lassification of Receipt of BPNT Social Assistance Using the C4.5 Algorithm (Case of the Grujugan Kidul Village Office)

by - -

Classification of Receipt of BPNT Social Assistance Using the C4.5 Algorithm (Case of the Grujugan Kidul Village Office)

Abdul Wahid
Institut Teknologi dan Sains Mandala
Jl. Sumatra. No.118-120, Jember
+62 852-5707-6330
abdulwahid@itsm.ac.id

Mas'ud Hermansyah
Institut Teknologi dan Sains Mandala
Jl. Sumatra. No.118-120, Jember
+62 813-3646-6463
masudhermansyah@itsm.ac.id

M. Faiz Firdausi
Institut Teknologi dan Sains Mandala
Jl. Sumatra. No.118-120, Jember
+62 852-5832-6262
faizfirdausi@itsm.ac.id

ABSTRACT

BPNT social assistance is a government program that is unsustainable to help the lives of people who are selected selectively and periodically, social assistance is intended for some people in villages who are considered financially unable, so they need help from the government. BPNT social assistance, commonly known as social assistance, is assistance charged by the regional Revenue and Expenditure Budget. The Municipal Government of Bondowoso provides various kinds of assistance programs, one of which is the BPNT Social Assistance, the distribution of which is carried out by the Grujugan Kidul Village Office. In the selection process to determine the beneficiaries of BPNT Social Assistance at the Grujugan Kidul Village Office, they still have not fully used information technology to support employee performance. So there are obstacles and it takes a long time. Therefore, we need a system that can help employees more easily determine beneficiaries. The application of Data Mining is a series of processes to find added value semi-manually in the form of unknown knowledge from a data set. In this study, the parameters are income, employment, and DTKS status. By implementing the C4.5 Data Mining Algorithm, it is hoped that it will make it easier and faster for employees to determine recipients of BPNT Social Assistance at the Grujugan Kidul Village Office. Paper, we describe the formatting guidelines for Proceedings. Maximum 200 words.

Keywords : Classification Method, C4.5 Algorithm, Eligibility for Receiving.

1. INTRODUCTION

BPNT social assistance is a government program that is unsustainable to help the lives of people who are selected selectively and periodically, social assistance is intended for some people in villages who are considered financially unable, so they need help from the government. BPNT social assistance, commonly known as social assistance, is assistance charged by the Regional Revenue and Expenditure Budget. According to the general provisions of Article 15 of the Government Regulation concerning Granting and Social Assistance in the Regional Revenue and Expenditure Budget of the Minister of Home Affairs Number 32 of 2011, the definition of Social Assistance refers to provisions for individuals, families and community groups which are unsustainable and selective in nature to prevent potential risks social. (Alkhairi & Situmorang, 2022).

The selection process for recipients of Pemko social assistance at the Martoba Kelurahan Office has obstacles in identifying recipients of Pemko social assistance, namely the time period from administrative selection and verification to the recipient's residence. From the beginning of the process, residents' files will be examined and those who are entitled to receive assistance will be selected according to the criteria specified in the social section. Thus, the process of deliberation in identifying potential recipients of Pemko social assistance is not systematic and sometimes inappropriate. Many branches of computer science can solve this problem. Based on the explanation above, the researcher uses Data Mining with the C4.5 Algorithm method to solve the problem.

In order to reduce poverty, one of which is by providing social assistance to those who are eligible so that it is right on target, this study discusses the classification of social assistance recipients using the C 4.5 algorithm in one of the Grujugan District areas, Bondowoso Regency, the results of which can be used for decision making by related parties whether residents is eligible to receive assistance or not. Algorithm C4.5 is an algorithm used to classify data using a decision tree technique (AD I Suradi, 2018), (Leonardi et al., 2021). The C4.5 algorithm is an extension of the ID3 algorithm and uses the decision tree principle (Hendriyanto, 2018). The main advantage of the C4.5 Algorithm is that it can produce models in the form of trees or rules that are easy to interpret, has an acceptable level of accuracy and can handle discrete and numeric attributes (Sholihah et al., 2018). In the C4.5 Algorithm, the model produced from the learning process from training data is in the form of a Decision Tree (Girsang et al., 2022). Thus the distribution of assistance by the authorities will not cause social jealousy for those who should receive assistance but instead do not receive assistance as at present.

