

ALGORITMA NAIVE BAYES UNTUK Mencari PERKIRAAN WAKTU STUDI MAHASISWA

NAIVE BAYES ALGORITHM FOR FINDING STUDENT ESTIMATED TIME STUDENTS

Yogi Ersan Fadrial
Universitas Lancang Kuning
yogiersan@unilak.ac.id

ABSTRACT

As the development of knowledge and technology has been very rapid, many have used and used automatic equipment for more efficient human work, one of which is using a computerized system. And this is currently needed in all aspects in various government and private agencies. One of them is Data Mining technology, the study program has the task of predicting the length of study of each student in order to determine which students will graduate on time and anticipate the occurrence of students who fail to complete their recovery, which will have an impact on the performance of study programs at the university. , with a good passing rate each year the predicate of the passing score is very satisfying, Each Study Program is obliged to see the development of student studies. Predicting can be done by means of data mining using the Naive Bayes Algorithm method.

Keywords: *Data Mining, Naive Bayes Algorithm, Estimated Time Student.*

ABSTRAK

Seiring perkembangan pengetahuan dan teknologi sudah sangat pesat ini, banyak yang memakai dan menggunakan peralatan secara otomatis untuk kerja manusia lebih efisien, yang salah satunya menggunakan sistem komputerisasi. Dan hal tersebut saat ini sangat dibutuhkan dalam segala aspek diberbagai instansi pemerintah maupun swasta. Salah satunya adalah teknologi Data Mining, program studi mempunyai tugas untuk memprediksi lama studi dari setiap mahasiswanya guna menentukan mahasiswa mana yang lulus tepat pada waktunya dan mengantisipasi terjadinya mahasiswa yang gagal dalam menyelesaikan perkuliahan, yang nantinya akan berdampak pada kinerja dari program studi yang ada di universitas, dengan tingkat kelulusan baik pada setiap tahunnya predikat nilai kelulusan sangat memuaskan, Setiap Program Studi berkewajiban untuk melihat perkembangan studi mahasiswanya, untuk memprediksi dapat dilakukan dengan data mining menggunakan metode Algoritma Naive Bayes.

Kata Kunci: *Data Mining, Naive Bayes Algorithm, Perkiraan Masa Studi.*

PENDAHULUAN

Beriringan perkembangan pengetahuan serta teknologi ini sangat cepat, sudah banyak yang memakai dan menggunakan dengan teknik otomatis, untuk aktivitas manusia lebih efisien yang salah satunya menggunakan sistem komputerisasi, dan mengenai itu dibutuhkan dalam semua perspektif diberbagai lembaga pemerintah maupun swasta, salah satunya ialah teknologi *Data Mining*, kegiatan yang melingkupi pengumpulan informasi, historis yang menghasilkan keteraturan pola dan hubungan dalam set data berukuran

besar, tujuan dari penafsiran ini adalah metode pencarian data yang tidak dikenal sebelumnya dari sekumpulan data besar. (Romero & Ventura, 2020).

Fakultas Ilmu Komputer Universitas Lancang Kuning ialah fakultas terfavorit diantara fakultas-fakultas yang terdapat di Universitas Lancang Kuning disaat ini, dengan tingkatan kelulusan yang bagus pada tiap tahunnya predikat angka kelulusan sungguh memadai, tiap Program Studi bertanggung jawab buat memantau kemajuan studi, setelah itu memiliki kewajiban buat memperkirakan lama

setiap mahasiswa untuk memastikan mana yang tercapai pas pada waktu, serta mengestimasi terbentuknya mahasiswa yang gagal dalam menuntaskan pembelajaran yang kemudian bakal berakibat pada prestasi dari program studi tersebut, buat memprediksi bisa dicoba dengan memakai data mining, menggunakan metode *Algoritma Naive Bayes* (Peling, dkk., 2017; Azahari, dkk., 2020; Widaningih, 2019).

Sukses tidaknya keahlian akademik Mahasiswa pada durasi studi salah satu dapat dicermati dari ketepatannya dalam menuntaskan waktu studi di universitas, Mengenai ini dapat dipantau dari mereka berada pada tingkatan 3 sampai 6, evaluasi dicoba dengan memisah data jadi 2 ialah selaku data latih (*training set*) yang diterima dari data histori alumni 150, dan juga (*testing set*) dari informasi mahasiswa di semester 3 atau 6 yang belum mengerjakan sebesar 61 data, Hasil dari studi dengan pembangunan model Naïve Bayes terbaik terdapat pada akurasi 76%, sesudah itu struktur digunakan untuk perkiraan kelulusan mahasiswa pada tingkatan 3, dengan hasil 48 mahasiswa berhasil pas waktu dan 13 mahasiswa lulus tidak tepat waktu. Pada penelitian ini karakteristik yang digunakan ialah jurusan, asal sekolah, area asal, tempat lahir, indikator hasil kuliah 1 hingga dengan 5, indikator hasil kumulatif, serta sasaran kelulusan (Nurliana, dkk., 2015).

