

BAB II

LANDASAN TEORI

Pada bab ini akan dijelaskan satu persatu tentang metode Naive Bayes Studi, tinjauan pustaka, *Entity Relationship Diagram* (ERD), Data Prediksi, *Data Flow Diagram* (DFD), Microsoft Visual Studio 2010 Ultimate, *Structured Query Language* (SQL).

2.1. Tinjauan Pustaka

Pada penelitian ini akan dipergunakan lima tinjauan pustaka yang ditinjau untuk mendukung penelitian yang dilakukan penulis, dimana tinjauan pustaka yang digunakan adalah :

1. Oleh M. Sabransyah, Yuki Novia Nasution dan Fidia Deny Tisna Amijaya (2017) dari Universitas Mulawarman dengan judul Aplikasi Metode Naïve Bayes dalam sistem Prediksi Resiko Penyakit Jantung. Di dalam jurnal ini menjelaskan tentang bagaimanakah implementasi metode naïve bayes ke dalam sebuah aplikasi untuk memprediksi resiko penyakit jantung. Hasil yang terdapat pada penelitian ini adalah metode naïve bayes dapat dipergunakan untuk memprediksi resiko seseorang terkena penyakit jantung.
2. Oleh Muchtarul Bari, Sampe Hotlan Sitorus dan Uray Ristian (2018) dari Universitas Tanjungpura dengan judul Implementasi Metode Naïve Bayes pada Aplikasi Penyebaran Wabah Penyakit ISPA. Jurnal ini menjelaskan tentang bagaimana cara membuat aplikasi prediksi menggunakan metode naïve bayes untuk memprediksi tentang penyebaran penyakit ISPA pada wilayah Pontianak. Dengan hasil penelitian tingkat keakurasian metode naïve bayes dalam memprediksi penyebaran penyakit ISPA adalah 82.97%.
3. Oleh Rahmania Shalihah (2016) dari Universitas Muhammadiyah Surakarta dengan judul Prediksi Perilaku Siswa SMKN 2 Surakarta Memakai Metode Naïve Bayes. Pada Jurnal ini menjelaskan tentang bagaimana cara memprediksi perilaku siswa menggunakan metode naïve bayes pada SMKN

- 2 Surakarta. Hasil penelitiannya dengan tingkat kakuratan yang tinggi, metode naïve bayes dikira cukup terbukti dalam memprediksi potensi perilaku siswa.
4. Oleh Lianny Wydiastuty Kusuma (2019) dari Universitas Budi Dharma Banten dengan judul Prediksi Kemampuan Lulusan SMK untuk Dapat Bersaing di Dunia Kerja dengan Menggunakan Naïve Bayes : Studi Kasus SMK Buddhi Tanggerang. Jurnal ini menjelaskan tentang penggunaan metode naïve bayes untuk memprediksi keahlian siswa lulusan SMK, apakah bisa bersaing dalam dunia kerja. Memiliki hasil yaitu metode naïve bayes mempunyai akurasi yang baik dengan poin akurasi sebesar 98% dan nilai AUC sebesar 0.980.
 5. Oleh Juli Sulaksono dan Danang Wahyu Widodo (2018) dari Universitas Nusantara PGRI Kediri dengan judul Aplikasi Penentuan Tiga Gaya Belajar Siswa Menggunakan Metode Naïve Bayes (Studi Kasus SDN Puhrubuh 1). Pada jurnal ini menjelaskan tentang bagaimana membuat aplikasi penent 3 gaya belajar siswa memakai metode naïve bayes. Memiliki hasil yaitu dapat mengetahui 3 jenis gaya belajar siswa dan solusi cara belajar sesuai dengan cara belajar masing-masing siswa.

2.2. Pengertian Sistem Prediksi

Prediksi merupakan proses meramalkan secara sistematis tentang suatu yang mungkin saja terjadi di masa mendatang berdasarkan data masa lampau dan masa sekarang yang dimiliki, agar tingkat kesalahan dapat diperkecil. Prediksi tidak selalu memberikan hasil secara tepat sesuatu yang mungkin saja terjadi, melainkan berusaha untuk mencari hasil sedekat mungkin dengan yang mungkin saja terjadi (Herdianto, 2013 : 8).

