



Perbandingan Kinerja LSTM dan Prophet untuk Prediksi Deret Waktu (Studi Kasus Produksi Susu Sapi Harian)

Alusyanti Primawati^{#1}, Imas Sukaesih Sitanggang^{#2}, Annisa^{#3}, Dewi Apri Astuti^{*4}

[#]Department of Computer Science, IPB University, Indonesia
Jl. Meranti Wing 20 Level 5 Kampus IPB Darmaga, Bogor 16680

¹1149alusyanti@apps.ipb.ac.id

²imas.sitanggang@apps.ipb.ac.id

³annisa@apps.ipb.ac.id

^{*}Department of INTP, Faculty of Animal Husbandry, IPB University, Indonesia
Jl. Agatis Kampus IPB Darmaga, Bogor 16680

⁴dewiapriastuti86@gmail.com

Abstrak— Prediksi deret waktu dibutuhkan untuk menjawab pertanyaan bisnis dimasa depan yang akurat sehingga perlunya membangun model prediksi yang memiliki kinerja bagus. Pendekatan *machine learning* seperti *long short term memory* (LSTM) dan Prophet menjadi populer saat ini untuk pemodelan prediksi deret waktu. Agribisnis susu segar saat ini salah satu studi kasus yang memerlukan peranan teknologi informasi seperti bisnis intelijen untuk memastikan ketersediaan pasokan susu dimasa depan. Upaya pertama yang perlu dilakukan adalah menyiapkan model prediksi yang tepat meskipun data awal yang dikumpulkan masih sedikit atau terbatas. Dataset produksi susu sapi selama 300 hari menjadi data penelitian yang dimodelkan kedalam LSTM dan Prophet. Keduanya dibandingkan kinerjanya terhadap data terbatas. Hasilnya uji koefisien determinasi R^2 keduanya yaitu 0.2, sehingga perlu dilakukan peningkatan kinerja melalui tahapan *revise and enhance*. Hasilnya, kedua model meningkat nilai R^2 menjadi 0.3 dan LSTM lebih baik dari Prophet. Meskipun demikian perbedaan keduanya tidak terlalu signifikan dan peningkatan juga tidak berbeda terlalu jauh karena data susu memiliki pola multi-periode dengan tren berbeda signifikan. Periode 90 hari pertama adalah masa klimaks laktasi sedangkan periode kedua setelah 90 hari adalah masa intervensi peternak menurunkan hasil perah untuk mempersiapkan ternak kambing perah ke masa kawin dan bunting.

Kata kunci— LSTM, Prediksi Deret Waktu, Prophet, Susu Sapi Harian

I. PENDAHULUAN

Prediksi deret waktu untuk peramalan saat ini menjadi penting, terutama jika dikaitkan dengan analitik prediktif dalam teknologi bisnis intelijen yang sedang populer saat ini. Pemodelan prediktif dalam bisnis intelijen akan menjawab pertanyaan bisnis dari perusahaan untuk

perencanaan [1]. Model prediktif bersamaan dengan forecasting merupakan salah satu dimensi utama dari bisnis analitik [2], [3]. Kinerja bisnis analitik dalam BI memiliki peranan penting dalam meningkatkan nilai bisnis [4]. Hasil peramalan digunakan untuk menggambarkan prediksi suatu nilai dimasa mendatang akan meningkat atau menurun. Ketika suatu nilai bisnis masa mendatang meningkat maka bisnis memiliki peluang keuntungan, sementara yang menurun akan mengakibatkan kerugian [5]. Dengan demikian terlihat pentingnya membangun model prediksi dengan kinerja bagus untuk perencanaan bisnis dengan tepat terkait data deret waktu. Perancangan model prediksi deret waktu berfokus pada penerapan algoritmanya [6]. Algoritma yang dipilih mampu membaca pola data deret waktu untuk peramalan pola data dimasa mendatang seperti pendekatan non-parametrik yaitu *machine learning*. Algoritma pendekatan *machine learning* yang saat ini populer dalam pemodelan prediksi deret waktu yaitu *long-short term memory* (LSTM) dan Prophet [6]. Prophet dianggap unggul terhadap data deret waktu yang memiliki tren musiman [7], sedangkan LSTM dianggap mampu mengatasi data multi-periode misalnya periode yang dibuat akibat munculnya *event* tertentu [8].

Kajian sudah dilakukan terdahulu terhadap kebutuhan agribisnis susu segar saat ini memerlukan peranan teknologi informasi untuk perencanaan ketersediaan pasokan susu masa mendatang [9]. Pekerjaan membangun teknologi bisnis intelijen merupakan perencanaan jangka panjang. Jika teknologi mulai dibangun, maka dataset deret nyata yang terkumpul masih terbatas. Oleh sebab itu perlu dikaji pemodelan prediksi yang tepat untuk data yang terbatas. Dataset produksi susu milik Righi et al berjumlah 300 hari merupakan dataset yang tepat sebagai eksperimen awal. Keunikan dataset susu sapi milik Righi, et al [10]

ditemukannya 2 pola berbeda yaitu masa 90 hari pertama masa produktif laktasi dan setelah 90 hari merupakan masa pengurangan pemerahan untuk mempersiapkan masa kawin dan bunting selanjutnya. Ini bisa menjadi dataset yang bagus yang dapat menunjukkan kemampuan model non-parametrik yang dikembangkan untuk memprediksi produksi susu sapi harian dengan jumlah dataset yang terbatas dan tren ganda tidak berulang dalam satu pola dataset.

Pemodelan prediksi produksi susu sapi paling populer dilakukan dengan pendekatan parametrik seperti *autoregressive integrated moving average* (ARIMA) [10], [11] dan algoritma lainnya [12]–[14]. Righi, et al, memodelkan prediksi susu sapi yang dikerjakannya menggunakan dataset deret waktu selama 300 hari [10]. Righie, et al sudah mencoba prediksi susu sapi menggunakan ARIMA terhadap hasil perhitungan EMA (*Exponential Moving Average*) menghasilkan akurasi 94,7% berdasarkan nilai MAE (*Mean Average Error*) [10]. Namun, prediksi tidak dilakukan langsung pada dataset, melainkan pada hasil perhitungan EMA. Ekperimen Righi tidak hanya menggunakan ARIMA, melainkan juga *artificial neural network* (ANN) dan *random forest* (RF). Hasilnya kinerja ketiganya tidak berbeda jauh secara signifikan, tetapi perbedaan utama dari aspek kecepatan kinerja bahwa ARIMA lebih cepat dibandingkan ANN dan RF. Namun, ANN dan RF bukanlah metode prediksi deret waktu melainkan metode machine learning dengan model klasifikasi.