Penelitian ini mengenakan metode Algoritma Naive Bayes untuk melakukan perkiraan kelulusan mahasiswa, kelulusan yang dimaksud ialah diterimanya seorang mahasiswa pada salah satu program studi di perguruan tinggi, agar mahasiswa tersebut dapat lulus pada periodenya (Muin, 2016).

Kasus yang kerap terjalin dalam industri asuransi merupakan banyaknya klien yang terlambat melunasi angsuran,

akhirnya dibutuhkan suatu sistem yang bisa mengklasifikasikan klien yang mana masuk kedalam jenis golongan lancar, kurang lancar serta yang tidak lancar, dalam pembayaran angsuran, pihak asuransi dapat menanggulangi semenjak dini kasus yang berlangsung (Bustami, 2013).

Listrik sangat penting bagi setiap lapisan masyarakat bahkan juga sangat dibutuhkan, sebagai sarana produksi dan untuk kehidupan sehari-hari, begitu penting tentu saja berdampak pada permintaan listrik yang semakin besar, tapi hal ini kiranya tidak merata dengan persediaan listrik yang belum mampu memenuhi permintaan konsumen yang begitu besar, untuk mengantisipasi hal tersebut perlu adanya bantuan dari pemerintah dan masyarakat dalam menggunakan listrik dengan bijak sehingga kebutuhan listrik tidak menjadi lebih besar dari persediaan, oleh karena itu setiap rumah haruslah paham cara penggunaan yang lebih efektif, penerapan metode naïve bayes diharapkan mampu untuk memprediksi besarnya penggunaan tiap rumah, agar lebih mudah mengatur penggunaan listrik. dari 60 rumah data penggunaan listrik, yang diuji dengan metode naïve bayes, maka diperoleh hasil persentase 78,3333% untuk keakuratan prediksi, di mana dari 60 data penggunaan listrik yang diuji, terdapat 47 data penggunaan yang berhasil diklasifikasikan dengan benar (Saleh, 2015).

Tujuan penelitian ini adalah untuk memahami, menganalisa dan menerangkan secara ilmiah hasil dari pengujian yang dilakukan serta mengetahui tingkat akurasi dari teknik penambangan data yang digunakan dalam percobaan (As'ad, 2016).

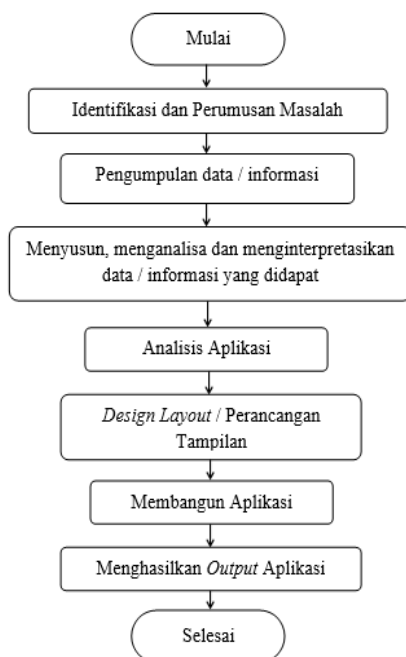
Bersumber pada hasil dari kuisioner yang sudah disebar pada mahasiswa sistem informasi, hendak dicoba

pengelompokan dari tingkatan bisa jadi kegemukan yang diakibatkan oleh gaya hidup yang tidak sehat. Bagi penelitian dari aspek pemicu dari kegemukan pada anak muda sungguh kompleks antara lain merupakan kurang tidur, banyak menghabiskan waktu pada aktivitas yang karakternya statis, pola makan yang tidak teratur, hasil riset nya melaporkan kalau pada umumnya anak muda kegemukan sebesar 10, 5% dari anak didik yang dijadikan subjek riset(Muslehatin, dkk., 2017).

Berdasarkan pandangan penelitian terdahulu, dapat disimpulkan teknik Data Mining dengan menggunakan Klasifikasi *Algoritma Naive Bayes* adalah sebuah sistem yang ditujukan untuk memprediksi suatu kemungkinan baik ataupun buruk serta mengelompokkan atau mengklasifikasikan masing-masing dari objek penelitian tersebut(Berrar, 2018).

METODE

Adapun metode penelitian ini adalah sebagai berikut:



Gambar 1. Kerangka Kerja

Pendekatan yang hendak dipakai dalam riset ini merupakan *System Development Live Cycle (SDLC)* dengan bentuk *waterfall*, bentuknya adalah pengembangan perangkat lunak sebagai bertingkat dengan menuntaskan tahapan-tahapan tadinya supaya dapat lanjut ke jenjang berikutnya, pada penelitian ini langkah mula adalah pengenalan permasalahan dengan mencari masukan yang diawasi lewat observasi, peneliti merumuskan jika bagaimana mendesain sesuatu aplikasi data mining dibuat memprediksi lama studi mahasiswa dengan memanfaatkan *Algoritma Naive Bayes*(Arar & Ayan, 2017).