Prediksi dapat didasari oleh metode penelitiannya maupun subjektif belaka. Sebagai contoh, prediksi ramalan cuaca akan berdasar pada data-data dan informasi aktual yang didasarkan pada pengamatan termasuk oleh satelit. Begitu juga dalam memprediksi gempa, gunung meletus, tsunami maupun bencana yang umum terjadi. Namun, prediksi seperti pertandingan sepakbola, olympiade, dll

biasanya berdasar pada pandangan subjektif dengan sudut pandang sendiri yang memprediksinya.

Peramalan (*forecasting*) merupakan suatu proses untuk memberikan data actual tentang situasi pada masa mendatang atas dasar informasi yang telah ada. Ramalan memiliki 3 bentuk utama yaitu: prediksi, proyeksi, dan perkiraan.

1. Prediksi ialah peramalan yang berdasarkan pada asumsi teoritik yang tegas. Asumsi ini dapat berupa hukum teoretis (contohnya hukum berubahnya nilai mata uang), proposisi teoritis (contohnya proposisi bahwa pecahnya masyarakat sipil diakibatkan oleh kesenjangan antara harapan dan kemampuan), atau analogi (contohnya analogi antara pertumbuhan organisasi pemerintah dan pertumbuhan organisme biologis).
2. Proyeksi ialah prediksi yang berdasarkan pada ekstrapolasi atas kecenderungannya pada masa lampau, maupun pada masa sekarang ke masa mendatang. Proyeksi membuat pernyataan yang tegas berdasarkan argumen yang didapat dari metode tertentu dan kasus yang paralel.
3. Perkiraan (*conjecture*) ialah prediksi yang berdasarkan pada penilaian yang informatif atau penilaian pakar tentang situasi masyarakat pada masa mendatang.

Tujuan dari dilakukannya prediksi kebijakan ialah untuk mendapatkan data tentang perubahan pada masa mendatang yang akan berpengaruh terhadap implementasi kebijakan serta konsekuensinya. Oleh karena itu, sebelum rekomendasi diformulasikan perlu adanya prediksi kebijakan sehingga dapat menghasilkan rekomendasi yang akurat untuk diberlakukan pada masa depan. Didalam pemrediksian kebutuhan mendatang dengan bertumpu pada masa lampau, dibutuhkan seseorang yang mempunyai daya sensitif yang tinggi dan dapat membaca kemungkinan pada masa mendatang.

2.3. Pendidikan Anak Usia Dini

Pendidikan anak pada usia dini merupakan suatu usaha pembelajaran yang ditunjukkan kepada anak umur empat tahun sampai dengan umur 6 tahun yang dilakukan melalui pemberian rangsangan pendidikan untuk membantu

pertumbuhan jasmani dan rohani agar anak menjadi mempunyai kesiapan dalam memasuki pendidikan selanjutnya.

Pendidikan anak pada usia dini dilakukan sebagai proses belajar dalam diri anak. Proses sosialisasi, anak diberikan pemahaman untuk dapat melatih diri menjadi anak yang bermoral, beretika dan bertanggung jawab. Proses pembentukan kerjasama peran, anak diberikan kesempatan untuk mengembangkan potensi sosialnya, supaya anak menyadari sebagai makhluk sosial yang selalu berinteraksi dengan orang lain. (Sujarwo, 2020).

Proses pembelajaran anak usia empat sampai enam tahun hendaknya memperhatikan sembilan keahlian anak yang meliputi:

1. Keahlian linguistic,
2. Keahlian matematik,
3. Keahlian visual spasial,
4. Keahlian musikal,
5. Keahlian kinestetik,
6. Keahlian naturalis,
7. Keahlian interpersonal
8. Keahlian intrapersonal, dan
9. Kecerdasan spiritual.

Dari 9 keahlian diatas, secara operasional disederhanakan kedalam 6 aspek pengembangan, yaitu:

1. Pengembangan moral dan nilai-nilai agama,
2. Fisik,
3. Bahasa,
4. Kognitif (proses berfikir),
5. Sosial emosional dan
6. Pengembangan seni.

