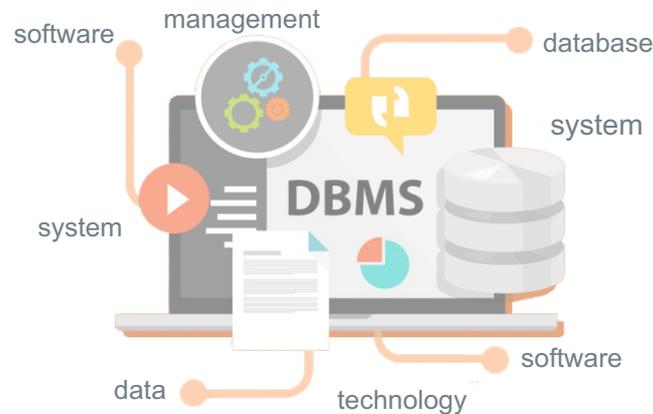


# **JURNAL SIMADA**

## **Sistem Informasi & Manajemen Basis Data**



### **Darmajaya Street View Menggunakan Teknologi Virtual Reality**

*Triowali Rosandi, Lia Rosmalia, M. Yajid Alfian*

### **Sistem Informasi Try Out Berbasis Web Pada SMK Darul Huda Campang Raya Sukabumi Bandar Lampung**

*Deni Impantri, Halimah*

### **Sistem Informasi Catatan Pelanggaran Berbasis Android (Studi Kasus : MTS Nurul Huda Madukoro)**

*Ferly Ardhy, Rustam*

### **Monitoring dan Evaluasi Tata Kelola Pelayanan Siswa Bimbel No Name Menggunakan Metode Cobit 4.1**

*Rini Nurlistiani, Kintan Imanita, Neni Purwati, Indera*

### **Sistem Informasi Keuangan Desa Pada Desa Tanjung Baru**

*Kurniasih, Satria Novari, Dian Sri Agustina*

### **Prediksi Kelulusan dan Putus Studi Mahasiswa dengan Pendekatan Bertingkat Pada Perguruan Tinggi**

*Hermanto*

### **Rencana Strategik Sistem Informasi dan Teknologi Informasi Akademik (Studi Kasus : Akademik Teknologi Pringsewu)**

*Alfian Nuziar, Abdi Darmawan*



Institut Informatika & Bisnis  
**DARMAJAYA**  
Yayasan Alfian Husin

**Pelindung**

Sriyanto, S.Kom., MM

**Pimpinan Redaksi**

Dr. Suhendro Yusuf Irianto, M.Kom

**Redaksi Pelaksana**

Fitria M.Kom

Rio Kurniawan, M.Cs

Yulmaini, S.Kom., M.Cs

**Editor Ahli (Mitra Bestari)**

Dr. Arta Moro Sundjaja (Univeristas Bina Nusantara)

DR. Deris Setiawan (Univetsitas Sriwijaya)

DR. Hustinawaty (Universitas Gunadarma)

Ramadiani, M.Kom., Ph.D (Universitas Mulawarman)

DR. Syifaun Nafisyah (UIN Sunan Kalijaga Yogyakarta)

**Editor Ahli**

Dr. Suhendro Yusuf Irianto, M.Kom

Dr. RZ. Abdul Aziz, ST., M.T

Joko Triloka, M.T., Ph.D

Dr. Sutedi, S.Kom., M.T.I

**Dewan Editor**

Hendra Kurniawan, S.Kom., M.T.I

Melda Agarina, S.Kom., M.T.I

Sri Karnila, S.Kom., M.Kom

Nurjoko, S.Kom., M.T.I

**Editor/Layout**

Dwi Lianiko, S.Kom

Febrian Eka Saputra, S.Kom

**Kesekretariatan**

Dona Yuliawati, S.Kom., M.T.I

Sushanty Saleh, S.Kom., M.T.I

Arman Suryadi Karim, S.Kom., M.T.I

**Bendahara**

Halimah, S.Kom., M.T.I

Ochi Marshella F, S.Kom., M.T.I

## **PENGANTAR REDAKSI**

Puji Syukur kehadiran Allah SWT, atas karunia dan rahmatnya sehingga Jurnal Ilmiah Sistem Informasi dan Manajemen Basis Data (SIMADA) Volume 3 Nomor 2 bulan Oktober 2020 dapat diterbitkan sesuai dengan periode yang telah ditetapkan.