Pengumpulan Informasi, langkah yang sistematis serta standar untuk mendapatkan data yang diinginkan.

Menyusun serta menganalisa informasi yang digabungkan butuh dikelompokkan maupun disusun maka data itu memiliki makna, sesudah berakhir dianalisis sehingga diberikan pengertian agar jelas keterangan yang digunakan.

Analisa aplikasi, dicoba dengan cara pengamatan langsung di tempat penelitian serta tanya jawab kepada subjek-subjek penelitian, dimulai dengan analisa kepada sistem objek penelitian, aktivitas ini bermaksud buat melihat serta menguasai kegiatan sistem yang sedang berjalan serta permasalahan yang lagi dialami.

Design Layout ataupun perencanaan bentuk, Membuat serta pengaturan isi yang terdapat di dalam aplikasi.

Membuat aplikasi, Pada bagian ini dicoba implementasi dari design layout yang sudah dicoba, alhasil pada langkah ini hendak menciptakan aplikasi yang cocok dengan keinginan industri serta yang sangat penting bisa menanggulangi kasus yang lagi dialami industri.

Output aplikasi peneliti hendak menciptakan sesuatu aplikasi data mining buat mencari perkiraan masa

studi mahasiswa yang bisa dipakai cocok dengan keperluan.

Data Yang dikenakan sebagai berikut:

Informasi Primer yang didapat ataupun digabungkan dengan cara langsung dari sumber aslinya, yang di dapat berbentuk pemantauan langsung ditepat penelitian Fakultas Ilmu Komputer Universitas Lancang Kuning serta tanya jawab dengan Pimpinan Program Studi.

Informasi Sekunder merupakan sumber informasi riset yang didapat lewat media perantara ataupun dengan cara tidak langsung yang berbentuk buku, catatan, fakta yang pernah ada, maupun arsip baik yang diterbitkan ataupun yang tidak diterbitkan. Dengan kata lain, peneliti menginginkan pengumpulan data dengan berkunjung ke perpustakaan, pusat analisis, pusat arsip ataupun membaca banyak buku yang berkaitan dengan penelitiannya.

Tools yang digunakan dalam pembuatan sistemnya menggunakan *software PHP, MySQL, Macromedia Dreamweaver*.

HASIL DAN PEMBAHASAN

Pengelompokan memanfaatkan Metode *Naive Bayes* untuk memastikan tingkatan kelulusan mahasiswa bisa dicoba dengan cara mengakulasi data kuesioner yang telah melalui waktu perkuliahan selama 6 semester, maksudnya mahasiswa itu telah memasuki semester 7 ataupun lebih, pada aplikasi ini diperlukan Data statistik serta probabilitas untuk sebagian kemungkinan yang jadi faktor bernilai dalam perhitungan Klasifikasi memakai Metode *Naive Bayes*, pada aplikasi pastinya mahasiswa bisa mengisi sendiri pertanyaan dari kuesioner serta langsung bisa memastikan apakah mahasiswa itu lulus tepat waktu ataupun tidak, pada aplikasi serta terdapat penjelasan yang berisikan

data- data mahasiswa yang telah mengisi kuesioner serta memperoleh hasil dari analisa dengan pengelompokan ataupun Klasifikasi memanfaatkan Metode *Naive Bayes*.

Menghitung nilai probabilitas setiap jawaban dengan rumus sebagai berikut :

Diketahui :

P = Peluang

Qi = Pertanyaan

Ci = Hipotesis/jawaban sementara (Status Lulus)

Maka :

$P(C_i)$

$$= P(C_i = \text{“Tepat”}) = 31/50$$

$$= P(C_i = \text{“Terlambat”}) = 19/50$$

$P(Q_1|C_i)$

$$= P(Q_1 = \text{“Y”} | C_i = \text{“Tepat”}) = 22/31 = \underline{0,71}$$

$$= P(Q_1 = \text{“Y”} | C_i = \text{“Terlambat”}) = 6/19 = \underline{0,32}$$

$$= P(Q_1 = \text{“T”} | C_i = \text{“Tepat”}) = 9/31 = \underline{0,29}$$

$$= P(Q_1 = \text{“T”} | C_i = \text{“Terlambat”}) = 13/19 = \underline{0,68}$$

$P(Q_2|C_i)$

$$= P(Q_2 = \text{“Y”} | C_i = \text{“Tepat”}) = 14/31 = \underline{0,45}$$

$$= P(Q_2 = \text{“Y”} | C_i = \text{“Terlambat”}) = 4/19 = \underline{0,2}$$

$$= P(Q_2 = \text{“T”} | C_i = \text{“Tepat”}) = 17/31 = \underline{0,55}$$

$$= P(Q_2 = \text{“T”} | C_i = \text{“Terlambat”}) = 15/19 = \underline{0,79}$$

Merepkan perhitungan pada studi kasus, pada penerapan ini dapat diketahui apakah mahasiswa lulus tepat waktu ataupun terlambat.

