

# MODEL KLASIFIKASI KEPUASAN MAHASISWA TEKNIK TERHADAP SARANA PEMBELAJARAN MENGGUNAKAN DATA MINING

Elga Mariati <sup>a,1,\*</sup>, Ariesta Lestari <sup>b,2</sup>, Widiatry <sup>c,3</sup>

<sup>a</sup> Jurusan Teknik Informatika, Universitas Palangka Raya

<sup>b</sup> Jurusan Teknik Informatika, Universitas Palangka Raya

<sup>c</sup> Jurusan Teknik Informatika, Universitas Palangka Raya

<sup>1</sup>[elgamariati@gmail.com](mailto:elgamariati@gmail.com), <sup>2</sup>[aries22ta@gmail.com](mailto:aries22ta@gmail.com), <sup>3</sup>[widiatry@gmail.com](mailto:widiatry@gmail.com)

\* corresponding author

## ARTICLE INFO

### Keywords

Data Mining  
Classification  
Decision Tree  
C4.5  
University

## ABSTRACT

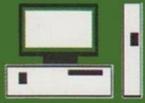
Student is one of the stakeholder in a university. Therefore, student's perception in the quality of learning facilities and infrastructures become important to ensure the university's performance. The Faculty of Engineering of University of Palangka Raya has not comprehensively evaluated the students' satisfactory of the learning's facilities. In this research, methods from data mining approach was implemented to classify whether the students satisfy or not with the quality of the learning's facility in Engineering Faculty. This research compared three data mining algorithm, Decision Tree C4.5, Support Vector Machine, and Naïve Bayes to obtain the best algorithm for the prediction system. 948 responses were collected, 61% of the respondent were satisfied with the quality of the learning facilities and infrastructures, while 39% of the respondents were dissatisfied. The Decision Tree c4.5 had the best performance with accuracy of 88% and precision of 98% compared to the Naïve Bayes and support vector machine. .

## 1. Pendahuluan

Pendidikan memiliki peran yang sangat penting untuk meningkatkan dan mempersiapkan sumber daya manusia (SDM) yang unggul dan berdaya saing tinggi. Disinilah peran institusi pendidikan tinggi menjadi sangat penting dalam menciptakan tenaga ahli yang mampu mengembangkan ilmu pengetahuan dan memberi sumbangan kepada pembangunan. Perguruan tinggi sebagai salah satu institusi pendidikan, dituntut untuk dapat menyelenggarakan pendidikan yang bermutu dan berkualitas kepada *stakeholder*-nya [1]. Kepuasan stakeholder, dalam hal ini mahasiswa, merupakan salah satu indicator keberhasilan perguruan tinggi dalam memastikan mutu dan kualitas tetap terjaga. Kepuasan mahasiswa terutama terhadap fasilitas perguruan tinggi menjadi keharusan agar suatu perguruan tinggi dapat bersaing dengan perguruan tinggi lainnya.

Salah satu perspektif tentang pelayanan pendidikan yang bermutu ialah tersedianya penunjang proses pembelajaran yang memadai [2]. Menurut [3] sarana dan prasarana pelayanan pendidikan adalah fasilitas langsung ataupun tidak langsung yang digunakan dalam proses memberikan materi pelajaran dari pendidik kepada peserta didik. Ketersediaan sarana dan prasarana pada institusi pendidikan merupakan salah satu faktor penunjang dalam proses pencapaian tujuan pendidikan. Oleh karena itu, evaluasi kepuasan mahasiswa terhadap sarana dan prasarana pembelajaran merupakan salah satu bagian penting yang perlu dilakukan agar menjadi bahan acuan peningkatan pelayanan pendidikan di perguruan tinggi.

Fakultas Teknik merupakan salah satu dari delapan fakultas yang ada di Universitas Palangka Raya, terdiri dari empat jurusan yaitu Teknik Sipil, Arsitektur, Teknik Informatika dan Teknik Pertambangan. Penilaian kepuasan mahasiswa terhadap sarana dan pembelajaran pada Fakultas Teknik sudah pernah dilakukan sebelumnya, namun penilaian masih dilakukan

