

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Penelitian mengenai penggunaan teknologi AI untuk memprediksi potensi kebakaran hutan dan lahan sudah banyak dilakukan. Pada penelitian sebelumnya, metode yang digunakan berbeda-beda sehingga menghasilkan nilai akurasi yang berbeda-beda pula. Namun, pada penelitian-penelitian yang sudah dilakukan sebelumnya menunjukkan bahwa penggunaan AI dapat membantu memprediksi potensi terjadi peristiwa kebakaran hutan dan lahan. Berikut adalah penelitian terdahulu yang menurut penulis mempunyai keterkaitan pada penelitian yang akan dijalankan oleh penulis.

Pertama, sebuah penelitian yang berjudul “Prediksi Potensi Kebakaran Hutan dengan Algoritma Klasifikasi C4.5 Studi Kasus Provinsi Kalimantan Barat” yang dilakukan oleh Aji Primajaya, Betha Nurina Sari dan Ahmad Khusaeri pada tahun 2020. Penelitian tersebut menggunakan algoritma klasifikasi C4.5 pada metode KDD (Knowledge Discovery in Database). Data yang berupa *hotspot* kebakaran, jarak *hotspot* ke pusat kota, jarak *hotspot* terhadap sungai terdekat serta data lainnya yaitu kecepatan rata-rata angin, suhu dan curah hujan pada setiap titik *hotspot*. Pengolahan data dilakukan dengan algoritma C4.5 pada tahap data *mining* menghasilkan model klasifikasi berupa pohon keputusan. Model terbaik didapatkan pada pembagian data 80:20 yaitu 80% untuk data *training* dan 20% untuk data testing dengan akurasi 89,78% [19].

Kedua, penelitian yang berjudul “Analisis Prediksi Harga Saham Sektor Perbankan Menggunakan Algoritma Long-Short Terms Memory (LSTM)” dilakukan oleh Prisma Aji Riyantoko, Tresna Maulana Fahrudin, Kartika Maulida Hindrayani dan Eristya Maya Safitri pada tahun 2020. Penelitian dilakukan dengan membandingkan tiga model optimasi (SGD, ADAM dan RMSprop) menggunakan LSTM pada empat *dataset* saham (BRI, BNI, BTN,

Mandiri). Dari penelitian ini didapatkan hasil bahwa optimasi Adam menunjukkan nilai RMSE yang stabil turun (akurasi stabil naik), sehingga dapat disimpulkan bahwa optimasi Adam menghasilkan nilai prediksi yang paling baik [20].

Ketiga, penelitian yang berjudul “Prediksi Data Time Series Saham Bank BRI dengan Mesin Belajar LSTM (Long-Short Term Memory)” dilakukan oleh Adhito Satyo Bayangkari Karno pada tahun 2020. Penelitian tersebut dilakukan untuk memprediksi data dengan *timestep* 7 menggunakan arsitektur LSTM dengan 4 *hidden layer* dimana tiap *layer* berisi 5 *node*, *dropout* 0,2 dan 1 *output layer* dengan 1 *node*. Hasil penelitian tersebut didapatkan hasil akurasi yang cukup tinggi dengan nilai RMSE (Root Mean Square Error) sebesar 227,47 [21].

Keempat, penelitian berjudul “Prediksi Lokasi Titik Panas Kebakaran Hutan Menggunakan Model Regresion SVM (Support Vector Machine) pada Data Kebakaran Hutan Daops Manggala Agni Oki Provinsi Sumatera Selatan Tahun 2019” oleh Jepri Yandi, Tri Basuki Kurniawan, Edi Surya Negara dan Muhamad Akbar pada tahun 2021. Penelitian tersebut menggunakan data berupa *hotspot* kebakaran hutan Sumatera Selatan selama tahun 2019 yang didapatkan jumlah *hotspot* sebanyak 16.379 titik disertai data atribut berupa suhu dan kelembapan udara. Data tersebut diolah dengan metode SVM (Support Vector Machine) dan hasilnya dipetakan dalam bentuk *heatmaps*. Hasil penelitian tersebut memperoleh nilai RSME 2.1 dan nilai R^2 0.83 [22].

Kelima, penelitian berjudul “Deep Learning Approach to Predict Forest Fires Using Meteorological Measurements” yang dilakukan oleh Naaman Omar, Adel Al-zebari dan Abdulkadir Sengur pada tahun 2021 untuk memprediksi kebakaran hutan. Penelitian tersebut menggunakan arsitektur Deep Learning berupa layer LSTM (Long-Short Term Memory) dengan dua *fully connected layers*, satu *dropout layer* dan satu *regression layer* sebagai *output*. Sedangkan data yang digunakan didapat melalui Fire Weather Index (FWI) Kanada yang memiliki 12 *features* dan 536 *instances*. Evaluasi dilakukan menggunakan Root Mean Squared Error (RMSE) dan sebagai perbandingannya dilakukan

prediksi pada beberapa teknik pembelajaran mesin (Machine Learning) tradisional atau sederhana. Hasil dari penelitian didapatkan bahwa prediksi dengan metode Deep Learning mendapatkan hasil *error* paling kecil pada nilai 0,021 sedangkan metode lain yaitu Decision Tree mendapatkan nilai *error* sebesar 0,274, regresi linear sebesar 0,025, Support Vector Machines (SVM) sebesar 0,023, dan Neural Network sebesar 0,023 [17].

