



STMIK
ATMA LUHUR

E-ISSN : 2581-0588

P-ISSN : 2301-7988

JURNAL SISFOKOM

(Sistem Informasi dan Komputer)

Volume 08 - No. 02 - September 2019

JURNAL SISFOKOM

(SISTEM INFORMASI DAN KOMPUTER)

Jurnal Sisfokom merupakan singkatan dari Jurnal Sistem Informasi dan Komputer. Jurnal ini merupakan kolaborasi antara sivitas akademika STMIK Atma Luhur dengan perguruan tinggi maupun universitas di Indonesia. Jurnal ini berisi artikel ilmiah dari peneliti, akademisi, serta para pemerhati system informasi dan teknik informatika. Jurnal Sisfokom diterbitkan 2 kali dalam setahun yaitu pada bulan Maret dan September. Jurnal ini menyajikan makalah dalam bidang ilmu sistem informasi dan komputer.

Penanggung Jawab

Ketua LPPM STMIK Atma Luhur

Ketua Dewan Penyunting

Fransiskus Panca Juniawan

Penyunting Pelaksana

Untung Rahardja

Henderi

Sujono

Dwi Yuny Sylfania

Eza Budi Perkasa

Rahmat Sulaiman

Mitra Bestari

Prof. Dr. Ir. Joko Lianto Buliali, M.Sc (Institut Teknologi Sepuluh November, Indonesia)

Dr. Ir. Djoko Soetarno, D.E.A (Universitas Bina Nusantara, Indonesia)

Dr. Setiawan Hadi, M.Sc. CS. (Universitas Padjajaran, Indonesia)

Dr. Indra Budi, S.Kom., M.Kom (Universitas Indonesia, Indonesia)

A'ang Subyakto, Ph.D (UIN Syarif Hidayatullah Jakarta, Indonesia)

Dr. Henderi, S.Kom, M.Kom (STMIK Raharja, Indonesia)

Novan Wijaya, S.Kom., M.Kom (AMIK MDP, Indonesia)

Silvester Dian Handy Permana, S.T., M.T.I. (Universitas Trilogi, Indonesia)

Molavi Arman, S.Kom., M.Kom (AMIK MDP, Indonesia)

Yustina Rada, S.Kom., MT (Universitas Kristen Wira Wacana Sumba, Indonesia)

Agus Perdana Windarto, S.Kom., M.Kom (STIKOM Tunas Bangsa, Indonesia)

P-ISSN: 2301-7988

E-ISSN: 2581-0588

Anjar Wanto, S.Kom., M.Kom (STIKOM Tunas Bangsa, Indonesia)

Usman Ependi, S.Kom., M.Kom (Universitas Bina Dharma, Indonesia)

Harrizki Arie Pradana, S.Kom, MT (STMIK Atma Luhur, Indonesia)

Harma Oktafia Lingga Wijaya, S.Kom., M.Kom (Universitas Bina Insan, Indonesia)

Muhammad Nur Faiz (Politeknik Negeri Cilacap, Indonesia)

SEKRETARIAT

LPPM STMIK Atma Luhur

Jl. Jend. Sudirman, Selindung Baru, Pangkalpinang

Kepulauan Bangka Belitung - Indonesia

Telp. (0717) 433 506 Fax. (0717) 433 506

e-mail : sisfokom@atmaluhur.ac.id - lppm@atmaluhur.ac.id

PENGANTAR REDAKSI

Jurnal Sisfokom (Sistem Informasi dan Komputer) adalah jurnal yang dikelola dan diterbitkan oleh LPPM STMIK Atma Luhur Pangkalpinang.

Jurnal Sisfokom (Sistem Informasi dan Komputer) Volume 08. Nomor 02 – September 2019 ini merupakan kolaborasi antara sivitas STMIK Atma Luhur dengan perguruan tinggi maupun universitas di seluruh Indonesia.

Redaksi mengucapkan terima kasih atas partisipasi dan kerja sama rekan – rekan dosen, sehingga Jurnal Sisfokom (Sistem Informasi dan Komputer) Vol.08 No.02 – September 2019 ini dapat terbit sesuai dengan yang telah kami rencanakan.

