# BAB II LANDASAN TEORI

## Tinjauan Studi

Pada sebuah upaya untuk melakukan analisis maka dibutuhkan sebuah panduan serta dukungan untuk setiap hasil analisis yang sudah ada sebelumnya yang akan berkaitan dengan sebuah analisis yang sedang dilakukan. Hasil penelitian-penelitian terdahulu ini terdiri dari topik dan pembahasan hasil penelitian yang telah dilakukan.

### Hasil Rangkuman Jurnal

1. Judul : Implementasi Model Regresi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Lama Bekerja

Penulis : Yayan Adrianova Eka Tuah, Anyan (Program Studi Pendidikan Komputer, STKIP Persada Khatulistiwa Sintang)

Perusahaan tidak dapat dipisahkan dari tenaga kerja. Salah satu faktor terpenting dalam mendongkrak kinerja tenaga kerja adalah pemberian gaji yang sesuai berdasar waktu lama karyawan tersebut bekerja. Penelitian ini bertujuan untuk mengetahui prediksi gaji karyawan berdasarkan tahun lama masa kerja seorang karyawan. Penelitian ini menggunakan metode pengumpulan data dan kuisioner. Teknik analisis data yang digunakan adalah analisis regresi linear dalam python *machine learning* untuk mengetahui pengaruh variabel bebas terhadap variabel terikat. Hasil yang didapatkan dari penelitian ini ialah adanya pengaruh positif serta signifikan antara gaji dan masa kerja terhadap kinerja karyawan. Berdasarkan pengujian menggunakan fungsi *linear regression* untuk memprediksi besarnya gaji dari pengalaman bekerja, diketahui bahwa semakin bertambahnya pengalaman bekerja seorang karyawan maka gaji karyawan tersebut juga akan bertambah setiap tahunnya. Hasil tersebut memperlihatkan bahwa pengalaman bekerja memberikan pengaruh positif terhadap besarnya gaji karyawan. [1]

1. Judul : Data Mining Menggunakan Regresi Linear Untuk Prediksi Harga Saham Perusahaan Pelayaran

Penulis : Ekka Puji Ariesanto Akhmad (Program Studi Ketatalaksanaan Pelayaran Niaga, Program Diploma Pelayaran, Universitas Hang Tuah)

Pergerakan harga penutupan saham PT. BULL cenderung mengalami variasi setiap harinya. Investor perlu melakukan tindakan yang tepat, sehingga risiko yang ada dapat dikurangi dengan mengetahui naik turunnya harga saham pada masa yang akan datang, dan mampu memprediksi langkah kebijakan yang optimal untuk membuat keputusan pembelian/penjualan saham yang sesuai. Tujuan penelitian ini untuk menerapkan data mining menggunakan regresi linear untuk prediksi harga saham pada perusahaan pelayaran. Penelitian ini menggunakan metodologi *Cross Industry Standard Process for Data Mining* (CRISP-DM). Hasil penelitian menunjukkan masih ada selisih antara harga penutupan saham luaran data testing dengan harga penutupan saham aktual yang ada di bursa saham. Evaluasi nilai *Root Mean Square Error* (RMSE) menunjukkan angka plus 7,522 dari data aktual harga penutupan saham periode harian PT. BULL. [2]

1. Judul : Machine Learning Untuk Model Prediksi Harga Sembako Dengan Metode Regresi Linier Berganda

Penulis : Kandari Puteri, Astried Silvianie (Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Institut Bisnis dan Informatika (IBI) Kosgoro 1957)

Harga sembilan bahan pokok atau yang biasa disebut sembako, setiap waktu dapat naik dan turun, serta kebutuhan akan informasi harga dari sembako harian. Oleh karena itu diperlukan suatu peramalan harga sembako harian untuk diketahui dibeberapa waktu ke depan. Salah satu metode yang dapat digunakan untuk memprediksi harga yang bertipe numerik kontinu ialah dengan menggunakan metode regresi. Penelitian ini bertujuan untuk mengimplementasikan metode regresi linier berganda dalam memprediksi harga dari sembako. Data yang digunakan adalah menggunakan sampel data sembako di DKI Jakarta. Hasil yang didapatkan, sistem machine learning dapat digunakan sebagai alat bantu untuk memprediksi harga sembako harian baik harga dimasa lampau, masa sekarang maupun masa yang akan datang. [3]

