

BAB II **LANDASAN TEORI**

2.1. Tinjauan Terhadap Penelitian Terkait

Berikut beberapa ulasan penelitian yang terkait sebelumnya yang menjadi referensi saya pada penelitian ini:

1. Ria Anggraini Walangadi, Irma Surya Kumala I dalam penelitian 2019 yang berjudul “Prediksi Penjualan Motor Dengan Menggunakan Metode Least Square” PT. Hasjrat Abadi merupakan usaha dagang yang bergerak dibidang jual beli sepeda motor. Motor yang didagangkan berupa sepeda motor Yamaha dengan berbagai jenis merk yamaha. PT. Hasjrat Abadi selama ini mengalami Penjualan Fluktuatif hal ini dikarenakan belum adanya sistem dalam peramalan yang diterapkan. Tujuan penelitian ini membuat sistem dengan teknik data mining yang akan digunakan untuk memprediksikan penjualan motor berdasarkan data transaksi penjualan serta menggunakan metode Least Square. Sehingga melalui system ini perusahaan dapat menyusun strategi penjualan. Hasil pengujian metode Least Square pada system tersebut menghasilkan tingkat akurasi dalam prediksi yang cukup baik dengan nilai akurasi sebesar 78,05% dan nilai Mean absolute Percentage Error (MAPE) sebesar 21,95%. Hasil akurasi tersebut dapat dikategorikan bahwa aplikasi yang dibuat layak untuk digunakan dalam memprediksi penjualan motor. Jumlah data yang bertambah dapat mengoptimalkan metode Least Square agar

menghasilkan prediksi yang lebih tepat dan akurat (Walangadi & Kumala, 2019).

2. Baginda Harahap dalam penelitian 2019 yang berjudul “Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)”. Pada penelitian ini, di kemukakan suatu metode pengolahan data dengan mengelompokkan data menggunakan algoritma K-Means clustering dari hasil stok penjualan UD. Toko Bangunan YD Indarung. Hasil observasi dibandingkan dengan identifikasi menggunakan algoritma K-Means clustering, sehingga di harapkan untuk penstokan bias ditambah dan bias dikurangi. Algoritma K-means, mungkin yang pertama dari algoritma pengelompokan yang diusulkan, didasarkan pada ide yang sangat sederhana: Diberi serangkaian kumpulan awal, menetapkan setiap titik ke salah satunya, lalu setiap pusat gugus diganti dengan titik rata-rata pada klaster masing-masing. Dua langkah sederhana ini diulang hingga konvergensi. Suatu titik ditugaskan ke cluster yang dekat dalam jarak Euclidean ke titik. Algoritma K-Means adalah salah satu algoritma analisis cluster yang paling umum digunakan. Inefisiensi yang digunakan K-Means untuk menangani data yang lebih besar harus diperhatikan. Seiring dengan set data yang diproses menjadi lebih besar, pendekatan host berbasis CPU tunggal tidak berdaya. Dalam tulisan ini, penulis ingin mencapai algoritma K-Means universal paralel yang mampu menangani kumpulan data yang lebih besar. Hal ini

diimplementasikan oleh CUDA. Penulis terutama memperhatikan fleksibilitas dan skalabilitas algoritma, implementasinya mungkin bukan metode yang paling efisien (Harahap, 2019).

3. Intan safira, Ratna salkawati, wowon priatna dalam penelitian pada 2022 berjudul "Penerapan Algoritma K-Means untuk Mengetahui Pola Persediaan Barang pada Toko Raja Bekasi" Permasalahan yang terjadi pada Toko Raja tersebut disebabkan karena mengalami kesulitan dalam menentukan persediaan minimum tiap barang yang harus dipenuhi berdasarkan minat konsumen. Untuk dapat mengatasi permasalahan yang terjadi, maka Toko Raja membutuhkan suatu metode dan sistem perancangan persediaan barang yang lebih baik sehingga dapat menentukan produk mana yang harus diproduksi secara banyak, sedang ataupun sedikit agar tidak lagi mengalami kekurangan atau bahkan kelebihan dalam pemenuhan persediaan produk tertentu. Untuk menyelesaikan permasalahan dengan pemanfaatan algoritma K-Means Clustering. Algoritma K-Means adalah algoritma clustering yang paling sederhana dibanding algoritma clustering yang lain. Algoritma ini mempunyai kelebihan mudah diterapkan dan dijalankan, relatif cepat, mudah untuk diadaptasi, dan paling banyak dipraktikkan dalam tugas data mining. Dalam penelitian ini dirumuskan suatu masalah yaitu bagaimana menggunakan algoritma K-Means dapat membantu clustering produk untuk penentuan prediksi persediaan barang?

Dengan tujuan membantu clustering produk untuk penentuan produksi persediaan barang, dan mengelompokkan barang yang ada di Toko Raja. Berdasarkan masalah bagaimana menggunakan algoritma K-Means dapat membantu clustering produk untuk penentuan prediksi persediaan barang? Hasil pengujian metode clustering K-Means menggunakan Davies-Bouldin Index (DBI) adalah sebesar 1,856 dimana nilai DBI mendekati nol cluster sudah cukup baik (Safira et al., 2022).

