Article

Performance Analysis of the CHAID Algorithm for Accuracy

Yeling Yang [1], **Feng Yi** [2,*], **Chuancheng Deng** [2] and **Guang Sun** [2,*]

[1] School of Physical Education, South China University of Technology, Guangzhou 510641, China; yelingyang2023@163.com

[2] Hunan University of Finance and Economics, Changsha 410021, China; 202005380235@mails.hufe.edu.cn

* Correspondence: yifeng@hufe.edu.cn (F.Y.); sunguang@hufe.edu.cn (G.S.); Tel.: +86-186-2754-5158 (G.S.)

Abstract: The chi-squared automatic interaction detector (CHAID) algorithm is considered to be one of the most used supervised learning methods as it is adaptable to solving any kind of problem at hand. We are keenly aware of the non-linear relationships among CHAID maps, and they can empower predictive models with stability. However, we do not precisely know how high its accuracy. To determine the perfect scope the CHAID algorithm fits into, this paper presented an analysis of the accuracy of the CHAID algorithm. We introduced the causes, applicable conditions, and application scope of the CHAID algorithm, and then highlight the differences in the branching principles between the CHAID algorithm and several other common decision tree algorithms, which is the first step towards performing a basic analysis of CHAID algorithm. We next employed an actual branching case to help us better understand the CHAID algorithm. Specifically, we used vehicle customer satisfaction data to compare multiple decision tree algorithms and cited some factors that affect the accuracy and some corresponding countermeasures that are more conducive to obtaining accurate results. The results showed that CHAID can analyze the data very well and reliably detect significantly correlated factors. This paper presents the information required to understand the CHAID algorithm, thereby enabling better choices when the use of decision tree algorithms is warranted.

Keywords: CHAID algorithm; chi-square detection; decision tree algorithm; branching principle

MSC: 65S99

1. Introduction

Since the 1990s, with the rapid development of information technology, the application of database systems has become more widespread, and at the same time database technology has entered a completely new stage, from solely managing some simple data to managing a wide variety of complex data such as images, videos, audio, graphics, and electronic files generated by various devices; therefore, the amount of data to process has become larger and larger [1].

In this era of such advanced information, the vast amount of information does not only bring us benefits but also us many negative effects. The most important factor in the influence on negative effects is that effective information is hard to refine, and too much meaningless data will inevitably cause issues in terms of the loss of meaningful knowledge. This is what John Nalsbert calls the "information-rich but knowledge-poor" dilemma [2]. With the original functions of database systems, people could not discover the relationships and rules implied in data, and could not predict future trends based on existing data. There is a lack of methods to uncover the hidden value behind data. To solve this problem, there is an urgent need for a technology that can analyze large amounts of information more deeply, obtain insight into its hidden value, and make seemingly useless data useful [3].

Decision trees are an effective way to generate classifiers from data, and represent the class of the most widely applied logical methods for this purpose [4,5]. In 1980, Kass

first proposed the chi-squared automatic interaction detector (CHAID), which is a tool used to discover relationships between variables, a decision tree technique based on an adjusted significance test (Bonferroni test) [6,7]. It divides the respondents into several groups according to the relationship between the underlying variable and the dependent variable, and then each group into several groups. Dependent variables are usually some key indicators, such as the level of use, purchase intention, etc. A dendrogram is displayed after each program run. The top is a collection of all respondents, the following is a subset of two or more branches, and the CHAID classification is based on a dependent variable [8]. Classification and regression tree (CART) is a decision tree algorithm based on the binary tree structure of the Gini coefficient that supports regression and classification. It adopts the method of pruning, which is mainly used in the classification prediction of small- and medium-sized data [9,10]. Iterative Dichotomiser (ID3) is a decision tree algorithm based on information entropy that can only deal with categorical variables and is easily overfit, which mainly uses classification prediction on small data sets [10,11].