1. Judul : Prediksi Harga Rumah Menggunakan Web Scrapping Dan Machine Learning Dengan Algoritma Linear Regression

Penulis : Andi Saiful, Septi Andryana, Aris Gunaryati (Universitas Nasional; Jl. Sawo Manila No.61, Pejaten Barat, Pasar Minggu, Jakarta Selatan, DKI Jakarta 12520, Jurusan Informatika, Universitas Nasional, DKI Jakarta)

Tempat tinggal atau rumah merupakan salah satu dari banyaknya kebutuhan primer bagi manusia. Maka, sangat penting untuk membuat perencanaan agar tiap keluarga dapat memiliki tempat tinggal pribadi. Dalam perencanaan tersebut dibutuhkan prediksi harga di masa mendatang. Maka, tujuan penelitian ini bertujuan untuk membuat model prediksi harga rumah menggunakan metode *machine learning* adapun algoritmanya adalah *linear regression*. Dengan melakukan *web scraping* untuk mengumpulkan data, melalui beberapa website yang bergelut dibidang jual beli rumah. Adapun menurut developer rumah yang ditanyakan dilapangan terkait variabel yang mempengaruhi harga rumah. Dari hasil penerapan prediksi harga rumah, disimpulkan bahwa pengolahan awal data yang dilakukan pada data set 7442 data menjadi 794 data sangat mempengaruhi dalam tingkat akurasi prediksi harga rumah tersebut. Dengan algoritma *linear regression* untuk memprediksi harga rumah dapat memberikan hasil keakuratan prediksi harga rumah dengan baik. [4]

1. Judul : Metode Regresi Linier Untuk Prediksi Pengadaan Inventaris Barang

Penulis : Melisa Winda Pertiwi, Richardus Eko Indrajit

Teknik prediksi data mining pada persediaan barang diperlukan di beberapa tempat, dalam beberapa kasus ada masalah pengadaan, misalnya di Dinas Pariwisata Pemuda dan Olahraga Kota Tasikmalaya. Hal ini disebabkan oleh kurangnya metode ilmiah untuk memprediksi pengadaan. Prediksi yang digunakan untuk pengadaan tahun depan diharapkan menghasilkan pengetahuan yang dapat digunakan untuk mengambil keputusan. Metode regresi linier merupakan salah satu metode yang dapat memprediksi (*forecasting*). Pada resource ini berbasis business intelligence yang berfokus untuk memprediksi persediaan barang dengan menggunakan metode regresi linier berbasis pada data sebelumnya, sehingga dapat diketahui keakuratan metode regresi linier terhadap data yang digunakan oleh RMSE (*Root Mean Square Error*). Penelitian ini berbasis business intelligence yang mana akan muncul suatu pengetahuan baru berdasarkan proses bisnis dari studi kasus untuk melakukan peramalan khususnya pada inventaris barang. Berdasarkan uji coba dataset inventaris barang menggunakan metode regresi linear, maka dapat disimpulkan bahwa metode ini baik terhadap dataset yang digunakan dengan menunjukan akurasi RMSE 0.94. [5]

1. Judul : Prediksi Kasus Covid-19 Di Indonesia Menggunakan Metode *Backpropagation* Dan Regresi Linear

Penulis : Wahyudin, Heri Purwanto (Konsentrasi Teknik Informatika, Program Studi Manajemen Informatika, STMIK LPKIA Bandung)