Keenam, penelitian berjudul “Intelligent computing based forecasting of deforestation using fire alerts: A deep learning approach” oleh Muhammad Ali Jamshed, Charalambos Theodorou, Tahera Kalsoom, Nadeem Anjum, Qammer H. Abbasi dan Masood Ur-Rehman pada tahun 2022. Penelitian tersebut menggunakan LSTM untuk memprediksi area deforestasi berdasarkan data riwayat kebakaran hutan. Model yang digunakan di-*train* sebanyak 150 *epochs* menggunakan beberapa konfigurasi layer dan ReLU sebagai *activation function*. Nilai *error* terkecil didapat pada konfigurasi 3 *hidden layers*, 2 *dense layers* dan mengatur nilai *dropout* pada 0,05 serta *learning rates* 0,01. Pada konfigurasi tersebut didapat nilai RMSE $4,051 \times 10^{-5}$ dengan akurasi sebesar 95% [18].

Ketujuh, penelitian yang berjudul “Prediksi Harga Saham Twitter dengan Long-Short Term Memory Recurrent Neural Network” dilakukan oleh Ibnu Akil dan Indra Chaidir pada tahun 2022. Penelitian dilakukan untuk membuat pemodelan Machine Learning yang dapat memprediksi naik turunnya harga saham dengan mengoptimalkan model menggunakan Recurrent Neural Network (RNN). Dalam prosesnya, konfigurasi model dilakukan sebanyak dua kali dengan konfigurasi yang pertama berupa *single step* dan kedua adalah *multiple step*. Hasil yang didapat dari konfigurasi pertama dengan menghitung nilai *metrics* menggunakan Mean Absolute Error (MEA) sebesar 0,9300 (*open*) dan 0,8042 (*close*), sedangkan pada konfigurasi model kedua didapatkan hasil sebesar 2,0540 (*open*) dan 2,0668 (*close*). Dari kedua konfigurasi model tersebut dapat disimpulkan bahwa model LSTM *single step* memiliki akurasi yang lebih baik, sedangkan pada model LSTM *multiple step*, model kehilangan dinamisme untuk melakukan prediksi nilai yang bersifat naik turun [23].

Tabel 2.1 Ringkasan Penelitian Sebelumnya

No.	Judul	<i>Comparing</i>	<i>Contrasting</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summary</i>
1	Prediksi Potensi Kebakaran Hutan dengan Algoritma Klasifikasi C4.5 Studi Kasus Provinsi Kalimantan Barat (2020)	Menggunakan algoritma klasifikasi C4.5 untuk Forecasting atau melakukan prediksi.	Menganalisis penerapan yang dilakukan untuk memprediksi kebakaran hutan di Provinsi Kalimantan Barat dengan algoritma C4.5.	Penelitian dilakukan untuk mengetahui hasil evaluasi dari kualitas model yang dihasilkan pada algoritma C4.5	Melakukan konfigurasi <i>percentage split</i> pada <i>dataset</i> untuk membandingkan nilai akurasi.	Proses prediksi menggunakan algoritma klasifikasi C4.5 dengan konfigurasi terbaiknya, yaitu <i>split</i> 80% data <i>training</i> dan 20% data <i>testing</i> mendapatkan nilai akurasi 89,7859%.
2	Analisis Prediksi Harga Saham Sektor Perbankan Menggunakan Algoritma Long-Short Terms Memory (LSTM) (2020)	Menggunakan algoritma LSTM untuk memprediksi beberapa saham sektor perbankan di Indonesia.	Membandingkan tiga model optimasi SGD, ADAM dan RMSprop pada algoritma LSTM.	Penelitian bertujuan untuk mendapatkan akurasi terbaik pada tiga model optimasi yang digunakan dengan algoritma LSTM.	Penelitian ini menggunakan empat <i>dataset</i> saham BRI, BNI, BTN dan Mandiri dari tahun 2015 hingga 2019. Masing-masing <i>dataset</i> dibagi menjadi 80% untuk <i>train</i> dan 20% untuk <i>test</i>	Hasil penelitian ini didapatkan bahwa optimasi Adam pada algoritma LSTM menunjukkan nilai RMSE yang stabil turun (akurasi stabil naik), sehingga dapat dikatakan memiliki akurasi

