

SKRIPSI

**PERBANDINGAN METODE REGRESI DAN JARINGAN
SARAF TIRUAN *BACKPROPAGATION*
(PREDIKSI HARGA MAKANAN LAUT)**

***COMPARISON OF REGRESSION METHODS AND
ARTIFICIAL NEURAL NETWORK BACKPROPAGATION
(SEAFOOD PRICE PREDICTION)***



**RAFLI ADIGUNA SAPUTRA
D0219525**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS SULAWESI BARAT
MAJENE
2023**

ABSTRAK

Skripsi ini ditulis oleh Rafli Adiguna Saputra dengan NIM D0219525 Program studi Teknik Informatika dibimbing langsung oleh Arnita Irianti, M.Si, S.Si dan Nurdina Rasjid, S.Pd, M.Pd untuk memenuhi salah satu syarat mendapatkan gelar Sarjana Teknik di Universitas Sulawesi Barat

Penggunaan data time series dalam melakukan prediksi dengan linear regresi adalah salah satu metode yang dapat digunakan untuk memahami dan meramalkan perilaku data yang berkaitan dengan waktu. Dalam penggunaan data time series, jaringan saraf tiruan dapat dilatih untuk mengidentifikasi pola, tren, serta siklus yang mungkin tersembunyi dalam data tersebut. Memanfaatkan informasi historis, jaringan saraf tiruan dapat merancang model yang mampu melakukan prediksi masa depan dengan akurasi yang tinggi. Metode Penelitian yang digunakan adalah metode penelitian kualitatif dengan Data time series penelitian ini diperoleh dari dataset time series harga makanan ikan dengan jumlah 9 spesies ikan, 2 spesies udang dan 1 jenis kepiting mulai dari Januari 2018 sampai Desember 2022 dengan teknik analisis data yaitu analisis plot time series untuk melihat pola dalam rentang waktu. Hasil penelitian ini yaitu hasil rata-rata akurasi prediksi seluruh spesies hewan menggunakan algoritma linear regresi adalah 95.26% sedangkan menggunakan JST atau jaringan saraf tiruan backpropagation adalah 96.50% sehingga dapat disimpulkan bahwa prediksi menggunakan algoritma JST lebih akurat dari linear regresi berganda.

Kata Kunci: *Prediksi, Linear Regresi Berganda, Jaringan syaraf tiruan, harga spesies hewan*

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Fluktuasi harga ikan dalam pasar perikanan merupakan fenomena yang kompleks dan signifikan. Harga ikan pada suatu periode dapat dipengaruhi oleh berbagai faktor, termasuk ketersediaan ikan pada periode sebelumnya. Hal ini bisa menyebabkan kenaikan harga pada periode berikutnya karena permintaan yang tetap tinggi. Selain itu, cuaca dan musim juga memainkan peran penting dalam siklus harga ikan. Musim tertentu bisa menghasilkan penangkapan ikan yang melimpah, yang kemudian memengaruhi harga ikan pada periode berikutnya. Selain faktor-faktor lokal, pasar ikan juga dipengaruhi oleh dinamika global, termasuk permintaan internasional, kebijakan perdagangan, dan perubahan ekonomi di negara-negara pengeksport dan pengeimpor. Studi-studi sebelumnya telah mengidentifikasi korelasi antara harga ikan pada periode sebelumnya dengan harga saat ini, memberikan dasar untuk pemahaman lebih lanjut. Di era teknologi, sistem informasi pasar dan perangkat lunak analitik memainkan peran penting dalam meramalkan fluktuasi harga ikan berdasarkan data historis. Oleh karena itu, pemahaman yang mendalam tentang bagaimana harga ikan pada periode sebelumnya memengaruhi harga saat ini sangat penting. Para pelaku industri perikanan dapat menggunakan pengetahuan ini untuk merencanakan produksi, penyimpanan, dan pemasaran ikan dengan lebih efektif, mengurangi risiko, dan mengoptimalkan proses bisnis mereka (Irawati & Syamsuddin, 2020). Di sisi lain, faktor-faktor ekonomi seperti biaya produksi, transportasi, dan inflasi juga

berperan. Jika biaya produksi dan distribusi ikan naik, produsen dan pedagang mungkin menaikkan harga jualnya untuk menjaga keberlanjutan bisnis mereka (Sepriano, S.Sos., M.Kom., CBPA, 2023). Melalui analisis harga ikan yang dipengaruhi oleh harga ikan pada periode sebelumnya, para pemangku kepentingan di sektor perikanan dapat mengambil keputusan yang lebih baik. Nelayan, pedagang, dan distributor ikan dapat menggunakan informasi ini untuk mengatur penangkapan dan pasokan ikan dengan lebih efisien. Mereka dapat mengantisipasi perubahan harga dengan lebih baik, menghindari penumpukan stok saat harga rendah, dan memaksimalkan keuntungan saat harga tinggi. Selain itu, pemahaman lebih mendalam tentang faktor-faktor yang memengaruhi harga ikan juga dapat membantu pembuat kebijakan dalam mengembangkan strategi yang mendukung keberlanjutan industri perikanan. Ini termasuk upaya untuk mengelola sumber daya perikanan secara berkelanjutan, melalui kuota penangkapan yang bijaksana dan perlindungan terhadap ekosistem laut. Dengan demikian, analisis harga ikan yang mempertimbangkan faktor-faktor sebelumnya adalah alat penting dalam menjaga stabilitas dan profitabilitas sektor perikanan, sambil melindungi lingkungan laut yang semakin penting (Muttaqin & Ismail, 2023).

Prediksi (*prediction*) adalah memperkirakan nilai-nilai data bertipe apa saja dan kapan saja (masa lalu, sekarang, dan masa depan). Terdapat satu istilah yang mirip dengan prediksi, yaitu peramalan (*forecasting*) adalah memperkirakan nilai-nilai data *time series* dimasa depan (Suyanto, 2018). Proses prediksi dalam *machine learning* yaitu dengan memahami sifat atau ciri dari objek yang tidak

dikenal dengan cara mengidentifikasi pola (*pattern*) dalam himpunan data (*dataset*) (Puteri & Silvanie, 2020).

Time series adalah sebuah jenis data yang merekam nilai dari suatu variabel sepanjang waktu. *Time series* dapat digunakan sebagai dataset untuk prediksi, yaitu untuk mengestimasi nilai variabel di masa depan berdasarkan pola-pola yang terdapat dalam data historis (Pradito & Purnia, 2022). *Time series analysis* adalah proses untuk membuat model matematis yang dapat menangkap pola-pola tersebut dan menghasilkan prediksi yang akurat dan andal. *Time series analysis* merupakan salah satu teknik analisis prediksi yang sangat berguna untuk berbagai bidang dan aplikasi, seperti ekonomi, keuangan, bisnis, kesehatan, cuaca, dan lain-lain sehingga dengan menggunakan *time series analysis*, kita dapat membuat keputusan yang lebih baik berdasarkan informasi yang didapat dari data *time series* (Yuliyanto et al., 2023).

