

Peramalan Produksi Pertanian Menggunakan Model Long Short-Term Memory

Mukhlis Mukhlis^{1,*}, Aziz Kustiyo², Aries Suharso²

¹ Ilmu Komputer; FMIPA; IPB/Universitas Bhayangkara Jaya; Jl. Dramaga Kabupaten Bogor; e-mail: clisfadhlmukhlis@apps.ipb.ac.id

² Ilmu Komputer; FMIPA; IPB; Jl. Dramaga Kabupaten Bogor; e-mail: kustiyoaziz@gmail.com, arso220377aries@apps.ipb.ac.id

* Korespondensi: e-mail: clisfadhlmukhlis@apps.ipb.ac.id

Diterima: 25 Mei 2021; Review: 02 Juni 2021; Disetujui: 08 Juni 2021

Cara sitasi: Muklis, Kustiyo A, Suharso A. 2021. Peramalan Produksi Pertanian Menggunakan Model Long Short-Term Memory. Bina Insani ICT Journal. Vol. 8 (1): 22-32.

Abstrak: Masalah yang timbul dalam peramalan hasil produksi pertanian antara lain adalah sulit untuk mendapatkan data yang lengkap dari variabel-variabel yang mempengaruhi hasil pertanian dalam jangka panjang. Kondisi ini akan semakin sulit ketika peramalan mencakup wilayah yang cukup luas. Akibatnya, variabel-variabel tersebut harus diinterpolasi sehingga akan menyebabkan bias terhadap hasil peramalan. (1) Mengetahui gambaran meta analisis penelitian peramalan produk pertanian menggunakan Long Short Term Memory (LSTM), (2) Mengetahui penelitian meta analisis cakupan wilayah, komoditi dan periode data terkait produk pertanian terutama gandum, kedelai jagung dan pisang, (3) Mengetahui praproses data antara lain menghilangkan data yang tidak sesuai, menangani data yang kosong, serta memilih variabel tertentu. Sebagai solusi dari masalah tersebut, peramalan hasil produksi pertanian dilakukan berdasarkan data historis hasil produksi pertanian. Salah model peramalan yang saat ini banyak dikembangkan adalah model jaringan syaraf LSTM yang merupakan pengembangan dari model jaringan syaraf recurrent (RNN). Tulisan ini merupakan hasil kajian literatur pengembangan model-model LSTM untuk peramalan hasil produksi pertanian meliputi gandum, kedelai, jagung dan pisang. Perbaikan kinerja model LSTM dilakukan mulai dari praproses, tuning hyperparameter, sampai dengan penggabungan dengan metode lain. Berdasarkan kajian tersebut, model-model LSTM memiliki kinerja yang lebih baik dibandingkan dengan model benchmark.

Kata kunci: jaringan syaraf, LSTM, peramalan, produksi pertanian, RNN.

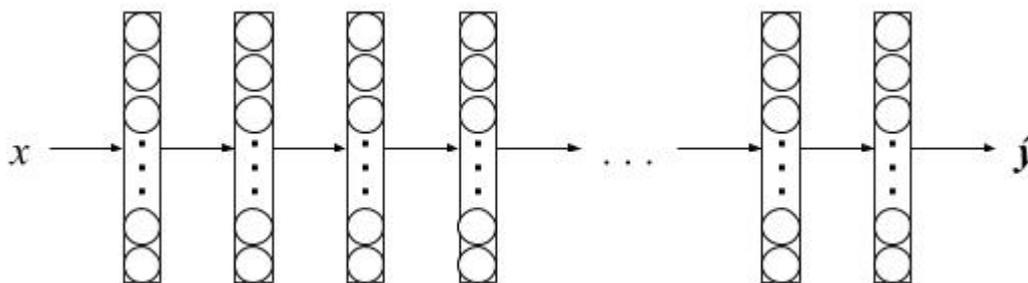
Abstract: Problems that arise in forecasting agricultural products include the difficulty of obtaining complete data on the variables that affect agricultural yields in the long term. This condition will be more difficult when the forecast covers a large area. As a result, these variables must be interpolated so that it will cause a bias towards the forecasting results. (1) Knowing the description of research maps for forecasting agricultural products using Long short term memory (LSTM), (2) Knowing Research Coverage areas, commodities, and data periods related to agricultural products, especially Wheat, Soybeans, corn, and bananas, (3) Knowing Preprocessing data between others remove inappropriate data, handle blank data, and select certain variables. This paper is the result of a literature review on the development of LSTM models for crop yields forecasting including wheat, soybeans, corn, and bananas. The Performance Improvements of the LSTM models were carried out by preprocessing data, hyperparameter tuning, and combining LSTM with other methods. Based on this study, LSTM models have better performance compared to the benchmark model.

Keywords: neural network, LSTM, forecasting, crop yield, RNN.

1. Pendahuluan

Peramalan hasil produksi pertanian sangat diperlukan bagi negara-negara yang sebagian besar penduduknya bergantung kepada hasil pertanian [16]. Prediksi hasil pertanian pada masa yang akan datang akan memberikan gambaran kecukupan hasil pertanian tersebut terhadap konsumsi dalam negeri. Dengan adanya gambaran tersebut maka pemerintah dapat melakukan langkah-langkah antisipasi untuk mengatasi masalah yang akan timbul. Masalah yang timbul dalam peramalan hasil produksi pertanian antara lain adalah sulit untuk mendapatkan data yang lengkap dari variabel-variabel yang mempengaruhi hasil pertanian dalam jangka panjang. Oleh karena itu dalam beberapa kasus, variabel-variabel tersebut harus diinterpolasi yang akan menyebabkan bias terhadap hasil peramalan. Di samping itu, dengan luasnya wilayah dalam satu negara akan menyebabkan keragaman yang sangat tinggi terhadap variabel-variabel tersebut. Sebagai akibatnya, peramalan hasil produksi pertanian berdasarkan variabel-variabel yang mempengaruhinya sangat sulit dilakukan.