Jurnal Sistem Informasi dan Manajemen Basis Data (SIMADA) merupakan jurnal yang diterbitkan oleh Jurusan Sistem Informasi Institut Informatika dan Bisnis (IIB) Darmajaya. Penerbitan jurnal ini sebagai wadah informasi berupa hasil penelitian, studi kepustakaan, gagasan, aplikasi teori dan kajian analisis kritis di bidang keilmuan Sistem Informasi dan Manajemen Basis Data.

Pada edisi ini terdapat 8 artikel dimana versi online dari jurnal tersebut dapat dilihat di [jurnal.darmajaya.ac.id](http://jurnal.darmajaya.ac.id). Kami ucapkan terima kasih banyak kepada semua pihak yang telah memberikan kontribusi dalam volume jurnal ini. Pada kesempatan ini kami kembali mengundang dan memberikan kesempatan kepada para peneliti, dibidang Sistem Informasi dan Manajemen Basis Data untuk kembali mempercayai jurnal SIMADA sebagai wadah bagi para peneliti dalam mempublikasikan hasil penelitiannya dalam jurnal ini.

Akhir kata redaksi berharap agar makalah dalam jurnal ini dapat memberikan kontribusi dan sumbangsih pemikiran yang bermanfaat dalam menjawab tantangan yang dihadapi khususnya bagi perkembangan ilmu dan teknologi dalam bidang Sistem Informasi dan Manajemen Basis Data.

Bandar Lampung, 13 Oktober 2020  
Redaksi Jurnal Simada

# Prediksi Kelulusan dan Putus Studi Mahasiswa dengan Pendekatan Bertingkat pada Perguruan Tinggi

Hermanto

UIN Raden Intan Lampung

Jl. Letnan Kolonel H Jl. Endro Suratmin, Sukarame, Kec. Sukarame, Kota Bandar Lampung, Lampung 35131

hermanto@radenintan.ac.id

## Abstract

Currently, the problem of college failure, its on-time graduation, and the factors that cause it is still an interesting research topic (C. Marquez-Vera, C. Romero and S. Ventura, 2011). This study compares three data mining classification algorithms namely Naive Bayes, Decision Tree and K-Nearest Neighbor to predict graduation and dropout risk for students to improve the quality of higher education and the most accurate algorithms to use Prepare graduation and dropout prediction Student studies. The best algorithm for predicting graduation and dropout is the decision tree with the best accuracy value of 99.15% with a training data ratio of 30%.

**Keyword :** Data Mining; Algoritma Naive Bayes; Decision Tree; K-Nearest Neighbor; Predict Graduation; Drop Out.

## Abstrak

Saat ini, masalah kegagalan studi, kelulusan tepat waktu dan faktor-faktor penyebabnya masih menjadi topik yang menarik untuk diteliti (C. Marquez-Vera, C. Romero, and S. Ventura, 2011). Dalam penelitian ini, akan dilakukan perbandingan tiga algoritma klasifikasi data mining yaitu *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbor* untuk prediksi kelulusan mahasiswa dan resiko putus studi bagi mahasiswa untuk peningkatan mutu perguruan tinggi dan dapat menghasilkan algoritma yang paling akurat dalam prediksi kelulusan dan putus studi mahasiswa. Algoritma yang terbaik untuk prediksi kelulusan mahasiswa dan putus studi mahasiswa adalah *Decision Tree* dengan nilai akurasi terbaik yaitu 99.15 % dengan rasio data training 30%.

