

KLASIFIKASI ASPIRASI MAHASISWA DENGAN NAIVE BAYES CLASSIFIER

CLASSIFICATION OF STUDENT'S ASPIRATION WITH NAÏVE BAYES CLASSIFIER

Tika Mahardhika P*¹, Ifan Rizqa M.Kom²

^{1,2}Jl. Imam Bonjol 207, telp/fax (+6224) 3517261/ (+6224) 3569684

³Jurusan Teknik Informatika, FIK UDINUS, SEMARANG

e-mail: 1tikamahardhika96@gmail.com, 2risqa.ifan@dsn.dinus.ac.id

Abstrak

Aspirasi mahasiswa adalah berbagai tuntutan dari mahasiswa yang dikemas dalam ide kreatif untuk mengusulkan proses perubahan atas suatu hal. Rata-rata aspirasi yang disampaikan berupa keluhan dan harapan. Aspirasi berguna sebagai bahan evaluasi dan deteksi dini terhadap kelemahan sistem kualitas perguruan tinggi menjadi lebih baik. kegiatan ini terjadi di UDINUS, DPM sebagai pihak yang mengelola aspirasi mahasiswa. Aspirasi didapatkan melalui sebuah mekanisme yang telah ditentukan seperti penyebaran kuisioner secara manual atau berbasis google form. Kuisioner yang disediakan mengharuskan mahasiswa mengisi isi aspirasi sesuai katagori yang disediakan. Namun, permasalahan yang dihadapi terkadang mahasiswa salah memilih katagori sesuai isi aspirasi yang ditulis. Maka dari itu perlu dibuat sistem yang dapat mengkatagorisasikan aspirasi mahasiswa secara otomatis. Klasifikasi teks dokumen menjadi cara yang paling baik untuk menentukan katagori berdasarkan isi aspirasi mahasiswa. Metode Naïve Bayes Classifier digunakan karena mampu menghasilkan akurasi yang tinggi. Dengan 1000 dokumen data latih yang masing-masing 250 dokumen setiap katagori yaitu "Sarana dan Prasarana", "Dosen", "Kepegawaian dan Sistem Akademik", "Saran dan Masukan" menghasilkan akurasi 90.20%. Dapat dikatakan metode ini layak implementasi dalam penelitian ini.

Kata kunci— Aspirasi Mahasiswa, Katagori, Naïve Bayes Classifier

Abstract

Students aspiration are various demands from the student that packed in creative idea to propose changing process of a thing. Mostly, aspiration delivered in complaints and expectation. Aspiration is used for evaluating the laxity and early detection in university quality system for the better. This activity took place in UDINUS, and DPM is the unit to manage the students aspiration. Aspiration is obtained through predetermined mechanism such as manual questionnaire distribution and or using google form. The provided questionnaire requires student to fill the content according to the provided aspiration categories. However, the problem is sometimes the student choose the wrong category according to the content. Therefore it is needed to create an application that can classified the students aspiration automatically. Document text classification become the best way to determine the category based on the content of the students aspiration. Naïve bayes classifier method is used because it is capable to produce high accuracy. With 1000 data training document of each category, "Sarana dan Prasarana", "Dosen", "Kepegawaian dan Sistem Akademik", "Saran dan Masukan" generate 90.20% accuracy. It can be said that this method is worth to implement in this research.

Keywords— Students Aspiration, Classification, Naïve Bayes Classifier

1. PENDAHULUAN

Aspirasi mahasiswa adalah berbagai tuntunan dari mahasiswa yang dikemas dalam ide kreatif untuk mengusulkan proses perubahan atas suatu hal [1]. Aspirasi mahasiswa dianggap baik manakala disampaikan tidak hanya secara lisan namun juga tertulis. Aspirasi tersebut disertai dengan argumentasi ilmiah dan mengedepankan norma serta kaidah keilmuannya [2]. Rata-rata aspirasi yang disampaikan berupa keluhan dan harapan. Aspirasi berguna sebagai bahan evaluasi dan deteksi dini terhadap kelemahan sistem kualitas perguruan tinggi menjadi lebih baik. Evaluasi tersebut dapat menjadi jaminan mutu perguruan tinggi agar mahasiswa memperoleh kepuasan dan tetap terjadi perbaikan yang berkelanjutan [3]. Oleh karena itu untuk memperbaiki standar kampus diperlukan aspirasi mahasiswa.

