

LAPORAN PROJECT BASED LEARNING
PREDIKSI HARGA JUAL RUMAH DI TANGERANG SELATAN
MENGGUNAKAN *BIDIRECTIONAL RECURRENT NEURAL*
NETWORKS (BRNN)

Fadillah Muhesa Latri¹⁾, Intan Melyani Yansori²⁾, Debora Sebrina Simanjuntak³⁾,
Nana Diana⁴⁾

Program Studi Sains Data, Jurusan Sains, Institut Teknologi Sumatera

Email : fadillah.120450003@student.itera.ac.id¹⁾,
intan.120450039@student.itera.ac.id²⁾, nana.120450042@student.itera.ac.id³⁾,
debora.120450073@student.itera.ac.id⁴⁾

Abstrak

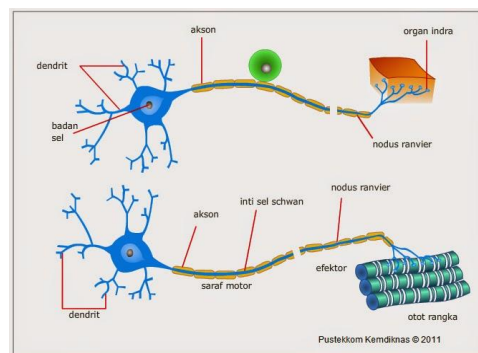
Recurrent Neural Networks (RNN) merupakan pendekatan dalam pemrosesan data sekuensial, di mana informasi dari langkah waktu sebelumnya digunakan untuk memahami data saat ini. Namun, RNN memiliki keterbatasan dalam memproses data masa depan. Oleh karena itu, *Bidirectional Recurrent Neural Networks* (BRNN) dikembangkan sebagai solusi, memungkinkan pemrosesan dua arah dalam data sekuensial. Penelitian ini mengevaluasi parameter dan preprocessing data untuk model BRNN dalam prediksi harga rumah. Percobaan menunjukkan bahwa model dengan parameter tertentu memberikan hasil yang baik, namun saran diberikan untuk mempertimbangkan arsitektur jaringan yang lebih sesuai, kembali memproses data, dan menggunakan matrik evaluasi yang relevan untuk meningkatkan kualitas model prediksi harga rumah.

Kata Kunci : BRNN, RNN, Matrik Evaluasi

I. Pendahuluan

1.1 Latar Belakang

Recurrent Neural Network adalah jenis *neural network* yang digunakan dalam *machine learning*. Model-model *neural network* bekerja menirukan sistem kerja saraf tetapnya kerja otak pada manusia. *Neural network* terdiri dari *neuron-neuron* (unit komputasi) yang saling terhubung dalam lapisan-lapisan. Ilustrasinya seperti berikut:



Gambar 1. Sistem Saraf Manusia

- Dendrit sebagai penerima *input* rangsangan dari sensor tubuh manusia
- Badan sel atau *nucleus* memprosesnya, dan
- *Output* dikeluarkan oleh akson berupa energi listrik dengan daya rendah menuju dendrit sel saraf lain yang ada di depannya.

Ilmuwan menirukan sel saraf tersebut menjadi konsep yang sederhana seperti ilustrasi berikut:

- Dendrit adalah ditirukan sebagai *input* dari beberapa saluran masukan dari x_1, \dots, x_n , dimana n itu adalah jumlah total inputan. Dendrit berperan sebagai jalur masukan yang mengirimkan sinyal ke sel saraf.
- *weight* adalah bilangan pengali, artinya setiap *input* dikali pada bilangan bobot atau *weight* dari w_1, \dots, w_n . *Weight* ini menggambarkan pentingnya setiap *input* dalam pengambilan keputusan oleh unit *neural*.
- *Nucleus* ditirukan sebagai pemroses data input dari x_1, \dots, x_n dilakukan operasi matematis dengan *weight* masing-masing atau dari w_1, \dots, w_n untuk menggabungkan input dengan bobotnya, mengaktifkan fungsi aktivasi, dan menghasilkan *output*.

Jadi, dalam RNN, kemampuan "mengingat" diimplementasikan melalui unit neural dan mekanisme rekuren yang memungkinkan jaringan untuk memproses data sekuensial dengan memperhitungkan informasi sebelumnya.

Recurrent Neural Networks (RNN) dikembangkan berdasarkan pemahaman bahwa dalam data sekuensial, seperti teks, audio, atau deret waktu, informasi dari langkah waktu sebelumnya sangat penting untuk memahami data di langkah waktu saat ini. RNN beroperasi dengan cara yang mirip dengan otak manusia, di mana unit komputasi menyimpan informasi dari langkah waktu sebelumnya dan menggunakannya untuk memahami konteks saat ini. Namun, RNN memiliki beberapa kekurangan, salah satunya adalah keterbatasan dalam memahami permasalahan dari data penelitian masa depan dalam data sekuensial, sehingga RNN melakukan pengembangan model salah satunya adalah *Bidirectional Recurrent Neural Networks* (BRNN).