Ada pendekatan *machine learning* yang memiliki kemampuan analisis prediktif menggunakan model regresi, seperti Long-Short Term Memory (LSTM) dan Prophet. Keduanya merupakan model prediksi deret waktu yang populer saat ini. LSTM memiliki keunggulan dalam memprediksi data nyata deret waktu [15], sedangkan Prophet unggul terhadap data historis yang memiliki efek musiman [16].

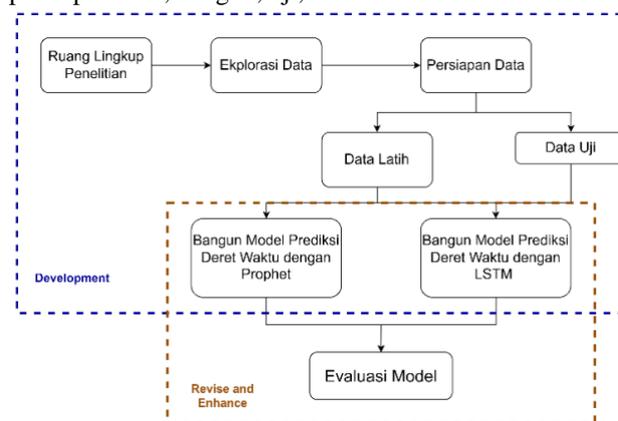
LSTM dan Prophet memiliki kemampuan untuk menangani karakteristik dataset produksi susu sapi milik Righi, et al [10] yang terbatas, univariate, non-linear dan tidak stasioner. Pemodelan dapat dilakukan terhadap data langsung. Dataset deret waktu produksi susu sapi dalam penelitian ini memiliki karakteristik data tidak stasioner, univariate, dan non-linear. Selain itu, berdasarkan hasil kajian pendahuluan yang kami lakukan sebelumnya bahwa kinerja LSTM dan Prophet belum ditemukan pernah dievaluasi dan dilakukan perbandingan kinerja keduanya untuk memprediksi data deret waktu produksi susu sapi, demikian juga dataset yang digunakan dalam penelitian ini [9]. Selanjutnya, hasil penelitian ini menjadi kontribusi untuk pengembangan penelitian selanjutnya untuk dataset yang serupa atau sejenis.

Singkatnya, penulis menyajikan hasil evaluasi kinerja model LSTM dan Prophet yang dibangun terhadap 300 dataset deret waktu produksi susu sapi milik Righi et al [10] digunakan untuk memilih model prediksi terbaik dari keduanya terhadap data yang serupa yang akan dimasukkan kedalam bagian teknologi bisnis intelijen pada agribisnis susu. Namun, hasil evaluasi bukan untuk dibandingkan terhadap hasil penelitian sebelumnya karena model

prediksi yang dibangun bukan dari hasil perhitungan EMA melainkan data original yang digabungkan dalam satu dataset. Sisa dari makalah ini disusun sebagai berikut: bagian 2 menjelaskan metode penelitian yang digunakan untuk membangun model prediksi deret waktu dengan menggunakan data deret waktu univariat yaitu susu sapi dengan horizon 300 hari. Pengaturan eksperimen, dataset dan hasil evaluasi dibahas di bagian 3. Terakhir, bagian 4 menyimpulkan makalah termasuk rekomendasi masa depan.

II. METODE PENELITIAN

Proses penelitian digambarkan keseluruhan kedalam Gambar 1, yang diambil sebagian dari metodologi proyek pemodelan prediktif [17]. Tahapan penelitian memiliki 5 proses yaitu cakupan dan definisi proyek, eksplorasi data, persiapan data, bangun, uji, dan evaluasi model.



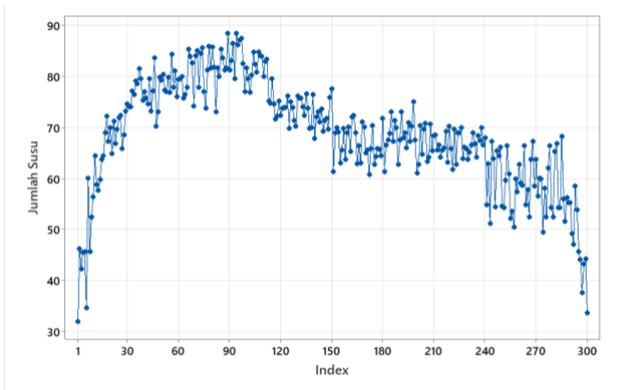
Gambar.1 Tahapan penelitian

A. Ruang Lingkup Penelitian

Cakupan proyek penelitian ini adalah membangun model terbaik untuk prediksi deret waktu univariat pada kasus peramalan produksi susu sapi harian selama 300 hari. Data produksi susu sapi yang digunakan dalam penelitian adalah dataset milik Righi, et al dan bersifat publik [10]. Ada 3 dataset produksi susu sapi dari 3 sapi yang berbeda selama 300 hari. Dalam penelitian ini, ketiga dataset digabungkan sebagai produksi susu sapi yang dihasilkan oleh satu peternakan. Hasil gabungan ketiga data susu sapi dimodelkan kedalam algoritma prediksi deret waktu LSTM dan Prophet. Penelitian ini dibatasi dengan tidak mengkaji kinerja algoritma berdasarkan kecepatan analisis, melainkan kinerja model berdasarkan evaluasi model baik atau buruk.

B. Eksplorasi Data

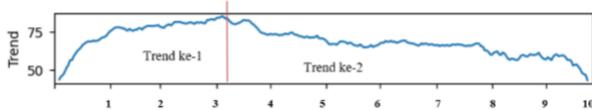
Proses untuk eksplorasi dataset menggunakan analisis statistika dasar dan visualisasi plot deret waktu. Hasil statistika dasar yaitu nilai mean dari dataset sebesar 68.727, sedangkan nilai minimumnya 31.9 dan maksimumnya 88.5. Nilai kuartil 1 (Q1) diperoleh 63.8 dan kuartil 3 (Q3) diperoleh 76.075 sehingga nilai median yang dihasilkan adalah 69.45. Visualisasi dataset dapat dilihat pada Gambar 2.



