

INSAND COMTECH

Information Science and Computer Technology Journal

**ALGORITMA K-NN DALAM MENGLASIFIKASIKAN DAERAH
POTENSI PEMAIN BULUTANGKIS TUNGGAL PUTRA DI
INDONESIA**

1-7

Adi Nur Rohkhim, Muhammad Hafidh Al-Ghifari, La Ode Muhammad
Zulkifli Rinjani, Natalia Damastuti

**SISTEM PAKAR MENDETEKSI KERUSAKAN SEPEDA MOTOR
MENGUNAKAN METODE NAIVE BAYES DI BENGKEL MOTOR
SMK ISLAM TANJUNG**

8-14

Nasa'i, Nindian Puspa Dewi

**ANALISA TRAFFIC JARINGAN MENGGUNAKAN SQUID PROXY
SERVER UNTUK PENINGKATAN PERFORMA AKSES INTERNET
DI UNIVERSITAS MADURA**

15-22

Khoiruddin, Ubadi

**SISTEM INFORMASI PONDOK PESANTREN DARUL ULUM
BANYUANYAR BERBASIS WEB DAN ANDROID**

23-29

Abdus Sakur, Ubaldi

**KLASIFIKASI SOAL PILIHAN GANDA BERBAHASA INDONESIA
BERDASARKAN LEVEL KOGNITIF PUSPENDIK DENGAN
SUPPORT VECTOR MACHINE**

30-37

Lilis Setyaningsih, Endang Setyati, Suhatati Tjandra

Jurnal Insand Comtech

Information Science and Computer Technology Journal

PROGRAM STUDI TEKNIK INFORMATIKA
UNIVERSITAS MADURA

Pelindung

Dekan Fakultas Teknik Universitas Madura

Penanggung Jawab

Ketua Program Studi Teknik Informatika, Fakultas Teknik
Universitas Madura

Mitra Bestari

Prof. Ir. Supeno Djanali, M.Sc., Ph.D

Dr. Arif Muntasa, S.Si., MT.

Dr. Ir. Francisca Haryanti Chandra, M.T.

Dewan Redaksi

Ketua Redaksi

Irwan Darmawan

Sekretaris Redaksi

Sandy Vikki Ariyanto

Tim Pelaksana

Sholeh Rachmatullah

Abd Wahab Syahroni

Indah Listiowarni

Alamat Redaksi:

Jurnal Insand Comtech, Program Studi Teknik Informatika, Fakultas Teknik,
Universitas Madura

Jl. Raya Panglegur Km. 3,5 Pamekasan, Kab. Pamekasan, 69317
Jawa Timur

Telp. (0324) 322231 psw. 114

Email: insandcomtech@unira.ac.id

Jurnal Insand Comtech

Information Science and Computer Technology Journal

**ALGORITMA K-NN DALAMMENGKLASIFIKASIKAN DAERAH
POTENSI PEMAIN BULUTANGKIS TUNGGAL PUTRA DI
INDONESIA**

**Adi Nur Rohkhim, Muhammad Hafidh Al-Ghifari, La Ode Muhammad
Zulkifli Rinjani, Natalia Damastuti 01 - 07**

**SISTEM PAKAR MENDETEKSI KERUSAKAN SEPEDA MOTOR
MENGUNAKAN METODE NAÏVE BAYES DI BENGKEL
MOTOR SMK ISLAM TANJUNG**

Nasa'i, Nindian Puspa Dewi08 – 14

**ANALISA TRAFFIC JARINGAN MENGGUNAKAN SQUID PROXY
SERVER UNTUK PENINGKATAN PERFORMA AKSES INTERNET
DI UNIVERSITAS MADURA**

Khoiruddin, Ubaidi 15 - 22

**SISTEM INFORMASI PONDOK PESANTREN DARUL ULUM
BANYUANYAR BERBASIS WEB DAN ANDROID**

Abdus Sakur, Ubaidi..... 23 - 29

**KLASIFIKASI SOAL PILIHAN GANDA BERBAHASA INDONESIA
BERDASARKAN LEVEL KOGNITIF PUSPENDIK DENGAN SUPPORT
VECTOR MACHINE**

Lilis Setyaningsih, Endang Setyati, Suhatati Tjandra 30 - 37

ALGORITMA K-NN DALAM MENGLASIFIKASIKAN DAERAH POTENSI PEMAIN BULUTANGKIS TUNGGAL PUTRA DI INDONESIA

Adi Nur Rohkhim ¹⁾, Muhammad Hafidh Al-Ghifari ²⁾, La Ode Muhammad Zulkifli Rinjani ³⁾, Natalia Damastuti ⁴⁾

^{1,2,3,4} Sistem Komputer, Fakultas Ilmu Komputer, Universitas Narotama Surabaya
Jln. Arief Rahman Hakim No.51 Surabaya

¹adinurrohkhim@gmail.com, ²Mh.alghifari007@gmail.com, ³oman.laode@gmail.com,

⁴natalia.damastuti@narotama.ac.id

ABSTRAK

Bulutangkis merupakan salah satu cabang olahraga yang dipertandingkan di ajang olimpiade musim panas. Tercatat pembulutangkis Indonesia telah memperoleh 7 medali emas di ajang tersebut. Hingga saat ini Indonesia belum bisa menambah medali emas satupun dari cabang olahraga lain yang dipertandingkan di Olimpiade. Tak heran bulutangkis menjadi olahraga yang sangat penting di Indonesia. Meskipun bulutangkis bukan berasal dari Indonesia, namun Indonesia telah melahirkan banyak legenda bulutangkis sejak tahun 1960-an hingga sekarang. Di era digital sekarang *sport science* telah dikembangkan di berbagai negara untuk mendukung kinerja atlet dan *official*, namun hal ini jika tidak didukung dengan regenerasi pemain muda, maka estafet prestasi akan menjadi terputus. Betapa pentingnya mempersiapkan pemain muda berpotensi untuk meneruskan tradisi prestasi di cabang bulutangkis. Indonesia merupakan salah satu negara dengan regenerasi pemain muda yang cukup lambat dibanding negara kompetitor yang lain seperti China, Korea Selatan dan Jepang. Implementasi algoritma *K-Nearst Neighbor* untuk mengklasifikasikan wilayah dengan potensi atlet bulutangkis tunggal putra di Indonesia menjadi salah satu solusi agar induk organisasi bulutangkis Indonesia lebih muda mendapatkan pemain tunggal putra berpotensi. Dengan menggunakan 1000 data peringkat nasional tunggal putra di Indonesia dan mengklasifikasikannya menjadi 3 daerah yaitu daerah Berpotensi, cukup berpotensi, dan tidak berpotensi.

Kata kunci : K-Nearst Neighbor, klasifikasi, bulutangkis, Indonesia, data mining

ABSTRACT

Badminton is one sport that is contested at the Summer Olympics. Noted Indonesian badminton player has won 7 gold medals in the event. Until now Indonesia has not been able to add any gold medal from other sports that were competed in the Olympics. No wonder badminton has become a very important sport in Indonesia. Even though badminton is not from Indonesia, but Indonesia has given birth to many badminton legends since the 1960s until now. In the digital age sport science has now been developed in various countries to support athlete and official performance, but if this is not supported by the regeneration of young players, the achievement relay will be interrupted. How important it is to prepare young players with the potential to carry on the tradition of achievement in the badminton branch. Indonesia is one of the countries with a relatively slow regeneration of young players compared to other competing countries such as China, South Korea and Japan. The implementation of the K-Nearst Neighbor algorithm to classify areas with the potential of male singles badminton athletes in Indonesia is one solution so that the parent of Indonesian badminton organizations get a potential single male player. By using 1000 national male singles ranking data in Indonesia and classifying them into 3 regions, Potential areas, enough potential, and no potential.

Keywords: K-Nearst Neighbor, classification, badminton, Indonesia, data mining.

PENDAHULUAN

Indonesia adalah salah satu negara kepulauan terbesar di dunia. Dengan jumlah penduduk lebih dari 250 juta jiwa. Indonesia seharusnya lebih mudah mencari bakat bakat muda berpotensi di bidang olahraga khususnya cabang olahraga bulutangkis sektor tunggal putra. Karena Indonesia mempunyai materi dengan jumlah penduduk yang banyak yang tersebar di seluruh kepulauan Indonesia. Kenapa harus tunggal putra ? . Sejak kesuksesan Taufik Hidayat di ajang olimpiade tahun 2004 silam dengan menyumbangkan medali emas, sampai sekarang Indonesia belum berbicara banyak di sektor ini, bahkan medali perunggu pun pembulutangkis tunggal putra Indonesia belum bisa tercapai.

Sebenarnya ada beberapa faktor yang menyebabkan prestasi atlet bulutangkis sektor tunggal putra Indonesia mengalami penurunan. Dan salah satu faktor yang berperan penting adalah masalah mempersiapkan pemain muda sebagai regenerasi. Karena Indonesia mempunyai wilayah kepulauan yang cukup luas. Sehingga wajar jika pengurus besar persatuan bulutangkis Indonesia mempunyai kendala dalam mencari bibit muda di daerah. Mulai dari masalah waktu, SDM, hingga anggaran. Meskipun memang telah di bangun klub bulutangkis besar dengan fasilitas yang sangat mempuni, namun klub ini hanya terbatas di pulau Jawa.

Berdasarkan latar belakang tersebut maka diperlukannya suatu pengklasifikasian berdasarkan daerah dengan potensi pemain bulutangkis tunggal putra di Indonesia, dengan menggunakan salah satu metode pengklasifikasian di *data mining*. Yaitu metode KNN atau *K-Nearest Neighbor*. Daerah di Indonesia akan di bagi menjadi 3 wilayah, yaitu wilayah dengan potensi pemiaan bulutangkis sektor tunggal putra, wilayah dengan cukup potensi pemiaan bulutangkis tunggal putra dan wilayah tidak berpotensi. Tujuan dari penelitian ini adalah membantu kinejra dari tim pencari bakat dari klub maupun dari pengurus besar persatuan bulutangkis seluruh Indonesia dalam mendapatkan pemain muda berpotensi disektor tunggal putra.

Bulutangkis Indonesia

Di Indonesia sendiri, sejarah bulutangkis dimulai pada tahun 1930-an. Di masa itu cabang olahraga ini berada di bawah naungan organisasi ikatan *sport Indoneisa* (ISI). Namun bulutangkis sempat dilupakan karena Indonesia menghadapi masa perang. Namun pada tahun 1974 presiden Soekarno kampanye “nation building” sehingga bulutangkis kembali berkembang. (Vannisa 2018)

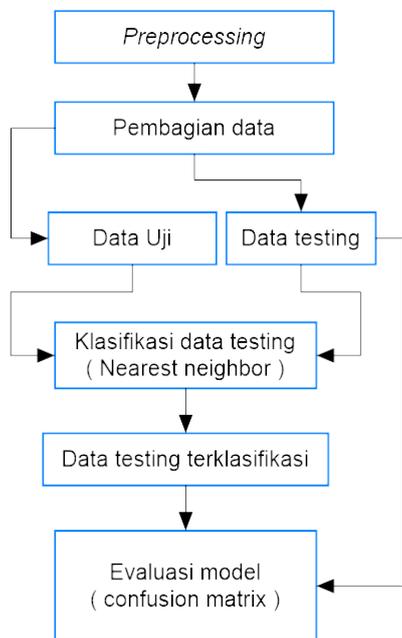
Sedangkan bulutangkis baru masuk menjadi salah satu cabang olahraga yang di pertandingan di olimpiade pada tahun 1992 di Barcelona. Kala itu bulutangkis hanya mempertandingkan 4 sektor ,yaitu sektor tunggal putra , tunggal putri , ganda putri dan ganda putra. Tim bulutangkis Indonesia sejak itu tak pernah nihil mendapatkan medali emas kecuali pada tahun 2012 di London. Tim bulutangkis Indonesia gagal menyambung tali estafet tradisi emas. Namun hal ini terbayar lunas ketika tahun 2016 di Brazil. Indonesia berhasil menyumbangkan medali emas dari sektor ganda campuran atas nama Tantowi Ahmad dan Liliana Natsir. Di final mereka mengalahkan ganda campuran asal negeri jiran Malaysia dengan 2 game langsung .(DjarumFoundation 2012)

Klasifikasi

Klasifikasi adalah suatu kinerja yang mempunyai sebuah penilaian objek data untuk memasukkannya ke dalam suatu kelas tertentu dari beberapa kelas yang telah tersedia. Dalam proses pengklasifikasian tersebut terdapat dua proses yang harus di jalankan, yaitu yang pertama dengan membangun model data untuk disimpan sebagai memori dan patokan nilai , kemudian yang kedua menggunakan model tersebut untuk melakukan pengidentifikasian atau pengklasifikasian atau prediksi pada suatu data lain, dan hasil akhirnya adalah agar data tersebut bisa diketahui di kelas mana objek data tersebut dimasukkan berdasarkan model data yang

telah disimpan dalam memori pada proses sebelumnya. (Riyan Eko Putri, Suparti 2014)

Metode klasifikasi merupakan salah satu metode dari data Mining. Data mining itu sendiri adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam database. Kinerja dari data mining yaitu memproses sebuah data dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning . Sehingga data mining berfungsi untuk mengekstrasi dan mengidentifikasi sebuah informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database yang mempunyai ukuran besar. (Sumarlin 2015)



Gambar 1. Pemodelan klasifikasi (Badu 2016)

K-Nearest Neighbor

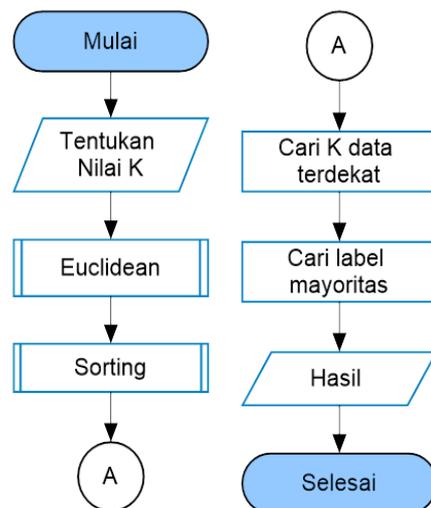
KNN adalah salah satu metode pengklasifikasian berdasarkan data *training* atau data pembelajaran dengan melihat dari posisi jarak yang paling dekat dengan objek berdasarkan nilai *k*. Metode ini bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut dan training sample.

Objek diberikan suatu titik query, kemudiantitik tersebut akan ditemukan sejumlah *K* objek atau titik training yang paling dekat dengan titik query. Nilai prediksi dari query akan tersebut akan dihitung berdasarkan klasifikasi tetangga .