Sebagai solusi dari masalah tersebut, peramalan hasil produksi pertanian dilakukan berdasarkan data historis hasil produksi pertanian. Salah model peramalan yang sedang banyak dikembangkan saat ini untuk kasus tersebut adalah model jaringan syaraf long sort-term memory (LSTM) yang merupakan pengembangan dari model jaringan syaraf recurrent (RNN) yang memiliki kinerja lebih baik dibandingkan rantai markov [6]. Model LSTM mampu mengingat data deret waktu masa lalu yang lebih panjang dibandingkan RNN [7] [8]. Tujuan dari penelitian ini adalah menelaah pengembangan beberapa model LSTM dalam peramalan hasil produksi pertanian pada berbagai komoditi tertama *gandum, kedelai jagung dan pisang*, untuk mengetahui meta analisis terkait cakupan wilayah, komoditi dan periode data. Penelaahan dilakukan mulai dari pengumpulan data sampai dengan pengukuran kinerja model.

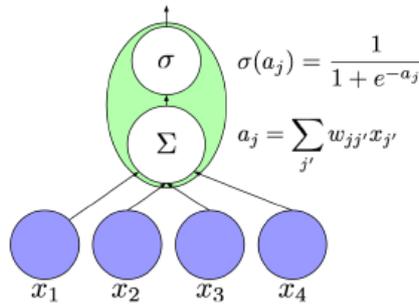


Sumber: Denny Britz/Wildlm (2018)

Gambar 1. Jaringan Syaraf Tiruan yang banyak lapis/berulang (RNN)

Jaringan saraf berulang atau *recurrent neural network* (RNN) adalah jenis arsitektur jaringan saraf tiruan yang pemrosesannya dipanggil berulang-ulang untuk memroses masukan yang biasanya adalah data sekuensial. RNN masuk dalam kategori *deep learning* karena data diproses melalui banyak lapis (*layer*). RNN telah mengalami kemajuan yang pesat dan telah merevolusi bidang-bidang seperti pemrosesan bahasa alami (NLP), pengenalan suara, sintesa musik, pemrosesan data finansial seri waktu, analisa deret DNA, analisa video, dan sebagainya. Permasalahan *Vanishing Gradient* dengan lapis yang banyak ini, sulit bagi simpul di lapis di bagian awal NN untuk “belajar” dari perubahan di \hat{y} , karena gradien yang timbul dari perubahan di \hat{y} tersebut semakin lama akan semakin mengecil dalam proses *back-propagation*, dan akhirnya menjadi kecil sekali mendekati nol atau dikatakan “menghilang” ketika sampai di lapis-lapis awal NN. Hal ini disebabkan oleh sifat perkalian antar bilangan pecahan. Bayangkan misalkan suatu bilangan pecahan $1/4$, dikalikan dengan bilangan pecahan lain misalnya $1/3$, maka dalam satu operasi ini nilainya sudah menjadi $1/12$. Dikalikan pecahan lain misalnya $1/5$, nilainya menjadi $1/60$, dst. Nilai ini akan mengecil secara eksponensial, dan dengan nilai pecahan yang kecil dan banyaknya operasi perkalian, nilainya akan mendekati nol.

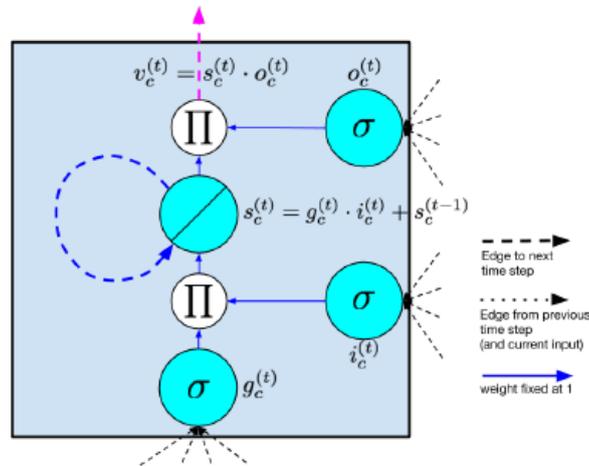
Demikian sebaliknya jika nilai gradiennya besar, maka akan menimbulkan masalah lain yaitu nilai gradien yang “meledak” (*exploding gradient*). Tapi masalah ini lebih mudah ditangani, misalnya dengan menerapkan suatu batas atas untuk nilai gradien. Jumlah lapis dalam RNN sendiri bisa panjang sekali, sepanjang jumlah deretan dalam masukan, dan ini menimbulkan permasalahan tersendiri karena RNN sering harus menyimpan ketergantungan dari input yang letaknya cukup berjauhan. Neuron buatan untuk menghitung fungsi nonlinier dari jumlah tertimbang masukannya dijelaskan pada gambar 2.



Sumber: Zachary C. Lipton (2015)

Gambar 2. Neuron buatan

Model LSTM terutama digunakan untuk mengatasi masalah gradien menghilang hal ini bagian perbaikan dari algoritma jaringan syaraf tiruan sebelumnya yang sebagian data belum terproses secara keseluruhan khususnya data pertanian informasinya sedikit tapi sangat penting untuk menghasilkan model, yang ditahun sebelumnya untuk diproses lagi. Model ini menyerupai jaringan neural berulang standar dengan lapisan tersembunyi, tetapi masing-masing biasa node pada gambar 2 di lapisan tersembunyi diganti dengan sel memori pada gambar 3 [4]. Setiap sel memori berisi node dengan tepi berulang yang terhubung sendiri dari bobot satu, memastikan bahwa gradien dapat melewati banyak langkah waktu tanpa menghilang atau meledak. Untuk membedakan referensi ke sel memori dan bukan node biasa, kami menggunakan subskrip c. Istilahnya adalah ingatan jangka pendek berasal dari intuisi data sebelumnya. Jaringan saraf berulang sederhana memiliki memori jangka panjang dalam bentuk beban. Bobot berubah perlahan selama pelatihan, menyandikan pengetahuan umum tentang data. Mereka juga memiliki ingatan jangka pendek dalam bentuk aktivasi fana, yang lolos dari setiap node ke node yang berurutan.



