

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Siti Mutiah (2019) melakukan penelitian prediksi kelayakan calon penerima bantuan bidik misi pengganti menggunakan algoritma *Naïve Bayes Classification* di Universitas Gadjah Mada. Dalam penelitian ini dilakukan pemodelan data mining dengan menggunakan algoritma *Naïve Bayes* untuk mendapatkan langkah-langkah sistematis dalam memprediksi kelayakan calon penerima bantuan bidikmisi pengganti. Data yang digunakan merupakan data primer yaitu dataset pelamar bantuan bidikmisi di Universitas Gadjah Mada. Penentuan kelayakan calon penerima bantuan bidikmisi pengganti menggunakan beberapa kriteria antara lain: penghasilan orang tua, jumlah tanggungan orang tua, ipk, dan status penerimaan beasiswa. Dari hasil pengujian model yang digunakan diperoleh nilai rata-rata akurasi yang cukup tinggi yaitu 93,4% dengan laju error 6,6%. Dengan demikian penerapan algoritma *Naïve Bayes* dapat dijadikan alternatif pengambilan keputusan dalam prediksi kelayakan calon penerima bantuan bidikmisi pengganti.

Adrian Satria Putra (2018) melakukan penelitian klasifikasi status gizi balita menggunakan *Naïve Bayes Classification* di Posyandu Ngudi Luruh. Dalam penelitian ini dilakukan pemodelan data mining dengan menggunakan algoritma *Naïve Bayes* untuk mendapatkan langkah-langkah sistematis dalam menentukan status gizi balita di Posyandu Ngudi Luruh, untuk menentukan gizi pada balita

maka akan ada beberapa parameter antara lain: umur, jenis kelamin, berat badan tinggi badan, gaji orang tua. Dari hasil pengujian model yang digunakan diperoleh nilai rata-rata akurasi yang cukup tinggi yaitu 95,1 dengan laju error 4,9%. Dengan demikian penerapan algoritma *Naïve Bayes* dapat dijadikan alternatif untuk menentukan gizi balita di Posyandu Ngudi Luhur.

Ulfah Pauziah (2017) dengan judul analisis penentuan karyawan terbaik menggunakan algoritma *Naïve Bayes* di PT.XYZ. Adanya bentuk dan fungsi aplikasi yang dapat dipakai untuk perusahaan-perusahaan, yang bisa digunakan untuk menghitung layak atau tidaknya seorang karyawan menjadi karyawan terbaik di perusahaan tersebut.

Tabel 2.1 menunjukkan perbandingan penelitian dalam algoritma *Naïve Bayes* yang sudah dilakukan dengan penelitian ini.

Tabel 2.1 Perbandingan Penelitian

Peneliti	Objek	Metode	Hasil
Adrian Satria Putra (2018)	Gizi balita	Algoritma <i>Naïve Bayes</i>	Hasil pengujian model yang digunakan diperoleh nilai rata-rata akurasi yang cukup tinggi yaitu 95,1 dengan laju error 4,9%.
Siti Mutiah(2019)	Mahasiswa pendaftar bantuan bidik misi pengganti.	Algoritma <i>Naïve Bayes</i>	Hasil pengujian model yang digunakan diperoleh nilai rata-rata akurasi yang cukup tinggi yaitu 93,4% dengan laju error 6,6%.

Tabel 2.1 Perbandingan Penelitian

Peneliti	Objek	Metode	Hasil
Ulfah Pauziah (2017)	Karyawan pada PT.XYZ	Algoritma Naïve Bayes	Mempermudah bagi pihak perusahaan dalam menentukan karyawan terbaik.
Zaqie Alfatah (2020)	Karyawan	Algoritma Naïve Bayes	Algoritma Naïve Bayes diharapkan mampu memberikan hasil penilaian kinerja karyawan di CV. SABIRIN SEJAHTERA.

