

## BAB II

### LANDASAN TEORI

#### 2.1 SISTEM PENDUKUNG KEPUTUSAN

Dalam bukunya yang berjudul “*Komputerisasi Pengambilan Keputusan*”, Dadan Umar Daihani (2001) menjelaskan bahwa konsep sistem penunjang keputusan (SPK)/*Decision Support System* (DSS) pertama kali diungkapkan pada awal tahun 1970-an oleh Michael S. Scott Marton dengan istilah *management decision system*. Terdapat banyak definisi mengenai SPK, antara lain :

a. Mann dan Waston .

Sistem berbasis komputer interaktif, yang membantu para pengambil keputusan untuk menggunakan data dan berbagai model untuk memecahkan masalah-masalah semi terstruktur dan tidak terstruktur.

b. Maran Alavi dan H. Albert Napier.

Suatu kumpulan prosedur pemrosesan data dan informasi yang berorientasi pada penggunaan model untuk menghasilkan berbagai jawaban yang dapat membantu manajemen dalam pengambilan keputusan. Sistem ini harus sederhana, mudah dan adaptif.

c. Little.

Suatu sistem informasi berbasis komputer yang menghasilkan berbagai alternatif keputusan untuk membantu manajemen dalam menangani berbagai permasalahan yang terstruktur ataupun tidak terstruktur dengan menggunakan data dan model.

d. Raymond McLeod, Jr.

Sistem penghasil informasi spesifik yang ditujukan untuk memecahkan suatu masalah tertentu yang harus dipecahkan oleh manager pada berbagai tingkatan.

Dari berbagai definisi di atas dapat dikatakan bahwa sistem pendukung keputusan adalah suatu sistem informasi spesifik yang ditujukan untuk membantu manajemen dalam mengambil keputusan yang berkaitan dengan persoalan yang

bersifat semi struktur. Sistem ini memiliki fasilitas untuk menghasilkan berbagai alternatif yang secara interaktif dapat digunakan oleh pemakai.

Kata kunci lainnya adalah penggunaan model sebagai dasar pengembangan alternatif. Penggunaan model ini berkaitan dengan sifat permasalahan yang harus dipecahkan pemakai yaitu semi terstruktur atau bahkan tidak terstruktur. Jadi semakin banyak pembendaharaan model yang dimiliki oleh sistem, maka alternatif keputusan yang dapat diciptakan juga semakin kaya.

Ciri lain dari sistem ini adalah pemanfaatan komputer sebagai motor penggeraknya. Oleh karena itu, sering kali disebutkan bahwa sistem pendukung keputusan adalah sistem yang berbasis komputer (*Computer based system*).

### **2.1.1 Karakteristik dan Nilai Guna SPK**

Dalam bukunya yang berjudul “*Komputerisasi Pengambilan Keputusan*”, Dadan Umar Daihani (2001) menjelaskan bahwa karakteristik sistem pendukung keputusan adalah :

- a. Sistem Pendukung Keputusan dirancang untuk membantu pengambil keputusan dalam memecahkan masalah yang sifatnya semi terstruktur ataupun tidak terstruktur dengan menambahkan kebijaksanaan manusia dan informasi komputerisasi.
- b. Dalam proses pengolahannya, sistem pendukung keputusan mengkombinasikan penggunaan model-model analisis dengan teknik pemasukan data konvensional serta fungsi-fungsi pencari / interogasi informasi.
- c. Sistem Pendukung Keputusan, dirancang sedemikian rupa sehingga dapat digunakan/dioperasikan dengan mudah.
- d. Sistem Pendukung Keputusan dirancang dengan menekankan pada aspek fleksibilitas serta kemampuan adaptasi yang tinggi.

Dengan berbagai karakter khusus diatas, SPK dapat memberikan berbagai manfaat dan keuntungan. Manfaat yang dapat diambil dari SPK adalah (Turban, 2005) :

- a. SPK memperluas kemampuan pengambil keputusan dalam memproses data/informasi bagi pemakainya.

- b. SPK membantu pengambil keputusan untuk memecahkan masalah terutama berbagai masalah yang sangat kompleks dan tidak terstruktur.
- c. SPK dapat menghasilkan solusi dengan lebih cepat serta hasilnya dapat diandalkan.
- d. Walaupun suatu SPK, mungkin saja tidak mampu memecahkan masalah yang dihadapi oleh pengambil keputusan, namun ia dapat menjadi stimulan bagi pengambil keputusan dalam memahami persoalannya, karena sistem pendukung keputusan mampu menyajikan berbagai alternatif .
- e. Sistem pendukung keputusan dapat menyediakan bukti tambahan untuk memberikan pembenaran sehingga dapat memperkuat posisi pengambil keputusan.

Di samping berbagai keuntungan dan manfaat seperti dikemukakan diatas, SPK juga memiliki beberapa keterbatasan, diantaranya adalah (Dadan Umar Daihani, 2001) :

- a. Ada beberapa kemampuan manajemen dan bakat manusia yang tidak dapat dimodelkan, sehingga model yang ada dalam sistem tidak semuanya mencerminkan persoalan sebenarnya.
- b. Kemampuan suatu SPK terbatas pada perbendaharaan pengetahuan yang dimilikinya (pengetahuan dasar serta model dasar).
- c. Proses-proses yang dapat dilakukan SPK biasanya juga tergantung pada perangkat lunak yang digunakan.
- d. SPK tidak memiliki kemampuan intuisi seperti yang dimiliki oleh manusia. Karena walau bagaimanapun canggihnya suatu SPK, SPK hanyalah suatu kumpulan perangkat keras, perangkat lunak dan system yang mengoprasikannya yang tidak dilengkapi dengan kemampuan berpikir.

