



TEKNOLOGI INFORMASI

UNIVERSITAS DARMA PERSADA

BAB II

LANDASAN TEORI

2.1 Gizi

Faktor yang mempengaruhi pertumbuhan, perkembangan dan kelangsungan hidup manusia adalah Gizi. Gizi sangat penting untuk kesehatan dan perkembangan manusia. Gizi yang baik sangat erat kaitannya dengan sistem kekebalan yang lebih kuat, risiko penyakit menular yang lebih rendah, dan umur panjang. Salah satu penentu faktor yang mempengaruhi kesehatan seseorang adalah gizi (Devi, 2010). Ilmu Gizi mempelajari proses yang terjadi sejak makanan masuk dalam tubuh, dicerna, diserap dan digunakan oleh tubuh berupa zat-zat yang bermanfaat bagi kesehatan (Forum Komunitas Pengembangan Ilmu Gizi, 1998).

Menurut (Indriani, 2015) dalam buku Gizi dan Pangan berpendapat tujuan utama dari ilmu gizi yaitu dengan mengkonsumsi makanan dapat mencapai, memperbaiki dan mempertahankan kesehatan tubuh. Maka dari itu penting untuk setiap lapisan masyarakat mempelajari dan mengetahui ilmu gizi.

Dengan makanan tubuh dapat memelihara/mempertahankan hidup, bertumbuh dan memperbaiki jaringan tubuh yang didapatkan dari kandungan bahan nutrisi. Ada 2 jenis nutrisi yang terdiri dari zat gizi makro dan zat gizi mikro. Karbohidrat, protein dan lemak termasuk dalam zat gizi makro. Vitamin dan mineral termasuk dalam zat gizi mikro. Untuk memenuhi kebutuhan zat gizi tersebut tidak dapat hanya dipenuhi oleh satu atau dua sumber makanan saja dan harus terdistribusi dalam jumlah tertentu.

2.1.1 Status Gizi Pada Bayi Usia 6-11 Bulan

Asupan gizi pada bayi perlu diperhatikan karena jika bayi kekurangan gizi akan menimbulkan gangguan pertumbuhan dan perkembangan mental, menurunnya kecerdasan, bahkan dapat menjadi penyebab kematian. Berdasarkan data yang dijabarkan oleh (Al Rahmad, 2016) hampir seperlima balita mengalami berat badan kurang (wasting), satu dari tiga bayi mengalami gangguan pertumbuhan (stunting) dan empat dari seratus bayi di Indonesia tidak bisa bertahan hidup lebih dari lima tahun. (United Nations Indonesia, 2020) memberikan ringkasan eksekutif dari Riset Kesehatan Dasar tahun 2018, bahwa di Indonesia dua juta balita mengalami kurang gizi, tujuh juta anak pertumbuhan yang kurang (stunting) dan dua juta kelebihan berat badan (overweight).

Berdasarkan (Peraturan Menteri Kesehatan Republik Indonesia Nomor 28 Tahun 2019 : Tentang Angka Kecukupan Gizi Yang Dianjurkan Untuk Masyarakat Indonesia, 2019) tabel kecukupan gizi yang dianjurkan Kemenkes RI untuk bayi 6-11 bulan dapat dideskripsikan berturut-turut pada tabel berikut :

Tabel 2.1 Rekomendasi atribut gizi (per orang per hari)

Atribut	Nilai
Berat Badan (kg)	9
Tinggi Badan (cm)	72
Energi (kkal)	800
Protein (g)	15
Lemak (g) Total	35
Omega 3	0,5
Omega 6	4,4
Karbohidrat (g)	105
Serat (g)	11
Air (ml)	900

Tabel 2.2 Rekomendasi kecukupan vitamin (per orang per hari)

Atribut	Nilai
Vit A (RE)	400
Vit D (mcg)	10
Vit E (mcg)	5
Vit K (mcg)	10
Vit B1 (mg)	0,3
Vit B2 (mg)	0,4
Vit B3 (mg)	4

Atribut	Nilai
Vit B5 (Pantotenat) (mg)	1,8
Vit B6 (mg)	0,3
Folat (mcg)	80
Vit B12 (mcg)	1,5
Biotin (mcg)	6
Kolin (mg)	150
Vit C (mg)	50

Tabel 2.3 Rekomendasi kecukupan mineral (per orang per hari)

