

## BAB II

### LANDASAN TEORI

#### 2.1 Sistem pendukung Keputusan

Menurut Raymond McLeod, sistem pendukung keputusan merupakan sebuah sistem yang menyediakan kemampuan untuk penyelesaian masalah dan komunikasi untuk permasalahan yang bersifat semi-terstruktur. Menurut Moore and Chang, SPK dapat digambarkan sebagai sistem yang berkemampuan mendukung analisis ad hoc data, dan pemodelan keputusan, berorientasi keputusan, orientasi perencanaan masa depan, dan digunakan pada saat-saat yang tidak biasa. Jadi sistem pendukung keputusan adalah sistem yang dipakai untuk mendukung pengambilan keputusan dalam menyelesaikan suatu masalah agar masalah yang ada dapat diselesaikan dengan baik. SPK adalah sistem interaktif berbasis computer yang membantu untuk mengambil keputusan memanfaatkan data dan model untuk memecahkan masalah yang tidak terstruktur. (Sinaga M.T & Melati, 2012)

##### a. Konsep SPK

SPK adalah sistem yang dapat dikembangkan, mampu mendukung analisis data dan pemodelan keputusan, berorientasi pada perencanaan serta tidak bisa direncanakan interval (periode) waktu pemakaiannya. SPK terdiri dari 3 komponen yang berinteraksi satu dengan yang lainnya, yaitu :

1. *Language System*, adalah suatu mekanisme untuk menjembatani pemakai dan komponen lainnya.
2. *Knowledge System*, adalah repository pengetahuan yang berhubungan dengan masalah tertentu baik berupa data maupun prosedur.

3. *Problem Processing System*, adalah sebagai penghubung kedua komponen lainnya, berisi satu atau beberapa kemampuan manipulasi atau menyediakan masalah secara umum yang diperlukan dalam pengambilan keputusan.

b. Karakteristik SPK

Karakteristik dari Sistem Pendukung Keputusan yang membedakan dari sistem lainnya adalah:

1. SPK dirancang untuk membantu pengambilan keputusan dalam memecahkan masalah yang sifatnya semi terstruktur maupun tidak terstruktur.
2. Dalam proses pengolahannya, SPK mengkombinasikan penggunaan model-model analisis dengan teknik pemasukan data konvensional serta fungsi-fungsi pencari.
3. SPK dirancang sedemikian rupa sehingga dapat digunakan dengan mudah oleh orang-orang yang tidak memiliki dasar kemampuan pengoperasian komputer yang tinggi.
4. SPK dirancang dengan menekankan pada aspek fleksibilitas serta kemampuan adaptasi yang tinggi. Sehingga mudah disesuaikan dengan berbagai perubahan lingkungan yang terjadi dan kebutuhan pemakai.

c. Keuntungan Sistem Pendukung Keputusan

Keuntungan dari Sistem Pendukung Keputusan antara lain:

1. Memperluas kemampuan pengambilan keputusan dalam memproses data/informasi bagi pemakainya.
2. Membantu pengambilan keputusan dalam hal penghematan waktu yang dibutuhkan untuk memecahkan masalah terutama berbagai masalah yang sangat kompleks dan tidak terstruktur.
3. Dapat menghasilkan solusi dengan lebih cepat serta hasilnya dapat diandalkan.

## **2.2 Data Mining**

Data mining adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Teknik dalam Data Mining yaitu bagaimana menelusuri data yang ada untuk membangun sebuah model. Model tersebut digunakan untuk mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan.