In practice, CHAID is often used in the context of direct selling, selecting consumer groups and predicting their responses, and determining how some variables affect other variables [12,13], while other early applications are in the research fields of medicine and psychiatry [14,15], as well as engineering project cost control, financial risk warning, and fire reception and handling analysis [16,17]. We are been starkly aware of the non-linear relationships among CHAID maps, and they can empower predictive models with stability [18]. However, we do not precisely know how high its accuracy is. To find out the perfect scope the CHAID algorithm fits into, this paper presented an analysis of the accuracy of the CHAID algorithm. of the analysis was based on the introduction of the CHAID algorithm in the second part to obtain insight into the causes of the accuracy of the CHAID algorithm and the difference between several other commonly used decision tree algorithms [19,20]. In the third part, we used IBM SPSS Statistics 26.0 software and the Python 3.7 language to realize the differences between multiple decision trees while further comparing the differences between the three on the big data set of bike-sharing requirements.

2. CHAID Algorithm and Chi-Square Detection

The core idea of the CHAID algorithm is to optimally divide the samples according to the given target variable and the selected feature index (such as a predictive variable), and group the contingency table to automatically judge according to the significance of the chi-square test. The field selection of the CHAID algorithm is performed by using the chi-square test.

2.1. Classification Process of the CHAID Algorithm

The target variables for categorization were first selected and then cross-categorized with the target variables to produce a series of 2-D taxonomic tables.

The chi-square values of the two-dimensional classification table were calculated separately, the size of the p-values was compared, the two-dimensional table with the lowest p-value was used as the best initial classification table, and then the categorical variable was used as the first-level variable of the CHAID decision tree.

Based on the best initial classification table, we continued to classify the target variables to obtain the second and tertiary variables of the CHAID decision tree.

The process was repeated until the p-value was greater than the set statistically significant alpha value or until the classification stopped when all variables were classified.

2.2. Introduction of Chi-Square Detection

2.2.1. The Concept and Significance of Chi-Square Detection

Chi-square detection is the deviation between the theoretical value and the actual value of the statistical sample. The degree of deviation determines the size of the chi-square value. If the degree of deviation is smaller, the chi-square value will be smaller. On

the contrary, the larger the chi-square value is, if the actual value and the theoretical value are calculated, the chi-square value is equal to 0.

2.2.2. The Basic Idea of Chi-Square Detection

The chi-square test is a commonly used hypothesis test based on the chi-square distribution.

Firstly, assume the following hypothesis: the expected frequency is not different from the observed frequency. Under this premise, the chi-squared values of the theoretical and actual values are calculated. The probability that the hypothesis holds under the current statistical sample can be determined based on the chi-squared distribution and degrees of freedom. Table 1 shows some chi-square value probability .

Table 1. A probability table of partial chi-square distributions.

$P(x^2 \geq k)$	k	$P(x^2 \geq k)$	k
0.50	0.455	0.05	3.841
0.40	0.708	0.025	5.024
0.25	1.323	0.010	6.635
0.15	2.072	0.005	7.879
0.10	2.706	0.001	10.828

If the *p*-value is small, then the probability of the hypothesis H_0 holding is small; the hypothesis H_0 should therefore be rejected, indicating a significant difference between the theoretical value and the actual value; if the *p*-value is large, the hypothesis H_0 cannot be rejected and there is a difference between the theoretical value and the actual situation represented by the actual value.

2.2.3. Formula for Chi-Square Detection

In this formula, where x^2 is the chi-square value obtained by the actual and theoretical values, *k* is the number of cells in the two-dimensional table, A_i is the actual value of i, E_i is the expected value of i, n is the total number of samples, p_i is the expected frequency of i, and E_i = (n: the total number of samples) × (p_i: the expected probability of i).

$$x^2 = \sum \frac{(A-E)^2}{E} = \sum_{i=1}^{k} \frac{(A_i - E_i)^2}{E_i} = \sum_{i=1}^{k} \frac{(A_i - np_i)^2}{np_i} \tag{1}$$

$$(i = 1,2,3,\ldots, k)$$

2.2.4. Steps for Chi-Square Detection

Firstly, assuming that H_0 holds, we determined the degree of freedom (degree of freedom = (row − 1) × (column − 1), where the row and the column are the number of rows and columns in the two-dimensional Table 2). Then the theoretical frequency number was obtained with the maximum likelihood estimation, At last, we substituted it into the formula to solve.