2. RESEARCH METHODE

2.1 Research Methodelogy

This research was conducted to implement Data Mining in determining the pattern of compiling files contained in the Martoba Pematangsiantar Village Office (Choiril Ichsan Damanik et al., 2021). The method used in research conducted by researchers is the C4.5 Algorithm method. In this research methodology, a process of scientific method will be described to obtain data that will be used in research to solve a problem (Nuraini et al., 2022).

2.2 Rancangan Penelitian

Tah The stages of the methodology in this study are shown in Figure 1 below:

Gambar 1. research stages

From the picture above, it can be described the stages of the research methodology with the following explanation:

a) Problem Identification

The author made direct observations in one of the areas in Grujugan Kidul Village, Grujugan District, Bondwoso Regency, and from the results of these observations the author mapped the process of distributing social assistance receipts. The distribution of social assistance has not been on target, thus hampering the administration process.

b) Needs Analysis

With the delay in the data collection process for determining who is eligible or not for social assistance, a system is needed that can make decisions quickly and accurately so that the resulting decisions are not subjective which could become a problem in the future.

The data preparation process that is carried out is by converting hard copy data into excel format (Sugianto & Maulana, 2019)

c) Classification C4.5

Through the analysis of the data collection, it can be seen whether residents are eligible or not to receive BPNT. From the existing problems, a classification method is used to predict the eligibility of BPNT beneficiaries, namely by using two models, the C4.5 classification algorithm model (Ermawati, 2019)

By using the existing training data, a decision tree is formed using the C4.5 algorithm. And a decision tree and rules are generated that represent the decision tree so that the results are for residents who deserve rice for the poor (Junaidi et al., 2023)

d) Implementation with Rapidminer.

After all the data is complete, then it is processed through the RapidMiner software with the stages of opening the data file in Excel, determining the role, breaking down the training and testing data, using the decision tree model, and measuring the performance of the data.

2.3 Data Analysis

The data analysis process is carried out after the data collection process is complete. In this study the data that was obtained by researchers was used for research data, the data used was primary data, primary data was data collected directly by researchers, usually primary data was collected through observation and interviews (Wahid & Agustina, 2018). Data analysis tools used by researchers in conducting this research are as follows:

a) Microsoft Excel

Microsoft Excel is software that functions to process data, which in this application is in the form of spreadsheets created and distributed by Microsoft Corporation to be used on Windows and Mac OS operating systems. The use of Microsoft Excel in this study is to process temporary data in order to get fast results.

b) RapidMiner

RapidMiner is software used in data mining processing, the use of RapidMiner is used to obtain valid test results with the results of previous calculations carried out in Micosoft Excel.

3. RESULT AND DISCUSSION

3.1 Result

In implementing the final results of applying the C4.5 Algorithm it is divided into two stages, namely the semi-manual calculation process using Microsoft Excel and the adjustment of the calculation results through data testing using the RapidMiner 5.3 software.

3.1 System Analysis

System analysis aims to find new patterns of determining BPNT social assistance beneficiaries using Data Mining Algorithm C4.5 classification. The process of selecting beneficiaries of BPNT social assistance has time constraints so that the settlement process does not take place systematically and sometimes is not on target. In this case, the data source was obtained from the Grujugan Kidul Village Office, Grujugan District, Bondowoso Regency. Based on the recapitulation results, 395 data samples were used for the classification of the C4.5 Algorithm in determining the recipients of Pemko social assistance. There are 3 criteria used, namely income, employment, and DTKS status. Table 1 shows the results of the collection of fields and data on recipients of BPNT social assistance used as training data in the data mining process.