Dalam konteks *machine learning*, *dataset time series* dapat digunakan untuk melatih model prediksi. Model ini dapat belajar dari pola dalam data historis dan kemudian digunakan untuk membuat prediksi tentang masa depan. Misalnya, model mungkin belajar bahwa penjualan cenderung meningkat menjelang liburan, dan kemudian dapat meramalkan peningkatan penjualan di masa depan berdasarkan pola ini. Penggunaan data *time series* dalam melakukan prediksi dengan linear regresi adalah salah satu metode yang dapat digunakan untuk memahami dan meramalkan perilaku data yang berkaitan dengan waktu. Regresi linear adalah teknik statistik yang digunakan untuk memodelkan hubungan antara variabel dependen (variabel yang ingin diprediksi) dan satu atau

lebih variabel independen (variabel yang digunakan sebagai prediktor) dalam bentuk persamaan linear. Dalam konteks data *time series*, variabel independen biasanya merupakan waktu atau indeks waktu (Kwok & Susanti, 2019). Dalam penggunaan data *time series*, jaringan saraf tiruan dapat dilatih untuk mengidentifikasi pola, tren, serta siklus yang mungkin tersembunyi dalam data tersebut. Memanfaatkan informasi historis, jaringan saraf tiruan dapat merancang model yang mampu melakukan prediksi masa depan dengan akurasi yang tinggi (Hauriza et al., 2021). Namun, penting untuk memastikan kualitas data yang baik, serta melibatkan pemahaman *domain* yang kuat dalam proses pelatihan dan evaluasi model agar hasil prediksi dari jaringan saraf tiruan dapat menjadi alat yang berharga dalam pengambilan keputusan (Trisna & Permana, 2019).

Berdasarkan latar belakang di atas, peneliti memiliki ide untuk melakukan prediksi harga ikan menggunakan dua algoritma yaitu regresi linear dan jaringan saraf tiruan. Regresi linear adalah metode statistik yang digunakan untuk memodelkan hubungan linier antara variabel *dependen* (*output*) dan satu atau lebih variabel *independen* (*input*). Dalam konteks prediksi *time series* dapat menggunakan regresi linear untuk memodelkan hubungan antara waktu sebagai variabel independen dan data *time series* sebagai variabel *dependen*. Jaringan saraf tiruan (*neural networks*) adalah model matematika yang terinspirasi dari struktur jaringan saraf manusia yang terdiri dari neuron buatan atau "*node*" yang terhubung dalam lapisan atau "*layer*." Dalam konteks prediksi *time series*, jaringan saraf tiruan dapat mempelajari hubungan antara waktu dan data *time series* dengan cara yang lebih fleksibel mampu menangkap pola yang non-linier

dan kompleks dalam data, yang sangat penting dalam analisis *time series*. Dua algoritma tersebut akan dibandingkan akurasi prediksinya berdasarkan *error analysis* dan akurasi prediksi. Dua algoritma tersebut dapat melakukan prediksi dengan tahapan-tahapan yang dapat mempelajari pola-pola dalam data *time series* yang digunakan dalam penelitian ini maka dari itu judul penelitian ini adalah PERBANDINGAN AKURASI LINEAR REGRESI DAN JARINGAN SARAF TIRUAN *BACKPROPAGATION* PADA PREDIKSI HARGA IKAN

B. Batasan Penelitian

Penelitian ini memiliki batasan berdasarkan kondisional penelitian. Batasan masalah penelitian ini adalah

1. Menggunakan algoritma linear regresi berganda sebagai implementasi sistem prediksi
2. Menggunakan algoritma jaringan saraf tiruan backpropagation sebagai implementasi sistem prediksi
3. Menggunakan dataset *open source time series* data harga ikan rentang waktu tahun 2018 – 2022 yang diperoleh dari data *kaggle* : <https://www.kaggle.com/datasets/niluhputuanitadewi/dataset-harga-ikan-gresik>

C. Rumusan Masalah

Berdasarkan latar belakang yang telah peneliti paparkan, penelitian ini memiliki rumusan masalah yaitu :

1. Bagaimana membangun sistem prediksi menggunakan algoritma linear regresi berganda?

2. Bagaimana membangun sistem prediksi menggunakan jaringan saraf tiruan *backpropagation*?
3. Apa hasil perbandingan dua algoritma linear regresi berganda dan jaringan saraf tiruan *backpropagation* dalam prediksi harga ikan?

D. Tujuan Penelitian

Penelitian ini memiliki tujuan sebagai berikut :

1. Untuk mengetahui sistem prediksi menggunakan algoritma linear regresi berganda
2. Untuk mengetahui sistem prediksi menggunakan jaringan saraf tiruan *backpropagation*
3. Untuk memperoleh hasil perbandingan dua algoritma linear regresi berganda dan jaringan saraf tiruan *backpropagation* dalam prediksi harga ikan

E. Manfaat Penelitian

1. Dapat mengetahui prediksi linear regresi berganda dan jaringan syaraf tiruan *backpropagation* dan membandingkan akurasi kedua algoritma dalam memprediksi harga ikan.
2. Dapat dijadikan sebagai sumber informasi terkait harga ikan di masa yang akan datang atau beberapa bulan kedepan

BAB II

DASAR TEORI

A. Data Mining

Data mining merupakan proses penggalian informasi dan pola yang bermanfaat dari data yang sangat besar. *Data mining* mencakup pengumpulan data, ekstraksi data, analisis data dan statistik data. *Data mining* juga dikenal sebagai *knowledge discovery*, *knowledge extraction*, *data/pattern analysis*, *information harvesting* dan lain-lain. *Data mining* juga merupakan proses logis untuk menemukan informasi yang berguna. Setelah ditemukan informasi dan pola yang dapat digunakan untuk alat pendukung dalam pengambilan keputusan dalam mengembangkan bisnis. Alat data mining dapat memberikan jawaban untuk berbagai pertanyaan yang terkait dengan bisnis dan terlalu sulit untuk diselesaikan. *Data mining* juga dapat digunakan untuk meramalkan tren masa depan yang memungkinkan pebisnis membuat keputusan yang efektif, proaktif dan dinamis. Data-data yang diolah dengan menggunakan teknik *data mining*, juga mampu menghasilkan pengetahuan yang sesuai dengan harapan. Misalnya pada bidang kesehatan, cukup banyak data yang dimiliki rumah sakit, seperti *medical record* dan radiologi, tetapi karena belum adanya standar koleksi data maka data-data tersebut sukar untuk diolah sehingga dengan kehadiran *data mining*, maka diharapkan data-data yang dimiliki oleh pihak kesehatan dapat diolah sesuai dengan keperluan hingga menghasilkan informasi dan pengetahuan

yang dapat dimanfaatkan oleh para pengambil kebijakan terutama pemerintah. (Arhami & Nasir, 2020).