Selain itu sejumlah pakar dari dalam maupun dari luar STMIK Atma Luhur telah memberikan kontribusi yang sangat berharga dalam menilai dan memberi koreksi perbaikan bagi makalah yang dimuat. Oleh karena itu redaksi mengucapkan banyak terima kasih kepada para pakar tersebut.

Pada kesempatan ini, redaksi juga mengundang dan membuka kesempatan seluas – luasnya bagi para peneliti, rekan – rekan dosen, dan pengamat / pemerhati Sistem Informasi maupun Teknik Informatika untuk ikut serta mempublikasikan hasil penelitiannya melalui jurnal ini.

Akhirnya, redaksi berharap makalah – makalah yang diterbitkan dalam jurnal ini dapat memberikan manfaat bagi seluruh elemen sivitas akademika di STMIK Atma Luhur khususnya, dan bagi perkembangan ilmu pengetahuan dan teknologi informasi pada umumnya.

Redaksi,

DAFTAR ISI

HALAMAN DEPAN.....	i
PENGANTAR REDAKSI.....	iii
DAFTAR ISI.....	iv
EVALUASI KEBERHASILAN APLIKASI QJUE MENGGUNAKAN MODEL ITPOSMO	98-104
<i>Kresna Ramanda, Muhammad Hilman Fakriza, Nico Dias Palasara</i>	
ANALISA DAN PERANCANGAN PELAPORAN HASIL KERJA DENGAN REMINDER SISTEM (STUDI KASUS PT INDOSAT)	105-115
<i>Putri Kusriyanti, Tarimantan Sanberto Saragih</i>	
PERBANDINGAN PERFORMANSI SINGLE WEB SERVER DAN MULTI WEB SERVER DENGAN METODE PAIRED SAMPLE T TEST.....	116-123
<i>Muhammad Ravis, Gian Muhammad, Molavi Arman</i>	
IMPLEMENTASI NAÏVE BAYESSIAN DENGAN LAPLACIAN SMOOTHING UNTUK PEMINATAN DAN LINTAS MINAT SISWA SMAN 5 PAMEKASAN	124-129
<i>Indah Listiowarni, Nilam Ramadhani</i>	
ANALISIS CLUSTERING K-MEANS PADA DATA INFORMASI KEMISKINAN DI JAWA BARAT TAHUN 2018	130-140
<i>Nugroho Irawan Febianto, Nico Dias Palasara</i>	
RANCANG BANGUN SISTEM INFORMASI MANAJEMEN PENGADUAN MASYARAKAT DAN MONITORING AKADEMIK PERGURUAN TINGGI	141-148
<i>Frits Gerit John Rupilele, Aram Palilu</i>	
RANCANGAN SISTEM INFORMASI PENGAJUAN DAN PELAPORAN TUNJANGAN KINERJA KEMENTERIAN KEUANGAN MENGGUNAKAN METODE PROTOTYPE	149-158
<i>Akhmad Syarifudin, Nur Ani</i>	
IMPLEMENTASI METODE EXTREME PROGRAMMING DALAM PENGEMBANGAN SISTEM INFORMASI IZIN PRODUK MAKANAN	159-164
<i>Fatoni, Dedi Irawan</i>	

PROTOTYPE E-KATALOG DAN PEMINJAMAN BUKU PERPUSTAKAAN BERBASIS MOBILE	165-171
<i>Vivi Sahfitri</i>	
REMINDER SYSTEM DAN MONITORING PROYEK UNTUK PENILAIAN KINERJA KARYAWAN BERBASIS WEB	172-183
<i>Nikko Putra Riyanto, Tarimantan Sanberto Saragih</i>	
PENERAPAN ALGORITMA SIMPLE ADDITIVE WEIGHTING (SAW) DALAM SISTEM PENDUKUNG KEPUTUSAN KELANJUTAN PROYEK PADA PT XYZ	184-190
<i>Arief Amarullah, Tarimantan Sanberto Saragih</i>	
APLIKASI PEMINJAMAN RUANGAN RAPAT KANTOR GUBERNUR PROVINSI KEPULAUAN BANGKA BELITUNG BERBASIS ANDROID	191-198
<i>Rendy Rian Chrisna Putra</i>	
ANALISIS DAN PENGEMBANGAN APLIKASI SIJAPTI KANTOR KOMISI APARATUR SIPIL NEGARA	199-206
<i>A.Yuni Lusfiani, Rinto Priambodo</i>	
SISTEM INFORMASI PERPUSTAKAAN MENGGUNAKAN METODE WATERFALL.....	207-214
<i>Mely Mailasari, Erma Delima Sikumbang</i>	