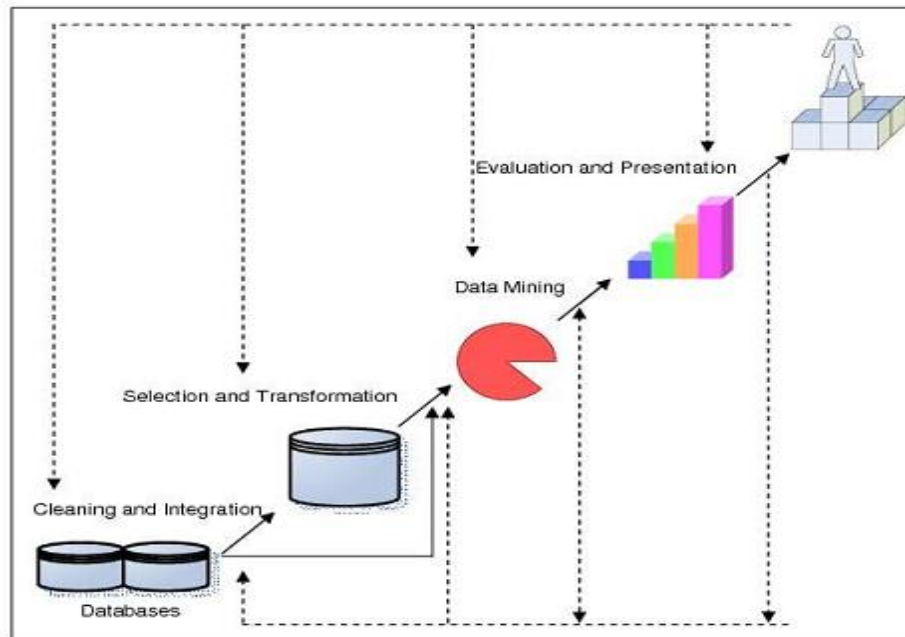
Penelitian terdahulu telah menggunakan data jenis item alat-alat kesehatan sebanyak 30 data untuk melakukan prediksi persediaan barang dengan pengujian pertama yaitu menghitung frekuensi itemset dengan parameter adalah minimal support sama dengan 16% dan maksimal support sama dengan 100%, support adalah nilai presentasi banyak jumlah jenis item dan jumlah presentasi banyak item dan pola kombinasi dua item didalam transaksi menggunakan Metode NBC dengan data sampel sebanyak 60 data, setelah dilakukan pengujian, terdapat 47 data penggunaan listrik rumah tangga berhasil di klasifikasi dengan benar dan sebanyak 13 data penggunaan listrik rumah tangga tidak berhasil diklasifikasi dengan benar, selain itu penelitian lain menggunakan Metode Regresi Linear dengan menggunakan

data sebanyak 6 bulan yaitu dari bulan July hingga Desember 2008 untuk memprediksi bursa yang ada di Negara Nigeria dan penelitian lainnya menggunakan data dari Badan Meteorology serta data dari dunia pendidikan yang dapat membantu dalam pengambilan keputusan (Andini, Witanti, & Renaldi, 2016).

Ada tiga ciri atau karakteristik dari data mining yaitu :

1. Data mining berkaitan dengan upaya menemukan sesuatu yang tersembunyi dan bentuk pola data yang tidak diketahui sebelumnya.
2. Data mining biasanya bekerja pada volume data yang besar. Volume data yang besar diperlukan untuk menghasilkan kesimpulan yang handal.
3. Data mining berguna dalam pengambilan keputusan strategis suatu organisasi.

Data mining merupakan bagian yang terintegrasi dengan proses *Knowledge Discovery in Databases* (KDD). KDD adalah menggali dan menganalisis informasi dan pengetahuan yang tersembunyi dari sejumlah besar himpunan data.



Gambar 2. 1. *Tahapan proses Knowledge Discovery in Databases*

*Tahapan proses Knowledge Discovery in Databases (KDD). adalah sebagai berikut :*

a. *Data cleaning (pembersihan data)*

*Pembersihan data merupakan proses menghilangkan atau mengubah data yang tidak konsisten atau noise*

b. *Data integration (integrasi data)*

*Integrasi data merupakan integrasi atau penggabungan data dari berbagai database atau gudang data ke dalam satu database baru.*

c. *Data selection (seleksi data)*

*Seleksi data merupakan proses pengambilan data dari database yang relevan untuk dianalisis*

d. *Data transformation (transformasi data)*

*Transformasi data merupakan proses dimana data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining*

e. *Data mining (penggalian data)*

*Penggalian data merupakan suatu proses dengan melibatkan metode untuk menemukan pengetahuan yang tersembunyi dari data*

f. *Pattern evaluation (Evaluasi pola)*

*Evaluasi pola merupakan proses mengenali pola yang menarik kemudian melakukan terhadap pola atau informasi yang ditemukan*

g. *Knowledge presentation*

*Presentasi pengetahuan merupakan penyajian dan visualiasi untuk menampilkan pengetahuan kepada user*

### **2.3 Prediksi**

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. (Ismail, 2014)

### **2.4 Produksi**

Produksi adalah setiap usaha yang menciptakan atau memperbesar daya guna barang. Untuk dapat melakukan proses produksi, orang tentu melakukan tenaga kerja, bahan baku, modal dalam segala bentuknya serta keahlian atau skill. Semua unsur-unsur tersebut dean faktor-faktor produksi. Sedangkan produksi merupakan kegiatan untuk meningkatkan manfaat suatu barang. Setelah hasil produksi makan didapat hasil produksi. Hasil produksi

adalah hasil akhir dari suatu proses produksi dalam memanfaatkan input adalah output atau produk (Rosyidi, 2012).