Table 1. BPNT Social Assistance Receipt Data

No	NAMA	NIK	PENDAPATAN	PEKERJAAN	STATUS DTKS
1	AHMAD BYRON PRAMATA	3511062806090001	1.000.000 - 1.500.000	Petani	Tidak
2	SULASTRI	3511064505730001	1.500.000 - 2.000.000	PNS	Ya
3	ULFA SAFITRI	3511065011040001	500.000 - 1.000..000	Buruh Tani	Ya
4	YANTO	3511061711830001	500.000 - 1.000..000	Ibu Rumah Tangga	Ya
5	AKIB	3511061505650011	1.000.000 - 1.500.000	Buruh Tani	Ya
6	MUZAIYANAH	3511065609710001	500.000 - 1.000..000	Ibu Rumah Tangga	Tidak
7	NUR KHOLISAH	3511066408990001	1.500.000 - 2.000.000	Wiraswasta	Tidak
8	SARAH EVELYNA	3511066705060002	> 2.000.000	Wiraswasta	Tidak
9	SUHARTONO	3511061201730001	> 2.000.000	PNS	Tidak
10	TEGAR	3511061102040002	1.000.000 - 1.500.000	Wiraswasta	Tidak
11	UMYANA	3511064207780001	0 - 500.000	Buruh Tani	Tidak
12	AHMAD ARDIS ABDUL MUGHNI	3511060909190002	1.000.000 - 1.500.000	Petani	Ya
13	MOHAMMAD MAHMUD	3511062912940003	1.500.000 - 2.000.000	Wiraswasta	Ya
14	YULIATUL HASANAH	3511067012950001	1.000.000 - 1.500.000	Buruh Tani	Tidak
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
391	ROHAMA	3511065510690001	500.000 - 1.000..000	Buruh Tani	Tidak
392	SATIP	3511060107600006	1.000.000 - 1.500.000	Petani	Ya
393	EKA NUR INTAN	3511066803030002	1.500.000 - 2.000.000	Wiraswasta	Ya
394	HANINA	3511065207600006	1.000.000 - 1.500.000	Petani	Tidak
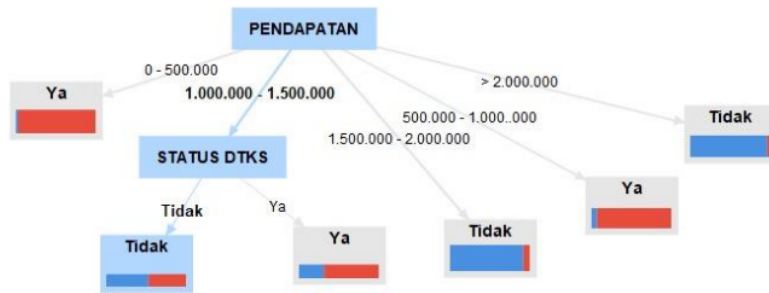
| 395 | INDRA DWI KURNIAWAN | 3511062803100001 | > 2.000.000 | PNS | Ya |

3.2 Algorithm Experiment Process C4.5

The following are the steps for implementing the C4.5 Algorithm on RapidMiner 5.3.
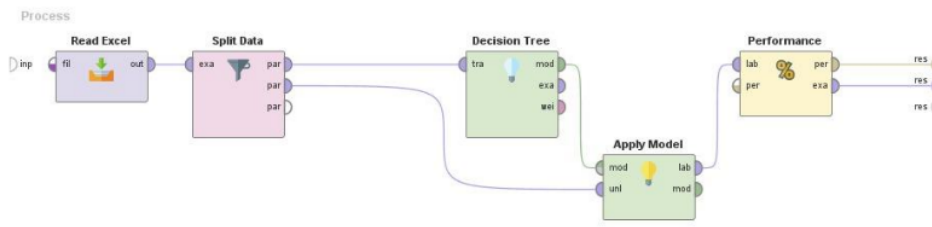
Gambar 2 Display Connect Port Decision Tree, Apply Model And Performance

Next, connect the ports of the Decision Tree operator, the Apply Model operator and the Performance operator as shown above, then click the Run icon on the toolbar to display the results.

