



# **MODEL KLASIFIKASI KOMODITAS DAN ESTIMASI HARI SETELAH TANAM UNTUK PADI DAN TEBU MENGUNAKAN *MACHINE LEARNING***

**FATMI AULIA HANUM**



**PROGRAM MAGISTER ILMU KOMPUTER  
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2026**



## PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul Model Klasifikasi Komoditas dan Estimasi Hari Setelah Tanam untuk Padi dan Tebu Menggunakan *Machine 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 2026

Fatmi Aulia Hanum  
G6501221023

## RINGKASAN

FATMI AULIA HANUM. Model Klasifikasi Komoditas dan Estimasi Hari Setelah Tanam untuk Padi dan Tebu Menggunakan *Machine Learning*. Dibimbing oleh AGUS BUONO dan IMAS SUKAESIH SITANGGANG.

Swasembada pangan merupakan kemampuan suatu wilayah dalam memenuhi kebutuhan pangan secara mandiri. Sektor pertanian merupakan penopang utama dari swasembada pangan. Walau demikian, sektor pertanian berpotensi mengalami penurunan kontribusi seiring peningkatan laju pertumbuhan populasi, degradasi lahan, dan perubahan iklim. Kemampuan adaptasi menjadi krusial dalam menghadapi perkembangan dan tantangan penurunan kontribusi pertanian. Salah satu tren penurunan kontribusi pernah terjadi di Provinsi Jawa Timur pada tahun 2019 hingga 2024. Provinsi Jawa Timur merupakan wilayah pertanian terluas di Indonesia, khususnya pada komoditas padi dan tebu. Akan tetapi, proses *monitoring* komoditas dan Hari Setelah Tanam (HST) di Provinsi Jawa Timur belum dilakukan menggunakan *machine learning* dan *deep learning*. Proses *monitoring* masih dilakukan berdasarkan survei oleh Penyuluh Pertanian Lapangan (PPL). Adaptasi proses *monitoring* dapat dilakukan dengan implementasi *smart farming* berbasis pertanian presisi dengan pendayagunaan *remote sensing* dan *Artificial Intelligent* (AI) untuk menghasilkan informasi yang tepat dan cepat. Penelitian ini bertujuan untuk pembuatan model klasifikasi komoditas dengan mengimplementasikan algoritma *Random Forest* (RF), *Support Vector Machine* (SVM), algoritma *Extreme Gradient Boosting* (XGBoost), serta *Convolutional Neural Network* (CNN). Hasil klasifikasi digunakan untuk estimasi HST menggunakan analisis *phenology* indeks vegetasi tanaman.

Penelitian dilakukan dengan empat tahapan inti yaitu, pengumpulan dan seleksi data, eksplorasi dan praproses data, pemodelan, serta evaluasi. Data yang digunakan dalam klasifikasi komoditas yaitu data survei lapangan, administrasi wilayah, Sentinel 2A, Sentinel 1A, GLCM (*Gray Level Co-occurrence Matrix*), SRTM (*Shuttle Radar Topography Mission*), *terrain*, dan CHIRPS (*Climate Hazards Center Infrared Precipitation*). Pengolahan data citra dilakukan dalam satu proses *stacking* GEE. Data dibagi menjadi data latih dan data uji. Pencarian parameter model terbaik dilakukan dengan *hyperparameter tuning*. Pemodelan komoditas dilakukan dengan algoritma RF, SVM, XGBoost, dan CNN. Hasil pemodelan dievaluasi dengan akurasi, *classification report* dan *confusion matrix*.

Proses estimasi HST dilakukan setelah pengklasifikasian komoditas dan deteksi masa tanam. Proses deteksi masa tanam dan estimasi HST dilakukan dengan identifikasi *phenology*. Identifikasi *phenology* dilakukan dengan membandingkan penggunaan Sentinel 1A, Sentinel 2A, dan gabungan Sentinel 1A dan Sentinel 2A. Akurasi estimasi HST dievaluasi dengan  $R^2$ , RMSE, dan MAE.

Hasil penelitian menunjukkan proses klasifikasi komoditas terbaik diperoleh dari algoritma XGBoost, yang memiliki akurasi tertinggi sebesar 91% dengan *recall* 93% untuk tanaman padi dan 91% untuk tanaman tebu. Sementara estimasi HST terbaik diperoleh dengan menggunakan citra Sentinel 1A dengan akurasi 90% dengan rentang perbedaan hari berkisar 3 hingga 18 hari.

