non-analytical background. It does not require any statistical knowledge to read and interpret them. Its graphical representation is very intuitive and users can easily relate their hypothesis.
* **Useful in Data exploration:**Decision tree is one of the fastest way to identify most significant variables and relation between two or more variables. With the help of decision trees, we can create new variables / features that has better power to predict target variable. It can also be used in data exploration stage. For e.g., we are working on a problem where we have information available in hundreds of variables, there decision tree will help to identify most significant variable.
* Decision trees implicitly perform variable screening or feature selection.
* Decision trees require relatively little effort from users for data preparation.
* **Less data cleaning required:**It requires less data cleaning compared to some other modeling techniques. It is not influenced by outliers and missing values to a fair degree.
* **Data type is not a constraint:**It can handle both numerical and categorical variables. Can also *handle multi-output problems.*
* **Non-Parametric Method:**Decision tree is considered to be a non-parametric method. This means that decision trees have no assumptions about the space distribution and the classifier structure.
* Non-linear relationships between parameters do not affect tree performance.
* The number of hyper-parameters to be tuned is almost null.

**2.7.5 Disadvantages of Decision Tree:**

* Overfitting**:** Decision-tree learners can create over-complex trees that do not generalize the data well. This is called overfitting. Over fitting is one of the most practical difficulty for decision tree models. This problem gets solved by setting constraints on model parameters and pruning.
* Not fit for continuous variables: While working with continuous numerical variables, decision tree loses information, when it categorizes variables in different categories.
* Decision trees can be unstable because small variations in the data might result in a completely different tree being generated. This is called variance, which needs to be lowered by methods like bagging and boosting (towardsdatascience.com, 2017).
* Greedy algorithms cannot guarantee to return the globally optimal decision tree. This can be mitigated by training multiple trees, where the features and samples are randomly sampled with replacement.
* Decision tree learners create *biased* trees if some classes dominate. It is therefore recommended to balance the data set prior to fitting with the decision tree.
* Information gain in a decision tree with categorical variables gives a biased response for attributes with greater no. of categories.
* Generally, it gives low prediction accuracy for a dataset as compared to other machine learning algorithms.
* Calculations can become complex when there are many class label
  + 1. **Are Tree Based Models Better Than Linear Models?**

If one can use logistic regression for classification problems and linear regression for regression problems, why is there a need to use trees? Actually, we can use any algorithm. It is dependent on the type of problem we are solving. Some key factors which will help us to decide which algorithm to use:

* If the relationship between dependent & independent variable is well approximated by a linear model, linear regression will outperform tree based model.
* If there is a high non-linearity and complex relationship between dependent & independent variables, a tree model will outperform a classical regression method.
* To build a model which is easy to explain to people, a decision tree model will always do better than a linear model. Decision tree models are even simpler to interpret than linear regression.**e.g. Decison Tree Classification**

**2.7.7 Common Terms Used With Decision Trees:**

* **Root Node:** It represents entire population or sample and this further gets divided into two or more homogeneous sets.
  1. **Splitting:** It is a process of dividing a node into two or more sub-nodes.
  2. **Decision Node:** When a sub-node splits into further sub-nodes, then it is called decision node.
  3. **Leaf/ Terminal Node:** Nodes do not split is called Leaf or Terminal node.
* **Pruning:** When we remove sub-nodes of a decision node, this process is called pruning. You can say opposite process of splitting.
* **Branch / Sub-Tree:** A sub section of entire tree is called branch or sub-tree.
* **Parent and Child Node:** A node, which is divided into sub-nodes is called parent node of sub-nodes whereas sub-nodes are the child of parent node.
  1. **RELATED WORKS**

Jayanthi et al (Jayanthi, Goyal, & Ahson, 2008) presented the role of data mining in Human Resource Management Systems (HRMS). A deep understanding of the knowledge hidden in Human Resource (HR) data is vital to a firm's competitive position and organizational decision making. Analyzing the patterns

nd relationships in HR data is quite rare. The HR data is usually treated to answer queries. Because HR data primarily concerns transactional processing (getting data into the system, recording it for reporting purposes) it is necessary for HRMS to become more concerned with the quantifiable data. They show how data mining discovers and extracts useful patterns from this large data set to find observable patterns in HR. The paper demonstrates the ability of data mining in improving the quality of the decision-making process in HRMS and gives propositions regarding whether data-mining capabilities should lead to increased performance to sustain competitive advantage.