Implementasi Naïve Bayessian dengan Laplacian Smoothing untuk Peminatan dan Lintas Minat Siswa SMAN 5 Pamekasan

Indah Listiowarni^[1], Nilam Ramadhani^[2]

Teknik Informatika, Universitas Madura^{[1], [2]}

Jl. Panglegur KM 3.5, Pamekasan, Jawa Timur

indah@unira.ac.id^[1], nilam_ramadhani@yahoo.com^[2]

Abstrak— Kurikulum 2013 memiliki beberapa perubahan dasar dari kurikulum sebelumnya, salah satunya adalah penyaluran dan penempatan siswa pada program peminatan. Setelah dilakukan klasifikasi peminatan, siswa akan diklasifikasikan lagi menggunakan nilai tes, yang disebut sebagai Lintas Minat Siswa. Penelitian ini berkonsentrasi untuk menerapkan metode Naive Bayessian pada sebuah sistem untuk menanggulangi permasalahan rumitnya proses klasifikasi dua tingkatan dan banyaknya data setiap tahunnya. Naive Bayes merupakan salah satu metode *machine learning* yang menggunakan perhitungan probabilitas, dan menggunakan laplacian *smoothing* untuk menghindari hasil akhir bernilai 0. Nilai perhitungan accuracy dan error rate pada 720 data training dengan pengambilan 5 kali jumlah data *testing* yang berbeda menggunakan naive bayessian dan laplacian *smoothing*, didapat nilai *accuracy* : 92,11% dan nilai *error rate* : 7,02%

Kata Kunci—*naive bayes, peminatan, lintas minat siswa, laplacian smoothing, data mining, klasifikasi*

I. PENDAHULUAN

Kurikulum 2013 merupakan kurikulum baru yang dijadikan acuan oleh beberapa SMA/MA/SMK atau sekolah yang sederajat, termasuk SMAN 5 Pamekasan. Kurikulum 2013 memiliki beberapa perubahan dasar dari kurikulum sebelumnya, salah satunya adalah peyaluran dan penempatan siswa pada program peminatan. Peminatan siswa adalah sebuah proses klasifikasi siswa pada sebuah bidang mata pelajaran sesuai dengan minat dan kemampuan siswa itu sendiri. Program peminatan siswa tersebut tidak lagi dilakukan saat siswa menginjak kelas XI melainkan mulai diklasifikasikan sejak siswa menjadi siswa baru (kelas X) di SMAN 5 Pamekasan, sehingga siswa diharapkan mampu menguasai bidang yang diinginkannya sejak mulai masuk SMA. Setelah siswa diklasifikasikan sesuai minat dan kemampuannya, semua siswa diberi kesempatan untuk mempelajari mata pelajaran di luar bidang minatnya yang diklasifikasikan lagi menggunakan nilai tes, yang disebut sebagai Lintas Minat Siswa yang akan dilakukan pada tahun pertama semester kedua, setelah Peminatan Siswa[1].

Peminatan dan Lintas Minat Siswa SMAN 5 Pamekasan diklasifikasi oleh tim BK yang terdiri dari 3 orang guru dan menangani lebih dari 300 orang siswa setiap tahunnya, jumlah data yang lumayan banyak dan proses klasifikasi yang terbilang rumit, akan banyak memakan banyak waktu dan tenaga. Sehingga dibutuhkan sebuah sistem yang bisa membantu tim guru BK (Bimbingan Konseling) dalam hal klasifikasi data nilai siswa, dan keakuratan data pada hasil akhir klasifikasi Peminatan dan Lintas Siswa.