### **2.1.2 Komponen Sistem Pendukung Keputusan**

Dalam bukunya yang berjudul “*Komputerisasi Pengambilan Keputusan*”, Dadan Umar Daihani (2001) menjelaskan bahwa sistem pendukung keputusan terdiri atas tiga komponen utama yaitu :

a. Sub sistem pengelolaan data (*database*).

Sub sistem pengelolaan data (*database*) merupakan komponen SPK yang berguna sebagai penyedia data bagi sistem. Data tersebut disimpan dan diorganisasikan dalam sebuah basis data yang diorganisasikan oleh suatu sistem yang disebut dengan sistem manajemen basis data (*Database Management System*).

b. Sub sistem pengelolaan model (*modelbase*).

Keunikan dari SPK adalah kemampuannya dalam mengintegrasikan data dengan model-model keputusan. Model adalah suatu tiruan dari alam nyata. Kendala yang sering dihadapi dalam merancang suatu model adalah bahwa model yang dirancang tidak mampu mencerminkan seluruh variabel alam nyata, sehingga keputusan yang diambil tidak sesuai dengan kebutuhan oleh karena itu, dalam menyimpan berbagai model harus diperhatikan dan harus dijaga fleksibilitasnya. Hal lain yang perlu diperhatikan adalah pada setiap model yang disimpan hendaknya ditambahkan rincian keterangan dan penjelasan yang komprehensif mengenai model yang dibuat.

c. Sub sistem pengelolaan dialog (*userinterface*).

Keunikan lainnya dari SPK adalah adanya fasilitas yang mampu mengintegrasikan sistem yang terpasang dengan pengguna secara interaktif, yang dikenal dengan subsistem dialog. Melalui subsistem dialog, sistem diimplementasikan sehingga pengguna dapat berkomunikasi dengan sistem yang dibuat.

Sedangkan fasilitas yang dimiliki oleh subsistem dialog dibagi menjadi tiga komponen :

- 1 Bahasa aksi (*action language*), yaitu suatu perangkat lunak yang dapat digunakan oleh user untuk berkomunikasi dengan sistem, yang dilakukan melalui berbagai pilihan media seperti *keyboard*, *joystick* dan *keyfunction* yang lainnya.
- 2 Bahasa tampilan (*display and presentation language*), yaitu suatu perangkat yang berfungsi sebagai sarana untuk menampilkan

sesuatu. Peralatan yang digunakan untuk merealisasikan tampilan ini diantaranya adalah printer, grafik monitor, plotter, dan lain-lain.

- 3 Basis pengetahuan (*knowledge base*), yaitu bagian yang mutlak diketahui oleh pengguna sehingga sistem yang dirancang dapat berfungsi secara interaktif.

## **2.2 PROFIL SUMMIT OTO FINANCE**

Didirikan tahun 1990, pada awalnya perusahaan ini bernama PT Summit Sinar Mas Finance, hasil kerjasama usaha antara PT Sinar Mas Multiartha dan Sumitomo Corporation Jepang. Awalnya PT Summit Sinar Mas Finance memfokuskan aktivitas usaha pada sewa guna usaha. Namun ditahun 2003 PT Summit Sinar Mas Finance mengubah aktivitas usahanya menjadi perusahaan pembiayaan kendaraan bermotor, sekaligus mengganti namanya menjadi PT Summit Oto Finance.

Sumitomo Corporation adalah perusahaan dagang Jepang yang terpadu (*sogo shosha*). Sebagai pemegang saham utama, Sumitomo Corporation memberikan dukungan dan mengendalikan semua aspek usaha dari manajemen, treasury, keuangan hingga operasional. Dengan dukungan dari Sumitomo Corporation, dalam lima tahun terakhir PT Summit Oto Finance telah berhasil tumbuh dan meningkatkan pembiayaan motor serta memiliki kantor jaringan yang tersebar di seluruh Indonesia.

Usaha utama PT Summit Oto Finance adalah pada pembiayaan kepemilikan motor baru. Target utama PT Summit Oto Finance lebih kepada pelanggan perorangan daripada perusahaan; dengan tujuan penyebaran risiko. Sebagai perusahaan pembiayaan yang independen, PT Summit Oto Finance tidak memiliki keterkaitan dengan pabrikan, sehingga perusahaan memiliki keleluasaan untuk membiayai semua merek motor yang ada di pasar. PT Summit Oto Finance juga telah menikmati pertumbuhan pasar motor domestik yang kuat dalam beberapa tahun terakhir, serta mampu mempertahankan posisinya sebagai salah satu pemain terkemuka dalam pembiayaan motor.