Gambar 2. Visualisasi plot dataset susu sapi gabungan 3 ekor sapi dalam 1 periode laktasi

Karakteristik dataset penelitian memiliki keunikan yaitu ditemukannya multi periode dengan 2 tren yang berbeda yaitu tren menaik dan menurun. Visualisasi plot data deret waktu dapat diinterpretasikan bahwa terdapat 2 pola tren. Hal ini membuktikan adanya tren produksi susu pada pemerahan 90 hari pertama dan setelah 90 hari. Produksi susu sapi meningkat pada 90 hari pertama masa laktasi. Namun, setelah periode 90 hari, sapi akan dipersiapkan untuk masuk masa kering. Masa kering terjadi saat sapi masuk masa kawin dan bunting selanjutnya. Masa kering dilakukan untuk meningkatkan produksi susu pada periode laktasi selanjutnya. Persiapan masuk masa kering dilakukan dengan menurunkan intensitas jumlah produksi susu selama pemerahan. Oleh sebab itu, tren setelah 90 hari memiliki tren menurun.

Tidak hanya memahami dataset dari visualisasi plot deret waktu tetapi juga dilakukan analisis berdasarkan dekomposisi deret waktu dari dataset yang disajikan dalam Gambar 3. Gambar 3 merupakan dekonstruksi data deret waktu produksi susu sapi 300 hari berdasarkan tren, seasonal, dan residual. Fokus pada tren sangat jelas ada 2 pola tren menaik dan menurun dan setiap pola hanya ada sample masing-masing 1 pola. Data deret waktu dalam penelitian tidak memiliki tren seasonal.



Gambar.3 Dekomposisi data deret waktu produksi susu sapi perah 300 hari dalam penelitian

Hasil eksplorasi data jika dibandingkan dengan dataset sapi perah selama 6 tahun [15], maka dataset yang digunakan adalah dataset jangka pendek hanya 1 tahun sehingga pola tren sesungguhnya belum terlihat jelas. Hal ini menjadi tantangan penelitian ini yaitu membuat model prediksi non-parametrik dari dataset dengan volume sedikit karena *machine learning* memiliki kinerja baik, apabila dataset bervolume besar.

C. Persiapan Data

Sebelum model dibangun perlu dilakukan tahap *pre-processing* untuk menghasilkan data yang bersih dan model dengan akurasi baik [18]. Tahap pertama, dataset yang tersedia memiliki 2 kolom yaitu kolom hari dan jumlah

susu. Sebelum membangun model prediksi deret waktu, bentuk data dipersiapkan dari kolom hari sebagai index berbentuk *date time*. Namun kolom hari belum berbentuk *date time*, sehingga kolom hari ditransformasikan menjadi *index date time*. Pemilihan *date time* dimulai 1 Januari 2020 sampai dengan 27 Oktober 2020 karena penelitian Righi dipublikasikan pada tahun 2020. Dataset tidak memiliki nilai null/ NaN/spasi sehingga tidak perlu ada data yang dihilangkan. Tahap kedua, data dibagi menjadi 2 yaitu data latih dan data uji. Data latih berjumlah 260 data awal dan data uji berjumlah 40 data terakhir.

D. Bangun Model Prediksi Deret Nyata

Model peramalan dibangun menggunakan alat bantu bahasa pemrograman python. Editor yang digunakan adalah Jupyter Notebook. Namun, khusus pendekatan parametrik juga menggunakan alat bantu dengan aplikasi analisis statistik yaitu MiniTab. Dataset yang digunakan untuk membangun model awal adalah dataset penuh yang memiliki 2 pola tren berbeda.

Pemodelan dilakukan dengan 2 skenario yaitu pemodelan parameter default dan *hyperparameter tuning* dengan *grid search* [19]. Teknik pengaturan parameter masing-masing algoritma diperlukan untuk memperoleh akurasi yang lebih baik (Tabel I).

TABEL I
PENGATURAN PARAMETER MODEL PREDIKSI YANG DIGUNAKAN DALAM PENELITIAN

No	Algoritma	Teknik Tuning Parameter
1	LSTM	Grid search (hidden layer=1-4; jumlah neuron=50,100; dropout=0.2,0.3; epoch=50,100; batch size=32,64)
2	Prophet	Grid search (<i>changepoint_prior_scale</i> dan <i>seasonality_prior_scale</i>)

Terakhir, proses perhitungan evaluasi berdasarkan nilai MAPE, RMSE dan MSE yang terkecil serta uji koefisien determinasi (R^2) yang mendekati 1 akan dibandingkan untuk pemilihan model terbaik.

E. Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja model berdasarkan kinerjanya. Kinerja model diukur dengan menggunakan hasil hitungan *Root Mean Square Error* (RMSE), *Mean Absolute Percentage Error* (MAPE) dan nilai uji koefisien determinan *R squared* (R^2). RMSE membandingkan nilai prediksi dan nilai sebenarnya. Jika RMSE semakin dekat nilai yang diprediksi dan amati, maka model semakin baik kinerjanya [5]. Tidak cukup mengukur dengan nilai RMSE, tetapi juga memerlukan nilai MAPE dan R^2 . MAPE digunakan untuk mengukur persentase kesalahan metode peramalan [20]. Jika nilai MAPE kurang dari 20% maka kemampuan model peramalan baik. Selain mengukur error, model diukur kecocokan nilai varian antara nilai prediksi dengan nilai sebenarnya dengan uji koefisien determinan (R^2). Meskipun pada umumnya evaluasi model prediksi menggunakan RMSE dan MAPE, tetapi R^2 lebih disarankan karena range penilaiannya jelas yaitu antara 0 dan 1[21]. Jika model memiliki nilai mendekati 1 maka determinan varian antara nilai prediksi dan sebenarnya

mendekati sesuai. Namun, ada beberapa kasus nilai R^2 menghasilkan nilai kurang 0 atau negatif, hal itu dikarena model yang dihasilkan sangat buruk [22]. Referensi [21] digunakan sebagai referensi perhitungan nilai RMSE, MAPE, dan R^2 , dimana variabel input x_i sebagai nilai aktual, sedangkan y_i adalah nilai prediksi dan m adalah jumlah total dataset.

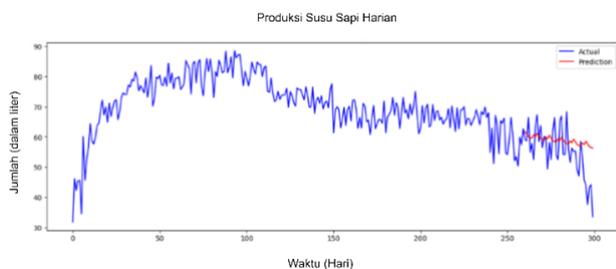
III. HASIL DAN PEMBAHASAN

Hasil dan pembahasan evaluasi pemodelan prediksi deret waktu pada dataset susu sapi harian menggunakan LSTM dan Prophet dijelaskan pada bagian ini. Dataset susu sapi harian dimasukan secara independen kedalam masing-masing model yang dioptimalkan dengan *hyperparameter tuning* menggunakan teknik *grid search*.

