



**PENERAPAN MODEL *CATEGORICAL BOOSTING* DALAM
MENGKLASIFIKASIKAN PENYAKIT PADA DAUN TOMAT**

SKRIPSI

**untuk memenuhi persyaratan
dalam menyelesaikan program sarjana Strata-1 Statistika**

**Oleh
FITRIA RAHMAH
NIM 2111017220022**

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JULI 2025**



**PENERAPAN MODEL *CATEGORICAL BOOSTING* UNTUK
KLASIFIKASI PENYAKIT PADA DAUN TOMAT**

PROPOSAL SKRIPSI

Oleh:
FITRIA RAHMAH
NIM. 2111017220022

**PROGRAM STUDI S-1 STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU
JULI 2025**

SKRIPSI

PENERAPAN MODEL CATEGORICAL BOOSTING TERHADAP
KLASIFIKASI PENYAKIT PADA DAUN TOMAT

Oleh
Fitria Rahmah
NIM. 2111017220022

Telah diseminarkan pada hari Jumat, tanggal 11 Juli 2025 dan disetujui oleh dosen pembimbing dan dosen penguji sebagai berikut:

Pembimbing I



Selvi Annisa, S.Si., M.Si
NIP. 199212262022032016

Penguji I



Dewi Sri Susanti, S.Si., M.Si
NIP. 197305161999032002

Pembimbing II



Prof. Dewi Angraini, S.Si.,
M.App.Sci., Ph.D
NIP. 199212262022032016

Penguji II



Triando Hamongan Saragih, S.Kom.,
M.Kom
NIP. 19890713201801213001

Banjarbaru, 25 Juli 2025

Mengetahui,

Koordinator Program Studi Statistika FMIPA ULM



Prof. Dewi Angraini, S.Si., M.App.Sci., Ph.D
NIP. 198303282005012001

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 11 Juli 2025



Fitria Rahmah
NIM. 2111017220022

ABSTRAK

Penerapan Model *Categorical Boosting* Dalam Mengklasifikasikan Penyakit pada Daun Tomat (Oleh: Fitria Rahmah; Pembimbing: Selvi Annisa dan Dewi Anggraini, 2025; 57 halaman)

Tomat merupakan tanaman hortikultura penting yang rentan terhadap berbagai penyakit daun, seperti *early blight* dan *late blight*, yang dapat menurunkan kualitas dan kuantitas hasil panen. Jenis penyakit tersebut sulit dikenali hanya melalui pengamatan gambar secara langsung, sehingga diperlukan metode klasifikasi citra berbasis *machine learning*. Penelitian ini menerapkan model *Categorical Boosting (CatBoost)* yang dikombinasikan dengan ekstraksi fitur dari arsitektur VGG16 guna mengklasifikasikan penyakit daun tomat. Data yang digunakan berasal dari platform PlantVillage melalui situs Kaggle dengan total 3.300 gambar berukuran 256×256 yang terdiri dari tiga kategori: daun sehat, terpapar *early blight*, dan *late blight*. Setelah dilakukan *data cleaning*, jumlah gambar yang tidak buram adalah 3285 dan di *resize* menjadi ukuran 224×224, lalu dilanjutkan proses ekstraksi fitur, data dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Model *CatBoost* kemudian dilatih dan dievaluasi menggunakan metrik akurasi. Hasil penelitian menunjukkan bahwa model *CatBoost* mampu mengklasifikasikan penyakit daun tomat dengan akurasi keseluruhan sebesar 94%. Akurasi per kelas juga menunjukkan hasil tinggi, yaitu 90% untuk *early blight*, 99.5% untuk *late blight*, dan 93% untuk daun sehat. Hasil ini menunjukkan bahwa kombinasi VGG16 dan *CatBoost* efektif dalam mendeteksi penyakit daun tomat berbasis citra digital, serta berpotensi mendukung deteksi dini dalam bidang pertanian.