Produksi merupakan salah satu aktivitas ekonomi yang sangat menunjang selain kegiatan industri. Tanpa adanya kegiatan produksi, konsumen tidak dapat memanfaatkan barang dan jasa yang dibutuhkannya.

Pada saat kebutuhan manusia masih sedikit dan masih sederhana, kegiatan produksi seringkali dilakukan sendiri, yaitu seseorang memproduksi untuk memenuhi kebutuhannya sendiri. Namun, seiring dengan semakin beragamnya kebutuhan dan keterbatasan sumber daya, sehingga seseorang tidak dapat lagi memproduksi sendiri barang dan jasa yang dibutuhkannya, sehingga ia membutuhkan pihak lain untuk memproduksi apa yang menjadi kebutuhannya tersebut. Secara teknis produksi dapat diartikan sebagai suatu proses mentransformasi input menjadi output, tetapi pengertian produksi dalam ilmu ekonomi mencakup tujuan kegiatan menghasilkan output serta karakter yang melekat padanya. (Arif & Amalia, 2013)

## **2.5 Naïve Bayes Classifier**

*Naïve Bayes Classifier* merupakan sebuah proses klasifikasian probabilistik sederhana yang berdasarkan pada penerapan Teorema Bayes (atau Aturan Bayes) dengan asumsi independensi (ketidaktergantungan) yang kuat, dengan kata lain, dalam Naïve Bayes, model yang digunakan adalah model fitur independen. Dalam Bayes (terutama Naïve Bayes), maksud independensi yang kuat pada fitur adalah sebuah fitur pada sebuah data tidak berkaitan dengan ada atau tidaknya fitur lain dalam data yang sama.

Keuntungan dari klasifikasi adalah metode ini hanya membutuhkan sejumlah kecil data pelatihan untuk memperkirakan parameter (sarana dan varians dari variabel) yang diperlukan untuk klasifikasi. Metode Naïve Bayes juga dapat digunakan dalam melakukan prediksi penjualan mobil di perusahaan manufaktur dan prediksi kinerja akademik mahasiswa dengan faktor yang berpengaruh meliputi Indeks Prestasi Kumulatif (IPK), Indeks Prestasi (IP) semester 1, IP semester 4 dan jenis kelamin dengan menghasilkan nilai akurasi sebesar 70%. (Andini, Witanti, & Renaldi, 2016)

NBC merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki kemampuan dan akurasi tinggi. (Saleh, 2015)

Teorema *Bayes* memiliki bentuk umum sebagai berikut :

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

Dalam hal ini:

- X : data dengan class yang belum diketahui
- H : hipotesis data X merupakan suatu *class* spesifik
- $P(H|X)$  : probabilitas hipotesis H berdasarkan kondisi X (*postetiori probability*)
- $P(H)$  : probabilitas hipotesis H (*priorprobability*)
- $P(X|H)$  : probabilitas X berdasarkan kondisi pada hipotesis H
- $P(X)$  : probabilitas dari X

Hasil dari rumus diatas untuk mencari probabilitas yang paling tinggi dengan membandingkan seluruh data yang ada yaitu  $P(H|X)$ .



Teorema *Bayesian* dengan asumsi bahwa setiap variabel bersifat bebas (*independence*). Sehingga dengan asumsi seperti itu menyimpulkan bahwa setiap atribut tidak saling terkait (*conditionally independence*) maka

$$P(X|C_i) = \sum_{k=1}^n P(X_k|C_i)$$

Setelah diperoleh hasil dari seluruh data pada setiap class, maka hasil akhirnya dapat menggunakan rumus:

$$P(X|C_i) = \arg \max P(X_i|C_i) * P(C_i)$$

## 2.6 Confusion Matrix

*Confusion matrix* adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Rumus ini melakukan perhitungan dengan 3 keluaran, yaitu: recall, precision dan accuracy (Raharjo, 2010).

