

**ANALISIS PREDIKSI VISIBILITAS HILAL BERBASIS
MULTINOMIAL LOGISTIC REGRESSION UNTUK
MENDUKUNG PENENTUAN AWAL BULAN
HIJRIAH DI INDONESIA,
MALANG RAYA**

TESIS

**Oleh:
TOMY IVAN SUGIHARTO
NIM. 240605210004**



**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

HALAMAN PENGAJUAN

**ANALISIS PREDIKSI VISIBILITAS HILAL BERBASIS *MULTINOMIAL*
LOGISTIC REGRESSION UNTUK MENDUKUNG PENENTUAN
AWAL BULAN HIJRIAH DI INDONESIA,
MALANG RAYA**

TESIS

Diajukan kepada:

**Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

Oleh:

**TOMY IVAN SUGIHARTO
NIM. 240605210004**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**ANALISIS PREDIKSI VISIBILITAS HILAL BERBASIS *MULTINOMIAL*
LOGISTIC REGRESSION UNTUK Mendukung Penentuan
AWAL BULAN HIJRIAH DI INDONESIA,
MALANG RAYA**

TESIS

Diajukan kepada:

**Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

Oleh:

**TOMY IVAN SUGIHARTO
NIM. 240605210004**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

HALAMAN PERSETUJUAN

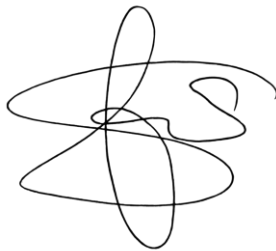
**ANALISIS PREDIKSI VISIBILITAS HILAL BERBASIS *MULTINOMIAL
LOGISTIC REGRESSION* UNTUK MENDUKUNG PENENTUAN
AWAL BULAN HIJRIAH DI INDONESIA,
MALANG RAYA**

TESIS

**Oleh:
TOMY IVAN SUGIHARTO
NIM. 240605210004**

Telah diperiksa dan disetujui untuk diuji:
Tanggal: 12 November 2025

Pembimbing I,



Dr. Mokhamad Amin Hariyadi, M.T
NIP. 19670018 200501 1 001

Pembimbing II,



Dr. Totok Chamidy, M.Kom
NIP. 19691222 200604 1 001

Mengetahui dan Mengesahkan
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. H. Muhammad Faisal, S.Kom., M.T.
NIP. 19740510 200501 1 007

HALAMAN PENGESAHAN





ANALISIS PREDIKSI VISIBILITAS HILAL BERBASIS *MULTINOMIAL LOGISTIC REGRESSION* UNTUK MENDUKUNG PENENTUAN AWAL BULAN HIJRIAH DI INDONESIA, MALANG RAYA

TESIS

Oleh:
TOMY IVAN SUGIHARTO
NIM. 240605210004

Telah Dipertahankan di Depan Dewan Penguji Tesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan\
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 12 November 2025

Susunan Dewan Penguji

Penguji I	:	<u>Dr. Irwan Budi Santoso, M.Kom</u> NIP. 19770103 201101 1 004	()
Penguji II	:	<u>Dr. Cahyo Crysdian, M.Cs</u> NIP. 19740424 200901 1 008	()
Pembimbing I	:	<u>Dr. Mokhamad Amin Hariyadi, M.T</u> NIP. 19670018 200501 1 001	()
Pembimbing II	:	<u>Dr. Totok Chamidy, M.Kom</u> NIP. 19691222 200604 1 001	()

Mengetahui dan Mengesahkan
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. H. Muhammad Faisal, S.Kom., M.T.
NIP. 19740510 200501 1 007

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Name : Tomy Ivan Sugiharto
NIM : 240605210004
Program Studi : Magister Informatika
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Tesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan Tesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 12 November 2025
Yang Membuat Pernyataan,



Tomy Ivan Sugiharto
240605210004

MOTTO

أَنْ تَكُونَ عَالِمًا أَمْرٌ مُهِمٌّ، وَلَكِنَّ أَنْ تَكُونَ ذَا أَدَبٍ أَهَمُّ؛ فَلَا عِلْمَ نَافِعًا بِغَيْرِ أُسَاسٍ مِنَ الْخُلُقِ وَالْأَدَبِ الْقَوِيِّ

“Menjadi berilmu itu penting, tetapi menjadi beradab jauh lebih penting; tidak ada ilmu yang bermanfaat tanpa pondasi akhlak dan adab yang kuat.”

HALAMAN PERSEMBAHAN

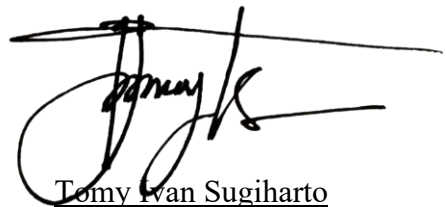
Assalamu'alaikum Wr. Wb.

Kupersembahkan karya Tesis terbaikku ini kepada:

1. Orang tua saya tercinta, TS dan IL, saudara serta seluruh keluarga besar, yang telah memberikan dukungan moral, materiel, dan doa yang tiada henti sehingga menjadi kekuatan terbesar bagi penulis.
2. AA yang selalu menjadi pendukung dan penyemangat terus-menerus.
3. Sahabat-sahabat dan budies friend Ahmad Zarkoni, serta rekan-rekan seperjuangan Syahreni, Fanny Brawijaya, Holidiyatul Amani yang telah menjadi teman diskusi dan pemberi semangat dalam suka maupun duka selama menempuh studi.

Wassalamu'alaikum Wr. Wb.

Malang, 12 November 2025



Tomy Ivan Sugiharto
NIM. 240605210004

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT, Tuhan semesta alam, yang atas rahmat, taufik, dan hidayah-Nya, penulis dapat menyelesaikan penyusunan Tesis, Tesis ini disusun sebagai salah satu syarat untuk memperoleh gelar Magister (S2) pada Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Penyusunan Tesis ini tidak akan terwujud tanpa bimbingan, arahan, motivasi, dan bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada:


1. Bapak Dr. H. Agus Mulyono, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang.
2. Bapak Prof. Dr. Muhammad Faisal, M.T., selaku Ketua Program Studi Informatika UIN Maulana Malik Ibrahim Malang.
3. Bapak Dr. Ir. M. Amin Hariyadi, M.T., selaku Dosen Pembimbing I, yang telah meluangkan waktu, tenaga, dan pikiran dengan penuh kesabaran dalam memberikan arahan, bimbingan, dan motivasi yang tak ternilai harganya selama proses penelitian ini.
4. Bapak Dr. Totok Chamidy, M.Kom., selaku Dosen Pembimbing II, yang telah memberikan masukan, koreksi, dan wawasan berharga demi kesempurnaan penelitian ini.
5. Bapak Dr. Irwan Budi Santoso, M.Kom., dan Bapak Dr. Cahyo Crysdian, M.Cs., selaku penguji pada sidang Tesis. Serta Ibu Dr. Ririen Kusumawati, M.Kom., selaku dosen wali, Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang, atas ilmu dan pelayanan yang telah diberikan.
6. Seluruh pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan bantuan baik secara langsung maupun tidak langsung.

Penulis menyadari sepenuhnya bahwa Tesis ini masih jauh dari kesempurnaan, baik dari segi substansi maupun metodologi. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan untuk perbaikan di masa depan.

Akhir kata, semoga Tesis ini dapat memberikan manfaat dan kontribusi, baik secara teoretis bagi pengembangan ilmu pengetahuan di bidang Informatika dan Astronomi, maupun secara praktis bagi masyarakat dan lembaga terkait.

Wassalamu'alaikum Wr. Wb.