Kata Kunci: *Early Blight*, *Late Blight*, Daun Tomat, Arsitektur VGG16, *Categorical Boosting*

ABSTRACT

Application of Categorical Boosting Model in Classifying Diseases of Tomato Leaves (By: Fitria Rahmah; Supervisors: Selvi Annisa and Dewi Anggraini, 2025; 57 pages)

Tomato is an important horticultural crop that is susceptible to various leaf diseases, such as early blight and late blight, which can reduce the quality and quantity of the crop. The type of disease is difficult to recognize only through direct image observation, so a machine learning-based image classification method is needed. This research applies the Categorical Boosting (CatBoost) model combined with feature extraction from the VGG16 architecture to classify tomato leaf diseases. The data used comes from the PlantVillage platform through the Kaggle site with a total of 3,300 images of 256×256 size consisting of three categories: healthy leaves, exposed to early blight, and late blight. After data cleaning, the number of images that are not blurry is 3285 and resized to a size of 224×224, then continued the feature extraction process, the data is divided into 80% training data and 20% test data. The CatBoost model is then trained and evaluated using accuracy metrics. The results showed that the CatBoost model was able to classify tomato leaf diseases with an overall accuracy of 94%. Accuracy per class also showed high results, namely 90% for early blight, 99.5% for late blight, and 93% for healthy leaves. These results show that the combination of VGG16 and CatBoost is effective in detecting digital image-based tomato leaf diseases, and has the potential to support early detection in agriculture.

Keywords: Early Blight, Late Blight, Tomato Leaf, VGG16 Architecture, Categorical Boosting

PRAKATA

Puji syukur saya haturkan atas kehadiran Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan penulisan Tugas Akhir dengan judul "Penerapan Model *Categorical Boosting* Dalam Mengklasifikasikan Penyakit pada Daun Tomat". Penyusunan Tugas Akhir ini bertujuan untuk memenuhi persyaratan dalam rangka menyelesaikan program sarjana di Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lambung Mangkurat. Selama proses penyusunan ini, tentunya terdapat banyak pihak yang senantiasa memberikan bantuan, perhatian, dan dukungan. Oleh karena itu, saya sampaikan ucapan terima kasih yang sebesar-besarnya kepada pihak-pihak terkait yaitu sebagai berikut.

1. Allah *Subhanahu Wa Ta'ala*, yang selalu memberikan kekuatan, kelancaran, dan kemudahan dalam proses pengerjaan Tugas Akhir;
2. Keluarga terutama orang tua yang senantiasa memberikan doa dan dukungan selama proses penyusunan Tugas Akhir;
3. Prof. Dewi Anggraini, S.Si., M.App.Sci. sebagai Kepala Jurusan Statistika sekaligus dosen pembimbing dan Ibu Selvi Annisa, S.Si., M.Si. Ph.D. selaku dosen pembimbing yang telah meluangkan tenaga, waktu dan pikiran untuk memberikan bimbingan selama proses penyusunan Tugas Akhir;
4. Ibu Dewi Sri Susanti, S.Si., M.Si. dan Bapak Triando Hamonangan Saragih S.Kom., M.Kom. selaku dosen penguji yang telah memberikan masukan dan bantuan dalam proses perbaikan Tugas Akhir;
5. Bapak dan Ibu dosen pengajar serta staf Program Studi Statistika FMIPA ULM yang telah memberikan ilmu, motivasi, nasihat, kelancaran, dan dukungan selama masa perkuliahan.
6. Teman-teman yang selalu memberikan dukungan, semangat, dan selalu menemani selama penyusunan Tugas Akhir, khususnya teman-teman "Taman Paud", "Jangan Suka Tantrum", dan "7 Pilar Smanduta";
7. Pacar saya yang telah membantu dan menemani saya selama pengerjaan Tugas Akhir;
8. Diri saya sendiri yang telah bertahan sekuat tenaga dalam menyelesaikan Tugas Akhir ini.

Tugas Akhir ini tentunya masih belum mendekati kata sempurna, sehingga kritik dan saran yang membangun dari berbagai pihak sangat diharapkan dalam perbaikan Tugas Akhir ini. Diharapkan Tugas Akhir ini dapat memberikan manfaat untuk penelitian selanjutnya serta pembaca atau pihak lainnya.

