

Parents' Sum of Salaries Analyses towards School Tuition Fee Arrears Potential with Decision Tree Method

I D Iskandar^{*1}

¹Universitas Bina Sarana Informatika Sistem Informasi Kampus Tasikmalaya, Jl. Tanuwijaya No.4, Tasikmalaya, 46113, Indonesia

E-mail: iqbal.iql@bsi.ac.id¹

Masuk: 17 April 2019, direvisi: 15 Juni 2019, diterima: 8 Agustus 2019

Abstrak. Dana Sumbangan Pembinaan Pendidikan (SPP) biasanya dipergunakan untuk pendanaan berbagai macam kegiatan di sekolah, mulai dari penggajian guru honorer yang terdapat pada sekolah swasta maupun negeri, pengadaan alat praktik, mencetak soal-soal ujian, serta operasional lainnya. Masalah yang terkait dengan SPP di lingkungan berdasarkan data yang diperoleh banyaknya siswa-siswa yang sering sekali menunggak pembayaran SPP sampai beberapa bulan, bahkan sampai siswa-siswa tersebut lulus. Atas masalah tersebut dilakukan penelitian untuk menganalisa potensi menunggaknya SPP siswa berdasarkan besaran pendapatan orangtua dengan pendekatan *Business Intelligence* menggunakan metode *decision tree*. Hasil analisis yang didapatkan siswa yang memiliki orangtua dengan pendapatan atau penghasilan kurang dari Rp 672.500,00 akan berpotensi menunggak pembayaran SPP lebih dari Rp 900.000,00 pada setiap bulannya, sedangkan orangtua siswa yang memiliki penghasilan diatas Rp 672.500,00 berpotensi memiliki tunggakan kurang dari Rp 900.000,00 atau tidak menunggak pembayaran SPP. Untuk evaluasi efektifitas algoritma *decision tree* terhadap pengolahan data, memiliki nilai akurasi sebesar 95.97%, dengan presisi 94.96% nilai tersebut mengartikan algoritma memiliki korelasi yang baik berdasarkan atribut dan data yang telah di proses oleh algoritma.

Kata kunci: Sumbangan Pembinaan Pendidikan; *Business Intelligence*; *decision tree*

Abstract. School tuition fee is typically used for funding school operational, i.e. paying honorary teachers in public and private schools, purchasing practical instruments, printing examination worksheets, and other net-operational costs. According to the discovered data in the research environment, the funding is unable to be acquired properly due to students' school tuition fees arrears for months even years until they graduate. Considering the condition, this research is conducted to identify the potential of students' school tuition arrears, based on the sum of their parents' salaries centered on the business intelligence approach, using the decision tree method. The analysis results show that, students whose parents' income is less than Rp 672.500,00 will be potentially in arrears with school tuition more than Rp 900.000,00 each month, while students whose parents' income is above Rp 672.500,00 are potentially in arrears of less than Rp 900.000,00 or not in arrears. To evaluate the effectiveness of the decision tree algorithm for data processing, it has an accuracy value of 95.97%, with a precision of 94.96% that means the algorithm has a good correlation based on attributes and the data that have been processed by the algorithm.

Keywords: School tuition fees; Business Intelligence; decision tree

1. Introduction

Each month, schools struggle with serious financial problems. According to the discovered data, the funding is unable to be acquired properly due to students' school tuition fees arrears. Normally, they must pay the tuition fees every month. Consequently, the schools find it difficult to develop since the funding of school operations, i.e. paying honorary teachers in public and private schools, and other net-operational costs, which are majorly sourced by school tuition fees, cannot be funded maximally. Government's school operational relief fund is also insufficient even unable to be allocated for covering the tuition fees and other operational expenses.

Solutions for the tuition arrears that have been done by the school are giving more time to parents to fully pay their tuition fees, and enlisting poor students for the scholarship. Nevertheless, the tuition arrears potential is still unidentified. Considering the condition, the research is conducted to identify the potential of students' school tuition arrears, based on the sum of their parents' salaries centered on business intelligence concept and the application of decision tree method with algorithm C4.5. The implementation of this algorithm classification is based on [1], which proves high accuracy algorithm C4.5 score in case of decision making for approving debt proposal in the bank. Other references related to the C4.5 algorithm are implemented for predicting loans in an economic cooperation. This research implements data mining for predicting the potential of loan approval, which falls behind schedule based on customers' salary category and its sum for each month [2].

Referring to [1] and [2], the C4.5 method would be tested at Angkasa Vocational School of Tasikmalaya environment since the same cases occur. Therefore, the aim of this research is observing the potential of students' school tuition fees arrears, based on the sum of parents' salaries as anticipation for Angkasa Vocational School in covering the deficit amounts of school operational funding, and a problem solving for future reference.