Untuk mendapatkan hasil dari tingkat kelulusan mahasiswa, maka ditentukan dengan rumus berikut:

a. Setiap nilai “Ya” dan “Tidak” pada tingkat kelulusan “Tepat” dikalikan

masing-masing pada nilai probabilitasnya. Maka :

$$= (Q1 \times Q2 \times Q3 \times Q4 \times Q5 \times Q6 \times Q7 \times Q8 \times Q9 \times Q10 \times Q11 \times Q12 \times Q13 \times Q14)$$

$$= (0.71 \times 0.55 \times 0.39 \times 0.45 \times 0.61 \times 0.42 \times 0.55 \times 0.55 \times 0.39 \times 0.35 \times 0.61 \times 0.68 \times 0.48 \times 0.48)$$

$$= 6.92877e^{-5}$$

b. Setiap nilai “Ya” dan “Tidak” pada tingkat kelulusan “Terlambat” dikalikan masing-masing pada nilai probabilitasnya. Maka :

$$= (Q1 \times Q2 \times Q3 \times Q4 \times Q5 \times Q6 \times Q7 \times Q8 \times Q9 \times Q10 \times Q11 \times Q12 \times Q13 \times Q14)$$

$$= (0.32 \times 0.79 \times 0.26 \times 0.53 \times 0.16 \times 0.58 \times 0.37 \times 0.53 \times 0.21 \times 0.63 \times 0.11 \times 0.11 \times 0.63 \times 0.21)$$

$$= 1.34263e^{-7}$$

Tabel 1. Hasil Perhitungan Studi Kasus

Nama	Hasil Perhitungan (Perkalian setian Qn)	Prediksi
A		
TEPAT	$6.92877e^{-5}$	Tepat Waktu
TERLAMBAT	$1.34263e^{-7}$	

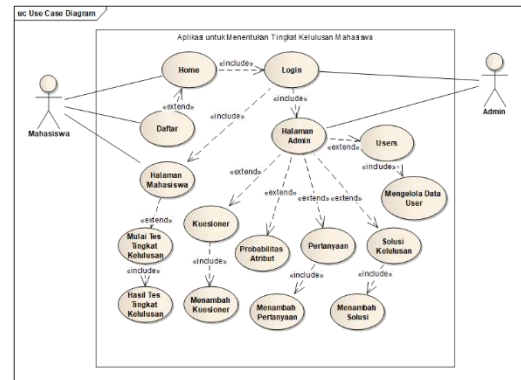
Jika Total nilai perhitungan tingkat kelulusan “Tepat” lebih tinggi daripada “Terlambat” maka prediksinya “Tepat Waktu”, begitu juga sebaliknya.

1. UML (Unifed Modeling Language)

UML (Unified Modelling Language) adalah metode pemodelan secara visual sebagai sarana untuk merancang dan membuat *software* berorientasi objek. Karena UML ini merupakan bahasa visual untuk pemodelan bahasa berorientasi objek, maka semua elemen dan diagram berbasiskan pada paradigma *object oriented*.

a. Use Case Diagram

Use case diagram adalah menggambarkan fungsionalitas yang diharapkan dari sebuah sistem, sebuah use case merepresentasikan interaksi antara actor dengan sistem. Berikut adalah pemodelan Use case dari aplikasi untuk menentukan tingkat kelulusan mahasiswa :



Gambar 2. Use Case Diagram

Use Case Diagram menggambarkan apa saja peran dari admin dan pengguna, dimana pengguna sebagai yang menggunakan sistem dan admin sebagai pengelola sistem.

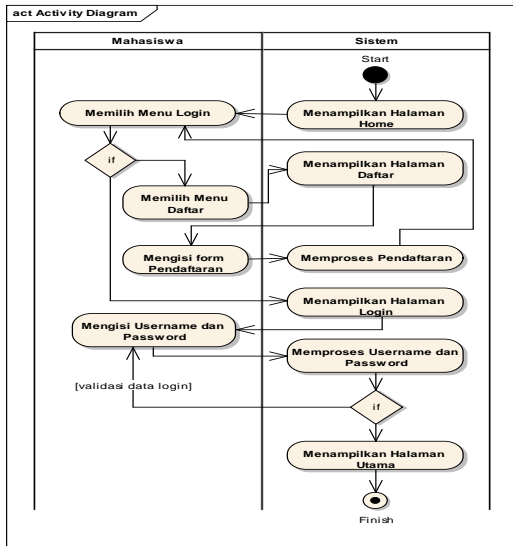
b. Activity Diagram

Activity diagram menggambarkan berbagai alur aktifitas dalam sistem yang sedang dirancang, bagaimana masing-masing alur berawal, *decision* yang mungkin terjadi, dan bagaimana mereka berakhir.