4. Indri Anatasya Alam merancang penelitian pada 2020 berjudul "PENERAPAN ALGORITMA K-MEANS UNTUK PENGELOMPOKAN DAN METODE LEAST SQUARE UNTUK PREDIKSI PENJUALAN BARANG (STUDI KASUS: BUANA MART KENDARI)" Toko Buana Mart merupakan toko yang bergerak dalam bidang penjualan bahan campuran yang menjual berbagai jenis barang atau produk serta kebutuhan pokok dan kebutuhan lainnya bagi konsumen. Dengan banyaknya jumlah data penjualan barang Buana Mart, pengelompokan jenis-jenis produk dapat digunakan untuk peningkatan stok pada produk secara tepat. Toko tersebut dalam memperkirakan stok penjualan barang masih menggunakan cara manual sehingga mengalami kesulitan dalam menentukan jumlah barang yang akan dijual. Oleh sebab itu toko tersebut memerlukan sebuah sistem yang dapat melakukan pengolahan data dan memprediksi stok penjualan barang. Untuk memudahkan dalam mengelola dan memprediksi stok barang maka

dibuat penelitian dengan menerapkan algoritma K-Means dan metode Least Square. Algoritma K-Means merupakan salah satu algoritma yang digunakan untuk melakukan pengelompokan data yang memiliki kesamaan karakteristik. Metode Least Square merupakan metode yang digunakan untuk melakukan proses perhitungan prediksi menggunakan data deret berkala pada suatu waktu yang akan datang dibuat berdasarkan data pada periode sebelumnya. Hasil penelitian ini dapat dibuat suatu aplikasi pengolahan data dengan algoritma K-Means dengan mengelompokkan barang untuk peningkatan stok. Hasil prediksi menggunakan metode least square mempunyai error (tingkat kesalahan) yang diukur dengan MAPE sebesar 18,14% (Alam et al., n.d.).

5. Wahyu brilian candra pratama merancang penelitian pada 2021 yang berjudul “RANCANG BANGUN APLIKASI PREDIKSI KEBUTUHAN STOK BARANG DAGANG DENGAN METODE LEAST SQUARE REGRESSION LINE DAN K-MEANS PADA KOPERASI PASAR KRANJI BARU BEKASI” Koperasi Pasar Kranji Baru merupakan sebuah koperasi yang bergerak di bidang penyediaan barang dan simpan pinjam yang di peruntukan kepada anggotanya. Pengeloaan penyediaan barang dan simpan pinjam pada Koperasi ini masih menggunakan sistem konvensional yaitu menggunakan buku dan Microsoft Excel. Sehingga dalam pengecekan persediaan barang, pengelolaan simpan pinjam,

pecatatan kebutuhan barang dagang anggota dapat membutuhkan waktu yang lama dan menyulitkan para petugas koperasi pasar dalam memperkirakan kebutuhan barang dagang anggota. Oleh karena itu, diperlukan sebuah aplikasi prediksi yang akan mempermudah para petugas koperasi dalam menentukan keputusan dalam menentukan berapa banyak kebutuhan barang dagang bagi para anggota koperasi. Dengan adanya aplikasi ini diharapkan dapat mempermudah dalam pengelolaan stok barang dagang, pengelolaan simpan pinjam dan mempermudah dalam menentukan prediksi kebutuhan stok barang dagang dan pengelompokan data kebutuhan stok barang dagang anggota dengan metode Least Square Regression Line dan K-Means pada Koperasi Pasar Kranji Baru Bekasi.

2.2. Peramalan (*Forecasting*)

Menurut Assauri (2016:73), prakiraan ramalan biasanya diklasifikasikan atas cakupan lamanya atau horizon waktu ke depan. Metode peramalan sendiri memiliki dua kategori utama yaitu metode kuantitatif dan metode kualitatif. Peramalan Kuantitatif yaitu memperkirakan secara kuantitatif mengenai apa yang terjadi dimasa yang akan datang Metode peramalan kuantitatif dapat dibagi menjadi dua bagian yaitu metode peramalan deret waktu dan metode kausal, sedangkan metode kualitatif dibagi menjadi metode eksploratoris dan normatif. Teknik peramalan kuantitatif sangat beragam, dikembangkan dari berbagai disiplin ilmu dan untuk berbagai maksud. Setiap teknik yang akan dipilih memiliki

sifat, ketepatan, tingkat kesulitan dan biaya tersendiri yang harus dipertimbangkan. (Dirpan,2007) menjelaskan bahwa pada umumnya peramalan kuantitatif dapat diterapkan bila terdapat tiga kondisi berikut:

1. Tersedia informasi tentang masa lalu (data historis)
2. Informasi tersebut dapat dikuantitatifkan dalam bentuk numerik
3. Dapat diasumsikan bahwa beberapa aspek pola masa lalu akan terus berlanjut di masa mendatang