Dalam usaha menyediakan layanan “one-stop service”, PT Summit Oto Finance mengembangkan web site ([www.otofinance.co.id](http://www.otofinance.co.id)). PT Summit Oto Finance juga terus memperkuat sistem Teknologi Informasi dengan cara meningkatkan efisiensi dan produktivitas di kantor-kantor cabang untuk melayani pelanggan. PT Summit Oto Finance senantiasa berkomitmen memberikan layanan terbaik kepada para pelanggan. Sampai akhir 2011 PT Summit Oto Finance telah mengoperasikan 177 kantor yang tersebar di Indonesia. Selain itu juga Perusahaan telah bekerjasama dengan bank-bank berjangkauan nasional dan PT Pos Indonesia untuk penerimaan pembayaran angsuran, sehingga para Pelanggan semakin mudah dan nyaman membayar angsurannya. Sumitomo Corporation sebagai pemegang saham utama PT Summit Oto Finance, berkomitmen untuk selalu memberikan dukungan pada PT Summit Oto Finance, baik dalam hal manajemen, pendanaan, pemasaran maupun operasional perusahaan.

Didukung modal yang kuat, tim manajemen yang handal, mitra usaha, pelanggan setia dan 12.313 karyawan, di tahun 2011 PT Summit Oto Finance telah berhasil tumbuh dengan Total Aset sebesar Rp 11,4 triliun, Total Ekuitas tumbuh menjadi Rp 2,8 triliun serta Laba Bersih tercatat sebesar Rp 22,0 miliar. Pada bulan Desember 2011 Pemeringkat Efek Indonesia (PEFINDO) memberikan peringkat idAA (Double A; Negative Outlook) untuk Perseroan dan Obligasi Summit Oto Finance. Pemegang Saham Pada 31 Desember 2011 komposisi pemegang saham Summit Oto Finance disajikan seperti **Tabel 2.1** :

**Tabel 2.1** Pemegang Saham Summit Oto Finance

<b>Pemegang Saham</b>	<b>%</b>
Sumitomo Corporation	85,00
PT Summit Auto Group	4.56
PT Sumitomo Indonesia	10,00
Djohan Marzuki	0.28
PT Sinar Mas Multiartha, Tbk	0.16
Jumlah	100,00

## **2.3 Prosedur Pengajuan Kredit**

### **2.3.1 Pengertian Kredit**

Pasal 1 (11) UU NO.10/1998, Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga.

Menurut Thomas Suyatno (1998), kredit ialah kredit adalah hak untuk menerima pembayaran atau kewajiban untuk melakukan pembayaran pada waktu diminta atau pada waktu yang akan datang karena penyerahan barang-barang sekarang.

Kredit menurut Rivai dan Veithzal (2007:4) adalah penyerahan barang, jasa atau uang dari satu pihak (kreditur/atau pemberi pinjaman) atas dasar kepercayaan kepada pihak lain (nasabah atau pengutang atau *borrower*) dengan janji membayar dari penerima kredit kepada pemberi kredit pada tanggal yang telah disepakati kedua belah pihak.

### **2.3.2 Pengajuan Kredit.**

Dalam pengajuan kredit kendaraan bermotor bisa dilakukan dengan dua cara yaitu, UINSA (<http://digilib.uinsby.ac.id/7960/6/BAB%20III.pdf> : 12 Januari 2015) :

- a. Dengan mendatangi stand atau dealer kendaraan

Disini mereka akan dibantu oleh pihak dealer dalam proses pengajuan kredit kendaraan bermotor baik itu dari apa-apa yang harus dipenuhi yang kemudian dari pihak dealer memberikan data pihak yang akan mengajukan kredit tersebut kepada pihak *leasing* untuk diproses dan dilakukan survei.

- b. Dengan mendatangi kantor *leasing*

Disini mereka akan langsung dilayani oleh bagian marketing dan diproses serta dilakukan survei, baik itu survei tempat tinggal ataupun lokasi sekitar.

Ada beberapa tahapan yang harus dipenuhi di mana jika di antara beberapa tahapan tersebut tidak terpenuhi maka akan memungkinkan tidak terkabulnya (ACC) pengajuan tersebut. Hal ini sangat penting terutama bagi mereka yang hendak mengajukan permohonan kredit kendaraan bermotor.

Adapun syarat atau tahapan yang harus dipenuhi oleh pihak pemohon (perorangan) adalah:

- a. Foto copy KTP pemohon dan penjamin
- b. Foto copy KK
- c. Foto copy Rekening Listrik / PDAM / PBB
- d. Slip gaji / Rekening tabungan / Surat Ket Penghasilan

Perlu diketahui, meski semua perlengkapan di atas terpenuhi semua hal itu bukan menjadi jaminan akan terkabulnya pengajuan kredit tersebut, akan tetapi dengan terpenuhinya beberapa perlengkapan tersebut akan mempermudah pihak *leasing* untuk memberikan kredit.

Dalam pengajuan kredit ini sangat mudah dan tidak harus menunggu beberapa hari, karena dengan adanya fasilitas yang memadai dari pihak perusahaan akan memperlancar jalannya proses tersebut, mungkin sebelumnya diketahui bahwa dalam pengajuan kredit butuh waktu beberapa hari dari proses pengajuan untuk mengetahui hasil dari pihak perusahaan, akan tetapi pada saat ini tidak harus menunggu lama, cukup sehari saja, dari proses pengajuan. Artinya kalau sekarang mengajukan kredit maka besok siang akan mengetahui bagaimana hasilnya.

