



PEMODELAN LONG SHORT TERM MEMORY (LSTM) UNTUK PRAKIRAAN PENJUALAN BERDASAR BASIS DATA PENJUALAN RETAIL PADA KONTROL PERSEDIAAN

Disusun Oleh :

Chamiyanti Nurkentjana Aju G651160724

Pembimbing :

1. Prof Dr Ir Agus Buono, Msi, Mkom
2. Dr Eng Annisa S. Kom, M. Kom



CONTENS

01 LATAR BELAKANG

02 METODE

03 HASIL DAN PEMBAHASAN

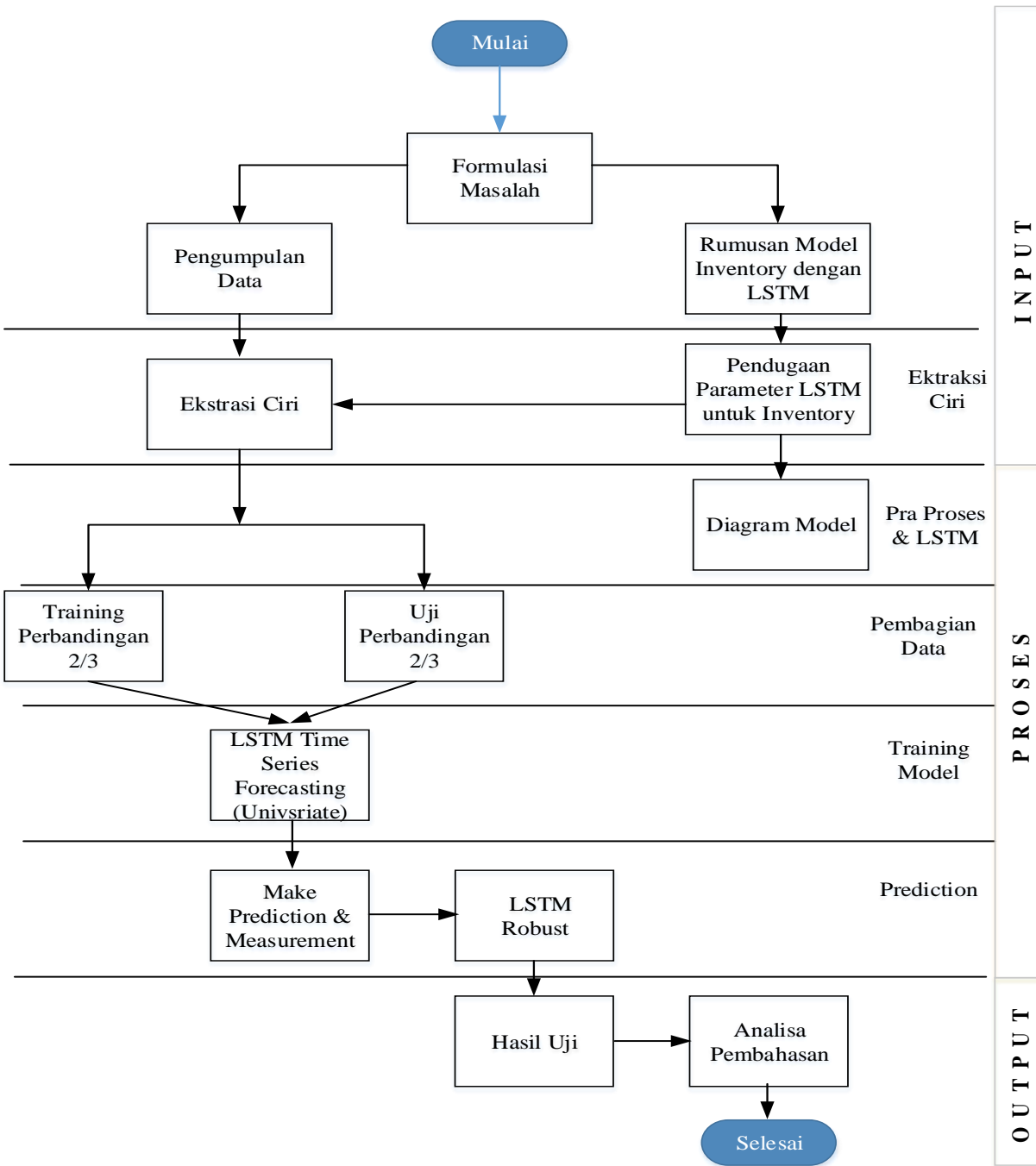
04 SIMPULAN DAN SARAN

LATAR BELAKANG



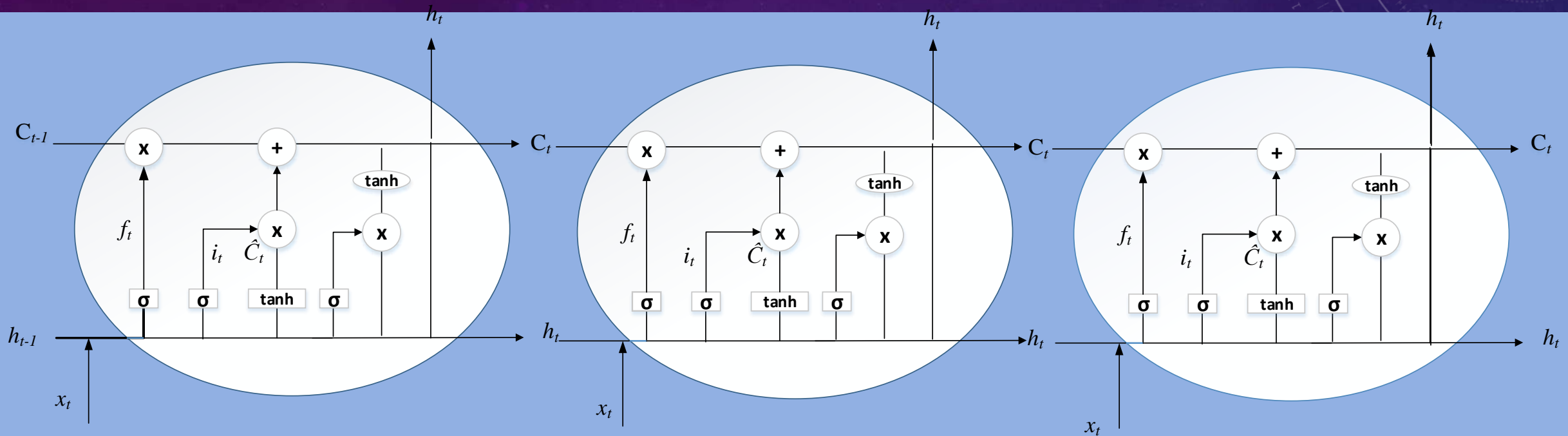
METODE

Tahapan Penelitian



METODE

LSTM

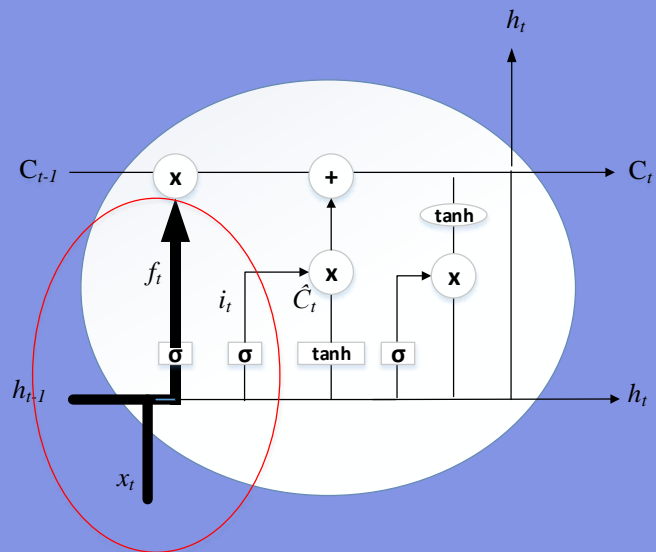


Sumber:

- S. Hochreiter and J. Schmidhuber, 1997.
- Felix A. Gers, Jürgen Schmidhuber, and Fred Cummins. 1999

METODE

LSTM



Gambar 1 (2)

Gambar 1 Persamaan (2) h_{t-1} dan x_t akan melewati gerbang Sigmoid pada (1). Langkah pertama LSTM memutuskan informasi apa yang akan dihapus dari *cell state*. Keputusan ini dibuat oleh sigmoid *layer* yang bernama “*forget gate layer*”. *Forget gate layer* akan memproses h_{t-1} dan x_t sebagai *input*, dan menghasilkan *output* berupa angka 0 atau 1 pada *cell state* C_{t-1} .