Malang, 12 November 2025



Tomy Ivan Sugiharto
NIM. 240605210004

DAFTAR ISI

HALAMAN PENGANTAR.....	i
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
ABSTRAK	xiii
ABSTRACT.....	xiv
مستخلص البحث.....	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Pernyataan Masalah	6
1.3 Tujuan Penelitian	7
1.4 Hipotesis	7
1.5 Manfaat Penelitian	8
1.6 Batasan Masalah	9
1.7 Sistematika Penulisan	9
BAB II STUDI PUSTAKA.....	12
2.1 Prediksi Visabilitas <i>Hilal</i>	12
2.2 Kerangka Teori	19
2.3 Integrasi Al-Qur'an, Tafsir dan Hadis	21
2.4 Sistem Kalender <i>Hijriah</i>	23
2.5 Astronomi Posisi (Gerak Bulan dan Matahari)	24
2.6 Kriteria Visibilitas <i>Hilal</i>	25
BAB III DESAIN DAN INTEGRASI SISTEM PENELITIAN	26
3.1 Kerangka Konsep.....	26
3.2 <i>Multinomial Logistic Regression</i> (MLR).....	27
3.3 Persiapan Data	29
3.4 Desain Sistem	31
3.4.1 Persiapan Dataset	32
3.4.2 Persiapan Pemodelan Dataset Baru.....	34
3.4.3 Tahap Pembangunan Arsitektur Model.....	35
3.4.4 Tahap Evaluasi Performa Model.....	37
3.4.5 Tahap Interpretasi dan Visualisasi	39
3.5 Instrumen Penelitian	39
3.5.1 Variabel Independen	39
3.5.2 Variabel Dependen	40
BAB IV UJI COBA	41
4.1 Skenario Uji Coba.....	41

4.2	Hasil Uji Coba	44
4.2.1	Analisis Awal dan Visualisasi Hubungan Antar Variabel (EDA)	44
4.2.2	Transformasi Waktu	47
4.2.3	Rekayasa Fitur (<i>Feature Engineering</i>).....	47
4.2.4	Finalisasi Dataset dan Seleksi Variabel	49
4.2.5	Pemodelan dan Evaluasi.....	52
4.3	Pembahasan Umum Model Formulasi Matematis.....	74
4.3.1	Model Logit (Linear Predictors):	75
4.3.2	Softmax (Hitung Probabilitas)	76
4.4	Diskusi Hasil Eksperimen dan Temuan Kunci	78
4.4.1	Dampak Parameter Regularisasi (C) terhadap Kinerja Model.....	78
4.4.2	Dampak Rasio Pembagian Data (Skenario A vs. B vs. C).....	79
4.4.3	Pembahasan Hasil	79
4.5	Validasi Eksternal	80
4.5.1	Validasi Parameter Elongasi terhadap Data BMKG	80
4.5.2	Uji Coba Model pada Data Proyeksi 1446 H.....	82
BAB V PENUTUP.....		85
5.1	Kesimpulan	85
5.2	Saran	88
5.3	Penelitian Masa Depan	88
DAFTAR PUSTAKA		90
LAMPIRAN I		94
LAMPIRAN II		97

DAFTAR GAMBAR

Gambar 2.1 Kerangka Teori.....	20
Gambar 2.2 Ilustrasi Posisi Bulan dan Matahari.....	24
Gambar 3.1 Kerangka Konsep	26
Gambar 3.2 Arsitektur <i>Multinomial Logistik Regression</i> (MLR).....	27
Gambar 3.3 Desain Sistem.....	32
Gambar 4.1 <i>Koefisien Matrikx</i> Uji Coba.....	42
Gambar 4.2 Distribusi Tinggi Bulan dan Tinggi Matahari	45
Gambar 4.3 Hubungan Beda Azimuth dengan Kategori <i>Hilal</i>	45
Gambar 4.4 Korelasi Elongasi dengan Tinggi Bulan.....	46
Gambar 4.5 Korelasi Matriks Berdasarkan Dataset Awal	46
Gambar 4.6 Bentuk Transformasi Pada Dataset	47
Gambar 4.7 Korelasi Berdasarkan Dataset Elongasi_B.....	51
Gambar 4.8 Peta Elongasi Awal Bulan Ramadan 1446 H.....	81
Gambar 4.9 Peta Elongasi Awal Bulan Syawal 1446 H	81
Gambar 4.10 Peta Elongasi Awal Bulan Zulhijah 1446 H	82
Gambar 4.11 Visual Prediksi Peta Perhitungan (Bulan Ramadhan 1446 H).....	83
Gambar 4.12 Visual Prediksi Peta Perhitungan (Bulan Syawal 1446 H)	83
Gambar 4.13 Visual Prediksi Peta Perhitungan (Bulan Zulhijah 1446 H)	83

DAFTAR TABEL

Tabel 2.1 Deretan Penelitian Pendukung Terdahulu.....	17
Tabel 3.1 Keterangan Variabel Dataset	29
Tabel 3.2 Dataset Visibilitas <i>Hilal</i>	30
Tabel 4.1 Hasil Uji Coba.....	41
Tabel 4.2 <i>Hilal</i> Dataset Akhir	50
Tabel 4.3 Variabel Input	52
Tabel 4.4 Variabel Fitur (X).....	52
Tabel 4.5 Variabel Target (y).....	53
Tabel 4.6 Proporsi Pembagian Dataset	53
Tabel 4.7 Hasil Standarisasi dengan <i>StandardScaler</i>	54
Tabel 4.8 Konfigurasi Model A	54
Tabel 4.9 Hasil Metrik Skenario A (60:40)	55
Tabel 4.10 Koefisien Regresi untuk Fitur yang Dipilih (Model A).....	55
Tabel 4.11 <i>Log-Odds</i> (Koefisien Regresi) untuk Kategori ‘Gelap’ Model A.....	56
Tabel 4.12 <i>Odds Ratio</i> untuk Kategori ‘Gelap’ Model A.....	57
Tabel 4.13 Detail Visual Model Skenario A	59
Tabel 4.14 Konfigurasi Model B	60
Tabel 4.15 Hasil Metrik Skenario B (70:30).....	61
Tabel 4.16 Koefisien Regresi untuk Fitur yang Dipilih (Model B)	61
Tabel 4.17 <i>Log-Odds</i> (Koefisien Regresi) untuk Kategori ‘Gelap’ Model B.....	62
Tabel 4.18 <i>Odds Ratio</i> untuk Kategori ‘Gelap’ Model B	63
Tabel 4.19 Detail Visual Model Skenario B	65
Tabel 4.20 Konfigurasi Model C	67
Tabel 4.21 Hasil Metrik Skenario C (80:20).....	67
Tabel 4.22 Koefisien Regresi untuk Fitur yang Dipilih (Model C)	68
Tabel 4.23 <i>Log-Odds</i> (Koefisien Regresi) untuk Kategori ‘Gelap’ Model C.....	69
Tabel 4.24 <i>Odds Ratio</i> untuk Kategori ‘Gelap’ Model C	70
Tabel 4.25 Detail Visual Model Skenario C	72

ABSTRAK

Sugiharto, Tomy Ivan. 2025. **Analisis Prediksi Visibilitas *Hilal* Berbasis *Multinomial Logistic Regression* Untuk Mendukung Penentuan Awal Bulan *Hijriah* Di Indonesia, Malang Raya**. Tesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing (I) Dr. Mokhammad Amin Hariyadi, M.T (II) Dr. Totok Chamidy, M.Kom

Kata Kunci: *Multinomial Logistic Regression*, Visibilitas *Hilal*, Model Pakar, Analisis Representasi Model, Klasifikasi, *Hisab-Rukyat*.