Berikut merupakan beberapa aktifitas yang terjadi pada sistem yang sedang dirancang :

1. Activity Diagram Login Mahasiswa

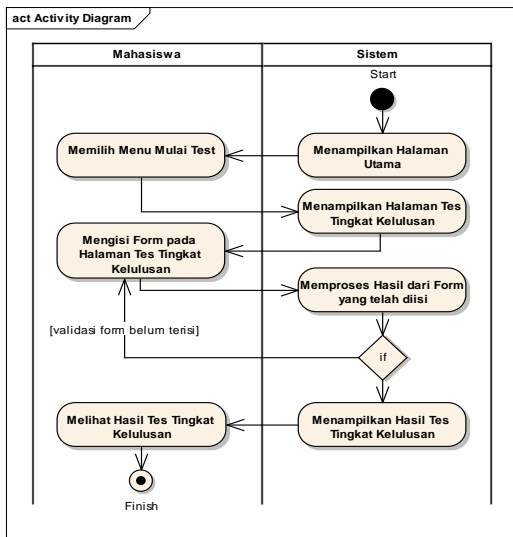
Activity Diagram Login Mahasiswa menggambarkan seorang Mahasiswa untuk masuk ke halaman utama agar dapat menggunakan aplikasi.



Gambar 3. Activity Diagram Login Mahasiswa

2. Activity Diagram Memulai Tes Tingkat Kelulusan

Activity Diagram Memulai Tes Tingkat Kelulusan merupakan gambaran bahwa dari Mahasiswa untuk melakukan tes tingkat kelulusan dan langsung dapat melihat hasil tes tingkat kelulusan mahasiswa.

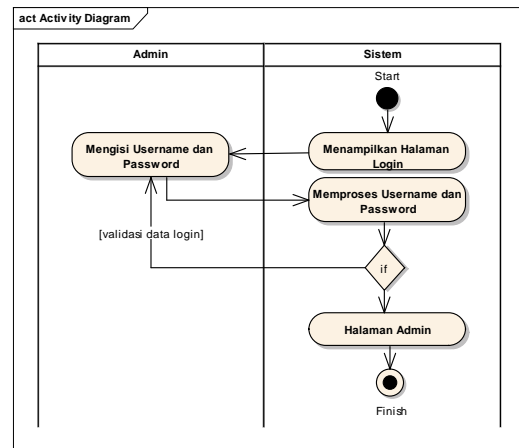


Gambar 4. Activity Diagram Memulai Tes Tingkat Kelulusan

3. Activity Diagram Login Admin

Activity Diagram Login Admin merupakan gambaran bahwa dari admin

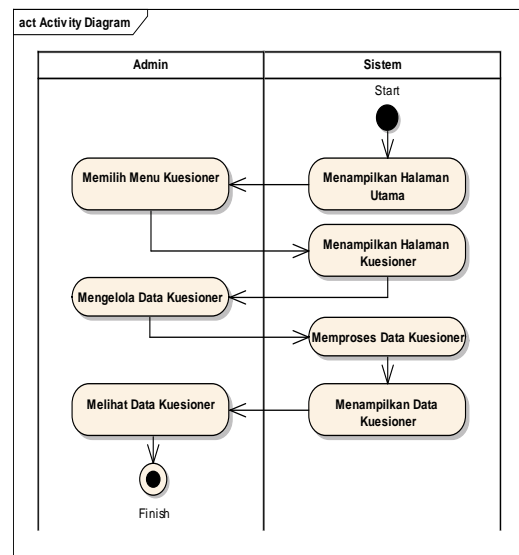
untuk masuk kehalaman utama harus melakukan login terlebih dahulu.



Gambar. 5 Activity Diagram Login Admin

4. Activity Diagram Mengelola Data Kuesioner

Activity Diagram Mengelola Data Kuesioner menggambarkan seorang admin untuk masuk kehalaman utama dan memilih menu Kuesioner untuk dapat mengelola dan melihat data-data kuesioner.