Kemunculan COVID-19 pertama kali menyerang manusia di Wuhan, China sedangkan di Indonesia sendiri mulai menyerang dalam lepas 2 Maret 2020 terdeteksi dua orang terkonfirmasi positif. Dari kasus tersebut tiap harinya mengalami pertambahan yang relatif signifikan. Hingga saat ini masih belum menemukan obat atau vaksin yang bisa digunakan untuk mengatasi penyebaran virus COVID-19. Adapun maksud dilakukannya penelitian ini yaitu untuk dapat memperkirakan jumlah kasus aktif pada penambahan kasus COVID-19 di Indonesia. Pada penelitian ini akan dicobakan dengan menggunakan metode *Backpropagation* dan Regresi Linear. Hasil prediksi kasus aktif dengan *Backpropagation* memberikan hasil penambahan dan penurunan yang tidak terlalu signifikan sedangkan hasil prediksi kasus aktif dengan Regresi Linear menunjukan bahwa penambahan kasus untuk tiap harinya mengalami penambahan kasus aktif. Setelah penjelasan serta penjabaran pada bab sebelumnya, dapat dijelaskan bahwa, secara umum hasil prediksi kasus aktif dengan Regresi Linear menunjukan bahwa penambahan kasus untuk tiap harinya mengalami pertambahan kasus aktif. [6]

1. Judul : Perbandingan Regresi Linear, Backpropagatian Dan Fuzzy Mamdani Dalam Prediksi Harga Emas.

Penulis : Nur Nafi’iyah

Tujuan dari penelitian ini untuk melihat hasil regresi linear, backpropagation dan fuzzy mamdani dalam memprediksi harga emas. Regresi linear merupakan persamaan garis dari data yang dikumpulkan. Fuzzy mamdani merupakan algoritma fuzzy yang menggunakan nilai yang crips(0-1). Sedangkan backpropagationmerupakan algoritma nueral network yang kompleks. Emas merupakan barang yang dapat digunakan untuk investasi. Sehingga jika kita dapat memahami bagaimana pergeseran harga emas maka kita mampu mendapatkan untung. Hasil dari ketiga metode menunjukkan bahwa korelasi dari regresi linear sangat bagus, yaitu 0,929. Dan nilai korelasi tertinggi dari ketiga metode berasal dari metode backpropagation. Hal ini terbukti bahwa dalam memprediksi harga emas menggunakan backpropagation lebih sedikit errornya ± 0,05. [7]

1. Judul : Implementasi Algoritma Regresi Linear Sederhana Dalam Memprediksi Besaran Pendapatan Daerah (Studi Kasus : Dinas Pendapatan Kab. Deli Serdang).

Penulis : Fransiskus Ginting, Efori Buulolo, Edward R Siagian.

Data Mining merupakan suatu penemuan informasi dengan melakukan penggalian pola informasi yang berisi pencarian trend dalam sejumlah data yang sangat besar serta membantu proses penyimpanan data dalam mengambil suatu keputusan diwaktu yang akan datang. Dalam menentukan pola dilakukan teknik klasifikasi mengumpulkan record (Training set). Pendapatan daerah umumnya berasal dari pajak dan retribusi daerah, pajak daerah merupakan salah satu sumber pendanaan bagi daerah secara rata-rata nasional belum mampu memberikan kontribusinya yang besar bagi pembentukan pendapatan asli daerah. Dengan memanfaatkan data Pendapatan Daerah dapat menghasilkan peramalan dan prediksi penghasilan Pendapatan Daerah kedepannya agar sesuai dengan kenyataan/realitas sehingga RAPBD yang telah direncanakan dapat berjalan dengan lancar. Regresi Linear Sederhana atau sering disingkat dengan SLR (Simple Linear Regression) merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun kuantitas untuk menggambarkan proses yang terkait dengan pengolahan data perolehan besaran pendapatan daerah. Sehingga dalam tahap pengujian dengan visual basic net dapat membantu dalam mengolah data Besaran Pendapatan Daerah yang valid. Berdasarkan proses hasil perhitungan yang telah dilakukan, maka dapat diketahui nilai prediksi besaran pendapatan daerah dinas pendapatan daerah kab. deli serdang. Dengan menggunakan algoritma regresi linear sederhana, dinilai dapat memprediksi besaran pendapatan daerah dinas pendapatan daerah kab. deli serdang untuk tahun berikutnya sehingga program-program yang telah direncanakan sebelumnya oleh dinas pendapatan dapat berjalan dengan lancar, dan juga dapat membuat programprogram yang baru agar dapat meningkatkan pendapatan daerah untuk memajukan daerah tersebut. [8]