A. Model Prediksi Deret Waktu dengan Prophet

Model prediksi Prophet adalah model prediksi berbasis jaringan syaraf tiruan yang populer dan dirancang untuk peramalan otomatis data deret waktu [23]. Tidak hanya itu, prophet juga memiliki keunggulan mampu mendeteksi titik perubahan dalam dataset dan model dapat beradaptasi dengan cepat terhadap fitur musiman [24], [25].

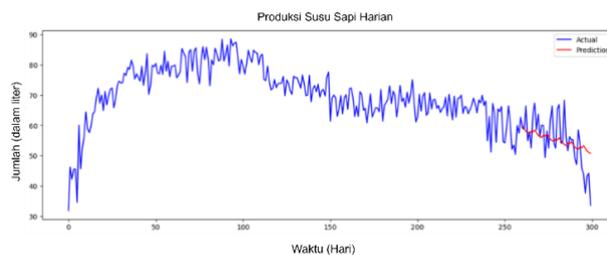
Model prophet dibangun dalam penelitian dengan 2 skenario yaitu model dengan parameter default (*changeoint_prior_scale* = 0.05 dan *seasonility_prior_scale* = 10.0) dan model dengan proses *hyperparameter tuning*. Parameter *changeoint_prior_scale* dan *seasonility_prior_scale* perlu diatur untuk meningkatkan kinerja prophet. Metode *grid search* digunakan untuk mencari parameter terbaik. Skala parameter *changeoint_prior_scale* yang digunakan yaitu 0.001, 0.01, 0.1, dan 0.5, sedangkan *seasonility_prior_scale* yaitu 0.01, 0.1, 1.0, dan 10.0. Tahap pertama yaitu memodelkan dataset dengan prophet dengan parameter default, hasilnya ditampilkan pada Gambar 4.



Gambar. 4 Plot prediksi model prophet

Tahap selanjutnya, validasi model dengan menggunakan *cross validation*. Hasil analisis *changeoint* membantu menetapkan jumlah data latih yang akan divalidasi dari model yaitu initial 160 hari, periode 40 hari dan horizon 100 hari. Hasil evaluasi model prophet skenario 1 yaitu RMSE 8,117; MAPE 8,7% dan R^2 sebesar -0,037. Model prophet skenario 2 mulai dibangun dengan memilih parameter terbaik yang dihasilkan *grid search*. Parameter terbaik *changeoint_prior_scale* yang terpilih adalah 0.5 dan *seasonility_prior_scale* yang terpilih adalah 0.01. Hasil evaluasi model prophet skenario 2 yaitu RMSE

6,813; MAPE 1.5% dan R^2 sebesar 0,0269. *Changeoint* yang ditangkap dari model ke 2 dijelaskan dalam plot prediksi model prophet skenario 2 (Gambar 5).



Gambar.5 Plot hasil prediksi deret waktu prophet dengan *grid search*

B. Model Prediksi Deret Waktu Long Short Term Memory (LSTM)

LSTM populer digunakan untuk memprediksi data deret waktu *real time* dan beberapa penelitian menghasilkan kinerja LSTM lebih baik dari yang algoritma lain karena kemampuannya menangani informasi yang tidak relevan dari data masa lalu [5]. Namun, pengaturan parameter saat membangun model LSTM merupakan hal yang perlu diperhatikan untuk membangun model yang baik.

Persiapan awal yang dilakukan adalah normalisasi dataset, membagi dataset menjadi data latih dan data uji, *shaping data* latih berdasarkan jumlah histori langkah waktu yang diamati (*window size*), dan menentukan parameter model. Kesulitan awal yang dialami adalah menentukan *shaping data* latih. Setelah diuji beberapa nilai, peneliti menggunakan *window size*= 100, maka *shape data* latih yang digunakan adalah 160,100,1. Data uji yang digunakan untuk validasi dan evaluasi model adalah 40 data terakhir dari 300 data.

Model prediksi LSTM yang dibangun menggunakan 2 skenario yaitu model LSTM dengan parameter *default* dan model yang dibangun melalui proses *hyperparameter tuning* dengan *grid search*. Parameter *default* pada model LSTM pertama yaitu 50 neuron, dropout 0.2, batch size 64 dan epoch sebanyak 20. Hasil evaluasi model pertama diperoleh RMSE 9,050; MAPE 12% dan R^2 senilai -0,290. Model skenario ke 2 dibangun dengan mencari parameter terbaik dari range parameter yang dimasukan kedalam *grid search*. Nilai parameter yang digunakan yaitu neuron: 16,32,50,100; batchsize: 32, 64; dropout: 0,2 dan 0,3; epcho: 10,20,30,40,50,100; dan validasi split: 0.2 dan 0.3. Total simulasi yang lakukan ada 12 disajikan dalam Tabel II. Hasil dari semua simulasi dijelaskan dalam Tabel III.

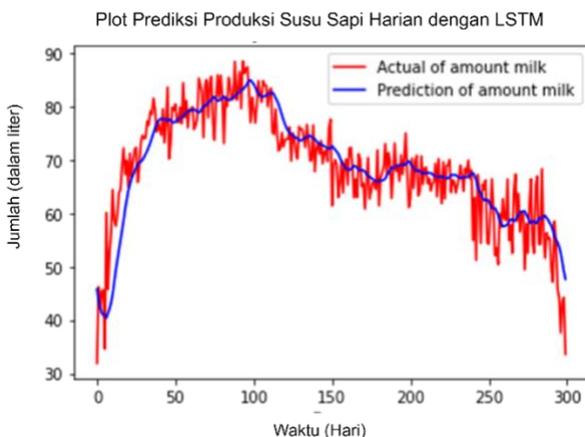
TABEL. II
DESKRIPSI SIMULASI TUNING PARAMETER LSTM

Simulasi	Tunning Parameter LSTM				
	Epoch	Val_split	Neuron	Dropout	Batch size
1	20	0.3	100	0.3	64
2	20	0.2	16	0.3	32
3	10	0.3	50	0.3	64
4	10	0.2	50	0.2	64
5	30	0.3	50	0.3	64
6	30	0.2	50	0.2	64
7	40	0.3	100	0.3	64
8	40	0.2	100	0.2	32
9	50	0.3	100	0.3	64
10	50	0.2	50	0.2	64
11	100	0.3	50	0.3	64
12	100	0.2	32	0.3	32