BRNN adalah perluasan dari RNN yang mengatasi kekurangan ini dengan menggunakan dua arah, yaitu maju (*forward*) dan mundur (*backward*), untuk memproses data sekuensial. Dengan cara ini, BRNN dapat melakukan prediksi informasi untuk masa akan datang dan masa lalu.

Bidirectional Recurrent Neural Networks (BRNN) yang memiliki sifat sekuensial dapat diimplementasikan dalam prediksi harga jual rumah di wilayah Tangerang Selatan karena kemampuannya untuk memproses data sekuensial dengan mempertimbangkan informasi masa lalu dan masa depan. Dalam konteks harga jual rumah, data historis mengenai penjualan rumah dan faktor-faktor yang mempengaruhi harga, seperti lokasi, ukuran, fasilitas, dan tren pasar, memiliki aspek sekuensial yang signifikan. Misalnya, tren harga rumah pada suatu lokasi dapat dipengaruhi oleh faktor-faktor seperti perubahan infrastruktur, kebijakan pemerintah, dan perubahan kebutuhan pembeli. Dengan BRNN, kita dapat memodelkan hubungan kompleks ini dengan mempertimbangkan informasi historis dan perkiraan harga di masa depan.

Perkembangan teknologi informasi, seperti BRNN menggunakan algoritma yang mendalam, membantu memproses data sekuensial dengan akurasi yang lebih tinggi. Dalam prediksi harga jual rumah, model BRNN dapat digunakan untuk memperkirakan harga rumah di masa depan dengan akurasi baik, yang berguna bagi pemilik rumah, investor, dan agen real estate dalam pengambilan keputusan.

Dengan demikian, BRNN yang sekuensial merupakan salah satu bentuk implementasi teknologi deep learning memberikan solusi yang efektif dalam memprediksi harga jual rumah di Tangerang Selatan, memanfaatkan informasi masa lalu dan masa depan untuk memberikan estimasi harga yang lebih akurat.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, maka terdapat rumusan masalah yang sesuai pada judul “Prediksi Harga Jual Rumah Di Tangerang Selatan Menggunakan *Bidirectional Recurrent Neural Networks* (BRNN)” sebagai berikut:

1. Bagaimana membangun model prediksi harga penjualan rumah yang akurat di Tangerang Selatan menggunakan algoritma *Bidirectional Recurrent Neural Networks* (BRNN)?
2. Bagaimana untuk menguji pengaruh penggunaan BRNN terhadap akurasi prediksi harga rumah?

1.3 Batasan Masalah

Berikut beberapa batasan masalah dari penelitian ini mencakup:

1. Data yang digunakan hanya data mengenai harga penjualan rumah di daerah Tangerang yang mencakup; properti, lokasi, harga, jumlah kamar tidur, jumlah kamar mandi, dan luas lantai.
2. Model yang dibangun hanya menggunakan *Bidirectional Recurrent Neural Networks* (BRNN).

1.4 Tujuan

Penelitian ini bertujuan untuk membangun model prediksi harga penjualan rumah yang akurat dengan menggunakan algoritma *Bidirectional Recurrent Neural Networks* (BRNN) dan menguji pengaruh penggunaan fungsi aktivasi ReLU dalam memprediksi performa model. Selain itu, hal ini dilakukan untuk memberikan hasil rekomendasi terkait penggunaan model prediksi harga penjualan rumah yang telah dibangun yang lebih akurat dan efektif.

II. Landasa Teori

2.1 Machine Learning

Machine Learning (ML) telah menjadi alat yang sangat diperlukan dalam sistem perangkat lunak, meningkatkan aplikasi yang sudah ada dan memungkinkan pembuatan aplikasi yang baru. Machine Learning (ML) atau Mesin Pembelajaran adalah cabang dari AI yang fokus belajar dari data (learn from data), yaitu fokus pada pengembangan sistem yang mampu belajar secara “mandiri” tanpa harus berulang kali diprogram manusia. Platform perangkat lunak yang tersedia untuk pembelajaran mesin mencerminkan akar akademis dan industri dari teknologi tersebut. Machine Learning menjawab pertanyaan tentang komputer yang secara otomatis membuat mesin belajar dari pengalaman. Ini adalah salah satu bidang teknis yang tumbuh paling cepat, yang berada di persimpangan ilmu komputer dan statistik, dan merupakan inti dari kecerdasan buatan dan ilmu data.

2.2 Deep Learning

Teknologi Deep Learning ini salah satu teknologi yang sangat populer untuk mengenali suatu kegiatan atau objek yang memiliki tingkat keakuratan yang sangat tinggi dibandingkan dengan metode mesin sebelumnya. Proses pembelajaran mesin dilakukan pada komputer yang berfungsi untuk mengklasifikasi data citra menjadi hasil klafikasi berupa prediksi. Deep Learning merupakan algoritma jaringan syaraf tiruan yang menggunakan data sebagai input dan kemudian memprosesnya dengan menggunakan sejumlah lapisan tersembunyi (hidden layer). CNN atau yang lebih dikenal dari Convolutional Neural Network, merupakan salah satu metode dalam Deep Learning yang dibuat untuk menutupi kelemahan dalam metode sebelumnya, akan tetapi dengan model ini sejumlah parameter bebas dapat dikurangi dan deformasi gambar input seperti translasi, rotasi dan skala dapat ditangani.