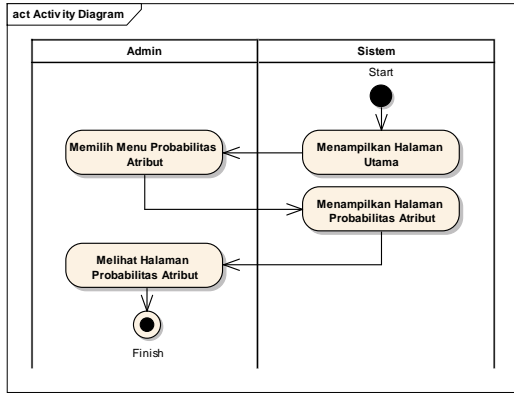


Gambar 6. Activity Diagram Mengelola Data Kuesioner

5. Activity Diagram Melihat Probabilitas Atribut

Activity Diagram Melihat Probabilitas Atribut merupakan

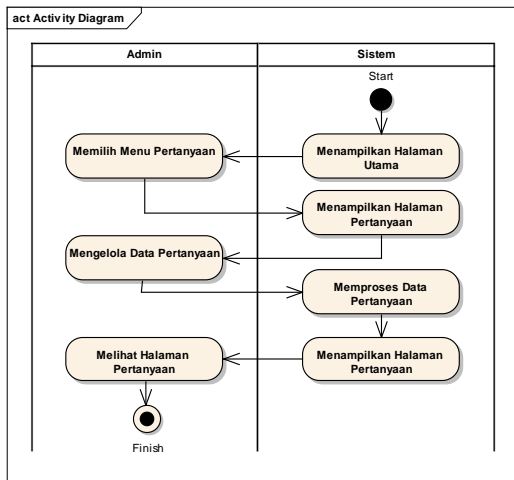
gambaran bahwa dari Admin untuk masuk kehalaman utama dan memilih menu Probabilitas Atribut untuk dapat melihat hasil perhitungan probabilitas dari beberapa data yang sudah diinputkan kehalaman Kuesioner.



Gambar. 7 Activity Diagram Melihat Probabilitas Atribut

6. Activity Diagram Mengelola Data Pertanyaan

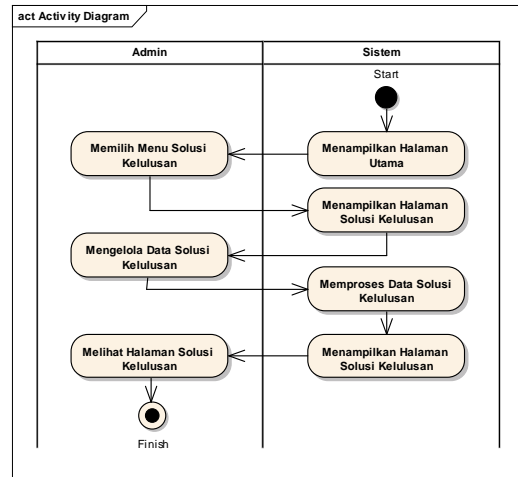
Activity Diagram Mengelola Data Pertanyaan merupakan gambaran bahwa dari Admin untuk masuk kehalaman utama dan memilih menu Pertanyaan untuk dapat mengelola dan melihat Data Pertanyaan untuk Tes Tingkat Kelulusan Mahasiswa.



Gambar. 8 Activity Diagram Mengelola Data Pertanyaan

7. Activity Diagram Mengelola Data Solusi Kelulusan

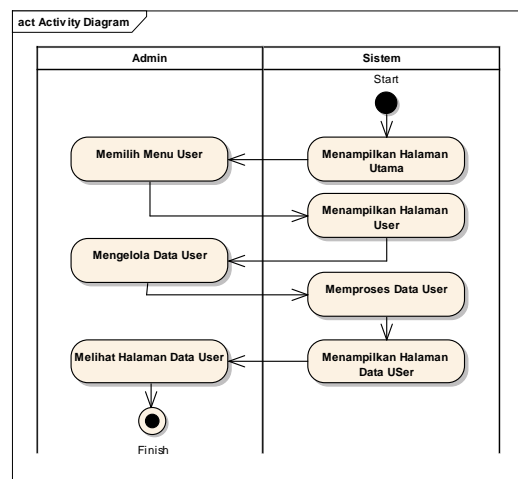
Activity Diagram Mengelola Data Solusi Kelulusan gambaran bahwa dari Admin untuk masuk kehalaman utama dan memilih menu Solusi Kelulusan untuk dapat mengelola dan melihat Data Solusi Kelulusan yang akan muncul di hasil Tes Tingkat Kelulusan Mahasiswa.



Gambar. 9 Activity Diagram Mengelola Data Solusi Kelulusan

8. Activity Diagram Mengelola Data User

Activity Diagram Mengelola Data User gambaran bahwa dari Admin untuk masuk kehalaman utama dan memilih menu User untuk dapat mengelola dan melihat Data User.