2.4. Penilaian Potensi Keahlian Anak

Penelitian ini menggunakan penelitian deskriptif, tujuan penelitian deskriptif bukan menjadi acuan suatu hipotesis tertentu, melainkan hanya berupa

gambaran suatu variabel yang dipergunakan, suatu variabel yang mendasari penelitian ini yaitu penilaian potensi pada anak usia dini. Subjek pada penelitian ini adalah anak usia 3 sampai 6 tahun di TK Putra Harapan 1 Karanglangit Lamongan, sebanyak 26 anak didik, jumlah jenis kelamin perempuan 11 anak dan laki-laki 14 anak didik.

Teknik mengumpulkan data pada penelitian ini melalui dokumentasi, laporan berisi 6 butir nilai yang diisi oleh guru melalui pembelajaran yang dilakukan setiap hari. Berikut ini penjabaran isi dari poin penilaian pada laporan :

Tabel 2.1 Laporan Penilaian

No.	Potensi Keahlian/ Kemampuan Anak	Deskripsi	Nilai
1	Pengembangan moral dan nilai-nilai agama	Selalu mengucapkan salam dan hormat ketika memulai atau sesudah pembelajaran.	4
		Terkadang ingat mengucapkan salam dan hormat ketika memulai atau sesudah pembelajaran.	3
		Harus selalu diingatkan mengucapkan salam dan hormat ketika memulai atau sesudah pembelajaran.	2
		Tidak mengucapkan salam dan hormat ketika memulai atau sesudah pembelajaran.	1
2	Fisik	Berse semangat dan antusias dalam mengikuti olahraga setiap 1 minggu sekali.	4
		Mengikuti olahraga setiap 1	3

		minggu sekali dengan motivasi.	
		Mengikuti olahraga setiap 1 minggu sekali dengan harus selalu diikuti oleh orang tua wali.	2
		Tidak pernah mengikuti olahraga setiap 1 minggu sekali.	1
3	Bahasa	Bisa menulis dan membaca huruf atau kata dengan baik dan benar.	4
		Bisa menulis dan membaca huruf atau kata dengan baik tetapi tidak selalu benar.	3
		Bisa membaca tetapi tidak bisa menulis, atau sebaliknya.	2
		Tidak bisa membaca dan menulis.	1
4	Kognitif (Proses Berfikir)	Bisa menjawab soal yang diberikan oleh guru dengan benar.	4
		Bisa menjawab soal yang diberikan oleh guru tetapi dengan motivasi.	3
		Menjawab dengan benar tetapi dengan bantuan orang lain.	2
		Tidak bisa menjawab menjawab soal yang diberikan oleh guru.	1
5	Sosial Emosional	Bisa berbicara dan berteman	4

		dengan siswa yang lain.	
		Berteman dengan siswa lain tetapi malu untuk berbicara.	3
		Berteman dengan siswa lain tetapi dengan bantuan orang tua.	2
		Tidak bisa berbicara maupun berteman dengan siswa lain.	1
6	Pengembangan Seni	Bisa menggambar atau bernyanyi bersama dengan baik.	4
		Bisa menggambar atau bernyanyi bersama tetapi dengan bantuan orang tua ataupun malu-malu.	3
		Bisa menggambar tetapi tidak bisa bernyanyi bersama ataupun sebaliknya.	2
		Tidak bisa menggambar atau bernyanyi bersama.	1

Tabel 2.1 diatas adalah poin perhitungan keahlian anak yang akan dinilai oleh guru setiap seminggu sekali, yang nantinya poin tersebut akan dijumlahkan dan dilakukan perhitungan setiap bulan untuk memprediksi keahlian setiap anak.

Tabel 2.2 Contoh Penilaian Tiap Minggu

	Nama Siswa	Moral	Fisik	Bahasa	Kognitif	Sosial	Seni
Minggu Pertama	Siswa 1	4	3	2	2	3	2
	Siswa 2	1	2	3	3	4	4
	Siswa 3	2	3	4	4	3	2

Minggu Kedua	Siswa 1	3	3	3	3	4	2
	Siswa 2	2	3	4	4	3	3
	Siswa 3	4	3	3	4	3	2
Minggu Ketiga	Siswa 1	4	3	3	2	2	3
	Siswa 2	1	3	4	4	3	3
	Siswa 3	2	3	3	4	2	4
Minggu Keempat	Siswa 1	3	4	3	3	2	2
	Siswa 2	1	3	3	4	3	3
	Siswa 3	2	4	3	4	2	2