2.3 Neural Network

Awal bermunculnya Akan tetapi, dalam penggunaannya RNN tidak mampu melakukan penyimpanan informasi sebelumnya dengan baik dikarenakan permasalahan gradien yang menghilang. Neural Network berasal dari penelitian pada otak manusia, dimana komponen dasarnya yaitu otak dan saraf, yang ditemukan pada tahun 1836. Kebanyakan Neural Network yang digunakan tidak 100% akurat. Solusi neural network digunakan bukan karena solusinya sempurna tetapi karena kinerjanya lebih baik daripada model komputasi alternatif.

2.4 Recurrent Neural Network

Recurrent Neural Network (RNN) adalah metode yang dirancang untuk menganalisis data sequential dengan memanfaatkan algoritma deep learning. RNN merupakan kelas dalam deep learning dengan proses pembelajaran yang terawasi. Dikarenakan kemampuan RNN dalam analisis data sequential, maka metode ini dapat digunakan untuk memproses data time series. Kelas deep learning ini memiliki arsitektur yang sama dengan deep learning yaitu input layer, hidden layer, dan output layer. Berdasarkan arsitektur yang dimiliki RNN, secara teori metode ini dapat menyelesaikan permasalahan ketergantungan dalam jangka panjang. Arsitektut jaringan ini adalah sebagai berikut:

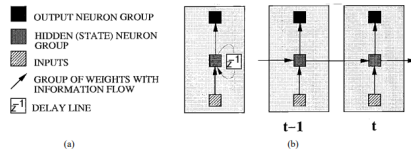


Fig. 1. General structure of a regular unidirectional RNN shown (a) with a delay line and (b) unfolded in time for two time steps.

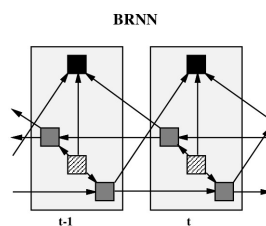
Gambar 2. Sistem RNN

1. Input layer (w) dan output layer (y) mempunyai dimensi yang sama sebagai kosakata.
2. Hidden layer (s) merupakan urutan yang besarnya lebih kecil.
3. U adalah matriks bobot antara hidden dan output layer.
4. Tanpa bobot berulang (W) (recurrent weight), model ini dapat menjadi jaringan bigram.

2.5. Bidirectional Recurrent Neural Network

Bidirectional Recurrent Neural Networks (BRNN) adalah jenis arsitektur jaringan syaraf tiruan yang memungkinkan informasi untuk mengalir maju dan mundur melalui jaringan. Dalam BRNN, input diberikan ke jaringan dari kedua arah, yaitu dari awal ke akhir dan dari akhir ke awal. Hal ini memungkinkan jaringan untuk mempelajari pola yang kompleks dan konteks yang lebih luas dalam data yang diolah.

BRNN terdiri dari dua lapisan jaringan syaraf tiruan, yaitu lapisan maju dan lapisan mundur. Lapisan maju memproses data urutan dari awal hingga akhir, sedangkan lapisan mundur memproses data urutan dari akhir hingga awal. Hasil dari kedua lapisan tersebut kemudian digabungkan untuk menghasilkan representasi yang lebih baik dari data urutan.



Gambar 3. Sistem BRNN

2.6 Matriks Evaluasi

Memiliki dua lapisan tersembunyi berulang yang berbeda pada BRNN, salah satunya memproses urutan masukan ke depan dan yang lainnya memprosesnya ke belakang. Setelah itu, hasil dari lapisan tersembunyi tersebut dikumpulkan dan dimasukkan ke dalam lapisan akhir pembuatan prediksi. Dibandingkan dengan jaringan saraf berulang searah konvensional, keakuratan BRNN lebih baik karena dapat memproses informasi dalam dua arah dan memperhitungkan konteks masa lalu dan masa depan. Karena dua lapisan Forward Stated dan Backward Stated dapat saling melengkapi dan memberikan lebih banyak data pada lapisan prediksi akhir, penggunaan dua lapisan yang berbeda juga menawarkan jenis regularisasi model.

Untuk memperbarui parameter model, gradien dihitung untuk lintasan maju dan mundur propagasi mundur melalui teknik waktu yang biasanya digunakan untuk melatih BRNN. Urutan masukan diproses oleh BRNN dalam satu lintasan maju pada waktu inferensi, dan prediksi dibuat berdasarkan keluaran gabungan dari dua lapisan tersembunyi.

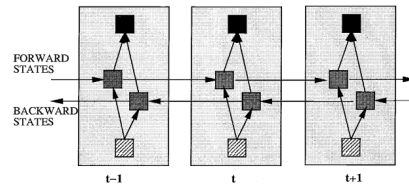


Fig. 3. General structure of the bidirectional recurrent neural network (BRNN) shown unfolded in time for three time steps.

