

KLASIFIKASI PENERIMA KIP SMPN 5 UNGARAN MENGUNAKAN ALGORITMA NAÏVE BAYES BERBASIS WEB

Dinia Kurnia Pratiwi¹, Muhammad Hizbullah²

Mahasiswa Program Studi Teknik Informatika-S1, Fakultas Ilmu Komputer, Universitas Dian
Nuswantoro Semarang Jl. Imam Bonjol No. 207, Jl.Nakula No.5-11, Semarang, Kode
Pos.50131, Telp.(024) 3517261

Email : 111201609772@mhs.dinus.ac.id, 111201610006@mhs.dinus.ac.id

ABSTRAK

Program Kartu Indonesia Pintar (KIP) diselenggarakan oleh Kementerian Pendidikan dan Kebudayaan dan Kementerian Sosial. Kartu ini bertujuan menghilangkan hambatan anak usia sekolah secara ekonomi untuk berpartisipasi di sekolah. Dengan demikian, mereka memperoleh akses pelayanan pendidikan yang lebih baik, mencegah murid mengalami putus sekolah, serta mendorong anak yang putus sekolah kembali bersekolah. Dalam pengkategorian ini akan di gunakan klasifikasi data mining metode naive bayes. Data mining bermanfaat dalam pengambilan informasi penting pada data yang berjumlah sangat besar. Tujuan penelitian ini adalah guna mengetahui kelayakan penerima KIP di SMPN 5 Ungaran dan mendapatkan akurasi data mining dengan metode naive bayes.

Kata Kunci : *Kartu Indonesia pintar, data mining, klasifikasi, metode naive bayes.*

BAB 1

PENDAHULUAN

KIP (Kartu Indonesia Pintar) merupakan kartu yang ditujukan bagi keluarga miskin dan rentan miskin yang ingin menyekolahkan anak yang berusia 7-18 tahun secara gratis. Mereka yang mendapat KIP ini akan diberikan dana tunai dari pemerintah secara reguler yang tersimpan dalam fungsi kartu KIP untuk bersekolah secara gratis tanpa biaya. Program KIP sendiri akan

ditujukan pada 15,5 juta keluarga kurang mampu di seluruh Indonesia yang memiliki anak usia sekolah 7 hingga 18 tahun baik yang telah terdaftar maupun yang belum terdaftar di sekolah maupun madrasah. Dengan program KIP ini diharapkan angka putus sekolah bisa turun dengan drastis.[1]

Penerima program KIP ini sendiri diprioritaskan pada :

- > Penerima BSM dari Pemegang KPS yang telah ditetapkan dalam SP2D 2014
- > Anak usia sekolah (6-21 tahun) dari keluarga pemegang KPS/KKS yang belum ditetapkan sebagai Penerima manfaat BSM
- > Anak usia sekolah (6-21 tahun) dari keluarga peserta PKH
- > Anak usia sekolah (6-21 tahun) yang tinggal di Panti Asuhan/Sosial
- > Siswa/santri (6-21 tahun) dari Pondok Pesantren yang memiliki KPS/KKS (khusus untuk BSM Mandrasah)
- > Anak usia sekolah (6-21 tahun) yang terancam putus sekolah karena kesulitan ekonomi dan/atau korban musibah berkepanjangan/bencana alam melalui jalur FUS/FUM;
- > Anak usia sekolah yang belum atau tidak lagi bersekolah yang datanya telah dapat direkapitulasi pada Semester 2 (TA) 2014/2015.

Dari uraian kasus diatas, untuk mengelompokkan sebuah layak tidaknya untuk mendapatkan KIP dapat menggunakan klasifikasi Naïve Bayes data mining. Proses klasifikasi merupakan sebuah proses penemuan model atau fungsi yang dapat menjelaskan atau membedakan konsep data, yang bertujuan untuk dapat memperkirakan konsep tersebut dari suatu objek yang labelnya tidak diketahui [2]. Data mining adalah proses untuk memperoleh keterkaitan dalam sebuah data yang belum diketahui, kemudian disajikan dengan suatu cara

yang dapat dipahami oleh pengguna sehingga keterkaitan tersebut dapat dijadikan sebagai dasar pengambilan keputusan [3]. Metode Naive Bayes dipilih karena hasil lebih akurat dibanding metode C4.5 dan teknik lebih sederhana, hal ini dibuktikan oleh Doreswamy, Hemanth. K. S dalam jurnalnya tahun 2012 “Performance Evaluation of Predictive Classifiers For Knowledge Discovery

From Engineering Materials Data Sets”.