Tabel 2.3 Contoh Jumlah Poin Satu Bulan

	Jumlah					
Nama Siswa	Moral	Fisik	Bahasa	Kognitif	Sosial	Seni
Siswa 1	14	13	11	10	11	9
Siswa 2	5	11	14	15	13	13
Siswa 3	10	13	13	16	10	10

2.5. Metode Naïve Bayes

Naïve Bayes adalah metode probabilitas pengklasifikasian sederhana berdasarkan Teorema Bayes yang mana pengklasifikasian dilakukan melalui training set sejumlah data secara efisien. Metode Naïve Bayes menggunakan peluang bersyarat dalam melakukan prediksi. (Muchtarul Bari, 2018).

Sebelum menjelaskan tentang peluang bersyarat pertama perlu diketahui tentang definisi peluang. Peluang dapat didefinisikan sebagai kemungkinan munculnya suatu kejadian. Teori peluang merupakan bagian integral dari ilmu statistik, yang juga merupakan salah satu bagian yang penting dalam teori statistik inferensial. Statistik inferensial berkaitan dengan metode perkiraan dan penarikan

kesimpulan terhadap karakteristik suatu populasi berdasarkan informasi yang diperoleh dari sampel. Cara menghitung peluang dapat dilihat pada Persamaan dibawah ini dimana peluang terjadinya A dilambangkan dengan $P(A)$ yaitu jumlah peluang A dari semua anggota titik contoh S.

$$P(A) = \frac{n(A)}{n(S)} \dots\dots\dots (2.1)$$

Keterangan :

$P(A)$ = Peluang kejadian A

$n(A)$ = Jumlah kejadian A

$n(S)$ = Jumlah titik contoh S

Peluang bersyarat ialah peluang munculnya suatu kejadian dengan syarat bahwa suatu kejadian yang lain telah muncul lebih dahulu. Peluang terjadi kejadian A dengan syarat bahwa kejadian B telah terjadi dihitung dengan Persamaan dibawah ini.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} ; P(B) > 0 \dots\dots\dots (2.2)$$

Keterangan :

$P(A|B)$ = Peluang kejadian A dengan syarat kejadian B telah terjadi

$P(A \cap B)$ = Peluang kejadian A irisan B

$P(B)$ = Peluang kejadian B

Naïve Bayes mengasumsikan bahwa nilai dari suatu masukan atribut pada kelas yang diberikan tidak tergantung dengan nilai atribut yang lain. Teorema Bayes merupakan teorema yang memprediksi peluang di masa mendatang berdasarkan pengalaman di masa lampau. Teorema tersebut dikemukakan oleh ilmuwan Inggris Thomas Bayes, dimana Persamaan dibawah ini merupakan persamaan Teorema Bayes.

$$P(C|F) = \frac{P(F|C).P(C)}{P(F)} \dots\dots\dots (2.3)$$

Keterangan :

F = Data dengan class yang belum diketahui

C = Hipotesis data F merupakan suatu class spesifik

$P(C|F)$ = Probabilitas hipotesis C berdasar kondisi F (*poteriori probability*)

$P(C)$ = Probabilitas hipotesis C (*Prior Probability*)

$P(F|C)$ = Probabilitas F berdasarkan kondisi pada hipotesis C

$P(F)$ = Probabilitas F

Untuk menjelaskan metode Naïve Bayes, perlu diketahui bahwa proses pengklasifikasian memerlukan sejumlah data untuk dapat menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, Teorema Bayes pada Persamaan diatas disederhanakan menjadi :

$$P(C|F_1, \dots, F_n) = \frac{P(C).P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)} \dots\dots\dots (2.4)$$

Dimana variable mempresentasikan kelas, sementara variabel $F_1 \dots F_n$ mempresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka persamaan diatas menjelaskan bahwa peluang terjadinya karakteristik tertentu dalam kelas posterior ialah peluang munculnya kelas *prior*, dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas (disebut dengan *likelihood*), dibagi dengan peluang kemunculan karakteristik sampel secara global (disebut juga *evidence*). Karena itu, Persamaan diatas dapat pula ditulis secara sederhana seperti :

$$\text{Posterior} = \frac{\text{Prior x likelihood}}{\text{evidence}} \dots\dots\dots (2.5)$$