Secara umum di beberapa universitas, mahasiswa dapat menyampaikan aspirasinya melalui kuesioner online. Melalui kuesioner tersebut aspirasi akan terekam dan tersusun dengan baik. Aspirasi akan dikaji guna menyaring mana yang harus disampaikan. Ketika sebuah aspirasi telah diputuskan, lalu disampaikan kepihak terkait untuk ditindak lanjuti [4]. Hal itu terjadi di Universitas Dian Nuswantoro Semarang (UDINUS), bahwa terdapat unit yang khusus menampung aspirasi mahasiswa yang selanjutnya hasil aspirasi tersebut dikaji dan disampaikan kebagian terkait oleh Dewan Perwakilan Mahasiswa (DPM) fakultas. Aspirasi mahasiswa, didapatkan melalui sebuah mekanisme yang telah ditentukan seperti penyebaran kuisisioner berupa kertas *form* yang telah disiapkan dan juga dengan antar muka berbasis *google form* yang susunan isiannya disesuaikan sama persis dengan format kuisisioner yang telah dibagikan secara manual.

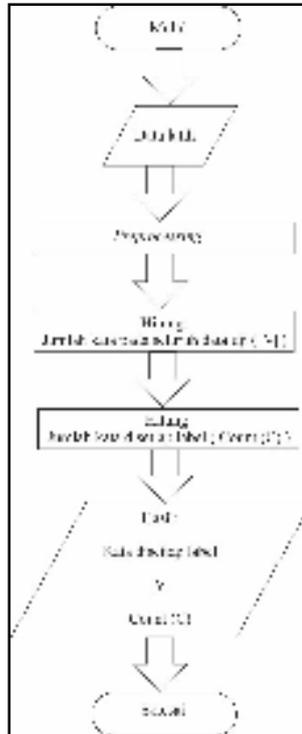
Beberapa kesulitan sering dihadapi DPM. Yang pertama, ketika mahasiswa mengisi kuesioner disediakan beberapa pilihan kategori aspirasi. Namun, terkadang antara kategori yang dipilih dan isi aspirasi tidak cocok. Sehingga DPM mengkatagorikan aspirasi tersebut secara manual. Yang kedua, kuesioner yang masih dalam bentuk kertas dan banyaknya aspirasi yang terkumpul tidak terekam serta tidak tersusun dengan baik. Sehingga DPM merasa kesulitan untuk mengkaji aspirasi yang ada.

Dari beberapa masalah yang disampaikan, klasifikasi teks dokumen menjadi cara yang paling baik untuk menentukan katagori berdasarkan isi aspirasi mahasiswa. Dengan menggunakan *information retrieval* dapat menyelesaikan persoalan tersebut. Ada beberapa metode klasifikasi didalamnya. Namun, *Naive Bayes Classifier* yang akan digunakan. Alasannya metode ini dapat mengklasifikasi teks dokumen dengan berbagai kategori, sebagai contoh metode *Naive Bayes Classifier* dapat mengklasifikasi postingan twitter tentang kemacetan lalulintas kota Bandung dengan menggunakan 100 data latih. Metode ini mampu menghasilkan 78 % serta 13106 data uji mampu menghasilkan keakurasian sebesar 91,6% [5]. Dengan adanya referensi seperti ini membutuhkan bahwa metode *Naive Bayes Classifier* dapat memberikan keakurasian yang tinggi walau dengan data latih yang sedikit. Penelitian juga dilakukan oleh Dyarsa dengan metode yang sama menghasilkan keakurasian sampai 91% dengan menggunakan data latih sebanyak 1000 data di masing-masing kategori [6].

Dengan beberapa uraian yang telah disampaikan, penulis ingin melakukan penelitian dengan membuat sistem untuk mengklasifikasi aspirasi mahasiswa dengan metode *Naive Bayes Classifier*. Di dalam sistem tersebut, mahasiswa dapat mengisi aspirasi tanpa memilih kategori yang biasanya disediakan. Karena nantinya ketika *form* aspirasi diisi maka program akan mendeteksi/ mengklasifikasi aspirasi tersebut. Selain itu disediakan *chart diagram* dan rekaman aspirasi mahasiswa yang akan memudahkan dan mempercepat DPM untuk mengkaji. Laporan aspirasi yang telah dikaji akan disampaikan kepihak terkait.