2. Theoretical Framework

School tuition fees are money paid for supporting the educational institution in the teaching and learning activities. School tuition fee is typically used for funding school operations, i.e. purchasing practical instruments, printing examination worksheets, and other operational costs. This research is conducted considering school tuition fees payments deficiency and delay influence school operational activities. Beside that, there is research that proves the benefits of the stipulated payment information system for school fees [3]. This system can do a quick identification process whether or not students have tuition fees arrears, and reports on payment of fast school fees. Generally, defrayal is a sacrificing economy source that is counted by the amount of money for achieving certain purposes. Besides, bad-debt expenses such as tuition arrears can be stated as a risk, which is causing disadvantage. As stated in Controlling Intern System of School Tuition Fees research [4], credit is demand towards costumers and another party to get money or certain services and commodities in the future. This theory is related to the conducted research since the school tuition fees arrears occur because of social asymmetry in the organization. Nonetheless, there are no analyses available regarding factors that caused the arrears.

Many types of research discuss school costs through analysis or information system development, such as building a tuition fee management software with Short Message Service-based notifications [5], and research building a tuition fee school information system to facilitate payment [3], [6]. However, none of them gives any correlation with parents' working salary. Thus, this research will discuss analysis related to the sum of parents' salaries towards tuition fees arrears potential. This research applies decision tree classification as a problem solving method.

2.1. Business Intelligence

Business is an effort to profit productively to meet everyone's needs. Business activities can be documented on several papers or in the form of electronic media. So, it is said as data. Business Intelligence is not a product nor the system, but an operational architecture and collection that are integrated into decision-making applications and databases that provide business people with easy access to business data [7]. Business Intelligence has attracted the attention of many organizations

regarding the usefulness and benefits for the organization. However, Business Intelligence still faces the challenges of obtaining maximum results from the implementation of Business Intelligence.

The main challenges of Business Intelligence are closely related to business patterns that are unique to each organization, as well as business policies and rules imposed by the company. This causes companies to not be able to buy Business Intelligence products such as finished goods in general and hopes to fulfill every solution of their business needs so Business Intelligence must be developed based on the needs and business processes of the company. All forms of transactions in business processes will produce unstructured data from customer responses either sourced from industry or institutions. The data can be analyzed using data mining to produce knowledge patterns reflecting transaction behavior in the long run.

Business Intelligence (BI) is a new technology for understanding the past and predicting the future. The technology intended here is a technology that can collect, store, access and analyze data to help decision-makers produce better decisions. As an application, BI includes activities in decision support systems, query, and reporting, OLAP (Online Analytical Processing), data mining in the enterprise risk management, statistical risk analysis model, forecasting, and data mining [8]. The following is the cycle of the Business intelligence and data cycles in Figure 1.

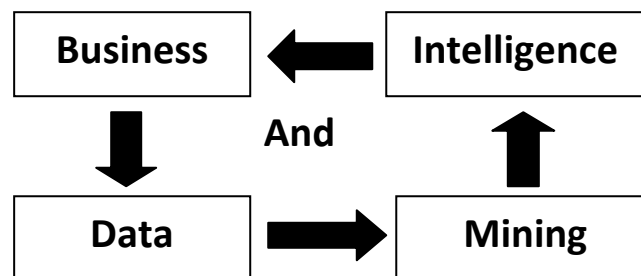


Figure 1. Business intelligence and data mining cycle [9]

In this research, data processing of business process results is carried out with a data mining approach, which aims to develop the business of providing education donations at the school. As a supporting theory of this research, a literature study has been conducted of several related researches such as prediction of bad credit through customer behavior using the C4.5 algorithm. The results of the research show accuracy = 91.06%, precision = 100.00% [10], it means that the validation of the algorithm for processing data has very good testing accuracy [11]. The data can be used as references of research concerning school tuition in arrears. The similar study analyses the collectibles of credit payment to minimize the risk of credit proposal mistake. The approach used in the study is data mining using algorithm of classification and regression tree to determine the creditor. From the analysis, it can be obtained the factors affecting the credit risks, such as tenor, amount of credit, profession, economic sector, and the spouse's profession. The accuracy of this prediction is 71.5%. Whereas, in similar credit card fraud prediction studies, the decision tree algorithm has a prediction accuracy of 100 % [12].

The relevance of the research with this research is in the data processing process with a data mining approach, which uses C.45 and references the level of accuracy in the effectiveness of the data processing process. C.45 can also be applied to create a decision support system that can provide an alternative decision for analysts in determining the level of risk of lending to prospective debtors [13]. The results of the credit feasibility study with C.45 obtained an accuracy value of 90.99% and an AUC value of 0.911 with a diagnosis level of Excellent Classification. Therefore it can be concluded that the application of the C4.5 classification algorithm is able to produce a pattern of creditworthiness with a good level of accuracy and diagnosis [14], [15]. Besides being applied in the banking environment, Business Intelligence can be implemented to support competitive universities by designing, measuring, managing, and developing intellectual capital, among others in the form of knowledge, skills and expertise in college human resources [16] related to management points in universities.

For example the application of Business Intelligence in the recapitulation of the value of lectures with a website-based information system produces an index of performance appraisal as decision support by the institution's management to improve the success parameters of the overall quality of assessment [17]. From the results of the presentation of some research related to this research. Then it can be used as a baseline to support the implementation of Business Intelligence with a data mining approach in predicting the payment of donations for the implementation of Education in vocational schools.