Nilai Evidence selalu tetap untuk setiap kelas pada suatu sampel. Nilai dari posterior tersebut akan dibandingkan dengan nilai posterior kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut persamaan Bayes tersebut dilakukan dengan menjabarkan $(C|F_1....F_n)$ menggunakan aturan perkalian seperti pada Persamaan berikut :

$$P(C|F_1....F_n) = P(C).P(F_1....F_n|C) = P(C).P(F_1|C).P(F_2....F_n|C,F_1) = P(C).P(F_1|C)....P(F_n|C,F_1....F_{n-1}) \dots\dots\dots (2.6)$$

Dapat dicermati bahwa hasil perumusan diatas mengakibatkan semakin banyak dan kompleksnya factor-faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Disinilah digunakan asumsi independensi yang sangat tinggi, bahwa masing-masing petunjuk $(F_1, F_2....F_n)$ saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan seperti pada Persamaan berikut :

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i).P(F_j)}{P(F_j)} = P(F_i) \dots\dots\dots (2.7)$$

Untuk $i \neq j$, sehingga :

$$P(F_i|C,F_j) = (P(F_i|C) \dots\dots\dots (2.8)$$

Dari rumus diatas dapat dicermati bahwa asumsi independensi naif tersebut membuat syarat peluang menjadi sederhana, sehingga perhitungan menjadi mungkin untuk dilakukan. Selanjutnya, penjabaran $P(C |F_1....F_n)$ dapat disederhanakan menjadi Persamaan berikut :

$$P(C|F_1....F_n) = P(C).P(F_1|C).P(F_2|C).... = P(C) \prod_{i=1}^n P(F_i|C) \dots\dots\dots (2.9)$$

Keterangan :

$P(C) \prod_{i=1}^n P(F_i|C)$ = Perkalian rating antar atribut

Rumus diatas merupakan model dari teorema Naïve Bayes yang berikutnya akan dipergunakan dalam proses prediksi potensi anak. Untuk menangani data numerik yang ada pada data penelitian, metode Naïve Bayes menggunakan persamaan Densitas Gauss seperti pada rumus berikut :

$$P(F_i = f_i | C = c_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(f_i - \mu_{ij})^2}{2\sigma_{ij}^2}} \dots\dots\dots (2.10)$$

Keterangan :

- P = Peluang
- F_i = Atribut ke i
- f_i = Nilai atribut ke i
- C = Kelas yang dicari
- C_j = Sub kelas C yang dicari
- μ = Mean (nilai rata-rata)
- σ = Standar Deviasi

Untuk menghitung nilai *mean* (nilai rata-rata) maka digunakanlah Persamaan berikut :

$$\mu = \frac{\sum_{i=1}^n X_i}{n} \text{ atau } \mu = \frac{X_1 + X_2 + X_3 + \dots + X_{10}}{n} \dots\dots\dots (2.11)$$

Keterangan :

- μ = Mean (nilai rata-rata)
- X_i = Nilai sampai ke i
- N = Jumlah sampel

Sedangkan untuk mencari nilai Standar Deviasi (simpangan baku), digunakanlah perhitungan berikut :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{n-1}} \dots\dots\dots (2.12)$$

Keterangan :

- σ = Standar Deviasi
- μ = *Mean* (nilai rata-rata)
- X_i = Nilai X ke-i
- n = Jumlah sampel

2.6. Contoh Penelitian Terdahulu

Contoh penelitian ini berjudul “Metode Naïve Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)” di buat oleh Asrul Ashari Muin, pada Jurnal Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar. Dalam penelitian ini ada dua dataset yang dipergunakan yaitu :

1. Data Penerimaan Mahasiswa Baru yang memiliki 15 atribut dan 363 record yang terdiri nama, tempat tanggal lahir, asal sekolah, tahun lulus, nama orang tua, nomor ujian, nilai tes tertulis, nilai tes wawancara, nilai rata-rata, ket. Lulus, lulus pilihan dan sebagainya.
2. Data nilai ujian masuk perguruan tinggi yang terdiri dari 5 atribut dan 362 record yang terdiri dari Nomor Ujian, Nilai tes tulisan, nilai tes wawancara, rata-rata dan keterangan Lulus.