Berdasarkan (Zai, 2022). sebagai salah satu bagian dari sistem informasi, data mining menyediakan perencanaan dari ide hingga implementasi akhir. Komponen-komponen dari rencana data mining adalah sebagai berikut

1. Analisa Masalah (*Analyzing the Problem*)

Data asal atau data sumber harus bisa ditaksir untuk dilihat apakah data tersebut memenuhi kriteria *data mining*. Kualitas kelimpahan data adalah faktor utama untuk memutuskan apakah data tersebut cocok dan tersedia sebagai tambahan. Hasil yang diharapkan dari dampak *data mining* harus dengan hati-hati dimengerti dan dipastikan bahwa data yang diperlukan membawa informasi yang bisa diekstrak.

2. Mengekstrak dan Membersihkan Data (*Extracting dan Cleansing The Data*)

Data pertama kali diekstrak dari data aslinya, seperti dari basis data, *text file*, *Microsoft Acces Database*, dan bahkan dari *spreadsheet*, lalu data tersebut diletakkan dalam *data warehouse* yang mempunyai struktur yang sesuai dengan data model secara khas. *Data Transformation Service (DTS)* dipakai untuk mengekstrak dan membersihkan data dari tidak konsisten dan tidak kompatibel dengan format yang sesuai.

3. Validitas Data (*Validating the Data*)

Sekali data telah diekstrak dan dibersihkan, ini adalah latihan yang bagus untuk menelusuri model yang telah kita ciptakan untuk memastikan bahwa semua data yang ada adalah data sekarang dan tetap.

4. Membuat dan Melatih Model (*Create and Training the Model*)

Ketika algoritma diterapkan pada model, struktur telah dibangun. Hal ini sangatlah penting pada saat ini untuk melihat data yang telah dibangun untuk memastikan bahwa data tersebut menyerupai fakta di dalam data sumber.

5. Query Data dari Model *Data Mining* (*Querying the Model Data*)

Ketika model yang telah cocok diciptakan dan dibangun, data yang telah dibuat tersedia untuk mendukung keputusan. Hal ini biasanya melibatkan penulisan *front end query* aplikasi dengan program aplikasi/suatu program basis data.

proses fase *data mining* dimulai dari data mentah dan berakhir dengan pengetahuan atau informasi yang telah diolah, yang didapatkan sebagai hasil dari tahapan-tahapan berikut:

1. *Data Cleansing*, juga dikenal sebagai data *cleansing*, ini adalah sebuah fase dimana data-data tidak lengkap, mengandung *error* dan tidak konsisten dibuang dari koleksi data, sehingga data yang telah bersih relevan dapat digunakan untuk diproses ulang untuk penggalian pengetahuan (*discovery knowledge*)
2. *Data Integration*, pada tahap ini terjadi integrasi data dimana sumber-sumber data yang berulang (*multiple data*), file-file yang

berulang(*multiple file*), dapat dikombinasikan dan digabungkan ke dalam suatu sumber.

3. *Data Selection*, pada langkah ini, data yang relevan terhadap analisis dapat dipilih dan diterima dari koleksi data yang ada.
4. *Data Transformation*, juga dikenal sebagai data *consolidation*. Pada tahap ini, dimana data-data yang telah terpilih, ditransformasikan ke dalam bentuk-bentuk yang cocok untuk prosedur penggalian (*mining procedure*) dengan cara melakukan normalisasi dan agregasi data.
5. *Data Mining*, tahap ini adalah tahap yang paling penting, dengan menggunakan teknik-teknik yang diaplikasikan untuk mengekstrak pola-pola potensial yang berguna.
6. *Pattern Evaluation*, pada tahap ini, pola-pola menarik dengan jelas mempresentasikan pengetahuan telah diidentifikasi berdasarkan *measure* yang telah diberikan.
7. *Knowledge Representation*, ini merupakan tahap terakhir dimana pengetahuan yang telah ditemukan secara visual ditampilkan kepada *user*. Tahap penting ini menggunakan teknik visualisasi untuk membantu *user* dalam mengerti dan menginterpretasikan hasil dari data *mining*. (Zai, 2022)

Dalam (Gimantra,2021), Secara garis besar kegunaan *data mining* dibagi menjadi dua yaitu: deskriptif dan prediksi. Secara deskriptif artinya *data mining* berarti menemukan pola-pola yang digunakan untuk memperjelaskan karakteristik data. Sedangkan dengan secara prediktif *data mining* berarti dapat digunakan

untuk menemukan model-model pengetahuan yang digunakan untuk melakukan prediksi.

Masih dalam Gimantra,2021, *Data mining* berdasarkan fungsionalnya 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 target, sedangkan yang lain disebut dengan atribut predator. Klasifikasi ini juga umum digunakan untuk permodelan 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 label nya. Klastering yang diorganisasi ke dalam struktur hierarkial akan mendefinisikan taksonomi dari data. Penerapan metode klastering yang tetap akan menghasilkan clustering yang berkualitas. Suatu cluster di karakterisasi 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 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 *devias/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 (Ginantra, 2021).

Berdasarkan teori yang peneliti paparkan sebelumnya, *data mining* merupakan proses pengambilan informasi dari sebuah data besar (*big data*) yang kemudian dilakukan ekstraksi data untuk dapat memperoleh informasi sesuai tujuan. Baik *supervised* maupun *unsupervised learning* adalah pendekatan yang dilakukan

algoritma komputer dalam mengenali pola pada data. *Supervised* mengenali data dari label khusus yang telah diberikan sebelumnya, sedangkan *unsupervised* mengenali data secara *real-time* begitu data disajikan

B. Prediksi

Prediksi merupakan suatu proses untuk meramalkan atau memperkirakan suatu variabel di masa yang akan datang. Prediksi sendiri terbagi atas 3 bagian, yaitu prediksi jangka panjang, jangka menengah dan jangka pendek. Prediksi jangka pendek merupakan prediksi yang dilakukan dengan memperhatikan pola data, dan membutuhkan jangka waktu yang pendek terhadap perubahan berdasarkan faktor-faktor yang membentuk pola data. Sedangkan prediksi jangka menengah dan jangka panjang digunakan untuk perencanaan strategis. Prediksi jangka menengah membantu untuk menyiapkan ekspansi dan mengantisipasi kebutuhan. Prediksi jangka panjang berfungsi untuk menjamin ketersediaan kebutuhan di masa depan.

Menurut Hadapiningradja dan Sarwido (2017), prediksi merupakan salah cara atau proses untuk memprediksi atau memperkirakan secara urut dan sistematis mengenai sesuatu yang mungkin dapat terjadi pada masa depan berdasarkan tentang informasi pada masa lalu dan sekarang yang dimiliki, agar tingkat error dan kesalahannya dapat diperkecil. Dalam prediksi tidak harus memberikan suatu jawaban secara pasti tentang kejadian yang nanti akan terjadi pada masa yang akan datang, melainkan berusaha untuk mencari jawaban yang akurat mungkin nanti akan terjadi.