Sebelum melakukan perhitungan dengan metode *K-Nearest Neighbor*, terlebih dahulu harus menentukan data latih dan data uji. Kemudian akan dilakukan proses perhitungan untuk mencari jarak menggunakan *Euclidean*. Teknik ini sangat sederhana dan mudah diimplementasikan. Dengan mengelompokkan suatu data baru berdasarkan jarak data baru itu ke beberapa data/tetangga terdekat. Pertama sebelum mencari jarak data ke tetangga adalah menentukan nilai *K* tetangga (neighbor). Lalu, untuk mendefinisikan jarak antara dua titik yaitu titik pada data training dan titik pada data testing. (Nugraha and Faraby 2018)

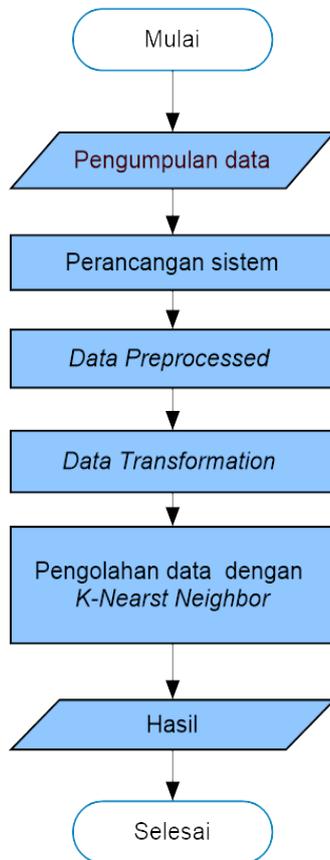
Algoritma *KNN* memiliki beberapa keunggulan yaitu kemampuan dalam mengatasi *training data* dengan banyak *noise* dan sangat efektif dalam menanagani *training data*-nya dengan data yang cukup besar. Sedangkan, kelemahan *KNN* yaitu *KNN* perlu menentukan nilai dari parameter *K* (jumlah dari tetangga terdekat). (Yustanti 2012)

Berikut flowchart metode *KNN*. (Nouvel 2015)



Gambar 2 .Flowchart KNN

METODE PENELITIAN



Gambar 3 .Flowchart metode penelitian

- **Pengumpulan data:** Pada tahap pengumpulan data, penulis menggunakan data primer dan data sekunder. Dimana data primer adalah data yang di hasilkan dari pengolahan sistem. Sedangkan data sekunder adalah data yang diperoleh dari website resmi organisasi persatuan bulutangkis seluruh Indonesia yaitu data daftar peringkat nasional bulutangkis sektor tunggal putra di Idnoneisa sebanyak 1000 data
- **Perancangan sistem:** Pada tahap perancangan sistem. Penulis melakukan perancangan pengaklasifikasian data dengan 3 keluaran yaitu daerah dengan potensi pemain bulutangkis tunggal putra, daerah dengan cukup potensi

bulutangkis, dan daerah tidak berpotensi pemain bulutangkis tunggal putra

- **Data Preprocessed :** Pada tahap ini penulis menggunakan pembersihan data di beberapa atribut, karena pada data daftar peringkat nasional pemain bulutangkis tunggal putra terdapat atribut yang masih belum lengkap (*incomplete*) seperti pada atribut "Id BWF". Atribut ID BWF adalah ID yang diberikan oleh organisasi badan bulutangkis dunia pada pemain yang telah berkarir di dunia pembulutankisan dunia, atau bisa di katakan pemain tersebut minimal pernah bermain atau berpartisipasi beberapa kali di turnamen kelas internasional sehingga mendapatkan point dan akhirnya BWF memberi ID pada pemain tersebut. Karena atribut ini tidak begitu berpengaruh pada penelitian ini. Akhirnya penulis melakukan pemersihan data. Agar data yang diolah bisa maksimal dan tidak mengalami eror
- **Transformation data:** Pada tahap ini. Penulis melakukan pengubahan tipe data pada beberapa atribut, agar pada saat data diolah dapat memenuhi asumsi asumsi yang mendasari analisa sistem. Serta penulis mengubah format file dari data yang diperoleh, yang awalnya file daftar peringkat nasional pemain tunggal putra bertipe ".exe", agar bisa diolah. Maka file tersebut di ubah formatnya menjadi "CSV"
- **Pengolahan data dengan KNN :** Pada tahap ini penulis menggunakan teknik pencarian tetangga terdekat yang umum dilakukan dengan menggunakan formula jarak *euclidean distance*. Jarak *euclidean distance* adalah formula untuk mencari jarak dengan 2 titik dalam ruang dua dimensi. Berikut rumus *euclidean*

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (1)$$

Dimana :
 $d(x_i, x_j)$ = Jarak *Euclidean*
 $(x_i), (x_j)$ = *record ke-i, record ke-j*
 (a_r) = data ke r
 i, j = 1,2,3.....
 n = dimensi objek

- **Hasil** : Hasil adalah tahapan akhir dari penelitian ini. Setelah proses pengolahan klasifikasi dengan KNN selesai. Maka akan dihasilkan sebuah informasi untuk dasar pengembangan penelitian ini.

Pada data daftar peringkat nasional pemain bulutangkis tunggal putra, terdapat beberapa atribut bertipe nominal. Atribut tersebut akan diinisialisasikan menjadi bentuk numerik. Berikut inialisasi atribut "Wilayah"

Tabel 1. Inisialisasi wilayah

Inisialisasi	Wilayah	Jumlah
1	DKI Jakarta	313
2	Jawa Barat	148
3	Jawa Tengah	146
4	Jawa Timur	116
5	Banten	71
6	D.I. Yogyakarta	26
7	Kalimantan Selatan	13
8	Bali	7
9	Sulawesi Utara	13
10	Sumatera Selatan	22
11	Kalimantan Tengah	9
12	Sulawesi Selatan	11
13	Kalimantan Barat	5
14	Riau	8
15	Kalimantan Timur	32
16	Sumatera Barat	5
17	Nusa Tenggara Barat	6
18	Bengkulu	2
19	Kepulauan 14	7
20	Jambi	2
21	Sumatera Utara	24
22	Sulawesi Tenggara	2
23	Kepulauan Bangka Belitung	1
24	Papua	1
25	Gorontalo	5
26	Maluku Utara	1
27	Nanggroe Aceh Darussalam	2

28	Lampung	2
----	---------	---

Berikut tabel penentuan kelas klasifikasi berdasarkan ranking nasional pembulutangkis tunggal putra dan poin turnamen yang di dapat oleh setiap pemain

Tabel 2. Penentuan klasifikasi

Rangking (poin turnamen)	Frekuensi	Inisialisasi
1 – 333 (73.090 – 720)	333	Berpotensi
334 – 667 (720 – 162)	334	Cukup benpotensi
668 – 998 (160 -25)	333	Tidak berpotensi

HASIL DAN PEMBAHASAN

Tools yang digunakan pada penelitian ini adalah "Weka versi 3.8.4". Weka merupakan salah tools data mining yang mampu melakukan beberapa metode pengolahan data dengan cara yang sederhana, salah satunya adalah klasifikasi menggunakan metode *K-Nearest Neighbor*. Pengujian menggunakan data awal akan di ubah dengan susunan sebagai berikut :

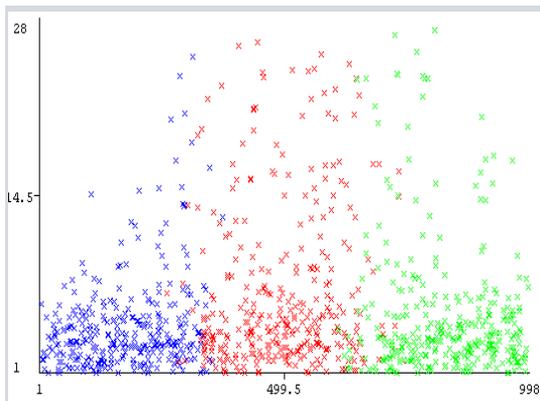
1. Jumlah data yang di gunakan adalah 1000 data yang di ambil dari web resmi organisasi persatuan bulutangkis republik Indonesia
2. Data awal mempunyai atribut sebanyak 7 atribut. Namun tidak semua atribut di gunakan untuk mengklasifikasikan. Hanya 4 atribut yang dipakai dalam penelitian ini.

Tabel 3. Data sampel daftar peringkat nasional

Rank-ing	poin	Wilayah	Klasifikasi
1	73090	1	Berpotensi
2	65680	2	Berpotensi
3	49000	2	Berpotensi

.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
998	25	2	Tidak berpotensi

Tabel 3. adalah data sampel dari data daftar peringkat pemain bulutangkis tunggal putra di Indonesia.



Gambar 4. Visualisasi data klasifikasi daftar peringkat bulutangkis tunggal putra

Pada tahap pengujian peneliti menggunakan data uji sebanyak 400 data dan menggunakan 600 data sebagai data latih, dimana masing masing mempunyai 4 atribut yaitu atribut ranking nasional, poin turnamen, kode wilayah dan klasifikasi potensi.

Tabel 4. Data sampel daftar data latih

Ranking	poin	Wilayah	Klasifikasi
1	24600	1	Berpotensi
2	21440	3	Berpotensi
3	21290	2	Berpotensi
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
600	50	2	tidak berpotensi

Tabel 4. adalah daftar data sampel data latih yang berjumlah 600 data. Data latih adalah data yang sudah ada sebelumnya berdasarkan fakta dan yang sudah terjadi. Nantinya data latih ini akan bertugas melatih data uji.

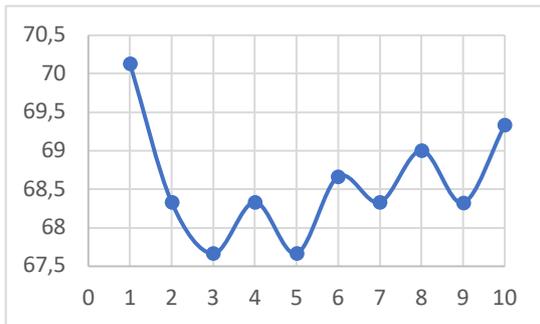
Tabel 5. Data sampel daftar data uji

Ranking	poin	Wilayah	Klasifikasi
1	16860	3	Berpotensi
2	16820	12	Cukup Berpotensi
.	.	.	.
.	.	.	.
.	.	.	.
.	.	.	.
400	45	2	tidak berpotensi

Tabel 5. adalah daftar data sampel data uji yang berjumlah 400 data. Data uji adalah data yang sudah berlabel yang tujuannya untuk menghitung nilai akurasi dari permodelan yang sudah dibentuk sebelumnya.

Tabel 6. Hasil pengujian

Kn	Akurasi (%)
K1	70.133
K2	68.333
K3	67.666
K4	68.333
K5	67.766
K6	68.666
K7	68.333
K8	69.003
K9	68.325
K10	69.333



Gambar 5. Grafik nilai akurasi

Dari hasil pengujian data testing terhadap data uji menggunakan KNN. Mulai dari K1 sampai dengan K10 didapatkan hasil. Nilai akurasi terendah terjadi pada K3 yaitu 67.666 % dan nilai akurasi tertinggi terjadi pada K1 yaitu 70.1333 %

PENUTUP

Setelah dilakukan klasifikasi daftar peringkat nasional pemain bulutangkis tunggal putra menggunakan metode KNN dapat disimpulkan bahwa bakat potensi pemain bulutangkis tunggal putra memang mayoritas terdapat di pulau Jawa. Namun hasil ini bukanlah patokan Organisasi persatuan bulutangkis Indonesia dalam mencari bibit pemain potensial tunggal putra di Indonesia.

DAFTAR PUSTAKA

- Badu, Zemi S. 2016. "PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI DANA DESA." (November).
- Djarum Foundation, Team. 2012. "Berharap Pada Mitos." 29 Juli 2012. <https://www.pbdjarum.org> (January 15, 2020).
- Nouvel, Ahmad. 2015. "Klasifikasi Kendaraan Roda Empat Berbasis Knn." 3(2): 66–69.
- Nugraha, Pratama Dwi, and Said Al Faraby. 2018. "Klasifikasi Dokumen Menggunakan Metode k -Nearest Neighbor (KNN) Dengan Information

Gain Document Classification Using k-Nearest Neighbor (k NN) Method with Information Gain." 5(1): 1541–50.

Riyan Eko Putri, Suparti, Rita Rahmawati. 2014. "Perbandingan Metode Klasifikasi Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Data Status Kerja Di Kabupaten Demak Tahun 2012." *GAUSSIAN* 3: 831–38.

Sumarlin. 2015. "Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA Dan BBM." 01: 52–62.

Vannisa. 2018. "Sejarah Bulu Tangkis." *June 24, 2018*. <https://perpustakaan.id> (January 15, 2020).

Tabel 4. Data sampel daftar data latih
2012. "Algoritma K-Nearest Neighbour Untuk Memprediksi Harga Jual Tanah." 9(1): 57–68.

Sistem Pakar Mendeteksi Kerusakan Sepeda Motor Menggunakan Metode Naïve Bayes di Bengkel Motor SMK Islam Tanjung

Nasa'i¹⁾, Nindian Puspa Dewi²⁾
^{1,2}Informatika, Teknik, Universitas Madura
^{1,2}Jln. Panglegur Km. 3,5 Pamekasan Jawa Timur Indonesia
¹nasha.lovers@gmail.com, ²nindianpd@unira.ac.id

ABSTARK

Sistem Pakar (expert system) adalah system atau program yang bertingkah laku seperti ahlinya atau pakarnya. Sistem pakar merupakan sistem yang berbasis pengetahuan yang bisa digunakan untuk membantu menyelesaikan masalah-masalah yang ada dalam dunia nyata. Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari sebuah dataset, sehingga naïve bayes dapat dijadikan metode pada aplikasi system pakar untuk mengetahui probabilitas suatu kerusakan yang mungkin terjadi pada sepeda motor. pada penelitian ini, dibuatakan sebuah aplikasi system pakar yang dapat mendeteksi kerusakan sepeda motor dan mampu memberikan solusi seperti seorang pakar dibengkel motor SMK Islam Tanjung yang dibangun menggunakan bahasa pemrograman php

Kata kunci: system pakar, naïve bayes, sepeda motor.

ABSTRACT

Expert systems are systems or programs that behave as experts. Expert systems are knowledge-based systems that can be used to help solve problems that exist in the real world. Naive Bayes is a simple probabilistic classification that calculates a set of probabilities by adding up frequencies and combinations of values from a dataset, so Naïve bayes can be used as a method in expert system applications to determine the probability of a damage that may occur on a motorcycle. in this study, an expert system application was created that could detect motorcycle damage and was able to provide solutions such as an expert at Islamic Vocational High School Tanjung motorcycle repair shop that built by using the php programming language.

Keywords: expert system, naïve bayes, motorbikes

PENDAHULUAN

Di zaman yang serba modern seperti saat ini kecepatan dan keakuratan dalam informasi sangatlah penting. Teknologi mempunyai peranan penting yang tentunya tidak terlepas kaitannya dengan Teknologi Informasi (TI). Komputer merupakan satu bagian yang paling penting dalam peningkatan Teknologi Informasi, salah satu faktornya yaitu kemampuan komputer dalam menyimpan dan mengingat informasi serta dapat dimanfaatkan semaksimal mungkin. Komputer memberikan kesimpulan atau pengambil keputusan yang kualitasnya sama dengan kemampuan seorang pakar dalam suatu bidang ilmu tertentu karena dapat menyimpan informasi aturan penalaran yang memadai. Salah satu cabang ilmu teknik informatika yang dapat mendukung hal tersebut adalah Sistem Pakar. Sistem pakar dapat diaplikasikan dalam berbagai bidang seperti dalam bidang perbengkelan yang berguna untuk mendiagnosa kerusakan berdasarkan gejala yang ditimbulkan seperti yang terjadi pada sepeda motor. Dalam hal ini penulis akan melakukan penelitian dibengkel motor SMK Islam Tanjung.

SMK Islam Tanjung merupakan sekolah menengah kejuruan yang berdiri di desa Dharma Tanjung kecamatan Camplong kabupaten Sampang pada tahun 2008 yang menjadi sekolah menengah kejuruan pertama di desa Dharma Tanjung. Sekolah ini sampai saat ini memiliki dua jurusan yaitu jurusan teknik komputer dan jaringan (TKJ) dan teknik bisnis sepeda motor (TBSM). TBSM merupakan jurusan teknik bisnis sepeda motor yang diharuskan oleh pemerintah untuk membuka bisnis secara mandiri sehingga SMK Islam Tanjung khusus jurusan teknik bisnis sepeda motor membuka bengkel motor yang diberi nama bengkel motor SMK Islam Tanjung. Usaha bengkel motor SMK Islam Tanjung ini didirikan pada pertengahan tahun 2015 yang beralamat di jalan Idaman Dh. Tanjung Camplong Sampang. Pada awalnya bengkel ini dijalankan oleh satu orang karyawan kemudian berkembang dan sekarang sudah memiliki tiga orang karyawan sebagai mekanik sepeda motor. Akan tetapi pihak bengkel motor SMK Islam Tanjung masih menggunakan cara manual dalam menganalisa kerusakan sepeda motor.

Berdasarkan latar belakang yang ada, maka dalam penelitian ini penulis membuat sistem pakar pendeteksi kerusakan sepeda motor menggunakan pemrograman berbasis web dan mysql dengan metode Naive Bayes. Diharapkan sistem dapat membantu memberikan saran dan solusi pada karyawan bengkel motor SMK Islam Tanjung dalam mendeteksi kerusakan sepeda motor pelanggannya..

TINJAUAN PUSTAKA

1. Sistem Pakar

Sistem Pakar (dalam bahasa Inggris : expert system) adalah sistem informasi yang berisi dengan pengetahuan yang berasal dari pakar. Pengetahuan di dalam sistem ini digunakan sebagai dasar oleh sistem pakar untuk menjawab pertanyaan (konsultasi).