Hamidah et al (Hamidah, AbdulRazak, & O, 2009), in their work described the background of data mining, data mining in human resource application and an overview of talent management. Their literature study reveals that most researchers have discussed HR applications from different type of application. However, there should be more HR applications and Data Mining techniques applied to different problem domains in HRM field research in order to broaden our horizon of academic and practice work on HR applications using Data Mining techniques. Due to these reasons, they proposed the suitable Data Mining techniques for performance prediction based on initial experiment. They suggested for future work that the data in HR can be tested using other Data Mining techniques to find out the best accuracy of the techniques, especially for talent management data. Besides that, the relevance of attributes should be considered as a factor to the accuracy of the classifier. It was also suggested that in future experiment, attribute reduction experiment should take place in order to choose the relevant attributes for each of the factor. Once the relevant attributes are attained, the next modeling steps can be established to recommend. Finally, the ability to continuously change and obtain new understanding is the power of HR application, and this can be the HR applications of future work.

Nagadevara et al (Nagadevara, Srinivasan, & Valk, 2008), explored the relationship of withdrawal behaviors like lateness and absenteeism, job content, tenure and demographics on employee turnover in a rapidly growing sector like the Indian software industry. The unique aspect of this research was the use of five predictive data mining techniques (artificial neural networks, logistic regression, classification and regression trees, classification trees (C5.0), and discriminant analysis) on a sample data of 150 employees in a large software organization. The results of the study clearly show a relationship between withdrawal behaviors and employee turnover. This study raised several issues for future research. First, further research could explicitly collect data on demographic variables across a large sample of organizations to examine the relationship between demographic variables and turnover. Second, large scale data on variables in the past academic research which have a relationship with turnover can be collected longitudinally.

Such a data set will allow for more rigorous analysis and also a refined prediction model. Third, the context specific variables of employee turnover which emerged from this study would warrant a deeper understanding of the phenomena. There is a need for more empirical research and in particular, longitudinal research using data within corporations to refine the model. Last, more research needs to be conducted on various samples to confirm the validation of the theoretical model and the prediction model proposed in the study.

Wei-Chiang and Ruey-Ming (Wie-Chiang & Ruey-Ming, 2007), in their work explored the feasibility of applying the *Logit* and *Probit* (Moore, 2007)models, which have been successfully applied to solve nonlinear classification and regression problems, to employee voluntary turnover predictions. A numerical example involving voluntary turnover data of 150 professional employees drawn from a motor marketing enterprise in central Taiwan was used with a usable sample size of 132. The data set was divided into two parts, the modeling dataset and the testing data set. The modeling data set was used to test the *logit* and *probit* models. The testing data set was not used for either model building or selection, and was used for estimating model performance when applied to future data. The empirical results of their investigation revealed that the proposed models have high prediction capabilities and that the two (*logit* and *probit) models* also provide a promising alternative for predicting employee turnover in human resource management. The authors suggested that turnover research should move in new directions based on new assumptions and methodologies, which would raise new issues and problems (such as the use of neural networks and support vector machines to conduct classification problem for detecting stayer or leaver).

# References

(2000, 04 12). Retrieved 04 12, 2019, from www.whatishumanresource.com: http://www.whatishumanresource.com/history--origin-of-performance-appraisal

Anuj, R., & Robin, P. (2013). “Survey on Decision Tree Classification algorithms for the evaluation of Student Performance. *International Journal Of Computer Technologies*.

Badr, H., & Abdelkarim, M. (2011). A comparative study of decision tree ID3 and C4.5. *IJACSA*.