Sumber: Hochreiter et al. (1997)

Gambar 3. Satu sel memori LSTM

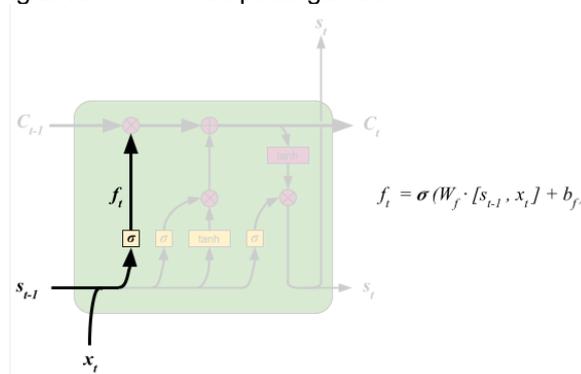
Keterangan Gambar:

Satu sel memori LSTM adalah, node yang terhubung sendiri adalah keadaan internal s. Garis diagonal menunjukkan linear, yaitu fungsi tautan identitas diterapkan. Biru putus-putus garis adalah tepi berulang, yang memiliki berat satuan tetap. Node menandai keluaran produk masukan mereka. Semua edge yang masuk dan dari node juga ada yang xed berat unit. seperti yang dikemukakan oleh [4].

Model LSTM memperkenalkan jenis penyimpanan perantara melalui sel memori. Sebuah sel memori adalah unit komposit, dibangun dari node yang lebih sederhana dalam pola konektivitas tertentu dengan dimasukkannya node perkalian baru, diwakili dalam diagram oleh surat. Semua elemen sel LSTM disebutkan dan dijelaskan di bawah ini. Perhatikan bahwa ketika kita menggunakan notasi vektor, kita mengacu pada nilai dari node di seluruh lapisan sel. Misalnya, s adalah vektor yang mengandung nilai dari s_c di setiap sel memori dalam sebuah

lapisan. Ketika subskrip c digunakan, itu adalah untuk mengindeks sel memori individu pada gambar 3.

Langkah pertama dalam LSTM adalah memutuskan informasi apa yang akan dibuang dari konteks C_{t-1} , dengan menggunakan gerbang sigmoid yang kita sebut “gerbang forgot” (forget gate, f_t). Gerbang ini membaca nilai s_{t-1} dan x_t , dan menghasilkan angka antara 0 dan 1 untuk setiap elemen dalam C_{t-1} . Nilai 1 artinya “benar-benar jaga elemen ini” sementara 0 artinya “benar-benar singkirkan elemen ini pada gambar 4.

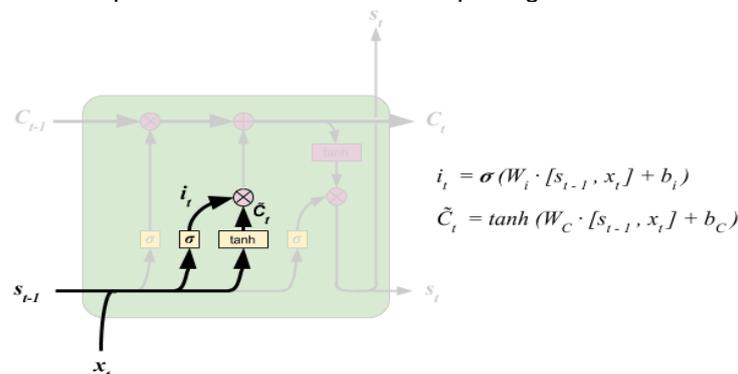


Sumber: Zachary C. Lipton (2015)

Gambar 4. Tahapan pertama LSTM

Catatan: notasi $[s_{t-1}, x_t]$ merupakan operasi konkatenasi; artinya kita menambahkan baris dari x_t dengan baris dari s_{t-1} . Suatu elemen dalam C_{t-1} mungkin menyimpan informasi gender subjek saat ini, sehingga kata ganti yang benar dapat digunakan. Ketika kita melihat subjek baru, maka elemen lama dalam C_{t-1} bisa kita forgotkan.

Langkah selanjutnya adalah memutuskan informasi baru apa yang akan kita gunakan di C_t . Proses ini memiliki dua bagian. Pertama, gerbang sigmoid yang disebut “gerbang masukan” (input gate, i_t) memutuskan nilai mana yang akan kita perbarui. Lalu sebuah lapis tanh menghasilkan kandidat vektor konteks baru, \tilde{C}_t , (dibaca: C tilde). Lalu kita gabungkan keduanya untuk membuat pembaruan ke konteks nanti pada gambar 5.

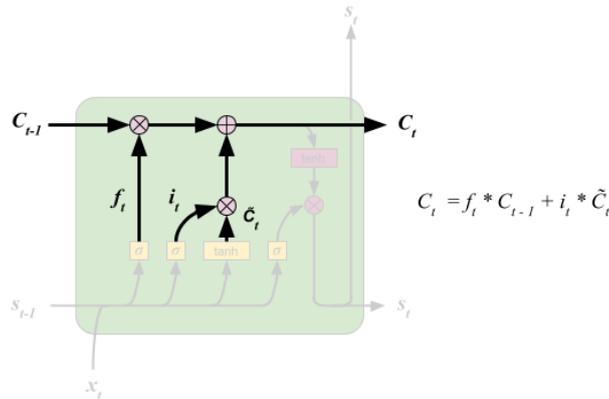


Sumber: Zachary C. Lipton (2015)

Gambar 5. Tahap kedua LSTM

Dalam contoh sebuah model peramalan produksi pertanian menambahkan jenis variabel subjek baru ke konteks, untuk menggantikan elemen yang lama yang kita forgotkan pada gambar 6.