Dengan adanya klasifikasi penerima KIP menggunakan metode naive bayes diharapkan dapat membantu pihak sekolah dalam mengelompokkan Kelayakan penerima KIP .

1.1 Identifikasi Masalah

sekolah juga turut memiliki peran dalam penentuan penerimaan KIP dengan mendaftarkan siswa yang kurang mampu dan meminta keluarga siswa membawa surat keterangan PKH, KIS atau SKTM .Dalam proses untuk meminta surat keterangan ini membutuhkan waktu yang cukup lama, karena harus melalui ketua RT hingga dapat di terbit kan oleh pihak kelurahan. Maka untuk dapat mengelompokkan sebuah layak atau tidak nya sesuai dengan kriteria yang ada dapat menggunakan klasifikasi Naïve Bayes data mining

Dari uraian kasus diatas, untuk mengelompokkan sebuah perkara penting pada kasus pidana umum yang terbagi dapat menggunakan klasifikasi Naïve Bayes data mining.

1.2 Lingkup Masalah

Pengumpulan data siswa SMPN 5 Ungaran tahun ajaran 2018-2019 dengan sample 60 dataset .

Pengklasifikasian penerima KIP menggunakan metode naïve bayes berdasarkan faktor yang mempengaruhinya yaitu jenis kelamin, alamat, pekerjaan orang tua, penghasilan orang tua, pendidikan orang tua, dan penerima KPS (Kartu Perlindungan Sosial).

1.3 Rumusan Masalah

Berdasarkan latar belakang di atas, dapat dirumuskan permasalahan yang akan diselesaikan adalah “Bagaimana kinerja naïve bayes dalam pengklasifikasian penerima KIP di SMPN 5 Ungaran dan menerapkannya ke dalam aplikasi data mining? “.

1.4 Tujuan dan Manfaat Penelitian

Adapun tujuan dan manfaat dari penelitian ini adalah sebagai berikut :

- a. Menambah wawasan dan referensi untuk mempelajari klasifikasi data mining menggunakan metode naïve bayes dan mengembangkan pada kasus lainnya.
- b. Dapat mengimplementasikan klasifikasi data mining menggunakan metode naïve bayes dalam dataset siswa SMPN 5 Ungaran

- c. Menambah literatur bagi pembaca dalam menerapkan algoritma Naïve Bayes Classifier.

BAB 2

LANDASAN TEORI

2.1 Tinjauan Pustaka

Hasil dari pencarian jurnal di Google Scholar, belum banyak penelitian yang mengangkat topik tentang pengklasifikasian Penerima KIP maka dari itu sebagai referensi dari penelitian yang di buat akan mengambil beberapa penelitian yang menggunakan Algoritma Naïve Bayes sebagai pengklasifikasian data masing-masing topik penelitian, di antaranya:

Dinda Ayu Muthia (2017) dalam penelitian ini digunakan metode pemilihan fitur, yaitu Genetic algorithm agar bisa meningkatkan akurasi pengklasifikasi Naïve Bayes. Penelitian ini menghasilkan klasifikasi teks dalam bentuk positif atau negatif dari review restoran. Pengukuran berdasarkan akurasi Naïve Bayes sebelum dan sesudah penambahan metode pemilihan fitur. Evaluasi dilakukan menggunakan 10 fold cross validation. Sedangkan pengukuran akurasi diukur dengan confusion matrix dan kurva ROC. Hasil penelitian menunjukkan peningkatan akurasi Naïve Bayes dari 86.50% menjadi 90.50%.