Penentuan awal bulan *Hijriah* di Indonesia kerap menghadapi perbedaan antara metode *hisab* dan *rukya* yang dipengaruhi oleh kompleksitas faktor astronomis seperti elongasi, tinggi bulan, dan beda azimuth, sehingga diperlukan pendekatan ilmiah yang objektif dan terukur untuk meminimalkan ketidakpastian visibilitas *hilal*. Penelitian ini menerapkan metode *Multinomial Logistic Regression* (MLR) untuk melakukan klasifikasi visibilitas *hilal* berdasarkan data observasi BMKG serta memodelkan hubungan antar variabel astronomis tersebut secara komputasional. Proses penelitian meliputi pengolahan dataset, rekayasa fitur, eksplorasi hubungan variabel, pelatihan model pada tiga skenario pembagian data (60:40, 70:30, 80:20), serta evaluasi menggunakan metrik *Accuracy*, *Precision*, *Recall*, *F1-Score*, *MAE*, dan *RMSE*. Hasil penelitian menunjukkan bahwa MLR mampu memodelkan dan memprediksi visibilitas *hilal* secara konsisten, dengan performa terbaik pada skenario 80:20 yang menghasilkan akurasi tinggi dan kestabilan prediksi, serta koefisien regresi yang menunjukkan pengaruh signifikan elongasi, tinggi bulan, dan beda azimuth terhadap kategori visibilitas (“Gelap”, “Samar”, “Terang”). Model ini juga terbukti dapat merepresentasikan pola Model Pakar (*rule-based*) melalui konsistensi probabilitas kelas dan tren koefisien logit, serta menunjukkan keandalan melalui validasi eksternal menggunakan data Elongasi BMKG dan proyeksi awal bulan *Hijriah* 1446 H. Temuan ini menegaskan bahwa MLR merupakan pendekatan komputasional yang efektif, *interpretable*, dan dapat direplikasi untuk mendukung keputusan penentuan bulan baru *Hijriah* secara lebih akurat, sekaligus memberikan dasar ilmiah bagi harmonisasi metode *hisab-rukya* pada tingkat nasional.

ABSTRACT

Sugiharto, Tomy Ivan. 2025. **Multinomial Logistic Regression-Based Prediction Analysis of Crescent Visibility to Support the Determination of the Beginning of the Hijri Month in Indonesia, Malang Raya**. Thesis. Master of Informatics study Program Faculty of Science and Technology State Islamic University Maulana Malik Ibrahim Malang. Advisor (I) Dr. Mokhammad Amin Hariyadi, M.T (II) Dr. Totok Chamidy, M.Kom

Keywords: Multinomial Logistic Regression, Crescent Visibility, Expert System, Model Representation Analysis, Classification, *Hisab-Rukyat*.

The determination of the beginning of the Hijri months in Indonesia often encounters discrepancies between the *hisab* (astronomical calculation) and *rukyat* (observation) methods, largely due to the complexity of astronomical factors such as elongation, lunar *altitude*, and azimuth difference, thereby necessitating an objective and measurable scientific approach to reduce uncertainties in crescent visibility. This study employs the Multinomial Logistic Regression (MLR) method to classify crescent visibility based on BMKG observational data and to computationally model the relationships among these astronomical variables. The research process includes dataset preprocessing, feature engineering, exploratory analysis, model training using three data-splitting scenarios (60:40, 70:30, 80:20), and evaluation through Accuracy, Precision, Recall, F1-Score, MAE, and RMSE metrics. The findings indicate that the MLR model is capable of modeling and predicting crescent visibility consistently, with the best performance produced under the 80:20 split, yielding high accuracy and prediction stability. The regression coefficients also show that elongation, lunar *altitude*, and azimuth difference significantly influence the visibility categories (“Dark”, “Vague”, “Bright”). The model successfully represents rule-based expert patterns through consistent class probabilities and logit coefficient trends, and demonstrates reliability through external validation using BMKG elongation data and projected dates for the Hijri year 1446 AH. These results confirm that MLR is an effective, interpretable, and replicable computational approach for supporting more accurate decisions in determining the new Hijri month, while also providing a scientific foundation for harmonizing *hisab-rukyat* methodologies at the national level.

مستخلص البحث

سوغيهارتو، تومي إيفان. 2025. تحليل التنبؤ برؤية الهلال اعتماداً على الانحدار اللوجستي المتعدد لدعم تحديد بداية الشهر الهجري في إندونيسيا، منطقة مالانغ راي. رسالة ماجستير. برنامج ماجستير علم المعلومات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانغ. المشرفان: (I) د. محمد أمين هاريايدي، م.ت. (II) د. توتوك شاميدي، م.كوم.

الكلمات المفتاحية: الانحدار اللوجستي متعدد الحدود، رؤية الهلال، نموذج الخبر، تحليل تمثيل النموذج، التصنيف، حسب الرقيات.

يواجه تحديد بداية الأشهر الهجرية في إندونيسيا اختلافات متكررة بين منهج الحساب الفلكي ومنهج الرؤية البصرية، وذلك نتيجة لتعقيد العوامل الفلكية مثل الاستطالة، وارتفاع القمر، وفارق السم، مما يستلزم اعتماد مقارنة علمية موضوعية وقابلة للقياس للحد من عدم اليقين في رؤية الهلال. تعتمد هذه الدراسة على نموذج الانحدار اللوجستي المتعدد الفئات (MLR) لتصنيف إمكانية رؤية الهلال بناءً على بيانات الرصد التابعة لهيئة الأرصاد الجوية والمناخية والجيوفيزيائية (BMKG)، وكذلك لنمذجة العلاقات بين المتغيرات الفلكية حسابياً. وتشمل منهجية البحث معالجة البيانات، وهندسة الخصائص، والتحليل الاستكشافي، وتدريب النموذج باستخدام ثلاثة سيناريوهات لتقسيم البيانات (60:40، 70:30، 80:20)، إضافة إلى التقييم باستخدام مقاييس الدقة، والاستدعاء، ومعامل F1، ومتوسط الخطأ المطلق (MAE)، والجذر التربيعي لمتوسط مربع الخطأ (RMSE). وتُظهر النتائج أنّ نموذج MLR قادر على تمثيل وتوقع رؤية الهلال بشكل مستقر، وأن أفضل أداء تحقق في سيناريو 80:20 مع دقة عالية وثبات في التنبؤ. كما تكشف معاملات الانحدار عن تأثير معنوي لكل من الاستطالة، وارتفاع القمر، وفارق السم في تحديد فئات الرؤية ("مظلم"، "خافت"، "واضح"). وقد نجح النموذج في تمثيل أنماط الخبراء المبنية على القواعد من خلال اتساق احتمالات الفئات واتجاهات معاملات اللوغيت، وأثبت موثوقيته من خلال التحقق الخارجي باستخدام بيانات الاستطالة الخاصة بـ BMKG وتوقعات مطالع الأشهر الهجرية لعام 1446 هـ. وتؤكد هذه النتائج أنّ MLR يمثل مقارنة حسابية فعالة وقابلة للتفسير وتكرار الاستخدام، تسهم في دعم القرارات المتعلقة بتحديد أوائل الشهور الهجرية بشكل أكثر دقة، وتوفر أساساً علمياً لتقريب منهجي الحساب والرؤية على المستوى الوطني..