Sekarang saatnya untuk memperbarui konteks lama C_{t-1} ke konteks baru C_t . Kita kalikan konteks lama dengan f_t untuk meforgtkan hal-hal yang kita putuskan untuk diforgotkan. Kita kalikan kandidat konteks baru kita \tilde{C}_t dengan i_t untuk memutuskan seberapa banyak kita akan menyertakan kandidat konteks baru. Lalu kita tambahkan keduanya.

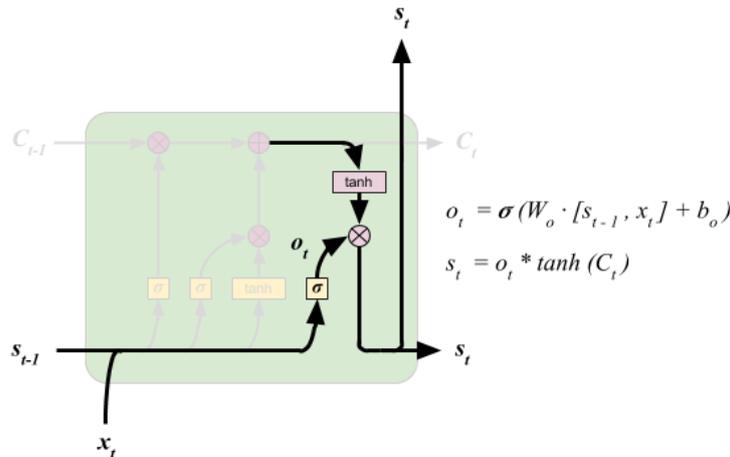


Sumber: Zachary C. Lipton (2015)

Gambar 6. Tahap ketiga LSTM

Dalam kasus model peramalan produksi pertanian adalah tempat kita benar-benar membuang informasi tentang gender subjek lama dan menambahkan informasi baru, seperti yang kita putuskan pada langkah sebelumnya.

Akhirnya, kita perlu memutuskan apa yang akan kita hasilkan. Output ini akan didasarkan pada nilai dalam konteks kita dan dilewatkan ke suatu filter. Pertama, kita jalankan gerbang sigmoid yang kita namakan gerbang output (*output gate, o_t*) untuk memutuskan bagian-bagian apa dari konteks yang akan kita hasilkan. Kemudian, kita lewatkan konteks melalui *tanh* untuk membuatnya menjadi antara -1 dan 1, dan kita kalikan dengan output gerbang sigmoid tadi sehingga kita hanya menghasilkan bagian yang kita putuskan pada gambar 7.



Sumber: Zachary C. Lipton (2015)

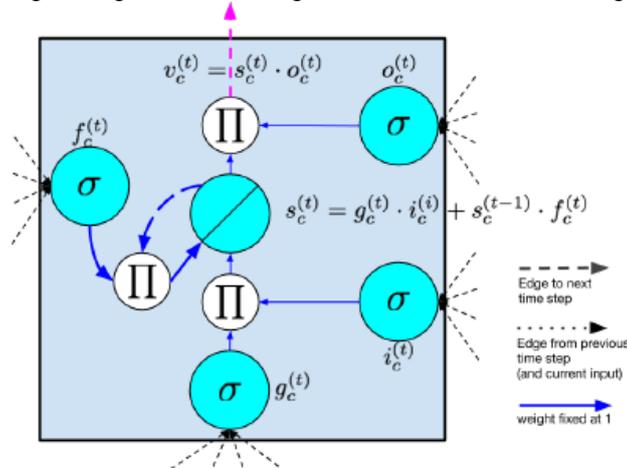
Gambar 7. Tahap akhir LSTM

2. Metode Penelitian

Ruang lingkup penelitian ini adalah penerapan LSTM pada peramalan hasil produksi pertanian meliputi gandum, kedelai, jagung dan pisang. Metode dalam penelitian ini adalah melakukan telaah pustaka terhadap beberapa penelitian yang telah dilakukan dalam mengembangkan model LSTM untuk peramalan hasil produksi pertanian. Telaah dilakukan terutama terhadap sumber data, pra-proses data, arsitektur dan parameter LSTM, proses pelatihan dan pengujian, kinerja serta kelebihan dan kekurangan model.

Di perhitungan manual dan di sebagian besar pekerjaan berikutnya, node input diberi label g. Kami mematuhi konvensi ini tetapi perhatikan bahwa hal itu mungkin membingungkan seperti halnya g tidak berdiri untuk gerbang. Di kertas asli, gerbang disebut y_{in} dan y_{out} tapi ini membingungkan karena y umumnya adalah singkatan dari output dalam pembelajaran mesin

literatur. Mencari pemahaman, kami memutuskan dengan konvensi ini dan menggunakan i, f, dan o untuk merujuk ke gerbang masukan, forgotkan dan keluaran masing-masing [20].