1. Judul : Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear

Penulis : Petrus Katemba, Rosita Koro Djoh

Kopi dari daerah ini menjadi andalan ekspor hasil perkebunan, yangtelah menembus pasar internasional dengan harga tinggi karena mutunya yang baik. Namun produksi kopi cenderung menurun yang mengakibatkan permintaan akan kopi mengalmi penurunan yangdisebabkan oleh beberapa faktor, baik faktor alam dan sistem yang digunakan masih tradisional.Upaya peningkatan produksi kopi telah dilakukan pemerintah dengan berbagai cara, namun lemahnya teknologi pendukung menjadi salah satu kendala peningkatan produksi kopi. Tujuanyang ingin dicapai adalah untuk mengetahuai apakah produksi kopi mengalami peningkatan ataupenurunan dari waktu ke waktu. Untuk memenuhi kebutuhan kopi maka dilakukan prediksi denganmenggunakan Regresi linear sederhana atau Simple Regresi Linear yang merupakan salah satumetode statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun kuantitas. Simple Regresi Linear terdiri dari satu buah variabelbebas (x) dengan satu buah variabel terikat (y). Dengan melakukan prediksi menggunakan Metode Regresi Linear dapat memberikan informasi yang membantu para petani dan pemerintah dalammengambil kebijakan guna meningkatkan produksi kopi di Kabupaten Manggarai. Hasil yangdiperoleh dari penelitian ini yang melibatkan 5 periode yaitu dari tahun 2011-2015 nilai tertinggipada tahun 2015 sebesar 1.537,38 ton dan nilai terendah pada tahun 2011 sebesar 1.109. Setelahdilakukan pengujian menggunakan MSE dan MAPE di peroleh nilai MSE 43,112% dan MAPE20,001% sehingga pengujian menggunakan MAPE jauh lebih baik dalam menghitung akurasiprediksi produksi kopi. [9]

1. Judul : Peramalan Penerimaan Mahasiswa Baru Universitas Samudra Menggunakan Metode Regresi Linear Sederhana.

Penulis : Tri Novriza Putri, Adam Yordan, Dara Havisha Lamkaruna.

Peramalan penerimaan mahasiswa baru universitas samudra menggunakan metode regresi linear sederhana bertujuan untuk meramalkan atau memprediksi jumlah penerimaan mahasiswa baru dengan menganalisis data-data yang ada pada tahun-tahun sebelumnya kemudian mempresentasikannya kedalam kejadian yang akan datang dengan suatu model matematis, perhitungan ini dapat berupa perhitungan yang menggunakan pendekatan baik kuantitatif atau kualitatif.Data yang digunakan pada penelitian ini adalah data penerimaan mahasiswa pada universitas samudra tahun 2014, 2015, 2016, 2017 dan 2019. Pada masalah ini menggunakan sebuah metode dalam data mining yaitu regresi linear sederhana.Hasil yang didapat pada penelitian ini adalah penggunaan metode regresi linear sederhana dapat dipertimbangkan karena jumlah error yang didapat pada hasil prediksi penerimaan mahasiswa baru 1 tahun kedepan tidak terlalu besar yaitu sebesar 21 dari 1428 mahsiswa hasil prediksi, dimana data ini di uji pada data tahun 2018 jika diprediksi pada tahun tersebut, banyak mahasiswa yang masuk pada tahun tersebut adalah 1449. [10]