Prediksi merupakan suatu usaha untuk meramalkan keadaan di masa mendatang melalui pengujian keadaan dimasa lalu. Hal ini dapat dilakukan dengan melibatkan pengambilan data masa lalu dan menempatkannya ke masa yang akan datang dengan suatu bentuk model matematis. Jadi bisa disimpulkan bahwa prediksi itu dapat memberikan hasil di masa yang akan datang dengan menghitung kejadian di masa sebelumnya. Jadi pada dasarnya prediksi ini sendiri tentunya tidak memiliki keakuratan 100%, kesalahan prediksi pada masa lalu dipakai untuk mengoreksi prediksi mendatang pada arah yang berlawanan dengan kesalahan tersebut. Penyesuaian tersebut tetap berlangsung sampai kesalahannya di koreksi. Prinsip ini yang tampaknya sederhana, yang memainkan peranan yang sangat penting dalam prediksi. Apabila dilihat berdasarkan sifat-sifat prediksi, maka teknik prediksi dapat dilakukan dengan 2 cara, yaitu (Hayami, R., & Oktaviandi, I. (2021) :

1. Secara kualitatif (non statistical method) adalah cara penaksiran yang menitikberatkan pada pendapat seseorang (judgement). Hal ini penting karena hasil prediksi tersebut ditentukan berdasarkan pemikiran yang bersifat intuisi, pendapat dan pengetahuan dari orang yang menyusunnya, Ada beberapa sumber pendapat yang dipakai sebagai dasar melakukan prediksi penjualan, antara lain pendapat salesman, pendapat manajer penjualan, pendapat para ahli, dan survei konsumen.
2. Secara kuantitatif (statistical method) adalah cara penaksiran yang menitikberatkan pada perhitungan-perhitungan angka dengan

menggunakan berbagai metode statistik. Hasil prediksi yang dibuat sangat bergantung pada metode yang dipergunakan dalam prediksi tersebut.

C. Regresi Linear

Berdasarkan (Bhirawa, 2020), Analisis regresi linier digunakan untuk peramalan, dimana dalam model terdapat variabel bebas X dan variabel bebas Y. Regresi linier itu menentukan satu persamaan dan garis yang menunjukkan hubungan antara variabel bebas dan variabel tak bebas yang merupakan persamaan penduga yang berguna untuk menaksir/meramalkan variabel tak bebas. Untuk mempelajari hubungan-hubungan antara variabel bebas, analisis ini terdiri dari dua bentuk, yaitu:

1. Analisis regresi sederhana
2. Analisis regresi berganda

1. Regresi Linear Sederhana (*Simple Linear Regression*)

Berdasarkan Daniya, 2020, Model regresi linier sederhana dikatakan sebagai model dengan regresor tunggal 'X' yang memiliki hubungan linier dengan respons 'Y'.

Model regresi linier sederhana adalah,

$$Y = \beta_0 + \beta_1 X + \varepsilon \text{ -----} \quad (2. 1)$$

Dimana, Y->response variable (atau variabel output), X ->regressor variable (atau input variable), β_0 -> intercept, β_1 ->slope, ε ->random error component. Di sini 'Y' adalah variabel acak dan 'X' tidak acak tetapi deterministik (Daniya et al., 2020).

Berdasarkan teoritis dari pakar diatas, Secara umum, persamaan rumus regresi linear sederhana (Daniya et al., 2020) :

$$Y = a + b x \quad (2. 2)$$

$$a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (2. 3)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} \quad (2. 4)$$

dimana :

Y = persamaan regresi linear

a = konstanta

b = koefisien

x = nilai estimasi

2. Regresi Linear Berganda (*Multiple Linear Regression*)

Berdasarkan (Loban, 2023), Regresi linier berganda adalah analisis regresi yang menjelaskan hubungan antara peubah respon (variabel dependen) dengan faktor-faktor yang mempengaruhi lebih dari satu prediktor (variabel independen. Ketika suatu hasil/keluaran, atau kelas berupa numerik, dan semua atribut adalah numerik, regresi linear adalah teknik yang tepat untuk menyelesaikan

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nx_n \quad (2. 5)$$

Dimana:

Y = variabel tidak bebas

x = Variabel bebas

a = konstanta (nilai Y apabila, $X_1, X_2 \dots X_n = 0$)

b = koefisien

nilai a dan b dapat dihitung dengan persamaan

$$b_1 = \frac{(\sum X_2^2)(\sum X_1 Y) - (\sum X_2 Y)(\sum X_1 X_2)}{(\sum X_1^2)(\sum X_2^2) - (\sum X_1 X_2)^2} \quad (2.6)$$

$$b_2 = \frac{(\sum X_1^2)(\sum X_2 Y) - (\sum X_1 Y)(\sum X_1 X_2)}{(\sum X_1^2)(\sum X_2^2) - (\sum X_1 X_2)^2} \quad (2.7)$$

$$a = \frac{\sum Y - (b_1 \sum X_1) - (b_2 \sum X_2)}{n} \quad (2.8)$$

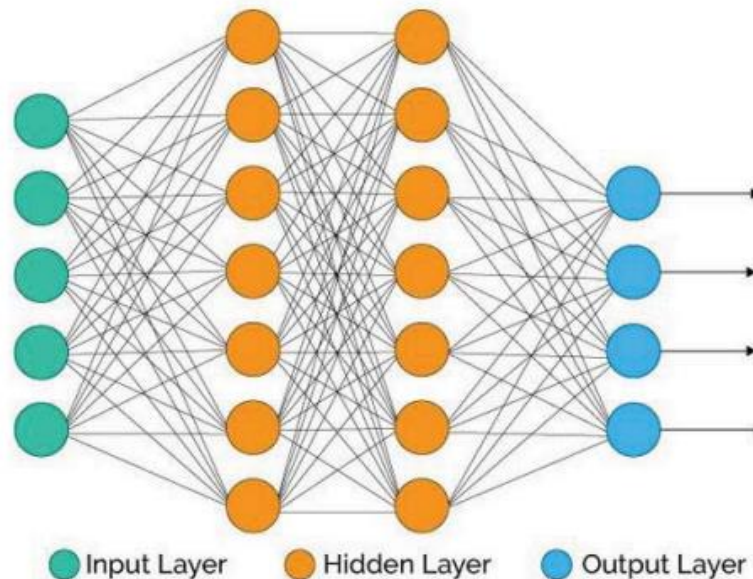
Ketika variable bebas lebih dari 2, nilai konstanta dan koefisien regresi setiap variabel bebas dapat diperoleh dengan menggunakan matriks determinan. Contohnya adalah ketika terdapat 3 persamaan dengan 3 variabel yang tidak diketahui nilainya, yaitu a, b1, b2 dan b3, persamaan tersebut dapat dinyatakan dalam persamaan matriks. (Triyanto et al., 2019)

Berdasarkan teori yang telah di paparkan, algoritma regresi linear adalah algoritma yang dapat memberikan keterkaitan antar variabel dengan menerapkan persamaan-persamaan matematis dengan hasil linear. Regresi linear dapat menerapkan konsep prediksi dengan berdasarkan data-data masa lalu dari sebuah peristiwa tertentu yang kemudian dapat diproses menggunakan persamaan matematis untuk prediksi.