TABEL. III
EVALUASI MODEL UNTUK SELEKSI PARAMETER TERBAIK MENGGUNAKAN GRID SEARCH

Simulasi	Ukuran Evaluasi Model		
	RMSE	MAPE	R ²
1	7,122	0,106	0,201
2	8,908	0,149	-0,250
3	7,323	0,109	0,156
4	7,161	0,111	0,192
5	9,509	0,161	-0,424
6	10,394	0,177	-0,701
7	8,428	0,142	-0,118
8	7,914	0,132	0,014
9	7,225	0,106	0,178
10	8,053	0,133	-0,021
11	10,135	0,173	-0,617
12	10,426	0,177	-0,712

Parameter terbaik yang terpilih yaitu model yang memiliki nilai R² yang mendekati 1. Simulasi 1 merupakan model terbaik yang dibangun berdasarkan parameter pilihan nilai R² sebesar 0,201 disertai RMSE 7,122 dan MAPE 10,6%. Hasil prediksi deret waktu dengan model LSTM skenario ke-2 disajikan kedalam Gambar 6.



Gambar. 6 Plot hasil prediksi deret waktu model LSTM dengan grid search.

C. Pembahasan Hasil Evaluasi Model Awal

Evaluasi kinerja semua model diringkas dalam Tabel IV. Model prediksi yang memiliki kinerja lebih baik yaitu Prophet dan LSTM setelah dilakukan parameter tuning dengan grid search. Hasil kinerja keduanya belum mampu

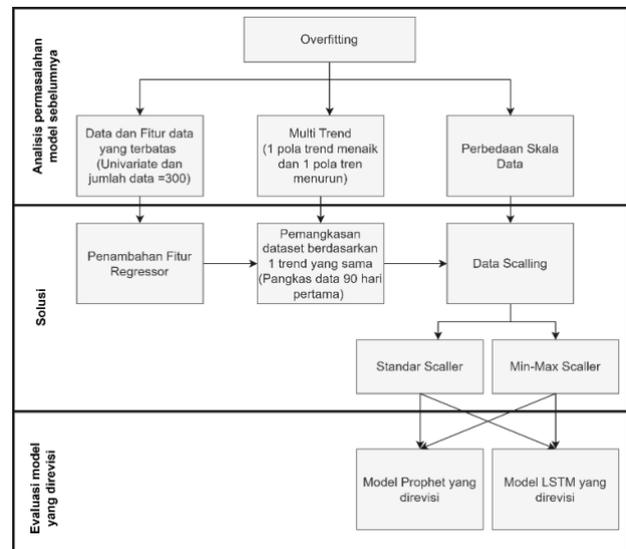
memprediksi mendekati data aktual, meskipun nilai error keduanya kecil. Oleh sebab itu keduanya hanya memiliki R² sekitar 0.2. Pada umumnya model prediksi yang baik memiliki R² yang mendekati 1, sedangkan R² dari model prophet dan LSTM masih jauh dari 1. Dengan demikian, pekerjaan dilanjutkan ke tahapan *revise and enhance*. Tahapan ini dilakukan untuk meningkatkan kinerja model yang sebelumnya.

TABEL. IV
HASIL EVALUASI KINERJA MODEL

Evaluasi	LSTM	LSTM+GS	Prophet	Prophet+GS
MSE	81,905	50,721	65,879	46,421
RMSE	9,050	7,122	8,117	6,813
MAPE	0,126	0,002	0,087	0,015
R ²	-0,290	0,201	-0,037	0,269

D. Tahapan Revise and Enhance Model

Tahapan *revise and enhance* yang dilakukan yaitu analisis masalah pemodelan sebelumnya dan mencari cara penanganannya (lihat Gambar 7). Tujuan tahapan ini adalah meningkatkan kinerja model Prophet dan LSTM

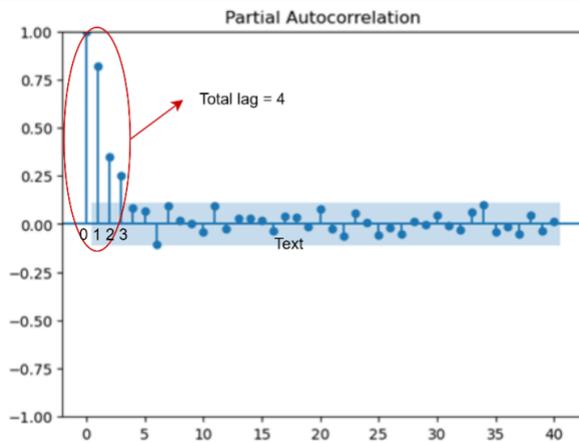


Gambar.7 Tahapan revise and enhance model

1) *Analisis Overfitting*: hasil evaluasi model prophet memiliki R² sebesar 0.269 lebih besar dibandingkan LSTM. Namun, kedua nilai R² dari Prophet dan LSTM masih tergolong kecil. Harapannya model prediksi yang dibangun memiliki R² yang mendekati 1. Peneliti mencoba membuat skenario untuk mengoptimalkan kinerja prophet dan LSTM. Tahapan pertama yang dilakukan adalah menganalisis *underfitting* dan *overfitting* pada model. Hasil pengujian model mengalami *overfitting* yaitu *error* pada data uji (RMSE 6.8) lebih tinggi daripada data latih (RMSE 3.7). Tidak hanya permasalahan penyetelan parameter, tetapi juga *overfitting* menjadi salah satu kelemahan model *non-linear* [20]. Dengan demikian, optimasi akan dilakukan dengan menangani *overfitting*. *Overfitting* diduga terjadi karena 3 permasalahan. Pertama, kumpulan data yang terbatas sehingga belum cukup untuk dimodelkan. Keterbatasan data seperti fitur prediksi hanya 1 yaitu jumlah susu harian, sedangkan kumpulan data

hanya berjumlah 300 data (belum 1 tahun). Kedua, dataset memiliki multi periode dengan pola tren yang berbeda dalam satu dataset, dan kedua pola hanya terjadi 1 kali (tidak berulang) karena data yang terbatas jumlahnya. Kumpulan data yang terbatas menyebabkan tidak cukup merepresentasikan tren dengan baik. Ketiga, model yang dibangun sebelumnya juga belum dilakukan *data scaling*.