1. Recall adalah proporsi kasus positif yang diidentifikasi dengan benar.  
Rumus dari recall =  $TP / (TP + FN)$
2. Precision adalah proporsi kasus dengan hasil positif yang benar. Rumus dari Precision =  $TP / (TP + FP)$
3. Accuracy adalah perbandingan kasus yang diidentifikasi benar dengan jumlah semua kasus Rumus dari accuracy =  $(TP + TN) / (TP + TN + FP + FN)$ .

## 2.7 Unified Modeling Language

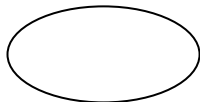

*Unified Modeling Language / UML* merupakan salah satu perangkat pemodelan. UML merupakan bahasa visual untuk pemodelan dan komunikasi mengenai sebuah sistem dengan menggunakan diagram dan teks – teks pendukung. (Munawar, 2012)

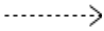
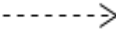



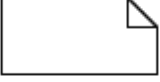
UML adalah sebuah bahasa pemrograman yang telah menjadi bahasa standar untuk merancang dan mendokumentasikan sistem perangkat lunak. UML menawarkan sebuah standar untuk merancang model sebuah sistem dan sudah digunakan secara luas dan menggunakan notasi yang sudah dikenal untuk analisa dan desain berorientasi objek.

### 2.7.1 Use Case Diagram

Diagram *Use case* merupakan pemodelan untuk kelakuan (*behavior*) sistem informasi yang akan dibuat. *Use case* mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat. *Use case* digunakan untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem informasi dan siapa saja yang berhak menggunakan fungsi-fungsi tersebut.

Tabel 2. 1. Simbol *Use case Diagram*






Gambar	Nama	Keterangan
	<i>Use case</i>	Fungsionalitas yang disediakan sistem sebagai unit – unit yang saling bertukar pesan antar unit atau aktor; biasanya dinyatakan menggunakan kata kerja di awal frase nama <i>use case</i> .
	<i>Actor</i>	Orang, proses atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat di luar sistem informasi yang akan dibuat itu sendiri sehingga walaupun simbol dari aktor adalah gambar orang, tapi aktor belum tentu merupakan orang; biasanya dinyatakan menggunakan kata benda di awal frase nama aktor

	<i>Ekstensi</i>	Relasi <i>use case</i> tambahan ke sebuah <i>use case</i> dimana <i>use case</i> yang ditambahkan dapat berdiri walau tanpa <i>use case</i> tambahan itu; mirip dengan prinsip inheritance pada pemrograman berorientasi objek; biasanya <i>use case</i> tambahan memiliki nama depan yang sama dengan <i>use case</i> yang ditambahkan
	<i>Include</i>	Relasi <i>use case</i> tambahan ke sebuah <i>use case</i> dimana <i>use case</i> yang ditambahkan memerlukan <i>use case</i> ini untuk menjalankan fungsinya atau sebagai syarat dijalankan <i>use case</i> ini
	<i>Generalization</i>	Hubungan generalisasi dan spealisasi (umum-khusus) antara dua buah <i>use case</i> dimana fungsi yang satu adalah fungsi yang lebih umum dari lainnya. Arah panah mengarah pada <i>use case</i> yang menjadi generalisasinya (umum)
	<i>Association</i>	Komunikasi antara aktor dan <i>use case</i> yang berpartisipasi pada <i>use case</i> atau <i>use case</i> memiliki interaksi dengan aktor.
	<i>System</i>	Menspesifikasikan paket yang menampilkan sistem secara terbatas.
	<i>Note</i>	Elemen fisik yang eksis saat aplikasi dijalankan dan mencerminkan suatu sumber daya komputasi

### 2.7.2 Activity Diagram

Diagram aktivitas menggambarkan aliran kerja dari sebuah sistem atau proses bisnis atau menu yang ada pada perangkat lunak. Diagram aktivitas menggambarkan aktivitas sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem.