$$\sigma(x) = 1/(1 + \exp^{-x}) \quad (1)$$

Dimana: x = data input, e = konstanta matematika (2,718281828459045235360287471352)

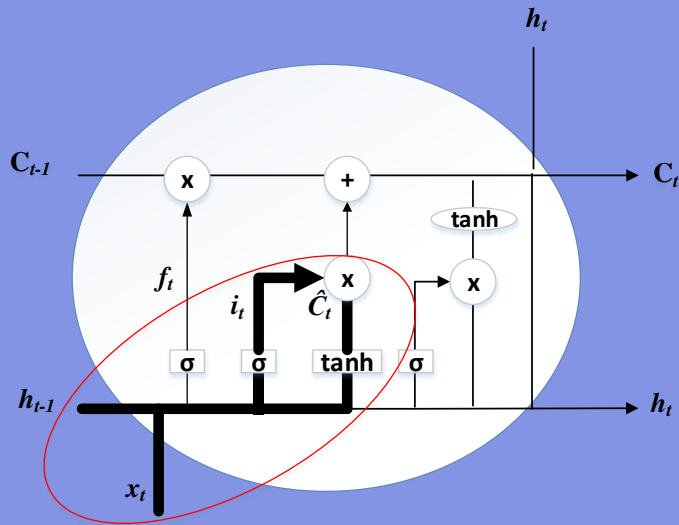
$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

Dimana *forget gate*, σ = fungsi sigmoid, W_f = nilai *weight* untuk *forget gate*, h_{t-1} = nilai *output* sebelum *orde* ke t , x_t = nilai input pada *orde* ke t , b_f = nilai bias pada *forget gate*
Nilai *weight* diuraikan pada persamaan (3)

$$W = (-1/d^{1/2}), 1/d^{1/2} \quad (3)$$

METODE

LSTM



Gambar 2 (4) dan (6)

Langkah selanjutnya adalah memutuskan informasi baru apa yang akan di simpan di *cell state* (4). Ada dua bagian. Pertama, lapisan sigmoid yang disebut "*input gate layer*" menentukan nilai mana yang akan diperbaharui. Selanjutnya, lapisan tanh membuat vektor nilai baru, \hat{C}_t , yang dapat ditambahkan ke *cell state* pada (6). Pada langkah selanjutnya, keduanya digabungkan untuk pembaharuan ke *cell state*.