Table 2. Chi-square detection sample data.

	Male	Female	
Reconcile	15 (55)	95 (55)	110
Do not reconcile	85 (45)	5 (45)	90
	100	100	200

Assume that, in the above-shown example, whether or not to reconcile has no relationship with gender.

Maximum likelihood estimation yields the expected value E_i: $E_1 = 100 \times 110/200 = 55$ (100 is the number of men surveyed and 110/200 is the number of people in all the surveys, from which the likelihood estimate of men reconciled).

$$E_2 = 100 \times 110/200 = 55$$
$$E_3 = 100 \times 90/200 = 45$$
$$E_4 = 100 \times 90/200 = 45$$

There is a significant difference between the value obtained in maximum likelihood estimation (inside of parentheses) and the actual value (outside of parentheses).

Generation formula:

$$x^2 = \sum_{i=1}^{k} \frac{(A_i - np_i)^2}{np_i} = \frac{(95-55)^2}{55} + \frac{(15-55)^2}{55} + \frac{(85-45)^2}{45} + \frac{(5-45)^2}{45} = 129.3 > 10.828 \quad (2)$$

Because the desired result, $x^2 > 10.828$, means that the null hypothesis of 0.001 might hold, there is a 99.9% probability is that no reconciliation is significantly associated with gender.

2.3. The Field Selection Procedure for CHAID Algorithm

After understanding the process of chi-square detection, we can take a look at the field selection process of the CHAID algorithm.

The CHAID field was selected using a chi-square statistic. The chi-square distribution stands for whether the two category fields are distributed or not. The numerical field will help discretizing the category fields which used to analyze whether they are related to the target field. Larger indicates a more significant relationship, and it is otherwise not obvious.

The row of the above-shown Table 3 is the income level, and the column shows customer intent to buy a computer. The calculation of chi-square statistics should first draw the frequency from the data, and then find their expectations. Then, the squared difference between the two is found and summed.

The Table 3 shows the actual data derived from the actual data statistics. The Table 4 presents seeking expectations, which is equivalent to thinking that the two options are independent, so one can multiply them directly by the probability and then multiply the total.

The Table 5 shows the variance. After summing, x^2 is equal to 0.57, and the corresponding probability is 75%, indicating that the two are relatively weak. The full chi-square probabilities are shown as follows (Table 6).

Table 3. Income and actual data.

Income	Yes	No	Total
high	2	2	4
medium	4	2	6
low	3	1	4
Total	9	5	14

Table 4. Income and computer purchase expectation data.

Income	Yes	No	Total
high	2.571429	1.428571	4
medium	3.857143	2.142857	6
low	2.571429	1.428571	4
Total	9	5	14

Table 5. The squared difference between the expected data and the actual data.

Income	Yes	No
high	0.126984	0.228571
medium	0.005291	0.009524
low	0.071429	0.128571

Table 6. All chi-square probabilities.

Age	Student	Credit_Rating	Income
x^2 = 3.54667	x^2 = 2.80000	x^2 = 0.93333	x^2 = 0.57037
P = 0.16977	P = 0.09426	P = 0.33400	P = 0.75188

After comparison, we can see that the age feature value is the largest in the x^2 = 3.54667 value, indicating that age is the most closely related to computer purchase intentions, so we chose age as the variable of the decision tree to produce the leaf node of the next level.

3. Comparison of Decision Tree Algorithm and Accuracy Analysis of the CHAID Algorithm

The three most commonly used decision tree algorithms are CHAID, CART, and ID3, including the latest C4.5 and even C5.0.