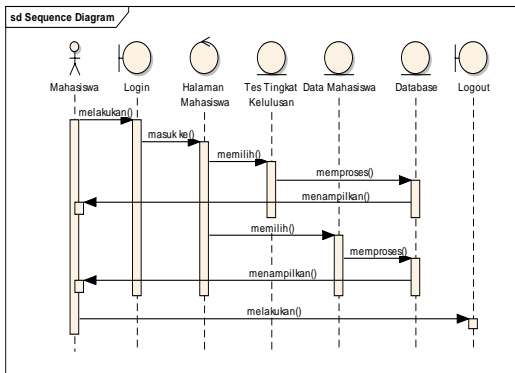


Gambar 10 Activity Diagram Mengelola Data User

c. Sequence Diagram

Sequence Diagram menggambarkan urutan aktifitas yang dilakukan pengguna aplikasi yaitu admin dan mahasiswa dalam langkah-langkah penggunaan aplikasi untuk menentukan tingkat kelulusan mahasiswa.

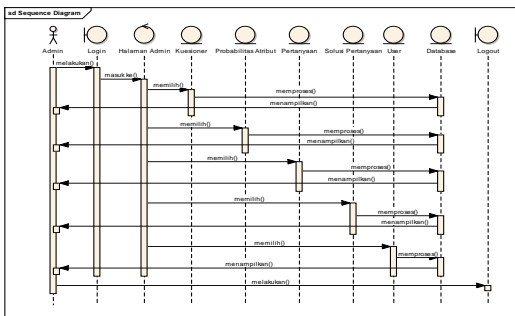
1) Sequence Diagram Mahasiswa



Gambar .11 Sequence Diagram Mahasiswa

Gambar diatas menjelaskan bahwa Mahasiswa terlebih dahulu melakukan login untuk dapat menggunakan aplikasi tersebut.

2) Sequence Diagram Admin

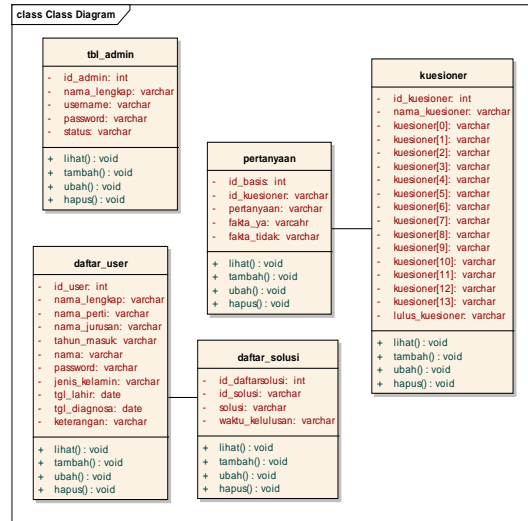


Gambar. 12 Sequence Diagram Admin

Gambar diatas menjelaskan bahwa Admin terlebih dahulu melakukan login untuk dapat mengelola aplikasi tersebut.

d. Class Diagram

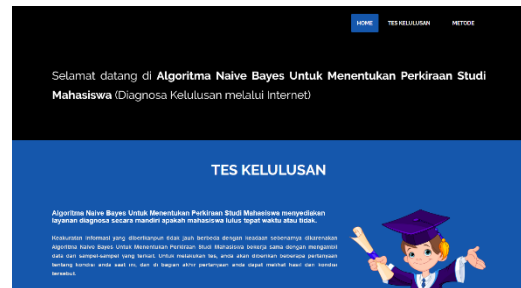
Class adalah sebuah spesifikasi yang diinstansiasi akan menghasilkan sebuah objek dan merupakan inti dari pengembangan dan desain berorientasi objek. Class menggambarkan keadaan (atribut/properti) suatu sistem.



Gambar. 13 Class Diagram

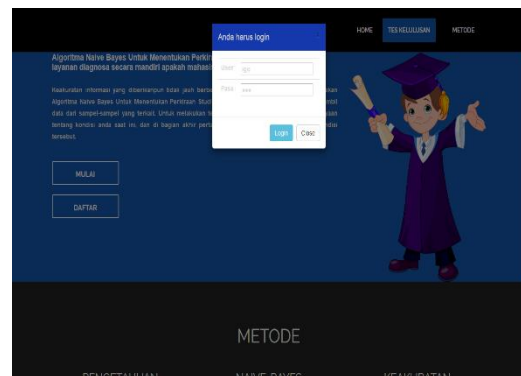
Class Diagram dimana dari gambar tersebut dapat dilihat relasi antar class yang saling berhubungan.