2.2 Dasar Teori

Dasar teori untuk memahami definisi, pengertian dasar dan istilah yang digunakan untuk penelitian ini. Berikut dasar teori yang digunakan.

2.2.1 Kinerja Karyawan

Kinerja seorang karyawan merupakan hal yang bersifat individual, karena setiap karyawan mempunyai tingkat kemampuan yang berbeda-beda dalam mengerjakan tugasnya. Pihak manajemen dapat mengukur karyawan atas unjuk kerjanya berdasarkan kinerja dari masing-masing karyawan. Kinerja adalah sebuah aksi, bukan kejadian. Aksi kinerja itu sendiri terdiri dari banyak komponen dan bukan merupakan hasil yang dapat dilihat pada saat itu juga. Pada dasarnya kinerja merupakan sesuatu hal yang bersifat individual, karena setiap karyawan memiliki tingkat kemampuan yang berbeda dalam mengerjakan tugasnya. Kinerja tergantung pada kombinasi antara kemampuan, usaha, dan kesempatan yang

diperoleh. Hal ini berarti bahwa kinerja merupakan hasil kerja karyawan dalam bekerja untuk periode (jurnal-sdm.blogspot.com)

Menurut Harsuko (2011), kinerja adalah sejauh mana seseorang telah memainkan baginya dalam melaksanakan strategi organisasi, baik dalam mencapai sasaran khusus yang berhubungan dengan peran perorangan dan atau dengan memperlihatkan kompetensi yang dinyatakan relevan bagi organisasi. Kinerja adalah suatu konsep yang *multidimensional* mencakup tiga aspek yaitu sikap (*attitude*), kemampuan (*ability*) dan prestasi (*accomplishment*).

2.2.2 CV.SABIRIN SEJAHTERA

CV. SABIRIN SEJAHTERA suatu perusahaan yang bergerak di bidang cuci mobil dan motor serta memelihara kondisi mobil dan motor agar tetap bagus, selama ini penilaian kinerja karyawan menggunakan keputusan pimpinan langsung. Pimpinan pun sulit dalam menilai kinerja masing-masing karyawan karena, kurang jelasnya kriteria penilaian karyawan. Penilaian karyawan hanya dilakukan sebagai referensi pimpinan saja, sehingga karyawan kurang termotivasi dalam menunjukkan kinerja terbaik mereka.

Pada perusahaan CV. SABIRIN SEJAHTERA mempunyai 7 orang karyawan dalam sehari, dalam sehari akan dibagi menjadi dua shift, shift pagi dan shift siang, pada shift pagi dimulai dari jam 07.00-12.00 WIB dan akan dilanjutkan dari jam 15.00-17.00 WIB, sedangkan shift yang kedua atau bisa dibilang shift siang dimulai dari jam 12.00 – 15.00 WIB dan akan dilanjutkan dari jam 17.00-21.00 WIB.

Penilaian kinerja karyawan yang dilakukan CV. SABIRIN SEJAHTERA akan memberikan umpan balik baik berupa teguran, kenaikan kompensasi, tanggung jawab, pelatihan, ataupun bentuk umpan balik lainnya kepada karyawan berkaitan dengan kinerja karyawan yang telah dicapai, dan juga akan memotivasi para karyawan untuk menjadi lebih baik tiap bulannya.

2.2.3 Data Mining

Data mining merupakan proses penggunaan teknik statistik, kecerdasan buatan atau *artificial intelligence* (AI), pengenalan pola, dan sistem basis data yang digunakan untuk memanfaatkan data dalam basis data yang diolah sehingga menghasilkan informasi baru yang berguna. (Prasetyo, 2014)

Data mining memindai basis data untuk menemukan pola yang tersembunyi, informasi prediktif yang mungkin dilewatkan oleh para ahli karena berada di luar dugaan. *Data mining* terdiri dari urutan iteratif dari langkah-langkah berikut: *data clean*, *data integration*, *data selection*, *data transformation*, *data mining*, *pattern evaluation*, dan *knowledge extraction*. (Kaur dan Aggarwal, 2010).