No.	Judul	<i>Comparing</i>	<i>Contrasting</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summary</i>
					kemudian diproses sebanyak tiga kali dengan tiga model optimasi yang berbeda.	paling baik dibandingkan kedua optimasi lainnya.
3	Prediksi Data Time Series Saham Bank BRI dengan Mesin Belajar LSTM (Long-Short Term Memory) (2020)	Menggunakan LSTM untuk memprediksi saham bank BRI.	Melakukan prediksi Time Series dengan LSTM pada saham bank BRI.	Penelitian dilakukan untuk menguji akurasi pada prediksi saham bank BRI dengan LSTM.	Pada penelitian digunakan nilai <i>timestamp</i> 7 pada <i>dataset</i> , kemudian model yang dibuat tersusun dari empat <i>hidden layer</i> LSTM dan <i>dropout</i> pada masing-masing layer sebesar 0,2.	Penelitian dengan menggunakan algoritma LSTM didapatkan hasil akurasi yang cukup tinggi dengan nilai RMSE (Root Mean Square Error) sebesar 227,47
4	Prediksi Lokasi Titik Panas Kebakaran Hutan Menggunakan Model Regresion SVM (Support Vector Machine) pada Data Kebakaran Hutan	Menggunakan SVM (Support Vector Machine) untuk melakukan prediksi.	Menganalisis penerapan metode SVM dalam melakukan prediksi <i>hotspot</i> atau titik panas kebakaran hutan.	Penelitian dilakukan untuk menguji akurasi pada algoritma SVM dalam melakukan prediksi.	Penelitian menggunakan data <i>hotspot</i> kebakaran sebagai fitur dan data suhu serta kelembapan udara sebagai atribut yang hasilnya	Hasil dari penelitian terbilang cukup baik dengan memperoleh nilai RMSE 2,1 dan R^2 0,83 pada model SVM yang dibuat.

No.	Judul	<i>Comparing</i>	<i>Contrasting</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summary</i>
	Daops Manggala Agni Oki Provinsi Sumatera Selatan Tahun 2019 (2021)				dipetakan dalam bentuk <i>heatmaps</i> .	
5	Deep Learning Approach to Predict Forest Fires Using Meteorological Measurements (2021)	Menggunakan arsitektur Deep Learning LSTM untuk melakukan prediksi.	Penerapan arsitektur Deep Learning berupa layer LSTM dengan dua <i>fully connected layers</i> , satu <i>dropout layer</i> dan satu <i>regression layer</i> sebagai <i>output</i> .	Penelitian dilakukan untuk membandingkan arsitektur Deep Learning dengan metode Machine Learning sederhana.	Melakukan perbandingan nilai RMSE pada algoritma Deep Learning, Decision Tree, Linear Regression, Support Vector Machines (SVM) Dan Neural Network.	Hasil dari penelitian menunjukkan bahwa metode Deep Learning merupakan metode yang paling baik dengan nilai <i>error</i> paling kecil dibandingkan beberapa metode Machine Learning lainnya.
6	Intelligent Computing Based Forecasting of Deforestation Using Fire Alerts: A Deep Learning Approach (2022)	Menggunakan metode LSTM untuk melakukan prediksi.	Pemilihan <i>hyperparameter</i> pada model LSTM untuk memprediksi area deforesatsi..	Penelitian dilakukan untuk menganalisis pemilihan <i>hyperparamter</i> pada model Deep	Melakukan <i>tuning hyperparameter</i> yang sesuai pada model LSTM kemudian dievaluasi	Hasil dari penelitian didapatkan model LSTM untuk memprediksi area deforestasi

No.	Judul	<i>Comparing</i>	<i>Contrasting</i>	<i>Criticize</i>	<i>Synthesize</i>	<i>Summary</i>
				Learning berbasis LSTM.	menggunakan RMSE.	dengan akurasi 95%.
7	Prediksi Harga Saham Twitter dengan Long-Short Term Memory Recurrent Neural Network (2022)	Menggunakan Algoritma Long-Short Term Memory (LSTM) pada Recurrent Neural Network (RNN) untuk melakukan prediksi.	Membandingkan penggunaan model regresi (RNN) <i>single</i> dan <i>multiple step</i> dalam melakukan prediksi.	Pada penelitian ini untuk menguji akurasi model RNN dengan layer LSTM pada regresi <i>single step</i> dan <i>multiple step</i> .	Pada penelitian ini <i>dataset</i> dibagi menjadi tiga, 70% untuk <i>training</i> , 20% untuk validasi dan 10% untuk <i>test</i> . Kemudian data tersebut diproses dua kali untuk mendapatkan hasil pada konfigurasi <i>single step</i> dan konfigurasi <i>multiple step</i> .	Model LSTM dengan RNN <i>single step</i> memiliki akurasi yang lebih baik, sedangkan pada model LSTM dengan RNN <i>multiple step</i> , model tampak kehilangan dinamisme dalam memprediksi nilai yang bersifat turun naik.