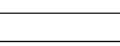
Tabel 2. 2. Simbol *Activity Diagram*


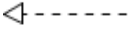
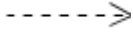
Gambar	Nama	Keterangan
	<i>Activity</i>	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja
	<i>Action</i>	State dari sistem yang mencerminkan eksekusi dari suatu aksi
	<i>Initial Node</i>	Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal
	<i>Activity Final Node</i>	Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status akhir
	<i>Fork Node</i>	Asosiasi penggabungan dimana lebih dari satu aktivitas digabungkan menjadi satu

### 2.7.3 Class Diagram

Diagram kelas menggambarkan struktur sistem dari segi pendefinisian kelas – kelas yang akan dibuat untuk membangun sistem. Kelas memiliki apa yang disebut atribut dan metode atau operasi. Diagram kelas dibuat agar dalam pembuatan program, kelas – kelas yang ada sesuai antara dokumentasi perancangan dengan perangkat lunak.

Tabel 2. 3. Simbol *Class Diagram*



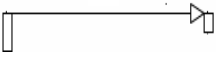
Gambar	Nama	Keterangan
	<i>Generalizatio</i>	Relasi antarkelas dengan makna generalisasi – spesialisasi ( umum khusus )
	<i>Association</i>	Relasi antarkelas dengan makna umum, asosiasi biasanya juga disertai dengan <i>multiplicity</i> .
	<i>Nary Association</i>	Upaya untuk menghindari asosiasi dengan lebih dari 2 objek.
	<i>Class</i>	Kelas pada struktur sistem


Gambar	Nama	Keterangan
	<i>Collaboration</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor
	<i>Realization</i>	Operasi yang benar-benar dilakukan oleh suatu objek.
	<i>Dependency</i>	Relasi antarkelas dengan makna kebergantungan antarkelas

#### 2.6.4 Sequence diagram

*Sequence diagram* menggambarkan kelakuan objek pada *use case* dengan mendeskripsikan waktu hidup objek dan message yang dikirimkan dan diterima antar objek. Oleh karena itu untuk menggambar diagram sekuen maka harus diketahui objek – objek yang terlibat dalam sebuah *use case* beserta metode – metode yang dimiliki kelas yang diintansiasi menjadi objek tersebut. Membuat diagram sekuen juga dibutuhkan untuk melihat scenario yang ada pada *use case*.

Tabel 2. 4. Simbol *Sequence Diagram*

Gambar	Nama	Keterangan
	<i>Aktor</i>	Orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat di luar sistem informasi yang akan dibuat itu sendiri, jadi walaupun simbol dari aktor adalah gambar orang, tapi aktor belum tentu merupakan orang; biasanya dinyatakan menggunakan kata benda di awal frase nama aktor
	<i>LifeLine</i>	Objek <i>entity</i> , antarmuka yang saling berinteraksi.
	<i>Message</i>	Spesifikasi dari komunikasi antar objek yang memuat informasi-informasi tentang aktifitas yang terjadi

	<i>Waktu Aktif</i>	Menyatakan objek dalam keadaan aktif dan berinteraksi, semua yang terhubung dengan waktu aktif ini adalah sebuah tahapan yang dilakukan di dalamnya
---	--------------------	---

## 2.8 PHP (Hypertext Preprocessor)

PHP atau yang memiliki kepanjangan *Hypertext Preprocessor* merupakan suatu bahasa pemrograman yang difungsikan untuk membangun suatu website dinamis. PHP menyatu dengan kode HTML, maksudnya adalah beda kondisi. HTML digunakan sebagai pembangun atau pondasi kerangkaan layout web, sedangkan PHP difungsikan sebagai prosesnya sehingga dengan adanya PHP tersebut, web akan sangat mudah *dimaintenance*. (Saputra, 2012)

## 2.9 MySQL

MySQL merupakan suatu database. MySQL dapat juga dikatakan sebagai database yang sangat cocok bila dipadukan dengan PHP. MySQL merupakan sebuah basis data yang mengandung satu atau sejumlah tabel. Tabel ini terdiri atas sejumlah baris dan setiap baris mengandung satu atau beberapa kolom. (Kusumaningsih & Rosa, 2014)

## 2.10 Database

Database adalah sekumpulan tabel – tabel yang berisi data dan merupakan kumpulan dari field atau kolom. Struktur file yang menyusun sebuah database adalah Data Record dan Field. Kumpulan data yang terintegrasi dan diatur sedemikian rupa sehingga data tersebut dapat dimanipulasi, diambil, dan dicari secara cepat. (Raharjo, 2012)