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penentuan awal bulan *Hijriah* memiliki pengaruh signifikan bagi komunitas Muslim global, termasuk di negara dengan populasi Muslim terbesar seperti Indonesia, dan secara spesifik di wilayah seperti Malang Raya, Jawa Timur. Penentuan ini merupakan bagian terpenting dari praktik keagamaan Islam, yang sangat bergantung pada pengamatan siklus bulan baru (Loucif et al., 2024). Sebagai negara dengan populasi Muslim terbesar di dunia, Indonesia memiliki peran serta tanggung jawab utama dalam memastikan akurasi penentuan awal bulan *Hijriah*, khususnya untuk bulan Ramadan dan Idul Fitri (Mohd Nawawi et al., 2024) di mana variabel seperti elongasi dan tinggi bulan memainkan peran kunci. Allah SWT berfirman dalam Q.S. Yunus (10) ayat 5:

هُوَ الَّذِي جَعَلَ الشَّمْسُ ضِيَاءً وَالْقَمَرَ نُورًا وَقَدَرَهُ ۖ مَنَازِلَ لِتَعْلَمُوا عَدَدَ السِّنِينَ وَالْحِسَابَ ۚ مَا خَلَقَ اللَّهُ ذَلِكَ إِلَّا بِالْحَقِّ يُفَصِّلُ الْآيَاتِ لِقَوْمٍ يَعْلَمُونَ

Artinya: “Dialah yang menjadikan matahari bersinar dan bulan bercahaya. Dialah pula yang menetapkan tempat-tempat orbitnya agar kamu mengetahui bilangan tahun dan perhitungan (waktu). Allah tidak menciptakan demikian itu, kecuali dengan benar. Dia menjelaskan tanda-tanda (kebesaran-Nya) kepada kaum yang mengetahui” (Q.S. Yunus: 5).

Menurut kitab *Al-Tafsir al-Wajiz* karya al-Wahidi menjelaskan, Allah tidak hanya menciptakan langit dan bumi sebagai bukti kekuasaan-Nya, tetapi juga menjadikan matahari sebagai sumber cahaya dan energi, serta bulan sebagai pemantul cahaya matahari. Dia pula yang menetapkan orbit bumi dan bulan sehingga manusia dapat mengetahui hitungan tahun dan waktu. Seluruh ciptaan ini hadir dengan hikmah dan kebenaran, menjadi tanda-tanda kebesaran-Nya bagi orang-orang yang berpengetahuan.

Secara tradisional, metode penentuan awal bulan *Hijriah* telah mengalami banyak perkembangan dan pembaruan berkelanjutan, terutama dalam integrasi *hisab* dan *rukyat*. Hal ini sejalan dengan penelitian oleh Rahmi (2025) yang mengatakan bahwa penentuan untuk metode perhitungan (*hisab*) dengan kriteria wujudul *hilal* dan metode Imkanur *Rukyat* dengan kriteria MABIMS menyebabkan perbedaan variasi dan mengganggu keseragaman praktik ibadah dan kalender nasional (Rahmi et al., 2025).

Sedangkan penentuan dengan metode pengamatan (*rukyat*) untuk penampakan bulan baru, yang juga dikenal sebagai *hilal* telah banyak dilakukan pembaharuan dengan bantuan alat dan cara yang beragam (Koto et al., 2024; Ulum, 2021; Yunus, 2011). Meskipun dalam praktiknya masih terus berlanjut, integrasi ilmu pengetahuan modern dan tradisi keagamaan semakin penting untuk mengatasi tantangan dan memastikan keakuratannya (Sultan, 2007).

Namun, metode tradisional seperti *rukyat* sering menghadapi tantangan akurasi dan konsistensi akibat kondisi cuaca buruk, lokasi geografis yang tidak optimal, serta umur bulan yang muda faktor yang memengaruhi variabel seperti beda azimuth dan elongasi (Hoffman, 2003; Hussain, M., & Al-Mostafa, 2021). Hal ini juga dijelaskan dalam penelitian Taruna (2017) yang menemukan bahwa salah satu faktor penting dalam menentukan penampakan *hilal* adalah keberadaan objek tinggi di lokasi yang berpotensi menghalangi *hilal* (Taruna & Prakoso, 2017).

Selanjutnya, penelitian oleh Djamaluddin juga menjelaskan bahwa ketidakpastian dalam pengamatan dapat menyebabkan perbedaan waktu shalat, sehingga menciptakan ketidakpastian bagi umat Islam (Djamaluddin, 2000).

Untuk mengatasi ketidakpastian tersebut, upaya peningkatan akurasi terus dilakukan secara maksimal, misalnya melalui penerapan teknologi oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG).

Salah satunya pemanfaatan jam atom untuk mengukur *Universal Time Coordinat* (UTC) dengan presisi tinggi dalam pengamatan *hilal* (Suprihatin et al., 2021). Hal ini guna untuk mendukung kebijakan pada Kementerian Agama (Kemenag) dalam pemahaman dan akurasi dalam pengamatan Visabilitas *Hilal*. Jam atom sendiri merupakan teknologi menggunakan standar frekuensi resonansi atom sebagai penghitungnya yang digunakan sebagai tanda waktu (BMKG, 2016).

Selain itu, masalah seperti polusi cahaya yang mengganggu pengamatan juga merupakan tantangan lain yang memerlukan pemetaan memadai dan pemahaman mendalam tentang kondisi atmosfer (Badescu, 2008). Dengan demikian, penerapan metode komputasi menawarkan solusi inovatif, yang mendasari pada penelitian Muztaba (2023) yang menjelaskan bahwa algoritma *machine learning* mampu meningkatkan akurasi deteksi bulan baru dan menciptakan program otomatis yang efisien (Muztaba et al., 2023).

Menurut hasil penelitian Al-Rajab (2023) menjelaskan pendekatan berbasis data dapat meningkatkan akurasi prediksi visibilitas *hilal* secara signifikan melalui algoritma canggih yang menggunakan berbagai metode pembelajaran mesin, salah satu yang digunakan adalah *Logistic Regression* (LR), yang merupakan algoritma pembelajaran mesin berbasis statistik tetapi hanya menggunakan model untuk masalah klasifikasi biner (Al-Rajab et al., 2023).

Model non-parametrik lainnya seperti *Support Vector Machine* (SVM) dan *Artificial Neural Network* (ANN) telah mencapai akurasi tinggi ($>90\%$), tetapi model-model ini sering dianggap sebagai *black-box* sehingga menyulitkan untuk mengidentifikasi faktor-faktor dominan yang mempengaruhi visabilitas (Al-Rajab et al., 2023). Penelitian ini berfokus pada pengamatan visabilitas *hilal* di Pesisir Pantai Ngliyep, Kabupaten Malang, Indonesia salah satu dari 34 titik pengamatan nasional di mana variabel beda azimuth dan tinggi bulan sangat dipengaruhi oleh topografi pantai.

Lokasi ini dipilih karena karakteristik geografisnya yang strategis, termasuk koordinat geografis $57^{\circ}23'$ Lintang Selatan dan $112^{\circ}36'29''$ Bujur Timur (Fanani, 2012), yang minim interferensi cahaya buatan dan garis pantai yang terbuka, menciptakan kondisi optimal pengamatan *hilal* (Hasan Faadillah et al., 2024). Hal ini sesuai dengan temuan yang menekankan pada pentingnya kondisi atmosfer dan topografi terhadap keakuratan pengamatan (Utama et al., 2023).

Karakteristik pantai yang memiliki *horizon* laut yang luas memungkinkan pengamat untuk meminimalkan potensi obstruksi visual dari pegunungan atau bangunan, meningkatkan akurasi pengukuran ketinggian bulan (Taruna & Prakoso, 2017). Namun, pengaruh variabel astronomis seperti umur *hilal*, elongasi sudut, dan kondisi atmosfer mengakibatkan kesulitan dalam pemodelan yang komprehensif (Allawi, 2022; Suprihatin et al., 2021).