Berdasarkan beberapa penelitian terdahulu, menunjukkan bahwa penggunaan AI dapat membantu memprediksi potensi timbulnya kebakaran hutan dan lahan. Pemodelan regresi linear menggunakan metode LSTM dinilai memiliki hasil yang bagus untuk memprediksi data sekuensial seperti pada pola kejadian kebakaran. Pola kebakaran disertai data seperti kecepatan angin, suhu, curah hujan digunakan pada penelitian sebelumnya. Pada penelitian yang akan dilangsungkan oleh penulis, data yang digunakan difokuskan pada Provinsi Kalimantan Barat. Dan data yang akan digunakan oleh penulis diperoleh dari citra satelit MODIS selama rentang tahun 2019 hingga 2023.

2.2 Landasan Teori

2.2.1 Kebakaran hutan

Kebakaran hutan bisa diartikan sebagai kondisi ketika hutan dilanda api, menyebabkan kerusakan hutan dan atau hasilnya dapat mengakibatkan kerugian ekonomi atau nilai lingkungan. berdasarkan Peraturan Menteri Kehutanan Nomor: P.12/Menhut-II/2009 tentang Pengendalian Kebakaran Hutan. Sementara itu, kebakaran dapat dibedakan pengertiannya pada kebakaran hutan dan kebakaran lahan berdasarkan lokasinya, dimana kebakaran hutan merupakan peristiwa yang terjadi pada lingkup hutan dan kebakaran lahan terjadi di luar lingkup hutan. Penyebab terjadinya kebakaran hutan dan lahan dapat disebabkan oleh pembakaran yang tidak terkendali karena proses alami (faktor alamiah) atau faktor kesengajaan oleh manusia [24]. Faktor alamiah yang sering menyebabkan terjadinya kebakaran adalah faktor musim kemarau panjang yang berakibat pada kekeringan. Sedangkan faktor kesengajaan seperti, pengosongan lahan dengan cara dibakar dan tindakan membuang puntung rokok secara sembarangan di area sekitar hutan [11].

2.2.1.1 Daerah rawan kebakaran

Daerah rawan kebakaran adalah daerah yang memiliki potensi tinggi untuk terbakar akibat faktor alam, manusia atau keduanya. Daerah rawan kebakaran dapat ditentukan berdasarkan beberapa parameter, seperti aktivitas manusia, keadaan tutupan lahan, kondisi cuaca, dan jenis tanah.

Tingkat kerawanan kebakaran dapat dibagi menjadi beberapa kelas, seperti rendah, sedang, tinggi, atau sangat tinggi, tergantung pada nilai rawan yang diperoleh dari analisis parameter tersebut [25].

2.2.1.2 Dampak kebakaran hutan

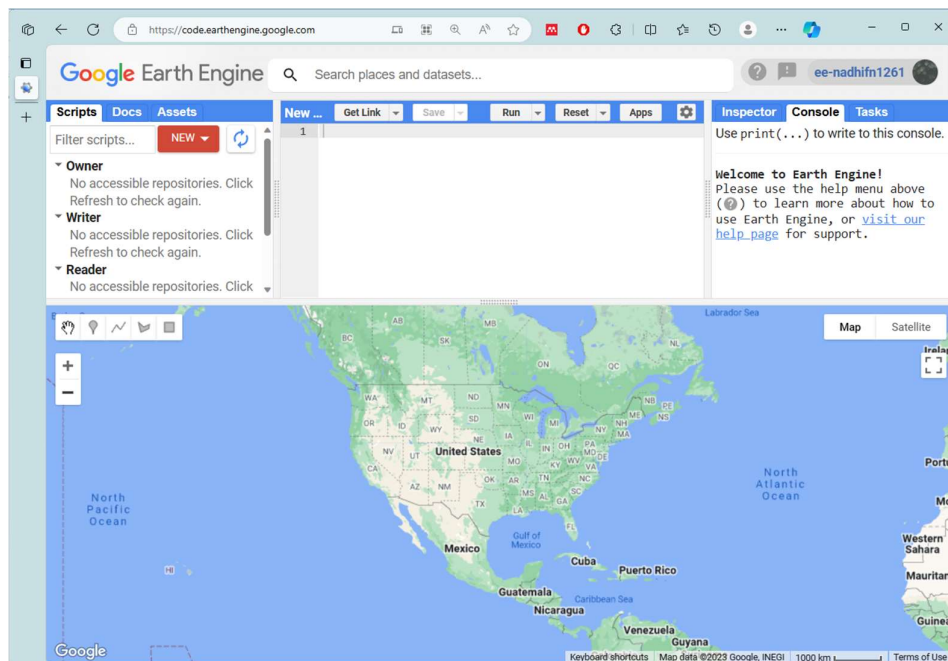
Kebakaran yang terjadi tentunya berdampak pada berbagai aspek kehidupan. Dampak yang diakibatkan pada peristiwa kebakaran hutan dan lahan yaitu sebagai berikut [11]:

1. Dampak pada aspek kesehatan adalah salah satu dampak yang terlihat langsung dengan jelas. Asap yang dihasilkan dari peristiwa kebakaran dapat berakibat pada masalah kesehatan, terutama infeksi saluran pernapasan akut (ISPA). Selain itu, kualitas udara juga dapat dipengaruhi oleh terjadinya peristiwa kebakaran.
2. Kerugian dalam aspek ekonomi juga dapat dirasakan akibat peristiwa kebakaran hutan dan lahan. Dengan adanya faktor alam musim kemarau, menjadikan kebakaran menjadi salah satu peristiwa yang rutin terjadi setiap tahunnya. Hal tersebut tentunya akan mempengaruhi kondisi alam terutama hutan dan lahan dapat mempengaruhi bahkan merugikan pertumbuhan ekonomi jika tidak segera dihentikan. Produk hutan serta ekowisata yang terdampak akan mempengaruhi devisa negara.
3. Aspek sosial budaya, aspek tersebut adalah salah satu hal yang tidak terlihat secara langsung, namun menimbulkan dampak yang dirasakan bagi masyarakat sekitar terjadinya peristiwa kebakaran hutan dan lahan. Terdampak dan hilangnya atau rusaknya hutan sebagai identitas masyarakat adat menjadikan kerugian dalam aspek sosial serta budaya bagi masyarakat tersebut.
4. Kerusakan lingkungan hidup juga menjadi dampak yang disebabkan oleh kebakaran hutan dan lahan. Kerusakan yang diakibatkan oleh kebakaran mempengaruhi bermacam-macam flora serta fauna pada lingkungan hidup hutan yang terdampak, seperti pada binatang,

tumbuhan, buah-buahan, kayu dan lain sebagainya yang ada di lingkungan tersebut.

2.2.2 Google Earth Engine (GEE)

Google Earth Engine disingkat GEE merupakan platform dari Google yang mampu melakukan pengolahan informasi *geospasial*. GEE dapat melakukan pengolahan data yang berbentuk citra penginderaan jauh untuk mengekstraksi informasi berbasis spasial [26]. Pengolahan data melalui platform GEE dapat dilakukan melalui aplikasi web melalui code editor yang disediakan.



Gambar 2.1 Tampilan aplikasi *code editor* GEE berbasis web

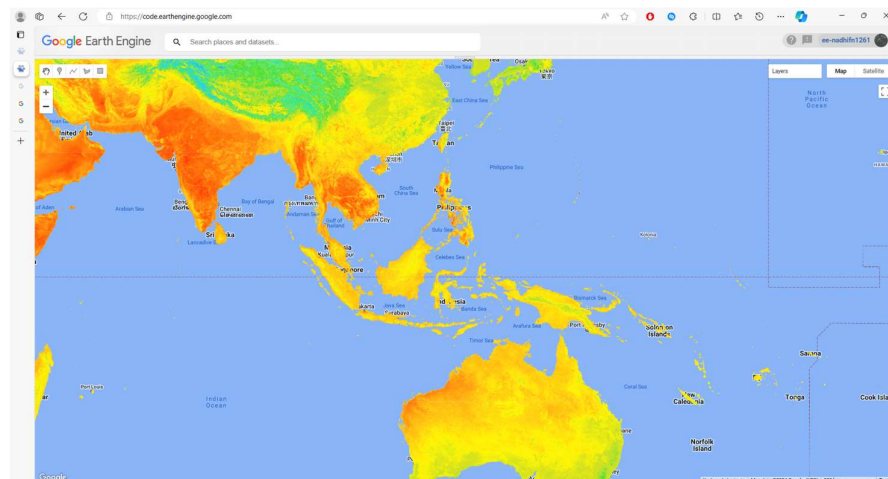
Gambar yang ditunjukkan pada Gambar 2.1 merupakan tampilan platform aplikasi GEE berbasis web. Aplikasi tersebut memiliki *code editor* yang terintegrasi dengan API menggunakan bahasa pemrograman JavaScript. Selain menggunakan aplikasi web, GEE juga dapat diakses melalui API yang tersedia pada bahasa pemrograman Python. API tersebut menyediakan *dataset* dari penyedia data citra satelit.

Data yang disediakan melalui GEE bersifat *open access* dari penyedia citra satelit utama. Penggunaan GEE mempermudah dalam proses

pengambilan data tanpa perlu mengunduh data dari penyedia utama [26]. *Dataset* yang disediakan merupakan data yang sudah diolah dari penyedia citra satelit.