2) *Penanganan Overfitting dengan Penambahan Regressor*: penanganan overfitting dilakukan dengan mempertimbangkan penambahan fitur karena fitur yang terdapat dalam dataset hanya ada 1 yaitu jumlah susu. Penambahan fitur dilakukan dengan menambahkan fitur regressor. Jumlah fitur regressor yang ditambahkan 4 berdasarkan jumlah lag yang berhasil ditemukan dari hasil analisis PCF dataset (lihat Gambar 8).



Gambar.8 Analisis PCF dataset

Gambar 8 menunjukkan bahwa setelah lag ke-3 ($k=0,1,2,3$), nilai PCF mendekati 0 (total lag = 4). Dengan demikian, peneliti ditetapkan 4 lag akan ditambahkan sebagai fitur regressor yaitu $y-1, y-2, y-3$, dan $y-4$ untuk memprediksi y (lihat Tabel V). Pada LSTM penambahan fitur regressor dilakukan dengan menambahkan input sequence sebanyak 4 dan reshape dataset pada data uji dan data latih menggunakan $time\ steps = 4$ sebagai panjang sample sebagai urutan sinyal dari setiap fitur regressor tambahan ($t-1, t-2, t-3$, dan $t-4$) sebagai variable input untuk memprediksi 1 waktu mendatang yaitu y (lihat Tabel VI). Penambahan fitur regressor menggunakan teknik *reframe* dengan teknik *lock_back* dan menyetel *time_steps* dari *convert* dataset deret waktu ke *supervised learning* sebanyak 4.

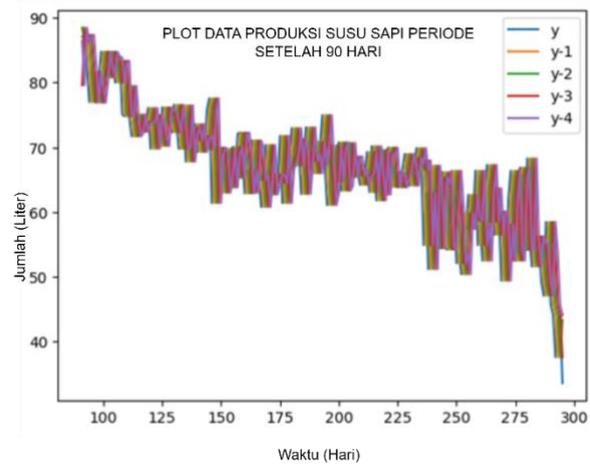
TABEL.V
DESKRIPSI DATASET DENGAN PENAMBAHAN 4 FITUR REGRESSOR PADA PROPHET

ds	y	y-1	y-2	y-3	y-4
2022-01-05	45,6	45,5	42,3	46,2	31,9
2022-01-06	34,6	45,6	45,5	42,3	46,2
2022-01-07	60,1	34,6	45,6	45,5	42,3
2022-01-08	45,7	60,1	34,6	45,6	45,5
2022-01-09	52,5	45,7	60,1	34,5	45,6

TABEL.VI
DESKRIPSI DATASET DENGAN PENAMBAHAN 4 FITUR REGRESSOR PADA LSTM

	var(t-4)	var(t-3)	var(t-2)	var(t-1)	var(t)
4	0,96	0,84	1,00	0,96	0,97
5	0,84	1,00	0,96	0,97	0,98
6	1,00	0,96	0,97	0,98	0,89
7	0,96	0,97	0,98	0,89	0,79
8	0,97	0,98	0,89	0,79	0,88

3) *Pemangkasan Dataset*: tahap selanjutnya adalah mempertimbangkan pemangkasan dataset yang akan dimodelkan untuk meningkatkan kinerja model. Pemangkasan dilakukan karena setelah dilakukan analisis ulang untuk memahami tren yang terkandung dalam dataset, peneliti menemukan 2 pola tren yang berbeda. Dataset susu yang terdiri dari 300 hari hasil susu akan dipangkas menjadi 2 yaitu dataset 90 hari pertama dan dataset setelah 90 hari. Pertimbangan ini dilakukan karena pemerahan susu sapi maksimal dilakukan pada 90 hari pertama masa laktasi, setelah 90 hari biasanya peternak akan mengurangi jumlah pemerahan karena sapi akan memasuki masa kering untuk persiapan masa kawin dan bunting. Selanjutnya model prediksi yang dibuat adalah model prediksi untuk dataset setelah 90 hari. Plot dataset baru divisualisasikan dalam Gambar 9.

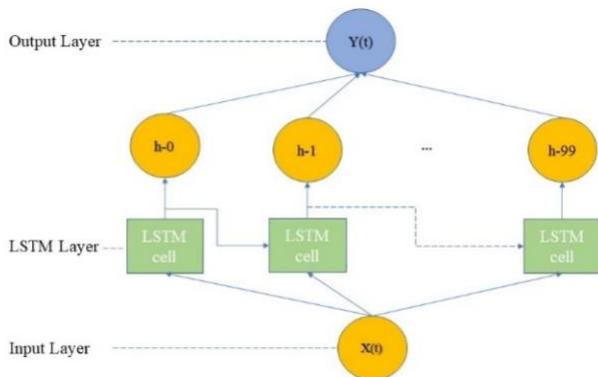


Gambar.9 Plot dataset deret waktu untuk tahapan *revise and enhance*

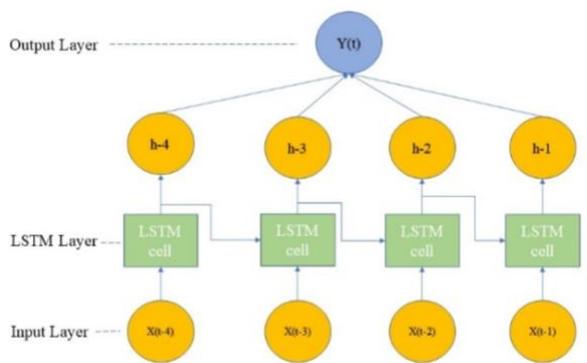
4) *Transformasi dan Normalisasi Data*: transformasi normalisasi data merupakan salah satu teknik pra-proses data. Transformasi data dilakukan dengan mengubah kolom tanggal menjadi index date time dengan nama kolom **ds**. Normalisasi data untuk mengubah data menjadi data yang memiliki skala yang sama. Normalisasi data dilakukan setelah melalui tahap EDA (*Exploratory Data Analysis*). Hasil EDA yaitu dataset tidak memiliki *missing value* ataupun outlier. Teknik normalisasi data digunakan adalah penskalaan data (*data scaling*) dengan *standar scaler* dan *min-max scaler*.