Sistem pakar adalah suatu sistem yang dirancang untuk dapat menirukan keahlian seorang pakar dalam menjawab pertanyaan dan memecahkan suatu masalah. Sistem pakar akan memberikan pemecahan suatu masalah yang didapat dari dialog dengan pengguna. Dengan bantuan sistem pakar seorang yang bukan ahli dapat menjawab pertanyaan, menyelesaikan serta mengambil keputusan yang biasanya diambil oleh seorang pakar.

2. Naive Bayes

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya.

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. Keuntungan penggunaan Naive

Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Bayes merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayes memiliki akurasi dan kecepatan yang sangat tinggi saat diaplikasi ke dalam database dengan data yang besar. Berikut teorema bayes.

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} + \dots$$

Keterangan :	
X	: Data dengan class yang belum diketahui
H	: Hipotesis data x merupakan suatu class spesifik
P(H X)	: Probabilitas hipotesis H berdasarkan kondisi X
P(H)	: Probabilitas hipotesis H (prior probability)
P(X H)	: Probabilitas X berdasar kondisi pada hipotesis H
P(X)	: Probabilitas dari X

METODE PENELITIAN

Metode ini merupakan metode yang dipraktikkan langsung untuk mengumpulkan data yang berhubungan dengan data kerusakan dan gejala – gejala kerusakan sepeda motor di bengkel motor SMK Islam Tanjung. Data – data tersebut dikumpulkan dengan cara:

1. Penelitian Langsung

Disini penulis mengadakan observasi secara langsung terhadap obyek yang di teliti. Hal ini yang dilakukan penulis adalah mengamati langsung bagaimana kegiatan yang terjadi dilapangan yaitu di Bengkel Motor SMK Islam Tanjung. Dalam hal ini objek yang diamati adalah bagaimana caranya karyawan bengkel motor SMK Islam Tanjung dalam menganalisa masalah atau kerusakan sepeda motor.

2. Penelitian Langsung

Studi pustaka yaitu pengumpulan data melalui hasil dokumentasi, baik bahan tersebut diperoleh dari buku, jurnal, ataupun hasil browsing dari internet. Pada tahap ini penulis melakukan pencarian dan pemahaman teori-teori yang berkaitan dengan tools dan bahasa pemrograman yang

akan digunakan dan teori-teori lainnya yang dapat membantu dalam proses analisis, perancangan, sampai implementasi dari program yang akan dibuat

3. Data Kerusakan Sepeda Motor

Tabel 1 Data kerusakan Sepeda motor dan solusi

NO	Id Kerusakan	Kerusakan	Solusi
1	K1	Torak atau seher sudah mulai aus	Di colter atau diperbesar kapasitas mesinnya
2	K2	Banjir pada sistem pelumas atas	Menggati seal pompa oli atas
3	K3	Kampas kopling Aus	Mengganti Kampas kopling
4	K4	Aus Pada Sentrifugal Clutch	Mengganti Kampas Ganda
5	K5	Connecting rood atau stang seher rusak	Ganti Connecting rood
6	K6	Katup longgar	Stel ulang katup
7	K7	Seher Aus	Ganti seher
8	K8	Setelan karbu kurang pas	Stel ulang karbu
9	K9	Rantai kamrat longgar	Ganti atau stel ulang rantai

4. Data Gejala Kerusakan Sepeda Motor

Tabel 2 data gejala kerusakan sepeda motor

NO	Id gejala	Gejala
1	G1	Motor terasa bergetar lebih keras
2	G2	Suara gemeretak dibagian cvt
3	G3	Akselerasi di putaran mesin atas terasa tertahan
4	G4	Akselerasi lambat dan tertahan
5	G5	Akselerasi tersendat-sendat di putaran mesin bawah

6	G6	Muncul suara kasar
7	G7	Bunyi tek tek di rpm rendah
8	G8	Akselerasi di putaran menengah terasa tertahan
9	G9	Muncul suara berisik di rumah cvt
10	G10	Munculnya suara dengung cukup keras
11	G11	Muncul suara berdecit
12	G12	Akselerasi awal terasa seperti slip
13	G13	Bunyi tek tek terus menerus

5. Data kerusakan Sepeda Motor dan Gejalanya.

1. K1 :Roller aus
G1 :Motor terasa bergetar lebih keras
G2 :Suara gemeretak dibagian cvt
G3 :Akselerasi di putaran mesin atas terasa tertahan
2. K2 :Rumah roller aus
G1 :motor terasa bergetar lebih keras
G2 :suara gemeretak dibagian cvt
G4 :Akselerasi lambat dan tertahan
3. K3 :Mangkuk kopling aus
G1 :motor terasa bergetar lebih keras
G4 :Akselerasi lambat dan tertahan
G5 :Akselerasi tersendat-sendat di putaran mesin bawah
4. K4 :Kampas kopling tipis
G4 :Akselerasi lambat dan tertahan
G11 :Muncul suara berdecit
G5 :Akselerasi tersendat-sendat di putaran mesin bawah
5. K5 :Secondary sliding sheave aus
G4 :Motor terasa bergetar lebih keras
G6 :Muncul suara kasar
G8 :Akselerasi di putaran menengah terasa tertahan
6. K6 : V-belt mellar
G9 : Muncul suara berisik di rumah CVT
G12 : Akselerasi awal terasa seperti slip
G4 : Akselerasi lambat dan tertahan
7. K7 : Gear rasio aus
G10 : munculnya suara dengung cukup keras
G6 : muncul suara kasar
8. Kruke as terlalu keluar
G1 : Motor terasa bergetar lebih keras

- G13 : Bunyi tek Tek terus menerus
G6 : Muncul suara kasar
9. K9 : Pislet longgar
G6 : muncul suara kasar
G7 : Bunyi tek Tek di rpm rendah

6. Rule Base

1. If (G1 and G2 and G3)
Then K1
2. If (G1 and G2 and G4)
Then K2
3. If (G1 and G4 and G5)
Then K3
4. If (G4 and G5 and G11)
Then K4
5. If (G4 and G6 and G8)
Then K5
6. If (G4 and G9 and G12)
Then K6
7. If (G6 and G10)
Then K7
8. If (G1 and G6 and G13)
Then K8
9. If (G6 and G7)
Then K9

7. Perhitungan Bayes

Misalnya gejala yang tampak pada sepeda motor ada 3 yaitu :
Motor terasa bergetar lebih keras (G1)
Suara gemeretak dibagian cvt (G2)
Akselerasi diputaran mesin atas teras tertahan (G3)
Berdasarkan gejala tersebut maka dapat dihitung :
Roller Aus
Tahap pencarian (X | H)
Pada tahap ini akan mencari probabilitas roller aus.

$$= \frac{\text{jumlah kemungkinan}}{\text{jumlah kerusakan}} = \frac{1}{9} = 0.11$$

Tahap pencarian P(X)
Probabilitas pencarian gejala adalah :
Motor terasa bergetar lebih keras
 $= \frac{\text{kemungkinan}}{\text{jumlah kemungkinan kerusakan akibat gejala}} = \frac{1}{3} = 0.33$

Suara gemeretak dibagian CVT
 $= \frac{\text{jumlah kemungkinan}}{\text{jumlah kemungkinan kerusakan akibat gejala}} = \frac{1}{3} = 0.33$

Akselerasi diputaran mesin atas terasa tertahan

$$= \frac{\text{jumlah kemungkinan}}{\text{jumlah kemungkinan kerusakan akibat gejala}} = \frac{1}{3} = 0.33$$

Perhitungan nilai bayes

$$K(K1 | G1) = [K (G1 | K1) * K (K1)] / [K (G1 | K1) * K (K1) + K (G1 | K2) * K (K2) + K (G1 | K3) * K (K3) + K (G1 | K4) * K (K4) + K (G1 | K5) * K (K5) + K (G1 | K6) * K (K6) + K (G1 | K7) * K (K7) + K (G1 | K8) * K (K8) + K (G1 | K9) * K (K9)]$$

$$K(K1 | G1) = \frac{0,0363}{0,181}$$

K(K1 | G1) = 0,200 (Hasil perhitungan bayes 1 pada kerusakan K1)

$$K(K1 | G2) = [K (G2 | K1) * K (K1)] / [K (G2 | K1) * K (K1) + K (G2 | K2) * K (K2) + K (G2 | K3) * K (K3) + K (G2 | K4) * K (K4) + K (G2 | K5) * K (K5) + K (G2 | K6) * K (K6) + K (G2 | K7) * K (K7) + K (G2 | K8) * K (K8) + K (G2 | K9) * K (K9)]$$

$$K(K1 | G2) = \frac{0,0363}{0,072}$$

K(K1 | G2) = 0,50 (Hasil perhitungan bayes 2 pada kerusakan K1)

$$K(K1 | G3) = [K (G3 | K1) * K (K1)] / [K (G3 | K1) * K (K1) + K (G3 | K2) * K (K2) + K (G3 | K3) * K (K3) + K (G3 | K4) * K (K4) + K (G3 | K5) * K (K5) + K (G3 | K6) * K (K6) + K (G3 | K7) * K (K7) + K (G3 | K8) * K (K8) + K (G3 | K9) * K (K9)]$$

$$K(K1 | G3) = \frac{0,0363}{0,0363}$$

K(K1 | G3) = 1 (Hasil perhitungan bayes 3 pada kerusakan K1)

Total Bayes Ke 1
= K (K1 | G1) + K (K1 | G2) + K (K1 | G3)
= 0,200 + 0,50 + 1 = 1,704 (total bayes ke tiga pada kerusakan (K1))

Langkah diatas diulang ulang sebanyak jumlah kerusakan yang ada, sampai ditemukan persentase kerusakan seperti dibawah ini.

Hasil = Total Bayes Pertama + Total Bayes Kedua + Total Bayes Ketiga + Total Bayes Keempat + Total Bayes Kelima + Total Bayes Keenam + Total Bayes Ketujuh + Total Bayes Kedelapan + Total Bayes Kesembilan

$$= K(K1 | G1) + K(K1 | G2) + K(K1 | G3) + K(K2 | G1) + K(K2 | G2) + K(K2 | G3) + K(K3 | G1) + K(K3 | G2) + K(K3 | G3) + K(K4 | G1) + K(K4 | G2) + K(K4 | G3) + K(K5 | G1) + K(K5 | G2) + K(K5 | G3) + K(K6 | G1) + K(K6 | G2) + K(K6 | G3) + K(K7 | G1) + K(K7 | G2) + K(K7 | G3) + K(K8 | G1) + K(K8 | G2) + K(K8 | G3) + K(K9 | G1) + K(K9 | G2) + K(K9 | G3)$$

= 1,7+0,7+0,2+0+0,2+0+1.0+0,2+0 = 3
= 3 hasil total dari seluruh nilai perhitungan bayes

Maka Perhitungan Probabilitas Kerusakan Sepeda Motor Adalah :

K1 : Kerusakan pada roller
= $\frac{\text{Total Bayes Pertama}}{\text{Hasil}} \times 100$
= $\frac{1,7}{3} \times 100 = 56,66\%$

K2: Kerusakan pada rumah roller aus
= $\frac{\text{Total Bayes Kedua}}{\text{Hasil}} \times 100$
= $\frac{0,7}{3} \times 100 = 23,33\%$

K3: kerusakan pada mangkuk kopling aus
= $\frac{\text{Total Bayes Ketiga}}{\text{Hasil}} \times 100$
= $\frac{0,2}{3} \times 100 = 6,66\%$

K4: Kerusakan pada kampas kopling tipis
= $\frac{\text{Total Bayes Keempat}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

K5: Kerusakan pada secondary sleading sheave aus
= $\frac{\text{Total Bayes Kelima}}{\text{Hasil}} \times 100$
= $\frac{0,2}{3} \times 100 = 6,66\%$

K6: Kerusakan Pada V-Belt mellar
= $\frac{\text{Total Bayes Keenam}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

K7: Kerusakan pada gear rasio aus
= $\frac{\text{Total Bayes Pertama}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

K8: Kerusakan pada kuke as terlalu keluar
= $\frac{\text{Total Bayes Kedelapan}}{\text{Hasil}} \times 100$
= $\frac{0,2}{3} \times 100 = 6,66\%$

K9: Kerusakan pada pislet longgar
= $\frac{\text{Total Bayes Kesembilan}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

Dari perhiitungan diatas dapat disimpulkan bahwa kerusakan yang mungkin terjadi adalah sebagai berikut

1. Roller aus 56,66 %
2. Rumah roller aus 23,33 %
3. Mangkuk kopling aus 6,66 %
4. Kampas kopling tipis 0 %
5. Secondary sleading sheave aus 6,66 %
6. V-Belt mellar 0 %
7. Gear rasio aus 0 %
8. Kruke as terlalu keluar 6,66 %
9. Pislet longgar 0 %

Jadi berdasarkan persentase hasil perhitungan bayes diatas dapat disimpulkan bahwa kemungkinan kerusakan terjadi pada :

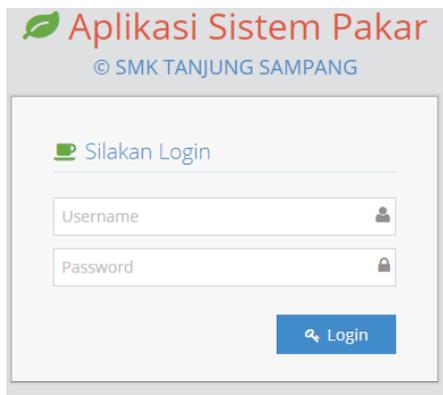
1. K1 (Roller aus) 56,66 %
2. K2 (Rumah roller aus) 23,33 %

HASIL DAN PEMBAHASAN

Berikut tampilan beranda. Pada halaman ini admin dapat login dengan mengklik tombol admin



Gambar 1 halaman beranda



Gambar 2 form login admin

Pada gambar 2 admin dapat melakukan login ke aplikasi dengan memasukkan user name dan password lalu click login, apabila login sukses maka akan diarahkan ke hakaman beranda admin seperti gambar 3 dan admin sudah dapat melakukan management seluruh data yang dibutuhkan di aplikasi system pakar ini.



Gambar 3 halaman admin

Pada gambar 1 user atau pelanggan dapat memberikan saran gejala yang tidak ada di dalam aplikasi dan pelanggan juga

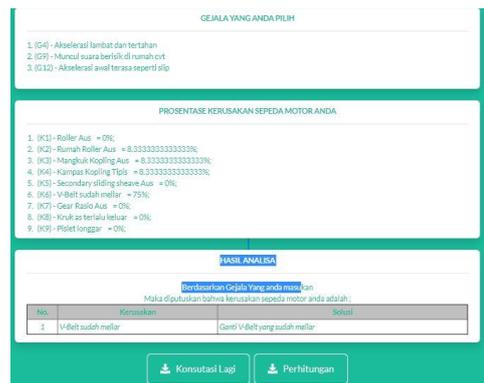
dapat melakukan konsultasi untuk mendapatkan informasi kerusakan sepeda motor dengan menekan tombol konsultasi dan akan diarah ke halaman konsultasi tetapi sebelumnya harus mengisi identitas seperti gambar 4 dibawah ini.



Gambar 4 form input identitas



Gambar 5 pertanyaan pertanyaan



Gambar 6 hasil akhir

KESIMPULAN

1. Kesimpulan

Berdasarkan hasil penelitian yang sudah dilakukan, peneliti dapat menyimpulkan bahwa system pakar yang dibuat dapat membantu karyawan bengkel motor SMK Islam tanjung dalam mendeteksi kerusakan sepeda motor pelanggan terutama pada saat jam jam sibuk.

Dengan adanya aplikasi ini siapapun bisa melakukan konsultasi meskipun ahlinya tidak ada di bengkel pada saat itu.