Alfa Saleh (2015) menerapkan algoritma naïve bayes classifier untuk

memprediksi besarnya penggunaan listrik tiap rumah tangga agar lebih mudah mengatur penggunaan listrik. dari 60 data penggunaan listrik rumah tangga yang diuji dengan metode naïve bayes, maka diperoleh hasil persentase 78,3333% untuk keakuratan prediksi, di mana dari 60 data penggunaan listrik rumah tangga yang diuji terdapat 47 data penggunaan listrik rumah tangga yang berhasil diklasifikasikan dengan benar.

Akhmad Pandhu Wijaya, Heru Agus Santoso (2016) Klasifikasi ini teknik dari data mining dan pertambangan teks juga digunakan untuk mencari atau mengatur kelas dibedakan dengan menggunakan beberapa fungsi dengan tujuan memungkinkan model untuk digunakan untuk data pengujian. Pada penelitian ini, objeknya adalah Situs Web Jawa Tengah dan diklasifikasikan oleh Naïve Bayes Classification (NBC). Dengan menggunakan metode ini diharapkan memfasilitasi klasifikasi dokumen bahasa Indonesia untuk identifikasi konten e-government.

Amrin, Hafdiarsya Saiyar (2018) Pada penelitian ini, menerapkan metode klasifikasi data mining, yaitu metode Naïve Bayes untuk mendiagnosa penyakit tuberkulosis. Berdasarkan hasil pengukuran performa dari model tersebut dengan menggunakan metode pengujian Cross Validation, Confusion Matrix dan Kurva ROC, diketahui bahwa metode Naïve Bayes dengan tingkat akurasi sebesar 94,18% dan nilai area under the curve (AUC) sebesar 0,977. Hal ini menunjukkan model yang dihasilkan termasuk

kategori klasifikasi sangat baik karena memiliki nilai AUC antara 0.90-1.00.

Andriana Candra Dewi, Arief Andy Soebroto, M. Tanzil Furqon (2015) Dalam sistem pakar ini menggunakan metode Naive Bayes sebagai metode inferensi untuk mendiagnosis penyakit. Jenis penyakit yang dapat dikenali lewat sistem pakar ini adalah 11 jenis penyakit sedangkan gejala yang dapat dikenali sistem pakar adalah 20 jenis gejala. Hasil pengujian keakuratan 26 data kasus uji, telah menghasilkan tingkat persentase kesesuaian 96,15%.

2.2 Kartu Indonesia Pintar

Program Indonesia Pintar Program Indonesia Pintar adalah salah satu program nasional yang bertujuan untuk meningkatkan angka partisipasi pendidikan dasar dan menengah, meningkatkan angka berkelanjutan pendidikan, menurunkan kesenjangan partisipasi pendidikan, meningkatkan kesiapan siswa pendidikan menengah untuk menghadapi pasar kerja. Program Indonesia Pintar diselenggarakan oleh Kementerian Pendidikan dan Kebudayaan dan Kementerian Sosial. Program Indonesia Pintar merupakan penyempurnaan dari program Bantuan Siswa Miskin (BSM) yang memberikan bantuan tunai pendidikan bagi anak usia sekolah yang berhak terutama dari keluarga pemegang KKS dan ditandai dengan Kartu

Indonesia Pintar (KIP) TNP2K (2016). Selain KKS dan KIP, pihak sekolah juga bisa mendaftarkan siswanya yang kurang mampu dengan meminta keluarga siswa membawa surat