2.2. Decision Tree Theory

A decision tree is classified as a data mining technique. According to [18], the decision tree is used for separating a data set into classes for its variable responses. Typically, variable responses have two classes: Yes or No (1 or 0). If the response variable has more than two categories, a variant of decision tree algorithm has been developed. In both cases such as the weather prediction and the classification of heart attack based on the factors that influence it, decision tree classification can be used when a response variable or target is in the form of categories [19].

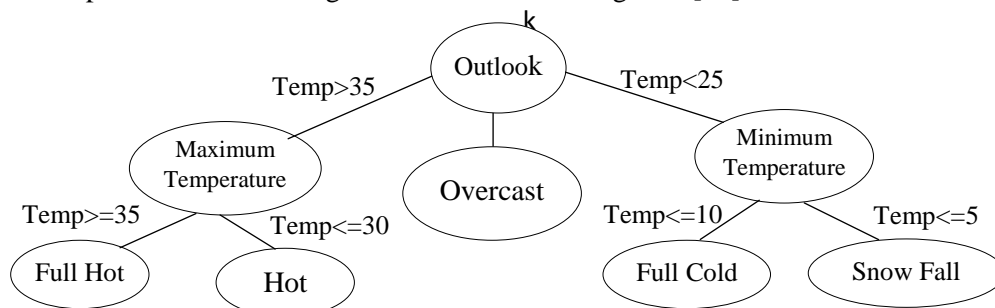


Figure 2. Decision Tree weather prediction [19]

Figure 2 shows a Decision Tree model that makes a decision by drawing flowchart depicted an upside-down tree, and the attributes will be tested after in every node. A node, or leaf which is positioned at the end of the decision tree, is a result extracted based on the target variable. Entropy or gain index is needed to build a decision tree because different criteria will create different tree through different refraction. Equation (1) shows a formula to get entropy.

$$H = - \sum P_k \text{Log}_2(P_k) \quad (1)$$

K=1,2,3,.. are representing the class of variable target while (P_k) is representing sample proportion that belongs to k class. (G) The gain index has identical characteristic with entropy measurement and can be defined by Equation (2).

$$G = \sum (1 - P_k^2) \quad (2)$$

“G value ranges from 0 to 0, 5 (max). This value has identical characteristic with H and one of the formulations can be applied for making data partition”. A decision tree can be applied as predicting implements algorithm C4.5 of decision tree based on ID3. C4.5 is able to handle attribute continuously, processing the set of the sample with missing values, and producing new rules and features. Thus, the C4.5 algorithm can be established to build a decision tree. [20]. “Algorithm C4.5 uses the information to acquire ratio as selection attribute criteria and classification sample which handles lack of information in choosing the attribute. The counting method can be explained as follow:

1. Set S as training set sample, since ISI is a representation from training sample total number, assuming it as property from S with divided value m non-repeat signed as $V = \{v_1, v_2, \dots, v_m\}$... Training sample value at A that is based on S is divided into m subset as $\{S_1, S_2, \dots, S_m\}$.
2. Values at ‘S_i’ is ‘v_i’ in all ‘A’ training sample. ‘U’ is a collection class that controls freq (u_i S) and it shows a total number of samples belong to ‘u_i’ class of ‘S’. Entropy is information that acts as a measurement of uncertain incident information sources that possibly happen. Entropy

information at set 'S' sample which is divided based on decision 'U' attributes is using Equation (3).

$$Info(U) = -\sum_i P(u_i) \log_2 P(u_i) \quad (3)$$

3. 'P(U_i) represents 'u_i' percentage at total sample value. Conditional enthalpy shows uncertainty when it starts to put real variables disorderly after a particular condition. The form of conditional enthalpy with set 'S' sample and are divided according to partition 'V' domain attribute is show in Equation (4).

$$Info(U|V) = \sum_j P(v_j) \sum_i P(u_i|v_j) \log_2 \left(\frac{1}{P(u_i|v_j)} \right) \quad (4)$$

4. P(u_i | v_j) shows conditional probability form at the 'u_i' category when 'A' is 'v_j'. Gain information, well known as mutual information, represents output set uncertainty for input set. Equation (5) is about Gain information as set 'S' training sample bases on 'A' partition attribute.

$$Gain(A) = Info(U) - (A/V) \quad (5)$$

5. V is all output set area of 'A'. It represents information of output status set 'V' about 'A'. What makes different with enthalpy in eliminating uncertainty is the mutual information at obtained data information. Training value sample set 'S' with different' value in attribute 'A' is {S₁, S₂,..., S_m}. Equation (6) is a formula for counting partition information [21].

$$SplitInfo(S, A) = -\sum_{i=1}^m \frac{|S_i|}{|S|} \log_2 \left(\frac{|S_i|}{|S|} \right) \quad (6)$$

6. The gain ratio is a ratio in gain rate information intended for value at gain information. Gain ratio counting formula for information level at training sample set 'S' can be classified by 'A' attribute as shown in Equation (7) [22], [23].

$$GainRation(A) = \frac{Gain(S, A)}{SplitInfo(S, A)} \quad (7)$$

Attribute tests of Algorithm C4.5 that has been chosen for each node at decision tree, are using gain ratio information. Thus, choose the node test attribute with the highest gain ratio information to be used. Examining theory, which supports this research, has been based on the previous researches. Accordingly, Algorithm decision tree data mining supporting theories are able to predict potentials may occur. The potentials are based on unstructured data that later would be processed with a data mining method. This approach would give a system that facilitates in obtaining required data. Various data contents occurred to have the potential to be new raw data in the form of unstructured one to be applied in other days.