DAFTAR PUSTAKA

- Saleh, A, 2015. *Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*. [pdf]. ISSN: 2354-5771. Januari 2015. Universitas Potensi utama. Medan. Indonesia 2015.
- Wahyudi, S., & Ratnasari, S. 2015 . *Sistem Pakar Diagnosis Penyakit Mata Menggunakan Naive Bayes Classifier*. [pdf] *Seminar Nasional Sains dan Teknologi*. 1 Februari 2015. Universitas Trunojoyo Madura.
- Wawan, P, & Yustina. 2015. *Sistem Untuk Deteksi Kerusakan Mesin Diesel Mobil Panther Dengan Metode Naïve Bayes*. [pdf] ISSN : 2338-4018. *Jurnal TIKomSIN*.
- Rodiyansyah, S. F., & Winarko, E. 2012. *Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification*. [pdf] ISSN: 1978-1520. Januari 2013. FMIPA UGM.
- Muhammad & Ardi, 2017. *Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit*. [pdf]. ISSN 2355-5920. September 2017. *Jurnal Pseudocode*, Volume IV Nomor 2.
- Adi, Haris. *Sistem Informasi Kerusakan Laptop Menggunakan Metode Naïve Bayes*. [pdf] ISSN: 2086-9479. *Jurnal Teknologi Elektro*. September 2017. Universitas Mercu Buana.
- Honda Motor. 2014 *Pedoman Reparasi Umum*. Penerbit Service Publication Office.
- Honda Motor. 2016 *Suplemen Pedoman Reparasi* . Penerbit Service Publication Office.
- Turban, Efraim, 1995, “*Decision Support Systems And Expert Systems*”, USA : Prentice
- Sutojo. 2011 *Kecerdasan Buatan*. Penerbit Andi: Yogyakarta

Sistem Pakar Mendeteksi Kerusakan Sepeda Motor Menggunakan Metode Naïve Bayes di Bengkel Motor SMK Islam Tanjung

Nasa'i ¹⁾, Nindian Puspa Dewi ²⁾
^{1,2}Informatika, Teknik, Universitas Madura
^{1,2}Jln. Panglegur Km. 3,5 Pamekasan Jawa Timur Indonesia
¹nasha.lovers@gmail.com, ²nindianpd@unira.ac.id

ABSTARK

Sistem Pakar (expert system) adalah system atau program yang bertingkah laku seperti ahlinya atau pakarnya. Sistem pakar merupakan sistem yang berbasis pengetahuan yang bisa digunakan untuk membantu menyelesaikan masalah-masalah yang ada dalam dunia nyata. Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari sebuah dataset, sehingga naïve bayes dapat dijadikan metode pada aplikasi system pakar untuk mengetahui probabilitas suatu kerusakan yang mungkin terjadi pada sepeda motor. pada penelitian ini, dibuatakan sebuah aplikasi system pakar yang dapat mendeteksi kerusakan sepeda motor dan mampu memberikan solusi seperti seorang pakar dibengkel motor SMK Islam Tanjung yang dibangun menggunakan bahasa pemrograman php

Kata kunci: system pakar, naïve bayes, sepeda motor.

ABSTRACT

Expert systems are systems or programs that behave as experts. Expert systems are knowledge-based systems that can be used to help solve problems that exist in the real world. Naive Bayes is a simple probabilistic classification that calculates a set of probabilities by adding up frequencies and combinations of values from a dataset, so Naïve bayes can be used as a method in expert system applications to determine the probability of a damage that may occur on a motorcycle. in this study, an expert system application was created that could detect motorcycle damage and was able to provide solutions such as an expert at Islamic Vocational High School Tanjung motorcycle repair shop that built by using the php programming language.

Keywords: expert system, naïve bayes, motorbikes

PENDAHULUAN

Di zaman yang serba modern seperti saat ini kecepatan dan keakuratan dalam informasi sangatlah penting. Teknologi mempunyai peranan penting yang tentunya tidak terlepas kaitannya dengan Teknologi Informasi (TI). Komputer merupakan satu bagian yang paling penting dalam peningkatan Teknologi Informasi, salah satu faktornya yaitu kemampuan komputer dalam menyimpan dan mengingat informasi serta dapat dimanfaatkan semaksimal mungkin. Komputer memberikan kesimpulan atau pengambil keputusan yang kualitasnya sama dengan kemampuan seorang pakar dalam suatu bidang ilmu tertentu karena dapat menyimpan informasi aturan penalaran yang memadai. Salah satu cabang ilmu teknik informatika yang dapat mendukung hal tersebut adalah Sistem Pakar. Sistem pakar dapat diaplikasikan dalam berbagai bidang seperti dalam bidang perbengkelan yang berguna untuk mendiagnosa kerusakan berdasarkan gejala yang ditimbulkan seperti yang terjadi pada sepeda motor. Dalam hal ini penulis akan melakukan penelitian dibengkel motor SMK Islam Tanjung.

SMK Islam Tanjung merupakan sekolah menengah kejuruan yang berdiri di desa Dharma Tanjung kecamatan Camplong kabupaten Sampang pada tahun 2008 yang menjadi sekolah menengah kejuruan pertama di desa Dharma Tanjung. Sekolah ini sampai saat ini memiliki dua jurusan yaitu jurusan teknik komputer dan jaringan (TKJ) dan teknik bisnis sepeda motor (TBSM). TBSM merupakan jurusan teknik bisnis sepeda motor yang diharuskan oleh pemerintah untuk membuka bisnis secara mandiri sehingga SMK Islam Tanjung khusus jurusan teknik bisnis sepeda motor membuka bengkel motor yang diberi nama bengkel motor SMK Islam Tanjung. Usaha bengkel motor SMK Islam Tanjung ini didirikan pada pertengahan tahun 2015 yang beralamat di jalan Idaman Dh. Tanjung Camplong Sampang. Pada awalnya bengkel ini dijalankan oleh satu orang karyawan kemudian berkembang dan sekarang sudah memiliki tiga orang karyawan sebagai mekanik sepeda motor. Akan tetapi pihak bengkel motor SMK Islam Tanjung masih menggunakan cara manual dalam menganalisa kerusakan sepeda motor.

Berdasarkan latar belakang yang ada, maka dalam penelitian ini penulis membuat sistem pakar pendeteksi kerusakan sepeda motor menggunakan pemrograman berbasis web dan mysql dengan metode Naive Bayes. Diharapkan sistem dapat membantu memberikan saran dan solusi pada karyawan bengkel motor SMK Islam Tanjung dalam mendeteksi kerusakan sepeda motor pelanggannya..

TINJAUAN PUSTAKA

1. Sistem Pakar

Sistem Pakar (dalam bahasa Inggris : expert system) adalah sistem informasi yang berisi dengan pengetahuan yang berasal dari pakar. Pengetahuan di dalam sistem ini digunakan sebagai dasar oleh sistem pakar untuk menjawab pertanyaan (konsultasi).

Sistem pakar adalah suatu sistem yang dirancang untuk dapat menirukan keahlian seorang pakar dalam menjawab pertanyaan dan memecahkan suatu masalah. Sistem pakar akan memberikan pemecahan suatu masalah yang di dapat dari dialog dengan pengguna. Dengan bantuan sistem pakar seorang yang bukan ahli dapat menjawab pertanyaan, menyelesaikan serta mengambil keputusan yang biasanya di ambil oleh seorang pakar.

2. Naive Bayes

Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya.

Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. Keuntungan penggunaan Naive

Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Bayes merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayes memiliki akurasi dan kecepatan yang sangat tinggi saat diaplikasi ke dalam database dengan data yang besar. Berikut teorema bayes.

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} + \dots$$

Keterangan :	
X	: Data dengan class yang belum diketahui
H	: Hipotesis data x merupakan suatu class spesifik
P(H X)	: Probabilitas hipotesis H berdasarkan kondisi X
P(H)	: Probabilitas hipotesis H (prior probability)
P(X H)	: Probabilitas X berdasar kondisi pada hipotesis H
P(X)	: Probabilitas dari X

METODE PENELITIAN

Metode ini merupakan metode yang dipraktikkan langsung untuk mengumpulkan data yang berhubungan dengan data kerusakan dan gejala – gejala kerusakan sepeda motor di bengkel motor SMK Islam Tanjung. Data – data tersebut dikumpulkan dengan cara:

1. Penelitian Langsung

Disini penulis mengadakan observasi secara langsung terhadap obyek yang di teliti. Hal ini yang dilakukan penulis adalah mengamati langsung bagaimana kegiatan yang terjadi dilapangan yaitu di Bengkel Motor SMK Islam Tanjung. Dalam hal ini objek yang diamati adalah bagaimana caranya karyawan bengkel motor SMK Islam Tanjung dalam menganalisa masalah atau kerusakan sepeda motor.

2. Penelitian Langsung

Studi pustaka yaitu pengumpulan data melalui hasil dokumentasi, baik bahan tersebut diperoleh dari buku, jurnal, ataupun hasil browsing dari internet. Pada tahap ini penulis melakukan pencarian dan pemahaman teori-teori yang berkaitan dengan tools dan bahasa pemrograman yang

akan digunakan dan teori-teori lainnya yang dapat membantu dalam proses analisis, perancangan, sampai implementasi dari program yang akan dibuat

3. Data Kerusakan Sepeda Motor

Tabel 1 Data kerusakan Sepeda motor dan solusi

NO	Id Kerusakan	Kerusakan	Solusi
1	K1	Torak atau seher sudah mulai aus	Di colter atau diperbesar kapasitas mesinnya
2	K2	Banjir pada sistem pelumas atas	Menggati seal pompa oli atas
3	K3	Kampas kopling Aus	Mengganti Kampas kopling
4	K4	Aus Pada Sentrifugal Clutch	Mengganti Kampas Ganda
5	K5	Connecting rood atau stang seher rusak	Ganti Connecting rood
6	K6	Katup longgar	Stel ulang katup
7	K7	Seher Aus	Ganti seher
8	K8	Setelan karbu kurang pas	Stel ulang karbu
9	K9	Rantai kamrat longgar	Ganti atau setel ulang rantai

4. Data Gejala Kerusakan Sepeda Motor

Tabel 2 data gejala kerusakan sepeda motor

NO	Id gejala	Gejala
1	G1	Motor terasa bergetar lebih keras
2	G2	Suara gemeretak dibagian cvt
3	G3	Akselerasi di putaran mesin atas terasa tertahan
4	G4	Akselerasi lambat dan tertahan
5	G5	Akselerasi tersendat-sendat di putaran mesin bawah

6	G6	Muncul suara kasar
7	G7	Bunyi tek tek di rpm rendah
8	G8	Akselerasi di putaran menengah terasa tertahan
9	G9	Muncul suara berisik di rumah cvt
10	G10	Munculnya suara dengung cukup keras
11	G11	Muncul suara berdecit
12	G12	Akselerasi awal terasa seperti slip
13	G13	Bunyi tek tek terus menerus

5. Data kerusakan Sepeda Motor dan Gejalanya.

1. K1 :Roller aus
G1 :Motor terasa bergetar lebih keras
G2 :Suara gemeretak dibagian cvt
G3 :Akselerasi di putaran mesin atas terasa tertahan
2. K2 :Rumah roller aus
G1 :motor terasa bergetar lebih keras
G2 :suara gemeretak dibagian cvt
G4 :Akselerasi lambat dan tertahan
3. K3 :Mangkuk kopling aus
G1 :motor terasa bergetar lebih keras
G4 :Akselerasi lambat dan tertahan
G5 :Akselerasi tersendat-sendat di putaran mesin bawah
4. K4 :Kampas kopling tipis
G4 :Akselerasi lambat dan tertahan
G11 :Muncul suara berdecit
G5 :Akselerasi tersendat-sendat di putaran mesin bawah
5. K5 :Secondary sliding sheave aus
G4 :Motor terasa bergetar lebih keras
G6 :Muncul suara kasar
G8 :Akselerasi di putaran menengah terasa tertahan
6. K6 : V-belt mellar
G9 : Muncul suara berisik di rumah CVT
G12 : Akselerasi awal terasa seperti slip
G4 : Akselerasi lambat dan tertahan
7. K7 : Gear rasio aus
G10 : munculnya suara dengung cukup keras
G6 : muncul suara kasar
8. Kruke as terlalu keluar
G1 : Motor terasa bergetar lebih keras

- G13 : Bunyi tek Tek terus menerus
G6 : Muncul suara kasar
9. K9 : Pislet longgar
G6 : muncul suara kasar
G7 : Bunyi tek Tek di rpm rendah

6. Rule Base

1. If (G1 and G2 and G3)
Then K1
2. If (G1 and G2 and G4)
Then K2
3. If (G1 and G4 and G5)
Then K3
4. If (G4 and G5 and G11)
Then K4
5. If (G4 and G6 and G8)
Then K5
6. If (G4 and G9 and G12)
Then K6
7. If (G6 and G10)
Then K7
8. If (G1 and G6 and G13)
Then K8
9. If (G6 and G7)
Then K9

7. Perhitungan Bayes

Misalnya gejala yang tampak pada sepeda motor ada 3 yaitu :
Motor terasa bergetar lebih keras (G1)
Suara gemeretak dibagian cvt (G2)
Akselerasi diputaran mesin atas teras tertahan (G3)
Berdasarkan gejala tersebut maka dapat dihitung :
Roller Aus
Tahap pencarian (X | H)
Pada tahap ini akan mencari probabilitas roller aus.

$$= \frac{\text{jumlah kemungkinan}}{\text{jumlah kerusakan}} = \frac{1}{9} = 0.11$$

Tahap pencarian P(X)
Probabilitas pencarian gejala adalah :
Motor terasa bergetar lebih keras
= $\frac{\text{kemungkinan}}{\text{jumlah kemungkinan kerusakan akibat gejala}} = \frac{1}{3} = 0,33$

$$= \frac{\text{kemungkinan}}{\text{jumlah kemungkinan kerusakan akibat gejala}} = \frac{1}{3} = 0,33$$

Suara gemeretak dibagian CVT
Akselerasi diputaran mesin atas terasa tertahan

$$= \frac{\text{kemungkinan}}{\text{jumlah kemungkinan kerusakan akibat gejala}} = \frac{1}{3} = 0,33$$

Perhitungan nilai bayes

$$K(K1 | G1) = [K (G1 | K1) * K (K1)] / [K (G1 | K1) * K (K1) + K (G1 | K2) * K (K2) + K (G1 | K3) * K (K3) + K (G1 | K4) * K (K4) + K (G1 | K5) * K (K5) + K (G1 | K6) * K (K6) + K (G1 | K7) * K (K7) + K (G1 | K8) * K (K8) + K (G1 | K9) * K (K9)]$$

$$K(K1 | G1) = \frac{0,0363}{0,181}$$

K(K1 | G1) = 0,200 (Hasil perhitungan bayes 1 pada kerusakan K1)

$$K(K1 | G2) = [K (G2 | K1) * K (K1)] / [K (G2 | K1) * K (K1) + K (G2 | K2) * K (K2) + K (G2 | K3) * K (K3) + K (G2 | K4) * K (K4) + K (G2 | K5) * K (K5) + K (G2 | K6) * K (K6) + K (G2 | K7) * K (K7) + K (G2 | K8) * K (K8) + K (G2 | K9) * K (K9)]$$

$$K(K1 | G2) = \frac{0,0363}{0,072}$$

K(K1 | G2) = 0,50 (Hasil perhitungan bayes 2 pada kerusakan K1)

$$K(K1 | G3) = [K (G3 | K1) * K (K1)] / [K (G3 | K1) * K (K1) + K (G3 | K2) * K (K2) + K (G3 | K3) * K (K3) + K (G3 | K4) * K (K4) + K (G3 | K5) * K (K5) + K (G3 | K6) * K (K6) + K (G3 | K7) * K (K7) + K (G3 | K8) * K (K8) + K (G3 | K9) * K (K9)]$$

$$K(K1 | G3) = \frac{0,0363}{0,0363}$$

K(K1 | G3) = 1 (Hasil perhitungan bayes 3 pada kerusakan K1)

Total Bayes Ke 1
= K (K1 | G1) + K (K1 | G2) + K (K1 | G3)
= 0,200 + 0,50 + 1 = 1,704 (total bayes ke tiga pada kerusakan (K1))

Langkah diatas diulang ulang sebanyak jumlah kerusakan yang ada, sampai ditemukan persentase kerusakan seperti dibawah ini.