5) *Penyederhanaan Parameter LSTM*: sebelumnya arsitektur LSTM dibangun dengan parameter default, kemudian arsitektur model LSTM disederhanakan menjadi jumlah input= 4 dan jumlah *cell LSTM*=4 dan ukuran *batch size* disederhanakan menjadi 1. *Overfitting* pada LSTM

juga dapat dilakukan dengan menyederhanakan nilai parameter model karena dataset yang digunakan belum cukup besar untuk olah dengan *machine learning* seperti LSTM. Penyederhanaan yang diterapkan yaitu meminimalkan jumlah *cell* LSTM dan menambahkan jumlah *input layer* (lihat Gambar 10 dan 11).



Gambar.10 Arsitektul model LSTM awal



Gambar.11 Model LSTM yang direvisi

E. Hasil Evaluasi Model Prophet dan LSTM yang Direvisi

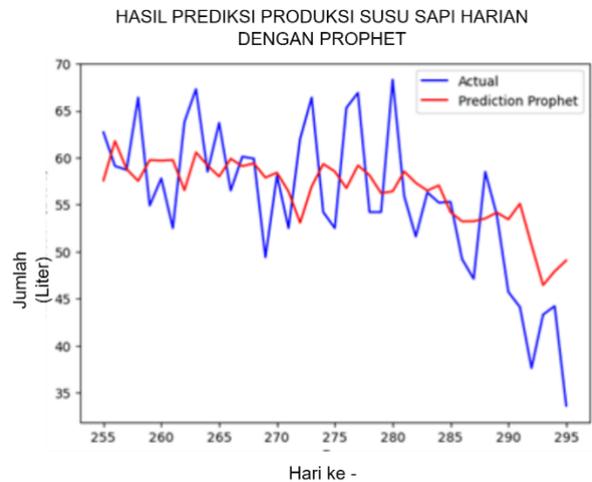
Evaluasi model dilakukan pada data yang melalui proses *data scalling* dan tidak melalui proses *data scalling*. Keduanya dilakukan untuk mencari peluang meningkatkan kinerja model. Komposisi data latih dan uji yang digunakan adalah 80:20. Hasil evaluasi kinerja model prophet ditunjukkan dalam Tabel VIII.

TABEL.VII
HASIL EVALUASI OPTIMASI MODEL PROPHET DAN LSTM

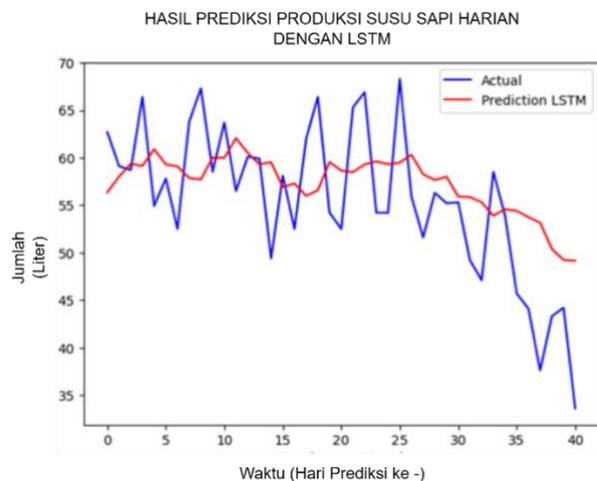
Evaluation Parameters	Prophet With Standar Scaler	Prophet With Min-Max Scaler	LSTM With Standar Scaler	LSTM With Min-Max Scaler
MSE	41.212	41.128	106.357	39.293
RMSE	6.419	6.413	10.312	6.268
MAPE	10.02%	9.99%	17.27%	9.98%
R ²	0.3483	0.3496	-0.681	0.3786

Hasil evaluasi revisi model LSTM kembali lebih unggul dibandingkan dengan model Prophet. Tabel IV juga menjelaskan bahwa model dengan penambahan fitur *regressor* dapat meningkatkan kinerja model dengan nilai R² lebih baik dari model sebelumnya. Kinerja model lebih baik saat dilakukan tranformasi data menggunakan teknik

penskalaan data yaitu *standar scaler* dan *min-max scaler*. R² terbesar diperoleh dari model LSTM dengan menggunakan *min-max scaler* yaitu sebesar 0.3786. Visualisasi plot prediksi deret waktu Prophet disajikan dalam Gambar 12, sedangkan visualisasi plot prediksi LSTM disajikan dalam Gambar 13.



Gambar.12 Visualisasi model prediksi deret waktu prophet yang direvisi



Gambar.13 Visualisasi model prediksi deret waktu LSTM yang direvisi

IV. KESIMPULAN

Berdasarkan hasil eksperimen beberapa skenario, LSTM memiliki kinerja model yang lebih baik dari Prophet untuk dataset deret waktu produksi susu selama laktasi 300 hari. Meskipun nilai R² LSTM dan Prophet tidak berbeda jauh secara signifikan, tetapi LSTM lebih unggul dari Prophet. Nilai uji koefisien determinasi (R²) dari model LSTM berhasil ditingkatkan menjadi 0.3786. Model LSTM dianggap cukup mampu menghasilkan nilai prediksi yang mengikuti nilai pengamatannya sebesar 37.86% dan persentase kesalahan LSTM juga dibawah 10% yaitu 9.98%. Dengan demikian LSTM dan Prophet dengan baik dapat memodelkan data yang tidak stasioner dan non-linear. Namun, LSTM dan Prophet lebih baik bekerja pada data univariate yang ditambahkan fitur regressor.

LSTM seharusnya mampu mengatasi data multi periode seperti dataset penelitian ini yang memiliki 2 periode tren yaitu tren menaik dan menurun. Namun, penyebab kecilnya nilai R^2 diduga karena pola tren data latih dan uji yang berbeda dan tidak berulang. Data historis yang dimodelkan belum cukup untuk merepresentasikan pola data untuk prediksi sehingga dilakukan tahapan *revise and enhance* untuk mengatasi permasalahan ini. Tahapan *revise and enhance* dilakukan dengan memangkas dataset setelah 90 hari (hanya periode ke 2 saat tren mulai menurun), penambahan regressor dan transformasi data menggunakan *mix-max scaler* berhasil meningkatkan kinerja model, meskipun tidak menaik secara signifikan. Peningkatan yang tidak besar juga diduga karena semakin terbatasnya jumlah setelah pemangkasan.