Umumnya horizon waktu ke depan dibedakan atas tiga kategori, yaitu:

1. Prakiraan ramalan jangka pendek, yang mencakup jarak waktu dari tiga bulan sampai dengan dengan satu tahun. Prakiraan ramalan jangka pendek ini digunakan dalam penyusunan rencana pembelian, penjadwalan tugas pekerjaan atau *job shceduling*, penetapan level tenaga kerja atau *workforce levels*, pemberian tugas (*job assigments*), dan tingkat produksi (*production levels*).
2. Prakiraan ramalan jangka menengah (*median range forecast*), umumnya prakiraan ramalan ini mencakup masa waktu dari satu tahun sampai dengan tiga tahun. Prakiraan ramalan jangka menengah ini digunakan dalam penyusunan rencana penjualan, perencanaan produksi dan *budgeting* atau penganggaran yang meliputi anggaran kas, dan analisis berbagai rencana produksi.
3. Prakiraan ramalan jangka panjang, umumnya prakiraan ramalan ini mencakup masa waktu tiga tahun atau lebih. Prakiraan ramalan jangka panjang ini digunakan untuk perencanaan produk baru, anggaran pengeluaran modal atau *capital*

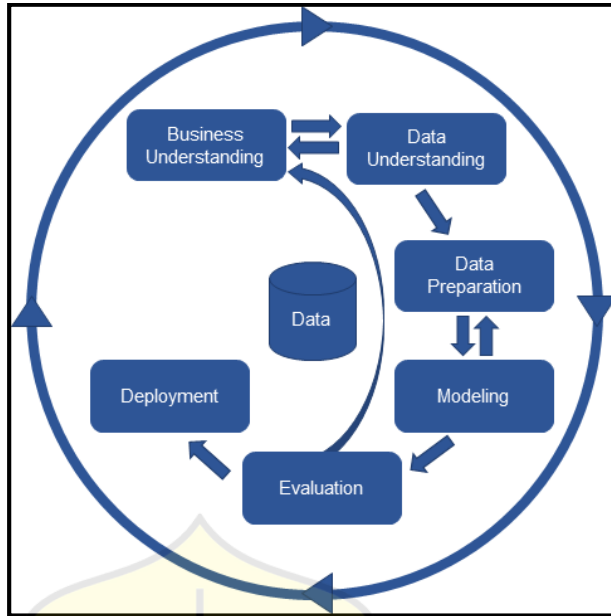
expenditure, perencanaan lokasi fasilitas ekspansi, dan riset & pengembangan (*Research & Development*).

2.3. Data Mining

2.3.1. Cross Industry Standard Process for Data Mining

Menurut Daniel T. Larose (2006:5) yang dikutip oleh Tubagus Riko Rivanthio dkk (2020), *Cross-Industry Standart Process for Data Mining* (CRISP-DM) yang dikembangkan tahun 1996 oleh analisis dari beberapa industri seperti Daimler Chrysler, SPSS dan NCR. CRISP-DM menyediakan standar proses data mining sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian

Dalam CRISP – DM, sebuah proyek *data mining* memiliki siklus hidup yang terbagi dalam enam fase. Keseluruhan fase berurutan yang ada tersebut bersifat adaptif. Fase berikutnya dalam urutan bergantung kepada keluaran dari fase sebelumnya. Hubungan penting antarfase digambarkan dengan panah. Sebagai contoh, jika proses berada pada fase modelling. Berdasar pada perilaku dan karakteristik model, proses mungkin harus kembali kepada fase *data preparation* untuk perbaikan lebih lanjut terhadap data atau berpindah maju kepada fase *evaluation*.



Gambar 2. 1 CRISP-DM
(sumber: Yogasetya, 2020)

Enam fase CRISP – DM (Chapman et al., 2000, yang di kutip oleh Yoga setya, 2020):

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)
 - a. Penentuan tujuan objek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
 - b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.
 - c. Menyiapkan strategi awal untuk mencapai tujuan.
2. Fase Pemahaman Data (*Data Understanding Phase*)
 - a. Mengumpulkan data.
 - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
 - c. Mengevaluasi kualitas data.

d. Jika diinginkan, pilih sebagian kecil grup data yang mungkin mengandung pola dari permasalahan.

3. Fase Pengolahan Data (*Data Preparation Phase*)

a. Siapkan dari data awal, kumpulkan data yang ingin digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.

b. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.

c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.

d. Siapkan data awal sehingga siap untuk perangkat pemodelan.

4. Fase Pemodelan (*Modelling Phase*)

a. Pilih dan aplikasikan teknik pemodelan yang sesuai.

b. Kalibrasi aturan model untuk mengoptimalkan hasil.

c. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan data mining yang sama.

d. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining* tertentu.

5. Fase Evaluasi (*Evaluation Phase*)

a. Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektifitas sebelum disebarkan untuk digunakan.