Tabel 2.4 Data Hasil Ujian Masuk Perguruan Tinggi

NO. UJIAN	Tes Tertulis		Tes Wawancara	Total	Keterangan Lulus
	Bahasa Indonesia & Inggris	Pengetahuan Umum & Matematika dasar			
0001	27.50	25.00	95.00	49.17	LULUS
0002	27.50	38.33	90.00	51.94	LULUS
0003	25.00	30.00	90.00	48.33	LULUS
0004	17.50	25.00	95.00	45.83	LULUS
0005	27.50	43.33	85.00	51.94	LULUS
0006	20.00	28.33	90.00	46.11	LULUS
0007	30.00	20.00	85.00	45.00	LULUS
0008	15.00	23.33	90.00	42.78	LULUS
0009	32.50	28.33	80.00	46.94	LULUS
0010	32.50	36.67	90.00	53.06	LULUS
0011	27.50	26.67	95.00	49.72	LULUS
0012	15.00	36.67	90.00	47.22	LULUS
0013	22.50	26.67	95.00	48.06	LULUS
0014	27.50	43.33	85.00	51.94	LULUS
0015	22.50	21.67	90.00	44.72	LULUS
0016	20.00	18.33	80.00	39.44	LULUS
0017	17.50	23.33	90.00	43.61	LULUS

2.6.1. Praproses

Tahap pembersihan data adalah langkah awal dari proses KDD. Proses pembersihan data mencakup diantaranya yaitu memeriksa data yang tidak konsisten, data dengan *missing value* dan *redundant* data. Pada tahapan ini tidak hanya membersihkan data yang mengandung *missing value* saja tetapi juga terhadap data yang tidak konsisten juga dilakukan. Data pada penelitian tersebut merupakan data yang sudah konsisten. Karena dua kelompok data (tabel) diambil seluruhnya tidak ada data yang di *cleaning*, maka jumlah atribut dan *record* pada kelompok data (tabel) adalah tetap.

2.6.2. Implementasi Perhitungan pada Naïve Bayes

Setelah data hasil pembersihan selesai maka data tersebut akan di terapkan pada metode naïve bayes klasifikasi. Cara kerja dari proses perhitungan Naïve Bayes klasifikasi adalah sebagai berikut : Tahap pertama yaitu dengan mengambil data sample atau contoh data seperti pada table 2.6 data tes mahasiswa.

Tabel 2.5 Data Hasil Cleaning

A	B	C	D	E	F	G
JURUSAN SEKOLAH	PILIHAN PERTAMA	PILIHAN KEDUA	Nilai Rata	Keterangan Lulus	Pilihan Lulus	Jurusan Lulus
IPA	Sistem Informasi	Ilmu Pemerintahan	55.83	LULUS	PIL 1	Sistem Informasi
IPA	Sistem Informasi	Ilmu Pemerintahan	51.94	LULUS	PIL 1	Sistem Informasi
IPA	Sistem Informasi	Ilmu Komunikasi	50.07	LULUS	PIL 1	Sistem Informasi
TKJ	Sistem Informasi	Teknik Informatika	52.50	LULUS	PIL 1	Sistem Informasi
IPA	Bahasa Indonesia	Teknik Informatika	51.94	LULUS	PIL 1	Bahasa Indonesia