Proses klasifikasi pada sebuah data tidak luput dengan penggunaan sebuah metode. Naive Bayes *classifier* merupakan metode yang paling banyak digunakan untuk proses klasifikasi. Naive Bayes *classifier* menggunakan teorema Bayes yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Ciri utama dari naive bayes *classifier* adalah asumsi yang kuat akan independensi dari masing-masing kondisi, artinya naive bayes *classifier* memiliki variabel yang independen. Dengan menerapkan Naive Bayes pada sebuah data terkadang juga menyebabkan *missclassification* jika data training hanya sedikit atau data-nya tidak kaya sehingga data testing tidak ditemukan pada data *training*, dan menyebabkan hasil probabilitas bernilai 0 (zero) dan menyebabkan *error* pada proses klasifikasi.

Kekurangan naive bayes *classifier* dapat diminimalisir dengan menggunakan metode *smoothing*, selain itu beberapa penelitian sebelumnya membuktikan bahwa metode *smoothing* dapat meningkatkan performance dari naive bayes *classifier*. Pada penelitian ini digunakan metode *smoothing* yaitu laplacian *smoothing*, atau yang biasa disebut *add-one smoothing* [2].

Penelitian yang berfokus pada klasifikasi peminatan pada siswa tingkat SMA, juga pernah dilakukan dalam penelitian lain, yaitu klasifikasi penjurusan siswa pada SMAN 1 Pontianak dengan menggunakan metode Algoritma C4.5[3], selanjutnya pada SMAN 2 Semarang dengan membandingkan Metode Regresi Logistik Biner dan *Multivariate Adaptive Regression Spline* [4]

II. METODE PENELITIAN

Metode penelitian digunakan untuk menjelaskan dan memberikan gambaran tentang penelitian yang akan digunakan, dataset yang akan digunakan, pembagian class pada setiap klasifikasi (peminatan dan lintas minat), alur sistem, metode yang digunakan dan pembahasan tahapan pre-processing yang dilakukan.

A. Dataset

Data yang digunakan dalam penelitian ini merupakan sebuah kesatuan data yang didapat dari instansi terkait yaitu SMAN 5 Pamekasan yang dijadikan tempat penelitian dan menambang data. Data yang diperlukan untuk proses penggalian data klasifikasi peminatan dan lintas minat siswa adalah berupa data siswa yang terdiri dari berbagai atribut yaitu nama siswa, nilai raport, nilai UNBK, IQ siswa, dan nilai tes pendukung lainnya.

Kumpulan data yang didapat akan digunakan untuk proses pengklasifikasian data siswa untuk mengetahui apakah siswa A diklasifikasikan sebagai siswa yang dipetakan dalam peminatan kelas IPA atau IPS, dan yang selanjutnya akan diklasifikasikan kembali menjadi kelas IPA ekonomi dan IPA geografi(kebumihan), sedangkan pada Lintas Minat IPS akan diklasifikasikan menjadi kelas IPS Kimia dan kelas IPS Fisika. Pada penelitian ini digunakan data siswa pada tahun 2018 untuk data testing dengan jumlah 180 data, dan data training diambil dari data siswa selama 4 tahun, mulai dari tahun 2014 hingga tahun 2017 dengan jumlah total 720 data siswa

B. Pre-Processing

Sebelum melakukan proses klasifikasi peminatan dan lintas minat siswa dengan menggunakan *naive bayesian* dengan *laplace smoothing*, perlu dilakukan *pre-processing* data untuk menyiapkan data[5]. Tahapan *pre-processing* yang dilakukan pada penelitian ini terdiri dari 3 tahapan pada setiap dataset yaitu data cleaning, data selection dan terakhir data *transformation*, dijelaskan pada gambar 1.



Gambar 1. Alur Pre-processing data

Data *cleaning* dilakukan pada *record* data yang tidak sesuai, seperti penghapusan karakter yang tidak diperlukan, mengisi data yang kosong (penyesuaian isi) dan pembulatan nilai pada atribut rata-rata rapor. Data *selection* dilakukan untuk memilih atribut yang akan digunakan pada proses selanjutnya dan membuang yang tidak diperlukan, seperti pada dataset peminatan siswa, atribut yang dihapus adalah atribut nama, jenis kelamin dan atribut rekomendasi BK, dan atribut pada dataset lintas minat siswa juga mengalami penghapusan atribut nama. Data *transformation* yaitu tahapan *pre-processing* untuk mengubah isi dan melakukan perubahan atribut, perubahan pada *dataset* setelah dilakukan *transformation* adalah sebagai berikut :