The CHAID algorithm has a long history. According to the principle of local optimization, CHAID uses the chi-square test to select the independent variables that affect the dependent variable the most. Then, because the independent variables may have many different categories, the CHAID algorithm will generate equal amounts of leaf nodes according to the number of categories of the independent variable, so the CHAID algorithm is a multi-fork tree.

The CHAID method is optimal when the predictor variable is a categorical variable. For continuous variables, CHAID automatically divides the continuous variables into 10 segments, but there may be omissions.

The CHAID algorithm uses the chi-square detection method in statistics, and because of the chi-square detection method, CHAID has a good mathematical theoretical basis in branch calculation, and its credibility and its accuracy are relatively high.

On this basis, the CHAID algorithm uses the pre-pruning method; pre-pruning is pruning before dividing and generating the decision tree under constant pruning, so pre-pruning not only reduces the training time overhead and testing time overhead of the CHAID decision tree but also reduces the risk of overfitting. On the other hand, some pre-pruning divisions may not improve the generalization performance, or may even cause a temporary decrease in generalization performance, but subsequent divisions based on this division may lead to a significant improvement in generalization performance, thus posing the risk of underfitting. Therefore, if the number of pruning is maintained in a good interval when the amount of data is sufficient and the types are mostly categorical variables, the risk of underfitting of the CHAID algorithm will be further reduced and the accuracy will be further improved.

In Figure 1, we can clearly see the significant correlation between features and car purchase intentions. In the first layer of the decision tree, we see that x^2 = 339.064 of safety terms is the largest chi-square value and the most significant correlation among all feature terms. The remaining feature terms calculate the chi-square value again on the basis of the safety term, so we can clearly see the correlation degree of each feature, and this can also be used as a model to predict whether someone will buy a certain kind of car according to their characteristics.

As for the CART (Classification and Regression Tree) algorithm, its segmentation logic is the same as that for CHAID, and the division of each layer is based on the test and

selection of all independent variables. However, the test standard used by CART is not the chi-square test, but the indicators of impurity, such as the Gini coefficient (Gini). The biggest difference between the two is that CHAID adopts the principle of local optimization, that is, the nodes are irrelevant to each other. After a node is determined, the following growth process is carried out completely within the node. CART, on the other hand, focuses on the overall optimization and adopts the post-pruning method, which makes the tree grow as much as possible, and then cuts the tree back and evaluates the non-leaf nodes in the tree from bottom to top, so the cost of training time is much larger than the pre-pruning decision tree.
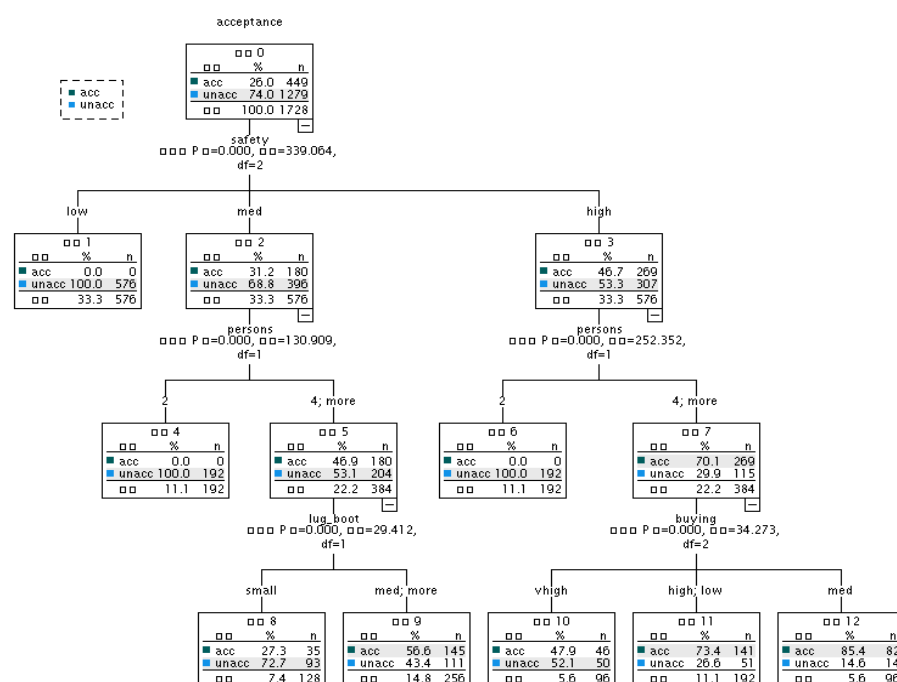

Figure 1. CHAID decision tree.