Dengan memanfaatkan data observasi visabilitas *hilal* dari BMKG melalui pendekatan berbasis data yang mendalam (Ma et al., 2022), penelitian ini akan menerapkan metode komputasi berbasis *Multinomial Logistic Regression* (MLR). MLR merupakan pengembangan model dari *Logistic Regression*, MLR

mengklasifikasi generalisasi regresi logistik ke masalah *multikelas*, yaitu dengan lebih dari dua kemungkinan hasil diskrit (Hadi et al., 2024).

Studi penelitian metodologis oleh Tabatabai (2025) menunjukkan bahwa pemilihan model regresi harus disesuaikan dengan karakteristik data (Tabatabai et al., 2025), apakah data cenderung memiliki *outliers* (cocok untuk *Logistic Regression*) atau mengikuti distribusi simetris (Battineni et al., 2019; Sharma et al., 2021). Penggunaan teknik ini diharapkan dapat memberikan solusi inovatif terhadap masalah yang dihadapi dalam pengamatan *hilal*, terutama dalam memahami hubungan kompleks antara faktor-faktor yang memengaruhi visabilitas *hilal* (Hadi et al., 2024; Li et al., 2017).

Hal ini juga didukung oleh penelitian Pinedo (2025) yang menganalisis hubungan antara fase bulan dengan kejadian kehamilan pada sapi *Holstein* dengan menerapkan inseminasi buatan menggunakan *Multinomial Logistic Regression* (Pinedo et al., 2025). Meskipun hasil penelitiannya hanya menunjukkan wawasan tentang faktor biologis pada hewan yang mungkin dipengaruhi oleh siklus bulan. MLR menawarkan interpretasi yang lebih jelas melalui analisis koefisien regresi, sehingga cocok untuk memberikan rekomendasi kebijakan kalender Islam.

Integrasi metode komputasional dalam penentuan visabilitas *hilal* tidak hanya memperkuat hasil pengamatan, tetapi juga membuka peluang untuk penerapan alat-alat observasi modern, seperti teleskop dan teknik pemrosesan *citra* (Hasan Faadillah et al., 2024; Umar et al., 2023). Hal ini menciptakan sinergi antara praktik keagamaan dan kemajuan teknologi, sehingga memungkinkan umat Muslim untuk mengikuti perkembangan dengan lebih cepat dan tepat waktu dalam penentuan hari-hari penting umat islam.

Dalam konteks ini, pengusulan pendekatan menggunakan *Multinomial Logistic Regression* untuk memprediksi visabilitas *hilal* berdasarkan klasifikasi elongasi, yang berlandaskan dasar hukum dari Kementerian Agama tentang kriteria bulan baru (Khoeron, 2022; Odeh, 2004). Tujuan akhirnya dapat memberikan solusi inovatif yang lebih canggih dan menjembatani kesenjangan antara tradisi dan inovasi dalam penentuan kalender *Hijriah*/Lunar serta mendukung pengambilan keputusan berbasis data dalam astronomi dan praktik keagamaan di Indonesia. Penelitian ini diharapkan meningkatkan akurasi prediksi hingga 90% di Malang Raya, mendukung unifikasi kalender *Hijriah* nasional.

1.2 Pernyataan Masalah

Metode tradisional sering tidak konsisten dalam mengenali faktor astronomis yang memengaruhi visabilitas *hilal*: Metode *hisab* dan *rukyat* sering kali tidak akurat karena perbedaan kondisi cuaca dan geografi di Indonesia, sehingga sulit menentukan hubungan antara faktor seperti beda azimuth, tinggi bulan, dan elongasi dengan visabilitas *hilal*.

Metode konvensional kurang mampu memprediksi visabilitas *hilal*: Metode tradisional tidak bisa memberikan prediksi yang baik untuk visabilitas *hilal*, terutama dengan variasi cuaca dan geografi Indonesia, sehingga perlu model seperti *Multinomial Logistic Regression* (MLR) untuk memprediksi berdasarkan elongasi, beda azimuth, dan tinggi bulan.

Belum ada pengukuran efektivitas metode komputasional dibanding metode tradisional: Efektivitas metode komputasional yang pakai data pengamatan *hilal* belum diukur dengan baik, dan belum dibandingkan dengan metode *hisab* dan *rukyat*, terutama di tengah variasi cuaca dan geografi Indonesia.

Diperlukan saran berbasis data untuk kalender *Hijriah* yang lebih tepat: Penentuan kalender *Hijriah* saat ini kurang tepat karena metode tradisional yang tidak konsisten, sehingga butuh saran dari model prediktif seperti MLR untuk praktik yang lebih baik, yang bisa digunakan Kementerian Agama dan organisasi Islam untuk keputusan bulan baru yang lebih akurat.

1.3 Tujuan Penelitian

1. Mengidentifikasi hubungan dan pengaruh faktor astronomis, khususnya beda azimuth, tinggi bulan, dan elongasi, terhadap visabilitas *hilal*.
2. Mengembangkan dan mengimplementasikan model klasifikasi berdasarkan metode *Multinomial Logistic Regression* (MLR) yang dapat mengukur dan memprediksi visabilitas *hilal* berdasarkan elongasi.
3. Mengukuran efektivitas metode komputasional yang menggunakan data observasi visabilitas *hilal*, serta melakukan perbandingan dengan metode konvensional (*hisab* dan *rukyat*).
4. Menyusun rekomendasi berbasis data untuk peningkatan praktis dalam penentuan kalender *Hijriah* yang lebih akurat, diharapkan dapat digunakan oleh Kementerian Agama dan organisasi Islam lainnya untuk keputusan yang lebih informatif dan akurat dalam penetapan bulan baru.

1.4 Hipotesis

Hipotesis penelitian ini didasarkan pada asumsi bahwa integrasi variabel astronomis utama: beda azimuth, tinggi bulan, dan elongasi dalam model *Multinomial Logistic Regression* (MLR) akan meningkatkan kehandalan model visabilitas *hilal* secara signifikan. Secara spesifik:

1. Model MLR yang dikembangkan akan mencapai kehandalan model visabilitas *hilal* lebih dari 85% berdasarkan data observasi BMKG tahun 2023-2025, dibandingkan dengan metode tradisional *hisab* dan *rukyat* yang sering terpengaruh ketidakpastian atmosfer (Al-Rajab et al., 2023).
2. Variabel elongasi akan memiliki pengaruh dominan (koefisien regresi >0.5) terhadap klasifikasi visabilitas *hilal*, diikuti oleh tinggi bulan dan beda azimuth, karena elongasi menentukan sudut iluminasi minimal sesuai kriteria MABIMS (minimal 3°), sementara beda azimuth dan tinggi bulan memodulasi faktor geofisika lokal di Malang Raya (Taruna & Prakoso, 2017).
3. Perbandingan antara model MLR dan metode konvensional akan menunjukkan penurunan eror prediksi (MAE dan RMSE <0.2) hingga 20-30%, sehingga mendukung rekomendasi berbasis data untuk Kementerian Agama dalam unifikasi kalender *Hijriah* nasional (Suprihatin et al., 2021).