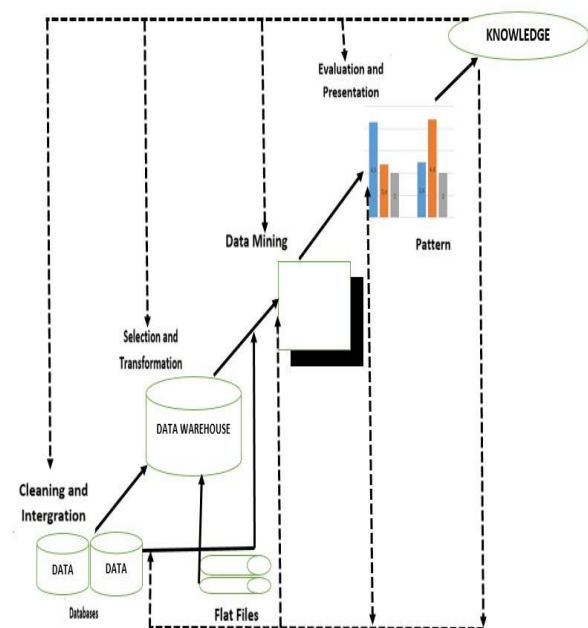
keterangan PKH, KIS atau SKTM. SKTM adalah surat keterangan tidak mampu yang ditujukan untuk keluarga miskin dan bisa didapat melalui kelurahan setempat Pemerintah Kabupaten Bangil (2016). KKS atau Kartu Keluarga Sejahtera adalah bantuan non tunai melalui pembukaan rekening simpanan bagi rakyat kurang mampu Azarine (2016). PKH adalah Program Keluarga Sejahtera yang merupakan bantuan non tunai bagi Rumah Tangga Sangat Miskin (RTSM) yang akan menerima bantuan jika keluarga tersebut menyekolahkan anaknya dengan ingkat kehadiran tertentu Kementerian Sosial (2016). KPS adalah Kartu Perlindungan Sosial yang merupakan kartu untuk mendapatkan manfaat dari Program Subsidi Beras, Bantuan Siswa Miskin dan Bantuan Langsung Sementara Masyarakat TNP2K (2016). KIS atau Kartu Indonesia Sehat adalah penanda masyarakat miskin agar mendapatkan pelayanan kesehatan melalui Jaminan Kesehatan Nasional TNP2K (2016).[4]

2.3 Data Mining

Pada dasarnya data mining tersusun atas algoritma atau metode yang memungkinkan seseorang untuk mendapatkan sebuah

wawasan dasar dan pengetahuan pada suatu data yang besar. Data mining juga dapat dikatakan sebagai penggabungan bidang sistem database, statistik, pembelajaran mesin, dan pengenalan pola. Dalam pemrosesannya adapun tahapan pra-pengolahan seperti ekstraksi data, pembersihan data, fusi data, reduksi data dan fitur konstruksi.

2.4 Tahap-Tahap Data Mining Berikut ini merupakan tahapan data mining yang sangat berpengaruh pada kualitas data :



Gambar 1 Tahap-tahap Data Mining

1. Pembersihan data (data cleaning)

Tahap pembersihan data yaitu proses menghilangkan noise dan data yang tidak konsisten atau tidak relevan. Selain itu, ada juga atribut-atribut data yang tidak relevan yang lebih baik dibuang. Pembersihan data akan mempengaruhi performansi dari teknik data mining karena data yang ditangani

akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data (data integration)

Tahap integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut usia, jenis pidana umum, nomor identitas dan lainnya. Hal ini dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang. **3. Seleksi Data (Data Selection)**

Tahap seleksi data yang terdapat pada database sering kali tidak semuanya dipakai, maka hanya data yang sesuai untuk dianalisis dan akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti pola kecenderungan konsumen dalam membeli kebutuhan pada sebuah supermarket, hasil analisis tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data (Data Transformation)

Tahap transformasi data yaitu data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis

asosiasi dan clustering hanya bisa menerima input data kategorikal.

5. Proses mining

Proses mining ini adalah proses yang utama saat metode diterapkan untuk menemukan sebuah pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi pola (pattern evaluation),

Tahap evaluasi pola dapat digunakan untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan. Apabila hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

7. Presentasi pengetahuan (knowledge presentation),

Presentasi pengetahuan merupakan proses visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan agar dipahami oleh orang tidak mengetahui data mining.

2.5 Klasifikasi

Klasifikasi adalah model atau fungsi yang memprediksi label kelas yang dapat menjelaskan atau membedakan konsep data, bertujuan untuk dapat memperkirakan konsep tersebut dari suatu objek yang

labelnya tidak diketahui. [4] Untuk membangun model yang membutuhkan satu set poin dengan label kelas yang benar, yang disebut training set. Setelah mempelajari model tersebut, maka otomatis memprediksi kelas untuk setiap titik baru. Berbagai jenis klasifikasi model telah diusulkan seperti *Decision tree*, *Naive Bayes classifier*, *Neural networks*, *Statistical analysis*, *K-Nearest Neighbor classifier*, *Support vector machines*.