$$i_t = \sigma. Wi [h_{t-1}, x_t] + b_i \quad (4)$$

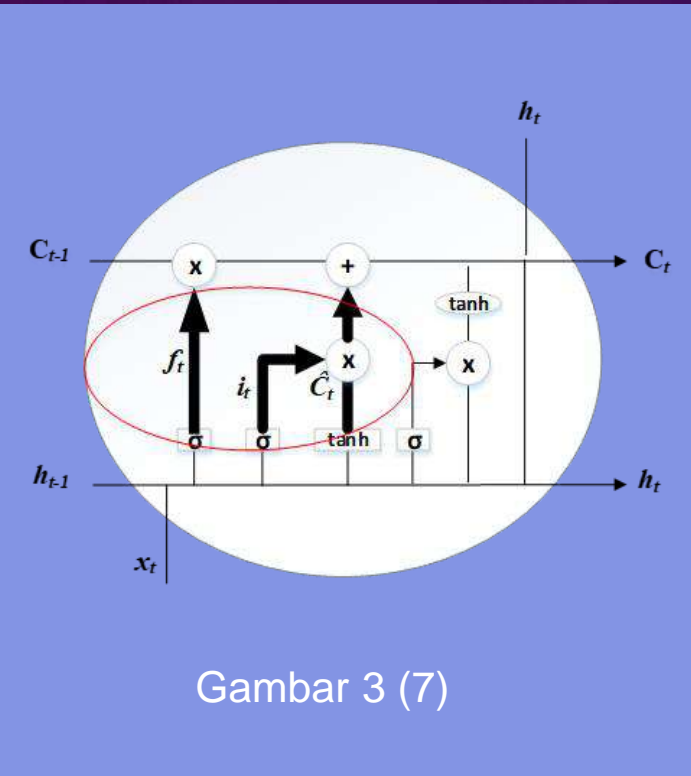
Dimana: i_t = *input gate*, σ = fungsi sigmoid, Wi = nilai *weight* untuk *input gate*, h_{t-1} = nilai output sebelum orde ke t , x_t = nilai input pada orde ke t , b_i = nilai bias pada *input gate*

$$\tanh(x) = 2\sigma(2x) - 1 \quad (5)$$

$$\hat{C}_t = \tanh (Wc .[h_{t-1}, x_t] + b_c) \quad (6)$$

Dimana: \hat{C}_t = nilai baru yang dapat ditambahkan ke *cell state*, \tanh = fungsi tanh, W_c = nilai *weight* untuk *cell state*, h_{t-1} = nilai output sebelum orde ke t , x_t = nilai input pada orde ke t , b_c = nilai bias untuk *cell state*

METODE LSTM



Gambar 3 (7)

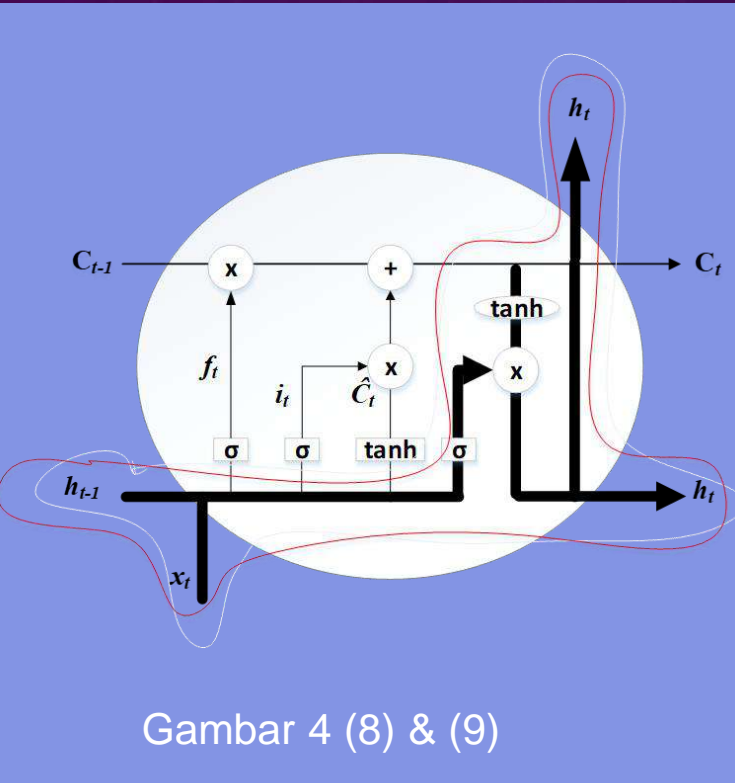
$$C_t = f_t \times C_{t-1} + i_t \times \hat{C}_t \quad (7)$$

Dimana: C_t = Cell state, f_t = forget gate, C_{t-1} = Cell state sebelum orde ke t, i_t = input gate, \hat{C}_t = nilai baru yang dapat ditambahkan ke cell state

Saatnya untuk memperbarui cell state lama (7), C_{t-1} , ke cell state baru C_t . Dengan mengalikan keadaan state lama dengan f_t , diputuskan untuk dilupakan pada *forget gate layer*. Lalu tambahkan $i_t * \hat{C}_t$. Ini adalah nilai baru, diskalakan dengan berapa banyak untuk diputuskan perbaharuan setiap cell state.

METODE

LSTM



Gambar 4 (8) & (9)

Langkah terakhir dari LSTM (8), output harus sesuai dengan *cell state* terlebih dahulu. Pertama menjalankan lapisan sigmoid yang menentukan bagian *cell state* mana yang menjadi *output* (8). Kemudian, menempatkan *cell state* melalui tanh (untuk mendorong nilai antara -1 dan 1) dan mengalikannya dengan output dari gerbang sigmoid, sehingga hanya mengeluarkan bagian-bagian yang diputuskan (9).