Hasil = Total Bayes Pertama + Total Bayes Kedua + Total Bayes Ketiga + Total Bayes Keempat + Total Bayes Kelima + Total Bayes Keenam + Total Bayes Ketujuh + Total Bayes Kedelapan + Total Bayes Kesembilan

$$= K(K1 | G1) + K(K1 | G2) + K(K1 | G3) + K(K2 | G1) + K(K2 | G2) + K(K2 | G3) + K(K3 | G1) + K(K3 | G2) + K(K3 | G3) + K(K4 | G1) + K(K4 | G2) + K(K4 | G3) + K(K5 | G1) + K(K5 | G2) + K(K5 | G3) + K(K6 | G1) + K(K6 | G2) + K(K6 | G3) + K(K7 | G1) + K(K7 | G2) + K(K7 | G3) + K(K8 | G1) + K(K8 | G2) + K(K8 | G3) + K(K9 | G1) + K(K9 | G2) + K(K9 | G3)$$

= 1,7+0,7+0,2+0+0,2+0+1.0+0,2+0 = 3
= 3 hasil total dari seluruh nilai perhitungan bayes

Maka Perhitungan Probabilitas Kerusakan Sepeda Motor Adalah :

K1 : Kerusakan pada roller
= $\frac{\text{Total Bayes Pertama}}{\text{Hasil}} \times 100$
= $\frac{1,7}{3} \times 100 = 56,66\%$

K2: Kerusakan pada rumah roller aus
= $\frac{\text{Total Bayes Kedua}}{\text{Hasil}} \times 100$
= $\frac{0,7}{3} \times 100 = 23,33\%$

K3: kerusakan pada mangkuk kopling aus
= $\frac{\text{Total Bayes Ketiga}}{\text{Hasil}} \times 100$
= $\frac{0,2}{3} \times 100 = 6,66\%$

K4: Kerusakan pada kampas kopling tipis
= $\frac{\text{Total Bayes Keempat}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

K5: Kerusakan pada secondary sleading sheave aus
= $\frac{\text{Total Bayes Kelima}}{\text{Hasil}} \times 100$
= $\frac{0,2}{3} \times 100 = 6,66\%$

K6: Kerusakan Pada V-Belt mellar
= $\frac{\text{Total Bayes Keenam}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

K7: Kerusakan pada gear rasio aus
= $\frac{\text{Total Bayes Pertama}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

K8: Kerusakan pada kuke as terlalu keluar
= $\frac{\text{Total Bayes Kedelapan}}{\text{Hasil}} \times 100$
= $\frac{0,2}{3} \times 100 = 6,66\%$

K9: Kerusakan pada pislet longgar
= $\frac{\text{Total Bayes Kesembilan}}{\text{Hasil}} \times 100$
= $\frac{0}{3} \times 100 = 0\%$

Dari perhiitungan diatas dapat disimpulkan bahwa kerusakan yang mungkin terjadi adalah sebagai berikut

1. Roller aus 56,66 %
2. Rumah roller aus 23,33 %
3. Mangkuk kopling aus 6,66 %
4. Kampas kopling tipis 0 %
5. Secondary sleading sheave aus 6,66 %
6. V-Belt mellar 0 %
7. Gear rasio aus 0 %
8. Kruke as terlalu keluar 6,66 %
9. Pislet longgar 0 %

Jadi berdasarkan persentase hasil perhitungan bayes diatas dapat disimpulkan bahwa kemungkinan kerusakan terjadi pada :

1. K1 (Roller aus) 56,66 %
2. K2 (Rumah roller aus) 23,33 %

HASIL DAN PEMBAHASAN

Berikut tampilan beranda. Pada halaman ini admin dapat login dengan mengklik tombol admin



Gambar 1 halaman beranda



Gambar 2 form login admin

Pada gambar 2 admin dapat melakukan login ke aplikasi dengan memasukkan user name dan password lalu click login, apabila login sukses maka akan diarahkan ke hakaman beranda admin seperti gambar 3 dan admin sudah dapat melakukan management seluruh data yang dibutuhkan di aplikasi system pakar ini.



Gambar 3 halaman admin

Pada gambar 1 user atau pelanggan dapat memberikan saran gejala yang tidak ada di dalam aplikasi dan pelanggan juga

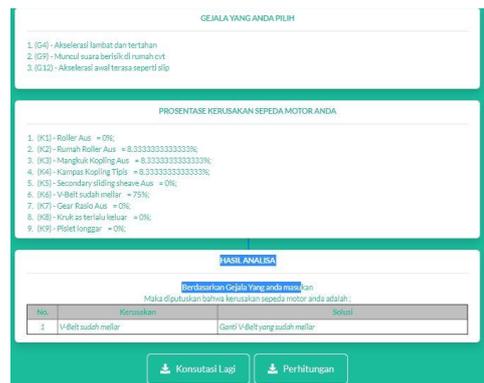
dapat melakukan konsultasi untuk mendapatkan informasi kerusakan sepeda motor dengan menekan tombol konsultasi dan akan diarah ke halaman konsultasi tetapi sebelumnya harus mengisi identitas seperti gambar 4 dibawah ini.



Gambar 4 form input identitas



Gambar 5 pertanyaan pertanyaan



Gambar 6 hasil akhir

KESIMPULAN

1. Kesimpulan

Berdasarkan hasil penelitian yang sudah dilakukan, peneliti dapat menyimpulkan bahwa system pakar yang dibuat dapat membantu karyawan bengkel motor SMK Islam tanjung dalam mendeteksi kerusakan sepeda motor pelanggan terutama pada saat jam jam sibuk.

Dengan adanya aplikasi ini siapapun bisa melakukan konsultasi meskipun ahlinya tidak ada di bengkel pada saat itu.

DAFTAR PUSTAKA

- Saleh, A, 2015. *Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*. [pdf]. ISSN: 2354-5771. Januari 2015. Universitas Potensi utama. Medan. Indonesia 2015.
- Wahyudi, S., & Ratnasari, S. 2015 . *Sistem Pakar Diagnosis Penyakit Mata Menggunakan Naive Bayes Classifier*. [pdf] *Seminar Nasional Sains dan Teknologi*. 1 Februari 2015. Universitas Trunojoyo Madura.
- Wawan, P, & Yustina. 2015. *Sistem Untuk Deteksi Kerusakan Mesin Diesel Mobil Panther Dengan Metode Naïve Bayes*. [pdf] ISSN : 2338-4018. *Jurnal TIKomSIN*.
- Rodiyansyah, S. F., & Winarko, E. 2012. *Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification*. [pdf] ISSN: 1978-1520. Januari 2013. FMIPA UGM.
- Muhammad & Ardi, 2017. *Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit*. [pdf]. ISSN 2355-5920. September 2017. *Jurnal Pseudocode*, Volume IV Nomor 2.
- Adi, Haris. *Sistem Informasi Kerusakan Laptop Menggunakan Metode Naïve Bayes*. [pdf] ISSN: 2086-9479. *Jurnal Teknologi Elektro*. September 2017. Universitas Mercu Buana.
- Honda Motor. 2014 *Pedoman Reparasi Umum*. Penerbit Service Publication Office.
- Honda Motor. 2016 *Suplemen Pedoman Reparasi* . Penerbit Service Publication Office.
- Turban, Efraim, 1995, “*Decision Support Systems And Expert Systems*”, USA : Prentice
- Sutojo. 2011 *Kecerdasan Buatan*. Penerbit Andi: Yogyakarta

SISTEM INFORMASI PONDOK PESANTREN DARUL ULUM BANYUANYAR BERBASIS WEB DAN ANDROID

Abdus Sakur ¹⁾, Ubaidi ²⁾

^{1,2}Teknik Informatika, Fakultas Teknik, Universitas Madura
Jl. Raya Panglegur No. Km 3,5 Pamekasan
¹abdusakur71@gmail.com, ²ubed@unira.ac.id

ABSTRAK

Selama ini pendataan santri dan kegiatannya dilakukan secara manual dengan menggunakan kertas sebagai media utamanya. Beberapa kegiatan yang memerlukan pendataan yang tertib adalah penerimaan santri baru, data tahfidz dan data kesehatan santri. Data-data santri tersebut diperlukan untuk dapat dijadikan informasi perkembangan santri bagi wali santri. Untuk pendataan tahfidz, masih dilakukan manual dengan menggunakan buku catatan yang dipegang oleh santri. Setiap santri selesai, maka akan dicatat pada buku tersebut dengan ditandatangani oleh ustadznya. Namun karena hafalan yang dilakukan untuk seluruh Al-Quran, maka cara seperti ini akan menyulitkan bagi ustadz dan santri untuk melakukan pencarian data. Apalagi jika buku catatan tersebut hilang. Dalam penelitian ini akan dibangun sebuah sistem informasi manajemen yang dapat digunakan untuk mengelola data kegiatan santri yang meliputi data penerimaan santri baru, data kondisi kesehatan, dan data tahfidz. Dengan adanya sistem ini, diharapkan proses pengelolaan, penjadwalan, serta pelaporan data kegiatan santri dapat dilakukan dengan lebih efisien. Hasil penelitian menunjukkan bahwa tiap user akan dipermudah dalam mengelola data karena menggunakan aplikasi berbasis android. Memudahkan dalam melakukan pencarian data. Wali santri akan mudah untuk mengetahui perkembangan santri tanpa harus memantau secara langsung ke pondok pesantren.

Kata kunci : Aplikasi, Android, Pondok Pesantren, Sistem Informasi, Tahfidz

ABSTRACT

For all these time, school boarding students do their activities manually. They use paper as main media. Some activities like school boarding registration, tahfidz and healthy. These information can be used to check school boarding students progress. Using paper for data collection will complicate because paper easy to damage and difficult to make searching. In this research, will built an application to manage school boarding registration, health and tahfidz data collection. The result show that every user will easier to manage data collection and searching data. Parents will easy to knows development of school boarding student whitout comes to school boarding.

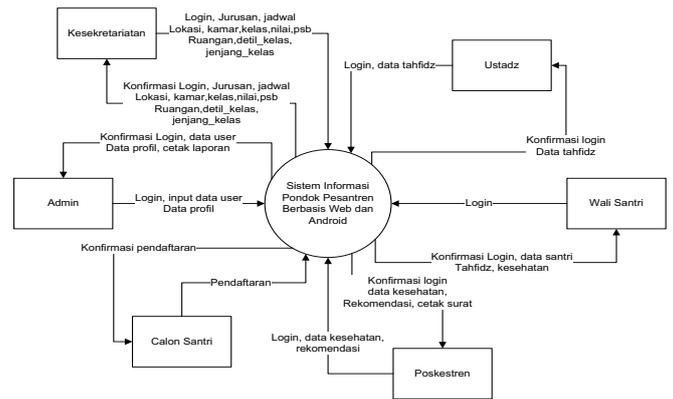
Keywords: *Aplication, android, school boarding, information system, tahfidz*

PENDAHULUAN

Semua kegiatan di dalam pesantren harus tercatat dengan rapi dan tertib. Hal ini mengingat jumlah santri yang banyak maka pendataan santri merupakan suatu masalah tersendiri bagi pondok pesantren. Selama ini pendataan santri dan kegiatannya dilakukan secara manual dengan menggunakan kertas sebagai media utamanya. Beberapa kegiatan yang memerlukan pendataan yang tertib adalah penerimaan santri baru, data tahfidz dan data kesehatan santri. Hal ini tentu saja tidak efisien, selain membutuhkan banyak kertas untuk menyimpan data kegiatan dari ratusan santri, model penyimpanan ini juga tidak memudahkan pengurus dalam proses pencarian data kegiatan santri. Sehingga dibutuhkan sebuah sistem terkomputerisasi yang dapat membuat penyimpanan data kegiatan santri menjadi lebih efisien. Kelebihan dari aplikasi ini adalah penggunaan perangkat telekomunikasi berbasis android yang akan memudahkan penggunaannya. Dengan aplikasi ini, pengguna dapat menggunakan aplikasi dimana saja dan kapan saja tanpa bergantung pada komputer atau laptop. Hal ini dilakukan karena perangkat telekomunikasi berbasis android sudah mewabah di masyarakat.

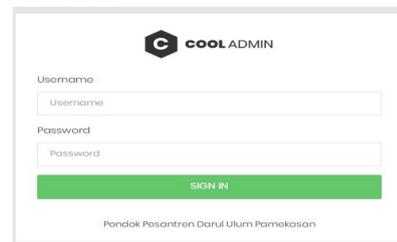
METODE PENELITIAN

Sistem terdiri dari beberapa user yaitu admin, kesekretariatan, poskestren, tahfidz, wali santri dan calon santri. Admin dapat melakukan proses login, kelola user, kelola hak akses dan kelola ustadz. Kesekretariatan dapat melakukan edit user, kelola jurusan, kelola ruangan, kelola jenjang kelas, kelola kelas, kelola kamar, kelola nilai, kelola jadwal, kelola nilai santri baru, kelola jadwal santri baru, edit dan hapus data santri. Tahfidz dapat melakukan edit user dan kelola tahfidz. Poskestren dapat melakukan edit user dan kelola data kesehatan santri. Wali santri dapat melakukan view data santri, tahfidz dan poskestren. Sedangkan calon santri dapat melakukan pendaftaran.

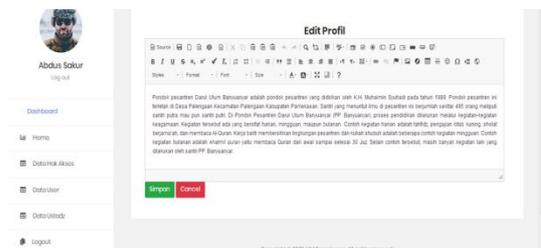


HASIL DAN PEMBAHASAN

Sistem terdiri dari beberapa user yaitu: admin, kesekretariatan, ustadz, poskestren dan wali santri.



Login Admin



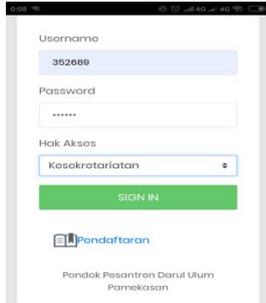
Menu Admin

Menu admin akan menampilkan form untuk mengubah data profil sistem. admin dapat mengubah data dengan cara mengganti data yang ada dan menekan tombol Simpan untuk menyimpan data. Menu Kesekretariatan digunakan untuk mengelola data yang berhubungan dengan data santri baik calon santri maupun santri lama. Semua user kecuali admin menggunakan aplikasi android untuk memudahkan dalam mengelola data. Untuk

itu, user harus menginstal aplikasi dalam smartphone berbasis android.

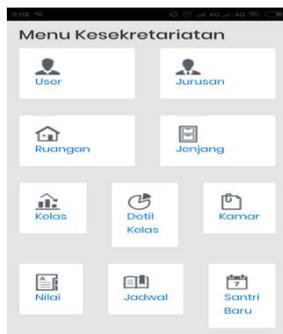


Instal Aplikasi



Login Kesekretarian

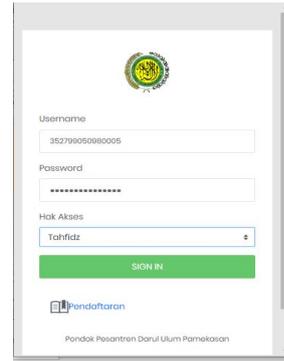
Menu ini akan form untuk melakukan login. Jika login berhasil, maka user akan masuk ke menu kesekretariatan.



Menu Kesekretarian

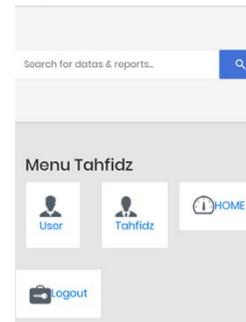
Menu ini akan menampilkan menu yang dapat diakses oleh user kesekretariatan. Untuk mengakses menu dapat dilakukan dengan menekan tombol menu yang ada.

Menu Tahfidz digunakan oleh ustadz untuk melakukan pendataan tahfidz santri. Untuk masuk ke menu tahfidz, ustadz harus melakukan login. Jika login berhasil, maka sistem akan menampilkan menu tahfidz.



Login Tahfidz

Login dilakukan dengan cara memasukkan username dan password dan memilih jenis hak akses sebagai tahfidz.



Menu Tahfidz

Setelah berhasil melakukan login, maka sistem akan menampilkan menu tahfidz yang terdiri dari menu user dan tahfidz.



Menu Tahfidz Santri

Menu ini akan menampilkan data santri yang ikut program tahfidz. Pada bagian bawah data tiap santri terdapat tombol Detil dan Tambah Data.