Sumber: Gers et al. (2000)

Gambar 8: Sel memori LSTM dengan forget gate seperti yang dijelaskan

Sejak LSTM pertama diperkenalkan, beberapa variasi telah diusulkan. Forget garbang, dijelaskan di atas, diusulkan pada tahun 2000 dan bukan merupakan bagian dari desain LSTM awal. Namun, mereka terbukti efektif dan standar di sebagian besar implementasi modern. Pada tahun yang sama, Gers dan Schmidhuber (2000) mengusulkan koneksi lubang intip yang lewat dari keadaan internal secara langsung ke gerbang input dan output dari node yang sama tanpa harus terlebih dahulu dimodulasi oleh gerbang keluaran. Mereka melaporkan bahwa koneksi ini meningkatkan kinerja pada tugas waktu di mana jaringan harus belajar mengukur dengan tepat interval antar acara. Intuisi koneksi lubang intip dapat ditangkap dengan contoh berikut. Pertimbangkan jaringan yang harus belajar menghitung objek dan memancarkan beberapa output yang diinginkan ketika objek n telah dilihat. Jaringan mungkin belajar untuk membiarkan sejumlah aktivasi tetap ke dalam keadaan internal setelah setiap objek terlihat. Aktivasi ini terjebak dalam status internal s_c by carousel error konstan, dan bertambah berulang setiap kali objek terlihat. Ketika objek ke-n terlihat, jaringan perlu mengetahuinya keluar konten dari keadaan internal sehingga dapat mempengaruhi keluaran. Untuk mencapai ini, gerbang keluaran o_c harus mengetahui isi dari keadaan internal s_c . Jadi s_c harus menjadi masukan ke o_c . Secara formal, perhitungan dalam model LSTM berjalan sesuai dengan perhitungan berikut, yang dilakukan pada setiap langkah waktu. Persamaan ini berikan algoritme lengkap untuk LSTM modern dengan forget gerbang:

$$g^{(t)} = \phi(W^{gx}x^{(t)} + W^{gh}h^{(t-1)} + b_g)$$

$$i^{(t)} = \sigma(W^{ix}x^{(t)} + W^{ih}h^{(t-1)} + b_i)$$

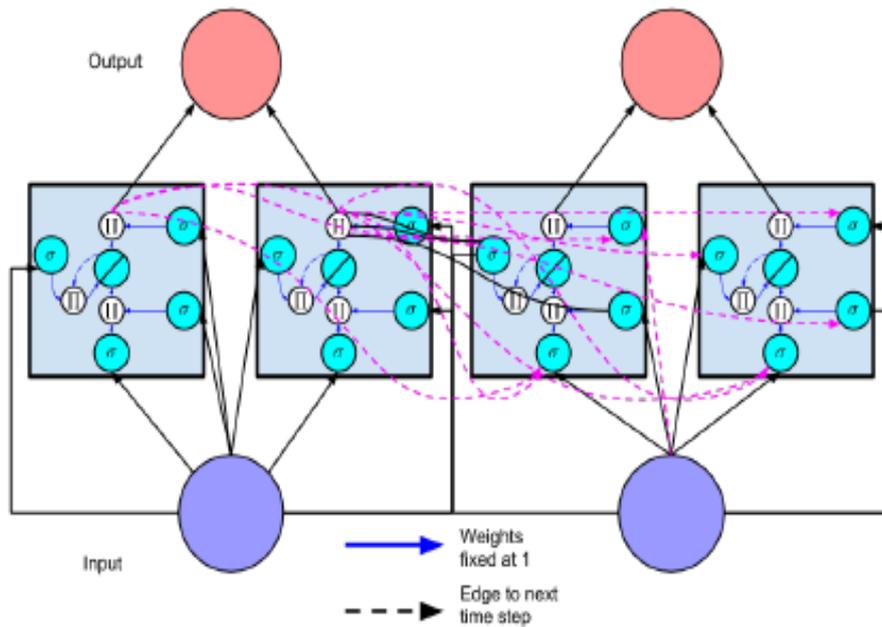
$$f^{(t)} = \sigma(W^{fx}x^{(t)} + W^{fh}h^{(t-1)} + b_f)$$

$$o^{(t)} = \sigma(W^{ox}x^{(t)} + W^{oh}h^{(t-1)} + b_o)$$

$$s^{(t)} = g^{(t)} \odot i^{(t)} + s^{(t-1)} \odot f^{(t)}$$

$$h^{(t)} = \phi(s^{(t)} \odot o^{(t)}).$$

Jaringan neural berulang dengan lapisan tersembunyi yang terdiri dari dua sel memori. Jaringan ditampilkan terbuka dalam dua langkah waktu [19].



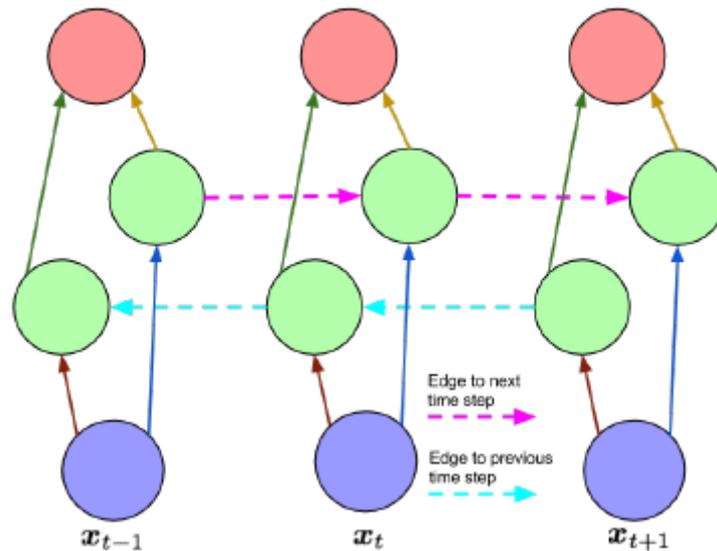
Sumber: Zachary C. Lipton (2015)

Gambar 9: Jaringan neural berulang

Nilai hidden layer dari LSTM pada waktu t adalah vektor $h^{(t)}$, sedangkan $h^{(t-1)}$ adalah nilai-nilai yang dikeluarkan oleh setiap sel memori di lapisan tersembunyi sebelumnya waktu. Perhatikan bahwa persamaan ini termasuk gerbang forgot, tetapi bukan lubang intip koneksi. Perhitungan untuk LSTM sederhana tanpa gerbang forgot adalah diperoleh dengan mengatur $f^{(t)} = 1$ untuk semua t . Kami menggunakan fungsi tanh untuk input node g mengikuti desain *state-of-the-art* [21]. Namun, dalam makalah LSTM asli, fungsi aktivasi untuk g adalah sigmoid σ .