In line with the explained examining theory and its approach, this concept is called 'Business Intelligence. Therefore, the classification method of the 'decision tree', will be implemented in this study, to predict the potential of student tuition arrears as anticipatory measures in the amount of school funding deficits, and problem solving for future reference. Angkasa Vocational School of Tasikmalaya would become an object of the research or example environment since the same cases may occur in other schools.

3. Methodology

The method proposed for the aspect of predicting school fee arrears prediction in this study uses a data mining approach with the decision tree method, the research model is found in Figure 3. Whereas for the process of applying prediction results, delinquency optimization and data retrieval using a business intelligence approach presented in Figure 4.



Figure 3. Proposed method model for analyzing the parents' sum of salaries toward tuition arrears potential.

Firstly, prepare a data set which is taken by a school administrator and treasurer. Secondly, split the data into training data and testing data. Thirdly, extract data mining to get the pattern. Decision tree algorithm and C. 4.5 are applied in this step. Finally, make an evaluation of algorithm effectiveness toward extraction data processing. The data set which users are students' tuition fee data from 2016 to 2017, and profiles of students' parents. For Business Intelligence approach steps, which is outlined into a model as presented in Figure 4

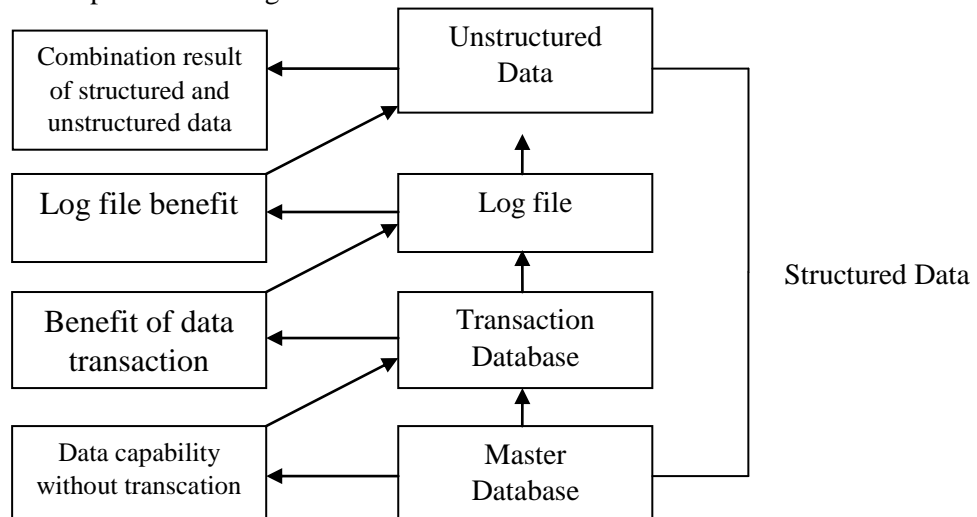


Figure 4. Business Intelligence approach method model

Figure 4 is an applied method in analyzing the sum of parents' salaries towards the potential of students' school tuition fees arrears potential established in Angkasa Vocational School of Tasikmalaya. A Business Intelligence approach method, can be explained in the following steps:

1. Master Database
Student profile, tuition fee data.
2. Data capability without transaction
Amounts of the student, information on tuition fee arrears, the data value of paid tuition fee, a data value of arrears tuition fee, a sum of salary total information data.
3. Transaction Database
Tuition fee payment data, amount of students who paid, amount of students who have not paid yet.
4. The benefit of data transaction
Unpaid tuition fee information, paid tuition fee information, operational cost total nominal to be covered information.
5. Logfile
Records, which involve: tuition fee payment time, expertise programs which have tuition fee arrears, expertise programs which pay the tuition fee on time.
6. Logfile benefit
To acknowledge the most favorable time for paying tuition fee, getting information related to expertise programs that pay tuition late, students with tuition fee arrears, students and expertise program that pay tuition fee on time.
7. Unstructured Data
Parents' sum of salary data.
8. Combination result of structured and unstructured data
To get information about students who have potential in tuition fee arrears based on parents' sum of salaries therefore, anticipation in covering school operational fund can be done. Through the collected data, aid funds can be allocated to students who have potential arrears. The eight steps would be processed directly by using Rapid Miner tools [21] for data processing efficiency and faster progress with decision tree algorithm data mining approach [24].

4. Result and Discussion

Processed data in this research are taken by two master databases school tuition data are from Angkasa Vocational school, and parents' sum of salary is gathered from education basic data of Angkasa Vocational School of Tasikmalaya.