Peminatan siswa

Data diubah atau digabung ke dalam *format* yang sesuai untuk diproses dalam *data mining*, Objek pada *dataset* yang awalnya berupa non-angka, akan dinotasikan menjadi angka, dalam hal ini akan terjadi pada atribut Prestasi Akademik. Tabel 1 menjelaskan bagaimana seharusnya data tersebut bertransformasi dari non-angka menjadi angka

Tabel 1. Aturan Kepemilikan Sertifikat Akademik

Range Nilai Tes	Keterangan
1	Ada
0	Tidak Ada

Selain itu, data set juga akan mengalami perubahan pada beberapa atribut, seperti pada satuan nilai pada rata-rata nilai raport SMP (Sekolah Menengah Pertama) akan diubah menjadi atribut lain mengingat syarat peminatan siswa yang digunakan oleh SMA Negeri 5 Pamekasan.

MIPA : atribut gabungan dari rata-rata mata pelajaran Matematika dan IPA

$$\frac{\text{Matematika} + \text{IPA}}{2} = \text{MIPA} \quad (1)$$

MIPS : atribut gabungan dari rata-rata mata pelajaran Matematika dan IPS

$$\frac{\text{Matematika} + \text{IPS}}{2} = \text{MIPS} \quad (2)$$

Rata-rata : atribut gabungan dari rata-rata mata pelajaran Matematika, IPA, IPS, Bahasa Indonesia, dan Bahasa Inggris

$$\frac{\text{Mat} + \text{IPA} + \text{IPS} + \text{BIN} + \text{BIG}}{5} = \text{rata-rata} \quad (3)$$

Rata-rata UNBK : atribut gabungan dari semua mata pelajaran UNBK, yaitu Matematika, IPA, Bahasa Indonesia, dan Bahasa Inggris

$$\frac{\text{Mat} + \text{IPA} + \text{BIN} + \text{BIG}}{4} = \text{rata UNBK} \quad (4)$$

Sehingga, dengan menggunakan persamaan (1)-(4) didapatkan data hasil transformasi/pengubahan isi atribut dari dataset sebelumnya. Persamaan (1)-(4) digunakan untuk mengubah dan menyesuaikan nilai pada tahap klasifikasi peminatan siswa, dimana MIPA adalah Matematika-IPA, MIPS adalah Matematika-IPS, IPA adalah Ilmu Pengetahuan Alam, IPS adalah Ilmu Pengetahuan Sosial, BIN adalah Bahasa Indonesia, BIG adalah Bahasa Inggris dan UNBK adalah Ujian Nasional Berbasis Komputer, hasil *data transformation* dapat terlihat seperti pada tabel 2.

Tabel 2. Hasil Data Transformation Peminatan Siswa (Data Training)

NISN	MI PA	MI PS	Rata - Rata	Rata-rata UN BK	IQ	prestasi	Kelas
0007203313	90	91	88	84	110	0	IPA
0001685099	89	90	87	91	105	1	IPA
9986741555	88	86	86	87	100	1	IPA

Lintas minat siswa

Untuk data set Lintas Minat juga perlu melewati proses *transformation* atau perubahan atribut sebelum masuk ke proses selanjutnya,

MIPA : atribut gabungan dari rata-rata raport SMP pada mata pelajaran Matematika dan IPA, atribut MIPA digunakan untuk siswa yang mengikuti lintas minat IPA, untuk me-transformasi nilai Matematika dan IPA menjadi MIPA dapat menggunakan persamaan (5).

$$\frac{\text{Matematika} + \text{IPA}}{2} = \text{MIPA} \quad (5)$$

MIPS : atribut gabungan dari rata-rata raport SMP pada mata pelajaran Matematika dan IPS, untuk me-transformasi nilai Matematika dan IPS menjadi MIPS dapat menggunakan persamaan (6).

$$\frac{\text{Matematika} + \text{IPS}}{2} = \text{MIPS} \quad (6)$$

Mafis : atribut gabungan (rata-rata) nilai raport semester I Matematika, fisika, dan biologi, untuk me-transformasi nilai Matematika, Biologi dan Fisika menjadi Mafis dapat menggunakan persamaan (7).

$$\frac{\text{Matematika} + \text{biologi} + \text{fisika}}{3} = \text{mafis} \quad (7)$$