Secara intuitif, dalam hal operan ke depan, LSTM dapat mempelajari kapan harus melepaskan aktivasi ke keadaan internal. Selama gerbang masukan bernilai nol, tidak ada aktivasi yang bisa masuk. Demikian pula, gerbang keluaran belajar kapan harus membiarkannya. Saat kedua gerbang ditutup, aktivasi terjebak dalam memori sel, tidak tumbuh atau menyusut, atau memengaruhi keluaran pada perantara langkah waktu. Dalam hal operan mundur, carousel error konstan memungkinkan gradien untuk disebarkan kembali ke banyak langkah waktu, baik meledak maupun menghilang. Dalam pengertian ini, gerbang belajar kapan harus membiarkan kesalahan masuk, dan kapan untuk membiarkannya keluar. Dalam praktiknya, LSTM telah menunjukkan kemampuan unggul untuk mempelajari jarak jauh dependensi dibandingkan dengan RNN sederhana. Akibatnya, mayoritas Makalah aplikasi mutakhir yang tercakup dalam ulasan ini menggunakan model LSTM.

Salah satu titik kebingungan yang sering terjadi adalah cara di mana banyak memori sel digunakan bersama untuk membentuk lapisan tersembunyi dari jaringan saraf yang berfungsi. Untuk mengurangi kebingungan ini, kami menggambarkan pada Gambar 10 sebuah jaringan sederhana dengan dua sel memori, analog dengan Gambar 9. Keluaran dari setiap sel memori ows pada langkah waktu berikutnya ke node input dan semua gerbang dari setiap sel memori. Hal ini umum untuk memasukkan beberapa lapisan sel memori [20]. Biasanya, dalam arsitektur ini setiap lapisan menerima masukan dari lapisan di bawahnya di langkah waktu yang sama dan dari lapisan yang sama di langkah waktu sebelumnya. Jaringan saraf berulang dua arah seperti yang dijelaskan oleh Schuster dan Paliwal (1997), terbentang dalam waktu dijelaskan pada gambar 10.



Sumber: Zachary C. Lipton (2015)

Gambar 10. Jaringan saraf berulang dua arah

3. Hasil dan Pembahasan

Penelitian yang ditelaah adalah penelitian yang dilakukan oleh Haider et al. (2019), Terliksiz dan Altylar (2019), Kaneko et al. (2019), Chao et al. (2018), Rebortera dan Fajardo (2019). Pada Tabel 1 disajikan satuan wilayah dan komoditi yang diteliti. Dari tabel tersebut dapat dilihat bahwa wilayah yang diteliti beragam mulai dari level Kabupaten (*county*), negara bagian, sampai wilayah di beberapa negara. Dari sisi komoditi, terdapat empat komoditi yaitu gandum, kedelai, jagung dan pisang dengan tingkat kepentingan yang berbeda-beda. Gandum merupakan tanaman pangan terpenting kedua di Pakistan [2] sedangkan jagung merupakan tanaman pangan utama di beberapa negara Afrika [5]. Di Filipina, pisang merupakan salah satu tanaman andalan untuk keperluan ekspor [10]. Dari tabel 1 juga dapat dilihat perbedaan periode datanya mulai dari 13 tahun sampai dengan 16 tahun.

Tabel 1 Cakupan wilayah, komoditi dan periode data yang ditelaah

Wilayah	Komoditi	Periode data	Pustaka
Pakistan	Gandum	1902 – 2018	Haider et al. (2019)
Lauderdale Country, Alabama	Kedelai	2003 – 2016	Terliksiz dan Altylar (2019)
Negara-negara di Afrika	Jagung (Afrika)	2003 – 2016	Kaneko et al. (2019)
99 Country di Iowa	Jagung (USA)	1980 – 2016	Chao et al. (2018)
Davao del Norte, Filipina	Pisang	2014 – 2018	Rebortera dan Fajardo (2019)

Sumber: Hasil Penelitian (2021)

Sumber data yang digunakan berasal dari berbagai instansi dan terdapat berbagai variabel pendukung. Sumber data penelitian disajikan pada tabel 2. Di samping data historis hasil produksi pertanian, beberapa penelitian menambahkan data pendukung seperti data satelit terkait suhu permukaan tanah dan tipe tutupan lahan, serta data cuaca dan data terkait kondisi tanah. Data tersebut ditambahkan dalam upaya untuk meningkatkan kinerja model peramalan yang dibangun.

Tabel 2. Sumber data penelitian

Komoditi	Sumber data	Jenis data	Pustaka
Gandum	Biro Statistik Federal Pakistan, Survei Ekonomi Pakistan 2017	Data historis produksi gandum Pakistan 1902 – 2018	Haider et al. (2019)
Kedelai	1. USDA NASS Quick Stat Lite tool 2. Satelit MODIS	1. Data produksi kedelai tahun 2003-2016 2. Citra satelit MODIS untuk suhu permukaan tanah dan tipe penutupan lahan	Terliksiz dan Altylar (2019)
Jagung (Afrika)	1. USAID (FEWS NET) 2. Satelit MODIS	1. Data produksi jagung tahun 2003-2016 2. 7 band data reflektansi permukaan dan 2 band untuk suhu permukaan tanah siang dan malam	Kaneko et al. (2019)

Komoditi	Sumber data	Jenis data	Pustaka
Jagung (USA)	1. National Agricultural Statistics Service (NASS) Quick Stats 2. Weather Underground 3. NOAA 4. Basis data SSURGO	1. Data produksi jagung 2. Data cuaca per jam 3. Data kelembaban tanah 4. Data kualitas tanah	Chao et al. (2018)
Pisang	Koperasi ARB Dapco, Davao del Norte, Filipina.	Hasil panen pisang tahun 2014 – 2018	Rebortera dan Fajardo (2019)

Sumber: Hasil Penelitian (2021)

Sebelum data mentah dapat digunakan untuk mengembangkan model peramalan, perlu dilakukan praproses agar data tersebut sesuai dengan kebutuhan dari model yang dibangun. Praproses data antara lain menghilangkan data yang tidak sesuai, menangani data yang kosong, serta memilih variabel tertentu. Di samping itu, dilakukan juga pemulusan data (*data smoothing*) dengan berbagai metode. Praproses data disajikan pada tabel 3.