1.5 Manfaat Penelitian

1. Mengukur akurasi prediksi visabilitas *hilal* menggunakan algoritma *Multinomial Logistic Regression* berbasis variabel beda azimuth, tinggi bulan, dan elongasi, dengan memanfaatkan data historis BMKG. Dengan memanfaatkan data historis dan informasi astronomis, model yang dikembangkan diharapkan dapat memberikan prediksi yang lebih tepat mengenai waktu dan kondisi yang diperlukan untuk observasi *hilal*.
2. Identifikasi hubungan kompleks antara umur bulan, ketinggian, dan faktor atmosferik lainnya yang mempengaruhi visabilitas *hilal*.
3. Pendekatan komputasional yang inovatif, penelitian ini diharapkan dapat memberikan kontribusi signifikan kepada komunitas ilmiah global,

memperkuat pemahaman tentang kedudukan dan perhitungan *hilal* dalam konteks astronomi dan budaya di berbagai belahan dunia.

1.6 Batasan Masalah

Berdasarkan batasan tersebut, hipotesis penelitian diformulasikan sebagai berikut:

1. Data yang digunakan merupakan hasil observasi *hilal* dari 34 titik Nasional Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) selama tahun 2023-2025 terakhir untuk menentukan awal bulan *Hijriah*.
2. 34 titik lokasi pengamatan difokuskan pengamatan visabilitas *hilal* di Indonesia bersama Kementrian Agama RI dan BMKG.
3. Pengukuran analisis kuantitatif dilakukan dengan model *Logistic Regression*, serta evaluasi kinerja menggunakan metrik *Accuracy*, MAE, RMSE, ROC, dan interpretasi koefisien. Serta, citra visual visabilitas *hilal* digunakan hanya sebagai validasi eksternal.

1.7 Sistematika Penulisan

Sistematika penulisan proposal Tesis ini disusun secara logis untuk membangun argumen dari konteks umum penentuan visabilitas *hilal* hingga solusi spesifik berbasis *Multinomial Logistic Regression* (MLR), dengan fokus pada variabel beda azimuth, tinggi bulan, dan elongasi. Struktur keseluruhan terdiri dari:

1. BAB I PENDAHULUAN

Menyajikan latar belakang, pernyataan masalah, tujuan, hipotesis, manfaat, batasan masalah, serta sistematika penulisan, untuk memberikan gambaran umum

tentang pentingnya prediksi visabilitas *hilal* di Indonesia, khususnya Malang Raya.

2. BAB II STUDI PUSTAKA

Membahas landasan teori meliputi sistem kalender *Hijriah*, astronomi posisi (gerak bulan dan matahari), kriteria visabilitas *hilal*, serta konsep klasifikasi dan MLR, termasuk rumusan matematis logit dan softmax untuk interpretasi variabel prediktor.

3. BAB III METODOLOGI PENELITIAN

Menjelaskan desain penelitian, instrumen (variabel independen: beda azimuth, tinggi bulan, elongasi; dependen: visabilitas *hilal*), sumber data (observasi BMKG 2023-2025), teknik pengumpulan dan analisis data menggunakan *Python* dengan *scikit-learn* untuk implementasi MLR, serta metrik evaluasi (*Accuracy*, *Precision*, *Recall*, *F1-Score*, MAE, RMSE).

4. BAB IV HASIL DAN PEMBAHASAN

Menyajikan hasil uji coba model MLR, analisis koefisien regresi untuk mengidentifikasi pengaruh variabel utama, perbandingan dengan metode tradisional, serta interpretasi pola data visabilitas *hilal* di Pantai Ngliyep, Malang Raya.

5. BAB V KESIMPULAN DAN SARAN

Merangkum temuan utama terkait akurasi prediksi berbasis MLR, implikasi untuk penentuan awal bulan *Hijriah*, serta rekomendasi untuk Kementerian Agama dan penelitian lanjutan, seperti integrasi *deep learning* untuk variabel tambahan.

6. DAFTAR PUSTAKA DAN LAMPIRAN

Mencantumkan referensi lengkap dan data pendukung, seperti kode *Python*, dataset BMKG, dan visualisasi hasil (misalnya, confusion *matrix*).

BAB II STUDI PUSTAKA

2.1 Prediksi Visabilitas *Hilal*

Tinjauan penelitian terkait bertujuan untuk mengidentifikasi landasan empiris dan celah penelitian yang mendukung analisis klasifikasi visabilitas *hilal*. Peninjauan ini mencakup studi prediktif, observasi, *policy*, dan analisis fiqh, dengan fokus pada variabel Beda Azimuth, Tinggi Bulan, dan Elongasi.

Penelitian oleh Rahmi (2025) membahas kontestasi penentuan Idul Adha antara Muhammadiyah dan Nahdatul Ulama di Indonesia, yang berakar pada perbedaan interpretasi hadis. Kontribusi peneliti terletak pada analisis perbedaan metodologi *Hisab Wujudul Hilal* dan Imkanur *Rukyat* MABIMS serta dampak sosialnya. Metode yang digunakan adalah *literature review* kualitatif dengan variabel tinggi *hilal* minimal 3° dan elongasi implisit. Hasil temuan utama menunjukkan variasi tanggal Idul Adha akibat perbedaan metode, memerlukan wisdom umat untuk mitigasi. Penelitian ini relevan dan mendukung penelitian saat ini dengan menyediakan konteks dispute lokal, namun gap penelitian ini terletak pada absennya solusi komputasi prediktif (Rahmi et al., 2025).

Penelitian oleh Gharaybeh (2025) membahas reliansi fiqh pada kalkulasi astronomi untuk penentuan awal bulan *Hijriah*, dengan fokus pada integrasi sains dan tradisi. Kontribusi peneliti terletak pada usulan integrasi kalkulasi dengan visual *sighting*, termasuk penggunaan *optical aids* yang diperbolehkan dalam fiqh. Metode yang digunakan adalah *literature review* deskriptif dengan variabel lunar *phases* dan *angles*. Hasil temuan utama menegaskan prioritas *rukya*t dengan dukungan kalkulasi sebagai pendamping. Penelitian ini berkaitan erat dan

mendukung penelitian saat ini dengan mengintegrasikan sains dan fiqh, namun gap penelitian ini muncul dari ketiadaan model prediktif untuk implementasi praktis (Gharaybeh, 2025).

Penelitian oleh Pramudya (2024) berfokus pada pengembangan teknik observasi *hilal* di lingkungan urban menggunakan instrumen modern untuk meningkatkan akurasi pengamatan. Kontribusi peneliti terwujud dalam inovasi *stacking* gambar CCD dan monitoring SQM/ASC sebagai solusi pengamatan di kondisi urban. Metode yang diterapkan meliputi observasi lapangan dengan variabel *altitude* $>5^\circ$, elongasi $6-18^\circ$, dan *Night Sky Brightness* (NSB) 18 mag/arcsec^2 . Hasil temuan utama menunjukkan efektivitas teknik tersebut dalam konteks urban Indonesia. Penelitian ini relevan dan mendukung penelitian saat ini dengan menyediakan data astronomi lokal, namun gap penelitian ini terletak pada ketiadaan pendekatan komputasi prediktif untuk mengatasi tantangan geofisika spesifik (Pramudya et al., 2024).

Penelitian oleh Koto (2024) menangani kontestasi antara *Rukyatul Hilal* dan *Hisab* Hakiki dalam penentuan awal bulan *Hijriah* di Indonesia, dengan fokus pada dampak sosialnya. Kontribusi peneliti terletak pada analisis solusi MUI fatwa 2004 dan sidang *Itsbat* untuk mencapai harmonisasi umat. Metode yang digunakan adalah *literature review* kualitatif dengan variabel Elongasi dan *altitude* sebagai titik analisis. Hasil temuan utama menegaskan validitas kedua metode, dengan dampak sosial yang signifikan memerlukan penanganan lebih lanjut. Penelitian ini memiliki hubungan kuat dan mendukung penelitian saat ini dengan menyediakan konteks dispute nasional, namun gap penelitian ini muncul dari absennya alat prediktif untuk mitigasi perbedaan tersebut (Koto et al., 2024).