Table 1. Students' school tuition fee master data

Name	PPDB	School Tuition fee arrears and School activities						Total Value	
Ade Sani Fitriani	-	1	110.000	110.000	1	70.000	70.000	180.000	180.000
Budi Pangestu	650.000	5	110.000	550.000	5	70.000	350.000	900.000	1.550.000
Kukuh Pambudi	-		110.000	-	-	70.000	-	-	-
Muhammad Andi N	650.000	1	110.000	110.000	1	70.000	70.000	180.000	830.000
Nurjaman	570.000	4	110.000	440.000	4	70.000	280.000	720.000	1.290.000
Syaida Ayumi	300.000	1	110.000	110.000	1	70.000	70.000	180.000	480.000
Wildan Badru Zaman	830.000	2	110.000	220.000	2	70.000	140.000	360.000	1.190.000
Widya Jannatun	950.000	4	110.000	440.000	8	70.000	560.000	1.000.000	1.950.000
Yanti Nuraeni	830.000	4	110.000	440.000	4	70.000	280.000	720.000	1.550.000
Yuni Wahyuni	550.000	5	110.000	550.000	5	70.000	350.000	900.000	1.450.000
Dede Restu	1.030.000	6	110.000	660.000	9	70.000	630.000	1.290.000	2.320.000
Sinta	-		110.000	-	-	70.000	-	-	-

Table 1 is an example of students' school tuition fees, which is occupied from the treasurer's database. The real data gathered are 276 samples.

Table 2. Parents' sum of salaries Data master

Mother's name	Father's name	Occupation	Salary
Enur	Udin	Labor	950.000
Nani	Ano	Labor	999.000
Nyi Kurniasih	Ian	Labor	850.000
Yuyun	Mamat	Civil servant/police/soldier	4.000.000
Enkoy Rukoyah	Ipin	Entrepreneur	750.000
Masriatin	Asep Saepulloh	Labor	880.000
Atin	Ujah Maman	Labor	760.000
Ida	Hani Royani Permana	Civil servant/police/soldier	5.000.000
Ika	Hayidin	Entrepreneur	1.850.000
Atik Sumiati	Ahmad Yani	Entrepreneur	890.000
Enong Yuyun	Omang	Labor	650.000

Table 2 is an example of parents' sum of salaries, which is taken from education basic data of Angkasa Vocational School of Tasikmalaya. The real data gathered are 276 samples. Both data Students' school tuition fee and Parents' sum of salaries are then processed with decision tree algorithm for obtaining model or pattern that can be applied as data prediction in the future.

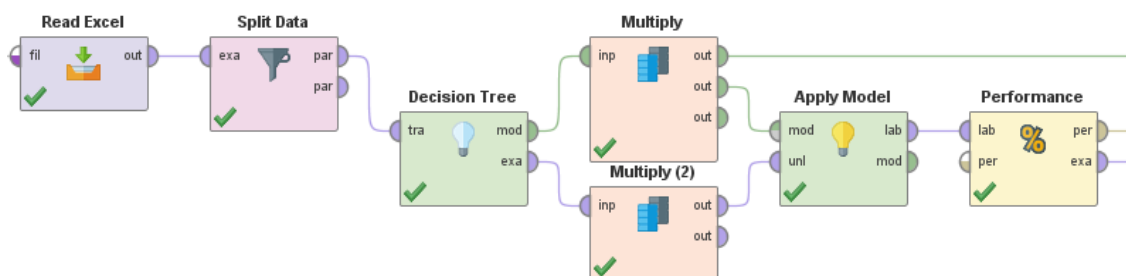


Figure 5. Decision tree method produces analyses pattern of parents' sum of salary towards school tuition arrears potential.

The processing is conducted structurally based on data mining method as proposed in Figure 5. Besides, Rapid miner is applied as a tool to produce the pattern. The prediction of parents' income on the potential arrears of school tuition payments, in each step, refers to the method proposed in Figure 3. Before the data processing is carried out by the algorithm, data is shared through the Split data operator. The operator is used to divide the dataset into two parts automatically, namely training data and data testing with a comparison of 10% testing data and 90% training data [25]. Testing data is used for comparative data between prediction results in the past with training data in the future if the results of predictions generated by the decision tree algorithm from data testing are the same as training data. Then predictions in the future are likely to be correct. The results of processing data by the algorithm can be seen in Figure 6.