Tabel 3. Praproses data

Komoditi	Praproses data	Pustaka
Gandum	Smoothing dengan Robust-LOWESS	Haider et al. (2019)
Kedelai	1. Memotong ukuran citra satelit sesuai batas wilayah 2. Memformat data reflektan menjadi 14x10x24 3. Atau 14 (tahun) x10 (spektral) x24 (3x8 pengukuran)	Terliksiz dan Altylar (2019)
Jagung (Afrika)	1. Menghapus titik data dengan entri yang hilang, 2. Menghapus titik data yang di luar dua standar deviasi, 3. Menghapus titik data dengan area yang ditanami kurang dari 5000 hektar	Kaneko et al. (2019)
Jagung (USA)	1. Menggunakan input data harian, bukan per jam 2. Memilih variabel dari 28 variabel menggunakan metode MRMR dan pakar diperoleh 10 variabel	Chao et al. (2018)
Pisang	1. Mengubah data deret waktu menjadi stasioner. 2. Lag differencing digunakan untuk menghilangkan tren peningkatan data. 3. Penskalaan data untuk memenuhi fungsi aktivasi tangen hiperbolik	Rebortera dan Fajardo (2019)

Sumber: Hasil Penelitian (2021)

Selanjutnya akan dibahas arsitektur dan parameter model LSTM yang dibangun. Perbaikan kinerja model LSTM dilakukan dengan melakukan praproses data seperti pada bagian sebelumnya atau melalui perbaikan pada model LSTM atau pun menggabungkan model LSTM dengan model yang lain. Selain itu, pemilihan parameter yang tepat juga akan meningkatkan kinerja model LSTM. Arsitektur dan parameter model LSTM yang dibangun disajikan pada tabel 4.

Tabel 4 Arsitektur dan parameter untuk model LSTM terbaik

Model LSTM	Hidden unit	Drop rate	Learning rate	Solver	Metode lain
Gandum	200	20	0.0032	SGDM	-
Kedelai	ta	No drop out	0.0010	Adam optimizer	CNN
Jagung (Afrika)	ta	Using drop out	Xavier initialitation	Adam optimizer	1. Gaussian Process 2. Random hyperparameter search
Jagung (USA)	2	Random	Random	ta	-
Pisang	2	No drop out	Ta	Adam optimizer	-

Sumber: Hasil Penelitian (2021)

Penentuan jumlah hidden unit untuk model LSTM terbaik secara umum dilakukan secara *trial-and-error* kemudian dipilih jumlah yang menghasilkan model dengan kinerja terbaik. Demikian juga dengan parameter *drop-out*, secara umum dicobakan untuk kedua kondisi yaitu dengan dan tanpa *drop-out*. Apabila menggunakan *drop-out* maka dicobakan beberapa nilai *drop-out* untuk mencari kinerja terbaik. Pada penentuan nilai learning rate, secara umum inialisasi ditentukan secara acak namun ada yang menggunakan metode lain yaitu inialisasi *xavier*. Pada pemilihan *solver*, hampir semuanya menggunakan *Adam optimizer*, di samping

SGDM (stochastic gradient descent with momentum). Adapun metode lain yang digabungkan dengan LSTM antara lain adalah CNN dan *Gaussian process* [18].

Tabel 5. Kinerja model LSTM terbaik

Model LSTM	RMSE	RMSE benchmark	R-value
Gandum	792.00	1065.00 ^c	0.81
Kedelai	0.81	5.32 ^d	Ta
Jagung (Afrika)	Ta	Ta	0.56 ^a
Jagung (USA)	13.82 ^b	14.55 ^b	Ta
Pisang	34.81	35.03	Ta

Sumber: Hasil Penelitian (2021)

Keterangan:

- a : R-value rata-rata untuk 6 negara
- b : data yang tersedia berupa MSE
- c : model ARIMA (1,2,2)
- d : 3D CNN
- ta : tidak tersedia data

Proses pelatihan dan pengujian model diawali dengan membagi data menjadi data pelatihan dan pengujian. Dari model-model LSTM yang dikembangkan tersebut terdapat variasi dalam membagi data pelatihan dan pengujian. Porsi untuk data pelatihan berkisar dari 80% sampai dengan kurang lebih 90%. Data pengujian biasanya dipilih beberapa data terakhir. Sebagian penelitian menambahkan data validasi sebesar 10%.

Kinerja model-model LSTM yang dibangun disajikan pada Tabel 5. Untuk model LSTM gandum dan LSTM kedelai dapat dibandingkan dengan RMSE benchmark karena data tersedia, sedangkan model untuk jagung (USA) dan pisang hanya dibandingkan dengan RMSE yang diperoleh dari penelitian yang dilakukan karena tidak tersedia data yang lain. Dari Tabel 5 dapat dilihat bahwa model LSTM yang dibangun memiliki nilai RMSE yang jauh lebih kecil dibandingkan nilai benchmark. Dengan demikian, model LSTM yang dibangun menghasilkan kinerja yang lebih baik dibandingkan dengan benchmark maupun model lain dari penelitian yang dilakukan. Dengan data yang ada pada Tabel 5, belum dapat dibandingkan kinerja model LSTM pada data komoditi yang berbeda.