**Kata Kunci :** Data Mining; Algoritma Naive Bayes; Decision Tree; K-Nearest Neighbor; Prediksi Kelulusan; Putus Studi

## 1. PENDAHULUAN

Saat ini, masalah kegagalan studi, kelulusan tepat waktu dan faktor-faktor penyebabnya masih menjadi topik yang menarik untuk diteliti (C. Marquez-Vera, C. Romero, and S. Ventura, 2011). Dalam mengatasi persaingan yang semakin tinggi dan salah satu faktor untuk meningkatkan mutu Perguruan Tinggi, manajemen memerlukan tindakan *preventif* untuk mengetahui faktor-faktor penyebab kegagalan mahasiswa dan kelulusan mahasiswa yang tidak tepat waktu. Beberapa penyebab kegagalan mahasiswa diantaranya rendahnya kemampuan akademik, faktor keuangan, tempat tinggal saat menempuh studi, tempat asal dan faktor lainnya. Hal tersebut menjadi sangat penting bagi manajemen perguruan tinggi mengingat persentasi mahasiswa lulus tepat waktu adalah salah satu elemen penilaian akreditasi yang ditetapkan oleh Badan Akreditasi Nasional.

Penelitian terkait permasalahan Putus Studi (*drop out*), mahasiswa *non* aktif dan kelulusan mahasiswa sudah banyak dilakukan oleh para peneliti sebelumnya seperti *Kotsiantis*, *Pierrakeas* dan *Pintelas* menyebutkan bahwa sangat penting bagi dosen untuk mendeteksi mahasiswa yang cenderung *drop out* sebelum mereka memasuki pertengahan masa studi. *Gerben W. Dekker* menyebutkan bahwa monitoring dan dukungan terhadap mahasiswa di tahun pertama sangat penting dilakukan. Kurikulum yang sulit dianggap sebagai salah satu penyebab tingginya jumlah mahasiswa *drop out*. Selain itu, nilai, prestasi, kepribadian, latar belakang sosial mempunyai peran dalam

kesuksesan akademik mahasiswa. Beberapa algoritma klasifikasi *data mining* telah digunakan untuk memprediksi perilaku mahasiswa yang berpotensi *drop out* diantaranya *decision tree*, *neural network*, *naïve bayes*, *instance-based learning*, *logistic regression* dan *support vector machine*.

Dalam penelitian ini, akan dilakukan perbandingan tiga algoritma klasifikasi data mining yaitu *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbor* untuk prediksi kelulusan mahasiswa dan resiko putus studi bagi mahasiswa untuk peningkatan mutu perguruan tinggi dan dapat menghasilkan algoritma yang paling akurat dalam prediksi kelulusan dan putus studi mahasiswa.

## 2. KERANGKA TEORI

### 2.1 Data Mining

*Data Mining* merupakan salah satu cabang ilmu komputer yang relatif baru yang memiliki keterkaitan dengan *machine learning*, kecerdasan buatan (*artiifial intelligence*), *statistic* dan *database*. *Data mining* menagacu kepada ekstraksi pengetahuan (*knowledge extraction*) analisis data atau pola (*patern analysis*), penggalian data dan lain sebagainya (Maimon, O. dan Last, M, 2000).

Terdapat beberapa pengertian yang berkaitan dengan *Data mining* dari beberapa referensi sebagai berikut.

1. *Data mining* adalah mencocokkan data dalam suatu model untuk menemukan informasi yang tersembunyi dalam basisdata.
2. *Data mining* merupakan aplikasi suatu algoritma untuk menggali informasi bermanfaat dari dalam basisdata.
3. *Data mining* adalah proses menemukan pola-pola didalam data, dimana proses penemuan tersebut dilakukan secara otomatis atau semi otomatis dan pola-pola yang ditemukan harus bermanfaat.
4. *Data mining* adalah proses penemuan informasi yang berguna pada penyimpanan data yang besar secara otomatis.
5. *Data mining* atau *Knowledge Discovery in Databases* (KDD) adalah pengambilan informasi yang tersembunyi, dimana informasi tersebut sebelumnya tidak dikenal dan berpotensi bermanfaat. Proses ini meliputi sejumlah pendekatan teknis yang berbeda, seperti *clustering*, *data summarization*, *learning classification rules*. (Maimon, O. dan Last, M, 2000).

### 2.2 Algoritma Naive Bayes

Teori keputusan *Bayes* adalah pendekatan *statistic* yang fundamental dalam pengenalan pola (*Pattern Recognition*). Pendekatan ini didasarkan pada kuantifikasi *trade-off* antara berbagai keputusan klasifikasi dengan menggunakan probabilitas dan ongkos yang ditimbulkan dalam keputusan-keputusan tersebut.