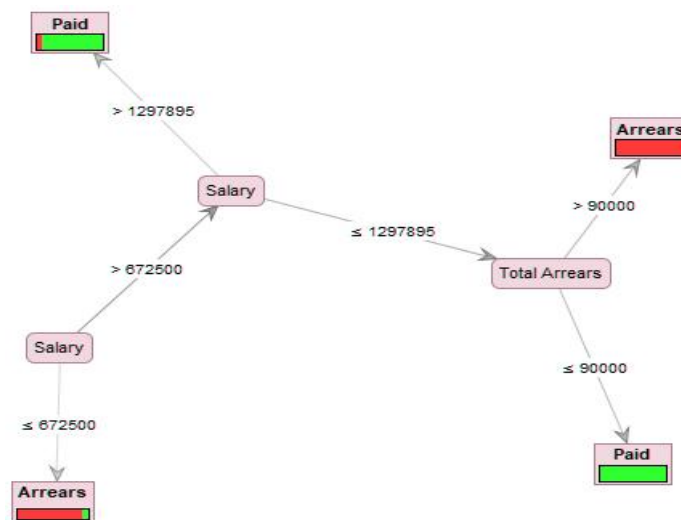


Figure 6. Pattern produced by the decision tree

The pattern or model in Figure 6 gives a reference as information in the form parents' sum salary data. The data obtained predicts tuition fee arrears potential. Explanation based on analyses is presented in Table 3.

Table 3. Prediction of school tuition fee arrears based on Parents' sum of salaries.

Salary	Tuition fee arrears	Status
< 672.500	-	Potentially arrears
> 672.500	-	Potentially paid
> 1.297.895	-	Potentially paid
-	< 900.000	Potentially paid
-	> 900.000	Potentially arrears

Table 3 explains the prediction patterns of parents' income towards potential tuition arrears. The explanation is as follows: Parents of students who have an income of less than Rp 672.500,00 will potentially have tuition arrears. While parents who have incomes range from Rp 672.500,00 to more than Rp 1.297.895,00. Then it will potentially tuition always fully-paid.

Table 4. Students who have arrears potential.

Salary	Tuition fee arrears	status	Total Number
< 672.500	-	arrears	36 students
> 672.500	-	paid	3 students
> 1.297.895	-	paid	60 students
-	< 900.000	paid	72 students
-	> 900.000	arrears	70 students

Table 4 explains the prediction of tuition arrears based on parents' sum of salary relat to a number of students that turned into 276 data samples. The data sample related to parents' sum of salary are then classified to find out whether they are able to make a full paid tuition fee or arrears. The results obtained from the research reveal 106 students of 276 have parents with a sum of salaries less than Rp 672.500,00 and arrears more than Rp 900.000,00. These conditions indicate students' tuition fee arrears potential in each month. Despite the fact, 135 students positively have no potential for the arrears.

Table 5. Confusion Matrix

	true arrears	true paid	class precision
Arrears pred.	106	3	97.25%
Paid pred.	7	132	94.96%
class recall	93.81%	97.78%	

Table 5 Confusion Matrix is a table based on the calculation of the level of correlation and accuracy of predictions between algorithms and data. In the delinquency prediction column, predictions of arrears potential will occur, the true variable in arrears has a value of 106, while true paid off has a smaller value of 3 with a corresponding level of 97.25%. This means that predictions of the potential arrears of payment will potentially occur. Whereas, in the prediction column paid off, the algorithm predicts the potential for paying off school fees will occur, the potential variable delinquent (true delinquency) has a value of 7 while the full potential (true paid), as much as 132 with a precision of 94.96%. This means that predictions of the potential payment of school fees will occur according to predictions.

Table 6. Algorithm evaluation

Evaluation	Value
Accuracy	95.97%
Precision	94.96%

Table 6 reveals the decision tree algorithm accuracy value towards data processing. It results in 95.97% accuracy, and 94.96% precision, which means the algorithm has a good correlation with the processed attribute.

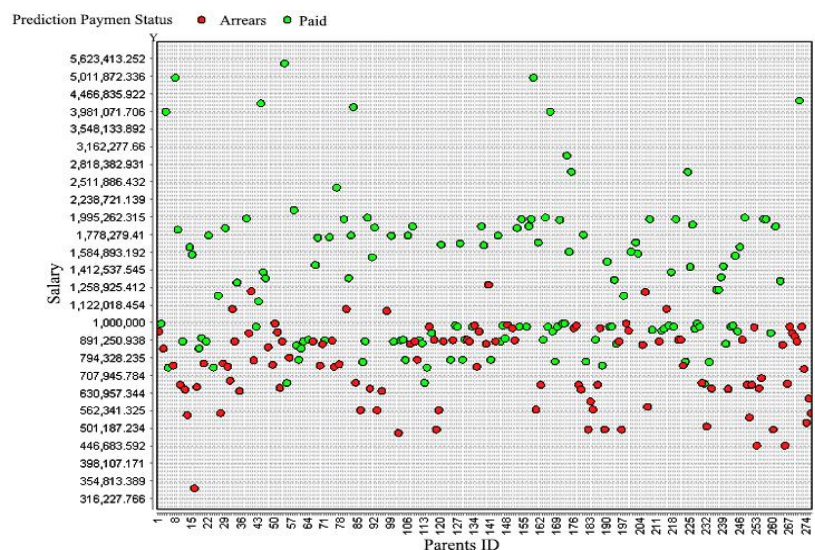


Figure 7. Scatter chart of school tuition fee arrears

The scatter chart in Figure 7 explains parents' sum salaries, which predicts to have potential of tuition fee arrears by representing the number of students based on their parents' salaries. Parent's sum of

salaries start from Rp 354.813,00 to Rp 1.100.000,00 are symbolized with red color clusters that gather in the area of potentially tuition arrears, while the sum of salaries starts from Rp 672.500,00 to Rp 5.523.413,00 are symbolized with green cluster color which represents full paid prediction.