Gambar 3 Display of Decision Tree Results

From Figure 3, information is obtained that in order to determine the eligibility of beneficiaries the order that has the top priority is income, employment and status of DTKS..

3.2 Discussion

The following are the rules generated from the decision tree:

Table 2. Generated Rules

No	Rules	Keputusan
1	Jika Penghasilan (0 – 500.000) dan Status DTKS = Tidak	Layak
2	Jika Penghasilan (0 – 500.000) dan Status DTKS = Ya	Layak
3	Jika Penghasilan (500.000 - 1.000.000) dan Status DTKS = Ya	Layak
4	Jika Penghasilan (500.000 - 1.000.000) dan Status DTKS = Tidak	Layak
5	Jika Penghasilan (1.000.000 - 1.500.000) dan Status DTKS = Ya	Layak
6	Jika Penghasilan (1.000.000 - 1.500.000) dan Status DTKS = Tidak	Tidak Layak
7	Jika Penghasilan (1.500.000 – 2.000.000) dan Status DTKS = Ya	Tidak Layak
8	Jika Penghasilan (1.500.000 – 2.000.000) dan Status DTKS = Tidak	Tidak Layak
9	Jika Penghasilan (> 2.000.000) dan Status DTKS = Tidak	Tidak Layak

The data used is 187 out of 396 data, the amount of data is reduced because it has gone through the cleansing stage which is part of the data mining process. The cleaning process is carried out when there are empty variable values for each attribute. So that the data comparison is obtained as shown below:

Decision Tree Method C4.5:

Comparison of data
6:4

accuracy: 85.33%

	true Tidak	true Ya	class precision
pred. Tidak	32	6	84.21%
pred. Ya	5	32	86.49%
class recall	86.49%	84.21%	

Comparison of data 7:3

accuracy: 89.29%

	true Tidak	true Ya	class precision
pred. Tidak	26	4	86.67%
pred. Ya	2	24	92.31%
class recall	92.86%	85.71%	

Comparison of data 8:2

accuracy: 91.89%

	true Tidak	true Ya	class precision
pred. Tidak	17	2	89.47%
pred. Ya	1	17	94.44%
class recall	94.44%	89.47%	

Comparison of data 9:1

accuracy: 94.44%

	true Tidak	true Ya	class precision
pred. Tidak	9	1	90.00%
pred. Ya	0	8	100.00%
class recall	100.00%	88.89%	

Based on data processing using RapidMiner software, the system accuracy value is 94.44%, meaning. Where the model that has been formed is tested for its level of accuracy by entering or testing derived from training data using split validation on the RapidMiner 5.3 application to test the level of accuracy..

4. Conclusion

Based on all the results of the research stages that have been carried out on the Application of Data Mining in Determining BPNT Social Assistance Recipients Using the C4.5 Algorithm (Case Study: Grujugan Kidul Village Office, Grujgan District, Bondowoso Regency) it can be concluded as follows..

a) Data Mining with the C4.5 Algorithm can be applied to Determine the Eligibility of BPNT Social Assistance Recipients in Grujugan Kidul Village, Grujgan District, Bondowoso Regency. The results obtained from testing the C4.5 Algorithm method into RapidMiner have imperfect validation values because the data received by researchers is less than perfect

b) The problem of determining the eligibility of BPPNT social assistance recipients can be solved using the Data Mining technique, namely the C4.5 Algorithm. The accuracy rate produced by this method is 94.44%.