*Bayesian classification* adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. *Bayesian classification* didasarkan pada *teorema Bayes* yang memiliki kemampuan klasifikasi serupa dengan *Decision Tree* dan *Neural Network*. *Bayesian Classification* terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar (Kusrini dan Emha Taufiq Luthfi, 2009).

Teorema Bayes memiliki bentuk umum sebagai berikut :

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

- X = Data dengan class yang belum diketahui
- H = Hipotesis data X merupakan suatu class spesifik
- $P(H|X)$  = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
- $P(H)$  = Probabilitas hipotesis H (prior prob.)
- $P(X|H)$  = Probabilitas X berdasarkan kondisi tersebut
- $P(X)$  = Probabilitas dari X

### 2.3 Decision Tree

Menurut Han et al. (2012), Pohon Keputusan (*Decision Tree*) merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon (tree) yang setiap internalnode (non-leaf node) merepresentasikan atribut, cabangnya merepresentasikan nilai dari atribut, dan daun (leaf node atau terminal node) merepresentasikan kelas. Node yang paling atas dari *decision tree* disebut sebagai root. *Decision Tree* merupakan metode klasifikasi yang paling populer digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami. Pada *decision tree* terdapat 3 jenis node (Cahyono, 2010), yaitu :

a. *Root Node*

*Root Node* merupakan node paling atas, pada node ini tidak ada input dan bisa tidak mempunyai output atau mempunyai output lebih dari satu.

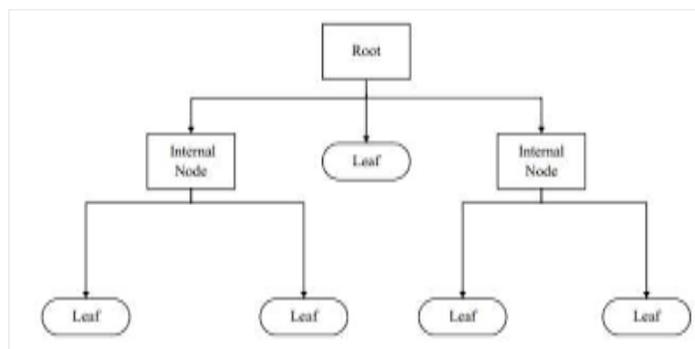
b. *Internal Node*

*Internal node* merupakan node percabangan, pada node ini hanya terdapat satu input dan mempunyai output minimal dua.

c. *Leaf Node atau Terminal Node*

*Leaf node* atau terminal node merupakan node akhir, pada node ini hanya terdapat satu input dan tidak mempunyai output.

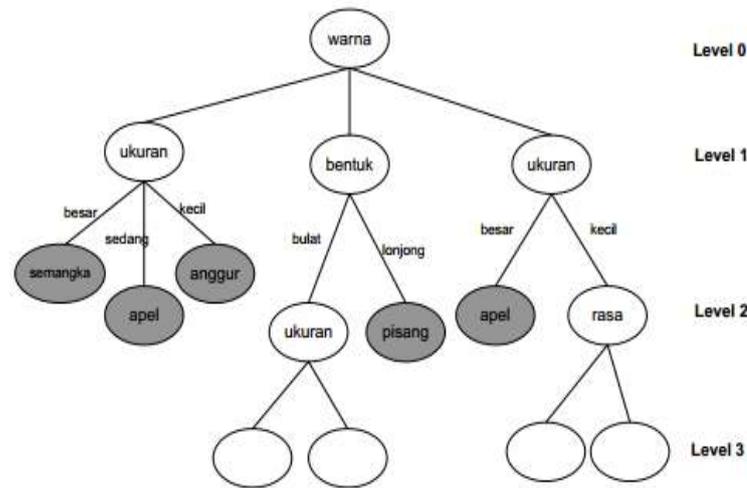
Contoh dari pohon keputusan dapat dilihat di gambar 1 berikut ini:



**Gambar 1.** Model Pohon Keputusan (Hamidah, 2012)

Dalam *decision tree* tidak menggunakan vector jarak untuk mengklasifikasikan obyek. Seringkali data observasi mempunyai atribut-atribut yang bernilai nominal. Seperti yang diilustrasikan pada gambar 2.6, misalkan obyeknya adalah sekumpulan buah-buahan yang bisa dibedakan berdasarkan atribut bentuk, warna, ukuran dan rasa. Bentuk, warna, ukuran dan rasa adalah besaran nominal, yaitu bersifat kategoris dan tiap nilai tidak bisa dijumlahkan atau dikurangkan. Dalam atribut warna ada beberapa nilai yang mungkin yaitu hijau, kuning, merah. Dalam atribut ukuran

ada nilai besar, sedang dan kecil. Dengan nilai-nilai atribut ini, kemudian dibuat *decision tree* untuk menentukan suatu obyek termasuk jenis buah apa jika nilai tiap-tiap atribut diberikan (B. Santoso, 2007).



**Gambar 2.** *Decision Tree*

#### 2.4 Algoritma K-Nearest Neighbor

*K-Nearest Neighbor* (KNN) termasuk kelompok *instance-based learning*. Algoritma ini juga merupakan salah satu teknik *lazy learning*. KNN dilakukan dengan mencari kelompok  $k$  objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing (Gerben W. Dekker, 2009). *Algoritma K-Nearest Neighbor* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. *Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dan kasus lama yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. Untuk mendefinisikan jarak antara dua titik yaitu titik pada data training ( $x$ ) dan titik pada data testing ( $y$ ) maka digunakan rumus *Euclidean*, seperti yang ditunjukkan pada persamaan (1)

$$D(x, y) = \sqrt{\sum_{k=1}^n (X_k - Y_k)^2}$$

Dengan  $D$  adalah jarak antara titik pada data training  $x$  dan titik data testing  $y$  yang akan diklasifikasi, dimana  $x = x_1, x_2, \dots, x_i$  dan  $y = y_1, y_2, \dots, y_i$  dan  $I$  merepresentasikan nilai atribut serta  $n$  merupakan dimensi atribut.

Pada *fase training*, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi data *training sample*. Pada *fase klasifikasi*, fitur-fitur yang sama dihitung untuk *testing data* (yang klasifikasinya tidak diketahui). Jarak dari vektor baru yang ini terhadap seluruh vektor *training sample* dihitung dan sejumlah  $k$  buah yang paling dekat diambil.

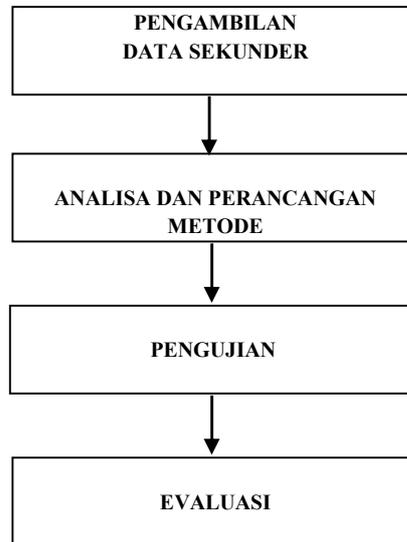
Langkah-langkah untuk menghitung metode *Algoritma K-Nearest Neighbor*:

- Menentukan Parameter  $K$  (jumlah tetangga paling dekat).
- Menghitung kuadrat jarak *Euclid* (*queri instance*) masing-masing objek terhadap data sampel yang diberikan.
- Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *Euclid* terkecil.
- Mengumpulkan kategori  $Y$  (*Klasifikasi Nearest Neighbour*)

Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *query instance* yang telah dihitung.

### 3. METODE PENELITIAN

Secara garis besar tahapan pelaksanaan penelitian yang dilakukan dapat dilihat pada gambar berikut ini :