Gambar 4. *Forward dan Backward States*

III. Metodologi

3.1 Data Understanding

Penelitian ini menggunakan data Prediksi Harga Jual Rumah di Tangerang Selatan yang Berisi Informasi *listing properti* dengan kurun waktu 01 mei 2023 hingga 19 juni 2023. Dataset bersumber dari kaggle.com, akan tetapi data tersebut merupakan hasil *scraping* dari Rumah.com. Pemilik dataset ini adalah Gerry Zani.

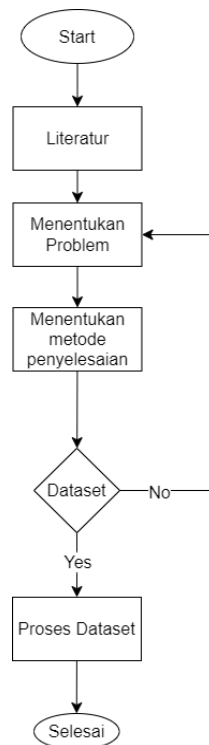
Data jual rumah di Tangerang Selatan berjumlah 29420 data dan memiliki 7 atribut yaitu *nav-link href*, *listing-location*, *price*, *bed*, *bath*, *listing-floorarea*, dan *listing-floor area 2*. Dataset dibagi menjadi dua bagian dengan 80:20, 80 adalah data *train* dan 20 data *test*. Dengan kumpulan data tersebut, kita dapat menganalisis harga dan karakteristik properti untuk mendapatkan informasi tentang pasar *real estate*. Seperti pada gambar dibawah ini.

	nav-link href	listing-location	price	bed	bath	listing-floorarea	listing-floorarea 2
0	https://www.rumah.com/listing-properti/dijual-...	Gading Serpong, Tangerang Selatan, Banten	6,9 M	5.0	5.0	420 m ²	Rp 20.720.721 per m ²
1	https://www.rumah.com/listing-properti/dijual-...	Gading Serpong, Tangerang Selatan, Banten	4.5 M	3.0	3.0	190 m	Rp 12.747.875 per m ²
...
29418	https://www.rumah.com/listing-properti/dijual-...	Bintaro, Tangerang Selatan, Banten	1,95 M	3.0	2.0	120 m ²	Rp 21.666.667 per m ²
29419	https://www.rumah.com/listing-properti/dijual-...	Bintaro, Tangerang Selatan, Banten	5,5 M	4.0	3.0	200 m ²	Rp 30.555.556 per m ²

Gambar 5. Tabel Dataset

1. *nav-link href*, memuat informasi alamat *website* setiap rumah
2. *listing-location*, memuat informasi alamat rumah
3. *price*, merupakan nominal rumah yang dijual
4. *bed*, merupakan jumlah kamar tidur yang ada pada rumah di jual
5. *bath*, merupakan jumlah kamar mandi atau toilet yang ada pada rumah
6. *listing-floor area*, luas area rumah dalam satuan m^2
7. *listing-floor area 2*, merupakan nominal jual rumah per m^2

3.2 Proses Mendapatkan Data



Gambar 6. *Flowchart* Metode Penelitian

Berikut Penjelasan Gambar 6. *Flowchart* Metode Penelitian:

1. Literatur

Tahap pertama dalam proses mendapatkan data yaitu menjalani penelitian literatur untuk memahami kerangka kerja, konteks yang relevan, dan melibatkan eksplorasi literatur sebelumnya serta pemahaman mendalam tentang prediksi harga penjualan rumah di Tangerang Selatan menggunakan algoritma *Bidirectional Recurrent Neural Networks* (BRNN).

2. Menentukan *Problem*

Setelah melakukan literatur tentang prediksi harga penjualan rumah di Tangerang Selatan menggunakan algoritma *Bidirectional Recurrent Neural Networks* (BRNN), langkah berikutnya adalah mengidentifikasi masalah atau merancang rumusan masalah agar dapat merinci permasalahan yang akan diselesaikan.

3. Menentukan Metode Penyelesaian

Setelah permasalahan diidentifikasi, dilakukan pemilihan dan perincian metode yang akan digunakan untuk menyelesaikan masalah atau menjawab rumusan masalah. Metode ini melibatkan pengumpulan data dan analisis menggunakan algoritma *Bidirectional Recurrent Neural Networks* (BRNN).

4. Dataset

Untuk mendukung metode penyelesaian yang telah dipilih, peneliti harus mendapatkan dataset yang relevan. Pada langkah ini, akan menghasilkan dua kemungkinan jawaban yaitu ya atau tidak. Apabila dataset yang didapat

belum memenuhi kriteria atau kurang relevan (tidak) maka peneliti akan kembali kelangkah sebelumnya untuk menentukan problem. Sedangkan, Apabila dataset yang didapat sudah relevan (ya) maka peneliti akan melanjutkan ke tahap proses dataset. Dataset yang digunakan yaitu harga penjualan rumah di Tangerang Selatan yang berasal dari sumber terbuka *Kaggle*.