Penelitian oleh Loucif (2024) membahas pengembangan kalender lunar global melalui pendekatan *machine learning* untuk mendukung unifikasi secara internasional. Kontribusi peneliti terwujud dalam *framework* ML dengan *Gradient Boosting* dan *Extra Trees*, yang mencapai *F1-score* 0.888 dan ROC 0.907. Metode yang digunakan meliputi *supervised learning* dengan *clustering regional* dan variabel *elongation*, *altitude*, dan *azimuth diff*. Hasil temuan utama menunjukkan peningkatan akurasi prediksi, terutama di wilayah Asia dan Afrika. Penelitian ini relevan dan mendukung penelitian saat ini dengan memvalidasi variabel astronomi, namun gap penelitian ini terletak pada keterbatasan model *biner* dan kurangnya adaptasi ke kondisi tropis lokal seperti Malang Raya (Loucif et al., 2024).

Penelitian oleh Syuhada (2023) adalah pemetaan posisi *hilal* terhadap penghalang geografis seperti Gunung Agung di lokasi *rukya*t Pantai Loang Baloq, Mataram. Kontribusi peneliti terletak pada identifikasi faktor geofisika yang memengaruhi visabilitas *hilal* dalam konteks tropis. Metode yang digunakan meliputi observasi lapangan selama 14 bulan dengan perhitungan segitiga bola, mengandalkan variabel Azimuth *hilal* (ΔAz), Tinggi *hilal*, dan Elongasi. Hasil temuan utama menunjukkan bahwa lokasi tersebut tidak ideal akibat penghalang dan keberawan, dengan rekomendasi perubahan lokasi untuk observasi yang lebih efektif. Penelitian ini berkaitan erat dan mendukung penelitian saat ini dengan mengintegrasikan geofisika tropis, namun gap penelitian ini muncul dari ketiadaan model prediktif untuk mengoptimalkan data tersebut (Syuhada et al., 2023).

Penelitian oleh Muztaba (2023) adalah pengembangan model deteksi *hilal* menggunakan *deep learning* untuk meningkatkan akurasi pengamatan lunar. Kontribusi peneliti terwujud melalui implementasi *Mask R-CNN* pada dataset observasi teleskop robotik. Metode yang digunakan meliputi analisis *deep learning* dengan variabel posisi *hilal* dan kondisi atmosfer, diuji pada data lunar historis. Hasil temuan utama menunjukkan kemampuan model dalam mendeteksi dan mengenali *hilal* dengan presisi tinggi. Penelitian ini relevan dan mendukung penelitian saat ini dengan menambahkan dimensi *deep learning* pada prediksi visabilitas, namun gap penelitian ini terletak pada fokus *black-box* tanpa interpretabilitas multinomial untuk konteks Malang Raya (Muztaba et al., 2023).

Penelitian oleh Mufid (2023) adalah implementasi kriteria MABIMS untuk unifikasi kalender *Hijriah* di negara-negara Brunei, Indonesia, Malaysia, dan Singapura. Kontribusi peneliti terwujud dalam analisis keberhasilan implementasi sejak 2022 serta rekomendasi dialog *socio-political* untuk mendukung harmonisasi. Metode yang diterapkan adalah *Systematic Literature Review* multidisiplin dengan variabel *Hilal Height*, *Elongation*, dan DAZ sebagai dasar analisis. Hasil temuan utama menunjukkan kesuksesan implementasi di Indonesia, meskipun memerlukan dukungan tambahan untuk mengatasi tantangan sosial. Penelitian ini relevan dan mendukung penelitian saat ini dengan mengkonfirmasi variabel astronomi dalam konteks nasional, namun gap penelitian ini terletak pada absennya model prediktif untuk mengatasi variasi lokal seperti di Malang Raya (Mufid & Djamaluddin, 2023).

Penelitian oleh Al-Rajab (2023) terkait pengembangan model prediktif visabilitas *hilal* menggunakan berbagai algoritma *machine learning* untuk

mendukung unifikasi kalender lunar secara global. Kontribusi peneliti terletak pada komparasi delapan algoritma, termasuk *Random Forest* dan *Neural Network*, yang meningkatkan akurasi prediksi dengan pendekatan *data-driven*. Metode yang digunakan meliputi analisis *supervised learning* dengan variabel *Moon Altitude*, DAZ, dan Elongasi, didukung oleh dataset global serta optimasi parameter. Hasil temuan utama menunjukkan akurasi melebihi 91% dan *Area Under Curve* (ROC) di atas 0.95, mengkonfirmasi kekuatan variabel astronomi tersebut sebagai prediktor utama. Penelitian ini memiliki hubungan kuat dan mendukung penelitian saat ini dengan memvalidasi variabel prediktif utama, namun gap penelitian ini terletak pada keterbatasan klasifikasi biner dan sifat *black-box*, yang tidak memadai untuk kompleksitas multinomial di Malang Raya. (Al-Rajab et al., 2023).

Penelitian oleh Allawi (2022) berfokus pada pengembangan model prediktif untuk visabilitas bulan sabit baru di Irak guna meningkatkan akurasi penentuan awal bulan *Hijriah*. Kontribusi peneliti terwujud melalui pengembangan model *Artificial Neural Network* (ANN) yang mampu memetakan tiga kategori visabilitas berdasarkan data lokal. Metode yang diterapkan meliputi penggunaan ANN dengan variabel Beda Azimuth (DAZ), lebar sabit (w), dan *Arc of Vision* (ARCV), diuji melalui dataset spesifik Irak. Hasil temuan utama menunjukkan akurasi sebesar 72.7%, dengan penegasan peran krusial DAZ dalam prediksi visabilitas. Penelitian ini berkaitan erat dan mendukung penelitian saat ini dengan menegaskan pentingnya DAZ sebagai variabel prediktif, namun gap penelitian ini muncul dari sifat model *black-box* yang sulit diinterpretasikan serta kurangnya adaptasi terhadap kondisi tropis Indonesia seperti Malang Raya. (Allawi, 2022).

Pada Tabel 2.1 menjelaskan deretan penelitian terkait sebagai pendukung pada kesenjangan dasar penelitian ini, sebagai rujukan yang digunakan untuk memahami literatur pada penelitian ini.