Atribut	Nilai
Kalsium (mg)	270
Fosfor (mg)	275
Magnesium (mg)	55
Besi2 (mg)	11
Iodium (mcg)	120
Seng3 (mg)	3
Selenium (mcg)	10

Atribut	Nilai
Mangan (mg)	0,7
Fluor (mg)	0,5
Kromium (mcg)	6
Kalium (mg)	700
Natrium (mg)	370
Klor (mg)	570
Tembaga (mcg)	220

2.1.2 Makanan Pendamping ASI (MPASI)

Menurut (Saraswati, 2017) sebagai penunjang kebutuhan bayi dan balita. MPASI merupakan salah satu kebutuhan penting. MPASI adalah makanan pendamping ASI yang diberikan oleh ibu kepada bayi atau balita. Anak berusia 6 bulan sampai 2 tahun harus diberikan MPASI. Adapun alasannya yaitu (Saraswati, 2017) :

1. ASI sudah tidak bisa memenuhi sepenuhnya kebutuhan nutrisi dan nafsu makan
2. Cadangan nutrisi dalam ASI telah habis dan tidak dapat dipenuhi oleh ASI (seperti zat besi) saja

3. Saat memperkenalkan makanan baru, rasa ingin tahu tentang berbagai tekstur dan rasa lingkungan membuat tahap eksplorasi ini sangat bermanfaat
4. Masih lemahnya imunitas bayi yang usia kurang dari 6 bulan

2.2 Data Mining

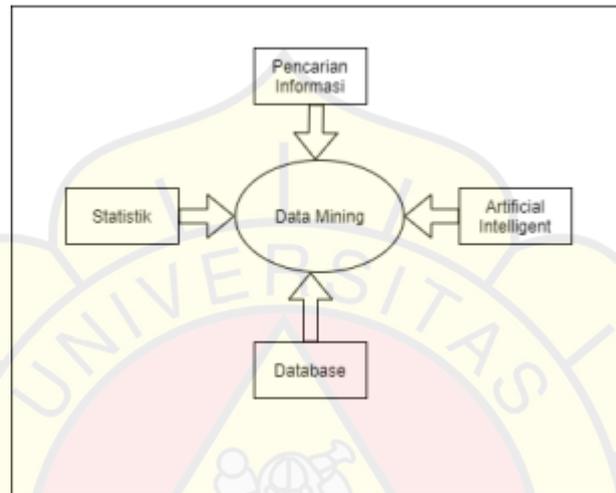
Menurut (Efori Buulolo, 2020 : 5) dalam e-book yang berjudul “Data Mining Untuk Perguruan Tinggi”. Data Mining atau kadang disebut juga *Knowledge Discovery in Database* (KDD) merupakan aktivitas yang berkaitan dengan pengumpulan data, pemakaian data historis untuk menemukan pengetahuan, informasi, keteraturan, pola atau hubungan dalam data yang berukuran besar. Output dalam data mining dapat dipergunakan sebagai alternatif dalam pengambilan keputusan atau untuk memperbaiki keputusan di masa yang akan datang.

Data mining salah satu bidang ilmu yang perkembangan sangat pesat, perkembangan tersebut disebabkan oleh beberapa faktor, antara lain :

1. Semakin tingginya kesadaran akan pentingnya data
2. Semakin tingginya pemanfaat output dari hasil pengolahan data dalam berbagai bidang contohnya bidang bisnis.
3. Perkembangan kumpulan data yang begitu cepat.
4. Peningkatan akses internet baik melalui navigasi web ataupun melalui smartphone.
5. Perkembangan hardware dan software khususnya yang berhubungan dengan data mining.
6. Perkembangan yang begitu cepat dalam bidang komputasi komputer.

7. Media penyimpanan yang semakin besar dengan harga yang semakin terjangkau.

Data mining bukanlah bidang ilmu yang berdiri sendiri, tetapi sangat berkaitan dengan bidang ilmu lain seperti database, statistic, pencarian informasi, dan artificial intelligent.