4. Kesimpulan

Peramalan hasil produksi pertanian merupakan hal yang sangat penting karena berkaitan dengan perencanaan kebijakan pemerintah pada masa yang akan datang. Pengembangan model-model LSTM untuk peramalan hasil produksi pertanian telah menunjukkan bahwa model LSTM merupakan model yang cukup menjanjikan untuk bidang ini. Kemampuan LSTM untuk mengingat pola data masa lalu yang cukup panjang tercermin dari nilai RMSE yang lebih kecil dari benchmark maupun dari model-model lain yang dibangun pada saat yang bersamaan. Penggabungan model LSTM dengan metode lain seperti CNN dan *Gaussian process* telah menunjukkan peningkatan kinerja model LSTM. peningkatan kerja tersebut merupakan akibat dari masuknya data pendukung ke dalam model LSTM. Peningkatan kinerja LSTM juga dapat diupayakan dengan melakukan praproses data seperti smoothing data. Penelitian lebih lanjut pengembangan model-model LSTM untuk peramalan hasil produksi pertanian masih sangat diperlukan baik dari sisi peningkatan kinerja LSTM maupun dari sisi komoditas pertanian.

Referensi

- [1] Chao ZJ, Liu C, Hendricks NP, Ganapathysubramanian B, Hayes DJ, Sarkar S. *Predicting County Level Corn Yields Using Deep Long Short-Term Memory Models*. 2018.
- [2] Haider SA, Naqvi SR, Akram T, Umar GA, Shahzad A, Sial MR, Khaliq S, Kamran M. *LSTM Neural Network Based Forecasting Model for Wheat Production in Pakistan*. *Agronomy* vol. 9, no.72, 2019, doi:10.3390/agronomy9020072.
- [3] Hava T. Siegelmann and Eduardo D. Sontag. Turing computability with neural nets. *Applied Mathematics Letters*, vol. 4, no.6, pp. 77-80, 1991.
- [4] Hochreiter dan Schmidhuber. Long Short-Term Memory, *Neural Computation* vol. 9, no.8, pp.1735-1780, 1997.

- [5] Kaneko A, Kennedy T, Mei L, Sintek C, Burke M, Ermon S, Lobell D. Deep Learning For Crop Yield Prediction in Africa. 2019, https://aiforsocialgood.github.io/icml2019/accepted/track1/pdfs/20_aishg_icml2019.pdf.
- [6] Krakovna, V., & Doshi-Velez, F. Increasing the interpretability of recurrent neural networks using hidden markov models. 2016, arXiv preprint arXiv:1606.05320.
- [7] Lipton Z.C., Berkowitz J, Elkan C. A Critical Review of Recurrent Neural Networks for Sequential Learning, 2015, <https://arXiv:1506.00019>.
- [8] Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzell, R. Learning to diagnose with LSTM recurrent neural networks, 2015, arXiv preprint arXiv:1511.03677.
- [9] Mike Schuster and Kuldeep K. Paliwal. Bidirectional recurrent neural networks. Signal Processing, IEEE Transactions on, vol.45, no. 11, pp. 2673-2681,1997.
- [10] Rebortera MA, Fajardo AC. An Enhanced Deep Learning Approach in Forecasting Banana Harvest Yields. IJACSA vol.10, no.9, 2019.
- [11] Richard Socher, Christopher D. Manning, and Andrew Y Ng. Learning continuous phrase representations and syntactic parsing with recursive neural networks. In Proceedings of the NIPS-2010 Deep Learning and Unsupervised Feature Learning Workshop, pp. 1-9, 2010.
- [12] Richard Socher, Cui C. Lin, Chris Manning, and Andrew Y. Ng. Parsing natural scenes and natural language with recursive neural networks. In Proceedings of the 28th international conference on machine learning (ICML-11), pp. 129-136, 2011.
- [13] Richard Socher, Eric H. Huang, Je_rey Pennington, Christopher D. Manning, and Andrew Y. Ng. Dynamic pooling and unfolding recursive autoencoders for paraphrase detection. In Advances in Neural Information Processing Systems, pp. 801-809, 2011.
- [14] Richard Socher, Je_rey Pennington, Eric H. Huang, Andrew Y. Ng, and Christopher D. Manning. Semi-supervised recursive autoencoders for predicting sentiment distributions. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 151-161. Association for Computational Linguistics, 2011.
- [15] Terlikisiz, A. S., & Altıylar, D. T. Use Of Deep Neural Networks For Crop Yield Prediction: A Case Study Of Soybean Yield in Lauderdale County, Alabama, USA. 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), 2019, doi:10.1109/agro-geoinformatics.2019.8820257
- [16] Wang, A. X., Tran, C., Desai, N., Lobell, D., & Ermon, S. Deep transfer learning for crop yield prediction with remote sensing data. In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, p. 50, 2018.
- [17] Yoram Singer and John C. Duchi. Efficient learning using forward-backward splitting. In Advances in Neural Information Processing Systems, pp. 495-503, 2009.
- [18] You, J., Li, X., Low, M., Lobell, D., & Ermon, S. Deep gaussian process for crop yield prediction based on remote sensing data. In Thirty-First AAAI Conference on Artificial Intelligence, 2017.
- [19] Zachary C. Lipton, John Berkowitz, Charles Elkan 2015. A Critical Review of Recurrent Neural Networks for Sequence Learning, 2015, arXiv:1506.00019v4.
- [20] Sutskever I., Vinyals O & Le Q. V. Sequence to Sequence Learning with Neural Networks. Advances in Neural Information Processing Systems pp.1-9, 2014, <https://arxiv.org/pdf/1409.3215.pdf>.
- [21] Zaremba W. & Sutskever I. Learning to Execute. International Conference on Learning Representations, pp.1-25, 2015, arXiv:1410.4615v3.
- [22] Gers F. & Schmidhuber. Recurrent nets that time and count. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Volume: 3, 2000, doi:10.1109/IJCNN.2000.861302.