2. METODE PENELITIAN

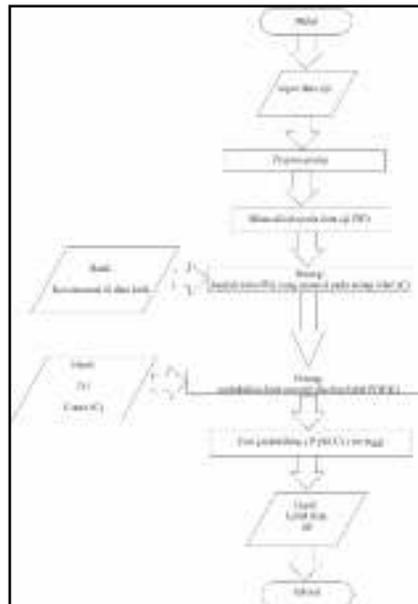
Pada penelitian ini menggunakan metode *Naive Bayes Classifier* sebagai proses klasifikasi aspirasi mahasiswa. Alasan menggunakan metode ini karena penelitian sebelumnya membuktikan memiliki keakurasian yang cukup tinggi walaupun dokumen yang digunakan hanya sedikit. Dalam menerapkan metode ini, proses klasifikasi dibagi menjadi 2 bagian yaitu pengolahan data latih dan data uji. Adapun proses pengolahan data latih disediakan flowchart dan penjelasan sebagai berikut:



Gambar 1. *Flowchart* Data Latih

1. Data latih telah diberi katagori dan tersimpan dalam database.
2. Semua data latih memalui proses preprocessing untuk membuat data lebih terstruktur. Hasil dari preprocessing menghasilkan kemunculan kata semua data latih (*term*).
3. Dari seluruh kata yang muncul dihitung jumlahnya ($|V|$).
4. Kemudian hitung semua kata disetiap katagori Count (C).
5. Proses pelatihan selesai.

Sedangkan pengolahan data uji disediakan flowchart dan penjelasan sebagai berikut:



Gambar 2. Flowchart Data Uji

1. Input data uji untuk mengetahui katagorinya.
2. Data uji masuk ke tahap preprocessing untuk menghilangkan karakter yang tidak bermakna.
3. Hasil preprocessing akan muncul *term* dari data uji. Kemudian hitung jumlah kata yang muncul tersebut.
4. Hitung jumlah term data uji yang muncul di *term* data latih setiap katagorinya.
5. Hitung probabilitas $P(W_i|C)$ dengan melibatkan jumlah seluruh *term* pada data latih ($|V|$) dan jumlah *kata* dari data latih setiap katagori ($Count(c)$)
6. Cari hasil probabilitas $P(W_i|C)$ yang tertinggi.
7. Maka outputnya data uji tersebut diketahui katagorinya.
8. Proses data uji selesai.

3. HASIL DAN PEMBAHASAN

3.1 Pengelolaan Data

Pengolahan data dilakukan guna untuk menguji kelayakan algoritma terhadap aspirasi mahasiswa. Adapun dalam penelitian ini dibutuhkan 2 jenis data yaitu data latih dan data uji. Berikut ini penjelasan masing- masing jenis data tersebut:

3.1.1 Data Latih

Data latih diperoleh dari DPM fakultas di UDINUS. Permasalahan yang terjadi setiap fakultas memiliki kategori yang berbeda. Maka dari itu perlu penyamaan katagori untuk mempermudah dalam mengujian algoritma pada aspirasi mahasiswa. Dari data latih yang didapat, 1000 dokumen aspirasi mahasiswa valid untuk digunakan. Dokumen dikatakan valid karena sebelumnya sudah memiliki katagori yang ditentukan. Sehingga dokumen dari data latih tidak dikelompokkan secara manual. Berikut kumpulan data latih yang didapatkan:

Tabel 1. Data Latih Katagori Sarana Dan Prasarana

NO	Dokumen (Aspirasi Mahasiswa)	Katagori
1	Mahasiswa membutuhkan lahan parkir yg lebih luas dari mobil maupun motor . Dan mahasiswa membutuhkan tempat jeda istirahat setelah selesai kuliah.dosen jangan seenaknya memberikan jam pengganti"	Sarana dan Prasarana
2	gedung H white bordnya kecil, dan berisik akibat pembangunan gedung yang tidak selesai selesai. penyebaran area wi-fi kurang luas dengan kecepatan yang lambat.	Sarana dan Prasarana
3	dosen: dosennya sering masuk terus dan tepat waktu jadi tidak bisa wifian sepenuhnya	Dosen
4	Banyak Dosen Udinus yang cara mengajar kurang asyik dan sulit dipahami	Dosen
7	Tukang parkir kurang detail (Motornya masih ada lahan tapi gak boleh masuk seharusnya menata dengan rapi)	Kepegawaian dan Sistem Akademik
8	Terkadang informasi beasiswa diberitahukan pada mahasiswa pada waktu mendekati tenggang	Kepegawaian dan Sistem Akademik
.	.	.
.	.	.
249	perbaiki yang ada dan yg sering menjadi keluhan mahasiswa ditampung dan dilaksanakan	Saran dan Masukan
250	sebaiknya mempunyai gedung sendiri untuk fak kesehatan , seperti univ lain	Saran dan Masukan

3.1.2 Data Uji

Data uji digunakan untuk mengetahui algoritma dapat digunakan untuk mendeteksi katagori secara otomatis. Data uji yang digunakan tidak memiliki kategori sebelumnya. Dengan disertakan NIM maka data latih tersebut valid untuk dipakai. Oleh karena itu data latih didapatkan ketika peneliti telah menyelesaikan sistem klasifikasi kategori aspirasi mahasiswa secara otomatis dengan menyetakan algoritma Naïve Bayes Classifier didalamnya. Sistem yang berbasis web tersebut dihosting agar seluruh mahasiswa UDINUS dapat mengakses sekaligus mencoba program tersebut. Berikut adalah hasil data uji yang diperoleh:

Tabel 2. Data Uji

NO	NIM	Dokumen
1	D22.2015.01612	membutuhkan pembelajaran yang lebih seru tidak monoton contohnya pembelajaran diluar kampus
2	D22.2015.01622	LCD Monitor ada beberapa kelas yang sulit konek dengan laptop
.	.	.
.	.	.
118	B11.2014.03365	Panas, kurang pohon.
119	B11.2014.03365	Liftnya kenapa sering rusak ya?

3.2 Preprocessing

Preprocessing dilakukan pada data latih dan data uji guna merubah teks menjadi *term index*. Sehingga sebuah *set term index* yang bisa mewakili dokumen dan lebih terstruktur. Pada proses ini *preposeccing* menggunakan *liblary* sastraawi. *Liblary* tersebut didapat dengan menginstal *composer* yang sebagai *vendor* dimana didalamnya terdapat sastraawi. Hal ini mempermudah dalam *mempreprocessing* dokumen dengan mudah dan cepat.

Sebelum *preprocessing* : “untuk fasilitas tolong untuk RUANG GALERI DIPERBESAR LAGI SEHINGGA TIDAK MEMBATASI KREATIVITAS MAHASISWA”

Sesudah *preprocessing* dengan *library* sastrawi: “untuk fasilitas tolong untuk ruang galeri besar lagi sehingga tidak batas kreativitas mahasiswa”

3.3 Hasil Naïve Bayes

Pada proses perhitungan ini, metode Naïve Bayes melibatkan data latih dan data uji menggunakan rumus sebagai berikut:

$$P(w_i|C) = \frac{\text{count}(w_i, C) + 1}{\text{count}(C) + |V|}$$

Algoritma ini terdapat proses menghitung jumlah persamaan kata antara data latih dan data uji. Dari hasil perhitungan tersebut menghasilkan nilai probabilitas setiap katagori. Nilai maksimal yang diperoleh merupakan hasil kategori yang diperoleh atau dideteksi. Berikut ini tabel hasil nilai probabilitas dan deteksi katagori dari dokumen uji yang telah dilakukan:

Tabel 3. Hasil Pengujian

No	Dokumen	Probabilitas				Hasil (max)
		SDP	DSN	KDSA	SDM	
1	Mudah input krs tp kadang kehabisan kelas	1,77E-09	6,27E-10	2,50E-07	1,04E-12	KDSA
2	Server sering down	5,09E+02	2,54E+02	5,69E+03	4,30E+02	KDSA
3	kadang materi yang diberikan kurang tersampaikan dengan baik,semoga kedepannya bisa menjelaskan dengan lebih baik agar mahasiswa mengerti dan paham	2,19E-42	1,54E-33	1,03E-39	3,96E-40	DSN
.
.
197	jam sholat dan jam kuliah sangat mepet, apalagi sholat isya	1,07E-19	2,25E-19	1,84E-19	1,36E-19	KDSA
199	Kalau bisa, lift di gedung c diperbaiki lagi. Soalnya sering macet, dan juga diberi pewangi. Karena tempatnya lembab, jadi mudah bau. Terimakasih	2,13E-48	7,80E-54	2,32E-53	3,79E-50	SDP

Keterangan:

SDP = Sarana dan Prasarana

DSN = Dosen

KDSA = Kepegawaian dan Sistem Akademik

SDM = Saran dan Masukan

3.4 Perhitungan Akurasi (*Confusion Matrix*)

Berdasarkan data yang didapat, perhitungan akurasi ini menggunakan data latih sebanyak 1000 dokumen. Berikut hasil perhitungannya:

Tabel 4. Perhitungan Confussion Matrix

	Aktual SDP	Aktual DSN	Aktual KDSA	Aktual SDM
Prediksi SDP	188	57	2	3
Prediksi DSN	1	249	0	0
Prediksi KDSA	3	17	226	4
Prediksi SDM	5	5	1	239

Dari tabel diatas menunjukkan tidak semua dokumen yang diprediksi sesuai dengan hasil aktual. Dengan tabel confusion matrik dari 1000 dokumen data latih menghasilkan masing-masing 188 “Sarana dan Prasarana”, 249 “Dosen”, 226 “kepegawaian dan sistem akademik” serta 239 “Saran dan Masukan” yang sesuai antara prediksi dan aktualnya. Maka dari itu nilai akurasi yang didapatkan sebagai berikut:

$$\text{akurasi} = \frac{188 + 249 + 226 + 239}{1000} \times 100\% = 90,2 \%$$

Akurasi yang dihasilkan sebesar 90,2%. Menurut Triowali jika memiliki tingkat nilai akurasi sebesar 0.90-1,00 dapat dikatakan excellent classification [7]. Sehingga kesimpulannya data latih ini memiliki akurasi yang excellent classification atau sangat baik.

4. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan pada tugas akhir ini dapat disimpulkan bahwa:

1. Metode *Naïve Bayes Classifier* dapat diimplementasikan pada pengkatagorian aspirasi mahasiswa UDINUS.
2. Menggunakan data latih sebanyak 1000 dokumen yang masing-masing katagori terdiri dari 250 dokumen menghasilkan keakurasian sebesar 90,2% yang dapat disebut dengan nilai diagnose yang *excellent classification* atau sangat baik dengan teknik pengujian *Confusion Matrix*.
3. Sistem ini dapat mengklasifikasikan aspirasi mahasiswa secara otomatis dan sistematis. Karena disediakan informasi jumlah data aspirasi yang disampaikan setiap katagori. Adanya hal tersebut mempermudah pihak DPM mengelola aspirasi mahasiswa.
4. Pada Data Laporan disediakan informasi *chart* jumlah data aspirasi setiap katagori pada masing-masing fakultas. Dengan maksud untuk menganalisa dan membandingkan katagori mana yang memiliki jumlah aspirasi terbanyak disetiap fakultas.

5. SARAN

Setelah melakukan penelitian ini, ditemukan beberapa hal ada yang kurang sempurna karena keterbatasan waktu dan kemampuan. Maka saran yang sesuai untuk kesempurnaan penelitian ini akan dijelaskan sebagai berikut:

1. Mengklasifikasi aspirasi mahasiswa menggunakan metode lain yang memiliki keakurasian yang lebih tinggi.

2. Semakin banyak data latih yang digunakan, kemungkinan akurasi yang diperoleh semakin tinggi.
3. Karena sebelumnya katagori telah ditentukan oleh pihak pengelola, dan menurut analisa ada kemiripan dokumen antara kategori satu dengan lainnya. Maka dari itu perlu juga dilakukan *clustering* data latih dengan metode tertentu agar mengetahui katagori yang pasti.
4. Pada sistem terdapat sub katagori yang masih diisi mahasiswa dengan manual dan sub kategori yang disediakan ditentukan sendiri oleh peneliti berdasarkan analisa lapangan. Maka dari itu dapat menggunakan metode *clustering* dan klasifikasi untuk penelitian kedepannya.