### **2.3.3 Penjelasan Kriteria**

Berikut ini dijelaskan beberapa kriteria yang akan digunakan dalam penelitian tugas akhir yaitu :

#### **1. Alamat Tinggal**

Kriteria alamat tinggal sangat penting pengaruhnya dalam penentuan nasabah itu diterima atau ditolak, karena kebanyakan nasabah yang diterima adalah nasabah yang tinggal rumah sendiri dan yang ditolak adalah yang kontrak atau kos.



## 2. Bukti Kepemilikan Rumah

Bukti kepemilikan rumah ini sangat penting untuk mengetahui siapa pemilik rumah yang di huni saat ini , apakah milik sendiri/miliksendiri tapi masih atas nama orang lain/ milik keluarga /kontrak .yang di prioritaskan dalam pemberian kredit ada rumah sendiri karena di sebabkan pemilik rumah otomatis tetap tinggal disitu dalam jangka waktu yang lama.

## 3. Pemakaian Unit

Pemakaian unit juga salah satu faktor yang mempengaruhi dalam pemberian kredit. Pemakaian sendiri merupakan kriteria yang ada pada nasabah dibandingkan pemakaian yang lainnya.

## 4. Penghasilan Perbulan

Berdasarkan fakta yang ada bahwa sebagian besar masyarakat ekonomi menengah kebawah banyak yang berminat dalam pembelian sepeda secara kredit,akan tetapi dalam pengajuan kredit besarnya angsuran harus 1/3 dari penghasilan tiap bulannya.

## 5. Usia Calon Debitur

Usia calon debitur berpengaruh dalam sebuah proses pengajuan kredit,usia 21-25 th adalah prioritas paling rendah ,25-40 adalah prioritas paling tinggi ,40-60 prioritas sedang.

## 6. Karakteristik Calon Debitur

Karakter seorang calon debitur sangat berpengaruh dalam sebuah pengajuan sebuah kredit, Surveyor harus pintar – pintar dalam menilai karakter seorang calon debitur . apakah seseorang itu layak di beri kredit apa tidak.

## 2.4 Teorima

### 2.4.1 Teorema FDR

Menurut Eko Prasetyo (2014:29) Fisher's Discriminant Ratio (FDR) umumnya dipakai untuk mengukur kekuatan diskriminasi fitur individu dalam memisahkan dua kelas berdasarkan nilai yang dimilikinya. Untuk  $\mu_1$  dan  $\mu_2$  masing-masing adalah nilai rata-rata dari dua kelas, sedangkan  $\sigma_1^2$  dan  $\sigma_2^2$  masing-masing adalah varian dari dua kelas dalam fitur yang akan diukur. FDR dirumuskan seperti persamaan berikut :

$$\text{FDR} = \frac{(\mu_1 - \mu_2)^2}{(\sigma_1^2 + \sigma_2^2)}$$

Hasil yang diberikan oleh FDR adalah untuk fitur yang mempunyai perbedaan yang besar pada rata-rata dari kelas dan varian kecil dari tiap kelas, maka nilai FDR yang tinggi akan didapatkan. Jika dua fitur mempunyai perbedaan absolut rata-rata yang sama tapi berbeda dalam jumlah varian ( $\sigma_1^2 + \sigma_2^2$ ), maka fitur dengan jumlah varian yang lebih kecil akan mendapatkan nilai FDR yang lebih tinggi. Di sisi lain, jika dua fitur mempunyai jumlah varian yang sama tetapi perbedaan absolut rata-rata berbeda, maka fitur dengan perbedaan absolut rata-rata yang lebih besar akan mendapatkan nilai FDR yang lebih tinggi.

Contoh pemilihan fitur dengan FDR pada satu fitur

Sebuah fitur dari dua kelas adalah sebagai berikut :

Kelas  $\omega_1$  : 3.5 3.7 3.9 4.1 3.4 3.5 4.1 3.8 3.6 3.7

Kelas  $\omega_2$  : 3.2 3.6 3.1 3.4 3.0 3.4 2.8 3.1 3.3 3.6

Nilai rata-rata yang didapat:

$\mu_1$  : 3.73 ,  $\mu_2$  : 3.25

Nilai varian yang didapat :

$\sigma_1^2$  : 0.0601 ,  $\sigma_2^2$  : 0.0672

$$\text{Maka didapat } FDR = \frac{(3.73 - 3.25)^2}{(0.0601 + 0.0672)} = 1.8099$$

Contoh pemilihan fitur dengan FDR pada empat fitur

Perhatikan data sintetik pada tabel dibawah ini. Data tersebut berisi 20 data yang terbagi menjadi 2 kelas, setiap data mempunyai 4 fitur : A, B, C, dan D. Dari keempat fitur tersebut, akan dihitung FDR sehingga bisa diketahui nama dari 4 fitur tersebut yang mempunyai diskriminasi kelas yang lebih baik.