Makim : atribut gabungan (rata-rata) nilai raport semester I Matematika, kimia, dan biologi, , untuk me-transformasi nilai Matematika, Biologi dan Kimia menjadi Makim dapat menggunakan persamaan (8)

$$\frac{\text{Matematika} + \text{biologi} + \text{kimia}}{3} = \text{makim} \quad (8)$$

Matgeo : atribut gabungan (rata-rata) nilai raport semester I Matematika, ekonomu, dan geografi, , untuk me-transformasi nilai Matematika, Ekonomi dan Geografi menjadi Matgeo dapat menggunakan persamaan (9)

$$\frac{\text{Matematika} + \text{ekonomi} + \text{geografi}}{3} = \text{matgeo} \quad (9)$$

Matsos : atribut gabungan (rata-rata) nilai raport semester I Matematika, ekonomi, dan sosiologi, , untuk me-transformasi nilai Matematika, ekonomi dan sosiologi menjadi Matsos dapat menggunakan persamaan (10)

$$\frac{\text{Matematika} + \text{ekonomi} + \text{sosiologi}}{3} = \text{matsos} \quad (10)$$

Dengan menggunakan persamaan (5)-(10) didapatkan Hasil data transformation pada lintas minat siswa dapat dilihat pada tabel 3 untuk lintas minat IPA (Ilmu Pengetahuan Alam) dan tabel 4 untuk lintas minat IPS (Ilmu Pengetahuan Sosial)

Tabel 3. hasil data transformation lintas minat IPA

NISN	MI PA	MI PS	MAT GEO	MATS OSIO	Minat
9995967299	88	89	77	77	Sosiologi
0008465630	88	89	75	76	Geografi
9983381489	87	86	76	75	Sosiologi

Tabel 4. hasil data transformation lintas minat IPS

NISN	MI PA	MI PS	MAT KIM	MAT FIS	Minat
9995962899	75	76	71	73	Fisika
0008465631	73	74	68	69	Fisika
9983381590	73	76	73	72	Bio-kim

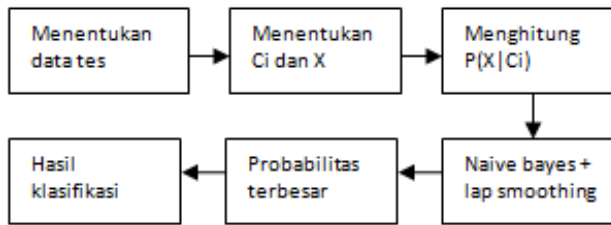
C. Metode

Proses penggalian data (data mining) adalah proses penentuan Pilihan Kelas yang tepat bagi seorang siswa dengan perhitungan didasarkan pada nilai mata pelajaran yang ada dan sesuai dengan probabilitas yang terbesar.

Pada proses penggalian data ini menggunakan Metode *Naïve Bayessian Classification* dengan *Laplacian smoothing* dengan beberapa tahapan, yaitu sebagai berikut [6]:

1. Menentukan Data Tes
2. Menentukan Nilai Ci dan Nilai X
3. Menghitung $P(X|Ci)$, $i=1,2$ untuk setiap kelas
4. Perhitungan *Naïve Bayes* dengan *Laplacian Smoothing*
5. Menentukan Nilai Probabilitas terbesar.
6. Hasil klasifikasi

Proses penggalian data dengan menggunakan Metode *Naïve Bayessian Classification* dengan *Laplacian smoothing*, dapat digambarkan dengan alur proses pada Gambar 2.



Gambar 2. proses mining

Naive bayes *classifier*, merupakan algoritma klasifikasi berdasarkan penerapan teorema bayes dengan asumsi independen yang kuat, *output* dari *classifier* berupa nilai probabilitas yang menunjukkan sebuah dokumen atau data harus diklasifikasikan atau dikelompokkan, sesuai nilai probabilitas tertinggi yang dihasilkan [7], nilai probabilitas kemunculan data dapat diketahui dengan menggunakan persamaan (11)

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (11)$$

Metode *smoothing* merupakan metode untuk menghindari hasil klasifikasi bernilai 0 dikarenakan data testing tidak ditemukan pada data training. Pada penelitian ini digunakan metode laplace *smoothing*, yaitu metode *smoothing* yang paling sederhana karena hanya menambahkan angka 1, tapi metode laplace *smoothing* memiliki performance yang cukup baik dibandingkan dengan metode *smoothing* lainnya[8], pada penerapan text atau numerik. Salah satunya pada perbandingan *smoothing* untuk pengkategorian soal ujian [2], sebagai penerapan pada data text (*text mining*).

$$P(H|X) = \frac{P(X|H)P(H)+1}{P(X)+|V|} \quad (12)$$

dimana |V| merupakan jumlah class yang digunakan, P(X) merupakan peluang data sampel yang diamati [9].