As for the CART (Classification and Regression Tree) algorithm, its segmentation logic is the same as that for CHAID, and the division of each layer is based on the test and selection of all independent variables. However, the test standard used by CART is not the chi-square test, but the indicators of impurity, such as the Gini coefficient (Gini). The biggest difference between the two is that CHAID adopts the principle of local optimization, that is, the nodes are irrelevant to each other. After a node is determined, the following growth process is carried out completely within the node. CART, on the other hand, focuses on the overall optimization and adopts the post-pruning method, which makes the tree grow as much as possible, and then cuts the tree back and evaluates the non-leaf nodes in the tree from bottom to top, so the cost of training time is much larger than the pre-pruning decision tree.

In Figure 2, we can see that the gini = 0.39 of the safety term in the first layer of the decision tree is the smallest and the most relevant among all the feature terms. However, the feature term appears repeatedly, so it is impossible to intuitively see the importance of each feature, and it is more suitable for the model to predict whether someone will buy a certain kind of car according to their characteristics.
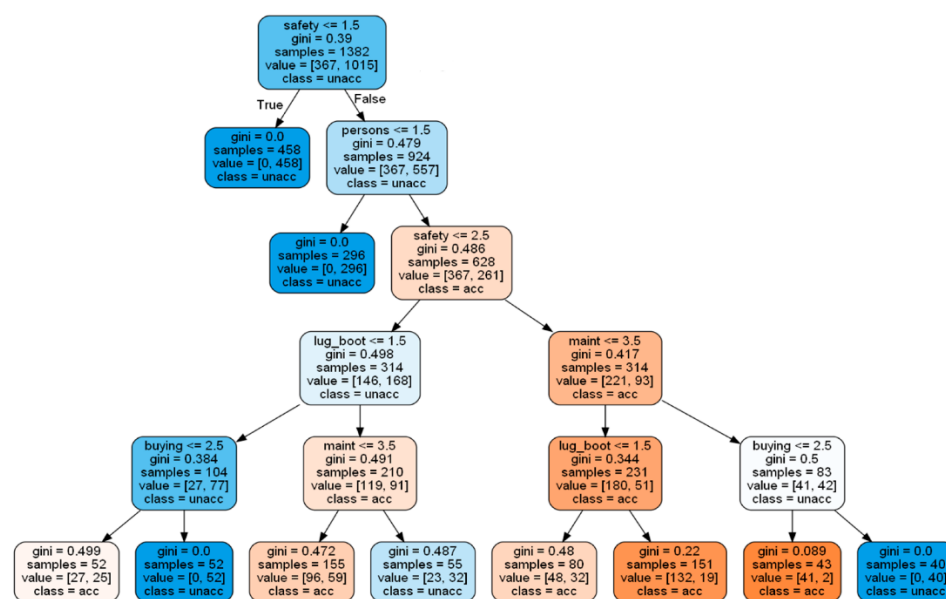
Figure 2. CART decision tree.

If there is missing data in the independent variable, CART can be used to find alternative data to replace the missing value, while CHAID takes the missing value as a separate type of value.

Among CART and CHAID, one is a binary tree and the other is a multi-fork tree; CART selects the best binary cut in each branch, so a variable is likely to be used multiple times in different trees; CHAID divides multiple statistically significant branches for one variable at a time, which will grow faster, but the support of the sub rapidly decreases compared with CART, approaching a bloated and unstable tree more quickly.