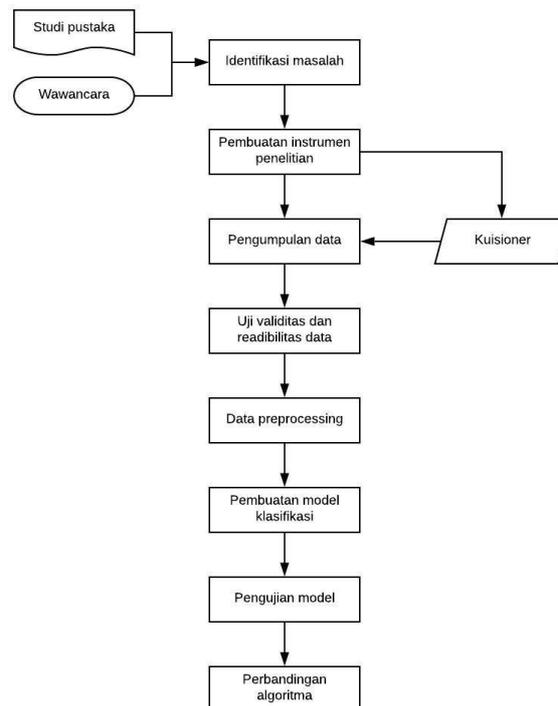
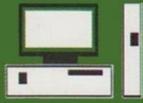


secara partial dan hasil pengumpulan data penilaian kepuasan masih belum pernah dievaluasi sebelumnya. Oleh karena itu, penelitian ini berfokus pada pengklasifikasian kepuasan mahasiswa terhadap mutu dan kualitas sarana prasana di Fakultas Teknik. Hasil klasifikasi ini dapat dipergunakan untuk melihat persepsi mahasiswa terhadap kualitas pendukung perkuliahan dan membantu pihak yang berwenang untuk mengambil keputusan dalam meningkatkan kualitas sarana dan prasarana di lingkungan Fakultas Teknik

Penelitian ini menggunakan teknik *data mining* dalam melakukan klasifikasi. Klasifikasi adalah suatu pengelompokan data dimana data yang digunakan tersebut mempunyai kelas label atau target [4]. Tiap kategori klasifikasi memiliki banyak pilihan algoritma beberapa algoritma yang sering digunakan adalah algoritma *naïve bayes*, *decision tree* C4.5, dan *Support Vector Machine (SVM)*. Dalam [5] menerapkan algoritma Naïve bayes untuk menganalisis tingkat mutu pelayanan (TMP) pada pelanggan: studi kasus PT PLN Batam. Dari 16 atribut yang digunakan hanya 8 atribut yang mempengaruhi tingkat kepuasan pelanggan karena memiliki presentasi jawaban lebih tinggi, sisanya memiliki presentasi relatif kecil. Penggunaan algoritma *naïve bayes* pada penelitian ini menghasilkan akurasi sebesar 74.46% yang menyatakan bahwa jumlah data yang benar lebih banyak daripada data yang salah. Akurasi terhadap prediksi kepuasan pelanggan diberikan oleh [6] menggunakan algoritma C4.5. Algoritma ini digunakan untuk menentukan pengujian prediksi kepuasan pelanggan taksi KOSTI. Atribut masukan kepuasan pelanggan yang digunakan dalam penelitian ini mencakup harga, fasilitas, layanan dan loyalitas. Penggunaan algoritma C4.5 memberikan hasil akurasi sebesar 88,01%. Artikel yang ditulis oleh [7] menggunakan algoritma Support Vector Machine (SVM) untuk mengklasifikasi kepuasan pelanggan Restoran bakso President. Tujuannya untuk mengetahui tingkat kepuasan pelanggan sebagai dasar penentuan strategi bisnis. Implementasi algoritma support vector machine memberikan nilai akurasi yang cukup baik yaitu 80%. Masing-masing penelitian sebelumnya hanya menerapkan satu algoritma saja, kemudian menyatakan bahwa algoritma yang dipergunakan sudah memberikan hasil yang baik dari segi akurasi. Oleh karena itu dalam penelitian akan dilakukan perbandingan ketiga algoritma tersebut berdasarkan keakuratan dalam melakukan klasifikasi.

## 2. Metodologi Penelitian

Gambar 1 memperlihatkan diagram alur terkait tahapan yang dilakukan dalam penelitian ini.



Gambar 1. Diagram Alur Penelitian

### 2.1. Identifikasi masalah

Pada tahapan identifikasi masalah dilakukan wawancara dengan Wakil Dekan Sarana dan Prasarana FT UPR mengenai penilaian kepuasan mahasiswa yang pernah dilakukan sebelumnya. Dari permasalahan yang ditemukan, dilakukan studi pustaka untuk menemukan cara atau teknik yang dapat dipergunakan untuk memecahkan permasalahan tersebut.