- b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
- c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
- d. Mengambil keputusan berkaitan dengan penggunaan hasil dari *data mining*.

6. Fase Penyebaran (*Deployment Phase*)

- a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.
- b. Contoh sederhana penyebaran: Pembuatan laporan.
- c. Contoh kompleks penyebaran: Penerapan proses data mining secara paralel pada departemen lain.

2.4. Pemrograman Aplikasi

2.4.1 Web

Menurut Ilka Zufria dan M. Hasan Azhari (2017:52), Website adalah sekumpulan halaman informasi yang disediakan melalui jalur internet sehingga bisa diakses di seluruh dunia selama terkoneksi dengan jaringan internet.

Menurut (Surajino, 2004) dalam jurnal Hendra Jaya (ISSN: 1829-7021) yang di kutip oleh (Novendri et al., 2019) Pengertian Web atau Situs Website atau situs dapat diartikan sebagai kumpulan halaman halaman yang digunakan untuk menampilkan informasi, teks, gambar diam atau bergerak, animasi, suara, dan atau gabungan dari semuanya itu, baik yang bersifat statis maupun dinamis yang membentuk satu rangkaian bangunan yang

saling berkait dimana masing masing dihubungkan dengan jaringan jaringan halaman/hyperlink .

2.4.2 HTML

Menurut Hidayatullah dan Kawistara (2015) dalam jurnal Fitri Ayu dan Nia Permata Sari (ISSN:2549-0222) “*Hypertext Markup Language* (HTML) adalah bahasa standard yang digunakan untuk menampilkan halaman *web*”. Yang bisa dilakukan dengan HTML yaitu: a. Mengatur tampilan dari halaman web dan isinya. b. Membuat tabel dalam halaman web. c. Mempublikasikan halaman web secara online. d. Membuat form yang bisa digunakan untuk menangani registrasi dan transaksi via web. Contoh: Setiap dokumen HTML diawali dan diakhiri dengan tag HTML (Novendri et al., 2019).

2.4.3 CSS

Menurut (Kurniawan, 2008:1) dalam Jurnal Siswanto dan Suwarni (ISSN:1858:2680) CSS merupakan kependekan dari Cascading Style Sheet yang berfungsi untuk mengatur tampilan dengan kemampuan jauh lebih baik dari tag maupun atribut standar HTML (Hypertext Markup Language). CSS sebenarnya adalah suatu kumpulan atribut untuk fungsi format tampilan dan dapat digunakan untuk mengontrol tampilan banyak dokumen secara bersamaan. Keuntungan menggunakan CSS yaitu jika ingin mengubah format dokumen, maka tidak perlu mengedit satu persatu (Novendri et al., 2019).

2.4.4 *JavaScript*

Menurut Jurnal yang di tulis oleh (R.H. Sianipar, 2018) *javascript* adalah bahasa scripting yang populer di internet dan dapat bekerja di sebagian besar browser populer seperti *Internet Explorer, Mozilla Firefox, Netscape dan Opera Mini*. Kode *javascript* dapat disisipkan dalam halaman web menggunakan tag *script*. Berikut ini beberapa sifat dari *javascript*:

1. Menambahkan interaktivitas ke halaman HTML.
2. Merupakan bahasa pemrograman *scripting*.
3. Bahasa *Scripting* merupakan bahasa yang ringan.
4. *JavaScript* merupakan bahasa terinterpretasi.

2.4.5 *Bootstrap*

Bootstrap adalah *framework front-end* yang intuitif dan *powerful* untuk pengembangan aplikasi *web* yang lebih cepat dan mudah. *Bootstrap* menggunakan HTML, CSS, dan *Javascript*. *Bootstrap* memiliki fitur-fitur komponen *interface* yang bagus seperti *Typografi, Forms, Buttons, Navigations, Dropdowns, Alerts, Modals, Tabs, Accordion, Carousel*, dan lain sebagainya. Dengan demikian dalam membuat *website* kita bisa menghemat waktu, fitur yang *responsive*, dan memiliki *design* yang konsisten menurut Jurnal yang di tulis oleh (Gregorius Agung, 2018). *Bootstrap* telah menyediakan kumpulan aturan dan komponen *class interface* dasar sebagai modal dalam pembuatan *web* yang telah dirancang sangat baik untuk memberikan tampilan yang sangat menarik, bersih,

ringan dan memudahkan bagi penggunanya. Dan penggunaan *bootstrap* ini kita juga diberikan keleluasan selama pengembangan *website*, anda bisa merubah dan menambah *class* sesuai dengan keinginan. *Bootstrap* memberikan kemudahan bagi anda, dengan menggunakannya dapat memangkas waktu, tenaga dalam proses pengerjaan suatu *website*. Kita selalu dituntut melakukan pekerjaan apapun dengan efisien dan efektif, dengan demikian penggunaan *framework twitter bootstrap* ini bisa anda pilih ketika membuat suatu *website* bagi anda maupun klien anda.