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = O_t \times \tanh(C_t) \quad (9)$$

Dimana: O_t = *output gate*, σ = fungsi sigmoid, W_o = nilai *weight* untuk *output gate*, h_{t-1} = nilai output sebelum orde ke t, x_t = nilai input pada orde ke t, b_o = nilai bias pada *output gate*, h_t = nilai output orde t, O_t = *output gate*. \tanh = fungsi tanh, C_t = *Cell state*

METODE



Root Mean Square Error (RMSE)

Rumus RMSE menurut (Montgomery *et al.* 2008) dapat dilihat pada (10).

$$\Sigma_t^n = (1((x_t - f_t))^2/n \quad (10)$$

Dimana x_t adalah nilai aktual pada waktu ke- t , f_t adalah nilai prediksi pada waktu ke- t n adalah jumlah data yang diprediksi.

Augmented Dicky Ad Fuller Test

- Augmented Dicky Ad Fuller Test adalah jenis uji statistik yang disebut Unit root test. Intuisi unit root test adalah menentukan seberapa kuat deret waktu ditentukan oleh tren.

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t \quad (11)$$

METODE

Rumusan Persediaan dari data penjualan



$$\alpha = (L \cdot A) / 365 \quad (12)$$

$$d = (n \mid v - \alpha \mid) / v \quad (13)$$

$$d = n - ((n \mid v - \alpha \mid) / \alpha) \quad (14)$$

$$d = n / 2 \quad (15)$$

A = Rata rata tahunan

L = Panjang rentang waktu dalam hari

n = skala (n=1 -> 0-1, n=5 -> 0-5)

v = jumlah yang terjual

Perhitungan sederhana dapat digunakan untuk memvalidasi persamaan ini.

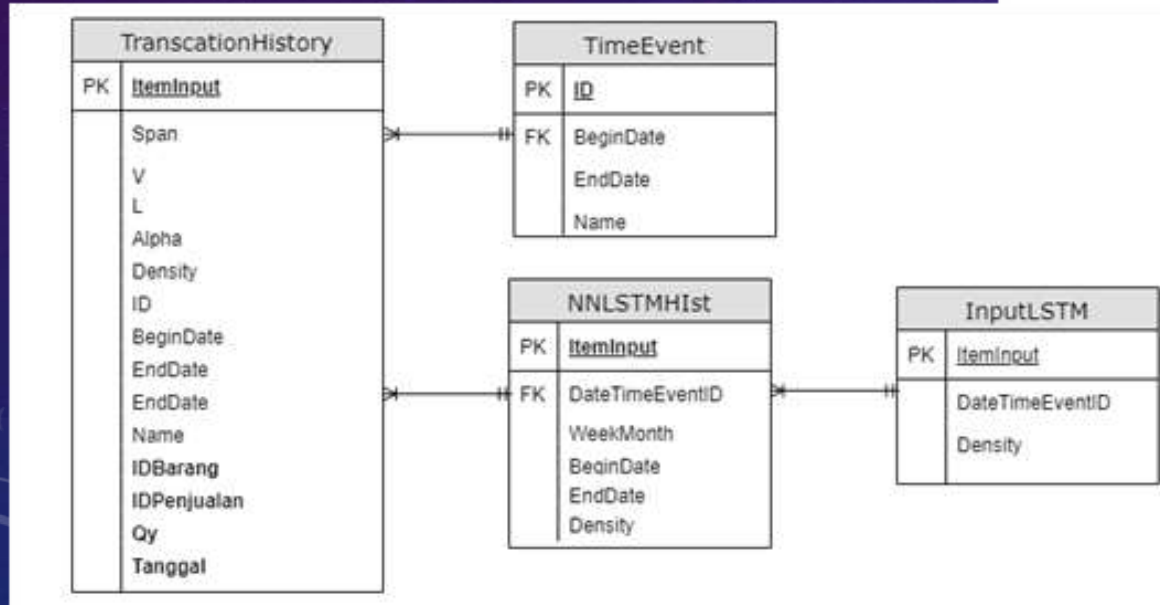
- Jika $(v) > (\alpha)$, karena v cenderung tak terhingga, nilai α cenderung diabaikan, menggunakan (13) cenderung ke n berarti berarti nilai tinggi.
- Jika $(v) < (\alpha)$, nilai v dapat diabaikan, menggunakan (14) yang cenderung 0 berarti nilai rendah.
- Jika $(v) = (\alpha)$, skala berada di tengah, yang membuatnya sama dengan $n/2$ menggunakan (15) berarti nilai sedang.

METODE

Parameter untuk prakiraan penjualan



Gambar Table Basis Data terkait



Gambar Table Basis Data terkait dengan LSTM

METODE: Pengumpulan Data

Pembagian Data



Basis Data	Tahun Data	Transaksi Penjualan	Prakiraan Vol. Penjualan	Observasi Data LSTM Network	Skala
Basis Data A	2003	Penjualan kendaraan bermotor (seperti motor x, mobil x)	Setiap kuantiti per produk 15 minggu ke depan	45 minggu dibagi menjadi 2 bagian training 1 bagian testing berarti pembagian data time series dibagi dengan 30 minggu data training dan 15 minggu data testing	1 dan 5
Basis Data B	2010-2013	Consumer goods (seperti perlengkapan rumah, buah buahan, pakaian, kosmetik),	Setiap kuantiti per kategori per 12 bulan kedepan	36 bulan dibagi menjadi 2 bagian training 1 bagian testing berarti	1 dan 5
Basis Data C	tahun 2011	Perlengkapan rumah (seperti lampu pojok, tempat tissue, pot meja)	Setiap kuantiti per produk dilakukan untuk per 15 minggu ke depan	45 minggu dibagi menjadi 2 bagian training 1 bagian testing berarti	1 dan 5