2.2.2.1 MODIS (Moderate Resolution Imaging Spectroradiometer)

Moderate Resolution Imaging Spectroradiometer, biasa disingkat MODIS merupakan salah satu instrumen atau sensor yang digunakan untuk mengambil citra yang ada dalam satelit Terra dan Aqua yang dioperasikan oleh NASA. MODIS dapat menangkap citra yang mencakup berbagai panjang gelombang, mulai dari ultraviolet hingga inframerah. Data citra yang dihasilkan dari MODIS tersebut banyak digunakan dalam pemantauan lingkungan, pertanian, bencana alam dan iklim global. Salah satu contohnya penggunaannya dalam pemantauan iklim global adalah dengan data Land Surface Temperature (LST) atau suhu permukaan bumi. Gambar 2.2 dibawah menunjukkan contoh data suhu permukaan bumi dari MODIS yang ditampilkan pada GEE.



Gambar 2.2 Contoh data LST dari MODIS yang ditampilkan pada GEE

2.2.3 Python

Python adalah sebuah *interpreted language*, yaitu bahasa pemrograman yang dapat dieksekusi tanpa memerlukan proses kompilasi (menggunakan interpreter). Interpreter pada bahasa pemrograman Python dapat digunakan secara interaktif sehingga membuatnya mudah untuk melakukan eksperimen

atau pengujian selama pengembangan program [27]. Pada awalnya Python digunakan untuk memanipulasi data bertipe *string*, namun sejalan dengan waktu Python juga masuk ke ranah pengolahan data untuk analisis. Python memiliki daya tarik sendiri karna memiliki banyak *library* tambahan untuk manipulasi dan visualisasi data [28].

2.2.4 API (Application Programming Interface)

API (Application Programming Interface) adalah antarmuka yang dibuat oleh pengembang suatu sistem agar sebagian atau seluruh fungsi sistem tersebut dapat diakses secara programatik (melalui program). API memungkinkan komunikasi antara dua komponen perangkat lunak yang berbeda. Komunikasi tersebut dilakukan berdasarkan program atau instruksi tertentu yang ditulis dalam bahasa pemrograman tertentu [29].

2.2.5 Folium

Folium merupakan sebuah *library* pada bahasa pemrograman Python yang berbasis pada leaflet.js yaitu *library* JavaScript yang memungkinkan untuk memanipulasi data dalam melakukan pemetaan [30]. Folium memungkinkan untuk membuat peta interaktif pada aplikasi web. Dengan menggunakan Folium dapat memanipulasi data geografis pada Python, seperti menambahkan penanda (*marker*), garis (*line*), poligon, serta informasi lainnya ke dalam peta interaktif.

2.2.6 GitHub Pages

GitHub Pages merupakan layanan *hosting* web dari GitHub yang memungkinkan pengguna untuk mempublikasikan situs atau aplikasi web. Layanan ini terintegrasi dengan GitHub yang memungkinkan pengguna untuk menyimpan data aplikasi web secara *online*. Dengan menggunakan layanan GitHub Pages, data aplikasi web yang tersimpan pada GitHub dapat dipublikasikan dalam situs web secara gratis. Penggunaan GitHub Pages memudahkan pengembang dalam penyebaran aplikasi web [31].

2.2.7 Time Series

Time Series merupakan sekumpulan data statistik yang berdasarkan dari pengamatan pada setiap interval waktu tertentu atau disebut sebagai deret

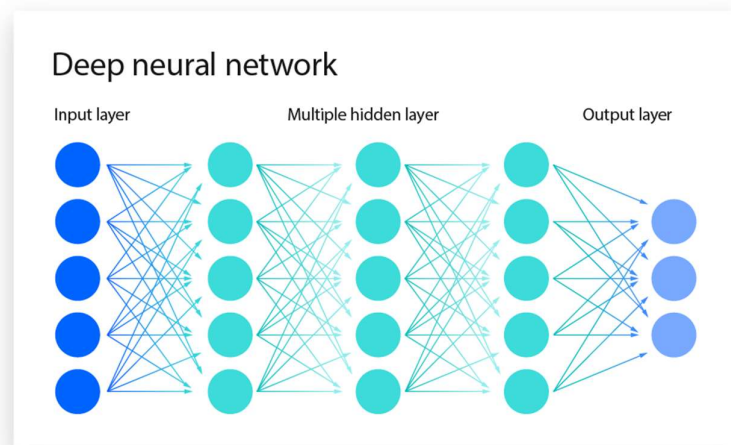
Gambar 2.3 Ilustrasi Sliding Window [35]

Gambar 2.3 menunjukkan ilustrasi penggunaan Sliding Window. Penggunaan metode Sliding Window memungkinkan penerapan algoritma Classical Supervised Learning pada data sekuensial. Sehingga, data *time series* dapat dilakukan prediksi pada algoritma regresi linear dengan menentukan *input window* sebagai atribut data dan *forecast horizon* sebagai fitur prediksi. Metode tersebut diperlukan pada penelitian yang akan dilakukan, sebagaimana *dataset* yang dipergunakan berupa data sekuensial.