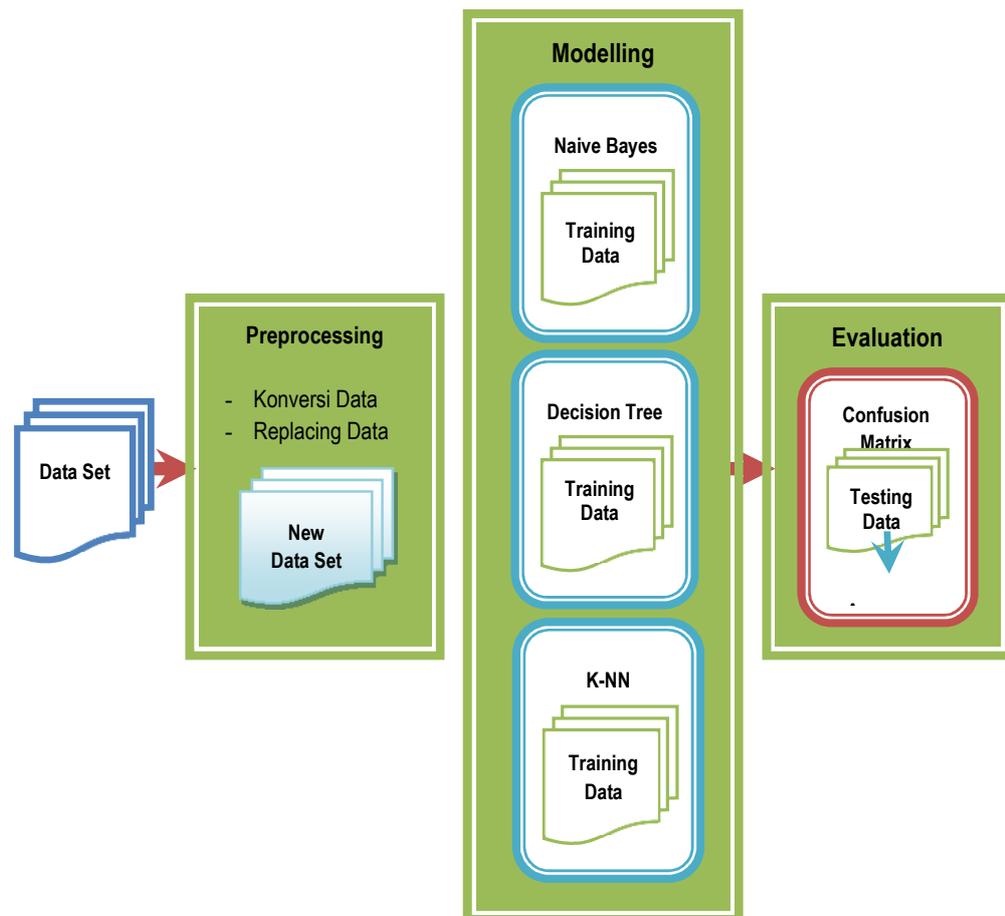


**Gambar 3.** Skema Alur Kegiatan Penelitian

Subjek yang akan diteliti dalam penelitian ini adalah tentang teknik data mining untuk memprediksi kelulusan dan putus studi mahasiswa dengan pendekatan bertingkat pada perguruan tinggi dengan menggunakan Algoritma *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbour*, penelitian ini dilakukan dengan mengambil *database* akademik yang ada di perguruan tinggi x yang merupakan salah satu perguruan tinggi swasta di Lampung. Data yang diperoleh dan akan digunakan dalam penelitian ini berupa data berkaitan dengan keterangan diri dari mahasiswa yaitu Jenis Kelamin, Umur, Jenis Sekolah, Kota Sekolah, Prodi, Angkatan, Kota Mahasiswa, SKS, IPK, Pekerjaan Orang Tua, dan Status Mahasiswa. Data-data tersebut yang kemudian akan dijadikan sebagai atribut dalam memprediksi Kelulusan dan Putus Studi Mahasiswa.

#### 3.1 Pemodelan (Modeling)

Metode yang akan di gunakan pada penelitian ini adalah *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbour*. Untuk melakukan pengukuran akurasi dalam penelitian ini akan menggunakan *framework RapidMiner Studio 7.0.000*



**Gambar 4.** Model Penelitian Algoritma *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbour*

#### 4. HASIL DAN PEMBAHASAN

##### 4.1 Persiapan Data

Dalam pengujian ini digunakan data sampel mahasiswa angkatan 2008-2012 yang sudah dinyatakan lulus dan angkatan 2008-2013 yang sudah di nyatakan putus studi. Data mahasiswa sebanyak 2767 record diambil 1300 record untuk digunakan sebagai data training. Berdasarkan hasil pengolahan data dari jumlah data yang digunakan tersebut dan dibagi kedalam 2 kelas maka diperoleh jumlah data mahasiswa dengan kategori Lulus sebanyak 517 Mahasiswa dan Putus Studi sebanyak 783 Mahasiswa.