D. Jaringan Syaraf Tiruan (*Artificial Neural Networks*)

Berdasarkan Dartres dan Soori, 2020, jaringan saraf tiruan dirancang dengan cara yang sama seperti otak manusia, dengan *node neuron* yang saling berhubungan dengan cara seperti *web*. *Neuron* adalah miliaran sel yang membentuk otak manusia. Setiap *neuron* terdiri dari tubuh sel yang memproses informasi dengan membawanya ke dan dari otak (*input* dan *output*). Gagasan utama dari jaringan tersebut adalah (sampai batas tertentu) terinspirasi oleh cara kerja sistem saraf biologis, untuk memproses data, dan informasi untuk

mempelajari dan menciptakan pengetahuan. Elemen kunci dari ide ini adalah menciptakan struktur baru untuk sistem pemrosesan informasi. (Dastres & Soori, 2021)



Gambar 1. 1 Pola arsitektur jaringan syaraf tiruan (Artificial Neural Network Systems. International Journal of Imaging and Robotics (IJIR). 2021)

Masih dari Dartres dan Soori (2021), sistem ini terdiri dari sejumlah besar elemen pemrosesan yang sangat saling berhubungan yang disebut neuron yang bekerja sama untuk memecahkan masalah dan mengirimkan informasi melalui sinapsis (koneksi elektromagnetik). *Neuron* saling berhubungan erat dan terorganisir menjadi beberapa lapisan. Lapisan input menerima data, sedangkan lapisan output menghasilkan hasil akhir. Di antara keduanya, satu atau lebih lapisan rahasia biasanya diapit. Pengaturan ini membuat memprediksi atau mengetahui aliran data yang tepat menjadi sulit.

Masih dalam Dartres dan Soori (2021), setiap koneksi memiliki bobot koneksi, dan setiap neuron memiliki nilai ambang batas dan fungsi aktivasi ini

dihitung jika setiap input memiliki bobot positif atau negatif berdasarkan tanda bobot *input*. Berat mempengaruhi intensitas sinyal pada koneksi. *Neuron* yang memiliki ambang batas di atasnya sinyal hanya ditransmisikan jika sinyal agregat melebihi itu. Nilai aktivasi adalah jumlah tertimbang dari unit penjumlahan, dan *output* dihasilkan berdasarkan sinyal dari nilai aktivasi ini.. (Dastres & Soori, 2021)

Terdapat 3 fase dalam pelatihan *Backpropagation*, yaitu fase maju (*feed forward*), fase mundur (*back propagation*), dan fase modifikasi bobot. Dalam fase *feed forward*, pola masukan dihitung maju dimulai dari lapisan input hingga lapisan *output*. Dalam fase *back propagation*, tiap-tiap unit *output* menerima target pola yang berhubungan dengan pola *input* untuk dihitung nilai kesalahan. Kesalahan tersebut akan dipropagasikan mundur, sedangkan fase modifikasi bobot bertujuan untuk menurunkan kesalahan yang terjadi. Ketiga fase tersebut diulang secara terus menerus hingga kondisi penghentian dipenuhi (Damanik et al., 2020).

Masih dalam Damanik, Secara rinci algoritma pelatihan jaringan *Backpropagation* dapat diuraikan sebagai berikut:

- Langkah 0: Inisialisasi bobot-bobot, konstanta laju pelatihan (α), toleransi error atau nilai bobot (bila menggunakan nilai bobot sebagai kondisi berhenti) atau set maksimal epoch (jika menggunakan banyaknya epoch sebagai kondisi berhenti).
- Langkah 1: Selama kondisi berhenti belum dicapai, maka lakukan langkah ke-2 hingga langkah ke-9

- Langkah 2: Untuk setiap pasangan pola pelatihan, lakukan langkah ke-3 sampai langkah ke-8
- Langkah 3: {Tahap I: Umpan maju (*feedforward*)}. Tiap unit masukan menerima sinyal dan meneruskannya ke unit tersembunyi di atasnya.
- Langkah 4: Masing-masing unit di lapisan tersembunyi (dari unit ke-1 hingga unit ke-p) dikalikan dengan bobotnya dan dijumlahkan serta ditambahkan dengan biasnya.
- Langkah 5: Masing-masing unit output ($y_k, k=1,2,3,\dots,m$) dikalikan dengan bobot dan dijumlahkan serta ditambahkan dengan biasnya.
- Langkah 6: {Tahap II: Umpan mundur (*backward propagation*)}. Masing-masing unit *output* ($y_k, k=1,2,3,\dots,m$) menerima pola target t_k sesuai dengan pola masukan/*input* saat pelatihan dan kemudian informasi kesalahan/*error* lapisan *output* (δ_k) dihitung. δ_k dikirim ke lapisan di bawahnya dan digunakan untuk menghitung besarnya koreksi bobot dan bias (ΔW_{jk} dan ΔW_{ok}) antara lapisan tersembunyi dengan lapisan *output*.
- Langkah 7: Pada setiap unit dilapisan tersembunyi (dari unit ke-1 hingga ke-p; $i=1\dots n; k=1\dots m$) dilakukan perhitungan informasi kesalahan lapisan tersembunyi (δ_j). δ_j kemudian digunakan untuk menghitung besar koreksi bobot dan bias (ΔV_{ji} dan ΔV_{jo}) antara lapisan *input* dan lapisan tersembunyi.
- Langkah 8: {Tahap III: Update bobot dan bias}. Masing-masing unit *output*/keluaran ($y_k, k=1,2,3,\dots,m$) dilakukan *update* bias dan bobotnya ($j=0,1,2,\dots,p$) sehingga menghasilkan bobot dan bias baru. Demikian juga

untuk setiap unit tersembunyi mulai dari unit ke-1 sampai dengan unit ke- p dilakukan update bobot dan bias.

- Langkah 9: Uji kondisi berhenti (akhir iterasi).

Berdasarkan teori yang telah dipaparkan, jaringan syaraf tiruan adalah sebuah metode atau algoritma yang meniru proses interaksi saraf manusia kedalam matematis agar dapat ditiru dan di implementasikan ke dalam sebuah program.

E. Data Time Series

Data *time series* merupakan serangkaian pengamatan yang terurut berdasarkan waktu dengan jarak yang sama. Jenis data ini sering ditemui dalam keseharian karena data tersebut dikumpulkan melalui waktu interval yaitu harian, mingguan atau bulanan dari data yang terkumpul dapat dilihat ada suatu pola di dalamnya. Dalam time series pola tersebut dibagi menjadi tiga yaitu pola trend, siklus dan musiman (*seasonal*). Pola musiman (*seasonal*) adalah pola yang mengalami pengulangan yang sama berkali-kali pada interval tertentu. Berdasarkan pembagian daerah (domain), data time series terbagi menjadi dua daerah (*domain*) yaitu daerah waktu (*time domain*) dan daerah frekuensi (*frequency domain*) (Al'afi et al., 2020). Daerah waktu menelaah tentang signifikansi auto korelasi, kestasioneran data, penaksiran parameter model regresi deret waktu dan peramalan (*forecasting*), sedangkan daerah frekuensi (*frequency domain*) menelaah frekuensi tersembunyi pada data musiman yang sulit diperoleh dalam daerah waktu (Aktivani, 2021).