Pada penelitian yang akan penulis lakukan, metode Sliding Window diaplikasikan pada *dataset*. *Dataset* sebelumnya diurutkan berdasarkan waktu pada tiap-tiap *point* atau titik koordinat. Kemudian Sliding Window dilakukan pada tiap-tiap *point* dengan urutan waktunya. *Input window* diberikan panjang sebanyak 3, sedangkan *forecast horizon* sepanjang 1.

2.2.8 Deep Learning

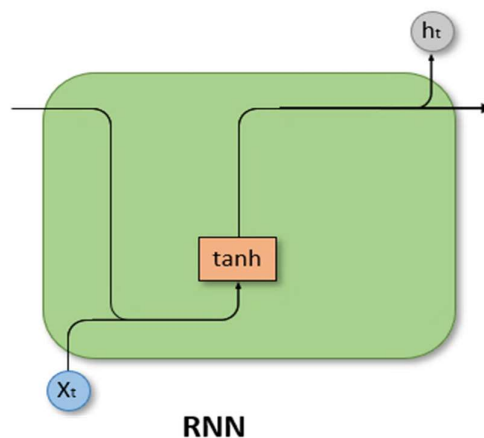
Deep Learning adalah sebuah pendekatan terhadap Machine Learning dengan menerapkan cara kerja otak manusia, ilmu statistik dan matematika terapan. Tujuan dari Deep Learning adalah untuk meningkatkan kemampuan Machine Learning dalam memecahkan masalah kecerdasan buatan. Konsep dasar dari Deep Learning yaitu dengan menirukan cara kerja neuron pada otak manusia yang disebut sebagai jaringan saraf tiruan (Neural Network). Dengan begitu, Deep Learning mampu memahami data berdimensi tinggi dengan struktur yang rumit. Sehingga algoritma pada Deep Learning mampu memecahkan berbagai macam masalah yang membutuhkan *input* data yang besar atau berdimensi tinggi [36]. Gambar 2.4 di bawah menunjukkan ilustrasi jaringan Deep Learning (Deep Neural Network).



Gambar 2.4 Ilustrasi jaringan *Deep Learning* [37]

2.2.8.1 RNN (Recurrent Neural Network)

RNN (Recurrent Neural Network) atau dapat disebut sebagai jaringan saraf berulang adalah jenis dari Neural Network yang mampu melakukan proses pada beberapa set data berkelanjutan (data sekuensial). Hal tersebut mengikuti kemampuan otak manusia yang secara umum tidak selalu membuat keputusan begitu saja namun selalu mempertimbangkan pengetahuan masa lalu saat melakukan pengambilan keputusan. Seperti halnya manusia, RNN mampu menyimpan informasi yang diperoleh pada masa lalu dengan mengimplementasikan *loop* dalam penerapan arsitekturnya yang secara otomatis menyimpan informasi masa lalu [38].

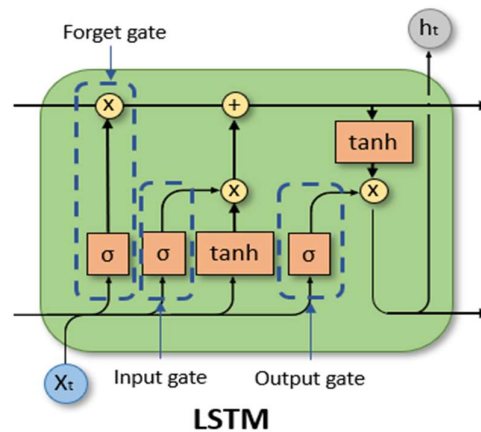


Gambar 2.5 Diagram RNN [39]

Gambar 2.5 merupakan diagram RNN yang tersusun dari *input* X_t , aktivasi fungsi tanh dan *output* h_t . X_t merupakan *input* sekuensial pada waktu t , fungsi tanh bertujuan untuk menentukan nilai *output* dari sebuah *neruon* berdasarkan *inputnya* dan h_t adalah *output* pada waktu t yang menyimpan informasi tentang konteks sekuensial. *Output* pada RNN disebut sebagai *state* atau memori.

2.2.8.2 LSTM (Long-Short Term Memory)

LSTM (Long-Short Term Memory) adalah salah satu jenis RNN sederhana yang dibuat untuk mengatasi masalah berkurangnya efektivitas dalam mempelajari tren jangka panjang pada data sekuensial [40].