2.6 Naive Bayes Classifier

Naive Bayes classifier adalah suatu pengklasifikasian data statistika yang dapat digunakan untuk memprediksi probabilitas keanggotaan label kelas. Berikut ini adalah bentuk umum dari naive bayes : $P(X|Ci) = \frac{P(X|Ci) * P(Ci)}{P(X)}$ Keterangan :

X = data dengan class yang belum diketahui.

Ci = hipotesis dari data X.

$P(X|Ci)$ = probabilitas kondisi X berdasarkan hipotesis Ci.

$P(Ci)$ = probabilitas hipotesis Ci.

$P(X)$ = probabilitas kondisi X.

Adapun cara pengujian data naive bayes, sebagai berikut : 1. Menghitung $P(Ci)$ untuk setiap kelas:

Nilai $P(Ci)$ untuk i

Menghitung peluang kelas (i) :

$$P(class(i)) = \frac{\text{jumlah kelas (i)}}{\text{jumlah keseluruhan data pada masing-masing atribut.}}$$

$P(X|Ci) = \text{jumlah } x / \text{jumlah kelas (i).}$

2. Menghitung $P(X|Ci)$ untuk semua kelas dengan mengalikan semua variabel.

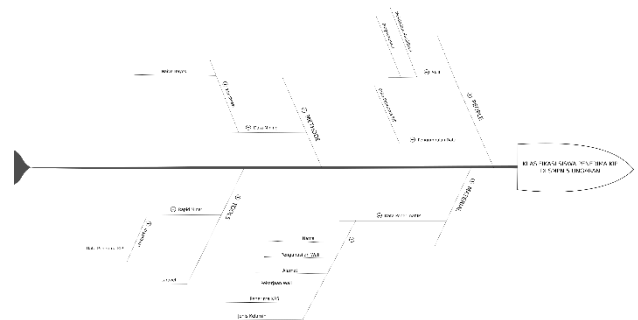
$$\sum P(X|Ci) * P(Ci)$$

3. Membandingkan hasil perkalian $P(X|Ci)$, setelah itu mengambil hasil yang terbesar sebagai hasilnya.

BAB 3

METODOLOGI PENELITIAN

3.1 Kerangka Pemikiran

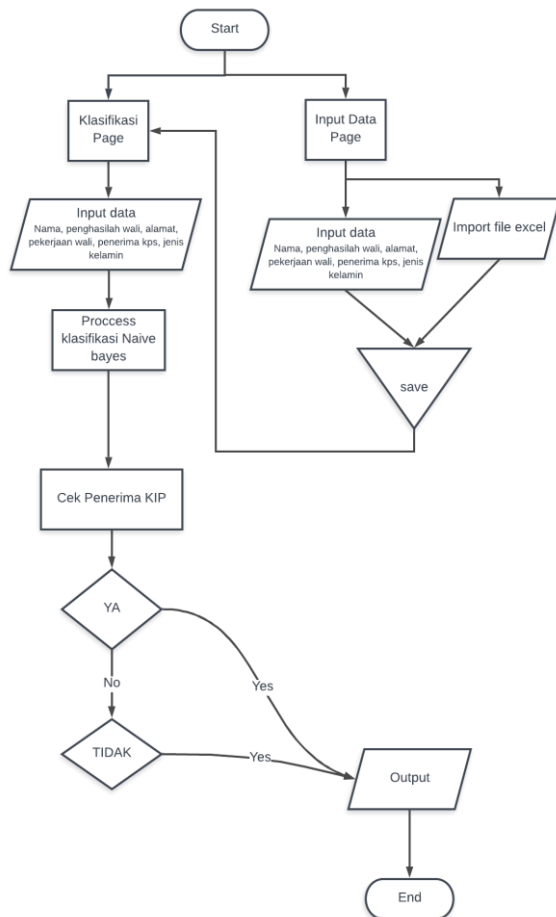


Gambar 2 Fishbone Diagram

Dari kerangka diatas, dapat dijelaskan bahwa masalah yang timbul yaitu tentang Klasifikasi Penerima KIP. Masalah ini bermula dari banyaknya kriteria yang ditentukan untuk dapat menerima KIP . yang sebelumnya dikelompokkan menjadi 2 jenis ya atau tidak layak dengan faktor yang mempengaruhinya yaitu Jenis kelamin, alamat, pendapatan orang tua , penghasilan orang tua, pendidikan orang tua, penerima KPS. Proses pengelompokan ini dinamakan klasifikasi yang akan menggunakan metode naive bayes

pada data mining dengan tools RapidMiner5 dan aplikasi berbasis web dengan framework laravel.

maka akan muncul hasil berupa layaknya untuk menerima KIP.