F. Normalisasi dan Denormalisasi

1. Normalisasi Data

Metode normalisasi data adalah proses membuat beberapa variabel memiliki rentang nilai yang sama, tidak ada yang terlalu besar maupun terlalu kecil sehingga dapat membuat analisis statistik menjadi lebih mudah. Adapun metode dalam melakukan normalisasi data yaitu *Min-Max normalization*, *Z-score Normalization*, dan *Decimal Scaling Normalization*. Salah satu metode normalisasi yang sering digunakan dalam melakukan normalisasi data adalah Metode decimal scaling yang merupakan metode transformasi data dengan normalisasi untuk menyamakan rentang nilai pada setiap atribut dengan skala tertentu dengan menggerakkan nilai desimal dari data ke arah yang diinginkan (Kusnaldi et al., 2022). Agar dapat menghasilkan data yang lebih baik. Berdasarkan penelitian terdahulu menyebutkan bahwa perbedaan rentang nilai pada setiap atribut dalam proses transformasi data menyebabkan tidak berfungsinya atribut yang memiliki nilai jauh lebih kecil dibandingkan dengan atribut-atribut lainnya, sehingga data yang telah dinormalisasi dengan metode *decimal scaling* lebih baik hasilnya secara signifikan daripada klasifikasi tanpa melakukan transformasi data (Nasution et al., 2019).

Min-max normalization metode normalisasi yang sering digunakan dalam hal ini keperluan analisis statistik data. metode ini *rescale* data dari suatu *range* ke *range* baru lain. Data di sama ratakan dalam range 0 dan 1. Diberikan nilai yang bersesuaian (dalam satu kolom) $\{sk\}$, $k=1,2,.. n$. Maka nilai normalisasi nya adalah:

$$S' = \frac{s - \min(s_k)}{\max(s_k) - \min(s_k)} \quad (2.9)$$

2. Denormalisasi Data

Denormalisasi merupakan alur/proses yang dilakukan untuk mengembalikan nilai ke nilai asli (awal) dengan tujuan agar mudah dipahami. Denormalisasi merupakan proses pengolahan data yang mengembalikan nilai data asli yang sebelumnya telah di normalisasi pada range tertentu. Proses ini dilakukan untuk mendapatkan nilai aktual prediksi dengan cara membandingkan dengan data sebenarnya. Rumus denormalisasi dalam range [0,1] dinyatakan dalam rumus (Gusti Ayu Rica Ananda,Luh Putu Ida Harini & Wijayakususma, 2023):

$$X_t = x(x_{max} - x_{min}) + x_{min} \quad (2.10)$$

G. Bahasa Pemrograman Python

Python adalah bahasa pemrograman yang memiliki keberagaman luas. Hanya diperlukan alat (*tool*) dan perpustakaan (*library*) yang tepat, lalu kamu bisa menjadi inovator sejati. Memulai belajar bahasa pemrograman membutuhkan nyali, kemauan, waktu, dan mungkin sejumlah minuman berenergi. Oleh karena itu, kamu harus mulai dari menetapkan tujuan dan mempelajari apa saja kegunaan *Python*. *Python* amat sangat mudah untuk dibaca. Sebagai interpreted language (bahasa pemrograman yang tidak perlu dikompilasi), *Python* tidak mengubah kodenya untuk menjadi terbaca oleh komputer. Bahasa ini juga merupakan bahasa pemrograman tujuan umum tingkat tinggi. Para pengembang mendesainnya untuk menjadi bunglon dari dunia pemrograman. Selain itu, *Python* bertujuan untuk

menghasilkan kode yang lebih jelas dan lebih logis tidak hanya untuk proyek skala kecil tetapi juga untuk yang lebih besar. Kamu dapat membandingkan Python dengan kubus Rubik: ia memiliki banyak sisi sehingga kamu dapat memutar dan bermain-main dengannya. Bahasa ini mampu menerapkan banyak prosedur komputer untuk menghasilkan teknologi yang dapat mengejutkanmu. Berikut beberapa fakta menarik yang menggambarkan dampak sebenarnya dari bahasa ini, dan apa saja kegunaan Python:

- a. *BitTorrent* yang terkenal dimulai sebagai sebuah Program *Python*.
- b. *NSA* (Badan Keamanan Nasional) menerapkan *Python* untuk analisis intelijen dan kriptografi.
- c. Para pengembang/*developer* menulis *Youtube* menggunakan *Python* (dan bahasa-bahasa lain).
- d. *Google* juga tidak asing dengan *Python*: perusahaan ini mendasarkan sistem pencarian web nya yang terkenal itu pada bahasa *Python*

Menurut Edwardo, 2018, *Python* adalah *scripting language* yang berorientasi objek. Bahasa pemrograman ini dapat digunakan untuk pengembangan perangkat lunak dan bisa dijalankan melalui berbagai sistem operasi. Saat ini, *Python* juga merupakan bahasa yang populer bagi bidang data *science* dan analisis. Hal ini dikarenakan oleh dukungan bahasa *Python* terhadap *library – library* yang di dalamnya menyediakan fungsi analisis data dan fungsi *machine learning*, data *preprocessing tools*, serta visualisasi data (Edwardo, 2018)

H. Akurasi Prediksi

1. RMSE

Root Mean Square Error atau yang sering disingkat RMSE adalah metrik evaluasi yang sering digunakan untuk mengukur perbedaan atau perbandingan antara nilai prediksi yang dihasilkan oleh suatu model dengan nilai aktual yang diamati. RMSE memiliki fungsi untuk mengumpulkan besarnya kesalahan dalam prediksi pada nilai data menjadi kesatuan keakurasian prediksi. RMSE memiliki akurasi yang tinggi dalam melakukan pengukuran dengan probabilitas hingga 50% (Sulaiman & Juarna, 2021). *Root Mean Square RMSE* merupakan besarnya tingkat kesalahan hasil prediksi, dimana semakin kecil (mendekati 0) nilai RMSE maka hasil prediksi akan semakin akurat. Sedangkan MSE merupakan nilai kuadrat dari RMSE (Rismala et al., 2023). Pada algoritma yang digunakan untuk memprediksi data, berikut persamaan rumus:

$$RMSE = \sqrt{\sum \frac{(Y' - Y)^2}{n}} \quad (2. 11)$$

Keterangan:

Y' = Nilali Prediksi

Y = Nilali Sebenarnya

n = Jumlah Data

2. MAPE

Mean Absolute Percentage Error (MAPE) dihitung dengan menggunakan kesalahan absolut pada tiap periode dibagi dengan nilai observasi yang nyata untuk periode itu. Kemudian, merata-rata kesalahan persentase absolut tersebut.