## Tinjauan Pustaka

### Pengertian Analisis

Menurut Jogiyanto (1999:129), Analisis dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh ke dalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan, kesempatan, hambatan yang terjadi dan kebutuhan yang diharapkan sehingga dapat diusulkan perbaikannya. [11]

### Pengertian Data

Data merupakan komponen utama data sistem informasi perusahaan karena semua informasi untuk pengambilan keputusan berasal dari data. Ada beberapa gagasan dalam pendefinisian istilah data, sebuah sumber menyebutkan bahwa data adalah fakta-fakta, pemikiran, atau pendapat yang tidak atau belum memiliki arti kegunaan. Sedangkan pengertian lain dari data dapat didefinisikan sebagai kelompok teratur simbol-simbol yang mewakili, tindakan, benda, dan sebagainya. Data terbentuk dari karakter yang dapat berupa alphabet, angka, maupun simbol khusus dan merupakan bentuk yang masih mentah sehingga perlu diolah lebih lanjut melalui suatu model untuk menghasilkan informasi. [12]

### Pengertian Pegawai

Soedaryono (2000:6) dalam bukunya yang berjudul “Tata Laksana Kantor”, menyebutkan bahwa pengertian pegawai adalah “Seseorang yang melakukan penghidupannya dengan bekerja dalam kesatuan organisasi, baik kesatuan kerja pemerinta maupun kesatuan kerja swasta”. Berikut, menurut Wijaya (2002:15), istilah pegawai mengandung pengertian sebagai berikut:

1. Menjadi anggota suatu kerjasama (organisasi) dengan maksud memperoleh balas jasa/imbalan kompensasi atas jasa yang telah diberikan.
2. Berada dalam sistem kerja yang sifatnya lugas/pamrih.
3. Berkedudukan sebagai penerima kerja dan berhadapan dengan pihak memberi kerja.
4. Kedudukan sebagai penerima kerja itu diperoleh setelah melalui proses penerimaan.
5. Dan akan menghadapi masa pemberhentian (pemutusan hubungan kerja antara pemberi kerja dengan penerima kerja).

Kemudian menurut Robbins (2006), dalam bukunya yang berjudul “Perilaku Organisasi”, pengertian pegawai adalah “Orang pribadi yang bekerja pada pemberi kerja, baik sebagai pegawai tetap atau tidak berdasarkan kesepakatan kerja baik tertulis maupun tidak tertulis, untuk melaksanakan suatu pekerjaan dalam jabatan atau kegiatan tertentu yang ditetapkan oleh pemberi kerja”. [13]

### Pengertian Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa yang akan datang berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar suatu kesalahan (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara tepat kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin di masa yang akan terjadi. [14]

Prediksi adalah salah satu unsur yang dapat digunakan untuk mendukung dalam pengambilan keputusan yang tepat. Prediksi dalam pengambilan keputusan didasarkan berdasar data yang ada pada waktu sekarang dan waktu lampau, sehingga mampu digunakan untuk menggambarkan kondisi yang sesuai dengan sasaran yang ingin dicapai. [15]

### Pengertian Gaji

Menurut Sukirno (2013:350), “Gaji dapat diartikan sebagai pembayaran kepada pekerja-pekerja tetap dan tenaga kerja professional, seperti pegawai pemerintah, dosen, guru, manajer, dan akuntan”. [16]

### Pengertian Sistem

Sistem merupakan suatu jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan suatu kegiatan atau untuk menyelesaikan suatu sasaran tertentu, sistem dapat juga diartikan sebagai kumpulan dari elemenelemen yang saling berinteraksi untuk mencapai suatu tujuan tertentu (Jogiyanto: 1995). [17]

### Pengertian Machine Learning

Machine learning adalah serangkaian algoritma komputer yang digunakan untuk mengoptimalkan kinerja komputer atau sistem berdasarkan data yang ada (Alpaydin, 2020). Kemampuan utama dari machine learning adalah melakukan modifikasi dan adaptasi keputusan dalam merespon perubahan (Marsland, 2015). Kegunaan machine learning antara lain sebagai berikut (Provost & Fawcett, 2013). [18]