Kata kunci: klasifikasi komoditas, *machine learning*, padi, tebu, umur tanaman.



## SUMMARY

FATMI AULIA HANUM. Commodity Classification Model and Days After Planting Estimation for Rice and Sugarcane Using Machine Learning. Supervised by AGUS BUONO dan IMAS SUKAESIH SITANGGANG.

Food self sufficiency is the ability of a region to independently meet its food needs. The agricultural sector serves as the primary pillar of food self-sufficiency. Nevertheless, the sector has the potential to experience a decline in contribution due to increasing population growth, land degradation, and ongoing climate change. Adaptive capacity becomes crucial in addressing the evolving challenges associated with the declining contribution of agriculture. A declining trend in agricultural contribution was observed in East Java Province from 2019 to 2024. East Java Province is the largest agricultural region in Indonesia, particularly for rice and sugarcane commodities. However, the monitoring process for commodities and Days After Planting (HST) in East Java has not yet utilized machine learning and deep learning approaches. The monitoring process is still conducted through surveys by Agricultural Extension Workers (PPL). Adaptation of the monitoring process can be achieved through the implementation of smart farming based on precision agriculture by leveraging remote sensing and Artificial Intelligence (AI) to produce accurate and timely information. This study aims to develop a commodity classification model by implementing Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Convolutional Neural Network (CNN) algorithms. The classification results are then used to estimate HST through phenological analysis of vegetation indices.

The study was conducted through four main stages: data collection and selection, data exploration and preprocessing, modeling, and evaluation. The data used for commodity classification included field survey data, administrative area data, Sentinel 2A, Sentinel 1A, GLCM (Gray Level Co-occurrence Matrix), SRTM (Shuttle Radar Topography Mission), terrain, and CHIRPS (Climate Hazards Center Infrared Precipitation). Image data processing was performed through a single stacking process in GEE. The dataset was divided into training and testing data. The search for the best model parameters was carried out using hyperparameter tuning. Commodity modeling was performed using RF, SVM, XGBoost, and CNN algorithms. The modeling results were evaluated by accuracy, classification report, and confusion matrix.

The HST estimation process was carried out after commodity classification and planting season detection. The planting season detection and HST estimation were conducted through phenology identification. Phenology identification was performed by comparing Sentinel 1A, Sentinel 2A, and a combination of Sentinel 1A and Sentinel 2A. The HST estimation was evaluated using  $R^2$ , RMSE, and MAE.

The results showed that the best commodity classification performance was achieved by the XGBoost algorithm, which attained the highest accuracy of 91%, with recall values of 93% for rice and 91% for sugarcane. Meanwhile, the best HST estimation was obtained using Sentinel 1A imagery, with an accuracy of 90% and a day-difference range of 3 to 18 days.

*Keywords:* commodity classification, crop age, machine learning, rice, sugarcane.



- 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.

© Hak Cipta milik IPB, tahun 2026  
Hak Cipta dilindungi Undang-Undang

*Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.*

*Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.*

# **MODEL KLASIFIKASI KOMODITAS DAN ESTIMASI HARI SETELAH TANAM UNTUK PADI DAN TEBU MENGUNAKAN *MACHINE LEARNING***

**FATMI AULIA HANUM**

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

**PROGRAM MAGISTER ILMU KOMPUTER  
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA  
INSTITUT PERTANIAN BOGOR  
BOGOR  
2026**



Penguji Luar Komisi Pembimbing pada Ujian Tesis:

Dr. Sony Hartono Wijaya, S.Kom., M.Kom.

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 mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Judul Tesis : Model Klasifikasi Komoditas dan Estimasi Hari Setelah Tanam untuk Padi dan Tebu Menggunakan *Machine Learning*

Nama : Fatmi Aulia Hanum

NIM : G6501221023

Disetujui oleh

Pembimbing 1:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.

---

Pembimbing 2:

Prof. Dr. Imas Sukaesih Sitanggang, S.Si M.Kom.

---

Diketahui oleh

Ketua Program Studi:

Dr. Toto Haryanto, S.Kom., M.Si.

NIP 19821117 201404 1 001

---

Dekan Sekolah Sains Data, Matematika, dan Informatika:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.