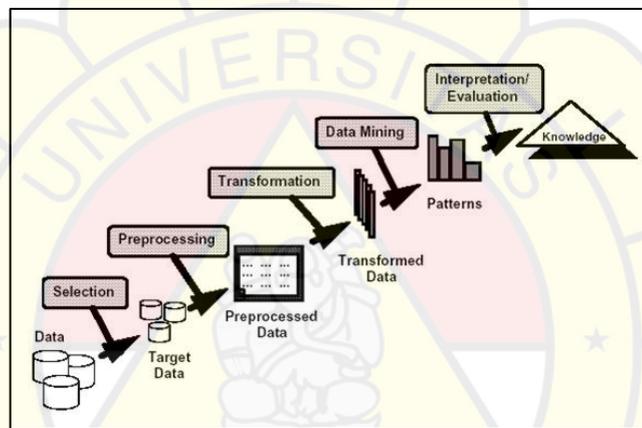


Gambar 2.1 Bidang Ilmu Data Mining

1. Database-Data Mining : Kumpulan data yang digunakan dalam data mining bersumber salah satunya dari database. Data yang digali/dicari informasinya dipisahkan dari data operasional yang di database.
2. Statistic-Data Mining : Dalam pengambilan keputusan, statistic membutuhkan data mulai dari pengumpulan data, pengambilan sampel data dan probabilitas. Data mining dalam penentuan sampel data, menganalisa, dan mempresentasikan output menggunakan teknik statistic.
3. Pencarian Informasi-Data Mining : Pencarian informasi merupakan salah satu kegiatan dalam proses data mining yang meliputi interpretasi, analisis dan penyimpanan data.

4. Artificial Intelligent-Data Mining : Salah satu cabang ilmu dari Artificial Intelligent adalah machine learning. Machine learning merupakan disiplin ilmu yang penting dalam data mining dimana system computer belajar dari training data yang digunakan.

Data Mining mempunyai beberapa model proses yang digunakan untuk mengarahkan pelaksanaan data mining, model proses yang biasa digunakan adalah *Knowledge Discovery Databases (KDD)*, *Cross Industry Standard for Data Mining (CRISP-DM)*.



Gambar 2.2 Tahapan KDD

1. Berikut adalah penjelasan dari setiap tahapan pada *Knowledge Discovery Databases (KDD)* :

- a. Domain Understanding and KDD Goals

Tujuan ditentukan dari sudut pandang user dan digunakan untuk mengembangkan dan pemahaman tentang domain aplikasi dan pengetahuan sebelumnya.

- b. Selection and Additions

Tahap kedua berfokus pada penentuan data target dan subset dari data sampel atau variabel.

c. Preprocessing : Data Cleaning etc

Pembersihan dan preprocessing data merupakan operasi dasar untuk menyelesaikan data yang konsisten tanpa noisy.

d. Transformation

Transformasi data dari satu bentuk ke bentuk lainnya sehingga data diimplementasikan dengan mudah.

e. Data Mining (*Chosing The Suitable Data Mining Task*)

Memilih metode data mining yang sesuai berdasarkan tujuan tertentu yang telah didefinisikan pada tahap pertama, contoh dari metode data mining adalah classification, regression, clustering dan summarization.

f. Data Mining (*Chosing The Suitable Data Mining Algorithm*)

Memilih algoritma yang tepat untuk pencarian pola-pola data, algoritma yang dipilih berdasarkan kecocokan kriteria dengan metode data mining.

g. Data Mining (*Implement Data Mining Algorithm*)

Pada tahap ini algoritma yang telah dipilih diimplementasikan

h. Evaluation and Interpretation

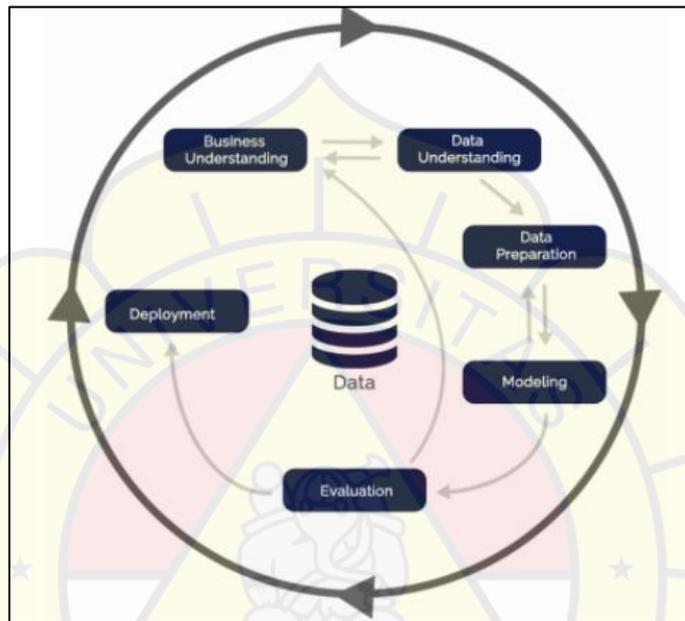
Tahap ini berfokus pada interpretasi dan evaluasi yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan hipotesa yang ada sebelumnya.

i. Discovered Knowledge

Penggunaan pengetahuan yang ditemukan dari proses KDD, dimana memutuskan apa yang akan dilakukan dengan pengetahuan dihasilkan.