**Tabel 1.** *Data Training* Lulusan dan Putus Studi (Sampling)

JK	UMUR	ASAL SEKOLAH	ALAMAT SEKOLAH	PRODI	ANGKTN	KOTA ASAL	SKS	IPK	KERJA ORTU	STATUS
L	27	SMA	Lain-lain	101	2011	120200	18	0.66	Lain-lain	Drop Out
L	27	SMA	Natar	102	2012	120900	38	3.00	Petani	Drop Out
L	29	SMA	Lampung Selatan	211	2011	120100	18	0.44	Pegawai Negeri	Drop Out
L	30	SMA	Lain-lain	212	2008	120900	54	0.88	Wiraswasta	Drop Out
L	35	SMA	Bandar Lampung	101	2008	126000	19	0.00	Petani	Drop Out
P	16	SMA	Wonorejo	101	2010	110800	148	3.15	Wiraswasta	Lulus
P	16	SMK	Lampung Tengah	212	2011	120200	80	2.95	Petani	Drop Out
P	17	MAN	Air Molek	101	2010	90400	18	1.44	Petani	Drop Out
P	17	MAN	Lampung Tengah	101	2010	120200	126	1.28	Pegawai Swasta	Drop Out

4.2 Hasil Percobaan

Berdasarkan dari percobaan-percobaan yang sudah dibuat, hasilnya sebagai berikut :

**Tabel 2.** Hasil 5 Kali Percobaan dengan 3 Algoritma

NO	Percobaan (%) Data Training	Akurasi Terbaik
1	10	<i>Naive Bayes</i>
		Akurasi Terbaik : 94.74 % Classification Rate :5.26 %
2	30	<i>Decision Tree</i>
		Akurasi Terbaik : 99.15 % Classification Rate 0.85 %
3	50	<i>Decision Tree</i>
		Akurasi Terbaik : 97.44 % Classification Rate 2.56 %
4	75	<i>Decision Tree</i>
		Akurasi Terbaik : 98.63 % Classification Rate 1.37 %
5	90	<i>Decision Tree</i>
		Akurasi Terbaik : 97.44 % Classification Rate 2.56 %

Dari hasil diatas dapat di simpulkan bahwa, Algoritma yang terbaik untuk prediksi kelulusan mahasiswa dan putus studi mahasiswa adalah *Decision Tree* dengan hasil akurasi terbaik yaitu dengan menggunakan data training sebanyak 30 % (390 record data) dan menghasilkan akurasi data terbaik sebesar 99.15 % dan *Classification Rate* 0.85 %.

4.3 Hasil Komparasi

**Tabel 3.** Hasil Komparasi 3 Algoritma

Metode	Sampling Type	Rasio Data Training (%)					$\bar{X}$
		10	30	50	75	90	
Naive Bayes	Linear	92.31	99.15	97.44	95.55	96.58	96.21
	Shuffled	94.87	97.44	97.44	96.23	94.87	96.17
	Stratified	97.44	96.58	96.92	96.23	95.16	96.47
Decision Tree	Linear	92.31	99.15	97.95	98.63	97.72	97.15
	Shuffled	97.44	96.58	96.41	96.23	98.58	97.05
	Stratified	97.44	99.15	97.44	98.63	97.44	98.02

**Tabel. 4.** Hasil Komparasi 3 Algoritma (Lanjutan)

Metode	Sampling Type	Rasio Data Training (%)					$\bar{X}$
		10	30	50	75	90	
Akurasi	Linear	84.62	99.15	94.87	82.19	64.96	85.16
	K-Nearest Neighbour	92.31	91.45	91.28	92.12	91.17	91.67
	Stratified	79.49	89.74	91.28	90.41	89.74	88.13

Komparasi dari 3 algoritma *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbour* dengan 3 *sampling type* (*Linear*, *Shuffled*, *Stratified*) dan menggunakan data training yang berbeda dari 10 % sampai dengan 90 % data, dapat disimpulkan bahwa *Decision Tree* dengan *Stratified Sampling* memiliki nilai akurasi rata-rata terbaik sebesar 98,02, kemudian *Naive Bayes* dengan akurasi rata-rata 96,47 dan akurasi terkecil di hasilkan *K-Nearest Neighbour* dengan nilai akurasi 91,67.