Banjarbaru, 11 Juli 2025

Fitria Rahmah

DAFTAR ISI

PERNYATAAN	Kesalahan! Bookmark tidak ditentukan.
ABSTRAK.....	iii
ABSTRACT	iv
PRAKATA	v
DAFTAR GAMBAR	vii
DAFTAR TABEL.....	viii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Kajian Penelitian Terdahulu	5
2.2 Kajian Teori.....	6
2.2.1 Daun Tomat Sehat.....	6
2.2.2 Daun Tomat Terpapar <i>Early Blight</i>	7
2.2.3 Daun Tomat Terpapar <i>Late Blight</i>	8
2.2.4 <i>Preprocessing Data</i>	10
2.2.5 Arsitektur VGG16	10
2.2.6 <i>Categorical Boosting</i>	14
2.2.7 Evaluasi Model.....	19
BAB III METODE PENELITIAN	22
3.1 Sumber Data.....	22
3.2 Variabel Penelitian	22
3.3 Prosedur Penelitian.....	22
BAB IV HASIL DAN PEMBAHASAN.....	26
4.1 <i>Preprocessing Data</i>	26
4.2 Analisis Deskriptif.....	28
4.3 Arsitektur VGG16 Sebagai Ekstraksi Fitur Gambar.....	29
4.4 <i>Training</i> dengan Model <i>Categorical Boosting</i>	34
4.5 Evaluasi Model pada Data Test.....	39
4.5.1 Akurasi Perkelas.....	39
4.5.2 Akurasi Keseluruhan.....	40
BAB 5 PENUTUP.....	41
5.1 Kesimpulan	41
5.2 Saran.....	41
DAFTAR PUSTAKA	43
LAMPIRAN.....	50

DAFTAR GAMBAR

Gambar 1.1 Total Produksi Tomat pada 10 Negara Teratas di Asia.....	1
Gambar 1.2 Jenis Daun Tomat	2
Gambar 2.1 Daun Tomat Sehat.....	7
Gambar 2.2 Daun Tomat yang Terserang Penyakit <i>Early Blight</i>	8
Gambar 2.3 Daun Tomat yang Terserang Penyakit <i>Late Blight</i>	9
Gambar 2.4 Ekstraksi Fitur pada Lapisan Konvolusi.....	11
Gambar 2.5 Proses Konvolusi Menggunakan <i>Padding</i>	12
Gambar 2.7 Proses <i>Max Pooling Layer</i> dengan <i>Stride 2</i>	13
Gambar 2.8 Iterasi Pohon Pertama dan Kedua pada <i>CatBoost</i>	16
Gambar 2.9 Iterasi Pohon ke-N	17
Gambar 3.1 Alur Penelitian.....	25
Gambar 4.1 Contoh Gambar Daun Tomat yang Buram.....	26
Gambar 4.2 Gambar Sebelum dan Sesudah di <i>Resize</i>	27
Gambar 4.3 Gambar yang Dikonversi Menjadi Saluran Warna RGB	27
Gambar 4.4 Proporsi Daun Tomat Sehat, <i>Early Blight</i> , dan <i>Late Blight</i>	28
Gambar 4.5 <i>Output Layer</i> Pertama.....	32
Gambar 4.6 <i>Output Max Pooling</i> pada <i>Layer</i> Pertama	33
Gambar 4.7 Pohon Pertama <i>Categorical Boosting</i>	36

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	5
Tabel 2.2 Arsitektur VGG16 yang Digunakan Sebagai Ekstraksi Fitur	14
Tabel 2.3 <i>Confusion Matrix</i> pada Data Jenis Daun Tomat	19
Tabel 2.4 Formula Pembentukan <i>Confusion Matrix</i> Data Multikelas.....	20
Tabel 3.1 Variabel Data Pengamatan	22
Tabel 4.1 Tabulasi Data setelah Penggunaan Arsitektur VGG16 pada Data Gambar.....	34
Tabel 4.2 <i>Tuning Parameter</i> pada Model <i>Categorical Boosting</i>	35
Tabel 4.3 <i>Confusion Matrix</i> Daun Tomat Sehat, Terpapar <i>Early Blight</i> , dan Terpapar <i>Late Blight</i>	39

PRODI STATISTIKA