METODE

Perangkat Penelitian



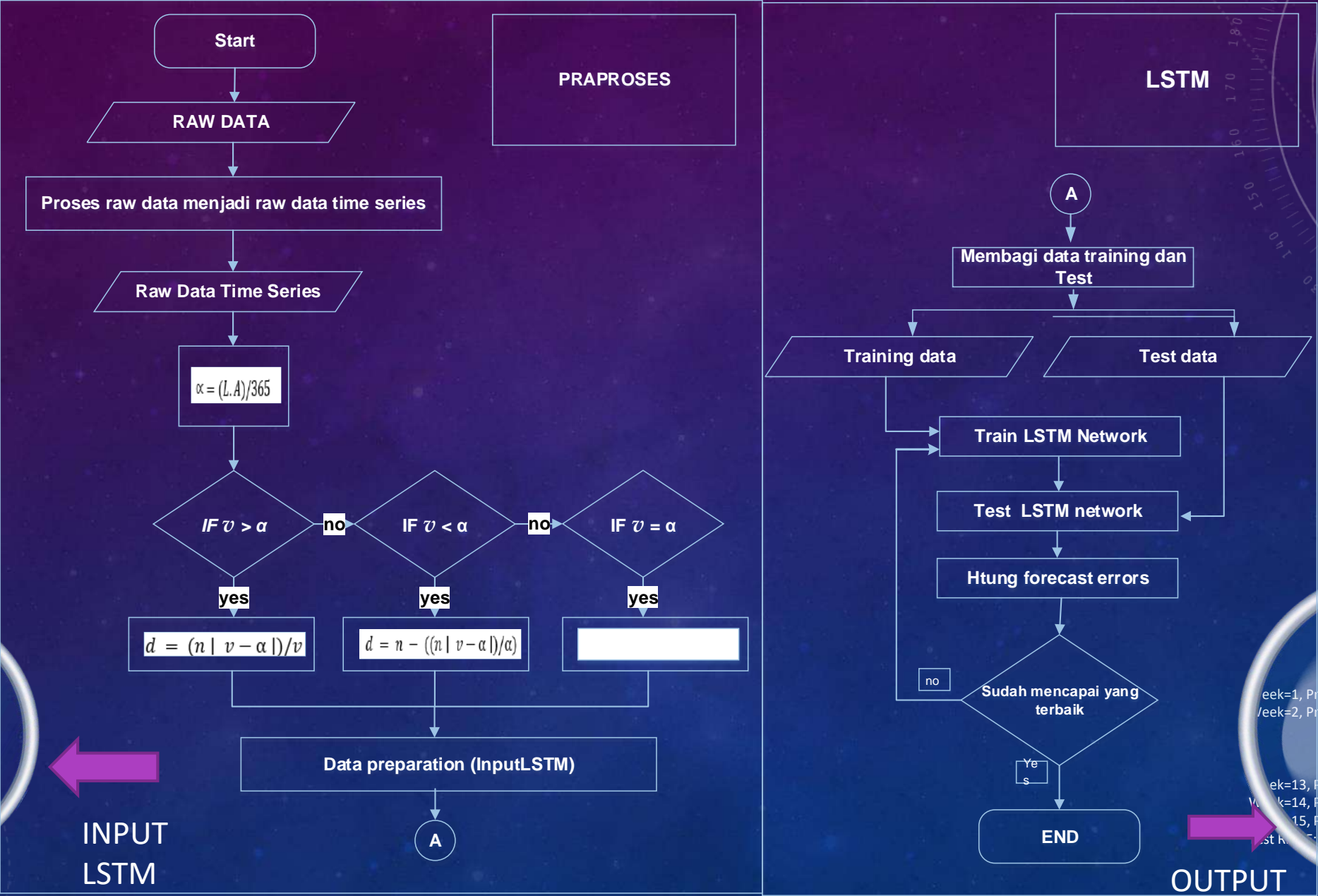
Perangkat keras yang digunakan dalam penelitian ini adalah computer personal dengan spesifikasi sebagai berikut:

- Prosesor : Intel® Core™ i7-3737U CPU @ 2.00GH (4CPUs), 2.0Ghz
- RAM : 16 GB
- Memori : 1GB

Perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut:

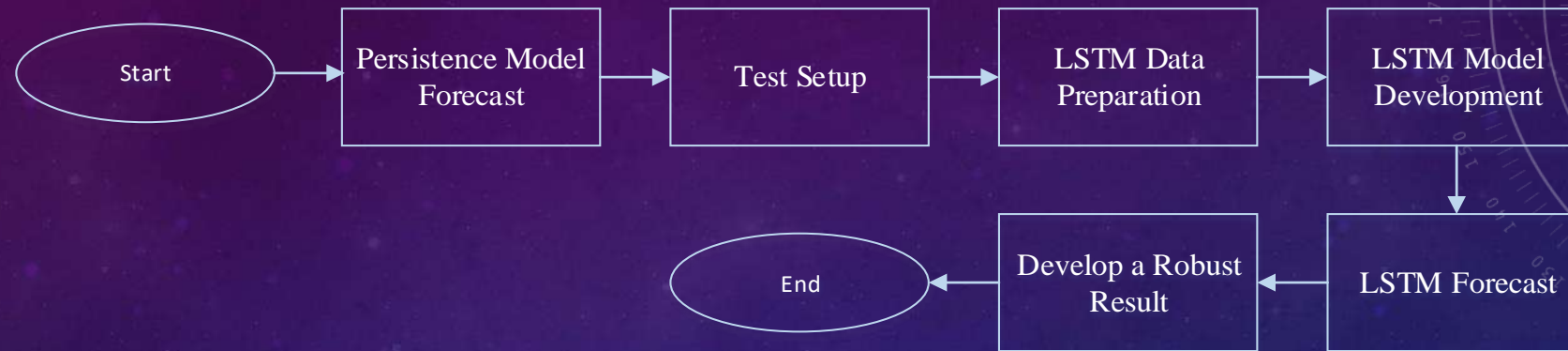
- Sistem operasi: Microsoft Windows Pro 10 (64-bit)
- SQL Server untuk praproses data
- Ms Excel untuk membaca dataset
- Anaconda 3, Jupiter Notebook 6.00, Bahasa pemrograman Python 3.7, *Package Keras* dan *Backend Tensorflow*