Penelitian masa depan dapat dilakukan dengan berfokus dalam mengatasi prediksi dengan dataset deret waktu yang terbatas dan multi periode. Kemunculan multi periode kemungkinan karena adanya periode tertentu dalam satu horizon waktu akibat intervensi dari *event* tertentu. Pengembangan kedepannya dimungkinkan dengan menambahkan metode *incremental* sebagai mekanisme khusus saat membangun model prediksi deret waktu dengan intervensi dari luar.

ACKNOWLEDGMENT

Dataset penelitian ini merupakan kontribusi hasil penelitian Righi et al [26] yang dipublikasikan sebagai dataset publik tahun 2020 dan dapat diunduh dari http://professor.unisinos.br/rafaelkunst/moocare_dataset.zip.

REFERENSI

- [1] H. Abbasimehr, M. Shabani, and M. Yousefi, "Computers & Industrial Engineering An optimized model using LSTM network for demand forecasting," *Comput Ind Eng*, vol. 143, no. March, p. 106435, 2020, doi: 10.1016/j.cie.2020.106435.
- [2] G. Cao, Y. Duan, and G. Li, "Linking Business Analytics to Decision Making Effectiveness: A Path Model Analysis," vol. 62, no. 3, pp. 384–395, 2015.
- [3] R. Sharda, D. Delen, and E. Turban, *Business Intelligence, Analytics, and Data Science_ A Managerial Perspective*, Fourth. Pearson, 2018.
- [4] S. Krishnamoorthi and S. K. Mathew, "Information & Management Business analytics and business value: A comparative case study," *Information & Management*, vol. 55, no. 5, pp. 643–666, 2018, doi: 10.1016/j.im.2018.01.005.
- [5] D. H. Fudholi, R. A. N. Nayoan, and S. Rani, "Stock Prediction Based on Twitter Sentiment Extraction Using," vol. 10, no. 1, pp. 187–198, 2022, doi: 10.52549/ijeei.10i1.3011.
- [6] A. R. S. Parmeza, V. M. A. Souza, and G. E. A. P. A. Batista, "Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model Antonio," *Inf Sci (N Y)*, 2019, doi: 10.1016/j.ins.2019.01.076.
- [7] C. Chandra and S. Budi, "Analisis Komparatif ARIMA dan Prophet dengan Studi Kasus Dataset Pendaftaran Mahasiswa Baru," vol. 6, pp. 278–287, 2020.
- [8] S. Jadon, J. Kanty, and A. Patnagar, *Challenges and Approaches to Time series forecasting: A Survey*. 2021.
- [9] A. Primawati, I. S. Sitanggang, and D. A. Astuti, "Business Intelligence and Analytics in Dairy Goat Livestock: Current and Future Challenge," no. 1, pp. 323–329, 2021.
- [10] R. da Rosa Righi, G. Goldschmidt, R. Kunst, C. Deon, and C. André da Costa, "Towards combining data prediction and internet of things to manage milk production on dairy cows," *Comput Electron Agric*, vol. 169, p. 105156, 2020, doi: <https://doi.org/10.1016/j.compag.2019.105156>.
- [11] M. D. Murphy, M. J. O. Mahony, L. Shalloo, P. French, and J. Upton, "Comparison of modelling techniques for milk-production forecasting Comparison of modeling techniques for milk-production forecasting," *J Dairy Sci*, vol. 97, no. 6, pp. 3352–3363, 2014, doi: 10.3168/jds.2013-7451.
- [12] S. S. Deshmukh and R. Paramasivam, "Forecasting of milk production in India with ARIMA and VAR time series models," *Asian Journal of Dairy & Food ...* researchgate.net, 2016.
- [13] R. Thakur and R. K. Gupta, "Prediction of milk yield using ARIMA," vol. 8, no. 3, pp. 1119–1123, 2020.
- [14] B. A. Taye, A. A. Alene, A. K. Nega, and B. G. Yirsaw, "Time series analysis of cow milk production at Andassa dairy farm, West Gojam Zone, Amhara Region, Ethiopia," *Model Earth Syst Environ*, 2020, doi: 10.1007/s40808-020-00946-z.
- [15] R. Chandra, S. Goyal, and R. Gupta, "Evaluation of Deep Learning Models for Multi-Step Ahead Time Series Prediction," *IEEE Access*, vol. 9, pp. 83105–83123, 2021, doi: 10.1109/ACCESS.2021.3085085.
- [16] S. J. Taylor and B. Letham, "Forecasting at Scale", doi: 10.7287/peerj.preprints.3190v2.
- [17] R. Sherman, *Business Intelligence Guidebook*. Elsevier Inc., 2015. doi: 10.1016/B978-0-12-411461-6.00015-0.
- [18] A. Perwitasari, Septiriana, and R. Tursina, "Data Preparation Structure untuk Pemodelan Prediktif Jumlah Peserta Ajar Mata Kuliah," *JEPIN*, vol. 9, no. 1, pp. 7–11, 2023.
- [19] H. Abbasimehr and R. Paki, "Improving time series forecasting using LSTM and attention models," *Journal of Ambient Intelligence and Humanized ...*, 2021.
- [20] S. Tepdang and R. Ponprasert, "Forecasting and Clustering of Cassava Price by Machine Learning (A study of Cassava prices in Thailand)," vol. 10, no. 4, pp. 825–836, 2022, doi: 10.52549/ijeei.v10i4.3946.
- [21] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," pp. 1–24, 2021, doi: 10.7717/peerj-cs.623.
- [22] N. M. M. C. Devi, I. P. A. Bayupati, and N. K. A. Wirdiani, "Prediksi Curah Hujan Dasarian dengan Metode Vanilla RNN dan LSTM untuk Menentukan Awal," *JEPIN*, vol. 8, no. 3, pp. 405–411, 2022.
- [23] K. Gajamannage and Y. Park, "Real-time forecasting of time series in financial markets using sequentially trained many-to-one LSTMs," *arXiv preprint arXiv:2205.04678*, 2022.
- [24] S. Arslan, "A hybrid forecasting model using LSTM and Prophet for energy consumption with decomposition of time series data," *PeerJ Comput Sci*, no. 2020, 2022, doi: 10.7717/peerj-cs.1001.
- [25] S. J. Taylor, M. Park, U. States, B. Letham, M. Park, and U. States, "Forecasting at Scale," pp. 1–25.
- [26] R. da Rosa Righi, G. Goldschmidt, R. Kunst, C. Deon, and C. André da Costa, "Towards combining data prediction and internet of things to manage milk production on dairy cows," *Comput Electron Agric*, vol. 169, p. 105156, 2020, doi: <https://doi.org/10.1016/j.compag.2019.105156>.