2.4.6 PHP

Menurut (Budi Raharjo 2012:41) dalam jurnal Yesi Susanti, dkk, (ISSN: 1858 – 2680) PHP adalah salah satu bahasan pemrograman skrip yang dirancang untuk membangun aplikasi web. Ketika dipanggil dari web browser, program yang ditulis dengan PHP akan di-parsing di dalam web *server* oleh interpreter PHP dan diterjemahkan ke dalam dokumen HTML, yang selanjutnya akan ditampilkan kembali web server (Novendri et al., 2019).

2.4.7 Database

Pengertian *Database* Menurut (Hesananda et al, 2017), *Database* ialah suatu wadah untuk menampung sebuah data yang ada pada sebuah sistem. *Database* juga bias diartikan sebagai kumpulan data. *Database* juga biasa dikenal formal dan tegas. *Database* juga bias diartikan dengan kumpulan data yang terintegrasi yang dapat dimanipulasi, diambil dan dicari secara cepat.

Jadi secara umum Basis data (*database*) adalah kumpulan informasi yang disimpan didalam komputer secara sistematis sehingga dapat diperiksa menggunakan suatu program komputer untuk memperoleh informasi dari basis data tersebut. Perangkat lunak yang digunakan untuk mengelola dan memanggil kueri (*query*) basis data tersebut disebut sistem manajemen basis data (*Database Management System*).

2.4.8 MySQL

Menurut Jurnal yang ditulis oleh (R.H. Sianipar, 2018) dalam Buku “Membangun Web dengan PHP & MYSQL untuk Pemula & Programmer”. MySQL bukan termasuk bahasa pemrograman. MySQL merupakan salah satu *database* populer dan mendunia. MySQL bekerja menggunakan *SQL Language (Structure Query Language)*. Pada umumnya, perintah yang paling sering digunakan dalam MySQL adalah *SELECT* (mengambil), *INSERT* (menambah), *UPDATE* (mengubah), dan *DELETE* (menghapus). Selain itu, SQL juga menyediakan perintah untuk membuat *database*, *field*, ataupun *index* untuk menambah atau menghapus data jurnal (Sianipar, R.H. 2015).

2.5. Algoritma Sistem

2.5.1. Least Square Regression Line

Metode *Least Square Regression Line* (LSRL) merupakan metode peramalan penjualan analisa trend linear. Pola data masa lalu yang dimiliki harus berpola tren untuk mendapatkan prediksi yang akurat. Terdapat 2 (dua) cara yaitu untuk jumlah data genap dan untuk jumlah data ganjil. Penerapan rumus metode LSRL terdapat perbedaan

yang besar untuk jumlah data genap dan ganjil, perbedaan terletak pada penentuan nilai X (periode) (Soepono, 2012:26 yang dikutip oleh (Rizqa & Jannah, n.d.))

Metode *Least Square Regression Line* (LSRL) merupakan metode peramalan penjualan analisa trend linear. Pola data masa lalu yang dimiliki harus berpola tren untuk mendapatkan prediksi yang akurat. Terdapat 2 (dua) cara yaitu untuk jumlah data genap dan untuk jumlah data ganjil. Penerapan rumus metode LSRL terdapat perbedaan yang besar untuk jumlah data genap dan ganjil, perbedaan terletak pada penentuan nilai X (periode) (Soepono, 2012:26). Rumus untuk perhitungan peramalan penjualan ditunjukkan pada persamaan.

$$Y = a + bX \quad (1)$$

Keterangan :

X = periode (waktu)

Y = jumlah penjualan pada periode X

a = bilangan konstan

b = koefisien kecondongan garis tren

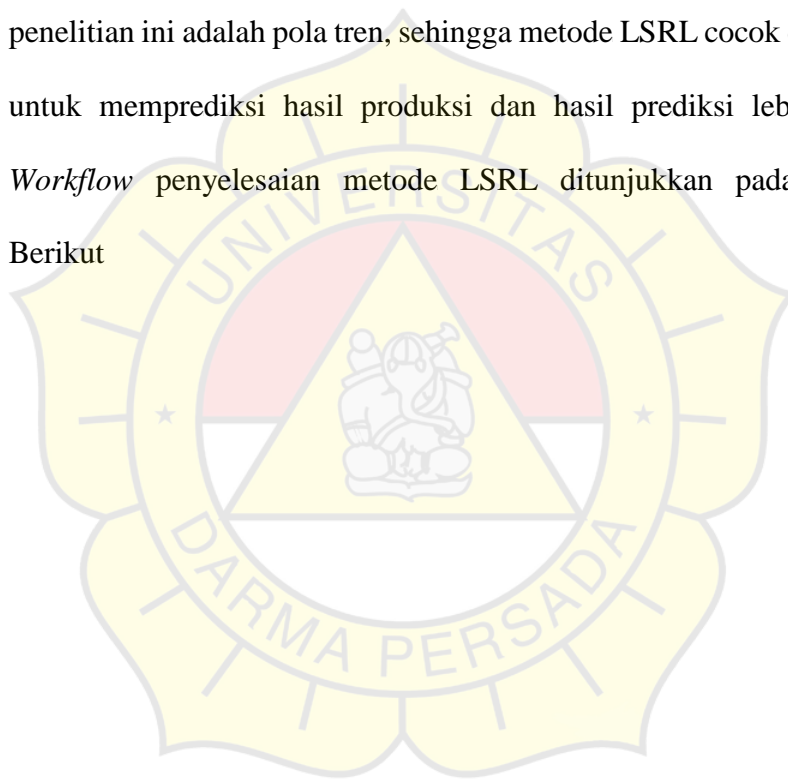
nilai a dan b diperoleh dari persamaan pembantu yang ditunjukkan pada persamaan 2 dan 3.