2. Tahapan *Cross Industry Standard for Data Mining* (CRISP-DM)

Cross Industry Standard for Data Mining (CRISP-DM) merupakan metode yang menggunakan model proses pengembangan data yang banyak digunakan para ahli untuk memecahkan masalah. Proses penelitian ini mengacu pada enam tahap CRISP-DM yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi dan penyebaran.



Gambar 2.3 Tahapan CRISP-DM

Berikut adalah penjelasan dari setiap tahapan pada *Cross Industry Standard for Data Mining (CRISP-DM)* :

a. Business Understanding

Business understanding adalah tahapan awal dari metodologi CRISP-DM. Tahapan business understanding berisi tentang menentukan tujuan bisnis, menilai situasi saat ini dan menetapkan tujuan dilakukan data mining. Tahapan ini sangat penting, namun sering diabaikan ketika seseorang terjun ke dunia data mining.

b. Data Understanding

Data understanding adalah kegiatan persiapan, mengevaluasi persyaratan data, dan termasuk pengumpulan data. Pada tahapan ini, data yang berhasil dikumpulkan kemudian dideskripsikan (bagian mana yang atribut, kelas, dan tipe data).

c. Data Preparation

Setelah data dikumpulkan, data-data tersebut perlu diidentifikasi, dipilih, dibersihkan, kemudian dibangun ke dalam bentuk/format yang diinginkan. Data preparation, disebut juga dengan data pre-processing.

d. Modelling

Model adalah aplikasi dari algoritma untuk mencari, mengidentifikasi, dan menampilkan pola. Pemilihan algoritma berdasarkan tipe data karena dari tipe data kita bisa mengetahui apakah data tersebut akan diestimasi, prediksi, klasifikasi, clustering atau melihat hubungan asosiatif.

e. Evaluation

Evaluation digunakan untuk membantu pengukuran evaluasi pada model. Kita bisa mengukur model mana yang paling baik digunakan untuk proses data mining. Pada penerapan klasifikasi, pengukuran evaluasi yang banyak digunakan adalah akurasi, sensitivity, specificity, g-mean, f-measure, dan lain sebagainya.

f. Deployment

Deployment adalah tahapan akhir dalam CRISP-DM. Setelah model dievaluasi dan dipilih algoritma dengan hasil pengukuran terbaik, dilanjutkan ke tahapan deployment. Tahapan deployment digunakan

untuk melakukan otomatisasi model atau pengembangan aplikasi, terintegrasi dengan sistem informasi manajemen atau operasional yang ada.

2.2.1 Kegunaan Data Mining

Secara garis besar kegunaan data mining dibagi menjadi dua yaitu : deskriptif dan prediktif. Secara deskriptif artinya data mining berarti menemukan pola-pola yang digunakan untuk menjelaskan karakteristik data. Sedangkan secara prediktif data mining berarti dapat digunakan untuk menemukan model-model pengetahuan yang digunakan untuk melakukan prediksi.

Menurut (Suyanto, 2017) Data mining berdasarkan fungsionalitasnya dapat dikelompokkan menjadi enam bagian yaitu :

1. Klasifikasi (*classification*)

Diterapkan pada data baru untuk mengelompokkan jenis objek. Klasifikasi termasuk pada model supervised. Pada persoalan klasifikasi kita memiliki sampel data dan memprediksi beberapa class yang ada berdasarkan sampel yang ada. Hanya satu atribut di antara banyak atribut yang disebut dengan atribut predator. Klasifikasi ini juga umum, digunakan untuk pemodelan bisnis dan lainnya. Misal klasifikasi menentukan penyakit tertentu atau menentukan customer berdasarkan model pembayarannya.

2. Klastering (*Clustering*)

Berbeda dengan klasifikasi, klastering termasuk model unsupervised. Klastering mengelompokkan data yang tidak diketahui labelnya. Klastering yang diorganisasi ke dalam struktur hierarkikal akan mendefinisikan

taksonomi dari data. Penerapan metode klustering yang tepat akan menghasilkan clustering yang berkualitas. Suatu cluster dikarakterisasi oleh centroid, atribut histogram dan clustering model hierarchial tree.