### Linear Regression (Regresi Linear)

Regresi linear (*linear regression*) adalah metode yang dipakai untuk memperoleh model hubungan antara variabel dependen dan variabel independent (Walpole et al., 2012). Terdapat dua jenis regresi linear, yaitu terdiri dari regresi linear sederhana (*simple linear regression*) dan regresi linear ganda (*multiple linear regression*). Perbedaan terdapat pada banyaknya variabel independen. Regresi linear sederhana merupakan metode yang dipakai untuk mendapatkan model hubungan antara satu variabel dependen dan satu variabel independen, sedangkan regresi linear ganda metode yang digunakan untuk memperoleh model hubungan antara satu variabel dependen dengan lebih dari satu variabel independen (Walpole et al., 2012). [19]

#### Simple Linear Regression

Analisis regresi sederhana menimbulkan suatu persamaan yang menampilkan hubungan variabel independen dan variabel dependen. Model linear sederhana persis seperti garis lurus sederhana yang menghubungkan satu variabel bebas X dengan variabel terkait Y. Berikut dapat dipaparkan melalui rumus matematika dari garis lurus tersebut.

*Y = β0 + β1X*

Dimana *β0* disebut sebagai intersep, dan *β1* disebut sebagai kemiringan. Intersep dan kemiringan adalah nilai yang sepenuhnya mencirikan garis lurus. Intersep ditafsirkan sebagai “memposisikan” regresi, sedangkan kemiringan ditafsirkan sebagai “mengukur hubungan” antara X dan Y. Jika kemiringannya 0, maka tidak ada hubungan antar variabel. Jika kemiringannya bernilai negative, maka asosiasinya negatif. Jika kemiringannya positif, maka asosiasinya positif. Semakin besar nilai absolut kemiringannya, semakin “kuat” asosiasinya. [22]

#### Scikit-Learn

Skicit-learn adalah modul pada bahasa pemrogaman Python yang menyediakan berbagai jenis algoritma machine learning. Bentuk library python merupakan bentuk yang tersedia dalam Scikit-Learn. Scikit-Learn memanfaatkan task-oriented interface yang konsisten sehingga memudahkan dalam membandingkan antarmetode. Scikit-Learn adalam modul Python yang mengintegrasikan berbagai algoritma machine learning untuk supervised learning dan unsupervised learning. [23]

## Metodologi Penelitian

Penelitian ini merupakan jenis penelitian deskriptif kualitatif dengan menggunakan simple linear regression yang menggunakan bahasa pemrograman Python. Penelitian deskriptif adalah penelitian yang digunakan untuk menemukan pengetahuan yang seluas-luasnya terhadap objek penelitian pada suatu masa tertentu. Penelitian deskriptif menyajikan satu gambar yang terperinci tentang satu situasi khusus. Penelitian deskriptif bertujuan untuk menjelaskan ataua mendeskripsikan suatu keaadan apa adanya dan menginterpretasi objek sesuai dengan apa adanya, peristiwa, atau segala sesuatu yang terkait dengan variabel-variabel yang bisa dijelaskan baik dengan angka-angka maupun kata-kata.[20]

Analisis data yang dilakukan adalah secara kualitatif. Analisi kualitatif merupakan analisis yang mencoba untuk memahami, mendalami, dan menerobos masuk di dalamnya terhadap suatu gejala-gejala. Kemudian menginterpretasikan dan menyimpulkan gejala-gejala tersebut sesuai dengan konteksnya. Sehingga dicapai suatu simpulan yang objektif dan alamiah sesuai dengan gejala-gejala pada konteks tersebut.[21] Fakta, masalah, dan gejala dapat dipahami oleh peneliti apabila diamati dan tidak dibatasi ide-ide awam. Oleh sebab itu, penelitian deskriptif kualitatif selalu digunakan untuk menganalisis peristiwa, fenomena maupun situasi sosial.