5. REFERENCES

A. D. I. Suradi. (2018). *"Penerapan data mining untuk menentukan rekomendasi beasiswa dengan metode algoritma c4.5.*

Alkhairi, P., & Situmorang, Z. (2022). Penerapan Data Mining Untuk Menganalisis Kepuasan Pegawai Terhadap Pelayanan Bidang SDM dengan Algoritma C4.5. *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, *7*(1), 40. https://doi.org/10.30645/jurasik.v7i1.414

Choiril Ichsan Damanik, Dedy Haratama, & Heru Satria Tambunan. (2021). Penerapan JST Backpropagation Untuk Memprediksi Data Penerimaan Mahasiswa Baru Pada Universitas Simalungun. *Bulletin of Computer Science Research*, *2*(1), 25–33. https://doi.org/10.47065/bulletincsr.v2i1.146

Ermawati, E. (2019). ALGORITMA KLASIFIKASI C4.5 BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK PREDIKSI PENERIMA BANTUAN PANGAN NON TUNAI. *SISTEMASI*, *8*(3), 513. https://doi.org/10.32520/stmsi.v8i3.576

Girsang, R., Ginting, E. F., & Hutasuhut, M. (2022). Penerapan Algoritma C4.5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, *1*(4), 449. https://doi.org/10.53513/jursi.v1i4.5727

Hendriyanto. (2018). *Manajemen Strategi Pengelolaan Pasar Dalam Meningkatkan Pendapatan Pedagang Perspektif Ekonomi Islam (Studi di Pasar Segamas Purbalingga)*,.

Junaidi, A., Yunita, Y., Agustyani, S., Agustyaningrum, C. I., & Arifin, Y. T. (2023). Klasifikasi Penerima Bantuan Sosial Menggunakan Algoritma C 4.5. *Jurnal Teknik Komputer*, *9*(1), 77–82. https://doi.org/10.31294/jtk.v9i1.14378

Leonardi, M., Emilda, R., Katrin, I., & Yulianto, A. (2021). Prediksi Penjualan Produk Rokok Pada PT. Indomarco Prismatama Menggunakan Algoritma C4.5. *Paradigma - Jurnal Komputer dan Informatika*, *23*(2). https://doi.org/10.31294/p.v23i2.11151

Nuraini, P., Tata Hardinata, J., & Pranayama Purba, Y. (2022). Penerapan Algoritma C4.5 Untuk Klasifikasi Pola Kepuasan Pelayanan E-KTP Di Kantor Camat Pematang Bandar. *Resolusi : Rekayasa Teknik Informatika dan Informasi*, *2*(3), 138–144. https://doi.org/10.30865/resolusi.v2i3.313

Sholihah, H., Satria, F., & Muslihudin, M. (2018). *Implementasi Algoritma C4.5 Klasifikasi Nasabah Potensial ADIRA Dinamika Multi Finance Pringsewu*.

Sugianto, C. A., & Maulana, F. R. (2019). Algoritma Naïve Bayes Untuk Klasifikasi Penerima Bantuan Pangan Non Tunai (Studi Kasus Kelurahan Utama). *Techno.Com*, *18*(4), 321–331. https://doi.org/10.33633/tc.v18i4.2587

Wahid, A., & Agustina, R. (2018). *RANCANG BANGUN PEMBAYARAN EKLEKTRIK DI KANTIN IBBRAHIMY MENGUNAKAN MODEL WATERFALL*.

lassification of Receipt of BPNT Social Assistance Using the C4.5 Algorithm (Case of the Grujugan Kidul Village Office)

7 Internet Source 1%

8 Submitted to University of South Australia
Student Paper 1%

9 Submitted to Universitas Jember
Student Paper 1%

10 Submitted to University of Lancaster
Student Paper 1%

11 ejournal.nusamandiri.ac.id
Internet Source 1%

Exclude quotes   Off   Exclude matches   < 1%
Exclude bibliography On