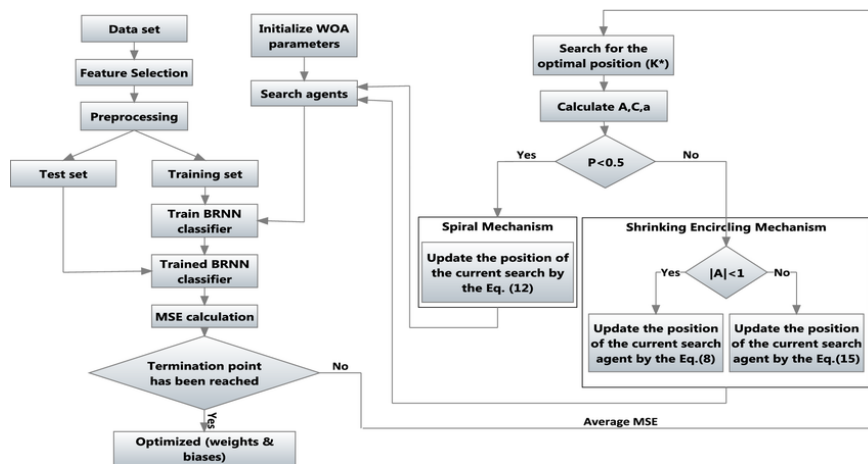
5. Proses Dataset

Proses data melibatkan pengolahan dan analisis dataset dengan menggunakan metode yang telah ditentukan. Hal ini mencakup pembersihan data, transformasi, penggabungan data, dan ekstraksi informasi yang relevan.

6. Selesai

Setelah analisis data selesai, peneliti dapat menggambarkan hasil penelitian, menginterpretasikan temuan, dan menyusun laporan penelitian yang mencakup temuan-temuan tersebut.

3.3 Proses Pengolahan Data



Gambar 7. Flowchart Pengolahan Data

Flowchart Pengolahan Data adalah diagram visual yang secara sistematis menguraikan tahapan yang diperlukan sehingga dapat menghasilkan informasi yang relevan dan bermanfaat. Flowchart ini membantu peneliti untuk memahami secara jelas langkah-langkah yang diperlukan dalam menjalankan proses pengolahan data, sehingga memudahkan pengambilan keputusan dan analisis data yang efektif. Berikut adalah penjelasan dari Gambar 7. Flowchart Pengolahan Data:

1. Start

Tahap awal proses pengolahan data, yaitu *input* dataset dan eksplorasi data untuk mendapatkan informasi dari dataset.

2. Feature Selection

Variabel-variabel pada dataset dapat disesuaikan kebutuhan penggunaannya, sehingga pemilihan variabel tersebut dapat menggunakan *feature selection*.

3. Processing

Proses pengolahan data seperti pembersihan data, transformasi data, integrasi. Pada proses ini, data dibagi menjadi *data train* sebagai data latih model, dan *data test* sebagai data uji model. *Processing* data adalah aspek fundamental

dan sangat penting untuk membuat keputusan berdasarkan data serta mengekstrak wawasan berharga dari informasi mentah.

4. Model data

Penerapan BRNN dalam prediksi harga penjualan rumah di Tangerang Selatan memungkinkan penggunaan BRNN untuk memprediksi harga penjualan rumah. Model ini memiliki keunggulan dalam pemrosesan data sekuensial dan pemahaman konteks temporal serta faktor kontekstual yang mempengaruhi harga rumah. Sebagai hasilnya, penggunaan BRNN sangat relevan dan efektif untuk mengoptimalkan pemahaman dan prediksi harga rumah.

5. *Mean Squared Error (MSE)*

Proses ini untuk metrik evaluasi model. Ketika metrik evaluasi nilainya baik dengan akurasi tinggi atau nilai error kecil maka dapat ditarik kesimpulan (*optimize*), ketika akurasi kecil dilakukan perulangan model, dengan *search for the optimal parameter*.

6. *search for the optimal parameter*.

Proses ini bertujuan untuk mencari dan menentukan parameter yang memberikan hasil terbaik atau optimal dalam suatu konteks tertentu. Proses pencarian parameter optimal dalam deep learning, terutama ketika menggunakan arsitektur jaringan saraf rekuren (RNN) *bidirectional* (BRNN), melibatkan pengoptimalan berulang dengan langkah-langkah maju (*forward*) dan mundur (*backward*), yang sesuai dengan proses pelatihan model.

Dalam arsitektur jaringan saraf rekuren (RNN) *bidirectional* (BRNN), perulangan *forward* adalah langkah yang digunakan untuk meneruskan informasi dari awal sampai akhir dalam sekuensial data. Pada awalnya, sekuensial data dimasukkan ke dalam BRNN melalui bagian maju. Ini bisa berupa sekuensial teks, audio, atau data lain yang diproses berdasarkan urutan. Kemudian data diumpankan melalui setiap elemen sekuensial dalam urutan maju. Dalam setiap langkah waktu (*time step*), BRNN memproses data yang masuk dan menghasilkan representasi atau output pada langkah waktu tersebut. Ini melibatkan komputasi dalam satu arah, yaitu dari awal hingga akhir sekuensial data. Selama perulangan *forward*, BRNN membangun representasi data yang memperhitungkan konteks dari awal sekuensial data hingga titik saat ini. Hal ini memungkinkan model untuk memahami sekuensial data dengan mempertimbangkan informasi yang ada di sebelah kiri atau sebelumnya. Setelah sekuensial data diolah dalam perulangan *forward*, hasil akhir dari bagian maju (*forward*) akan digunakan untuk berbagai tugas seperti klasifikasi, analisis, atau prediksi.