9. Implementasi Tampilan Program
a. Halaman Utama

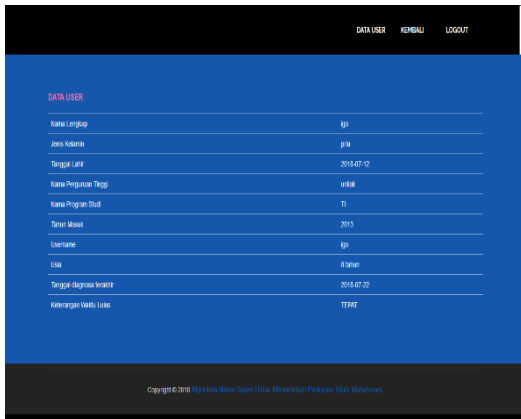


Gambar. 13 Tampilan Halaman Utama

b. Login User

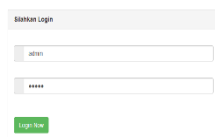


Gambar. 14 Tampilan Halaman Login User
c. Data User



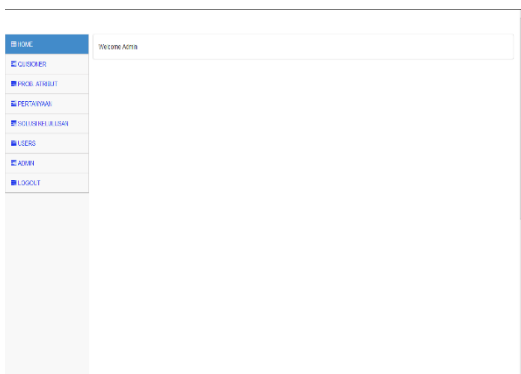
Gambar. 15 Tampilan Halaman Data User

d. Login Admin



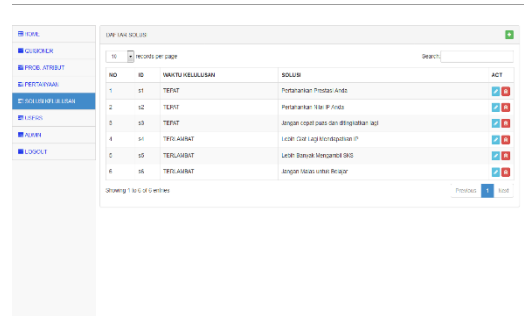
Gambar. 16 Tampilan Halaman Login Admin

e. Utama Admin



Gambar. 17 Tampilan Halaman Utama Admin

f. Solusi Kelulusan



Gambar. 18 Tampilan Halaman Menu Solusi Kelulusan.

SIMPULAN

Aplikasi yang sudah dibuat bermaksud untuk mengenali prediksi lama studi dari tiap mahasiswa yang sudah menempuh kuliah sampai minimal semester VI dengan mengenakan data nilai matakuliah yang telah tercapai sebelumnya dan jumlah keseluruhan SKS yang telah didapat. Aspek perkiraan lama studi mencakup data nilai-nilai matakuliah yang telah ditempuh sampai semester VI dan juga jumlah SKS yang didapat hingga semester VI serta faktor data induk berbentuk pembelajaran terakhir dari orang tua serta rute masuk perguruan tinggi apakah jalur prestasi maupun regular. Kalkulasi yang dipakai yakni Klasifikasi dengan memanfaatkan Algoritma Naive Bayes. Output yang diperoleh pada aplikasi berbentuk informasi tingkatan kelulusan dari mahasiswa yang telah melaksanakan pengecekan kelulusan

DAFTAR PUSTAKA

- Arar, Ö. F., & Ayan, K. (2017). A feature dependent Naive Bayes approach and its application to the software defect prediction problem. *Applied Soft Computing*, 59, 197-209.
- As' ad, B. (2016). Prediksi Kehadiran Menggunakan Metode Klasifikasi Naïve Bayes, One-r, Decision Tree. *Jurnal Penelitian*

- Komunikasi dan Opini Publik*, 20(1).
- Azahari, A., Yulindawati, Y., Rosita, D., & Mallala, S. (2020). Komparasi Data Mining Naive Bayes dan Neural Network memprediksi Masa Studi Mahasiswa S1. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(3), 443-452.
- Berrar, D. (2018). Bayes' theorem and naive Bayes classifier. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics; Elsevier Science Publisher: Amsterdam, The Netherlands*, 403-412.
- Muin, A. A. (2016). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, 2(1), 22-26.
- Muslehatin, W., Ibnu, M., & Mustakim, M. (2017, May). Penerapan Naive Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau. In *Seminar Nasional Teknologi Informasi Komunikasi dan Industri* (pp. 250-256).
- Nasution, N., Djahara, K., & Zamsuri, A. (2015). Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes (Studi Kasus: Fasilkom Unilak). *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, 6(2), 1-11.
- Peling, I. B. A., Arnawan, I. N., Arthawan, I. P. A., & Janardana, I. G. N. (2017). Implementation of Data Mining To Predict Period of Students Study Using Naive Bayes Algorithm. *International Journal of Engineering and Emerging Technology*, 2(1), 53-57.
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355.
- Saleh, A. (2015). Implementasi metode klasifikasi naive bayes dalam memprediksi besarnya penggunaan listrik rumah tangga. *Creative Information Technology Journal*, 2(3), 207-217.
- Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4, 5, Naive Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16-25.