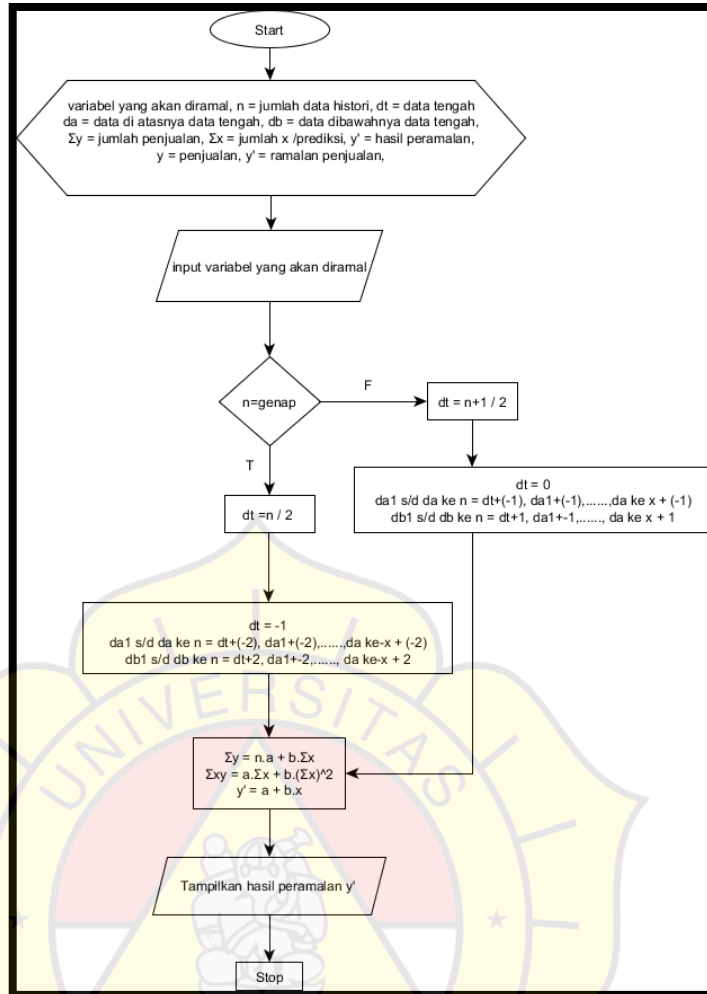
$$a = \frac{\sum Y}{N} \quad (2)$$

$$b = \frac{\sum XY}{\sum X^2} \quad (3)$$

Data periode dengan jumlah data genap dari jumlah data dibagi 2 (dua), selanjutnya nomor pertengahan diberi angka permulaan -1, penetapan angka berikutnya untuk data di atasnya (nomor lebih kecil) +

(ditambah) dengan angka -2 (minus dua) dan untuk nomor dibawahnya (nomor lebih besar) ditambah $+2$ (dua). Data periode dengan jumlah data ganjil dari jumlah data $+1$ lalu dibagi 2 (dua), selanjutnya nomor pertengahan diberi angka permulaan 0 (nol), penetapan angka berikutnya untuk data diatasnya (nomor lebih kecil) + (ditambah) dengan angka -1 (minus satu) dan untuk nomor dibawahnya (nomor lebih besar) ditambah 1 (satu) (Soepono, 2012:26). Pola data pada penelitian ini adalah pola tren, sehingga metode LSRL cocok digunakan untuk memprediksi hasil produksi dan hasil prediksi lebih akurat. *Workflow* penyelesaian metode LSRL ditunjukkan pada Gambar Berikut





Gambar 2. 2 Alur Metode Least Square Regression Line

Sumber: Tesis (Rizka Fahiratul,2015)

Keterangan :

1. Inialisasi semua variabel yang dibutuhkan untuk proses prediksi.
2. Menginputkan variabel yang akan diprediksi.
3. Mengecek jumlah data historis apakah berjumlah ganjil atau genap.
4. Jika data historis yang tersedia berjumlah genap,

diselesaikan dengan rumus $dt = n/2$ (jumlah data dibagi dua), namun apabila data historis yang tersedia berjumlah ganjil, diselesaikan dengan rumus $dt = n+1/2$ (jumlah data historis ditambah satu lalu dibagi dua)