Tabel 2.2 data empat fitur

<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>Kelas</b>
154.833	29.95	1.424	1.133	1
163.464	31.30	1.554	1.476	1
164.703	21.6	1.764	1.605	1
148.176	28.88	1.644	1.386	1
147.567	30.04	2.014	1.881	1
151.011	26.02	1.674	1.333	1
149.835	27.63	2.074	2.195	1
123.942	27.07	1.704	1.195	1
141.456	25.71	1.484	1.295	1
152.082	30.56	2.154	2.533	1
211.785	34.06	2.344	2.986	2
234.717	35.54	1.424	1.090	2
239.673	35.12	1.294	1.167	2
207.207	29.84	1.474	1.290	2
203.616	30.01	1.554	1.171	2
233.793	29.61	1.514	1.314	2
240.996	32.27	2.765	1.210	2
257.691	37.50	2.134	2.252	2

223.524	31.23	1.504	1.157	2
217.056	31.54	1.414	1.271	2

Rata-rata dari keempat fitur untuk kelas  $\omega_1$  didapat:

[149.7069 27.8760 1.7490 1.6032]

Rata-rata dari keempat fitur untuk kelas  $\omega_2$  didapat:

[227.0058 32.6720 1.7421 1.4908]

Varian dari keempat fitur untuk kelas  $\omega_1$  didapat:

[131.8451 8.5727 0.0636 0.2123]

Varian dari keempat fitur untuk kelas  $\omega_2$  didapat:

[298.3515 7.4954 0.2431 0.3881]

Maka **FDR** sebagai berikut:

[13.8893 1.4315 0.0002 0.0210]

**Secara urut, fitur yang paling informatif: A-B-D-C**

Dari hasil FDR yang didapat, dapat diamati bahwa fitur A mempunyai nilai FDR tertinggi. Artinya fitur A mempunyai informasi diskriminasi yang paling baik terhadap kedua kelas, kemudian fitur B dan D adalah fitur terbaik berikutnya. Sementara untuk fitur C tidak mempunyai informasi diskriminasi yang cukup baik pada kedua kelas.

### 2.4.2 Teorema Gain

Algoritma 1D3 (Iterative Dichotomiser 3) pertama kali diperkenalkan oleh Quinlan (Quinlan, 1986) yang digunakan untuk menginduksi decision tree. Algoritma 1D3 dapat bekerja dengan baik pada semua fitur yang mempunyai tipe data kategorikal (nominal atau ordinal). Dalam perkembangannya, 1D3 banyak mengalami perbaikan pada versi berikutnya seperti C4.5, atau C5.0.

#### Algoritma Induksi Decision Tree 1D3

---

1. Dimulai dari node akar.
  2. Untuk semua fitur, hitung nilai entropy untuk semua sampel (pola latih) pada node.
  3. Pilih fitur dengan informasi gain yang maksimal.
  4. Gunakan fitur tersebut sebagai node pemecahan menjadi cabang.
  5. Lakukan secara rekursif pada setiap cabang yang dibuat dengan mengulangi langkah 2 sampai 4 hingga semua data dalam setiap node hanya memberikan satu label kelas. Node yang tidak dapat dipecah lagi merupakan daun yang berisi keputusan (label kelas).
- 

Entropy dapat dihitung menggunakan persamaan berikut :

$$E(s) = - \sum_{i=1}^m p(\omega_i | s) \log_2 p(\omega_i | s)$$

Dimana untuk  $P(\omega_i | s)$  adalah proporsi kelas ke- $i$  dalam semua data latih yang diproses di node  $s$ , didapatkan dan jumlah semua baris data dengan label kelas  $i$  dibagi jumlah baris semua data. Sedangkan  $m$  adalah jumlah nilai berbeda dalam data. Entropy digunakan untuk menentukan yang manakah node yang akan menjadi pemecah data latih berikutnya. Nilai entropy yang lebih tinggi akan meningkatkan potensi klasifikasi. Yang perlu diperhatikan, adalah jika entropy untuk node bernilai 0 berarti semua data vektor berada pada label kelas yang sama dan node tersebut menjadi daun yang berisi keputusan (label kelas).

Yang juga perlu diperhatikan dalam perhitungan entropy adalah jika salah satu dari elemen  $\omega_i$  jumlahnya 0 maka entropy dipastikan 0 juga. Jika proporsi semua elemen  $\omega_i$  sama jumlahnya maka dipastikan entropy dipastikan bernilai 1.

Gain digunakan untuk memperkirakan pemilihan fitur yang tepat untuk menjadi pemecah pada node tersebut. Gain sebuah fitur  $ke-j$  dihitung menggunakan persamaan berikut.

$$G(s, j) = E(s) - \sum_{i=1}^n p(v_i | s) \times E(s_i)$$

Dimana  $P(v_i|s)$  adalah proporsi nilai  $v$  muncul pada kelas dalam node. Sedangkan  $E(s_i)$  adalah entropy komposisi nilai  $v$  dan kelas  $ke-j$  dalam data ke- $i$  node tersebut.

Sedangkan  $n$  adalah jumlah nilai berbeda dalam node.

### Contoh

Data latih pola pada tabel 2.3 digunakan untuk melakukan prediksi “apakah harus bermain baseball ?” dengan menjawab Ya atau Tidak. Fitur yang digunakan ada 4: cuaca, suhu, kelembaban, dan angin. Nilai fitur menggunakan tipe kategorikal, rinciannya sebagai berikut:

- Cuaca = {cerah, mendung, hujan}
- Suhu = {panas, lembut, dingin}
- Kelembaban = {tinggi, normal}
- Angin = {pelan, kencang}