Tambah Data Tahfidz

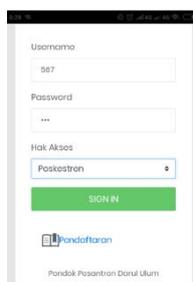
Menu ini digunakan jika terdapat santri yang akan melanjutkan tahfidznya. Ustad memasukkan data yang ada pada form. Tombol Simpan akan menyimpan data yang telah dimasukkan.



Detil Tahfidz

Menu ini akan memberikan informasi pada user tentang progress tahfidz dari tiap santri. Ustadz dapat mengubah dan menghapus data dengan cara menekan tombol Edit dan Hapus.

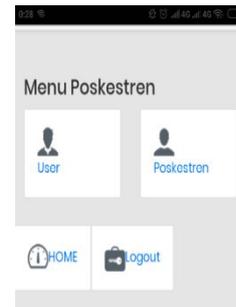
Menu Poskestren digunakan oleh ustadz untuk melakukan pendataan kesehatan santri. Untuk masuk ke menu poskestren, ustadz harus melakukan login. Jika login berhasil, maka sistem akan menampilkan menu poskestren.



Login Poskestren

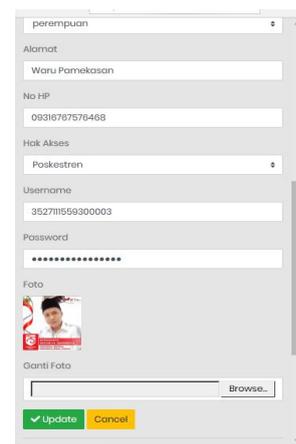
Login dilakukan dengan cara memasukkan data username dan password dan memilih hak akses sebagai poskestren. Jika login

berhasil, maka system akan menampilkan menu poskestren.



Menu Poskestren

Menu poskestren terdiri dari dua menu yaitu User dan Poskestren.



Poskestren Santri

Menu ini akan menampilkan data semua santri. Pada bagian bawah terdapat tombol Detil dan Tambah Data.



Detil Poskestren Santri

Menu ini akan menampilkan informasi kesehatan santri. Pada bagian bawah terdapat tombol Unduh Surat Rekom, Edit dan Hapus. Tombol Unduh Surat Rekom akan mengunduh file surat rekom. Surat rekom terdiri dari Ijin, Sakit dan Pulang. Sedangkan

tombol Edit dan Hapus untuk mengubah dan menghapus data yang telah dimasukkan.

Tambah Data Poskestren Santri

Menu ini digunakan untuk memasukkan data kesehatan santri dan memberikan surat rekom yang dapat digunakan untuk disampaikan pada pengurus pondok pesantren.

Calon santri adalah orang yang akan menjadi santri di Pondok Pesantren Darul Ulum. Proses awal dilakukan dengan cara melakukan pendaftaran dengan cara menekan link Pendaftaran yang ada pada saat login.

Form pendaftaran

Menu ini akan menampilkan form untuk melakukan pendaftaran.

Login Calon Santri

Login calon santri diperlukan jika calon santri ingin mengetahui informasi yang berhubungan dengan penerimaan santri baru.

Menu Calon Santri

Menu ini akan menampilkan data user calon santri. Untuk mengetahui informasi tentang penerimaan santri baru, calon santri dapat menekan tombol User.

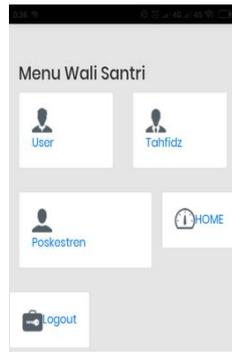
Data Calon Santri

Menu ini akan menampilkan data calon santri yang sudah dilengkapi oleh operator kesekretariatan seperti data jurusan, jadwal.

Wali santri adalah orang tua atau wali dari santri. Wali santri dapat mengetahui informasi yang berhubungan dengan santrinya. Untuk masuk ke menu wali santri harus melakukan login.

Login Wali Santri

Login dilakukan dengan cara memasukkan username dan password dan hak akses sebagai wali santri.



Menu Wali Santri

Jika login berhasil, maka sistem akan menampilkan menu wali santri. Menu wali santri terdiri dari menu user, menu tahfidz dan menu poskestren.



Menu User Wali Santri

Menu ini akan menampilkan data pribadi santri dan lokasi tempat tinggalnya di pondok pesantren.



Tahfidz Wali Santri

Menu ini akan menampilkan informasi tahfidz santri.

PENUTUP

Adapun kesimpulan yang dapat diambil dari pembahasan dari bab sebelumnya adalah:

1. Tiap user akan dipermudah dalam mengelola data karena menggunakan aplikasi berbasis android.

2. Memudahkan dalam melakukan pencarian data.
3. Wali santri akan mudah untuk mengetahui perkembangan santri tanpa harus memantau secara langsung ke pondok pesantren.

Adapun saran untuk peningkatan sistem adalah:

1. Melengkapi fitur cetak pada semua data yang ditampilkan.
2. Ditambahkan petunjuk penggunaan bagi semua user.

DAFTAR PUSTAKA

- Triansah, A., Cahyadi, D., & Astuti, I. F. 2015. *Membangun Aplikasi Web dan Mobile Android Untuk Media Pencarian Kost Menggunakan Phonegap Dan Google Maps API*. Jurnal Informatika Mulawarman, 10(1), 58–61
- Ahmad Bagus Setiawan, Juli Sulaksono, 2019. *Sistem Informasi Manajemen Santri Di Pondok Pesantren Al Ishlah Kota Kediri*. Jurnal Ilmiah NERO Vol. 4, No.2
- Erliyah Nurul Jannah, Inayatul Karimah, Chandra Sukma Anugrah, 2018. *Sistem Informasi Manajemen Kegiatan Santri Pondok Pesantren Berbasis Web*. Seminar Nasional Teknologi Informasi dan Komunikasi
- Fathansyah.2015.*Basis Data Revisi Kedua*.Bandung:Informatika Bandung.
- Haryanto, E. M. O. 2013. *Perancangan Aplikasi Mobile Pencarian Lokasi Penyewaan Rumah Kost Berbasis Android Di Kota Yogyakarta*. Naskah Publikasi Sekolah Tinggi Manajemen Informatika Dan Komputer Amikom Yogyakarta.
- Indra, Yatini.2010.*Flowchart, Algoritma, dan Pemrograman Menggunakan C++ Builder*.Yogyakarta:Graha Ilmu.

Peranginangin, Kasiman. 2006. *Aplikasi WEB
dengan PHP dan MySQL*.
Yogyakarta: C.V Andi Offset.

KLASIFIKASI SOAL PILIHAN GANDA BERBAHASA INDONESIA BERDASARKAN LEVEL KOGNITIF PUSPENDIK DENGAN SUPPORT VECTOR MACHINE

Lilis Setyaningsih¹⁾, Endang Setyati²⁾, Suhatati Tjandra³⁾
^{1,2,3}Teknologi Informasi, Institut Sains dan Teknologi Terpadu Surabaya
 Jl. Ngagel Jaya Tengah No. 73-77, Baratajaya, Gubeng, Surabaya
¹liestyaaq@gmail.com,²endang@stts.edu,³tati@stts.edu

ABSTRAK

Bank soal merupakan kebutuhan penting di sekolah, sebagai media evaluasi pembelajaran. Bank soal yang berkualitas, dapat membantu dalam menganalisa kemampuan kognitif peserta didik. Pada akhirnya meningkatkan kualitas pembelajaran di sekolah. Penelitian ini menggunakan dataset berupa pokok soal dari soal PG, yang di dalamnya terdapat stimulus yang berupa text. Soal akan diklasifikasi berdasarkan level kognitif Puspendik, yaitu level pengetahuan dan pemahaman (L1), level aplikasi (L2) dan level penalaran atau HOTS (L3). Tidak semua soal menggunakan kata kerja operasional. Proses penelitian dilakukan dalam 3 tahap, yaitu pengumpulan dan labeling dataset, preprocessing, training model, dan testing data. Penelitian dilakukan dengan *Bag Of Words*, *TF-IDF*, *Latent Semantic Analysis* dan metode *Support Vector Machine*. Hasil uji coba menunjukkan tingkat akurasi dalam mengklasifikasi soal L1 sebesar 64,67%, L2 75,42% dan L3 25,83%. Tingkat akurasi secara keseluruhan dicapai, pada saat ujicoba 30% data testing, yaitu 61,54%.

Kata kunci : Klasifikasi Soal Pilihan Ganda, Level Kognitif Puspendik, Latent Semantic Analysis, Support Vector Machine.

ABSTRACT

A question bank is an important need in schools as a medium for evaluating learning. A quality question bank, can help in analyzing the cognitive abilities of students. In the end, it will improve the quality of learning in schools. This study uses a dataset in the form of subject matter from PG questions, in which there is a stimulus in the form of text. Questions will be classified based on Puspendik's cognitive level, namely the level of knowledge and understanding (L1), the level of application (L2) and the level of reasoning or HOTS (L3). Not all questions use operational verbs. The research process was carried out in 3 stages, namely collecting and labeling datasets, preprocessing, training models, and testing data. The research was conducted with Bag of Words, TF-IDF, Latent Semantic Analysis and Support Vector Machine methods. The test results showed that the level of accuracy in classifying L1 questions was 64.67%, L2 75.42% and L3 25.83%. The overall level of accuracy was achieved, at the time of testing 30% of the testing data was 61.54%.

Keywords: *Multiple Choice Question Classification, Puspendik Cognitive Level, Latent Semantic Analysis, Support Vector Machine.*

PENDAHULUAN

Evaluasi pembelajaran merupakan salah bentuk pengukuran tingkat kognitif peserta didik. Ranah kognitif berkaitan dengan kemampuan berpikir, kemampuan memahami, mengimplementasi dan melakukan penalaran. Evaluasi pembelajaran akhir semester atau ujian sekolah biasanya

dilaksanakan dengan menggunakan soal bertipe pilihan ganda (PG).

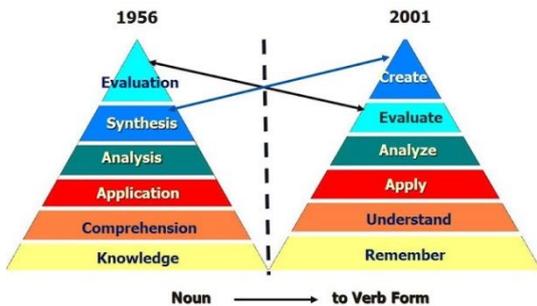
Pada penelitian ini diharapkan mampu mengklasifikasikan soal PG, sehingga sekolah mempunyai bank soal yang berkualitas agar dapat mengukur tingkat kognitif peserta didik.

Dataset menggunakan soal PG yang digunakan di tingkat X SMK. Bagian soal yang digunakan hanya bagian pokok soal (*stem*) tanpa option jawaban. Stimulus soal yang digunakan, hanya stimulus soal yang berupa text.

Klasifikasi soal akan menghasilkan 3 kelas, yaitu L1, L2 dan L3. Klasifikasi dilakukan dengan menggunakan metode Support Vector Machine, dan 3 kernelnya yaitu Linear, Poly dan RBF.

Taksonomi Bloom Revisi

Taksonomi Bloom (Revisi) membagi kognitif dalam 6 level, yaitu (C1) *remember* atau mengingat, (C2) *understand* atau memahami, (C3) *apply* atau menerapkan, (C4) *analyze* atau menganalisa, (C5) *evaluate* atau mengevaluasi, dan (C6) *create* atau mencipta (Kusuma, S. F. 2015)



Gambar 1.Level Kognitif Taksonomi Bloom Revisi

<https://dreamlikechild.weebly.com/uploads/3/9/0/4/39042447/827490737.jpg>

Pada Taksonomi Bloom Revisi, terdapat pergantian dari penggunaan kata benda menjadi kata kerja yang mampu menggambarkan proses kognitif.

Level Kognitif Puspendik

Puspendik (Pusat Penilaian Pendidikan Kemendikbud RI) sejak tahun 2015 telah membagi level kognitif menjadi 3 level, yaitu level pengetahuan dan pemahaman atau L1 (C1,C2), level aplikasi atau L2 (C3) dan level penalaran atau L3 (C4, C5, C6 atau HOTS) (Asmira, O. 2018)

Pada setiap level tersebut mempunyai ciri-ciri katakerja operasional yang sering digunakan seperti pada tabel 1.

Tabel 1.Contoh KKO Level Kognitif Puspendik

Level I	KKO
L1	Mengutip Menyebutkan Menjelaskan Membilang Mengidentifikasi Membaca Menamai Menandai Menghafal Meniru Mencatat Mengulang Meninjau Menulis Menyatakan Memperkirakan Menceritakan Mendiskusikan Menerangkan Membedakan
L2	Mengurutkan Menerapkan Mengkalkulasi Memodifikasi Menghitung Mencegah Menentukan Menggambarkan Menggunakan Mengemukakan Mengadaptasi Mempersoalkan Melaksanakan Melakukan Mensimulasikan Mentabulasi Mengklasifikasi Mengoperasikan Meramalkan
L3	Mengaudit Mengatur Memecahkan Menegaskan Menganalisis Merinci Menominasikan Mendiagramkan Mengkorelasikan Menguji Memerintahkan Mengaitkan Menransfer Melatih Mengedit Menemukan Menyeleksi Mendeteksi Menelaah Membangkakan Merancang Membatas Mereparasi Mengkonstruksi Merumuskan Menghubungkan Menciptakan Menampilkan

Soal Pilihan Ganda

Soal pilihan ganda terdiri dari bagian pokok soal (*stem*) dan option jawaban. Pada pokok soal terdapat stimulus dan pertanyaan/pernyataan tidak lengkap (Puspendik. 2019). Lihat gambar 2 di bawah ini, bagian-bagian dari pokok soal pilihan ganda.

Contoh Soal Pilihan Ganda

1. Tono, Budi, dan Andi membeli alat-alat tulis yang sama di sebuah toko. Tono membeli dua buku tulis, satu pena, dan satu pensil dengan membayar Rp11.000,00, Budi membeli satu buku tulis, satu pena dan satu pensil dengan harga Rp7.000,00, Andi membeli tiga buku tulis dan dua pena, ia membayar seharga Rp16.000,00.

Jika banyak buku tulis, pena, dan pensil yang mereka beli berturut-turut dinyatakan dengan x , y , dan z , model matematika dari permasalahan tersebut adalah ...

A. $2x + y + z = 11.000; x + y + 2z = 7.000; 3x + 2y = 16.000$ } Pengecoh

B. $2x + y + z = 11.000; x + y + z = 7.000; 3x + y = 16.000$ } **Kunci Jawaban**

C. $2x + y + z = 11.000; x + y + z = 7.000; 3x + 2y = 16.000$ } Pengecoh

D. $2x + y + z = 11.000; x + y + z = 7.000; 2x + 3y = 16.000$ } Pengecoh

E. $x + 2y + z = 11.000; x + y + z = 7.000; 3x + 2y = 16.000$ } Pengecoh

Labels in image: Pokok Soal (Stem), Stimulus, Pertanyaan/pernyataan tidak lengkap, Kunci Jawaban, Pengecoh.

Gambar 2. Bagian Pokok Soal Pilihan Ganda

Pembobotan Term dengan TF-IDF

TF-IDF merupakan suatu algoritma yang digunakan dalam pembobotan terhadap kemunculan suatu kata (token) yang dijadikan sebagai fitur dalam suatu dokumen seleksi (Sudin, S. 2018)

Langkah perhitungan TF-IDF

1. Menghitung Term Frequency, yaitu frekwensi kemunculan term dalam sebuah dokumen

$$tf(t, d) = \left(\frac{\text{jumlah kemunculan term pada satu dokumen}}{\text{jumlah seluruh term dalam satu dokumen}} \right) \dots (1)$$

2. Menghitung Nilai *idf*. *Idf* adalah hubungan antara banyak jumlah kata dengan dokumen

$$idf(t, D) = \log \left(\frac{\text{jumlah seluruh dokumen}}{\text{jumlah dokumen suatu term muncul}} \right) \dots (2)$$

3. Menghitung bobot (w) kata dalam dokumen.

$$W = (tf * idf) \dots (3)$$

Latent Semantic Analysis

Latent Semantic Analysis (LSA) merupakan sebuah metode yang memanfaatkan model statistik matematis untuk menganalisa struktur semantik suatu teks (Suhartono, D. 21015). LSA menerapkan matriks SVD (*Singular Value Decomposition*), untuk mereduksi dimensi nilai kompleksitas dalam pemrosesan term dokument matriks.