3. Regresi (*Regression*)

Merupakan suatu fungsi yang digunakan untuk memodelkan data untuk meminimalkan hasil kesalahan prediksi. Umumnya regresi dilakukan dengan data yang bersifat time series.

4. Association Rule

Merupakan pemodelan kebergantungan. Fungsi asosiasi ini biasanya kita kenal dengan istilah “market basket analysis” yang merupakan fungsi untuk menemukan relasi atau korelasi antara himpunan item-item. Aturan asosiasi diartikan pada basket data yang digunakan untuk keperluan promosi, desain katalog untuk meningkatkan penjualan. Contoh penerapan asosiasi adalah ketika customer membeli pamper maka ada kemungkinan membeli bir.

5. Anomaly Detection

Mengidentifikasi data yang tidak umum. Bisa berupa outlier, perubahan deviasi/bias yang penting dan perlu investigasi lebih lanjut.

6. Summarization

Menyediakan representasi data yang lebih sederhana meliputi pelaporan, visualisasi data yang dipergunakan untuk menunjang informasi dan penguatan keputusan.

2.2.2 Teknik Data Mining

Menurut (Vulandari, 2017) teknik yang digunakan dalam data mining erat kaitannya dengan “penemuan” (*discovery*) dan “pembelajaran” (*learning*) yang terbagi dalam tiga metode utama pembelajaran yaitu :

1. Supervised Learning

Supervised Learning adalah teknik yang paling sering digunakan. Teknik ini sama dengan “programming by example”. Teknik ini melibatkan fase pelatihan dimana data pelatihan historis yang karakter-karakternya dipetakan ke hasil yang telah diketahui diolah dalam algoritma data mining. Proses ini melatih algoritma untuk mengenali variabel-variabel dan nilai-nilai kunci yang nantinya akan digunakan sebagai dasar dalam membuat perkiraan-perkiraan ketika diberikan data baru.

2. Unsupervised Learning

Teknik pembelajaran ini tidak melibatkan fase pelatihan seperti yang terdapat pada supervised learning. Teknik ini bergantung pada penggunaan algoritma yang mendeteksi semua pola seperti associations yang muncul dari kriteria penting yang spesifik dari data masukan. Pendekatan ini mengarah pada pembuatan banyak aturan (rules) yang mengkarakterisasikan penemuan seperti associations, clusters dan segments. Aturan-aturan ini kemudian dianalisis untuk menemukan hal-hal penting.

3. Reinforcement Learning

Teknik pembelajaran ini jarang digunakan dibandingkan dengan dua teknik lainnya, namun memiliki penerapan-penerapan yang terus dioptimalkan

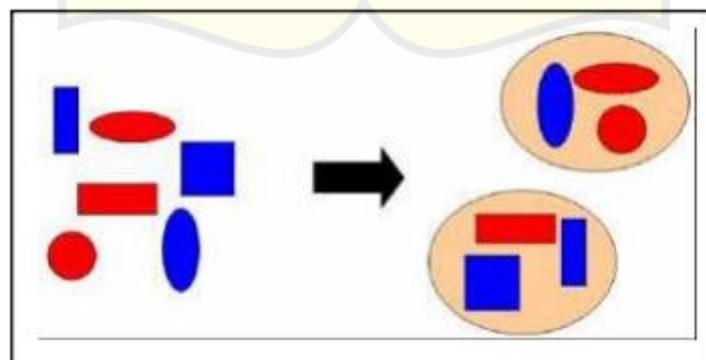
dari waktu ke waktu dan memiliki kontrol adaptif. Reinforcement learning sangat tepat digunakan untuk menyelesaikan masalah-masalah yang sulit serta bergantung pada waktu.

2.2.3 Teknik Clustering

Clustering adalah salah satu model yang dilakukan dengan cara melakukan proses segmentasi terhadap populasi yang heterogen ke dalam sejumlah cluster yang homogen. Proses clustering ini berbeda dengan klasifikasi dimana pada clustering adalah tidak diketahui waktu dimana algoritma dimulai. (Deny, William, Muhammad, 2020 : 53).