Perulangan *forward* dalam BRNN memungkinkan model untuk memahami sekuensial data dengan mempertimbangkan konteks dari bagian awal hingga titik saat ini. Namun, BRNN juga memiliki bagian mundur (*backward*) yang memproses data dalam arah sebaliknya, dari akhir ke awal. Dengan menggunakan kombinasi kedua arah ini, BRNN memiliki kemampuan untuk memahami sekuensial data secara lebih komprehensif dengan memperhitungkan konteks dari kedua arah, yang seringkali sangat berguna dalam tugas pemrosesan bahasa alami, pemrosesan teks, atau analisis data sekuensial lainnya.

Proses perulangan *backward* dimulai dengan mengambil hasil dari bagian maju (*forward*) dari BRNN. Data hasil dari bagian *forward* dianggap sebagai masukan awal untuk bagian *backward*. Data masukan dari hasil *forward* diumpankan kembali

melalui BRNN dalam arah mundur. Dalam setiap langkah waktu (time step), BRNN memproses data masukan dari akhir sekuensial data ke awal. Selama perulangan *backward*, BRNN membangun representasi data yang memperhitungkan konteks dari akhir sekuensial data ke titik saat ini. Hal ini memungkinkan model untuk memahami sekuensial data dengan mempertimbangkan informasi yang ada di sebelah kanan atau setelahnya. Hasil akhir dari bagian *backward* (mundur) adalah representasi yang mencerminkan konteks sekuensial data dari akhir ke awal. Output ini dapat digunakan untuk berbagai tugas, seperti klasifikasi, analisis, atau prediksi[4].

3.4 Pseudocode BRNN

Pseudocode adalah suatu notasi atau cara penulisan algoritma yang bersifat informal dan mudah dimengerti oleh manusia, yang digunakan untuk merencanakan dan menggambarkan langkah-langkah suatu algoritma atau program komputer tanpa harus mengikuti sintaksis tertentu dari bahasa pemrograman tertentu. *Pseudocode* digunakan sebagai alat bantu dalam merancang, mengkomunikasikan, dan mendokumentasikan pemikiran dan logika algoritma sebelum diimplementasikan dalam sebuah bahasa pemrograman yang sebenarnya. Berikut *Pseudocode* BRNN yang digunakan:

```
def forward_pass(xo):
    h_f= xo
    for i in range(numlayers-1):
        y_f= dot(W_f[i], h_f) + B_f[i]
        h_f= activation[i](y_f)
    return h_f

def backward_pass(xo):
    h_b= xo
    for i in range(numlayers-1):
        y_b= dot(W_b[i], h_b) + B_b[i]
        h_b= activation[i](y_b)
    return h_b

def predict(xo):
    h_f= forward_pass(xo)
    h_b= backward_pass(xo)

# Gabungkan output forward & backward)
comb_output = Concatenate(h_f, h_b)
y = dot(W_o, comb_output) + B_o
return
activation[numlayers-1](output)
```

Keterangan :

f : *Forward*

b: *Backward*

h: *Hidden state*

o : *Output*

B: Bias *Output*
W : *Weight*
y: *Output* dari *input*
xo : Input

IV. Hasil dan Pembahasan

4.1 Data

Data dalam penelitian ini menggunakan data sekunder diperoleh dari kaggle.com yang diakses publik dengan judul “*Housing Price in South Tangerang City (Indonesia)*”. Pemilik dataset ini adalah Gerry Zani . Dataset ini berjumlah 29420 data, dengan atribut dataset berjumlah 7 sebagai berikut penjelasan dari dataset:

- nav-link href, memuat informasi alamat *website* setiap rumah.
- listing-location, memuat informasi alamat rumah.
- price, merupakan nominal rumah yang dijual
- bed, merupakan jumlah kamar tidur yang ada pada rumah di jual
- bath, merupakan jumlah kamar mandi atau toilet yang ada pada rumah
- listing-floorarea, luas area rumah dalam satuan m^2
- listing-floorarea2, merupakan nominal jual rumah per m^2 .

	nav-link href	listing-location	price	bed	bath	listing-floorarea	listing-floorarea 2
0	https://www.rumah.com/listing-onpersidijak...	Gading Serpong, Tangerang Selatan, Banten	6,9 M	5,0	5,0	420 m ²	Rp 20.720.721 per m ²
1	https://www.rumah.com/listing-onpersidijak...	Gading Serpong, Tangerang Selatan, Banten	4,5 M	3,0	3,0	180 m ²	Rp 12.747.815 per m ²
2	https://www.rumah.com/listing-onpersidijak...	Gading Serpong, Tangerang Selatan, Banten	3,95 M	3,0	3,0	132 m ²	Rp 30.859.375 per m ²
3	https://www.rumah.com/listing-onpersidijak...	Gading Serpong, Tangerang Selatan, Banten	3,3 M	4,0	3,0	220 m ²	Rp 18.333.333 per m ²
4	https://www.rumah.com/listing-onpersidijak...	Gading Serpong, Tangerang Selatan, Banten	3,5 M	3,0	2,0	180 m ²	Rp 27.777.778 per m ²