5. Data tengah pada data historis berjumlah genap diberi angka permulaan -1, penetapan angka berikutnya untuk data di atasnya data tengah ditambah -2 dan untuk data dibawahnya ditambah 2.
6. Data tengah pada data historis berjumlah ganjil diberi angka permulaan 0, penetapan angka berikutnya untuk data di atasnya data tengah ditambah -1 dan untuk data dibawahnya ditambah 1.
7. Proses selanjutnya yaitu menghitung peramalan penjualan dengan rumus $y' = a + bX$, nilai a dan b didapatkan dengan melakukan eliminasi dua persamaan linear yaitu $\sum Y = n.a + b \sum X$ dan $\sum XY = a \sum X + b \sum X^2$.

2.5.2. K-Means

K-Means Clustering merupakan algoritma pengklasteran yang paling sederhana dibanding algoritma pengklasteran yang lain. K-Means Clustering mempunyai kelebihan mudah diterapkan dan dijalankan, relatif cepat, mudah untuk diadaptasi, dan paling banyak dipraktikkan dalam tugas data mining. Pengklasteran merupakan suatu metode untuk mengelompokkan dokumen di mana dokumen dikelompokkan dengan

konten untuk mengurangi ruang pencarian yang diperlukan dalam merespon suatu query (Grossman dan Frieder, 2004)

Cluster Analysis merupakan salah satu metode objek mining yang bersifat tanpa latihan (*unsupervised analysis*), sedangkan *K-Means Cluster Analysis* merupakan salah satu metode *cluster analysis* non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam *cluster* yang lain. Tujuan pengelompokan adalah untuk meminimalkan *objective function* yang di set dalam proses *clustering*, yang pada dasarnya berusaha untuk meminimalkan variasi dalam satu *cluster* dan memaksimalkan variasi antar *cluster*.

Metode *cluster* ini meliputi *sequential threshold*, *parallel threshold* dan *optimizing threshold*. *Sequential threshold* melakukan pengelompokan dengan terlebih dahulu memilih satu objek dasar yang akan dijadikan nilai awal *cluster*, kemudian semua *cluster* yang ada dalam jarak terdekat dengan *cluster* ini akan bergabung, lalu dipilih *cluster* kedua dan semua objek yang mempunyai kemiripan dengan *cluster* ini akan digabungkan, demikian seterusnya sehingga terbentuk beberapa *cluster* dengan keseluruhan objek yang terdapat didalamnya.

Jika diberikan sekumpulan objek $X = (X_1, X_2, \dots, X_n)$ maka algoritma *K-Means Cluster Analysis* akan mempartisi X dalam k buah *cluster*, setiap *cluster* memiliki centroid dari objek-objek dalam *cluster*

tersebut. Pada tahap awal algoritma K-Means *Cluster Analysis* dipilih secara acak k buah objek sebagai centroid, kemudian jarak antara objek dengan centroid dihitung dengan menggunakan jarak euclidian, objek ditempatkan dalam *cluster* yang terdekat dihitung dari titik tengah *cluster*. Centroid baru ditetapkan jika semua objek sudah ditempatkan dalam *cluster* terdekat. Proses penentuan centroid dan penempatan objek dalam *cluster* diulangi sampai nilai centroid konvergen (centroid dari semua *cluster* tidak berubah lagi). Secara umum metode K-Means *Cluster Analysis* menggunakan algoritma sebagai berikut:

7. Tentukan k sebagai jumlah *cluster* yang di bentuk. Untuk menentukan banyaknya *cluster* k dilakukan dengan beberapa pertimbangan seperti pertimbangan teoritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster*.
8. Bangkitkan k Centroid (titik pusat *cluster*) awal secara random. Penentuan centroid awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak k *cluster*, kemudian untuk menghitung centroid *cluster* ke-i berikutnya, digunakan rumus sebagai berikut:

$$v = \frac{\sum_{i=1}^n X_i}{n}; i = 1, 2, 3, \dots, n$$

dimana:

v: centroid pada *cluster*

I: objek ke-i

n: banyaknya objek/jumlah objek yang menjadi anggota *cluster*

9. Hitung jarak setiap objek ke masing-masing centroid dari masing-masing *cluster*. Untuk menghitung jarak antara objek dengan centroid penulis menggunakan Euclidian Distance.

$$d(x, y) = ||x - y|| = \sqrt{\sum_{i=1}^n x_i - y_i^2 ; i = 1, 2, 3, \dots n}$$

Dimana

X_i : objek x ke-i

Y_i : daya y ke-i

n: banyaknya objek

10. Alokasikan masing-masing objek ke dalam centroid yang paling terdekat. Untuk melakukan pengalokasian objek kedalam masing-masing *cluster* pada saat iterasi secara umum dapat dilakukan dengan dua cara yaitu dengan hard k-means, dimana secara tegas setiap objek dinyatakan sebagai anggota *cluster* dengan mengukur jarak kedekatan sifatnya terhadap titik pusat *cluster* tersebut, cara lain dapat dilakukan dengan fuzzy C-Means.