Tabel 2.3 Data pola klasifikasi bermain baseball

Cuaca	Suhu	Kelembaban	Angin	Bermain
Cerah	Panas	Tinggi	Pelan	Tidak
Cerah	Panas	Tinggi	Kencang	Tidak
Mendung	Panas	Tinggi	Pelan	Ya
Hujan	Lembut	Tinggi	Pelan	Ya
Hujan	Dingin	Normal	Pelan	Ya
Hujan	Dingin	Normal	Kencang	Tidak
Mendung	Dingin	Normal	Kencang	Ya
Cerah	Lembut	Tinggi	Pelan	Tidak
Cerah	Dingin	Normal	Pelan	Ya
Hujan	Lembut	Normal	Pelan	Ya
Cerah	Lembut	Normal	Kencang	Ya
Mendung	Lembut	Tinggi	Kencang	Ya
Mendung	Panas	Normal	Pelan	Ya
Hujan	Lembut	Tinggi	Kencang	Tidak

Dimulai dari node akar, harus dihitung dulu entropy untuk node akar (semua data) terhadap komposisi kelas.

$$\begin{aligned} E(\text{semua}) &= -\left( \begin{aligned} &(p(\text{ya} \mid \text{semua}) \times \log_2 p(\text{ya} \mid \text{semua})) + \\ &(p(\text{tidak} \mid \text{semua}) \times \log_2 p(\text{tidak} \mid \text{semua})) \end{aligned} \right) \\ &= -\left( (9/14) \times \log_2 (9/14) + (5/14) \times \log_2 (5/14) \right) \\ &= 0.9403 \end{aligned}$$

Selanjutnya dihitung entropy untuk setiap nilai fitur terhadap kelas, untuk entropy nilai dalam Cuaca, didapat:

$$\begin{aligned} E(\text{semua}_{\text{cerah}}) &= -\left( \begin{aligned} &(p(\text{ya} \mid \text{cerah}) \times \log_2 p(\text{ya} \mid \text{cerah})) \\ &+ (p(\text{tidak} \mid \text{cerah}) \times \log_2 p(\text{tidak} \mid \text{cerah})) \end{aligned} \right) \\ &= -\left( (2/5) \times \log_2 (2/5) + (3/5) \times \log_2 (3/5) \right) \\ &= 0.9710 \end{aligned}$$

$$\begin{aligned} E(\text{semua}_{\text{mendung}}) &= -\left( \begin{aligned} &(p(\text{ya} \mid \text{mendung}) \times \log_2 p(\text{ya} \mid \text{mendung})) \\ &+ (p(\text{tidak} \mid \text{mendung}) \times \log_2 p(\text{tidak} \mid \text{mendung})) \end{aligned} \right) \\ &= -\left( (4/4) \times \log_2 (4/4) + (0/4) \times \log_2 (0/4) \right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} E(\text{semua}_{\text{hujan}}) &= -\left( \begin{aligned} &(p(\text{ya} \mid \text{hujan}) \times \log_2 p(\text{ya} \mid \text{hujan})) \\ &+ (p(\text{tidak} \mid \text{hujan}) \times \log_2 p(\text{tidak} \mid \text{hujan})) \end{aligned} \right) \\ &= -\left( (3/5) \times \log_2 (3/5) + (2/5) \times \log_2 (2/5) \right) \\ &= 0.9710 \end{aligned}$$

Entropy tersebut hitung pada setiap fitur. Selanjutnya dihitung Gain untuk setiap fitur. Dan perhitungan, didapat Gain seperti dibawah ini. Selengkapnya hasil perhitungan entropy dan gain untuk node akar disajikan di tabel 2.7.



$$\begin{aligned}
G(\text{semua}, \text{cuaca}) &= E(\text{semua}) - \sum_{i=1}^n p(v_i | \text{semua}) \times E(\text{semua}_{\text{cuaca}}) \\
&= E(\text{semua}) - \left( \begin{aligned} &(p(\text{cerah} | \text{semua}) \times E(\text{semua}_{\text{cerah}})) + \\ &(p(\text{mendung} | \text{semua}) \times E(\text{semua}_{\text{mendung}})) + \\ &(p(\text{hujan} | \text{semua}) \times E(\text{hujan}_{\text{cerah}})) \end{aligned} \right) \\
&= 0.9403 - ((5/14) \times 0.9710) + ((4/14) \times 0) + ((5/14) \times 0.9710) \\
&= 0.2467
\end{aligned}$$

$$G(\text{semua}, \text{suhu}) = 0.0292$$

$$G(\text{semua}, \text{kelembaban}) = 0.1518$$

$$G(\text{semua}, \text{angin}) = 0.0481$$

Tabel 2.4 Hasil perhitungan entropy dan gain untuk node akar

Node			Jumlah	Ya	Tidak	Entropy	Gain
1	Total		14	9	5	0.9403	
	Cuaca						0.2467
		Cerah	5	2	3	0.9710	
		Mendung	4	4	0	0	
		Hujan	5	3	2	0.9710	
	Suhu						0.0292
		Panas	4	2	2	1.0000	
		Lembut	6	4	2	0.9183	
		Dingin	4	3	1	0.8113	
	Kelembaban						0.1518
		Tinggi	7	3	4	0.9852	
		Normal	7	6	1	0.5917	
	Angin						0.0481
		Pelan	8	6	2	0.8113	
		Kencang	6	3	3	1.0000	

Hasil yang didapat di tabel 2.4 menunjukkan bahwa gain tertinggi ada di fitur cuaca.

### 2.4.3 Teorema Bayes

Menurut Bustami (2013) Algoritma *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya.