Istilah *data mining* dan *knowledge discovery in database* (KDD) sering digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Kedua istilah tersebut mempunyai konsep yang berbeda, tetapi saling berkaitan dan digunakan secara bergantian karena *data mining* merupakan salah satu tahapan dalam keseluruhan proses KDD (Kusrini dan Luthfi). Berikut adalah tahapan dalam proses KDD.

1. Pembersihan data

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.

2. Seleksi data

Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis atau nilai yang akan masuk pada komputasi yang akan diambil dari database.

3. Transformasi (*Transformation*)

Pengkodean adalah proses transformasi pada data yang telah dipilih. Proses pengkodean dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data dengan menggunakan teknik atau metode tertentu. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. Interpretasi/Evaluasi (*Interpretation/Evaluation*)

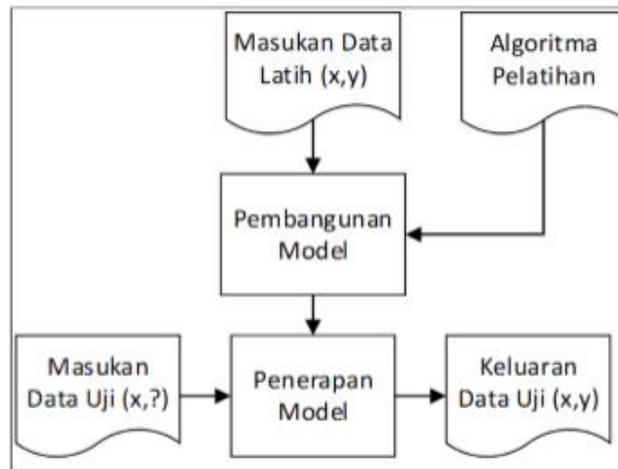
Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap-tahap ini disebut dengan interpretasi yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

2.2.4 Klasifikasi

Klasifikasi adalah proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui

Classification (Hand dan Kamber, 2006:285) adalah suatu model dalam data mining dimana, *classifier* dikonstruksi untuk memprediksi *categorical* label, seperti “aman” atau “beresiko” untuk data aplikasi peminjaman uang; ”ya” atau ”tidak” untuk data marketing; atau “treatment A”, ”treatment B” atau “treatment C” untuk data medis. Kategori tersebut dapat direpresentasikan dengan nilai yang sesuai dengan kebutuhannya, dimana pengaturan dari nilai tersebut tidak memiliki arti tertentu.

Kerangka kerja klasifikasi ditunjukkan pada Gambar 2.1. Pada gambar tersebut, disediakan sejumlah data latih (x,y) untuk digunakan sebagai data pembangun model, kemudian menggunakan model tersebut untuk memprediksi kelas dari data uji $(x,?)$ sehingga data uji $(x,?)$ diketahui kelas y yang seharusnya (Prasetyo, 2014).



Gambar 2.1 Proses pekerjaan klasifikasi

2.2.5 Algoritma Naïve Bayes

Algoritma Naïve Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi, Naïve Bayes merupakan teori pendekatan statistika yang fundamental dalam data mining. Pendekatan ini didasarkan pada kuantitatif *trade-off* antara berbagai keputusan klasifikasi dengan menggunakan probabilitas (Santoso,2007).

Klasifikasi Naïve Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya (Bustami, 2014)

Persamaan teorema Bayes adalah :

$$P(C|X) = \frac{P(X|C).P(C)}{P(X)} \dots\dots\dots(2.1)$$

Keterangan :

X : Data dengan class yang belum diketahui

C : Hipotesis data X merupakan suatu class spesifik

$P(C|X)$: Probabilitas hipotesis C berdasarkan kondisi X (posteriori probability)

$P(C)$: Probabilitas Hipotesis C (prior probability)

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

$P(X)$: 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.