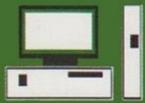
### 2.2. Pengumpulan Data

Pengumpulan data dilakukan dengan membagikan kuesioner kepada responden yaitu mahasiswa di lingkungan FT UPR. Kuesioner ini bersifat tertutup, dimana jawabannya dibatasi atau sudah ditentukan oleh peneliti sehingga responden hanya tinggal memilih jawaban terbaik. Kuesioner ini terdiri atas sembilan pertanyaan terkait kepuasan mahasiswa terhadap sarana dan prasarana di FT UPR. Terdapat satu pertanyaan yang dipergunakan untuk mengklasifikasi jawaban mahasiswa yaitu pertanyaan “Apakah secara umum mahasiswa puas dan tidak puas terhadap kondisi sarana dan prasaran di lingkungan Fakultas Teknik UPR”.

### 2.3. Uji Validitas dan Reabilitas

Secara teknik untuk melihat baik tidaknya alat pengukur penelitian, dapat diketahui dengan melihat kriteria utamanya, yaitu (a) validitas, (b) reabilitas atau kehandalan [8]. Teknik mengetahui suatu instrument valid atau tidak dapat dilakukan dengan menghitung koefisien korelasi uji instrumen dengan uji kriteria/ *product moment* untuk setiap butir menggunakan persamaan 1.

$$r_{xy} = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{(n(\sum x_i^2) - (\sum x_i)^2)(n(\sum y_i^2) - (\sum y_i)^2)}}$$



Hasil perhitungan koefisien korelasi kemudian dibandingkan dengan nilai nilai tabel. Jika  $r_{hitung} >$  atau  $= r_{tabel}$  maka buti dikatakan valid. Jika  $r_{hitung} <$   $r_{tabel}$  maka butir dikatakan tidak valid. Pada penelitian ini digunakan 30 data hasil kuesioner yang telah diisi mahasiswa untuk dihitung nilai validitas dan reabilitasnya.

Pada pengujian reabilitas instrument menggunakan perhitungan koefisien Alfa Cronbach dengan persamaan 2. Hasil perhitungan koefisien *alfa* kemudian dibandingkan dengan nilai nilai tabel. Jika *alfa*  $>$  atau  $= r_{tabel}$  maka instrument penelitian dikatakan reliable. Jika nilai *alfa*  $<$   $r_{tabel}$  maka instrument penelitian dikatakan tidak reliable (Poppy dan Rully, 2017:126).

$$r_i = \frac{k}{(k - 1)} \left\{ 1 - \frac{\sum s_i^2}{s_t^2} \right\}$$

#### 2.4. Data Preprocessing

Proses pre-proses yang dilakukan adalah dengan membersihkan data tidak valid yang ada di dataset. Kriteria data valid tersebut antara lain data dengan jawaban lebih dari 75% tidak setuju/sangat tidak setuju dan pada kesimpulan jawaban memilih puas terhadap kondisi sarana dan prasana pembelajaran di Fakultas Teknik UPR dan sebaliknya. Selain *cleaning data* pada proses ini juga dilakukan pengubahan data kuesioner kedalam format *excel*.

Pada proses data mining dataset jumlah instance di kelas target harus seimbang agar tidak mengganggu kinerja dari data mining yang dibuat. Jumlah *instance* yang tidak seimbang pada tiap-tiap kelas akan mempengaruhi kebenaran prediksi kelas minoritas [9]. Untuk menghindari kesalahan tersebut dalam penelitian ini, jumlah dataset diatur secara acak menjadi 360 data pada kelas puas dan 350 data pada kelas tidak puas sehingga masing-masing kelas seimbang. Proses ini disebut *undersampling*, dimana jumlah instance pada kelas mayoritas dikurangi agar jumlahnya sebanding dengan class minoritas [9].

#### 2.5. Pembuatan Model Klasifikasi

Dataset dibagi menjadi data latih dan data uji dengan proporsi masing-masing yaitu 80% data latih dan 20% data uji. Kemudian dataset siap diolah ke dalam *machine learning* dan dibuat model klasifikasi dengan algoritma *decision tree C4.5*, *Naïve Bayes*, dan *Support Vector Machine* menggunakan data latih.

#### 2.6. Pengujian Model

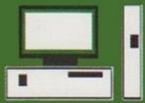
Model dari algoritma *decision tree C4.5*, *support vector machine* dan *naïve bayes* yang dibuat dari data latih. Data testing dipergunakan untuk melihat performa dari model tersebut dengan cara menghitung nilai akurasi dan presisi menggunakan persamaan berikut [10].