METODE: Diagram Model



Training Model

LSTMs untuk Univariat *Time Series*

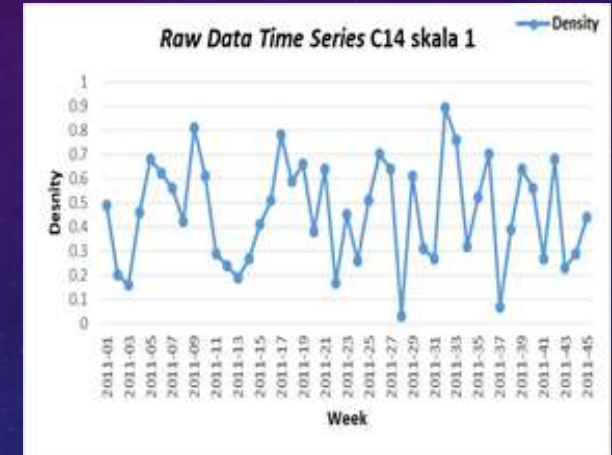
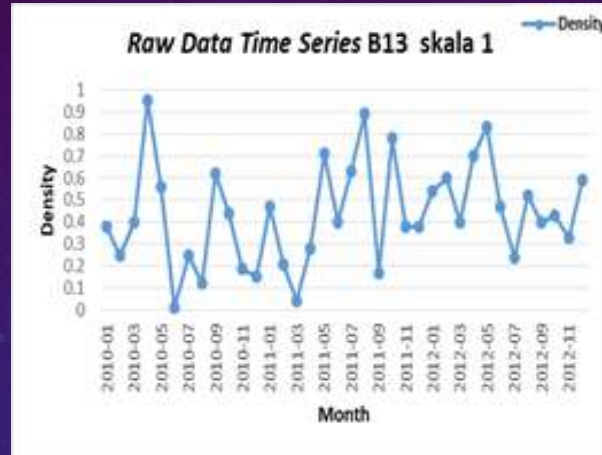
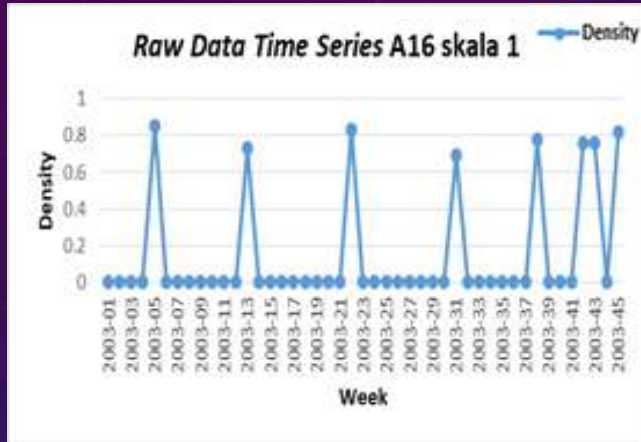


Jason Brownlee PhD

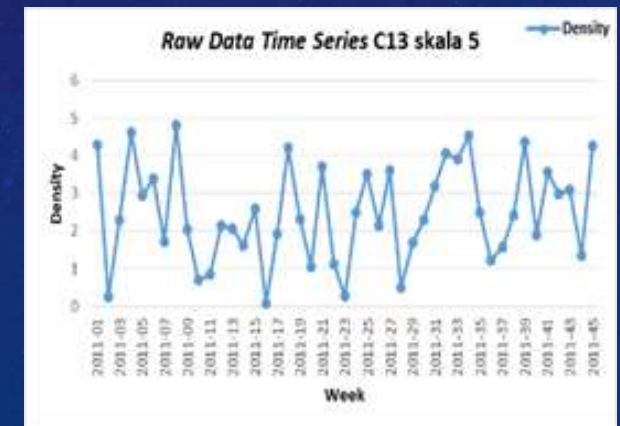
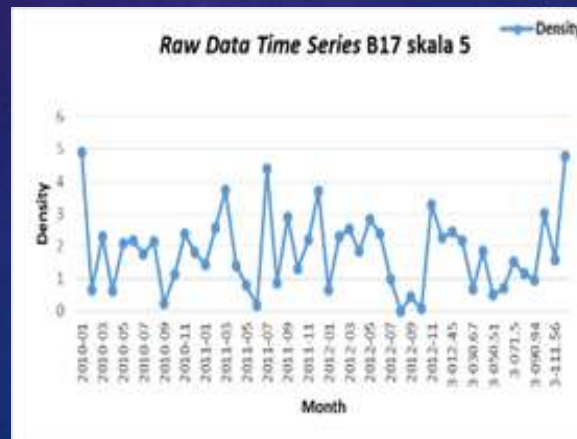
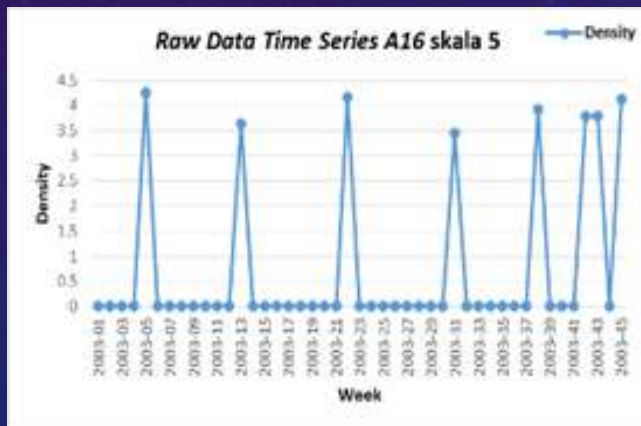
- **Persistence Model Forecast:** tempat pengamatan dari langkah waktu sebelumnya ($t-1$) digunakan untuk memprediksi pengamatan pada langkah waktu saat ini (t).
- **Test Setup:** Model meniru dari dunia nyata dimana membagi data train dan data uji pada pengamatan penjualan yang tersedia setiap bulan atau setiap minggu dan digunakan pada peramalan berikutnya.
- **LSTM data preparation :** Sebelum masuk ke dalam LSTM, data harus diubah terlebih dahulu dibagi menjadi 3 bagian:
 1. Ubah deret waktu menjadi masalah pembelajaran yang diawasi.
 2. Ubah data deret waktu sehingga stasioner.
 3. Ubah pengamatan menjadi skala spesifik.
- **LSTM Forecast:** Setelah model LSTM cocok dengan data pelatihan, dapat digunakan untuk membuat perkiraan. memprediksi setiap langkah waktu baru satu per satu dari data uji (the fixed approach)
- **Develop Robust:** Develop Robust adalah mengulangi percobaan dari bagian sebelumnya beberapa kali, kemudian menggunakan RMSE rata-rata sebagai indikasi seberapa baik konfigurasi akan dilakukan pada rata-rata data yang tidak terlihat.