III. HASIL DAN PEMBAHASAN

Proses penggalian data dengan menggunakan metode Naive Bayesian Classification dengan Laplacian Smoothing, mengikuti proses alur pada gambar 2 dimulai dengan menentukan data tes hingga terbentuk proses klasifikasi pada setiap tujuan klasifikasi, yaitu klasifikasi peminatan siswa dan lintas minat siswa

Tahap evaluasi juga diperlukan untuk mengukur performance naive bayessian menggunakan laplacian *smoothing*, menggunakan perhitungan *accuracy* dan *error rate* [10] pada hasil klasifikasi peminatan siswa.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

$$error\ rate = \frac{FP+FN}{FP+FN+TP+TN} \quad (14)$$

Evaluasi dilakukan dengan pengambilan 5 kali jumlah data testing yang berbeda terhadap 720 data training, untuk

mengetahui pengaruh jumlah data tes, terhadap hasil klasifikasi. Tabel 5 menjelaskan hasil *accuracy* dan *error rate* untuk setiap pengambilan jumlah data tes yang berbeda, sekaligus memperlihatkan jumlah rata-rata *accuracy* dan *error rate*, dengan menggunakan persamaan (13) maka didapatkan nilai *accuracy* dan persamaan (14) digunakan untuk mengetahui nilai *error rate*

Tabel 5. hasil evaluasi pada hasil klasifikasi peminatan siswa

Jumlah data testing	Jumlah Data Training	Accuracy (%)	Error Rate (%)
35	720	90,33	6,86
60	720	96	4
80	720	90	8,5
100	720	92	8
180	720	92,22	7,77
		92,11	7,026

Setelah melakukan perhitungan data mining dan didapatkan sebuah hasil berupa klasifikasi, sebuah sistem akan diuji keakuratan data yang dihasilkan, dan kemungkinan *error rate* yang muncul. Dalam hal ini sistem akan di evaluasi menggunakan perhitungan *confusion matrix*, dengan begitu nilai akurasi dan *error* dapat diketahui.

Seperti yang terlihat pada tabel 5, Jumlah dan macam data testing yang digunakan dalam proses *mining* mempengaruhi hasil yang didapat dari perhitungan akurasi dan *error rate*, pada gambar 4.21 diketahui hasil akurasi adalah 93,33% dan *error rate* 6,66%, hasil tersebut didapat dari hasil proses *mining* yang terdiri dari 15 data testing dan 100 data training, kemudian hasil yang berbeda didapatkan ketika sistem mengeksekusi data testing dengan jumlah yang berbeda

Berikut adalah hasil tampilan implementasi *naive bayes* pada data siswa untuk mengklasifikasikan peminatan dan lintas minat siswa. Mulai dari data *training*, data *testing*, hasil hingga evaluasi.

Gambar 6. Data Training Peminatan siswa

Gambar 7. Data Testing Peminatan siswa

Gambar 6 dan gambar 7 secara berurutan merupakan tampilan halaman pengolahan data *training* dan data *testing*. Pada halaman data *training* dan data *testing*, terdapat beberapa menu yang dapat digunakan yaitu menambah dan menghapus data. Data *testing* yang digunakan pada penelitian ini (gambar 7) merupakan data *testing* dan bukan merupakan bagian dari data *training*.

Pre-processing data secara otomatis dilakukan di dalam sistem, sehingga memudahkan *user*, hingga didapatkan hasil *mining* yang dapat dilihat rinci perhitungannya menggunakan naive bayes dan laplace *smoothing* pada gambar 8.