Berdasarkan semua percobaan yang dilakukan, maka dapat disimpulkan bahwa algoritma yang terbaik untuk memprediksi Kelulusan dan Putus Studi Mahasiswa di perguruan tinggi x dengan data training sebanyak 1300 *record* yang dilakukan dengan jumlah data training yang berbeda, menghasilkan akurasi terbaik adalah Algoritma *Decision Tree* dengan nilai akurasi terbaik 99,15 dengan data training sebanyak 30 % (390 *record*).

## 5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan maka dapat diambil beberapa kesimpulan, antara lain:

1. *Ratio data training* yang digunakan mempengaruhi nilai akurasi pada setiap percobaan. Pada percobaan ke-1 nilai akurasi terbaik di hasilkan oleh *Naive Bayes* dengan 94,74 % dengan *ratio data training* 10%. Pada percobaan ke-2, ke-3, ke-4 dan ke-5 nilai akurasi terbaik dihasilkan oleh *Decision Tree* dengan nilai akurasi dan rasio data training berturut-turut : 99,15 % *ratio data training* 30%, 97,44 % *ratio data training* 50%, 98,63 % *ratio data training* 75% dan 97,44 % *ratio data training* 90%.
2. Variabel yang paling berpengaruh terhadap kelulusan mahasiswa dan Putus Studi adalah Indeks Prestasi Semester (IPK) dan Satuan Kredit Semester (SKS) yang diambil dan Program Studi yang diambil berdasarkan Algoritma *Decision Tree*.
3. Berdasarkan penelitian yang dilakukan ini, dari ketiga algoritma yang digunakan *Naive Bayes*, *Decision Tree* dan *K-Nearest Neighbour* untuk prediksi kelulusan dan putus studi mahasiswa, maka algoritma yang terbaik adalah *Decision Tree* dengan nilai akurasi terbaik yaitu 99,15 % dengan rasio data training 30%.

## DAFTAR PUSTAKA

- C. Marquez-Vera, C. Romero, and S., 2011. Ventura, "Predicting School Failure Using Data Mining," *Journal of Educational Data Mining*.
- Cahyono, A. B., 2009. Analisis Pemanfaatan Small Disjunct Pada Decision Tree Dengan Algoritma Genetika. *Jurnal Seminar Nasional Aplikasi Teknologi Informasi*.
- Gerben W. Dekker, 2009. Predicting Students Drop Out: A Case Study. *In International Conference on Educational Data Mining, Cordoba, Spain*, 41-50, 2009.

- Hamidah, I., 2012. Aplikasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5. Digilab Umikom.
- Han, J., Kamber, M., & Pei, J., 2012. Data Mining: Concepts And Techniques. Waltham: Morgan Kaufmann.
- Kusrini dan Emha Taufiq Luthfi., 2009. Algoritma Data Mining. Penerbit Andi, Yogyakarta.
- Kusrini, L dan Taufiq Emha., 2009. Algoritma Data Mining, Edisi Pertama. Penerbit Andi, Yogyakarta.
- Maimon, O. dan Last, M., 2000. Knowledge Discovery and Data Mining, The Info Fuzzy Network (IFN) Methodology. Dordrecht: Kluwer Academic.
- Santoso, Budi., 2007. Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis, 1st ed. Graha Ilmu, Yogyakarta.





**Diterbitkan :**  
**LEMBAGA PENGEMBANGAN PEMBELAJARAN, PENELITIAN, DAN PENGABDIAN MASYARAKAT (LP4M)**  
**INSTITUT INFORMATIKA & BISNIS DARMAJAYA**

**Alamat :** Jalan Zainal Abidin Pagar Alam No.93 Gedong Meneng, Bandar Lampung 35142  
**Telp. 0721-787214 Fax. 0721- 700261**  
**email : [simada@darmajaya.ac.id](mailto:simada@darmajaya.ac.id)**  
**Website : [jurnal.darmajaya.ac.id](http://jurnal.darmajaya.ac.id)**