

Job Industry Analysis based on Decision Tree Algorithm

Yunyang Liao, Yunxiao He, Juanjuan Mu, Jiawen Chen, Juan Li*

Wuhan Business University, Wuhan, China

Abstract: Influenced by various factors such as market economic policies, the employment problem of college graduates has become more and more significant. Therefore, the design of employment job industry analysis based on decision tree algorithm is significant for job seekers. Aiming at the influencing factors such as region, specialty, salary and so on, combined with the prediction model data preparation, data pre-processing, correlation analysis of data, visualization technology, to realize the job industry analysis based on the decision tree algorithm, to provide certain reference basis for the job seekers in the job-seeking project.

Keywords: Data Visualization, Decision Tree Algorithm, Post Data Analysis, Data Preprocessing, Word Cloud Analysis.

1. Introduction

In recent years, with the development of the economy and technological progress, more and more industries are changing, and with the increase in the number of college graduates, the employment of college students has become very difficult, so job industry analysis has become more and more necessary. Job industry analysis has an important role for enterprises and individuals. For enterprises, job industry analysis can help them understand the market trend and competition pattern, so as to make more effective strategies and plans. In addition, by analyzing the flow of talent within the industry, companies can better recruit and train talent to improve their competitiveness and performance. For individuals, job industry analysis can help them better understand the job market and career trends so that they can make more informed career choices. By understanding the working environment and treatment within the industry, individuals can better plan their career development and improve their career competitiveness.

To summarize, job industry analysis has become increasingly important in the current society. It can help enterprises and individuals better understand the market trends and competitive landscape, so as to make more informed decisions and planning. Therefore, this paper focuses on the research and application of job industry analysis, will crawl from the major job boards to pre-process

and categorize the information, the processed recruitment data imported into the database, extract the number of industries and the number of specific people needed for each recruitment position, the use of decision tree algorithms modeling, through the statistics to get the number of people need to be recruited in different industries and the proportion of them. Finally, the pie chart form to the user to visualize. It provides an important reference for enterprise development and personal employment, so as to better cope with future challenges and opportunities.

2. Introduction to Decision Tree Algorithms and Models

A decision tree is a predictive model that represents a mapping relationship between object attributes and object values. Each node in the tree represents a particular object, while each forked path represents a possible attribute value, and each leaf node corresponds to the value of the object represented by the path taken from the root node to that leaf node.

The machine learning technique of generating decision trees from data is called decision tree learning, commonly known as a decision tree. A decision tree contains three types of nodes: decision nodes: usually represented by rectangular boxes; opportunity nodes: usually represented by circles; and end nodes: usually represented by triangles.

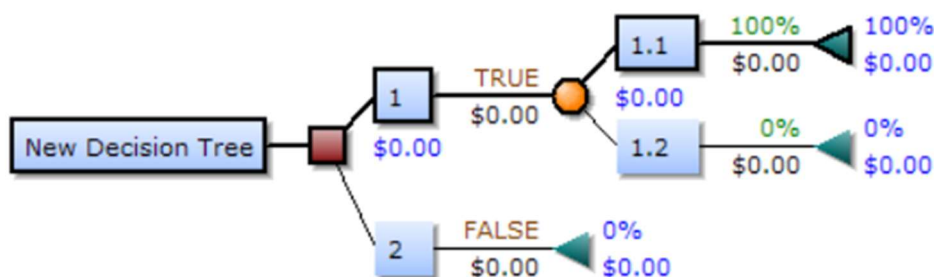


Figure 1. Decision tree view

A decision tree is a tree-like structure in which each leaf node corresponds to a classification, and a non-leaf node corresponds to a division on an attribute that divides the samples into a number of subsets based on their different

values on that attribute. For non-pure leaf nodes, the majority class labeling gives the class to which the samples arriving at this node belong. The central problem in constructing a decision tree is how to choose the appropriate attribute to do

- [5] Yang Zhengan. Exploration of the use of big data visualization and analysis technology[J]. Science and Technology Innovation and Application, 2023, 13 (32): 46-49.