Gambar 9. Dataset Harga Rumah di Tangerang Selatan

Dataset dibagi menjadi dua bagian dengan 80:20, 80 adalah data *train* dan 20 data *karang*. Berdasarkan korelasi data, bahwa hasil korelasi Pearson, Spearman, dan Kendall adalah

	Pearson		Spearman		Kendall	
	bed	bath	bed	bath	bed	bath
bed	1.000000	0.799451	1.000000	0.810829	1.000000	0.730314
bath	0.799451	1.000000	0.810829	1.000000	0.730314	1.000000

Gambar 10. Korelasi Data

Gambar 10 artinya yang memiliki korelasi kuat antara kolom adalah "bed" dan "bath." Sebagai indikasi bahwa jumlah kamar tidur ("bed") berkorelasi positif dengan jumlah kamar mandi ("bath"). Sehingga penulis melakukan *feature selection* untuk pemodelan data berdasarkan korelasi tersebut.

Hasilnya sebagai berikut:

Percobaan 1

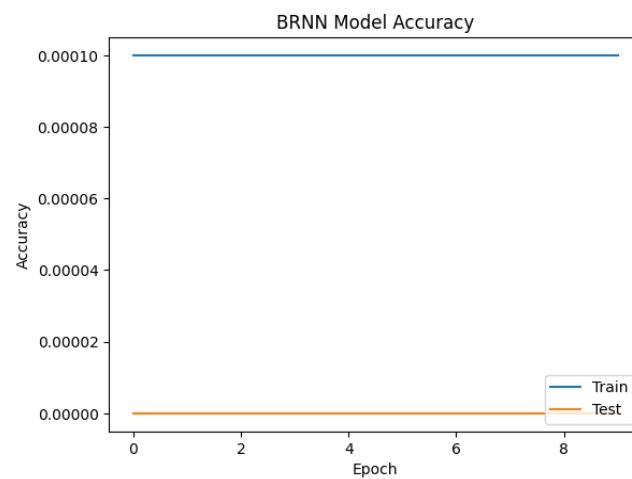
Dataset terdapat *missing value*, yaitu atribut bath = 69,68 %, bed=46,90 % dan listing-floorarea 2 = 12,57%

Parameter yang digunakan pada percobaan 1 adalah

- Layer SimpleRNN dalam Bidirectional
- Fungsi aktivasi lapisan Dense = Relu layer pertama dan sigmoid yang kedua
- Epoch/iterasi= 10

- Hidden Layer =
- Fungsi Aktivasi = Sigmoid
- Optimizer = adam
- loss= binary crossentropy
- metrics = accuracy

Evaluasi Model	BRNN
loss	nan
acc	9.9975e-5
val loss	-301798
vall acc	0.0000e_oo



Gambar 11. Percobaan 1

Percobaan 2

Parameternya

Layers LSTM dalam mode Bidirectional

Hidden Layer 64, 128, 1

Fungsi aktivasi LSTM = tanh

Fungsi aktivasi lapisan Dense =Relu layer pertama dan sigmoid yang kedua

Optimizer = Adam

Metrik = Accuracy

Epoch = 5

Regularisasi = L2 pada layer LSTM dengan kernel-regularizer=12

Evaluasi Model	BRNN
loss	nan

acc	9.9975e-5
val loss	-301798
vall acc	0.0000e_oo

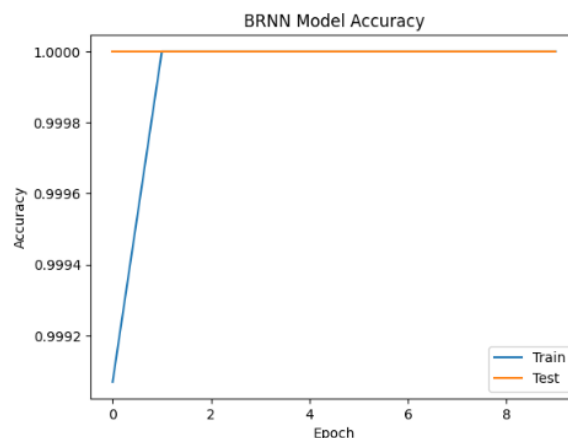
4.2 Evaluasi Model

a. percobaan 1

- Model ini memiliki banyak lapisan, dan dalam pelatihan yang disertakan, dilatih selama 50 epoch.
- Anda menghadapi masalah dengan loss yang sangat besar dalam bentuk negatif, yang mungkin disebabkan oleh pengaturan yang salah dalam model atau tugas yang tidak sesuai dengan arsitektur model ini. Mungkin juga ada masalah dengan pre-processing data.
- Akurasi model sangat rendah (sekitar 0.0001) dan tidak ada peningkatan yang signifikan selama pelatihan.
- Loss validasi juga sangat besar dan tidak menunjukkan peningkatan yang signifikan. Ini menunjukkan overfitting atau masalah serius dalam arsitektur model atau data yang digunakan.
- Hasil ini menunjukkan bahwa model BRNN yang digunakan dalam percobaan ini tidak efektif dalam tugas prediksi harga rumah. Perlu untuk mengevaluasi dan memperbaiki arsitektur model atau memeriksa masalah dalam pre-processing data.

