



MODEL PREDIKSI PM2.5 DARI KEBAKARAN HUTAN DAN LAHAN DI WILAYAH SUMATERA MENGGUNAKAN ALGORITMA *DEEP LEARNING*

NAUFAL IBNU SALAM



PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Model Prediksi PM2.5 dari Kebakaran Hutan dan Lahan di Wilayah Sumatera menggunakan Algoritma *Deep Learning*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, April 2024

Naufal Ibnu Salam
G6501201024

RINGKASAN

NAUFAL IBNU SALAM. Model Prediksi PM_{2.5} dari Kebakaran Hutan dan Lahan di Wilayah Sumatera menggunakan Algoritma *Deep Learning*. Dibimbing oleh IMAS SUKAESIH SITANGGANG dan MUSHTHOFA.

Kebakaran hutan dan lahan (karhutla) merupakan salah satu bencana penyebab pencemaran udara di Indonesia. Data per 12 Juli 2022, Kementerian Lingkungan Hidup dan Kehutanan (KLHK) mencatat sebanyak 13.854 kasus kebakaran aktif selama tahun 2021. Dampak langsung yang disebabkan oleh karhutla antara lain: menyebabkan penyakit infeksi saluran, mengganggu aktivitas manusia, terganggunya moda transportasi, timbulnya persoalan internasional asap dari kebakaran hutan tersebut seperti kerugian materiil dan imateriil. Asap karhutla mengandung polutan berbahaya khususnya partikel PM_{2.5} yang bersifat karsinogenik dan berperan dalam Infeksi Saluran Pernafasan Akut (ISPA). Pemerintah Indonesia telah menerbitkan aturan batasan PM_{2.5} dalam Peraturan Pemerintah RI No. 41 Tahun 1999 tentang Pengendalian Pencemaran Udara, dengan baku mutu udara ambien nasional untuk PM_{2.5} sebesar 65 μm^3 dalam rata-rata 24 jam.

Salah satu upaya untuk meminimalisir meningkatnya jumlah penderita penyakit ISPA adalah dengan mendeteksi partikel PM_{2.5} yang dihasilkan dari asap karhutla. Penelitian ini bertujuan untuk membangun model prediksi partikel PM_{2.5} menggunakan algoritma *machine learning* yaitu *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU). Data yang digunakan adalah dataset titik panas yang diperoleh dari sensor *Visible Infrared Imaging Radiometer Suite* (VIIRS) milik *National Aeronautics and Space Administration* (NASA) yang diunduh pada website *Fire Information for Resource Management System* (FIRMS) NASA dan data PM_{2.5} yang diunduh dari website OpenAQ. Nilai *missing value* yang terdapat pada data diisi menggunakan metode *Interpolasi Linear* (IL), *Last Observation Carried Forward* (LOCF), *Next Observation Carried Backward* (NOCB), *Interpolasi Nearest* (IN), *Exponential Moving Average* (EMA)(*Span*=5), *Exponential Moving Average* (EMA)(*Half-life*=5). Penelitian ini berhasil melakukan praproses data dan pembentukan model prediksi PM_{2.5} dari karhutla di wilayah Pulau Sumatra. Hasil uji keseluruhan dari model ini memuaskan. Nilai R^2 yang diperoleh menggunakan algoritma LSTM dan GRU mendapatkan nilai 0,920 hingga 0,997. Performa model dari setiap skenario sangat dipengaruhi oleh penanganan *missing value* pada dataset. Algoritma LSTM menghasilkan model prediksi yang lebih baik dibandingkan algoritma GRU ketika diterapkan pada dataset yang lebih besar dan kompleks. Model prediksi PM_{2.5} sangat penting untuk mendeteksi secepat mungkin partikel PM_{2.5} untuk peringatan dini polusi udara dan perlindungan kesehatan masyarakat.

Kata kunci: *Gated Recurrent Unit*, Karhutla, *Long Short-Term Memory*, PM 2.5, Prediksi



SUMMARY

NAUFAL IBNU SALAM. Prediction Model of PM2.5 from Forest and Land Fires in the Sumatra Region using Deep Learning Algorithm. Supervised by IMAS SUKAESIH SITANGGANG and MUSHTHOFA.

Forest and land fires were one of the disasters causing air pollution in Indonesia. As of July 12, 2022, the Ministry of Environment and Forestry (MoEF) had recorded a total of 13,854 active fire cases during the year 2021. Direct impacts caused by karhutla included causing respiratory infections, disrupting human activities, disturbing transportation modes, and the emergence of international issues due to the smoke from forest fires, leading to material and immaterial losses. Karhutla smoke contained hazardous pollutants, especially PM2.5 particles, which were carcinogenic and played a role in Acute Respiratory Infection (ARI). The Indonesian government had issued regulations limiting PM2.5 in Government Regulation No. 41 of 1999 concerning Air Pollution Control, with the national ambient air quality standard for PM2.5 set at 65 $\mu\text{m}/\text{m}^3$ in a 24-hour average.

One effort to minimize the increasing number of respiratory disease patients was by detecting PM2.5 particles produced from karhutla smoke. The study aimed to build a predictive model for PM2.5 particles using machine learning algorithms, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). The data used included hotspot datasets obtained from the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor owned by the National Aeronautics and Space Administration (NASA), downloaded from the NASA Fire Information for Resource Management System (FIRMS) website, and PM2.5 data downloaded from the OpenAQ website. Missing values in the data were filled using Linear Interpolation (IL), Last Observation Carried Forward (LOCF), Next Observation Carried Backward (NOCB), Nearest Interpolation (IN), Exponential Moving Average (EMA)(Span=5), and Exponential Moving Average (EMA)(Half-life=5). The study successfully conducted data preprocessing and the formation of a PM2.5 prediction model from forest fires in the Sumatra Island region. The overall test results of this model were satisfactory. The obtained R^2 values range of this model LSTM and GRU from 0.920 to 0.997. The model's performance in each scenario was highly influenced by the handling of missing values in the dataset. The LSTM algorithm model outperformed the GRU algorithm model when applied to larger and more complex datasets. The PM2.5 prediction model was crucial for detecting PM2.5 particles as quickly as possible for early air pollution warnings and public health protection.