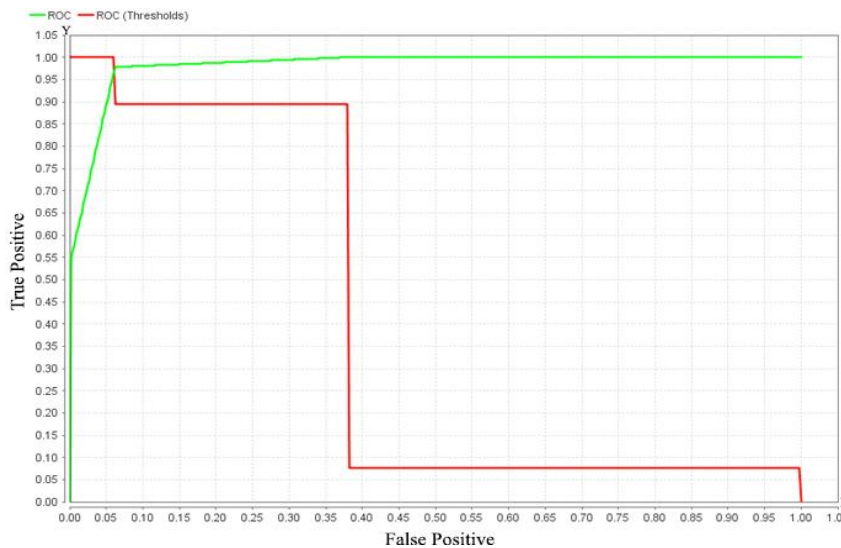


Figure 8. The area under curve chart

Figure 8 is an Area under curve diagram (AUC). The overall accuracy of the test can be explained in the area under the curve. The y-axis is said to be true positive while the x-axis is said to be a false positive. The red line shows the classification value 1.0 against the x-axis leading to the lower right and the y-axis leads to the left part with a value of 1.0 which means the prediction on the false positive or in predicting the truth is said to be true. While the green line shows the value 0 on the x-axis in the lower-left position and 0.981 on the y positive true axis that leads to the right side. Deciphering the classification of predictions close to 0.1 means perfect. So that the process of classification and prediction made by the decision tree towards attributes is very good or said to be true.

5. Conclusion

The research, which discusses Parents' sum of Salaries Analyses towards School Tuition Fee Arrears Potential is able to be established in Angkasa Vocational School of Tasikmalaya since the Decision Tree is applied as its method. By applying data mining decision tree, school tuition fee arrears can be identified through the form of pattern. Based on the patterns obtained in the case study at Angkasa Vocational School, parents with a salary less than Rp 672.500,00 tend to have arrears potential. Meanwhile, parents with an amount of salary more than Rp 672.500,00 until Rp 1.297.895,00 and over potentially pay the school tuition fee. The data samples are obtained from the numbers of students. The data samples are classified based on parents' sum of salary to find out whether the parents are able to make full payment or arrears. As a result, 106 students of 276 have parents with the sum of salaries less than Rp 672.500,00 and arrears more than Rp 900.000,00. These conditions indicate students' tuition fee arrears potential in each month. Despite the fact, there are 135 students whose parents have a sum of salaries more than Rp 672.500,00. It means they positively have no potential for the arrears. Algorithms decision tree and Confusion Matrix have algorithm value accuracy towards data processing. It results in 95.97% accuracy and 94.96% precision, which mean the algorithms have a good correlation with the algorithm processed attribute.

In the research of predicting delinquent payments for school fees, the scope of variables used only covers the income of parents. So for the development of research in the future can be added a variable influence the number of children that parents have on the ontime-tuition payment of school fees. As well as implementing the framework proposed in the website or mobile-based information system.