IPA	Teknik Informatika	Peternakan	46.11	LULUS	PIL 2	Peternakan
IPA	Teknik Informatika	Peternakan	45.00	LULUS	PIL 2	Peternakan
IPS	Ilmu Pemerintahan	Sistem Informasi	42.78	LULUS	PIL 1	Ilmu Pemerintahan
KESEHATAN	PPKn	Kesehatan Masyarakat	50.28	LULUS	PIL 1	PPKn
KESEHATAN	PPKn	Teknik Informatika	53.06	LULUS	PIL 1	PPKn
KESEHATAN	PPKn	Teknik Informatika	51.06	LULUS	PIL 1	PPKn
KESEHATAN	Kesehatan Masyarakat	Peternakan	47.22	LULUS	PIL 1	Kesehatan Masyarakat
KESEHATAN	Bahasa Indonesia	Ekonomi Islam	48.06	LULUS	PIL 2	Ekonomi Islam
KESEHATAN	Teknik Informatika	Ilmu Pemerintahan	51.94	LULUS	PIL 1	Teknik Informatika
BAHASA	Ilmu Pemerintahan	Matematika	44.72	LULUS	PIL 1	Ilmu Pemerintahan
IPA	Bahasa Indonesia	Sistem Informasi	50.88	LULUS	PIL 1	Bahasa Indonesia
IPS	PPKn	Ilmu Pemerintahan	43.61	LULUS	PIL 2	Ilmu Pemerintahan
IPA	Agribisnis	Ekonomi Islam	51.39	LULUS	PIL 1	Agribisnis
IPA	Matematika	Sistem Informatika	50.11	LULUS	PIL 1	Matematika
KESEHATAN	Kesehatan Masyarakat	Sistem Informasi	52.22	LULUS	PIL 1	Kesehatan Masyarakat
IPS	Sistem Informasi	Ilmu Pemerintahan	50.28	LULUS	PIL 1	Sistem Informasi
KESEHATAN	Peternakan	Teknik Informatika	19.44	TIDAK LULUS	TIDAK LULUS	TIDAK LULUS
IPS	Teknik Informatika	Ilmu Komunikasi	51.00	LULUS	PIL 1	Teknik Informatika
IPA	PPKn	Kesehatan Masyarakat	51.39	LULUS	PIL 1	PPKn
IPA	PPKn	Kesehatan Masyarakat	53.33	LULUS	PIL 1	PPKn
TKJ	Sistem Informasi	Kesehatan Masyarakat	50.83	LULUS	PIL 2	Kesehatan Masyarakat
IPS	Sistem Informasi	Matematika	50.34	LULUS	PIL 1	Sistem Informasi

Tabel 2.6 Data Tes Mahasiswa

JURUSAN SEKOLAH	PILIHAN PERTAMA	PILIHAN KEDUA	Nilai Rata	Keterangan Lulus	Pilihan Lulus
IPA	Sistem Informasi	Ilmu Pemerintahan	55.83	LULUS	PIL 1
BAHASA	Sistem Informasi	Ilmu Pemerintahan	42.78	LULUS	PIL 2
IPS	Sistem Informasi	Ilmu Pemerintahan	50.28	LULUS	PIL 1
TKJ	Sistem Informasi	Teknik Informatika	52.50	LULUS	PIL 1
IPS	Ilmu Pemerintahan	Sistem Informasi	42.78	LULUS	PIL 1
KESEHATAN	Kesehatan Masyarakat	Peternakan	47.22	LULUS	PIL 1
BAHASA	Ilmu Pemerintahan	Matematika	44.72	LULUS	PIL 1
PERKANTORAN	Ilmu Pemerintahan	Ekonomi Islam	47.50	LULUS	PIL 1
IPS	PPKn	Ilmu Pemerintahan	43.61	LULUS	PIL 2
IPS	Bahasa Indonesia	Ilmu Pemerintahan	44.44	LULUS	PIL 2
IPA	Matematika	Sistem Informasi	50.11	LULUS	PIL 1
KESEHATAN	Kesehatan Masyarakat	Sistem Informasi	52.22	LULUS	PIL 1
IPS	Sistem Informasi	Ilmu Pemerintahan	50.28	LULUS	PIL 1

Selanjutnya dari tabel diatas bisa ditentukan dengan memakai metode naïve bayes klasifikasi, cara kerjanya sebagai berikut :

Rumusnya adalah :

$$P(X|H) = \frac{P(X|H)P(H)}{P(X)} \dots\dots\dots (2.13)$$

Sehingga,

- P(Pil1) = 10/13 = 0.77

- $P(\text{Pil2}) = 3/13 = 0.23$
- $P(\text{Prodi=IPA|Pil1}) = 2/10 = 0.20$
- $P(\text{Prodi=IPA|Pil2}) = 0/3 = 0.00$
- $P(\text{Pilihan Pertama = Sistem Informasi|Pil1}) = 4/10 = 0.40$
- $P(\text{Pelican Pertama = Sistem Informasi|Pil2}) = 1/3 = 0.33$
- $P(\text{Pilihan Kedua = Ilmu Pemerintahan|Pil1}) = 3/10 = 0.30$
- $P(\text{Pilihan Kedua=Ilmu Pemerintahan|Pil2}) = 3/3 = 1.00$
- $P(\text{Nilai Rata = 55.83|Pil1}) = 1/10 = 0.10$
- $P(\text{Nilai Rata = 55.83|Pil1}) = 0/3 = 0.00$