MAPE merupakan pengukuran kesalahan yang menghitung persentase penyimpangan antara data aktual dengan data peramalan (Krisma Alviani, Azhari Muhammad, 2019). *Mean Absolute Percentage Error* (MAPE) adalah nilai rata – rata perbedaan absolut yang ada diantara nilai dari prediksi dan nilai realisasi yang disebutkan sebagai hasil persenan dari nilai realisasi. Penggunaan *Mean Absolute Percentage Error* (MAPE) pada evaluasi dari hasil peramalan dapat melihat tingkat akurasi terhadap angka peramalan dan angka realisasi (Krisma Alviani, Azhari Muhammad, 2019). Berikut persamaan rumus:

$$MAPE = \sum \left| \frac{y_i - y'_i}{y_i} \right| \times 100 \quad (2.12)$$

Keterangan:

y_i = harga aktual data ke-i

y'_i = harga prediksi data ke-i

rumus persamaan 2.12 di atas akan menghasilkan nilai akurasi yang kecil berdasarkan data nya karena MAPE fokus nya mencari nilai kesalahan yaitu error maka penentuan tingkat keberhasilan prediksi dengan proses pengurangan 100 dari nilai akurasi, berikut rumus nya:

$$Akurasi\ keberhasilan = |nilai\ MAPE - 100| \quad (2.13)$$

I. Penelitian Terkait

Tabel 2. 1 Penelitian Terkait

No	Tahun	Judul	Penelitian Sekarang	Perbedaan
1	2019	Sistem Prediksi Kebutuhan Obat di Puskesmas Menggunakan Metode <i>Least Square</i>	Menggunakan metode Least Square dalam memprediksi produksi obat dengan rata-rata error 12,70%	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
2	2022	Prediksi Harga Minyak Mentah Menggunakan Jaringan Syaraf Tiruan	Nilai variable pelatihan pada alfa 0,5 merupakan nilai MSE yang paling rendah dari hasil model prediksi yang didapatkan dari jumlah lapisan tersembunyi pada 12 neuron dengan nilai momentum yang didapat 0,5 dan epoch yang optimum sebesar 10000 serta error toleransinya sebesar 0,001	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
3	2021	Comparison of Neural Network and Recurrent Neural Network to Predict Rice Productivity in East Java	kedua metode NN dan RNN dapat melakukan prediksi dengan tingkatan error yang rendah dan metode RNN lebih cepat memprediksi dibandingkan NN karena <i>hidden layer</i> RNN lebih sedikit daripada NN. Jumlah <i>node hidden layer</i> dan <i>learning speed</i> sangat signifikan terhadap perhitungan komputasi	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE

4	2021	Improving Tourist Arrival Prediction: A Big Data and Artificial Neural Network Approach	Menggunakan 2 metode untuk perbandingan prediksi yaitu ARIMA dan ANN. Data yang digunakan berdasarkan <i>autoregressive only</i> dan <i>google trends data</i> yang hasilnya <i>autoregressive only</i> lebih akurat. Pengujian menggunakan RMSE	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
5	2019	Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi	Menggunakan regresi linear berganda dengan hasil sebesar 94,51% dan pengujian RMSE 0,432	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
6	2022	Prediksi Penetapan Tarif Penerbangan Menggunakan Auto-Ml Dengan Algoritma Random Forest	Membandingkan akurasi R2 dan Random Forest dengan hasil algoritma akurasi Random Forest dengan RMSE 1,96%	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
7	2020	Stock price prediction using artificial neural network integrated moving average	Evaluasi hasil eksperimen dilakukan dengan menentukan nilai Root Mean Square Error (RMSE). Nilai RMSE terbaik yang diperoleh adalah 0,004 dan nilai RMSE rata-rata adalah 0,0121.	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
8	2020	Forecasting of Covid-19 cases based on prediction using artificial neural network curve fitting technique	menggunakan ANN metode <i>network curve fitting technique</i> dalam memprediksi kasus covid-19 di 5 negara periode 10 bulan.	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
9	2020	Prediction prices of Basrah light oil	periode prediksi 2008-2018 dengan tingkatan	Membandingkan algoritma linear

		using artificial neural networks	akurasi untuk tahun 2019 MRE 0.024 dan MMRE 0.016	regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
10	2021	Pemanfaatan ANN untuk Prediksi Penjualan Online Industri Rumahan selama Pandemi Covid-19	Hasil goal sesuai nilai error toleransi yang sudah ditentukan yaitu 0,02 pada epoch ke 33, dan nilai Mean Absolute Percentage Error 0,33%, artinya tingkat kesalahan jaringan dalam melakukan pelatihan tergolong kecil yaitu 0,33%	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
11	2021	Prediksi jumlah produksi air pdam menggunakan metode ann dengan optimasi pso	Metode Artificial Neural Network dengan optimasi Particle Swarm Optimization secara ilmiah dapat menurunkan nilai RMSE yang sebelumnya dari 4,943 menjadi 3,797.	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
12	2019	Peramalan Jumlah Produksi Air Dengan Algoritma Backpropagation	Pola arsitektur 5-6-1 dengan akurasi 99,97%	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
13	2022	Prediction Of Used Car Prices Using Artificial Neural Networks And Machine Learning	Hasil eksperimen menunjukkan bahwa model Random Forest dengan nilai Mean Absolute Error sebesar 1,0970472 dan nilai error R2 sebesar 0,772584 telah memberikan error yang lebih sedikit diantara semua algoritma lainnya.	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE

14	2020	Real Estate Price Range Prediction Using Artificial Neural Network And Grey Wolf Optimizer	Algoritma ANN-GWO menggunakan MSE memiliki tingkat akurasi prediksi 98,7951%	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE
15	2022	Pork Price Prediction Using Bi-RNN-LSTM Artificial Neural Network	Kesalahan prediksi yang dicapai pada kumpulan data kami untuk Mean Square Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) dan Symmetric Mean Absolute Percentage Error (SMAPE) masing-masing adalah 0,48, 0,69, 0,53, 3,37%, 3,37%.	Membandingkan algoritma linear regresi berganda dan jaringan saraf tiruan backpropagation dengan akurasi prediksi MAPE

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Kesimpulan dari penelitian ini, membangun model linear regresi berganda dalam prediksi dengan mencari nilai koefisien yaitu nilai b_1 , b_2 , b_3 dan b_4 setelah itu mencari nilai intercept atau nilai a sehingga dapat dibuat model persamaan linear regresi berganda dalam proses prediksi yaitu $y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$. Hasil akurasi prediksi yang diperoleh dari model yang dibuat adalah 97.70%. Model JST dibangun dengan memisahkan data latih dan data uji sebagai data pelatihan dan data prediksi ujicoba. Proses pelatihan data menjadi faktor yang paling utama dalam prediksi JST karena di dalam terdapat parameter yang menjadikan pelatihan data lebih mendalam. Hasil rata-rata akurasi prediksi seluruh spesies hewan menggunakan algoritma linear regresi adalah 95.26% sedangkan menggunakan JST atau jaringan saraf tiruan *backpropagation* adalah 96.50%.