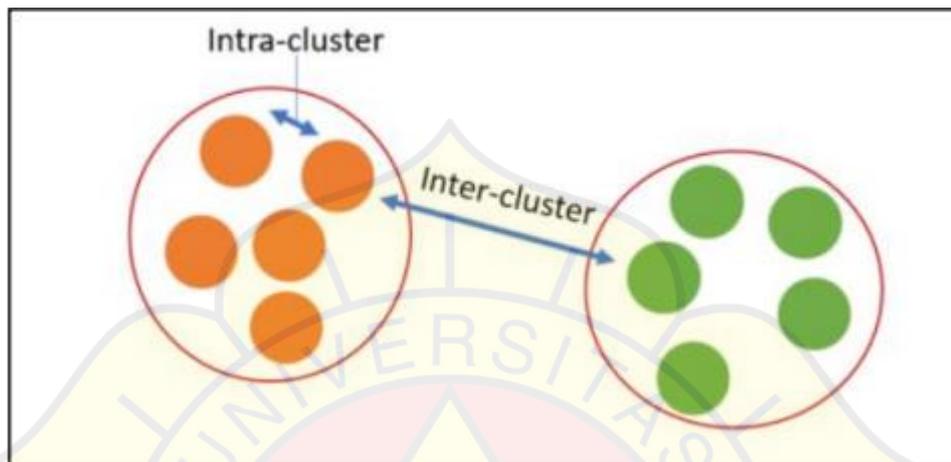
Clustering dilakukan dengan cara mengumpulkan data yang serumpun dari sebuah data set yang lebih besar. Teknik ini menyingkapkan sejumlah kelompok yang digunakan sebagai masukan datanya. Dengan clustering, kelompok minoritas yang tersebar dikelompokkan dalam sebuah kelompok besar yang memiliki kemiripan entitas.

Clustering dapat juga di gunakan untuk mendeteksi secara otomatis cluster dari record-record yang berdekatan yang memiliki pengertian tertentu di dalam keseluruhan variable-variabel.



Gambar 2.4 Penggambaran Teknik Clustering

Pada gambar 2.4 Mengilustrasikan data yang dikelompokkan berdasarkan kriteria bentuknya. Cluster yang baik ketika jarak antara objek data yang ada pada cluster yang sama (intra-cluster) dapat diminimalkan dan jarak antara objek yang ada pada cluster yang berbeda (inter-cluster) dapat dimaksimalkan.



Gambar 2.5 Perbedaan Intra-Cluster Dengan Inter-Cluster

2.2.4 Teknik Klasifikasi

Klasifikasi adalah proses pengumpulan data bersama-sama yang didasarkan atas sekumpulan kesamaan yang awalnya telah ditentukan oleh seorang analis sebelum analis dimulai. Teknik ini memeriksa data yang telah diklasifikasikan dan dikumpulkan dalam group bersama-sama sesuai dengan keanggotaannya. Aturan keanggotaan bisa mempunyai komponen waktu, komponen geografis, dan komponen kualitatif. (Deny, William, Muhammad, 2020 : 52).

Klasifikasi adalah salah satu model yang paling sering digunakan sebagai model pembelajaran di dalam data mining. Tujuannya adalah untuk membangun sebuah model untuk memprediksi perilaku konsumen di masa mendatang melalui klasifikasi record di dalam database ke dalam beberapa kelompok kelas berdasarkan kriteria tertentu.

2.3 Metode Pengembangan Sistem

2.3.1 Model Agile

Dalam pengembangan aplikasi Perancangan Sistem Informasi Aplikasi Gizi Balita Berbasis Web pada Puskesmas Cibaregbeg Cianjur ini penulis menggunakan metodologi Agile. Metodologi Agile adalah sekelompok metodologi pengembangan perangkat lunak yang didasarkan pada prinsip-prinsip yang sama atau pengembangan sistem jangka pendek yang memerlukan adaptasi cepat dari pengembang terhadap perubahan dalam bentuk apapun.



Gambar 2.6 Metodologi agile

Agile development merupakan salah satu dari metode pengembangan perangkat lunak yang digunakan dalam pengembangan perangkat lunak. Adapun penjelasan urutan dari tahapan-tahapan yang dimiliki metodologi Agile adalah sebagai berikut:

1. *Requirement*

Dalam langkah ini merupakan analisa terhadap kebutuhan sistem. Pengumpulan data dalam tahap ini bisa melakukan sebuah penelitian, wawancara atau study literatur. Seorang system analisis akan menggali

informasi sebanyak-banyaknya dari user sehingga akan tercipta sebuah sistem komputer yang bisa melakukan tugas-tugas yang diinginkan oleh user tersebut. Tahapan ini akan menghasilkan dokumen user requirement atau bisa dikatakan sebagai data yang berhubungan dengan keinginan user dalam pembuatan sistem. Dokumen inilah yang akan menjadi acuan system analisis untuk menterjemahkan kedalam bahasa pemrograman.