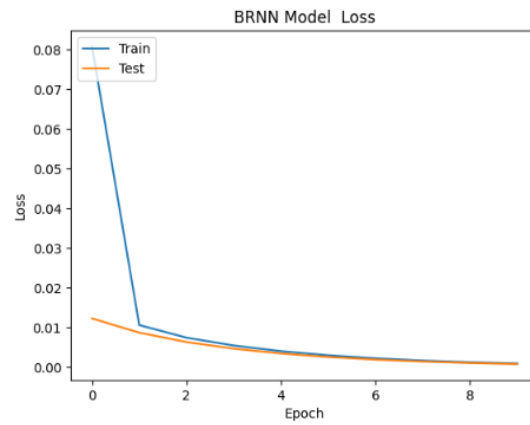
b. percobaan 2

Hasil percobaan menunjukkan bahwa model telah mencapai akurasi yang sangat tinggi (100%) pada data *test*, dan 99% untuk data *train* seperti ditampilkan pada Gambar 12 dan loss yang sangat rendah yaitu kurang dari 1% seperti pada gambar 13. Ini dapat mengindikasikan overfitting, di mana model telah "menghafal" data pelatihan, tetapi hasil yang baik pada data validasi juga menunjukkan kinerja yang baik. Penting untuk melakukan evaluasi lebih lanjut untuk memastikan bahwa model tersebut tidak mengalami overfitting dan dapat digeneralisasi ke data baru dengan baik.



Gambar 12. Akurasi Percobaan 2

Percobaan pertama mengalami akurasi yang sangat kecil atau buruk, karena ada kesalahan dalam pembuatan parameter. Percobaan 1 untuk hidden layer sama semua yaitu 32,32,32, selain itu fungsi aktivasi setiap hidden dan lapisan dense sama.



Gambar 13. Loss Percobaan 2

Berdasarkan Gambar 13 semakin banyak *epoch* nilai *loss* semakin kecil, artinya model semakin terlatih.

V. Penutup

5.1 Kesimpulan

Berdasarkan hasil percobaan, dapat disimpulkan bahwa model prediksi harga penjualan rumah menggunakan algoritma *Bidirectional Recurrent Neural Networks* (BRNN) dengan parameter yang tepat memberikan hasil yang baik untuk memprediksi harga rumah di Tangerang Selatan. Namun perlu dipertimbangkan lagi dalam penggunaan arsitektur jaringan yang lebih baik. Keoptimalan model ditentukan parameter dan preprocessing data. Berdasarkan percobaan 2 bahwa dengan model 2 dengan parameternya memberikan hasil yang bagus untuk *recurrent* harga rumah untuk kedepannya.

5.2 Saran

Berdasarkan hasil penelitian, disarankan perlu mempertimbangkan penggunaan arsitektur jaringan yang lebih sesuai, seperti LSTM atau GRU, yang lebih cocok untuk tugas seperti prediksi harga rumah. Selain itu, penting untuk melakukan kembali *pre-processing* seperti *missing value* dihapus atau diisi dengan nilai selain *mean*, termasuk normalisasi yang selain *standar scaller* data dan pemrosesan atribut.

Evaluasi metrik yang sesuai juga diperlukan untuk memahami kualitas model. Metrik yang dicoba selain *binary crossentropy* seperti MAE (*Mean Absolute Error*), MSE (*Mean Squared Error*), atau metrik lain yang relevan dengan tugas prediksi harga rumah. Selain itu, diharapkan penelitian selanjutnya dapat mempertimbangkan penggunaan data yang lebih lengkap dan representatif untuk meningkatkan akurasi prediksi harga rumah di daerah Tangerang Selatan.

VI. Referensi

- [1] A. Saiful, "Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 41–50, 2021, doi: 10.35957/jatisi.v8i1.701.
- [2] M. Schuster, "Bidirectional recurrent neural networks," no. December 1997, 2016, doi: 10.1109/78.650093.
- [3] E. F. Rahayuningtyas, F. N. Rahayu, and ..., "Prediksi harga rumah menggunakan general regression neural network," *Jurnal pdfs.semanticscholar.org*, 2021. [Online]. Available: <https://pdfs.semanticscholar.org/8093/4ad739aa088f74823cff162c5a831c59f78a.pdf>
- [4] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," in *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, Nov. 1997, doi: 10.1109/78.650093.
- [5] Haqi, Moh. *Penerapan Chatbot Untuk Mendeteksi Kemiripan Gambar Asli Atau Manipulasi Menggunakan Metode Metadata Dan Error Level Analysis (Ela) Sebagai Komputasi Cnn*. Diss. Program Studi Teknik Informatika, 2020.
- [6] LINDA, SUNDARI. "PEMODELAN TIME SERIES UNTUK PERAMALAN SUHU UDARA MENGGUNAKAN METODE LONG SHORT TERM MEMORY (LSTM)(Studi Kasus: Stasiun Klimatologi Lampung)." 2023.