B. Saran

Algoritma yang digunakan masih dapat di optimalkan dengan berbagai metode algoritma yang dapat memberikan tingkat akurasi lebih akurat dari penelitian ini. Penelitian selanjutnya disarankan menggunakan data langsung dari sumber pemerintah yang memiliki dataset dan wewenang dalam pengolahan informasi harga ikan dalam bentuk *time series* per bulan, per minggu atau per hari.

DAFTAR PUSTAKA

- Aktivani, S. (2021). Uji Stasioneritas Data Inflasi Kota Padang Periode 2014-2019. *Jurnal Statistika Industri Dan Kompetensi*, 6(1), 26–33.
- Al’afi, A. M., Widiart, W., Kurniasari, D., & Usman, M. (2020). Peramalan Data Time Series Seasonal Menggunakan Metode Analisis Spektral. *Jurnal Siger Matematika*, 1(1), 10–15. <https://doi.org/10.23960/jsm.v1i1.2484>
- Arhami, M., & Nasir, M. (2020). *DATA MINING Algoritma dan Implementasi*. Penerbit Andi.
- Bhirawa, W. T. (2020). Proses Pengolahan Data Dari Model Persamaan Regresi Dengan Menggunakan Statistical Product and Service Solution (SPSS). *Statistika*, 71–83.
- Damanik, S. M., Lubis, M. R., Saputra, W., & Parlina, I. (2020). Analisis Jaringan Saraf Tiruan Metode Backpropagation Dalam Memprediksi Ekspor Menurut Kelompok Barang Ekonomi Di Provinsi Sumatera Utara. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 4, 403–412. <https://doi.org/10.30865/komik.v4i1.2729>
- Daniya, T., Geetha, M., Santhosh Kumar, B., & Cristin, R. (2020). Least square estimation of parameters for linear regression. *International Journal of Control and Automation*, 13(2), 447–452.
- Dastres, R., & Soori, M. (2021). Artificial Neural Network Systems. *International Journal of Imaging and Robotics (IJIR)*, 2021(2), 13–25. www.ceserp.com/cp-jour
- Ginantra, N. L. W. S. R. D. (2021). *Data Mining dan Penerapan Algoritma*. Yayasan Kita Menulis.
- Gusti Ayu Rica Ananda, Luh Putu Ida Harini, I. G. L., & Wijayakusuma. (2023). *Klasifikasi Calon Nasabah Debitur Ksp. Samudra Harta Dengan Recurrent Neural Network*. 8(4).
- Hauriza, B., Muladi, M., & Wirawan, I. M. (2021). Prediksi Tingkat Inflasi Bulanan Indonesia Menggunakan Metode Jaringan Saraf Tiruan. *Jurnal Teknologi Dan Informasi*, 11(2), 152–167.

<https://doi.org/10.34010/jati.v11i2.4924>

- Irawati, I., & Syamsuddin, M. (2020). Strategi Pengembangan Usaha Budidaya Ikan Kerapu Sistem Keramba Jaring Apung Di Teluk Ambon Dalam. *Jurnal Akuakultur Sungai Dan Danau*, 5(2), 30. <https://doi.org/10.33087/akuakultur.v5i2.66>
- Krisma Alviani, Azhari Muhammad, W. P. P. (2019). Perbandingan Metode Double Exponential Smoothing Dan Triple Exponential Smoothing Dalam Parameter Tingkat Error Mean Absolute Percentage Error (MAPE) dan Means Absolute Deviation (MAD) Alviani Krisma Putut Pamilih Widagdo Kata kunci-forecasting, Double Ex. *Prosiding Seminar Nasional Ilmu Komputer Dan Teknologi Informasi*, 4(2), 81–87.
- Kusnaldi, M. R., Gulo, T., & Aripin, S. (2022). Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal. *Journal of Computer System and Informatics (JoSYC)*, 3(4), 330–338. <https://doi.org/10.47065/josyc.v3i4.2112>
- Kwok, E., & Susanti, W. (2019). Penerapan Metode Regresi Linier dalam Aplikasi Sistem Peramalan Jumlah Bahan Baku untuk Produksi Tahu. *Mahasiswa Aplikasi Teknologi Komputer Dan Informasi*, 1(2), 1–8.
- Loban, J. M. (2023). *ANALISIS REGRESI FAKTOR-FAKTOR YANG MEMPENGARUHI*. 8(1), 138–145.
- Muttaqin, A., & Ismail, I. (2023). Kemiskinan Kultural Kemiskinan Kultural Masyarakat Nelayan di Desa Panipahan Kecamatan Pasir Limau Kapas. *Jurnal Sosiologi Agama Indonesia (JSAI)*, 4(2), 279–292. <https://doi.org/10.22373/jsai.v4i2.2815>
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN. *Computer Engineering, Science and System Journal*, 4(1), 78. <https://doi.org/10.24114/cess.v4i1.11458>
- Pradito, B., & Purnia, D. S. (2022). Komparasi Algoritma Linear Regression dan Neural Network Untuk Memprediksi Nilai Kurs Mata Uang. *EVOLUSI*:

- Jurnal Sains Dan Manajemen*, 10(2), 64–71.
<https://doi.org/10.31294/evolusi.v10i2.13284>
- Puteri, K., & Silvanie, A. (2020). Machine Learning untuk Model Prediksi Harga Sembako. *Jurnal Nasional Informatika*, 1(2), 82–94.
- Rismala, Ali, I., & Rizki Rinaldi, A. (2023). Penerapan Metode K-Nearest Neighbor Untuk Prediksi. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 585–590.
- Sepriano, S.Sos., M.Kom., CBPA, D. (2023). *Pengantar Ekonomi Bisnis*. PT. Sonpedia Publishing Indonesia.
- Sulaiman, A., & Juarna, A. (2021). Peramalan Tingkat Pengangguran Di Indonesia Menggunakan Metode Time Series Dengan Model Arima Dan Holt-Winters. *Jurnal Ilmiah Informatika Komputer*, 26(1), 13–28.
<https://doi.org/10.35760/ik.2021.v26i1.3512>
- Trisna, P., & Permana, H. (2019). PERBANDINGAN PERAMALAN INFLASI KOTA DENPASAR MENGGUNAKAN METODE FUZZY TIME SERIES DAN MULTILAYER PERCEPTRON. *Jurnal Teknologi Informasi Dan Komputer*, 5(3), 285–296.
- Triyanto, E., Sismoro, H., & Laksito, A. D. (2019). Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 4(2), 66–75.
<https://doi.org/10.36341/rabit.v4i2.666>
- Yuliyanto, M. R., Wuryandari, T., & Utami, I. T. (2023). Peramalan Pendapatan Bulanan Menggunakan Fuzzy Time Series Chen Orde Tinggi. *Jurnal Gaussian*, 12(1), 61–70. <https://doi.org/10.14710/j.gauss.12.1.61-70>
- Zai, C. (2022). Implementasi Data Mining Sebagai Pengolahan Data. *Portal Data*, 2(3), 1–12.