Persamaan dari teorema *Bayes* adalah :

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

Keterangan:

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$  : Probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$  : Probabilitas hipotesis H (prior probability)

$P(X|H)$  : Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$  : Probabilitas X

Untuk menjelaskan teorema *Naive Bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema *bayes* di atas

disesuaikan sebagai berikut :

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

Dimana Variabel  $C$  merepresentasikan kelas, sementara variabel  $F_1 \dots F_n$  merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas  $C$  (*Posterior*) adalah peluang munculnya kelas  $C$  (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas  $C$  (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga *evidence*). Karena itu, rumus diatas dapat pula ditulis secara sederhana sebagai berikut:

$$\text{Posterior} = \frac{\text{Prior x likelihood}}{\text{evidence}}$$

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai - nilai *posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus *Bayes* tersebut dilakukan dengan menjabarkan  $(C|F_1, \dots, F_n)$  menggunakan aturan perkalian sebagai berikut :

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C) P(F_1, \dots, F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2), P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned}$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin 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 (*naif*), bahwa masing - masing petunjuk ( $F_1, F_2 \dots F_n$ ) saling bebas (*independen*) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

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

Untuk  $i \neq j$  sehingga

$$P(F_i|C, F_j) = P(F_i|C)$$

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

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C)P(F_1|C)P(F_2|C)P(F_3|C) \dots \\ &= P(C) \prod_{i=1}^n P(F_i|C) \end{aligned}$$

Persamaan diatas merupakan model dari teorema *Naive Bayes* yang selanjutnya akan digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus *Densitas Gauss* :

$$P(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi} \sigma_{ij}} \exp -\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}$$

Keterangan :

P : Peluang

$X_i$  : Atribut ke i

$x_i$  : Nilai atribut ke i

Y : Kelas yang dicari

$Y_j$  : Sub kelas Y yang dicari

$\mu$  : Mean, menyatakan rata – rata dari seluruh atribut

$\sigma$  : Deviasi standar, menyatakan varian dari seluruh atribut

Menurut Ahmad Basuki (2006) tahap-tahap dari pengambilan keputusan dengan naïve bayes adalah sebagai berikut :

1. Memasukan parameter atau input data yang akan diproses.
2. Mencari nilai atribut hasil terakhir yang akan menjadi hasil diagnosis kelompok sesuai banyaknya yaitu yang beresiko dan yang normal
3. Mencari nilai probabilitas dari setiap setiap atribut berdasarkan pengelompokan hasil akhir
4. Melakukan pengelompokan (dalam hal ini melakukan perkalian) pada masing – masing probabilitas tiap – tiap atribut sesuai dengan jenis hasil diagnosis
5. Dari langkah no 4 kemudian mencari nilai probabilitas terbesar sehingga dihasilkan kesimpulan atau keputusan diagnosis.

### Algoritma Klasifikasi Naïve Bayes

Algoritma Klasifikasi Naïve Bayes dihitung sesuai dengan rumus Naïve Bayes  $P(Y) \prod_{i=1}^q p(X_i|Y)$  yang langkah – langkah perhitungannya dijelaskan sebagai berikut (Meinggian vilian sari,2014):

1. Menghitung nilai probabilitas kelas berdasarkan data nilai  $\rightarrow P(Y)$
2. Menghitung nilai probabilitas tiap fitur berdasarkan data latih  $\rightarrow$

$$\prod_{i=1}^q P(X_i | Y)$$

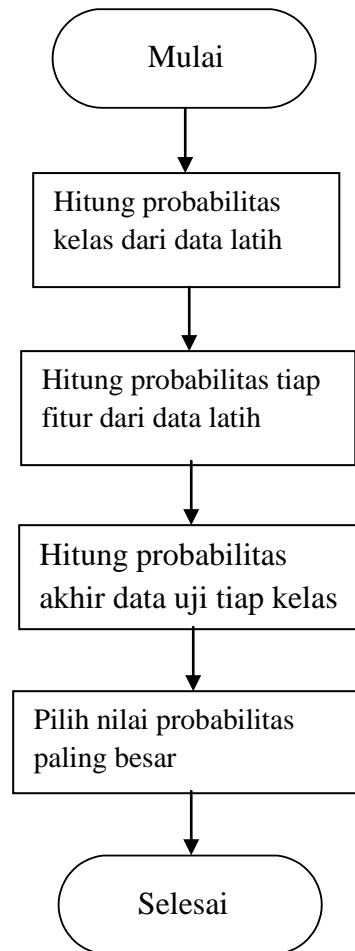
- Untuk fitur bertipe numeric menggunakan rumus berikut

$$P(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp \frac{-(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}$$

Fitur numerik berikut ini dihitung tiap data uji.

3. Menghitung nilai probabilitas akhir
  - Mengalikan hasil dari  $P(Y)$  dan  $\prod_{i=1}^q P(X_i | Y)$  pada masing-masing kelas data uji.