Keywords: Forest and Land Fires, Gated Recurrent Unit, Long Short-Term Memory, PM2.5, Prediction



MODEL PREDIKSI PM2.5 DARI KEBAKARAN HUTAN DAN LAHAN DI WILAYAH SUMATERA MENGGUNAKAN ALGORITMA *DEEP LEARNING*

NAUFAL IBNU SALAM

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Ilmu Komputer

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2024**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



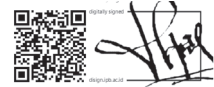
Judul Tesis : Model Prediksi PM2.5 dari Kebakaran Hutan dan Lahan di Wilayah Sumatera menggunakan Algoritma *Deep Learning*
Nama : Naufal Ibnu Salam
NIM : G6501201024

Disetujui oleh

Pembimbing 1:
Prof. Dr. Imas Sukaesih Sitanggang, S.Si M.Kom



Pembimbing 2:
Dr. Mushthofa, S.Kom, M.Sc



Diketahui oleh

Ketua Program Studi:
Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom
NIP 19750130 199802 2 001



Dekan Fakultas Matematika Dan Ilmu Pengetahuan Alam:
Dr. Berry Juliandi, S.Si., M.Si.
NIP 19780723 200701 1 001



Tanggal Ujian:
29 Februari 2024

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2022 sampai bulan Juni 2023 ini ialah meminimalisir dampak dari kebakaran hutan dan lahan di wilayah Sumatera, dengan judul “Model Prediksi PM2.5 dari Kebakaran Hutan dan Lahan di Wilayah Sumatera menggunakan Algoritma *Deep Learning*”.

Terima kasih penulis ucapkan kepada orang tua, Bapak Moh. Sodikin, dan Ibu almarhumah Mahmudah Astiani yang telah merawat dan membimbing saya sejak kecil. Ungkapan terima kasih juga disampaikan Ibu Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom dan Bapak Dr. Mushtofa, S.Kom, M.Sc yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada, Bapak Dr. Aziz Kustiyo S.Si., M.Kom atas saran dan masukkannya selaku dosen penguji sidang akhir tesis penulis. Ungkapan terima kasih juga disampaikan kepada Fiatuszahro, keluarga, serta teman seperjuangan kuliah yang telah memberikan dukungan, doa, dan kasih sayangnya.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, April 2024

Naufal Ibnu Salam

DAFTAR ISI

DAFTAR TABEL	viii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	viii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
II TINJAUAN PUSTAKA	4
2.1 Kebakaran Hutan dan Lahan	4
2.2 <i>Long Short-Term Memory</i>	4
2.3 <i>Gated Recurrent Unit</i>	7
III METODE	10
3.1 Data Penelitian	10
3.2 Tahapan Penelitian	11
3.3 Lingkungan Pengembangan	16
IV HASIL DAN PEMBAHASAN	17
4.1 Seleksi PM2.5 Berdasarkan Data Titik Panas	17
4.2 Praproses Data	18
4.3 Pembagian Data	19
4.4 Pembuatan Model	20
4.5 Evaluasi Model	23
V SIMPULAN DAN SARAN	27
5.1 Simpulan	27
5.2 Saran	27
DAFTAR PUSTAKA	28
LAMPIRAN	32
RIWAYAT HIDUP	40



DAFTAR TABEL

1	Luas karhutla dalam lima tahun terakhir (Ha) (SiPongi KLHK 2023)	1
2	Kasus karhutla di Indonesia (BPS 2022)	1
3	Skenario percobaan	12
4	Data titik panas di Pulau Sumatera	17
5	Jumlah data PM2.5 di Pulau Sumatera	18
6	Jumlah data agregasi PM2.5 di Pulau Sumatera	18
7	Jumlah data latih dan data uji PM2.5	20
8	Skenario model LSTM dan GRU	20

DAFTAR GAMBAR

1	Luas kebakaran hutan di Indonesia (Data Indonesia 2022)	4
2	Contoh arsitektur LSTM (Olah 2015)	5
3	Contoh arsitektur GRU (Chung <i>et al.</i> 2014)	8
4	Tahapan penelitian	11
5	Persebaran data titik panas dan data PM2.5 di Pulau Sumatera	17
6	Visualisasi imputasi PM2.5 dengan berbagai metode	19
7	Plot perbandingan nilai aktual (oranye) dan prediksi (hijau) PM2.5 model LSTM dengan berbagai metode	22
8	Plot perbandingan nilai aktual (oranye) dan prediksi (hijau) PM2.5 model GRU dengan berbagai metode.	23
9	Hasil evaluasi data uji model prediksi PM2.5 semua <i>confidence</i>	24
10	Hasil evaluasi data uji model prediksi PM2.5 <i>confidence</i> tinggi sedang	25
11	Hasil evaluasi data uji model prediksi PM2.5 <i>confidence</i> tinggi	26

DAFTAR LAMPIRAN

1	<i>Source code</i> model prediksi dengan menggunakan LSTM	33
2	<i>Source code</i> model prediksi dengan menggunakan GRU	36