Hasil dan Pembahasan

Raw Data Time Series



Gambar Contoh raw data penjualan A16, B13, C14 dengan skala 1



Gambar Contoh raw data penjualan A16, B17, C14 dengan skala 5

Hasil dan Pembahasan

Augmented Dickey–Fuller test dan Persistence Model Forecast per produk

- Contoh Augmented Dickey–Fuller test untuk C13 n-5 dibawah ini:

ADF Statistic: -6.44789515

p-value: 0.00000002,

Critical Values:

1%: -3.589

5%: -2.930

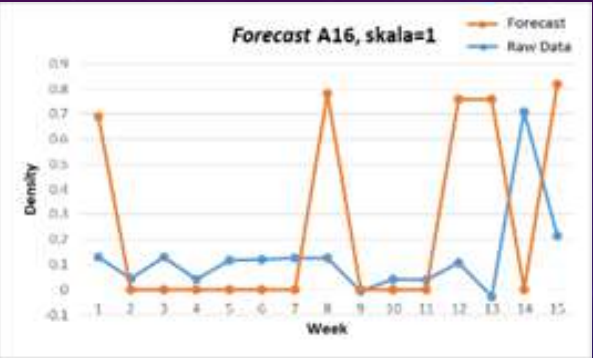
10%: -2.603

- Contoh RMSE Persistence RMSE: 1.507 untuk C13 n-5.



Hasil dan Pembahasan

Hasil forecast



RMSE A: 0.426
B: 0.189
C: 0.380

Gambar Contoh Forecast Data penjualan A16, B13, C14 dengan skala 1



RMSE A: 2.120
B: 1.220
C: 1.293

Gambar Contoh Forecast Data penjualan A16, B17, C13 dengan skala 5

Hasil dan Pembahasan

Hasil Robust LSTM



Table Nilai RMSE dari 3 basis data dengan skala 1

Basis Data A		Basis Data B		Basis Data C	
Produk	RMSE	Produk	RMSE	Produk	RMSE
A11	0.44	B11	0.37	C11	0.39
A12	0.56	B12	0.28	C12	0.33
A13	0.49	B13	0.24	C13	0.39
A14	0.43	B14	0.39	C14	0.30
A15	0.43	B15	0.32	C15	0.38
A16	0.43	B16	0.27	C16	0.32
A17	0.51	B17	0.42	C17	0.35
A18	0.45	B18	0.35	C18	0.35