Dimulai dari Input data siswa pada SMPN 5 Ungaran berupa Jenis kelamin, Alamat, Penghasilan Orang tua, Pekerjaan Orang tua, Pendidikan Orang tua, Penerima KPS Kemudian, penerima akan di klasifikasikan menggunakan metode naive bayes layak atau tidak nya untuk menerima KIP.

Apabila penerima tidak sesuai dengan kriteria penerima KIP maka akan di identifikasi kembali menuju peninjauan data Penerima KIP. Apabila penerima telah cocok dengan salah satu kriteria

DAFTAR PUSTAKA

- [1] <https://www.cermati.com/artikel/kartu-indonesia-pintar-apa-itu-manfaat-dan-siapa-berhak-menerima> diakses 11 April 2019
- [2] E. U. A. S. Bonifacius Vicky Indriyono, “Klasifikasi Jenis Buku Berdasarkan Judul Dan Sinopsis Menggunakan Metode Naïve Bayes Classifier (Studi Kasus: STMIK KADIRI),” *J. Sist. Inf.*, vol. 5, no. 4, 2015.
- [3] M. Ridwan, H. Suyono, and M. Sarosa, “Penerapan Data Mining Untuk Evaluasi

Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier,” *Eeccis*, vol. 7, no. 1, pp. 59–64, 2013.

- [4] M. J. Zaki and M. J. Meira, *Data Mining and Analysis: Fundamental Concepts and Algorithms*. 2014.

- [5] Dinda Ayu Muthia, “Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naïve Bayes,” Vol.2. No. 2 Februari 2017

E-ISSN: 2527-4864

- [6] Alfa Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga" Citec Journal, Vol. 2, No. 3, Mei 2015 –Juli 2015 ISSN: 2354-5771
- [7] A. P. Wijaya and H. A. Santoso, "Naïve Bayes Classification Pada Klasifikasi Dokumen Untuk Identifikasi Konten E-Government," *J. Appl. Intell. Syst.*, vol. 1, no. 1, pp. 48–55, 2016.
- [8] Amrin, Hafdiarsya saiyar "Aplikasi Diagnosa Penyakit Tuberculosis Menggunakan Algoritma Naïve Bayes" Jurnal Riset Komputer (JURIKOM), Vol. 5 No. 5, Oktober 2018 ISSN 2407-389X (Media Cetak) Hal: 498-502
- [9] Andriana Candra Dewi, Arief Andy Soebroto, M. Tanzil Furqon, "Sistem Pakar Diagnosa Penyakit Sapi Potong Dengan Metode Naïve Bayes" *Journal of Environmental Engineering & Sustainable Technology JEEST* Vol. 02 No. 02, November 2015, Pages 72-78
- [10] Nofriansyah, Dicky. 2014. *Konsep Data Mining Vs Sistem Pendukung Keputusan*. Yogyakarta: Deepublish.
- [11] <http://mbahsecond.blogspot.co.id/2013/10/metodologi-pengembanganwaterfall.html> diakses pada 1 April 2019
- [12] <http://ariplie.blogspot.co.id/2015/04/pengertian-dan-tahap-tahap-datamining.html> diakses tanggal 4 April 2019.
- [13] <http://nur-statistika-unpad.blogspot.co.id/2015/07/crisp-data-mining.html> diakses tanggal 5 April 2019.
- [14] <http://yusnitadewi.ilearning.me/2013/10/01/eksplorasi-data-mining-software-rapid-miner/> diakses tanggal 7 April 2019.
- [16] <http://yusnitadewi.ilearning.me/2013/10/01/eksplorasi-data-mining-software-rapid-miner/> diakses tanggal 4 Juni 2016.