NIP. 19660702 199302 1 001

---

Tanggal Ujian: 13 Maret 2026

Tanggal Lulus:

## PRAKATA

Puji dan syukur penulis panjatkan kehadiran Allah SWT karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan karya ilmiah ini yang berjudul Model Klasifikasi Komoditas dan Estimasi Hari Setelah Tanam untuk Padi dan Tebu Menggunakan *Machine Learning*. Walaupun banyak kesulitan yang penulis hadapi ketika menyusun tesis ini, namun berkat bantuan dan dorongan dari berbagai pihak akhirnya tesis ini dapat diselesaikan dengan baik. Untuk itu, perkenankanlah penulis mengucapkan terima kasih kepada:

- 1 Prof. Dr. Ir. Agus Bueno, M.Si., M.Kom, sebagai ketua komisi pembimbing dengan kesabaran dan keikhlasannya telah meluangkan waktu untuk mengarahkan dan membuka wawasan penulis dalam menyelesaikan tesis ini.
- 2 Prof. Dr. Imas Sukaesih Sitanggung, S.Si., M.Kom, sebagai anggota komisi pembimbing yang juga dengan kesabaran dan keikhlasannya telah meluangkan waktu untuk mengarahkan dan membuka wawasan penulis dalam menyelesaikan tesis ini.
- 3 Segenap dosen dan staf program studi Ilmu Komputer IPB yang telah mengajar dan membantu penulis selama mengikuti studi.
- 4 Tim Digitalisasi dan Data Science (DDS) dan Tim *Indonesia Food and Research Institute* (IFFRI) PT Pupuk Indonesia yang telah memberikan izin dan dukungan dalam penyelesaian studi.
- 5 Ayah Ibunda terkasih, Bapak Suherman dan Ibu Yurhasna yang selalu memberikan dukungan, doa, semangat, dan material dalam pembuatan tesis ini.
- 6 Hendrick Kasma Dihadja, Leli Herma Suryani, Fitri Kamella Sari, Zul Akmal Agung, Han Xiao Dong, dan Wang Xiaoyan selaku kakak dan adik yang selalu memberikan dukungan dan motivasi untuk terus belajar, berkreasi, dan berprestasi.
- 7 Sahabat seperjuangan Helda Yunita Taihutu, Hilyatul Mustafidah, Zalana Endesnah, dan teman-teman Ilmu Komputer angkatan 2022 yang selalu mendukung dan memberi semangat selama pengerjaan tesis.
- 8 Seluruh rekan – rekan yang tidak dapat penulis sebutkan, atas bantuan, kritik, saran, arahan, dan bimbingannya yang telah diberikan kepada penulis.

Penulis sepenuhnya menyadari bahwa tesis ini masih banyak kekurangan dan ketidaksempurnaan. Kritik dan saran yang bermanfaat sangat diharapkan penulis untuk lebih menyempurnakan karya tulis ini. Semoga karya ilmiah memberikan manfaat.

Bogor, April 2026

*Fatmi Aulia Hanum*



## DAFTAR ISI

DAFTAR TABEL	ix
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	x
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	3
1.4 Manfaat	3
1.5 Ruang Lingkup	3
II TINJAUAN PUSTAKA	4
2.1 Pertanian Presisi	4
2.2 Penginderaan Jauh	5
2.3 <i>Phenology</i>	6
2.4 Indeks Vegetasi	8
2.5 Model Data Spasial	12
2.6 <i>Data Time Series</i>	13
2.7 <i>Random Forest Classifier</i>	14
2.8 <i>Extreme Gradient Boosting</i>	18
2.9 <i>Support Vector Machine</i>	20
2.10 <i>Convolutional Neural Network</i>	24
2.11 Evaluasi	26
III METODE	28
3.1 Data Penelitian	28
3.2 Lingkungan Pengembangan	28
3.3 Tahapan Penelitian	29
IV HASIL DAN PEMBAHASAN	36
4.1 Pengumpulan dan Seleksi Data	36
4.2 Eksplorasi dan Preproses Data	41
4.3 Pemodelan	49
4.4 Evaluasi	55
V SIMPULAN DAN SARAN	58
5.1 Simpulan	58
5.2 Saran	58
DAFTAR PUSTAKA	59
LAMPIRAN	63
RIWAYAT HIDUP	68