SVD merupakan teorema aljabar linier yang menyebutkan bahwa persegi panjang dari term-document matrix dapat dipecah/didekomposisikan menjadi tiga matriks, yaitu : matriks ortogonal U, matriks diagonal S, dan transpose dari matriks ortogonal V.

$$A_{mn} = U_{mm} \times S_{mn} \times V_{mn}^T \dots (4)$$

A_{mn} = Matrik awal

U_{mn} = Matrik Ortogonal U

S_{mn} = Matrik Diagonal S

V_{mn}^T = Transpose Matrik Ortogonal V

Cosine similarity digunakan untuk menghitung nilai kosinus sudut antara vektor dokumen dengan vektor kueri.

Semakin kecil sudut yang dihasilkan, maka tingkat kemiripan kata semakin tinggi.

$$\cos \alpha = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \dots (5)$$

A = vektor dokumen

B = vektor kueri

A . B = perkalian dot vektor A dan vektor B

|A| = panjang vektor A

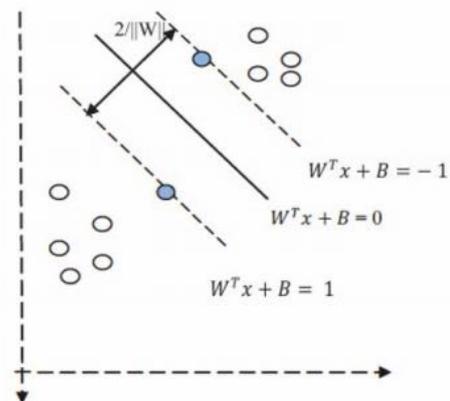
|B| = panjang vektor B

|A| |B| = cross product antara |A| dan |B|

α = sudut yang terbentuk antara vektor A dan B

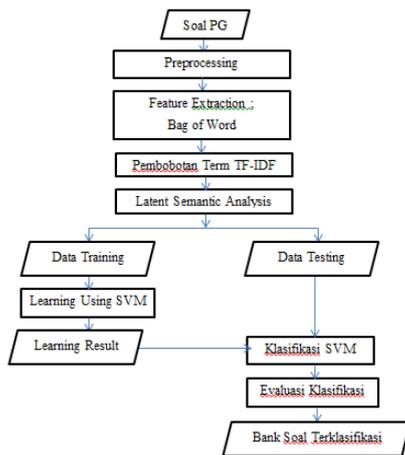
Support Vector Machine

SVM (Support Vector Machine) adalah metode klasifikasi untuk data linier dan nonlinier. Pemetaan nonlinear untuk mengubah data pelatihan menjadi dimensi yang lebih tinggi, dan kemudian menemukan hyperplane optimal linier untuk pemisahan kategori (Valtemir, 2018).



Gambar 3. Proses Klasifikasi SVM

METODE PENELITIAN



Gambar 4. Alur Sistem

Pada penelitian ini dilakukan tahapan-tahapan proses seperti terlihat pada gambar 4 di bawah ini :

Dataset Penelitian

Soal PG dikumpulkan dari bank soal sekolah dan buku pelajaran. Soal-soal tersebut diklasifikasi secara manual berdasarkan KKO (L1, L2 dan L3) oleh Musyawarah Guru Mata Pelajaran (MGMP) Sekolah yang terdiri dari 2 guru yang mempunyai pengalaman mengajar lebih dari 10 tahun serta divalidasi oleh Tim Kurikulum.

Tabel 2. Contoh Dataset

Kelas	Mapel	Soal
L1	INDO	Asap kendaraan bermotor yang berwarna putih disebabkan adanya kebocoran dibagian ring seker. Jika asapnya hitam berarti terlalu banyak campuran bensin daripada udara di dalam karburator. Laporan di atas menjelaskan tentang jenis laporan
L2	SJRH	Adakalanya sejarah menggunakan ilmu sosial dan sebaliknya ilmu sosial menggunakan sejarah. Artinya, terjadi persilangan antara konsep diakronik dan sinkronik. Penerapan persilangan Konsep tersebut dapat dilakukan pada pilihan ...
L3	PKN	Perhatikan karakteristik berikut! 1) Daerah yang berbatasan dengan laut bebas. 2) Wilayah perairan ini diukur sejauh 200

Kelas	Mapel	Soal
		mil dari garis pantai yang paling jauh menjorok ke laut (garis dasar). 3) Pada daerah ini Negara pantai memiliki kedaulatan dan prioritas utama dalam mengelola sumber daya alam yang terkandung di dalamnya. Berdasarkan karakteristik tersebut, dapat disimpulkan bahwa wilayah yang dimaksud adalah ...

Soal PG terdiri dari 3 mata pelajaran yaitu Bahasa Indonesia (INDO), Pendidikan Pancasila dan Kewarganegaraan (PKN), dan Sejarah Indonesia (SJRH). Jumlah dataset 135 soal, yang terdiri dari 45 soal level pengetahuan dan pemahaman (L1), 45 soal level aplikasi (L2) dan 45 soal level penalaran (L3). Jumlah soal per mapel adalah 45 soal.

Preprocessing

Pada Text Mining, preprocessing data digunakan untuk mengekstraksi fitur-fitur menarik dan pengetahuan dari data teks yang tidak terstruktur. Text data biasanya mengandung beberapa format seperti angka, tanggal, kata-kata umum yang tidak digunakan pada Text Mining seperti preposisi, pronominal dapat dihilangkan. (Kannan, S. 2014)

Berikut adalah penjelasan dari alur preprocessing yang terdapat pada gambar 4:

- a. **Cleaning data**
Pada tahap ini dilakukan penghapusan URL, spasi ganda, angka dan tanda baca (*Punctuation*) kecuali tanda hubung (-). Kata berulang seperti manfaat-manfaat, akan diperlakukan sebagai satu kata.
- b. **Casefolding**
Tahapan mengubah huruf besar menjadi huruf kecil. Proses casefolding akan mempermudah pencarian kata yang sama.
- c. **Tokenisasi**
Tokenisasi adalah langkah merubah kalimat menjadi dalam bentuk kata (*token*). Setiap kata dipisahkan dengan spasi (*white space*).
- d. **Pos tagging**

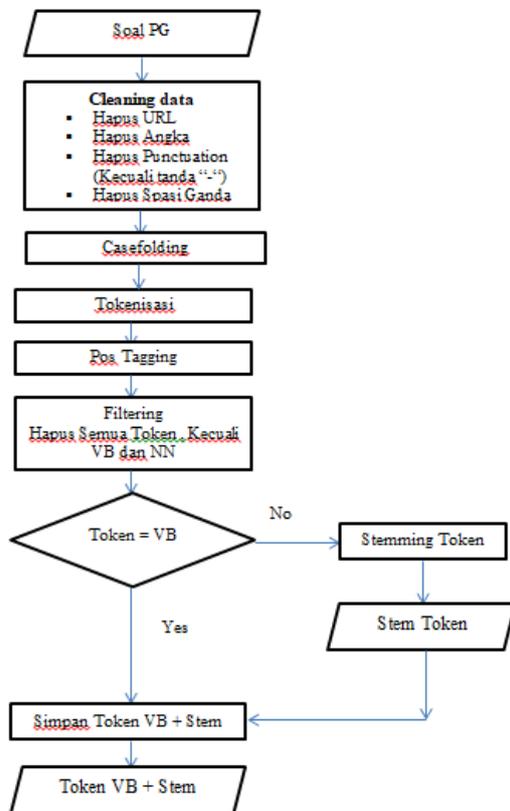
Proses pos tagging merupakan proses pemberian label jenis kata. Contoh : “VB” = kata kerja, “NN” = kata benda, “ADV” = Adverb, IN = preposisi, dan lain-lain. Proses ini diperlukan untuk proses filtering.

e. *Filtering/ Stopword*

Filtering dilakukan untuk menghapus kata-kata yang tidak mempengaruhi klasifikasi. Klasifikasi ini berkaitan dengan penggunaan KKO dan kata yang terkait dengannya, maka seluruh jenis kata kecuali VB dan NN akan dihapus.

f. *Stemming*

Proses stemming akan merubah kata ke dalam bentuk kata dasar (Silva, 2018). Contoh : “penjelasan” setelah di stemming akan menjadi kata “jelas”. Kata hasil proses stemming akan di simpan dan digabung dengan kata kerja.



Gambar 5. Alur Preprocessing

Pada tabel 3, dapat dilihat hasil dari proses preprocessing.

Tabel 3. Hasil Preprocessing

Soal	Hasil Preprocessing
Asap kendaraan bermotor yang berwarna putih disebabkan adanya kebocoran dibagian ring seker. Jika asapnya hitam berarti terlalu banyak campuran bensin daripada udara di dalam karburator. Laporan di atas menjelaskan tentang jenis laporan	[berwarna, disebabkan, berarti, menjelaskan, asap, kendar, motor, ada, bocor, bagi, campur, bensin, udara, dalam, karburator, lapor, atas, jenis, lapor]
Adakalanya sejarah menggunakan ilmu sosial dan sebaliknya ilmu sosial menggunakan sejarah. Artinya, terjadi persilangan antara konsep diakronik dan sinkronik. Penerapan persilangan Konsep tersebut dapat dilakukan pada pilihan ...	[menggunakan, ilmu, menggunakan, terjadi, dilakukan, sejarah, ilmu, sejarah, silang, konsep, sinkronik, terap, silang, konsep, pilih]
Perhatikan karakteristik berikut! 1) Daerah yang berbatasan dengan laut bebas. 2) Wilayah perairan ini diukur sejauh 200 mil dari garis pantai yang paling jauh menjorok ke laut (garis dasar). 3) Pada daerah ini Negara pantai memiliki kedaulatan dan prioritas utama dalam mengelola sumber daya alam yang terkandung di dalamnya. Berdasarkan karakteristik tersebut, dapat disimpulkan bahwa wilayah yang dimaksud adalah...	[perhatikan, berbatasan, diukur, mil, menjorok, memiliki, terkandung, berdasarkan, disimpulkan, dimaksud, adalah, karakteristik, ikut, daerah, laut, wilayah, air, garis, pantai, laut, garis, dasar, daerah, negara, pantai, daulat, prioritas, kelola, sumber, daya, alam, karakteristik, wilayah]

Bag of Words

Bag of word merupakan kumpulan kantung (*bag*) kata yang mewakili dokumen tanpa memandang urutan kata dan tata bahasa tapi tetap mempertahankan keberagamannya (Putri, W.T. H.

2018). Hasil Bag of Word merupakan gabungan dari seluruh hasil preprocessing.

Tabel 4. Hasil Bag of Words

Bag of Words	
berwarna, disebabkan, berarti, menjelaskan, asap, kendaraan, motor, ada, bocor, bagi, campur, bensin, udara, dalam, karburator, lapor, atas, jenis, lapor, menggunakan, ilmu, menggunakan, terjadi, dilakukan, sejarah, ilmu, sejarah, silang, konsep, sinkronik, terap, silang, konsep, pilih, perhatikan, berbatasan, diukur, mil, menjorok, memiliki, terkandung, berdasarkan, disimpulkan, dimaksud, adalah, karakteristik, ikut, daerah, laut, wilayah, air, garis, pantai, laut, garis, dasar, daerah, negara, pantai, daulat, prioritas, kelola, sumber, daya, alam, karakteristik, wilayah	

Pembobotan Term TF-IDF

Pembobotan term menghasilkan tabel bobot masing-masing term, seperti pada tabel 5.

Tabel 5. Pembobotan Term

	abad	abang	abdul	abris	abstrak	abstraksi	acara	ada	adalah
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.051495
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.354163	0.205586
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.186389
...
130	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000
131	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.040033
132	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.076183
133	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.068373
134	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000

Latent Semantic Anlysis

Pada tahap ini dilakukan pengelempokan term yang mempunyai frekwensi tertinggi di masing-masing level soal. Pada tabel 6, merupakan contoh hasil pemrosesan LSA pada soal L1.

Tabel 6. Hasil LSA pada L1

Hasil LSA pada L1
sejarah, peristiwa, lembaga, perintah, adalah, kata, nonkementrian, disebut, teks, erti, bagi, manusia, melaksanakan, negara, tugas, berasal, hasil, contoh, kelompok, hidup, memiliki, merupakan, presiden, penggunaanya, ditunjukkan, berbeda, dibentuk, observasi, masa, berarti, syajaratun, pembabakan, klasifikasi, pohon, berdasarkan, eksposisi, selenggara, waktu, dilakukan, definisi, sebu, gambar, lapor, orang, struktur, dikategorikan, bidang, nilai, paragraf, indonesia, tuju, isi, dicontohkan, berulang, pola, aspek, jalan, utuh, konsep, memberikan, arti, pilih, pemerintah,

nonkemetrian, arsip, awas, sifat, purba, mengubah, kepala,

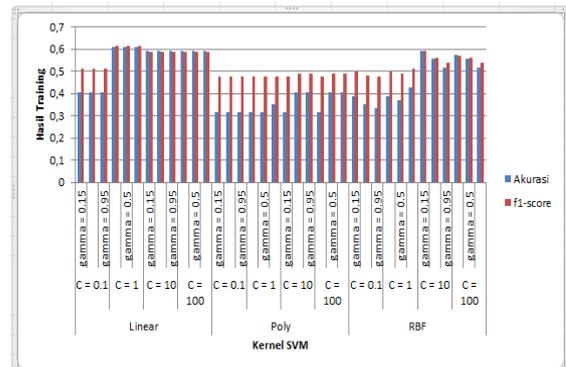
Perhitungan matriks SVD dan Cosine Similarity, seperti pada tabel 7.

Tabel 7. Matriks Cosine Similarity

	0	1	2	3	4
0	1.000000	0.000000	0.000000	0.000000	0.000000
1	0.000000	1.000000	0.065798	0.020031	0.036194
2	0.000000	0.065798	1.000000	0.079971	0.144498
3	0.000000	0.020031	0.079971	1.000000	0.479051
4	0.000000	0.036194	0.144498	0.479051	1.000000

Training Model SVM

Data training yang digunakan dalam ujicoba sebesar 60% dari dataset. Klasifikasi dilakukan dengan menggunakan metode Support Vector Machine dengan 3 kernel, yaitu Linier, Poly dan RBF. Nilai parameter C yang digunakan dalam ujicoba dalam range 0.1; 1; 10; 100, sedangkan nilai γ dalam range 0.15; 0.5 ; 0.95. Hasil training dapat dilihat pada gambar 6.



Gambar 6. Grafik Hasil Training Model SVM

Hasil training dengan menggunakan 60% data training, menghasilkan nilai akurasi paling tinggi di kernel linier dengan nilai C = 1, yaitu dengan akurasi 61,11% dan f1-score 61,52%.

HASIL DAN PEMBAHASAN

Model terbaik yang dihasilkan dari data training, akan diterapkan dalam pengujian data testing. Uji coba dilakukan dengan berulang menggunakan data test 20%, 30% dan 40 %, dengan nilai C =1 pada kernel linear.

Tabel 8 menjelaskan tentang jumlah soal berdasarkan mata pelajaran yang digunakan sebagai data test. Penggunaan data test ini dilakukan secara acak.

Tabel 8. Jumlah Data Testing Berdasarkan Mata Pelajaran

Data Uji	Mata Pelajaran		
	Indo	Sejarah	PKN
20%	8	9	10
30%	13	13	15
40%	13	22	19

Jumlah Data testing berdasarkan Level soal, terlihat pada Tabel 9.

Tabel 9. Jumlah Data Testing Berdasarkan Level Soal

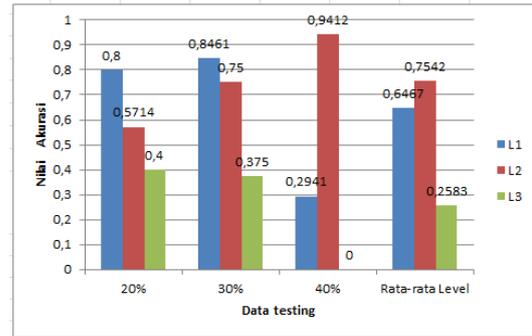
Data Uji	Level		
	L1	L2	L3
20%	10	7	10
30%	13	12	16
40%	17	17	20

Setelah diuji coba dengan kernel linear dan nilai C=1, menghasilkan nilai akurasi dan f1-score seperti pada grafik di bawah ini.