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

#### 2.7. Perbandingan Model

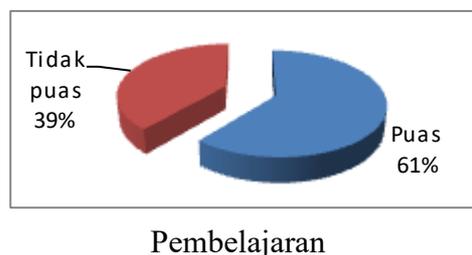
Model dari algoritma *decision tree C4.5*, *support vector machine* dan *naïve bayes* yang dibuat dari data latih, hasil pengujian terhadap data uji dilakukan untuk mendapatkan nilai akurasi dan presisi. Nilai akurasi dan presisi dari ketiga algoritma akan dibandingkan



sehingga diketahui algoritma mana yang dapat diterapkan lebih baik pada evaluasi kepuasan mahasiswa terhadap sarana dan prasarana pembelajaran di Fakultas Teknik UPR.

### 3. Hasil dan Pembahasan

Pengumpulan data penelitian dilaksanakan dengan menyebarkan kuesioner kepada mahasiswa Fakultas Teknik UPR. Kuesioner disebar pada periode tanggal 4 April 2020 sampai dengan tanggal 10 April 2020. Jumlah hasil kuesioner yang berhasil dikumpulkan adalah 981 sampel. Dari 981 data yang terkumpul terdapat 948 data valid yang dapat digunakan dan 33 data tidak valid. Data diklasifikasi tidak valid apabila pada jawaban kuesioner terdapat lebih dari 75% tidak setuju/sangat tidak setuju, namun pada kesimpulan responden memilih puas terhadap kondisi sarana dan prasarana pembelajaran di Fakultas Teknik UPR dan sebaliknya. Grafik hasil analisis kepuasan mahasiswa digambarkan pada Gambar 2.



Gambar 2. Hasil Mahasiswa Terhadap

Pembelajaran

Kuesioner Kepuasan Sarana dan Prasarana

Berdasarkan jawaban pada soal no. 9, 580 sampel diklasifikasikan sebagai puas dan 368 sampel masuk dalam klasifikasi tidak puas. Dikarenakan dataset yang tidak seimbang, maka dataset diatur seimbang secara acak menjadi 360 data dengan nilai puas dan 350 data dengan nilai tidak puas sehingga masing-masing kelas

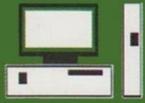
#### 3.1. Uji Validitas dan Readibilitas

Uji validitas dan readibilitas dilakukan pada 30 sampel data yang dipilih secara acak. Hasil uji validitas menunjukkan bahwa semua instrument penelitian valid karena  $r_{hitung}$  lebih besar dari  $r_{tabel}$  (0,361) pada taraf signifikansi 5%. Artinya tiap pernyataan yang diberikan oleh responden berkorelasi dengan skor totalnya dan semuanya dinyatakan valid [11]. Berdasarkan perhitungan uji reabilitas diperoleh nilai  $r_i$  sebesar 0,787. Instrument penelitian disimpulkan *reliable* karena nilai  $r_i$  lebih besar daripada nilai pada tabel R (0,361) untuk 30 sampel.

#### 3.2. Pengujian Klasifikasi

Pada proses ini dibuat model klasifikasi dengan algoritma *decision tree C4.5*, *Naïve Bayes*, dan *Support Vector Machine* menggunakan data latih. Klasifikasi adalah salah satu teknik data mining yang mempelajari sekumpulan data sehingga dihasilkan aturan yang bisa mengklasifikasi atau mengenali data-data baru yang belum pernah dipelajari (Suyanto, 2017:115).

Dari hasil implementasi perhitungan *confusion matrix*, didapatkan keluaran performa akurasi model klasifikasi untuk tiga algoritma seperti yang terlihat pada Tabel 1. Dari tabel tersebut terlihat bahwa model klasifikasi menggunakan algoritma C4.5 menghasilkan prediksi benar terhadap kelas puas sebanyak 56 dan jumlah prediksi benar terhadap kelas tidak puas sebanyak 69. Algoritma ini juga melakukan kesalahan prediksi sebanyak 16 terhadap kelas puas dan 1 terhadap kelas tidak puas. Hasilnya, algoritma C4.5 memiliki akurasi 88% dan presisi 98%. Model klasifikasi menggunakan SVM menghasilkan prediksi benar terhadap kelas puas sebanyak 58 dan jumlah prediksi benar terhadap kelas tidak puas sebanyak 67. Algoritma ini melakukan kesalahan prediksi sebanyak 14 terhadap kelas puas dan 4 terhadap kelas tidak puas. Hasilnya akurasi dari algoritma SVM sebesar 88% dan presisi 98%. Model klasifikasi menggunakan algoritma *naïve bayes* melakukan prediksi benar



terhadap kelas puas sebanyak 53 dan jumlah prediksi benar terhadap kelas tidak puas sebanyak 68. Algoritma ini juga melakukan kesalahan prediksi sebanyak 19 terhadap kelas puas dan 2 terhadap kelas tidak puas. Hasilnya, algoritma C4.5 memiliki akurasi 88% dan presisi 98%.