Gambar 2.6 Diagram LSTM [39]

Pada Gambar 2.6 dapat terlihat diagram LSTM yang umumnya tersusun dari sebuah *cell state*, *input gate*, *forget gate* dan *output gate*, susunan tersebut merupakan modifikasi dari RNN sederhana yang hanya memiliki satu gerbang dengan fungsi tanh. *Input gate* menggunakan aktivasi fungsi *sigmoid* (σ) untuk menentukan informasi apa saja yang harus masuk ke *cell state* sebagai sel utama dengan mengalikannya pada hasil luaran fungsi *tanh* pada gerbang utama RNN dan menjumlahkan hasil kali ke sel utama. *Forget gate* menentukan informasi yang harus dihapus dari *cell state* sebelumnya menggunakan aktivasi fungsi *sigmoid*, sehingga hanya informasi relevan yang akan disimpan. *Output gate* menggunakan fungsi *sigmoid* kemudian dikalikan dengan keluaran fungsi tanh dari sel

utama untuk menentukan seberapa banyak *internal state* yang harus diteruskan ke layer selanjutnya yang lebih besar. Dengan begitu, sebuah *neural network* yang menggunakan LSTM dapat mempelajari dependensi jangka panjang dan dapat memproses seluruh rangkaian data dalam data sekuensial [41].

2.2.9 Akurasi dan kesalahan (*error*)

Akurasi dan kesalahan atau biasa disebut *error* adalah dua hal yang sering digunakan untuk mengevaluasi kinerja suatu metode, alat atau sistem. Kedua nilai tersebut saling berkaitan dalam melakukan pengukuran suatu variabel oleh sebuah sistem. Nilai yang tepat dari sebuah variabel disebut nilai sebenarnya. Sedangkan nilai yang dihasilkan dari pengukuran variabel oleh suatu sistem disebut sebagai nilai pengukuran. Akurasi mengacu pada kedekatan antara nilai pengukuran dan nilai sebenarnya. Namun, nilai sebenarnya dari suatu variabel jarang dapat diukur dengan tepat, selisih antara nilai yang diukur dengan nilai sebenarnya disebut sebagai *error* [42]. Dengan begitu, *error* dapat diperoleh dengan persamaan (2.1).

$$e = \text{nilai pengukuran} - \text{nilai sebenarnya} \quad (2.1)$$

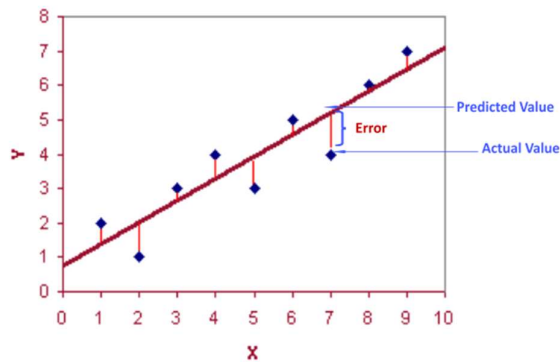
2.2.9.1 *Loss*

Nilai *loss* identik dengan *error*, yaitu nilai yang dapat digunakan untuk mengukur selisih nilai yang diukur dengan nilai sebenarnya. Pada dasarnya *loss* adalah nilai yang berdasar pada *error*. Namun, *loss* memiliki sedikit perbedaan dengan *error*. *Error* digunakan untuk menghitung selisih pada kesalahan pada hasil pengukuran, sedangkan *loss* digunakan untuk mengukur kesalahan yang terjadi pada model atau metode yang digunakan. Pada perhitungannya, nilai *loss* dihitung menggunakan fungsi *loss* sedangkan *error* adalah nilai selisihnya.

2.2.9.2 RMSE (Root Mean Squared Error)

RMSE atau Root Mean Squared Error adalah salah satu metode untuk pengukuran nilai *error* untuk mengevaluasi model. RMSE termasuk kedalam *loss function* atau fungsi *loss*. Metode tersebut biasa digunakan untuk pengukuran performa pada model yang digunakan untuk melakukan

prediksi regresi linear [43]. Pengukuran *error* pada RMSE diperoleh dengan mengukur jarak pada nilai pengukuran (*predicted value*) dengan nilai sebenarnya (*actual value*) pada garis linear. Gambar 2.7 menunjukkan ilustrasi perhitungan nilai *error* pada metode RMSE.



Gambar 2.7 Ilustrasi nilai *error* pada RMSE [44]

Nilai RMSE didapatkan dengan mencari nilai akar yang didapatkan dari rata-rata jumlah selisih nilai sebenarnya yang dikuadratkan. Persamaan (2.2) di bawah menunjukkan perhitungan RMSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.2)$$

Keterangan:

y_i = nilai sebenarnya pada data ke-i

\hat{y}_i = nilai pengukuran pada data ke-i

n = jumlah data