4. Data uji akan diklasifikasikan pada kelas dengan nilai probabilitas akhir terbesar.



Gambar 2.5 Flowchart Naïve Bayes

## 2.5 RISET SEBELUMNYA

Penelitian pertama dilakukan oleh Siregar, Setri Hiyanti dengan judul “Pengaruh Standardisasi Kelayakan Pemberian Kredit Terhadap Pemberian Kredit Usaha Mikro, Kecil, dan Menengah (UMKM) Pada PT. Bank Mandiri - Micro Banking District Center Medan Pulau Pinang” Kredit untuk pengusaha kecil tidak terlepas dari peranan bank-bank pemerintah maupun bank swasta salah satunya adalah Bank Mandiri dalam mengembangkan bisnis pembiayaan UMKM. Bank Mandiri menyederhanakan proses pemberian kredit dengan merancang standardisasi formulir aplikasi khusus bagi usaha mikro dan kecil Penelitian ini bertujuan untuk mengetahui pengaruh standarisasi kelayakan

pemberian kredit Usaha Mikro, Kecil dan Menengah (UMKM) di Jalan Pulau Pinang Medan. Jenis penelitian adalah deskriptif (kuantitatif). Populasi dalam penelitian mengenai pengaruh standarisasi kelayakan pemberian kredit Usaha Mikro, Kecil dan Menengah (UMKM) di Jalan Pulau Pinang Medan. Populasi adalah seluruh debitor di bank Mandiri sebanyak 8900 orang. Besar sampel sebanyak 99 orang. Teknik pengambilan sampel menggunakan metode purposive sampling. Pengumpulan data dengan menggunakan kuisioner. Hasil penelitian diperoleh terdapat pengaruh antara standarisasi kelayakan pemberian kredit (dokumen ( $p = 0,000$ ), agunan ( $p = 0,002$ ), kondisi ( $p = 0,012$  dan aspek keuangan ( $p = 0,009$ ) terhadap keputusan pemberian kredit UMKM. Untuk itu disarankan bagi pihak Bank Mandiri agar memberikan kredit bagi usaha UMKM dengan persyaratan yang lebih ringan, serta menyesuaikan nilai aset yang menjadi agunan dengan nilai kredit yang diberikan.

Penelitian kedua Dilakukan Oleh Hermanto Undjila Dengan Judul “ Penerapan Metode *Profile Matching* Untuk Menentukan Kelayakan Pemberian Kredit Sepeda Motor “ Penentuan kelayakan pemberian kredit sepeda motor sering kali kurang tepat sehingga mengakibatkan resiko kredit macet yang dilakukan oleh nasabah. Untuk itu diperlukan suatu aplikasi sistem pendukung keputusan yang bertujuan menganalisa nasabah yang layak mendapatkan kredit. Adapun manfaat yang diperoleh dengan adanya sistem ini yaitu membantu pihak perusahaan dalam menentukan pemberian kelayakan kredit. *Profil matching* merupakan suatu metode sistem pendukung keputusan yang dapat digunakan dalam sistem ini. Nasabah yang mengajukan kredit akan dinilai berdasarkan kriteria-kriteria yang telah ditentukan oleh perusahaan antara lain : jenis kelamin, bukti kepemilikan rumah, jenis pekerjaan, penghasilan perbulan, pengeluaran perbulan, usia calon debitor dan karakteristik calon debitor. Kemudian nilai profil yang dimiliki oleh seorang nasabah akan dicocokkan dengan nilai profil pencapaian yang telah ditentukan. Hasil dari proses ini berupa prediksi yang dapat dijadikan rekomendasi bagi

pihak pengambil keputusan untuk menentukan nasabah yang layak untuk menerima kredit. Naive bayes mempunyai banyak keunggulan dari profil matching antara lain perhitungannya lebih cepat, akurat, hasilnya langsung bisa dilihat.

Penelitian ketiga dilakukan oleh **Yudi Ardiyanto , Eko Prasetyo , Ilham** Dengan Judul “Sistem Pendukung Keputusan Kelayakan Pengajuan Kredit Sepeda Motor Menggunakan Metode *Naïve Bayes Classifier* ( Study Kasus Di Summit Otto Finance)“ Summit OTTO Finance dalam menentukan kelayakan pengajuan kredit sepeda motor selama ini masih dengan cara manual dan hanya menggunakan keputusan dari hasil survey para surveyor, sehingga tingkat kesesuaian data hasil seleksi masih belum maksimal. Sehingga perlu dilakukannya pembuatan sebuah sistem yang dapat menunjang dalam pengambilan suatu keputusan.

Untuk menyeleksi kelayakan pengajuan kredit sepeda motor digunakan metode *Naïve Bayes Classifier* karena metode ini mendeskripsikan suatu metode pemrosesan berdasarkan klasifikasi tiap kriteria, sehingga metode ini tepat digunakan dalam proses pencarian data. Sistem pendukung keputusan dengan metode *Naïve Bayes Classifier* ini menggunakan tujuh kriteria yang paling menentukan dalam pengambilan keputusan menurut pengambilan kuisioner dari para surveyor antara lain jenis kelamin, bukti kepemilikan rumah, jenis kelamin, penghasilan perbulan, pengeluaran perbulan, usia calon debitur dan karakteristik calon debitur . Dengan hasil dari proses menggunakan metode *Naïve Bayes Classifier* ini berupa kecocokan laporan antara hasil dari data acuan dan perhitungan dari sistem pendukung keputusan kelayakan pengajuan kredit sepeda motor menggunakan metode naïve bayes classifier sebagai rekomendasi untuk menentukan penerimaan kredit sepeda motor. Tingkat kesesuaiannya besar prosentase perhitungan yang dihasilkan sistem secara keseluruhan sebesar 73% bagi pengguna sehingga bisa ditarik kesimpulan bahwasannya sistem yang dibangun telah berjalan dengan baik.