Ketiga algoritma, decision tree C4.5, SVM dan naïve bayes menunjukkan akurasi yang cukup baik dalam melakukan klasifikasi secara benar baik untuk kelas puas dan kelas tidak puas cukup baik yaitu diatas 80%, dimana akurasi algoritma C4.5 dan algoritma SVM lebih baik daripada akurasi yang diberikan oleh algoritma naïve bayes. 88%. Hasil ini sesuai dengan penelitian-penelitian sebelumnya yang mengklaim bahwa masing-masing algoritma memiliki performa yang bagus pada dataset yang digunakan.

**Tabel 1** Tabel *Confusion Matrix Decision Tree Testing dataset*

A\P	C4.5		SVM		Naïve Bayes	
	1	2	1	2	1	2
1	56	16	58	14	53	19
2	1	69	3	67	2	68

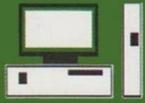
#### 4. Kesimpulan

Kuesioner kepuasan mahasiswa terhadap sarana dan prasana pembelajaran di Fakultas Teknik disebarkan kepada 948 mahasiswa. Hasil kuesioner menunjukkan bahwa 61% mahasiswa menjawab puas terhadap sarana dan prasana pembelajaran dan 39% menjawab tidak puas. Hasil uji validitas instrument penelitian menunjukkan bahwa semua item dalam koesioner valid dengan taraf signifikansi 5% (0,361). Uji reabilitas juga menunjukkan bahwa bahwa instrument penelitian reliable karena nilai (0,787) lebih besar daripada nilai pada tabel R (0,361).

Algoritma Decision Tree C4.5 memiliki performa terbaik dengan akurasi 88% dan presisi 98% dibandingkan algoritma Naïve Bayes dan Support Vector Machine ketika dibuat model dari data latih dan diuji menggunakan data uji dari dataset kuesioner kepuasan mahasiswa terhadap sarana dan prasana pembelajaran.

#### Daftar Pustaka

- [1] D. Purwandani, C. Sutarsih, and Sururi, "Pengaruh Mutu Layanan Sarana dan Prasarana Terhadap Kepuasan Mahasiswa di Fakultas Pendidikan Teknologi dan Kejuruan Universitas Pendidikan Indonesia," *J. ADPEND Tata Kelola Pendidik.*
- [2] A. Sunandar, "Efektivitas Keberadaan Komite Sekolah Untuk Peningkatan Mutu Layanan Sekolah," Skripsi. Universitas Negeri Malang, 2013.
- [3] N. Legiwati, "Pengaruh Pemanfaatan Sarana Dan Prasarana Pembelajaran Dan Motivasi Belajar Terhadap Hasil Belajar Siswa Kelas Vii Smpn 3 Grati Satap Kabupaten Pasuruan," *J. Penelit. dan Pendidik. IPS*, vol. 10, no. 2, pp. 294--309, 2016.
- [4] M. Bharati and M. Ramageri, Data mining techniques and applications. *Indian Journal of Computer Science and Engineering*, vol. 1 no. 4, pp. 301-305. 2010.
- [5] R. Harmann, "Penerapan Metode Data Mining (Naïve Bayes) Untuk Menganalisis Tingkat Mutu Pelayanan (TMP) Pada Pelanggan Studi Kasus," Pada Pelanggan Studi Kasus, Batam: PT PLN Batam. Universitas Batam, 2013.



- [6] M. I. R. Saputra, "Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Kepuasan Pelanggan Taksi Kosti,". Skripsi. Fakultas Ilmu Komputer UDINUS, 2016.
- [7] J. Ferlin, F. A. Bachtiar, and A. N. Rusydi, "Klasifikasi Customer Intent untuk Mengetahui Tingkat Kepuasan Pelanggan menggunakan Metode Support Vector Machine pada Restoran Bakso President.," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 10, 2020.
- [8] R. Indrawan and P. Yaniawati, *Metodologi penelitian kuantitatif, kualitatif, dan campuran untuk manajemen, pembangunan, dan pendidikan*. Bandung: Refika Aditama, 2014.
- [9] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, "Learning from class-imbalanced data: Review of methods and applications," *Expert System Application*, vol. 73, pp. 220--239, 2017.
- [10] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37--63, 2011.
- [11] Sugiyono, *Statistika untuk penelitian*. Bandung: Alfabeta, 2014.