Therefore, after the number of data categories in the data set increases to a certain extent, the accuracy of the CHAID algorithm will have a large decrease compared with the CART algorithm. The number of features of the data can be reduced by removing some data irrelevant to the target data during the data cleaning so as to improve the accuracy of the CHAID algorithm.

The ID3 (Iterative Dichotomiser) algorithm and CART are in the same period; its biggest feature is that the independent variable selection criteria are based on the measure of information gain: the attribute with the highest information gain is selected as the split attribute of the node, and the result is the minimum information required to classify the segmented node, which is also an idea of division purity. As for the later development of C4.5, which can be understood as the development version of ID3, the main difference between the two is that C4.5 uses the information gain rate instead of the information gain measure in ID3. The main reason for such a replacement is that the information gain measure has a disadvantage, that is, it tends to choose attributes with a large number of values. As an extreme example, for the division of Member_Id, each Id is a pure group, but such a division has no practical significance. The information gain rate adopted by C4.5 can overcome this disadvantage. It adds a piece of split information to normalize constraints on the information gain. Additionally, C5.0 is the latest version. Compared with C4.5, C5.0 uses less memory and builds a smaller rule set than C4.5, while also being more accurate.

In Figure 3, we can see that the features of the first layer of the decision tree term are persons, where persons entropy = 0.827 is the largest entropy among all, but the features will repeat, so can not intuitively see the importance of features; it is therefore more suitable for the model to make predictions according to the characteristics of an individual.
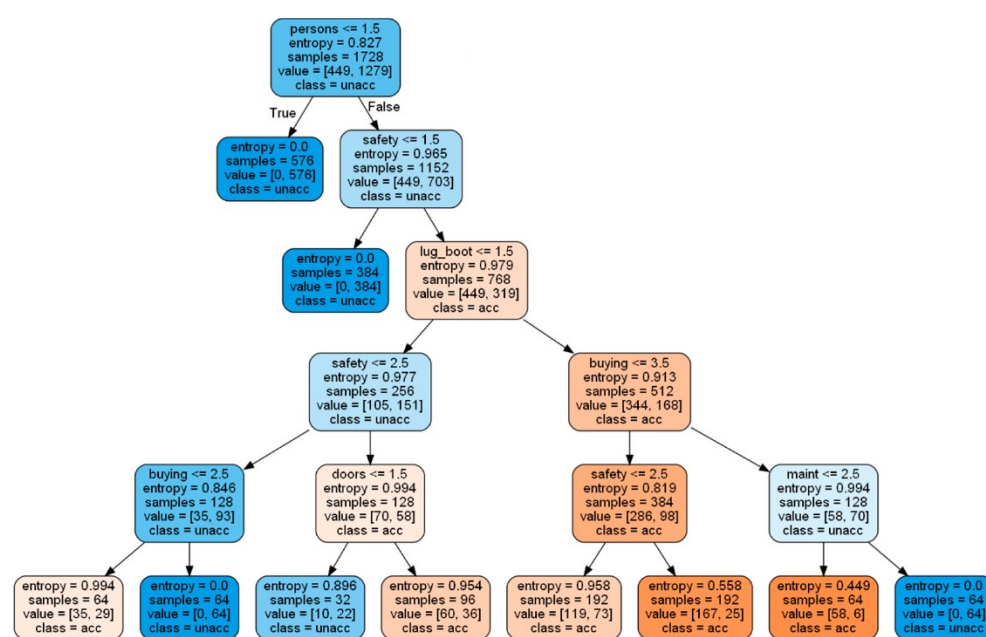
Figure 3. ID3 decision tree.

We also compared classification modeling using CHAID, CART, and ID3 decision tree algorithms on the large data set of bike-sharing requirements.