UCAPAN TERIMA KASIH

Dengan memanjatkan puji syukur kehadirat Allah SWT. Tuhan yang Maha Pengasih dan Maha Penyayang yang telah melimpahkan segala rahmat, hidayah serta inayah-Nya kepada penulis sehingga laporan tugas akhir dengan judul "**KLASIFIKASI ASPIRASI MAHASISWA DENGAN NAIVE BAYES CLASSIFIER**" dapat penulis selesaikan dengan rencana dukungan dari berbagai pihak yang tidak ternilai besarnya. Oleh karena itu penulis menyampaikan terimakasih kepada:

1. Prof. Dr. Ir Edi Noersasongko, M.Kom, selaku Rektor Universitas Dian Nuswantoro Semarang.
2. Dr. Drs. Abdul Syukur, MM, selaku Dekan Fasilkom Universitas Dian Nuswantoro.
3. Dr. Heru Agus Santosa, Ph.D, selaku Ka.Progdi Teknik Informatika-S1.
4. Bapak Ifan Rizqa, M.Kom selaku pembimbing tugas akhir yang memberikan ide penelitian, memberikan informasi referensi yang penulis butuhkan dan bimbingan yang berkaitan dengan penelitian penulis.
5. Dosen-dosen pengampu di Fakultas Ilmu Komputer Teknik Informatika Universitas Dian Nuswantoro Semarang yang telah memberikan ilmu dan pengetahuannya masing-masing.
6. Orang tua penulis tercinta serta keluarga besar yang telah memberikan motivasi penulis dalam menyelesaikan Tugas Akhir ini.
7. Ketua DPM setiap Fakultas tahun periode 2017-2018 yang mempercayai saya untuk menggunakan data aspirasi mahasiswa UDINUS dalam penelitian ini.
8. Keluarga besar HMTI yang telah membantu memotivasi dan memberikan semangat hingga terselesaikannya Tugas Akhir ini.
9. Teman-teman kuliah saya yang telah membantu memberikan masukan dalam menyelesaikan program untuk menyelesaikan tugas akhir ini.

Semoga Tuhan Yang Maha Esa memberikan balasan yang lebih besar kepada beliau-beliau, dan pada akhirnya penulis berharap bahwa penulisan laporan tugas akhir ini dapat bermanfaat dan berguna sebagaimana fungsinya.

DAFTAR PUSTAKA

- [1] "Pengerian Menurut Para Ahli," 2016 April 9. [Online]. Available: <http://www.pengertianmenurutparaahli.net/pengertian-aspirasi/>. [Accessed 17 Januari 2018].
- [2] ORMAWA, PANDUAN ORGANISASI MAHASISWA (ORMAWA) UNIVERSITAS DIAN NUSWANTORO, Semarang, 2015.
- [3] A. Hamzah, "SENTIMENT ANALYSIS UNTUK MEMANFAATKAN SARAN KUESIONER DALAM," *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST)*, 2014.
- [4] DPM, "Alur Penyampaian Aspirasi Keluarga Mahasiswa Telkom University," 18 Mei 2017. [Online]. Available: <http://dpm.orgs.telkomuniversity.ac.id/dpm-kema/alur-penyampaian-aspirasi-keluarga-mahasiswa-telkom-university/>. [Accessed 17 January 2018].
- [5] S. F. Rodiyansyah, "Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification," *IJCCS (Indonesia Journal of Computing and Cybernetics Systems)*, vol. 7, p. 1, 2013.

- [6] D. S. Pamungkas, "ANALISIS SENTIMENT PADA SOSIAL MEDIA TWITTER MENGGUNAKAN NAIVE BAYES CLASSIFIER TERHADAP KATA KUNCI "KURIKULUM 2013"," p. 1, 2014.
- [7] T. Rosandy, "PERBANDINGAN METODE NAIVE BAYES CLASSIFIER DENGAN," Jurnal TIM Darmajaya, vol. 02, p. 58, 2016 .