Gambar 7. Grafik Hasil Testing dengan Kernel Linear

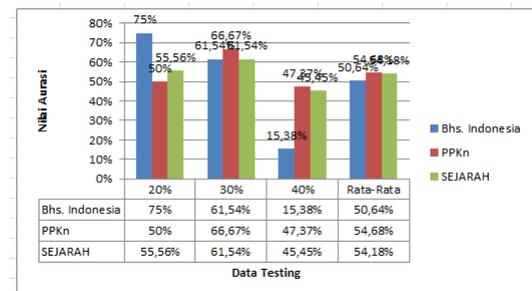
Nilai akurasi dan *f1-score* tertinggi, diperoleh saat ujicoba data testing 30%, yaitu sebesar 63,41% dan 62,13%. Nilai rata-rata dari ujicoba data testing diperoleh nilai akurasi 53,85% dan *f1-score* 50,38%.



Gambar 8. Grafik Hasil Testing Berdasarkan Level Soal

Pada gambar 8, nilai akurasi tertinggi berdasarkan level soal pada L1, bernilai 84,61% saat uji coba data testing 30%. L2 diperoleh nilai akurasi 94,12% saat ujicoba data testing 40%. L3 mempunyai nilai akurasi paling rendah, saat data uji 40% karena tidak ada satupun soal yang bisa terklasifikasi. Akhirnya dari 3 kali ujicoba, rata-rata akurasi L1 adalah 64,67%, L2 75,42% dan L3 25,83%.

Analisa lebih lanjut akurasi hasil ujicoba data testing berdasarkan mata pelajaran, seperti data pada grafik di bawah ini.

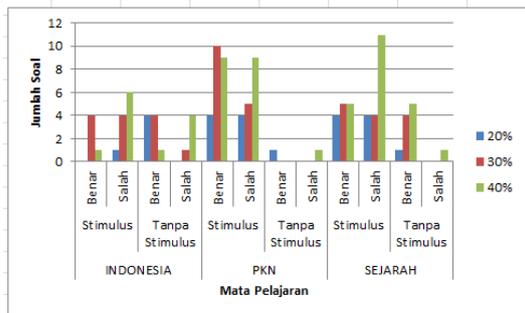


Gambar 9. Grafik Hasil Testing berdasarkan Mata Pelajaran

Soal Bahasa Indonesia pada saat data testing 20%, terdapat 5 soal yang diuji, dimana 4 soal tanpa ada stimulus terklasifikasi dan 1 soal dengan stimulus yang tidak terklasifikasi dengan benar.

Pada testing 30%, terdapat 8 soal dengan stimulus, dan 5 soal tanpa stimulus. Soal Bahasa Indonesia yang berstimulus mempunyai akurasi yang sama antara yang terklasifikasi dan yang tidak terklasifikasi dengan benar. Sedangkan yang tanpa stimulus, 4 soal diantaranya terklasifikasi dengan benar.

Pada saat testing 40%, soal bahasa indonesia yang diuji lebih banyak dari kedua level sebelumnya. Jumlah soal yang berstimulus 7 soal, 6 soal tidak terklasifikasi dengan benar. Sedangkan soal yang tanpa stimulus, 4 soal tidak terklasifikasi dengan benar. Lihat grafik di gambar 10, analisa soal mata pelajaran terkait ada atau tidaknya stimulus, sehingga bisa terklasifikasi dengan benar atau tidak.



Gambar 10. Analisa Soal Mata Pelajaran Berdasarkan Stimulus

Pada grafik di atas terlihat pada soal-soal yang berstimulus, terdapat banyak kesalahan prediksi, dan soal yang tanpa stimulus tingkat kesalahannya jauh lebih kecil.

PENUTUP

Pada penelitian ini klasifikasi soal PG berdasarkan Level Puspendik mempunyai tingkat akurasi L1 64,67%, L2 75,42% dan L3 25,83% dengan menggunakan Support Vector Machine. Tingkat Akurasi klasifikasi paling tinggi pada saat uji coba data testing 30%, yaitu 65,70%. Keberadaan stimulus sangat mempengaruhi hasil prediksi soal.

Berdasarkan analisa hasil testing dan analisa soal, perlu dilakukan penelitian tentang pengaruh stimulus pada klasifikasi soal PG dan penambahan feature extraction lainnya sehingga hasil klasifikasi lebih maksimal.

DAFTAR PUSTAKA

Asmira, O., Ariyana, Y., Bestary, R., Pudjiastuti, A., Setiawati, W. 2018.

Buku Pegangan Pembelajaran Berorientasi pada Keterampilan Berpikir Tingkat Tinggi. Direktorat Jendral Guru dan Tenaga Kependidikan dan Kebudayaan. Jakarta.

Kannan, S., Gurusamy, V. 2014. *Preprocessing Techniques for Text Mining*. https://www.researchgate.net/publication/273127322_Preprocessing_Techniques_for_Text_Mining. Diakses tanggal 14 September 2020

Kusuma, S. F., Siahaan, D., Yuhana, U. L. 2015. *Automatic Indonesia's Questions Classification Based On Bloom's Taxonomy Using Natural Language Processing*. Bandung :International Conference on Information Technology Systems and Innovation (ICITSI). Hal 1-6.

Puspendik. 2019. *Penulisan Soal Pilihan Ganda*. <https://www.slideshare.net/MuslihatunSyarifah/ppt-penulisan-soal-pilihan-ganda>. Diakses tanggal 16 September 2020

Putri, W. T. H., Hendrowati, R. 2018. *Penggalan Teks Dengan Model Bag Of Words terhadap Data Twitter*. Jurnal Muara Sains, Teknologi, Kedokteran, dan Ilmu Kesehatan. Vol 2 No 1. Hal .129-138.

Suhartono, D. 2015. *Penggunaan Latent Semantic Analysis dalam Pemrosesan Text*. <https://socs.binus.ac.id/2015/08/03/penggunaan-latent-semantic-analysis-lsa-dalam-pemrosesan-teks/>. Diakses tanggal 20 September 2020.

Sudin, Sakina. 2018. Analisis Jenis Pertanyaan Berbahasa Indonesia pada Question and Answering System Menggunakan Metode Support Vector Machine (SVM). Jurnal Ilmiah Teknologi Sistem Informasi

Valtemir A. Silva, Ig I. Bittencourt, and José C. Maldonado. 2018. *Automatic Question Classifiers: a Systematic Review*, IEEE Transactions on Learning Technologies. Vol. 12. Hal 485-502.

JUDUL TULISAN : SESUAI DENGAN VARIABEL UTAMA MAKSIMAL 12 KATA (16 TNR)

Penulis ¹⁾, Penulis ²⁾, ... dst

¹Nama Prodi, Nama Fakultas, Nama Perguruan Tinggi
Alamat Perguruan Tinggi
Alamat Email

²Nama Prodi, Nama Fakultas, Nama Perguruan Tinggi
Alamat Perguruan Tinggi
Alamat Email

.... dst

ABSTRAK

Abstrak merupakan ringkasan dari penelitian yang telah dibuat. Panjang abstrak antara 150 sampai dengan 200 kata. Berbentuk 1 paragraf yang berisi latar belakang, tujuan, metode dan hasil penelitian. Font menggunakan Times New Roman 10 poin dan spasi tunggal (1). Kata kunci bersesuaian dengan variabel utama dan berjumlah 3 – 5 buah. Khusus untuk abstrak dalam bahasa Inggris harus dicetak miring semua. Abstrak harus memenuhi fungsinya, sehingga penulis harus memperhatikan ketentuan-ketentuan dalam menulis abstrak termasuk pemilihan kata yang sesuai dengan standar EYD, penyusunan kalimat baku sehingga menjadi paragraf yang komprehensif.

Kata kunci : Kata kunci bersesuaian dengan variabel utama dan berjumlah 3-5 buah, pisahkan dengan koma diantaranya.

ABSTRACT

An abstract is a summary of the research that has been made. Abstract length is between 150 to 200 words. Abstract must be contains in a paragraph that contains the background, objectives, methods and research results. Font Times New Roman size 10 point and single spaced. Keywords and key variables corresponding to numbered 3-5 pieces. An abstract in English must italicized all. In order to fulfill its functions, the authors should consider the terms in writing abstract included the words in the Indonesian language in accordance with the EYD standards, making of raw sentences that make a comprehensive paragraph.

Keywords: *keywords and key variables corresponding to numbered 3-5 pieces, separate them with commas.*

PENDAHULUAN

Pendahuluan menjelaskan latar belakang permasalahan, tinjauan pustaka yaitu dengan mengulas naskah ilmiah yang menimbulkan gagasan dan yang mendasari penelitian yang dilakukan, tujuan penelitian dan batasan masalah, secara singkat dan jelas untuk menghindari kerancuan pengertian pemahaman dan penafsiran makna bagi pembacanya.

Pendahuluan menunjukkan substansi permasalahan yang diangkat artikel serta alasan-alasan yang melatarbelakangi kegiatan

penulisan artikel baik teoritis maupun empiris.

Kalimat-kalimat awal seharusnya merupakan hasil pemikiran sendiri bukan kutipan. Penggunaan dan kembangkan kata-kata kunci sesuai dengan topik dan permasalahannya kemudian rangkai menjadi kalimat-kalimat dengan menggunakan tata bahasa yang baku. Penyajian harus runut antara alinea pertama dengan berikutnya harus jelas. Bagian pendahuluan diuraikan secara deskriptif, font ukuran 11 poin, spasi tunggal.

METODE PENELITIAN

Metode penelitian merupakan prosedur dan teknik penelitian. Bagian metode adalah untuk mendeskripsikan langkah-langkah yang dilakukan untuk memperoleh hasil penelitian. Setiap hasil pengukuran yang dilaporkan harus diketahui metode yang digunakan untuk memperoleh hasil tersebut. Bagian ini juga berisi prosedur dan langkah-langkah bagaimana hasil penelitian dapat diperoleh.

Penjelasan metode harus disertai kutipan yang mengacu pada Daftar Pustaka. Metode terdiri dari rancangan sistem dan prosedur penelitian (dalam bentuk algoritma atau yang lainnya). Bagian ini dapat dibagi menjadi beberapa sub bab, tetapi tidak perlu mencantumkan penomorannya.

HASIL DAN PEMBAHASAN

Bagian ini memuat data (dalam bentuk ringkas) dan analisa data serta interpretasi terhadap hasil pengolahan data. Deskripsikan hasil dengan urutan yang logis dan jelas. Narasi berisi informasi yang disarikan dari data, tidak duplikasi dengan teks.

Pembahasan berdasarkan pengolahan dan analisa data dilakukan dengan mengkaitkan studi empiris atau teori untuk interpretasi. Jika dilihat dari proporsi tulisan, bagian ini harusnya mengambil proporsi terbanyak. Bagian ini bisa dibagi menjadi beberapa sub bab, tetapi tidak perlu mencantumkan penomorannya.

PENUTUP

Penutup umumnya berisi kesimpulan dan saran dari penulis. Kesimpulan berfungsi untuk mengingatkan pembaca akan argumen penting dalam pembahasan, fokus pada tujuan atau pertanyaan penelitian. Tarik semua perincian menjadi pertanyaan umum dari semua argumen.

Saran merupakan penelitian lanjutan yang dirasa masih diperlukan untuk penyempurnaan hasil penelitian supaya berdaya guna. Penelitian tentunya tidak selalu berdaya guna bagi masyarakat dalam satu kali penelitian, tapi merupakan rangkaian penelitian yang berkelanjutan.

DAFTAR PUSTAKA

Bagian ini hanya memuat referensi yang benar-benar dirujuk, dengan demikian, referensi yang dimasukkan pada bagian ini akan ditemukan tertulis pada bagian-bagian sebelumnya. Daftar pustaka ditulis sesuai dengan aturan penulisan. Urutan referensi yang dirujuk disesuaikan dengan huruf abjad serta tidak perlu dikelompokkan berdasarkan jurnal, buku, prosiding atau skripsi/tesis/desertasi.

Sistematika penulisan serta contohnya sebagai berikut :

a. Artikel Jurnal

Nama penulis (nama belakang(.) dan nama depan(.) dipisahkan dengan (,) jika ada penulis berikutnya atau tanpa (,) dan (,) jika dipisahkan dengan dan/and). Tahun publikasi. *Judul Jurnal*. Nama jurnal. Volume Nomor dari jurnal. Hal.

Franses, P., Ghijsel, H. 1999. *Additive Outliers, GARCH and Forecasting. International Journal of Forecasting*. Vol 15 No 1. Hal 1–9.

b. Buku

Penulisan nama seperti diatas (poin a). Tahun publikasi. *Judul buku*. Penerbit. Kota.

Mulyono, Sri. 2002. *Riset Operasi*. LPFE UI. Jakarta.

c. Skripsi/Tesis/Disertasi

Penulisan nama seperti diatas (poin a). Tahun publikasi. *Judul skripsi / tesis / disertasi*. Jenis penelitian dan gelarnya Nama perguruan tinggi. Kota.

Heinzelman, W. 2000. *Application specific protocol architectures for wireless network*. Ph.D dissertation, Massachusetts Institute of Technology. Cambridge.

d. Artikel dalam Prosiding

Penulisan nama seperti diatas (poin a). Tahun publikasi. *Judul artikel/paper*. Nama prosiding diikuti nama dan nomor konferensinya. Kota penyelenggaraan. Penyelenggara seminar. Nomor halaman dalam prosiding.

Dewi, A. R. 2003. *Pengaruh Konservatisme Laporan keuangan Terhadap Earnings Response Coeficient*. Prosiding Simposium Nasional Akuntansi VI Surabaya.

Surabaya: Universitas Airlangga. Hal 119 – 159.

e. Artikel dalam Internet

Penulisan nama seperti diatas (poin a). Tahun publikasi. *Judul artikel/paper*. Alamat website. Tanggal akses.

Suhendra, Maman. 2004. *Evaluasi atas Penerapan Balanced Scorecard sebagai Sistem Pengukuran Kinerja Perusahaan*.

<http://www.balancedscorecard.com>.

diakses tanggal 4 Maret 2014.

f. Artikel dalam Majalah/Koran

Penulisan nama seperti diatas (poin a). Tahun publikasi. *Judul artikel/paper*. Nama Majalah/Koran. Bulan publikasi. Halaman artikel.

Monke, L. 1997. *Infusing technology into a school: Tracking the unintended consequences*. Bulletin of Science, Technology and Society. 4 Februari 1997. Hal 25.

Aturan Penulisan Lainnya:

1. Referensi menggunakan aturan author, hanya mencantumkan nama belakang penulis dan tahun tulisan (contoh: Kotler, 2000) dan mohon dicek ulang dengan daftar pustaka.
2. Tidak menggunakan catatan kaki.
3. Ukuran kertas yang digunakan ukuran A4.
4. Panjang tulisan maksimum 8 halaman.
5. Margin atas dan kiri 3 cm, bawah dan kanan 2.5 cm
6. Spasi 1
7. Dalam bentuk 2 kolom, kecuali judul sampai abstrak 1 kolom dan tabel yang berkapasitas besar.
8. Semua jenis rumus ditulis menggunakan Mathematical Equation (bagi pengguna MS Word ada di bagian Insert => Equation), termasuk pembagian/fraksi, Zigma, Akar, Matriks, Integral, Limit/Log, Pangkat, dsb. Semua jenis simbol menggunakan simbol standar yang ada di Word Processor (bagi pengguna MS Word ada di bagian Insert => Symbol). Persamaan matematika ditulis rata kanan dan diberi nomor secara berurutan.
9. Judul tabel dan gambar ditulis di tengah, title case, dengan jarak 1 spasi dari tabel atau gambarnya. Tulisan "Tabel" atau "Gambar" dengan nomornya diletakkan satu baris sendiri. Judul tabel diletakkan di atas tabel (sebelum tabel) dan judul gambar diletakkan di bawah gambar (setelah gambar). Penulisan sumber tabel atau gambar diletakkan di bawah tabel dan gambar (center pada gambar dan sejajar tabel pada tabel dengan huruf 10 pt). Pada gambar, penulisan sumber diletakkan setelah judul gambar dengan jarak 1 spasi. Tulisan dalam tabel 10 pt.