2. *Design*

Proses *design* akan menterjemahkan syarat kebutuhan sebuah perancangan perangkat lunak yang dapat diperkirakan sebelum dibuat koding. Proses ini berfokus pada struktur data, arsitektur perangkat lunak, representasi *interface*, dan *detail* (algoritma) prosedural. Tahapan ini akan menghasilkan dokumen yang disebut *software requirement*. Dokumen inilah yang akan digunakan programmer untuk melakukan aktivitas pembuatan sistemnya.

3. *Development*

Tahapan inilah yang merupakan tahapan secara nyata dalam mengerjakan suatu sistem. Dalam artian penggunaan computer akan dimaksimalkan dalam tahapan ini. Setelah pengkodean selesai maka akan dilakukan testing terhadap system yang telah dibuat tadi.

4. *Testing*

Disini Perangkat Lunak yang telah di buat di tes oleh bagian control kualitas agar Bug yang ditemukan bisa segera diperbaiki dan kualitas perangkat lunak terjaga.

5. *Deployment*

yaitu proses yang dilakukan oleh penjamin kualitas untuk menguji kualitas sistem. Setelah sistem memenuhi syarat maka perangkat lunak siap dideployment.

6. *Review*

Tidak ada perangkat lunak yang 100% bebas dari bug, oleh karena itu sangatlah penting agar perangkat lunak dipelihara secara berkala.

2.4 **Algoritma Sistem**

2.4.1 **Algoritma Naïve Bayes Classification (NBC)**

Algoritma Naive Bayes Classification adalah sebuah metode klasifikasi yang didasarkan pada probabilitas dan statistic yang dikemukakan oleh ilmuwan inggris Thomas Bayes. Algoritma Naïve Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naïve Bayes Classifier ini adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi/kejadian. (Deny, William, Muhammad, 2020 : 105).

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

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class spesifik

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

P(H) : Probabilitas hipotesis H (prior probabilitas)

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

P(X) : Probabilitas X

2.4.2 Algoritma K-Means Clustering

K-Means merupakan salah satu algoritma dalam teknik clustering non-hierarki yang berusaha memartisi data yang ada ke dalam bentuk satu atau lebih cluster. Algoritma ini memartisi data ke dalam cluster sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam cluster yang lain. (Deny, William, Muhammad, 2020 : 119).

2.4.3 Tahapan Metode K-Means

Di dalam metode K-Means memiliki beberapa tahapan yang harus dilakukan saat melakukan perhitungan untuk mencari data cluster hingga iterasi menjadi sama dengan hitungan yang sebelumnya, berikut adalah tahapan K-Means :

- a. Menentukan k sebagai jumlah cluster (C1, C2 dan C3) yang ingin dibentuk.
- b. Tentukan nilai untuk pusat cluster (centroid).
- c. Hitung jarak setiap data input terhadap masing-masing centroid menggunakan rumus Euclidean Distance hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Dibawah ini adalah persamaan

Euclidean Distance :

$$d(x||y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan :

$d(x,y)$: Jarak antara data pada posisi titik x dan y

x : Posisi titik data pertama (pusat klaster) y : posisi titik data kedua (data dari N)

n : Jumlah atribut data

Jarak terpendek antara centroid dengan object menentukan titik klaster antar object.

Adapun rumus iterasi lainnya didefinisikan sebagai berikut :

$$v_{ij} = \frac{1}{N_k} X_{ij}$$

Keterangan :

v_{ij} : Centroid, cluster pertama pada variabel ke-j

N_k : Jumlah data yang menjadi anggota cluster ke-i i,k = Indeks dari cluster

J : Indeks dari variabel

X_{ij} : Nilai data ke-k yang ada di dalam cluster ini untuk variable ke-j

- d. Kemudian lakukan iterasi selama partisi data masih bergerak (tidak ada lagi objek yang bergerak ke partisi lain), bila masih maka ke poin 3.
- e. Bila grup data sekarang sama dengan grup data sebelumnya, maka hentikan iterasi. Data telah dipartisi sesuai nilai centroid akhir.