6. References

- [1] R. K. Amin, Indwiarti, and Y. Sibaroni, "Implementation of decision tree using C4.5 algorithm in decision making of loan application by debtor (Case study: Bank pasar of Yogyakarta Special Region)," *2015 3rd Int. Conf. Inf. Commun. Technol. ICoICT 2015*, vol. 0, pp. 75–80, 2015.
- [2] K. Arun, G. Ishan, and K. Sanmeet, "Loan Approval Prediction based on Machine Learning Approach," *IOSR J. Comput. Eng.*, vol. 18, no. 3, pp. 18–21, 2016.
- [3] Hendria Tony Fristanto, "Rancang Bangun Sistem Informasi Pembayaran Sumbangan Pembinaan Pendidikan (SPP) Dan Insidental Pada Sekolah Menengah Kejuruan Muhammadiyah Tinatar Punung Hendria Tony Fristanto," *Indones. J. Netw. Secur.*, vol. 3, no. 4, pp. 24–28, 2014.
- [4] H. P. Dewi, "Sistem Pengendalian Interen Atas Piutang Sumbangan Pembinaan Pendidikan (Spp) Sebagai Upaya Peningkatan Kinerja Pada Yayasan Islam Al - Hamidiyah Depok Tangerang," *J. Din. Ekon. Bisnis*, vol. 12, pp. 76–88, 2015.
- [5] R. K. A N Hak, Cepy Slamet, "Implementasi Sistem Informasi Keuangan Siswa Berbasis Sms Gateway Di Smk Muhammadiyah 2 Kadungora," *J. Algoritm.*, vol. 11, no. 1, pp. 1–8, 2014.
- [6] I. S. Ma'rifati, "Sistem Informasi Akuntansi Pendapatan Sumbangan Pembinaan Pendidikan (SPP) Pada SMU XYZ," *J. Evolusi*, vol. 3, no. 1, pp. 1–8, 2015.
- [7] S. Darudiato, S. W. Santoso, and S. Wiguna, "Business Intelligence : Konsep Dan Metode," *J. CommIT*, vol. 4, no. 9, pp. 63–67, 2010.
- [8] D. D. Wu, S. H. Chen, and D. L. Olson, "Business intelligence in risk management: Some recent progresses," *J. Inf. Sci.*, vol. 256, pp. 1–7, 2014.
- [9] A. K. Maheshwari, *Intelligence and Disease.*, 1st ed., vol. 219, no. 5655. New York: Business Expert Press, LLC, 2015.
- [10] A. Sucipto, "Pada Koperasi Simpan Pinjam Dengan Menggunakan METODE Algoritma Klasifikasi C4.5," *J. DISPROTEK*, vol. 6, no. 1, pp. 75–87, 2015.
- [11] Ivandari, "Model Pohon Keputusan Untuk Klasifikasi Persetujuan Kredit," *J. Litbang Kota Pekalongan*, vol. 9, pp. 124–136, 2015.
- [12] R. S. Oetama, "Enhancing Decision Tree Performance in Credit Risk Classification and Prediction," *J. Ultim. J. Tek. Inform.*, vol. 1, no. 7, pp. 50–55, 2015.
- [13] M. H. Rifqo and T. Arzi, "Implementasi Algoritma C4.5 untuk Menentukan Calon Debitur dengan Mengukur Tingkat Risiko Kredit pada Bank Bri Cabang Curup," *J. Pseudocode*, vol. III, no. 2, pp. 83–90, 2016.
- [14] Abdussomad, "Penerapan Algoritma Klasifikasi C4.5 Untuk Menghasilkan Pola Kelayakan Kredit," *J. Inov. Inform.*, vol. II, no. 1, pp. 19–27, 2017.
- [15] Z. Nur and I. Mukhlash, "Implementasi Business Intelligence Pada Manajemen Report Bank XYZ," *J. Sains Dan Senni Pomits*, vol. 3, no. Bisnis Intelijen, pp. 16–21, 2014.
- [16] M. Mustafid, "Sistem business intelligence untuk mendukung perguruan tinggi yang kompetitif," *J. Sist. Inf. Bisnis*, vol. 3, no. 1, pp. 18–24, 2013.
- [17] U. Rahardja, Q. Aini, A. Khoirunisa, and B. Intelligence, "Implementasi Business Intelligence Menggunakan Highchart pada Sistem Penilaian Absensi berbasis YII Framework," *J. Comput. Sci. Res. its Dev.*, vol. 9, no. 2, pp. 115–124, 2017.
- [18] A. K. Hamoud, A. S. Hashim, and W. A. Awadh, "Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 5, no. 2, p. 26, 2018.
- [19] S. S. Bhatkande and R. G. Hubballi, "IJARCCE Weather Prediction Based on Decision Tree Algorithm Using Data Mining Techniques," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 5, pp. 483–487, 2016.
- [20] H. Jiang and R. Wang, "Fault mode prediction based on decision tree," *IEEE Adv. Inf. Manag. Commun. Electron. Autom. Control Conf.*, pp. 1729–1733, 2016.
- [21] T. Wiyana, Hari;Kurniawan, Sigit; S. Ali, "Memprediksi Transaksi Nasabah Pada Koperasi Simpan Pinjam PT. Asuransi Staco Mandiri Dengan Algoritma C4.5," in *sentika*, 2018, pp.

- 97–106.
- [22] G. Lukhayu Pritalia, “Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce,” *Indones. J. Inf. Syst.*, vol. 1, no. 1, pp. 47–56, 2018.
 - [23] W. Said, M. M. Hassan, and A. M. Fawzy, “Data mining techniques for database prediction: Starting point,” *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 23, pp. 7723–7738, 2018.
 - [24] R. K Rai, “Intricacies of Unstructured Data,” *J. ICST Trans. Scalable Inf. Syst.*, vol. 4, no. 14, pp. 1–5, 2017.
 - [25] E. S. Rahayu, R. S. Wahono, and C. Supriyanto, “Penerapan Metode Average Gain, Threshold Pruning dan Cost,” *J. Intell. Syst.*, vol. 1, no. 2, pp. 91–97, 2015.