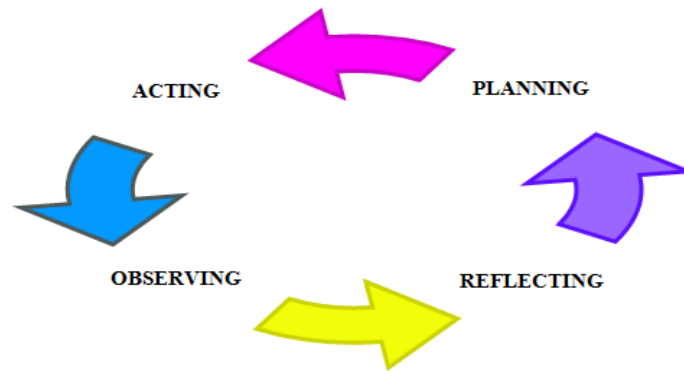
### 2.11 Penentuan Jenis Metode Penelitian

Dalam melakukan penentuan jenis metode Penelitian, maka Penulis dalam hal ini akan melakukan Cara mengkategorisasikan penelitian bisa dilakukan dengan melihat metode penelitian ataupun dengan melihat riset desainnya atau ada juga yang membaginya berdasarkan dikotonomi penelitian dasar dan penelitian aplikatif. Penelitian kuantitatif pada dasarnya merupakan suatu pengamatan yang melibatkan suatu ciri tertentu, berupa perhitungan, angka atau kuantitas. Penelitian kuantitatif ini didasarkan pada perhitungan persentase, rata-rata, kuadrat, dan juga perhitungan statistik lainnya (Hasibuan, 2011)

penelitian tindakan merupakan usaha perbaikan pemahaman, cara dan kondisi yang dilakukan secara kolaboratif dan diperkuat oleh (Stringer, 2010) mendefinisikan penelitian tindakan merupakan pendekatan sistematis dalam melakukan penyelidikan yang memungkinkan orang mendapatkan solusi yang efektif terhadap permasalahan yang sedang dihadapi.

Penelitian tindakan bertujuan untuk menyelesaikan masalah praktis yang dijumpai dalam organisasi atau komunitas dengan mengikutsertakan para pihak terkait ( stakeholders) dengan menggunakan pendekatan ilmiah untuk mencapai perbaikan dan perubahan yang diinginkan.

Menurut Kurt Lewin , terdapat empat komponen pokok dalam penelitian tindakan yaitu *Planning* (Perencanaan), *Action* (Tindakan), *Observing* (Pengamatan) dan *Reflecting* (Refleksi). Model penelitian tindakan Kurt Lewin di tunjukan pada gambar 2.2 :



Gambar 2. 2. Model penelitian tindakan

## 2.12 Review Penelitian

Untuk menunjang keterkaitan dan keterbaruan ilmu pengetahuan, berikut akan dijelaskan beberapa pendapat dari penelitian sebelumnya yang dianggap mendukung penelitian ini.

*Naive Bayes Classifier* merupakan sebuah proses klasifikasi probabilistik sederhana yang berdasarkan pada penerapan Teorema Bayes (atau aturan Bayes) dengan asumsi independensi (ketidaktergantungan) yang kuat, dengan kata lain dalam *Naive bayes* model yang digunakan adalah model fitur independen. Data penjualan dapat dimanfaatkan dengan Data Mining, Representasi Data Mining diyakini dapat berguna dalam melakukan pengelolaan, memberikan informasi dan meningkatkan nilai akurasi mengenai prediksi penjualan produk baru dan daerah pemasarannya hal tersebut didasarkan pada penelitian sebelumnya menggunakan Metode Naive Bayes Classifier dan Regresi Linear yang dapat melakukan klasifikasi terhadap data baru. (Andini, Witanti, & Renaldi, 2016)

Klasifikasi *Naive Bayes* berjalan sangat baik dibandingkan dengan model klasifikasi lainnya. Dalam penelitian yang dilakukan oleh (Purwanto



& Hastuti, 2013) dan diperkuat oleh penjelasan (Saleh A. , 2015) Penelitian yang dilakukan untuk komparasi algoritma *C4.5*, *naïve bayes*, *neural network* dan *logistic regression* Dari hasil pengujian dengan mengukur kinerja dari keempat algoritma tersebut menggunakan metode pengujian *Confusion Matrix*, diketahui bahwa algoritma *naïve bayes* memiliki nilai akurasi paling tinggi, yaitu 91,61% diikuti oleh algoritma *C4.5* dengan akurasi sebesar 89,77% metode *neural network* sebesar 84,07% dan yang terendah adalah metode *logistic regression* dengan nilai akurasi 80,02%.

Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

### Review Paper Penunjang Penelitian

	Paper I	Paper II	Paper III	Penelitian yang Diusulkan
<b>Peneliti</b>	<b>Heru Purwanto, Khafiizh Hastuti (2011)</b>	<b>Taghsya Izmi Andini, Wina Witanti, Faiza Renaldi (2016)</b>	<b>Bahrawi As'ad (2016)</b>	<b>Felida Nafiri Novitasari</b>
<b>Isi Penelitian</b>	Dalam penelitian yang dilakukan oleh (Purwanto & Hastuti, 2011) Membahas mengenai pola prediksi dari setiap atribut-atribut dari data set dengan menggunakan Algoritma <i>Naïve Bayes</i> dan melakukan	Dalam penelitian yang dilakukan oleh (Andini, Witanti & Renaldi, 2016) membahas mengenai konsep penggabungan Metode <i>Regresi Linear</i> dan <i>Naïve Bayes Classifier</i> yang dapat digunakan	Dalam penelitian yang dilakukan oleh (As`ad, 2016) membahas mengenai pengujian dengan menggunakan tiga metode klasifikasi, yaitu <i>naïve bayes</i> , <i>decision tree</i> , <i>oneR</i> Pengujian ini juga dilakukan dengan menggunakan bantuan tools	Dalam penelitian prediksi ini menggunakan metode <i>Naive Bayes Classifier</i> dan pengujian kinerja sistem menggunakan metode <i>Confusion Matrix</i> untuk menentukan keakurasian dalam memprediksi jangka waktu

	<b>Paper I</b>	<b>Paper II</b>	<b>Paper III</b>	<b>Penelitian yang Diusulkan</b>
	pengujian data training terhadap data testing untuk melihat pemodelan data sudah baik atau belum dengan Menggunakan aplikasi IDE Visual Basic 2008	untuk melakukan klasifikasi terhadap data baru.	mesin pembelajaran yang disebut “WEKA ( <i>Wakaito Environment for Knowledge Analysis</i> )”.	penyelesaian pesanan dan Dalam proses perancangan aplikasi prediksi peneliti menggunakan bahasa pemrograman berbasis <i>web</i> dan database <i>Mysql</i> .
<b>Tujuan</b>	merancang dan membangun sebuah aplikasi data mining untuk memprediksi kelayakan pemohon kredit dengan menggunakan algoritma <i>naïve bayes</i> .	Untuk memahami, memberikan informasi dan meningkatkan nilai akurasi mengenai prediksi penjualan produk baru dan daerah pemasarannya.	Untuk memahami, menganalisa dan menerangkan secara ilmiah hasil dari pengujian yang dilakukan serta mengetahui tingkat akurasi dari teknik teknik penambangan data yang digunakan dalam percobaan.	Merancang dan membangun aplikasi untuk membantu Gajendra Furniture dalam proses jangka waktu penyelesaian pesanan menerapkan perbandingan perhitungan manual dengan prediksi <i>naive bayes</i>
<b>Kekurang an</b>	program untuk melakukan pengujian prediksi masih dilakukan satu per satu tiap record, dan belum dapat melakukan pengujian dan prediksi secara menyeluruh	Sistem ini memerlukan pengembangan dengan metode yang dapat mempelajari data latih dengan cepat dan dapat membaca data latih dengan atribut yang memiliki arti yang sama, supaya dapat meningkatkan nilai akurasinya	Dalam penelitian ini menggunakan mesin pembelajaran “WEKA” dan data yang dimasukkan terlalu sedikit untuk untuk memprediksi suatu pola, akan lebih baik performansinya jika data yang dianalisis atau diolah sangat banyak.	algoritma <i>naïve bayes</i> memiliki nilai akurasi paling tinggi dibandingkan dengan algoritma <i>C4.5</i> , <i>naïve bayes</i> , <i>neural network</i> dan <i>logistic regression</i> dengan kekurangan algoritma <i>Naive Bayes</i> secara umum yaitu mengasumsikan variabel bebas.