Kemudian,

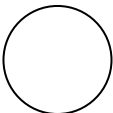

- $P(\text{Pil1}) P(\text{IPA|Pil1}) P(\text{Sistem Informasi|Pil1}) P(\text{Ilmu Pemerintahan|Pil1}) P(55.83|Pil1) = 0.77 \times 0.20 \times 0.40 \times 0.30 \times 0.10 = 0.00203$
- $P(\text{Pil2}) P(\text{IPA|Pil2}) P(\text{Sistem Informasi|Pil2}) P(\text{Ilmu Pemerintahan|Pil2}) P(55.83|Pil2) = 0.23 \times 0 \times 0.33 \times 1 \times 0 = 0$

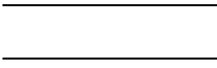
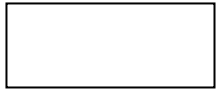
Jadi hasil dari data Lulus adalah Pil 1 (Pertama)

2.7. Data Flow Diagram (DFD)

Data Flow Diagram (DFD) adalah model logika atau proses yang dibuat untuk menggambarkan darimana asal data dan kemana tujuan data yang keluar dari sistem, yang mana data disimpan, proses apa yang menghasilkan data tersebut dan interaksi antara data yang tersimpan dan proses yang dikenakan pada data tersebut (Waden, 2010).

Tabel 2.7 Data Flow Diagram (DFD)

Simbol	Keterangan
	Process/buble Mempunyai fungsi untuk Menggambarkan transformasi aliran data masuk menjadi aliran data keluar.
	Data Flow Mempunyai fungsi untuk menggambarkan aliran

	data atau paket informasi dari satu bagian sistem ke bagian lain. Arah panah menggambarkan aliran data.
	Data Store Memiliki fungsi untuk menggambarkan model data yang tersimpan
	Eksternal Entity/terminator Memiliki fungsi untuk menggambarkan kesatuan luar yang berhubungan dengan system.

2.8. Sybase Power Designer

Power Designer merupakan salah satu alat yang dapat digunakan untuk membuat atau merancang sebuah basis data melalui ER-diagram. Dalam Power Designer, ada bermacam-macam pemodelan data yang dipergunakan untuk merancang basis data diantaranya yaitu, *Conceptual Data Model* (CDM) merupakan model yang dibuat berdasarkan anggapan bahwa dunia nyata terdiri dari koleksi obyek dasar yang dinamakan entitas (*entity*) serta hubungan (*relationship*) antara entitas itu. *Physical Data Model* (PDM) merupakan model yang menggunakan sejumlah tabel untuk menggambarkan data serta hubungan antara data tersebut. (NM Rosyiddin, 2016)

2.9. Microsoft Visual Basic Studio

Microsoft Visual Basic Studio adalah sebuah bahasa pemrograman yang berupa *Integrated Development Environment* (IDE) tampilan untuk membuat program *software* berbasis sistem operasi Microsoft Windows dengan menggunakan model pemrograman (COM). Visual Basic merupakan turunan bahasa pemrograman BASIC dan menawarkan pengembangan perangkat lunak komputer berbasis grafik dengan cepat. Beberapa bahasa skrip seperti *Visual Basic for Applications* (VBA) dan *Visual Basic Scripting Edition* (VBScript), mirip seperti halnya Visual Basic, tetapi cara kerjanya yang berbeda (Phil, Jones, 2001).

2.10. Structured Query Language (SQL)

SQL atau bisa disebut dengan *Structured Query Language* SQL merupakan suatu bahasa pemrograman yang dipergunakan untuk mengakses data di dalam sebuah database. SQL disebut juga dengan istilah *query*, dan bahasa SQL secara praktiknya dipergunakan sebagai bahasa standar untuk manajemen database. Hingga saat ini hampir seluruh server database atau software database mengenal dan mengerti bahasa SQL. (Dedi, 2017)