11. Lakukan iterasi, kemudian tentukan posisi centroid baru dengan menggunakan persamaan.

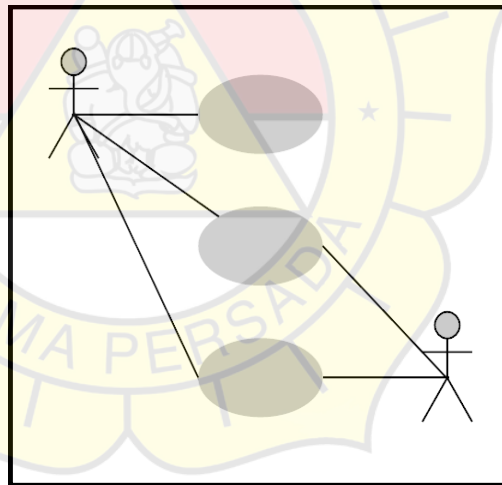
Ulangi langkah 3 jika posisi centroid baru tidak sama. Pengecekan konvergensi dilakukan dengan membandingkan matriks group assignment pada iterasi sebelumnya dengan matrik group assignment pada iterasi yang sedang berjalan. Jika hasilnya sama maka algoritma k-means *cluster*

analysis sudah konvergen, tetapi jika berbeda maka belum konvergen sehingga perlu dilakukan iterasi berikutnya

2.6. Pemodelan Sistem

Unified Modelling Language (UML) adalah sebuah bahasa yang telah menjadi standar dalam industri untuk evaluasi, merancang, dan mendokumentasikan sistem peranti lunak. UML menawarkan sebuah standar untuk merancang model sebuah sistem. UML juga dapat digunakan untuk aplikasi *modeling procedural* seperti VB atau C. jurnal (Yuni Sugiarti, 2018).

2.6.1 UseCase Diagram

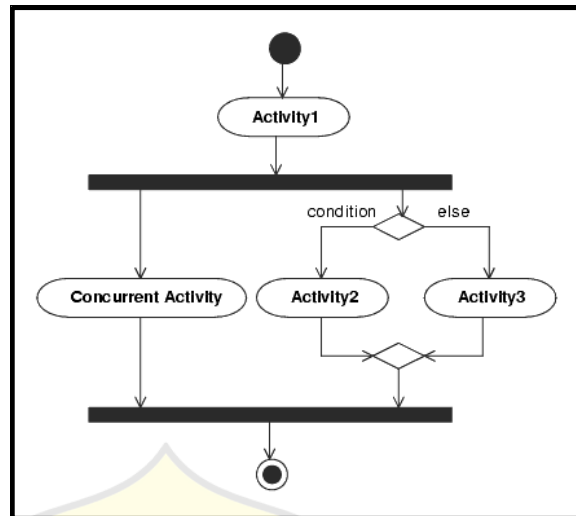


Gambar 2. 3 Use Case Diagram

Sumber: Jurnal (Yuni Sugiarti, 2018)

UseCase diagram merupakan pemodelan untuk menggambarkan *behavior* dan mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem yang akan dibuat. Jurnal (Yuni Sugiarti,2018).

2.6.2 Activity Diagram

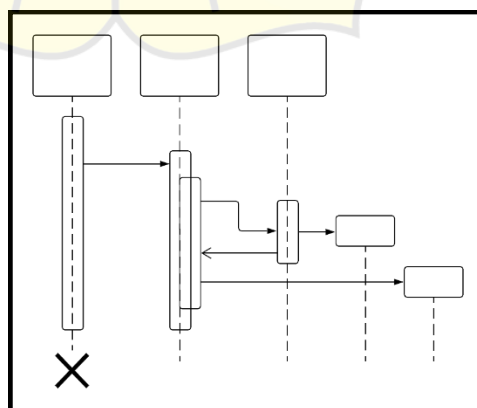


Gambar 2. 4 Activity Diagram

Sumber: Jurnal (Yuni Sugiarti,2018)

Activity diagram menggambarkan *workflow* (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis. Hal yang perlu diperhatikan di sini adalah bahwa diagram aktivitas menggambarkan kegiatan sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem. (Yuni Sugiarti,2018)

2.6.3 Sequence Diagram



Gambar 2. 5 Sequence Diagram

Sumber: Jurnal (Yuni Sugiarti,2018)

Diagram sekuens (*sequence*) menggambarkan *behavior* objek pada *Usecase* dengan mendeskripsikan waktu hidup dan *message* yang dikirimkan dan diterima antarobjek. Banyaknya diagram sekuens yang harus digambar adalah sebanyak pendefinisian *Usecase* yang memiliki proses sendiri atau yang penting semua *Usecase* telah didefinisikan interaksi jalannya pesan sudah dicakup pada diagram sekuens. (Yuni Sugiarti,2018).