Hasil dan Pembahasan

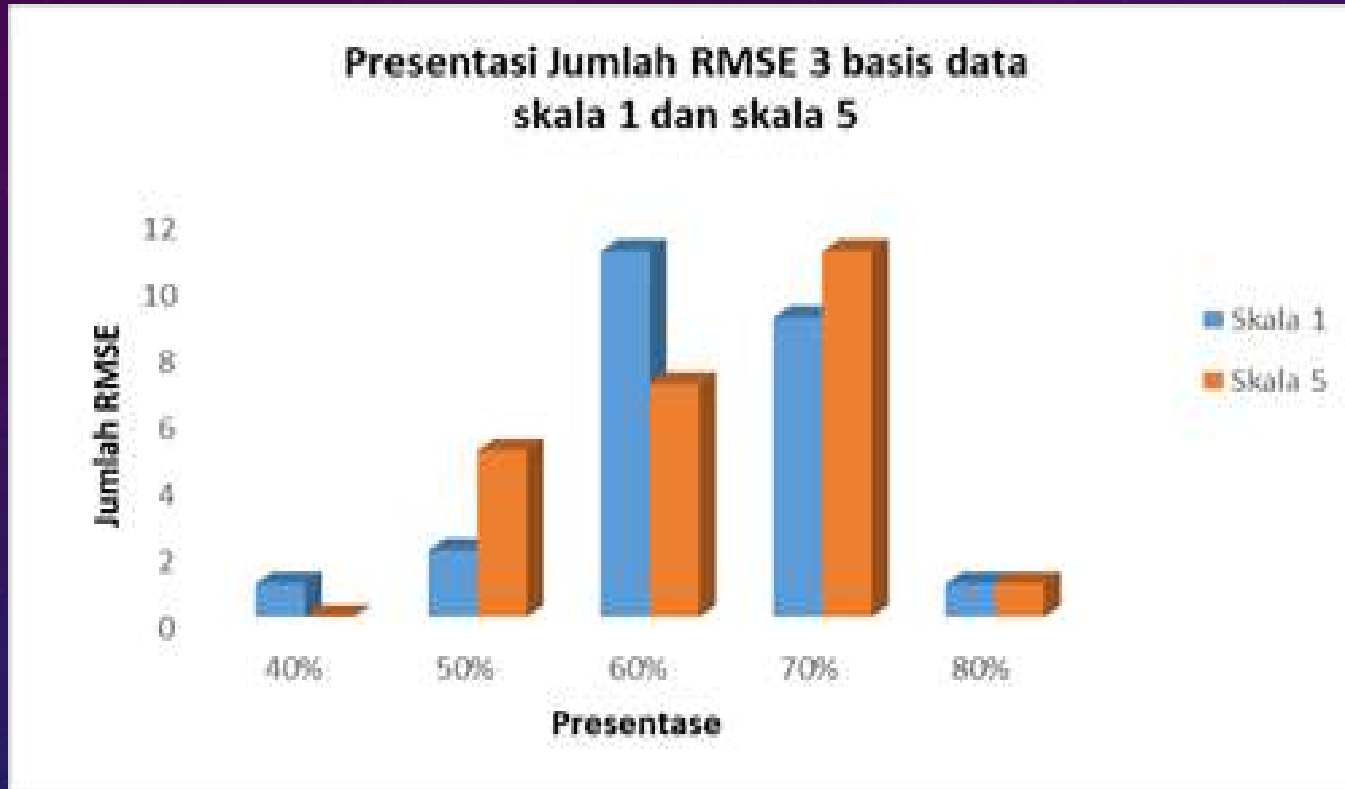
Hasil Robust LSTM



Tabel 3 Nilai RMSE dari 3 basis data dengan skala 5

Basis Data A		Basis Data B		Basis Data C	
Produk	RMSE	Produk	RMSE	Produk	RMSE
A11	2.10	B11	1.68	C11	1.93
A12	2.69	B12	1.47	C12	1.74
A13	2.39	B13	1.24	C13	1.36
A14	2.40	B14	2.07	C14	1.53
A15	2.22	B15	1.71	C15	1.86
A16	2.21	B16	1.38	C16	1.67
A17	2.58	B17	1.19	C17	1.74
A18	2.66	B18	1.72	C18	1.89

Hasil dan Pembahasan Robust LSTM



RMSE produk skala 1 banding skala 5 untuk 40% ada 1:0, 50% 2:5, 60% 11:7, 70% 9:11, 80% 1:1, total RMSE 24 produk.

Gambar Presentasi dari jumlah RMSE 3 basis data dengan skala 1 dan skala 5

KESIMPULAN DAN SARAN

Kesimpulan



Basis Data B Nilai RMSE cenderung relative lebih kecil

- 12 Bulan ke depan
- 36 bulan observasi
- Qty per Kategori
- Lebih banyak data musiman



Basis Data A dan C Nilai RMSE cenderung relative lebih besar

- 15 minggu ke depan
- 45 minggu observasi
- Qty per barang
- Sedikit data musiman



Basis Data A RMSE cenderung relative lebih besar dari B & C

- 15 minggu ke depan
- 45 minggu observasi
- Qty per barang
- Sedikit data musiman
- Data Jarang

Tanpa mempertimbangkan dampak musiman dan histori yang rendah setengah dari data, RMSE mencapai 70-80% Nilai skala 1 dan nilai skala 5 cenderung serupa sehingga untuk nilai skala bisa digunakan berapun nilainya

KESIMPULAN DAN SARAN

Saran

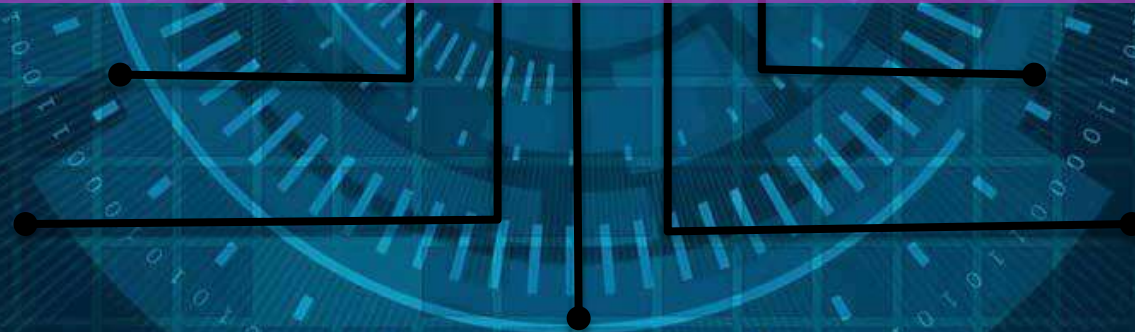


- LSTM network memerlukan histori yang kuat dan waktu musiman yang fluktuatif sebagai observasi pada data *time series* untuk hasil prakiraan yang baik.
- Secara keseluruhan LSTM mempunyai potensi yang baik pada data yang terbatas.

.



ALHAMDULILLAH



THANK YOU