Table 7 clearly illustrates the distinctions among the three decision tree algorithms (CHAID, CART, and ID3), as well as the modeling and detection outcomes of these algorithms on large datasets of bike-sharing demand. The detection accuracy column shows that CHAID had a 92.3% accuracy on the shared bike data test set, indicating that it has excellent results in classification and prediction modeling on large data sets. CART, on the other hand, had an 85.7% accuracy, suggesting that its classification and prediction modeling on large data sets are not as good as CHAID and it is better suited for small and medium-sized data sets. ID3 had a 69.1% accuracy on the shared bike data test set, which indicates that its classification and prediction modeling on large data sets are prone to overfitting and it is better suited for small data sets.

Table 7. Comparing the performance of three decision tree models and their experimental results on large datasets of bike-sharing demand.

Decision Tree Algorithm	Branch Principle	Tree Structure	Support Model	Continuous Value Processing	Missing Value Processing	Pruning	Detection Accuracy
CHAID	Chi-square value	multiway tree	Classification, regression	support	support	prepruning	92.3%
CART	Gink coefficient	binary tree	Classification, regression	support	support	post-pruning	85.7%
ID3	information gain	multiway tree	Classification	nonsupport	nonsupport	nonsupport	69.1%

4. Discussion

The decision tree algorithm belongs to the supervised learning machine learning method, and it is a commonly used technology in data mining. It can be used to classify the analyzed data, and it can also be used for prediction. Common algorithms are CHAID, CART, ID3, C4.5, C5.0, and so on. The second part is the study of the core idea of the

CHAID decision tree algorithm and the classification process, the specific steps of the classification process, and the principle formula of the CHAID decision tree algorithm in the branching process. The third part is a comparison between the CHAID decision tree algorithm and other commonly used decision tree algorithms and a partial analysis of the accuracy of the CHAID algorithm. In this study, we provided an example of factor analysis for automobile satisfaction and implement an in-depth comparison using multiple decision tree algorithms. Additionally, we modeled and evaluated the accuracy of three decision tree algorithms—CHAID, CART, and ID3—on large datasets of bike-sharing demand. Through a thorough analysis of their accuracy, we examined the performance of these three decision tree algorithms on large datasets.

The CHAID algorithm uses chi-square detection and pre-pruning in the branch method, the CART is the Gini coefficient (Gini) pruning in the branch method. ID3 is a measure based on information gain, and C4.5 and C5.0 adopt information gain rate. CHAID and CART can process continuous data, while ID3 cannot; ID3 cannot process data with missing values, CHAID and CART can; with too many features in the data, CART and ID3 easily overfit, and CHAID is more stable than the previous two, and obtain more accurate results.

This paper leads us toward an in-depth understanding of these algorithms, letting us choose a relatively good decision tree algorithm for data mining according to our specific data. The application of the CHAID algorithm can also enable some countermeasures to address some factors affecting accuracy, so it is more conducive to obtaining a more accurate and better result.

Although there are some works have analyzed and compared CHAID, CART, and ID3, none of them based on the big data sets. We tested the CHAID, CART, and ID3 algorithms with a shared bike system which provides the big data set. This data set can enable a more in-depth understanding of the CHAID algorithm instead of the practical improvement on the CHAID algorithm.

In the next stage of research, we will find more data to compare these several decision tree algorithms, and further discuss the accuracy of the CHAID algorithm for use with different data. According to the experimental results, we can specifically summarize the differences between these decision tree algorithms and the influence of different data on the accuracy of the CHAID algorithm, as well as the differences between the accuracy of each algorithm. When we choose a decision tree algorithm according to the specific situation of the data, we can achieve a clearer and more intuitive understanding of the data and have a better understanding of the accuracy analysis of the CHAID algorithm.