## DAFTAR TABEL

1	<i>Bands</i> citra Sentinel 2	5
2	Sensitifitas penggunaan <i>band</i> (Giovos <i>et al.</i> 2021)	8
3	Penelitian terkait penggunaan algoritma RF untuk klasifikasi	16
4	Penelitian terkait penggunaan algoritma XGBoost untuk klasifikasi	19
5	Penelitian terkait penggunaan algoritma SVM untuk klasifikasi	22
6	Penelitian terkait penggunaan algoritma CNN untuk klasifikasi	25
7	Tabel atribut hasil deliniasi citra Sentinel 2A	28
8	Tabel atribut hasil deliniasi citra Sentinel 2A	32
9	Format data survei	36
10	Data survei hasil <i>cleansing</i>	38
11	Visualisasi indeks vegetasi	42
12	Jumlah data <i>label annotation</i>	43
13	Identifikasi jumlah <i>outlier</i>	45
14	Parameter pemodelan	49
15	Uji harmonik komoditas padi	51
16	Uji coba harmonik pada komoditas tebu	52
17	Perbandingan HST survei dengan estimasi HST <i>phenology</i>	55
18	Akurasi pemodelan	56
19	<i>Classification report</i> pemodelan pada keseluruhan data	57
20	Akurasi estimasi HST	57

## DAFTAR GAMBAR

1	Siklus pengolahan data pertanian presisi (Gursoy dan Akin 2024)	4
2	Gambaran <i>phenology</i> tanaman padi (Sukojo dan Kurniawan 2021)	6
3	Gambaran <i>phenology</i> tanaman tebu (Cahyono <i>et al.</i> 2022)	7
4	Identifikasi <i>key phenology phase</i> (Zhao <i>et al.</i> 2024)	7
5	Gambaran hasil proses <i>harmonic</i> (Philipp <i>et al.</i> 2021)	11
6	Visualisasi pola spektral indeks. (a) Vegetasi pada tanaman padi. (b) Vegetasi pada tanaman lainnya (Triscowati <i>et al.</i> 2020)	12
7	Data vektor (Campbell dan Shin 2011)	13
8	Data raster (Campbell dan Shin 2011)	13
9	Karakteristik utama permasalahan <i>time series</i> (Marshall dan Cai 2019)	14
10	Struktur <i>tree</i> pada model <i>Random Forest</i> (Chowdhury 2024)	15
11	Tahapan algoritma XGBoost (Bushenkova <i>et al.</i> 2024)	18
12	Representasi <i>hyperplane</i> keputusan optimal permasalahan linear	20
13	Arsitektur CNN-1D (Youssef <i>et al.</i> 2022)	24
14	Arsitektur CNN-2D (Youssef <i>et al.</i> 2022)	25
15	Tahapan inti klasifikasi komoditas dan prediksi HST	29
16	Tahapan klasifikasi komoditas	30
17	Diagram input dan <i>output</i> klasifikasi komoditas	30
18	Tahapan estimasi HST	31
19	Diagram <i>input</i> dan <i>output</i> prediksi HST	31
20	Proses <i>cleansing</i> dan <i>stacking</i> pada GEE	33



21	Data survei komoditas dan umur tanaman	36
22	Contoh kesalahan pengambilan data	37
23	Contoh kesalahan format tanggal	37
24	Penambahan informasi HST	37
25	Deliniasi label kelas klasifikasi	38
26	Administrasi wilayah dan pembuatan <i>grid</i>	39
27	<i>Band</i> citra Sentinel 2A dan Sentinel 1A	39
28	Citra Sentinel 1A	40
29	Data SRTM	40
30	Data CHIRPS	40
31	Praproses <i>speckel noise</i> Sentinel 1A	41
32	Perbandingan proses <i>cloud masking</i>	41
33	Proses penambalan citra dengan <i>Harmonic Smoothing</i>	42
34	Visualisasi distribusi data	44
35	<i>Outlier</i> pada distribusi data	45
36	<i>Outlier</i> data setelah <i>isolation forest</i>	46
37	Visualisasi distribusi data setelah <i>isolation forest</i>	47
38	Hasil klasifikasi komoditas padi dan tebu	50
39	Penghitungan <i>harmonic smoothing</i> dan median bulanan	50
40	Titik koordinat uji jumlah harmonik	51
41	Pendeteksian nilai minimal dan maksimal berdasarkan H1 hingga H6	54
42	Visualisasi masa tanam	54
43	Perbedaan musim tanam dalam 1 lahan	54
44	<i>Confussion matrix</i> pemodelan	56

## DAFTAR LAMPIRAN

1	Format dan distribusi data lapangan	64
2	Korelasi Fitur	65
3	Uji <i>mutual information</i> dan ANOVA fitur	66
4	Uji <i>effect size</i> dan <i>kruskal wallis</i> fitur	67