Bhumika Gupta, A. R. (2017). Analysis of Various Decision Tree Algorithms for Classification in Data Mining. *International Journal of Computer Applications (0975 – 8887)*, 5.

Brid, R. S. (2018, October 2). *https://medium.com/greyatom/decision-trees-a-simple-way-to-visualize-a-decision-dc506a403aeb*. Retrieved from medium.com: https://medium.com/greyatom/decision-trees-a-simple-way-to-visualize-a-decision-dc506a403aeb

Chen, J. (2018, January 22). *Performance-Based Compensation*. Retrieved from www.investopedia.com: https://www.investopedia.com/terms/p/performancecompensation.asp

Devinder, K., Rajiv, B., & Sunil, K. G. (2015, September 15). Implementation of Enhanced Decision Tree Algorithm on Traffic Accident Analysis. *IJSRT*.

Dulewicz, V. (1989). Performance appraisal and counselling. In V. Dulewicz, *Performance appraisal and counselling* (pp. 645-649). New York: John Wiley & Sons.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 327-358.

Gary Roberts, M. P. (2007, May). Why Employees Dislike Performance Appraisals. *Regent Global Business Review*, p. 8. Retrieved from www.regent.edu.

Hamidah, J., AbdulRazak, H., & O, Z. A. (2009). Towards Applying Data Mining Techniques for Talent Managements, 2009. *International Conference on Computer Engineering and Applications.* *volume 2.* Singapore: IACSIT Press.

Hemlata, C. (2013). ID3 Modification and Implementation in Data Mining. *International Journal of Computer Applications*, 0975-8887.

*https://en.wikipedia.org/wiki/Ross\_Quinlan*. (2019). Retrieved from wikipedia.com.

*https://sites.google.com/site/whatishumanresource/traditional-methods-of-performance-appraisal*. (2000). Retrieved 04 15, 2019, from https://sites.google.com/site/whatishumanresource: https://sites.google.com/site/whatishumanresource/traditional-methods-of-performance-appraisal

Jayanthi, R., Goyal, D., & Ahson, S. (2008). Data Mining Techniques for Better Decisions. *International Journal of Business Information Systems*, 464-481.

*modern methods of performance appraisal*. (2000, june). Retrieved from whatishumanresource.com: https://sites.google.com/site/whatishumanresource/modern-methods-of-performance-appraisal

Moore, C. (2007). An Introduction to Logistic and Probit Regression Models. Texas, United States of America.

Nagadevara, V., Srinivasan, V., & Valk, R. (2008). Establishing a Link between Employee Turnover and Withdrawal Behaviours. *Application of Data Mining Techniques ,Research and Practice in Human Resource Management,*, 81-99.

*Performance appraisals: What India Inc. plans to offer to its employees*. (2019, Feb). Retrieved from timesjobs.com: https://content.timesjobs.com/performance-appraisals-what-india-inc-plans-to-offer-to-its-employees/articleshow/68112507.cms

Qing-Yun, D., Chun-ping, Z., & Hao, W. (2016). “Research of Decision Tree Classification Algorithm in Data Mining. *International Journal of Database Theory and Application Vol.9*, 1-8.

*towardsdatascience.com*. (2017). Retrieved from towardsdatascience.com: https://towardsdatascience.com/boosting-the-accuracy-of-your-machine-learning-models-f878d6a2d185

Wie-Chiang, H., & Ruey-Ming, C. (2007). A Comparative Test of Two Employee Turnover Prediction Models. *International Journal of Management, Vol.24 No.2*, 216-229.

*wiki/C4.5\_algorithm.* (2017, march 4). Retrieved from wikipedia: https://en.wikipedia.org/wiki/C4.5\_algorithm

wikipedia. (2019). *https://en.wikipedia.org/wiki/Ross\_Quinlan*. Retrieved from wikipedia.com.

*wikipedia.org/wiki/Decision\_tree\_pruning*. (2019, june 9). Retrieved from wikipedia: https://en.wikipedia.org/wiki/Decision\_tree\_pruning