Author Contributions: Conceptualization, Y.Y.; Methodology, C.D.; Validation, G.S.; Formal analysis, F.Y. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: See address to get sample data set https://tianchi.aliyun.com/dataset/54174.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ture, M.; Tokatli, F.; Kurt, I. Using Kaplan–Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4. 5 and ID3) in determining recurrence-free survival of breast cancer patients. *Expert Syst. Appl.* **2009**, *36*, 2017–2026.
2. Liu, Y.; Zhang, L.; Chen, Y. A comparative study of decision tree algorithms for predicting stock prices. *J. Intell. Fuzzy Syst.* **2021**, *40*, 7459–7470.
3. Zheng, X.; Sun, H.; Lu, X.; Xie, W. Rotation-Invariant Attention Network for Hyperspectral Image Classification. *IEEE Trans. Image Process.* **2022**, *31*, 4251–4265.
4. Li, Y.; Ren, J.; Yan, Y.; Liu, Q.; Ma, P.; Petrovski, A.; Sun, H. CBANet: An End-to-end Cross Band 2-D Attention Network for Hyperspectral Change Detection in Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **2023**, *in press*.
5. Akin, M.; Ecevit, E.; Barbara, M.R. Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. *Plant Cell Tissue Organ Cult.* **2017**, *128*, 303–316.
6. Zheng, X.; Chen, X.; Lu, X. Visible-Infrared Person Re-Identification via Partially Interactive Collaboration. *IEEE Trans. Image Process.* **2022**, *31*, 6951–6963.
7. Kristof. C.; Stefan. L.; Geert. V. A comparative study of decision tree algorithms for predicting customer churn in mobile tele-communications industry. *Decis. Support Syst.* **2017**, *95*, 27–36.
8. Zhang, J.; Zhang, Y.; Liu, J. A comparative study of decision tree algorithms for predicting customer purchase behavior. *Int. J. Ind. Eng. Comput.* **2019**, *10*, 299–310.
9. Abidin, A.Z.; Abdullah, L. A comparative study of decision tree algorithms for predicting student academic performance. *Int. J. Emerg. Technol. Learn.* **2019**, *14*, 190–203.
10. Xie, G.; Ren, J.; Marshall, S.; Zhao, H.; Li, R.; Chen, R. Self-attention Enhanced Deep Residual Network for Spatial Image Steganalysis. *Digit. Signal Process.* **2023**, *in press*.
11. Gao, Y.; Zhu, S.; Wang, J. A comparative study of decision tree algorithms for predicting employee turnover. *Int. J. Hum. Resour. Manag.* **2019**, *30*, 3084–3103.
12. Shiju. S. and Remya. R. N. Comparative Analysis of Decision Tree Algorithms: ID3, C4.5 and Random Forest. *Comput. Intell. Data Min.* **2014**, *1*, 549-562.
13. Liu, Y.; Zhang, L. A comparative study of decision tree algorithms for predicting the success rate of crowdfunding projects. *Appl. Soft Comput.* **2021**, *101*, 107074.
14. Zhang, Y.; Li, Y.; Liu, J. A comparative study of decision tree algorithms for credit scoring. *Expert Syst. Appl.* **2021**, *157*, 113466.
15. Isra'a. A. Zriqat; Ahmad. M. A. and Mohammad. A. A Comparative Study for Predicting Heart Diseases Using Data Mining Classification Methods. *Int. J. Comput. Sci. Inf. Secur. (IJCSIS)* **2016**, *14*, 868–879.
16. Prajwala, T.R. A Comparative Study on Decision Tree and Random Forest Using R Tool. *Int. J. Adv. Res. Comput. Commun. Eng.* **2015**, *4*, 196–199.
17. Al-Masri, E.; Shatnawi, R. A comparative study of decision tree algorithms for intrusion detection systems. *J. Netw. Comput. Appl.* **2021**, *185*, 103040.
18. Yang, Y.; Li, S.; Wang, J. A comparative study of decision tree algorithms for predicting traffic accidents. *J. Adv. Transp.* **2020**, *2020*, 21–35.
19. Liu, P.; Wang, F. A comparative study of decision tree algorithms for predicting online shopping behavior. *Int. J. Ind. Eng. Comput.* **2019**, *10*, 115–126.
20. Chen, L.; Zhang, X.; Zhang, Y. A comparative study of decision tree algorithms for predicting customer churn. *J. Intell. Fuzzy Syst.* **2019**, *36*, 5517–5529.