Gambar 8. Hasil klasifikasi Peminatan siswa

Selanjutnya, tampilan pada gambar 9 akan diringkas lagi tampilan nya (menampilkan data yang dibutuhkan untuk proses *mining* selanjutnya atau kebutuhan pelaporan/berkas), agar memudahkan pengguna untuk mengetahui hasil klasifikasi yang direkomendasikan, pada gambar 8.

Gambar 9. Hasil mining

Gambar 10 hasil evaluasi

Tampilan pada gambar 10 merupakan hasil evaluasi yang berguna untuk mengetahui sejauh mana tingkat keakuratan hasil klasifikasi sesuai dengan data asli, sehingga dapat meyakinkan pihak SMAN 5 Pamekasan untuk menggunakan hasil klasifikasi.

IV. KESIMPULAN

Metode Laplace *smoothing* berhasil mengurangi kemungkinan kegagalan klasifikasi, karena bisa mencegah hasil klasifikasi bernilai 0, dari hasil perhitungan klasifikasi dengan menggunakan naive bayes dengan laplacian *smoothing* nilai perhitungan *accuracy* dan *error rate* menggunakan *confusion matrix* pada 720 data *training* dengan pengambilan 5 kali jumlah data *testing* yang berbeda, didapat nilai *accuracy* : 92,11% dan nilai *error rate* : 7,02%. Sehingga hasil klasifikasi dapat dijadikan pertimbangan penentuan kelas peminatan dan lintas minat siswa SMAN 5 Pamekasan

Penelitian selanjutnya bisa dengan menambahkan fitur perbandingan untuk metode *smoothing* lainnya selain dengan metode laplace *smoothing*.

DAFTAR PUSTAKA

- [1] Kementerian Pendidikan Dan Kebudayaan RI, "Pedoman Peminatan Peserta Didik," pp. 1–74, 2013.
- [2] I. Listiowarni and E. R. Setyaningsih, "Analisis Kinerja Smoothing pada Naive Bayes untuk Pengkategorian Soal Ujian," *J. Teknol. Manaj. Inform.*, vol. 4, 2018.
- [3] B. Novianti, T. Rismawan, and S. Bahri, "Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus: Sma Negeri 1 Pontianak)," *J. Coding, Sist. Komput. Untan*, vol. 04, no. 3, pp. 75–84, 2016.
- [4] R. Binadari, Y. Wilandari, and Suparti, "Perbandingan Metode Regresi Logistik Biner dan Multivariate Adaptive Regression Spline (MARS) pada Peminatan Jurusan SMA (Studi Kasus SMA Negeri 2 Semarang)," *J. Gaussian*, vol. 4, no. 4, pp. 987–996, 2015.
- [5] S. Garcia, "Effectiveness of Data preprocessing for data mining," *Int. J. Curr. Eng. Technol.*, vol. 4, no. 5, pp. 3480–3483, 2014.
- [6] B. Said, "DATA SMS CENTER BUPATI PAMEKASAN MENGGUNAKAN NAÏVE BAYES DAN WINNOWER," *J. LINK*, vol. 26, no. 2, pp. 1–5, 2017.
- [7] E. Prasetyowati and N. Ramadhani, "Sistem Evaluasi Dan Klasifikasi Kinerja Akademik Mahasiswa Universitas Madura Menggunakan Naive Bayes Dengan Dirichlet Smoothing," *JUTI J. Ilm. Teknol. Inf.*, vol. 16, no. 2, p. 192, 2018.

- [8] R. A. Ramadhani, F. Indriani, and D. T. Nugrahadi, "Comparison of Naïve Bayes smoothing methods for Twitter sentiment analysis," *2016 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSYS 2016*, no. October 2016, pp. 287–292, 2017.
- [9] V. Cherian and M. S. Bindu, "Heart Disease Prediction Using Naïve Bayes Algorithm and Laplace Smoothing Technique," *Int. J. Comput. Sci. Trends Technol.*, vol. 5, no. 2, pp. 68–73, 2017.
- [10] S. Rachmatullah and A. P. Wijaya, "Rekomendasi Disposisi Surat dengan Metode Naïve Bayes Pada Arsip Surat di Kantor Bakorwil Kabupaten Pamekasan," *DoubleClick J. Comput. Inf. Technol.*, vol. 2, no. 2, pp. 50–59, 2019.