Tabel 2.1 Deretan Penelitian Pendukung Terdahulu

No	Penulis	Topik	Metode	Hasil
1	Nailur Rahmi et al. (2025)	Menganalisis kontestasi penentuan <i>hisab</i> dan <i>rukyat</i> yang berdampak pada keseragaman ibadah nasional.	Kualitatif: Analisis sekunder dari artikel dan data organisasi.	Perbedaan metode menyebabkan variasi tanggal hingga 1-2 hari, mengganggu kalender nasional.
2	Mohammed Gharaybeh (2025)	Mengintegrasikan fiqh dengan kalkulasi astronomi untuk awal bulan <i>Hijriah</i> , rekonsiliasi antara <i>rukyat</i> dan <i>hisab</i> .	Deskriptif: Analisis historis, Qur'an/hadis, dan fatwa.	Korelasi 67% antara <i>rukyat</i> dan <i>hisab</i> ; rekomendasi integrasi untuk unity.
3	Yudhiakto Pramudya et al. (2024)	Mengembangkan teknik observasi <i>hilal</i> di observatorium urban, status instrumen untuk akurasi.	Kualitatif: Pengembangan instrumen (teleskop, CCD) Analisis data astronomi: <i>Hisab</i>	Instrumen meningkatkan akurasi observasi hingga 85% di kondisi urban.
4	Ismail Koto et al. (2024)	Perbedaan penentuan awal bulan <i>Hijriah</i> (Ramadan, Syawal, Dzulhijjah)	Kualitatif: Tafsir Qur'an/hadis, Analisis konten.	<i>Rukyat</i> kuat secara spiritual, <i>hisab</i> hakiki lebih ilmiah. Pendekatan imkanur <i>rukyat</i> dan sinergi antarlembaga penting untuk kalender nasional.
5	Samia Loucif et al. (2024)	Prediksi Visabilitas <i>Hilal</i>	Supervised ML: SVM, DT, RF, LR, dll.; Feature selection: (ANOVA, PCA), hyperparameter tuning (GA, PSO).	Model Extra Trees dan ensemble menunjukkan performa terbaik (F1-score 0.8881, ROC 0.9075). Machine learning efektif untuk prediksi <i>hilal</i> global dan adaptif terhadap variasi atmosfer dan lokasi.
6	Nurnadiyah Syuhada et al. (2023)	Pemetaan posisi <i>hilal</i> terhadap Gunung Agung di lokasi <i>rukyat</i> Pantai Loang Baloq Mataram, analisis ketertampakan <i>hilal</i> .	Analisis data astronomi: <i>Hisab</i>	Lokasi tidak ideal (terhalang Gunung Agung); <i>hilal</i> kiri/kanan potensial terlihat, tepat pada gunung bergeser ke Tanjung Mina; pemetaan 12 bulan (Muharram-Dzulhijjah).

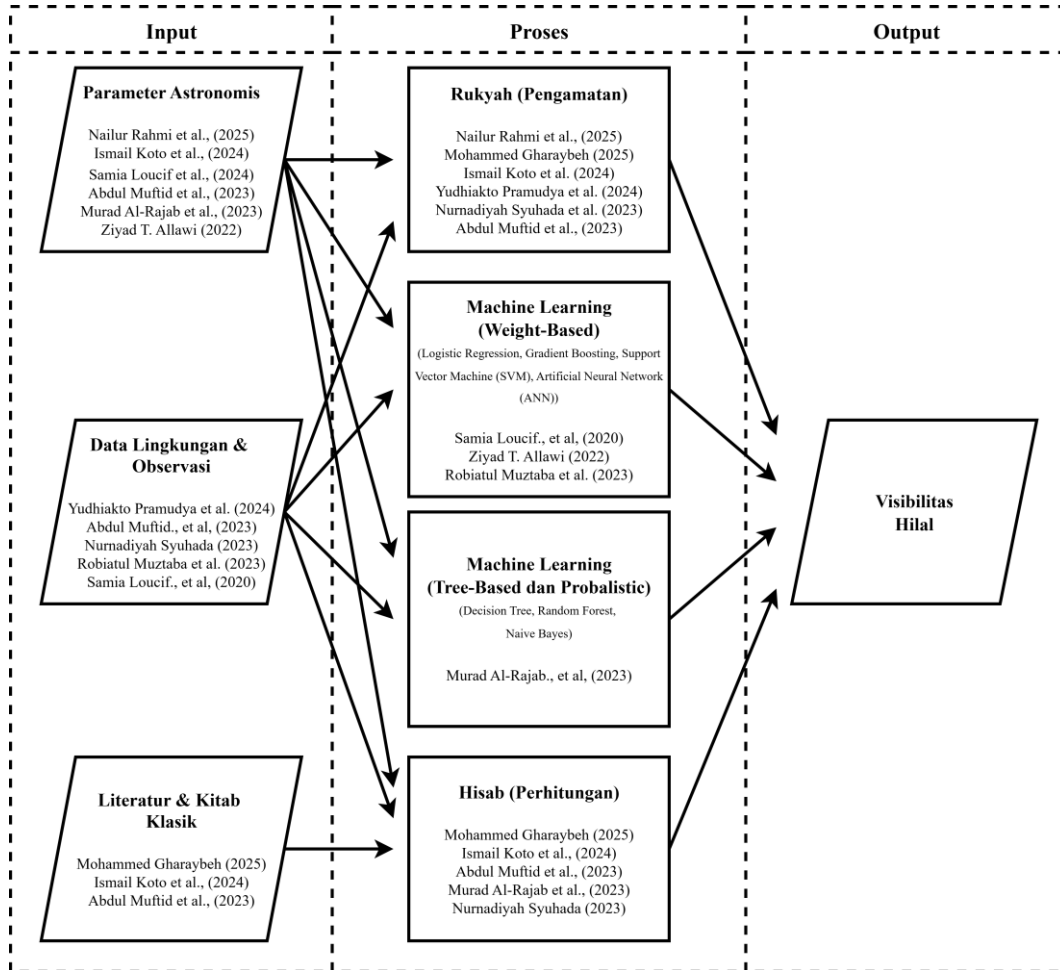
No	Penulis	Topik	Metode	Hasil
7	Robiatul Muztaba et al. (2023)	<i>Deep learning</i> Mask R-CNN untuk deteksi dan recognisi <i>hilal</i> dari dataset teleskop robotik, fokus kondisi ekstrem seperti tertutup awan.	Unsupervised DL: Mask R-CNN	mAP 0.99 pada IOU 0.5 untuk kondisi ekstrem/normal; 99% akurasi fase bulan; framework untuk monitoring otomatis.
8	Abdul Mufid (2023)	Implementasi kriteria baru MABIMS (3° tinggi <i>hilal</i> , 6.4° elongasi) untuk unifikasi kalender Hijri; strategi multidisiplin (fiqh, astronomi, sosiopolitik).	Kualitatif: Review literatur hadith/fiqh, Analisis data astronomi: <i>Rukyat</i>	Unifikasi dapat dipercepat melalui kesepakatan kriteria, dateline, otoritas; implementasi Indonesia 2022 sukses.
9	Murad Al-Rajab et al. (2023)	Prediksi Visabilitas <i>Hilal</i>	Binary klasifikasi: LR, SVM, ANN, RF	Decision Tree dan Random Forest mencapai akurasi 93%, RF mencatat F1-score tertinggi 0,88, SVM dan Neural Networks akurasi 92%. Model ML efektif meningkatkan akurasi penentuan <i>hilal</i> dan berkontribusi bagi harmonisasi awal Ramadan secara global.
10	Ziyad T. Allawi (2022)	Prediksi Visabilitas <i>Hilal</i>	Binary klasifikasi: ANN	ANN berhasil digunakan untuk membangun kalender Hijri empat tahun (1440–1443 H) di Baghdad.
11	Penelitian ini	Klasifikasi-Prediksi Visabilitas <i>Hilal</i>	Kategorikal klasifikasi: Multinomial Logistic Regression	Klasifikasi <i>hilal</i> berdasarkan elongasi pada kategori “gelap”, “samar”, dan “terang”. Rekomendasi unifikasi kalender nasional.

Berdasarkan tinjauan terhadap penelitian terkait, terdapat sejumlah celah yang menjadi landasan kontribusi penelitian ini. Studi berbasis *machine learning*, seperti pada penelitian Allawi (2022), Al-Rajab et al. (2023), dan Loucif et al. (2024), menawarkan pendekatan prediktif canggih, namun terbatas pada model *black-box* dengan klasifikasi biner, yang tidak mampu menangani kompleksitas multinomial (Tidak Terlihat, Tipis, Nyata) yang relevan untuk konteks lokal Malang Raya. Penelitian observasi lapangan, seperti Syuhada et al. (2023) dan Pramudya et al. (2024), menyediakan data empiris yang kuat, tetapi terhambat oleh ketiadaan model prediktif untuk mengintegrasikan variabel geofisika

spesifik, seperti kabut dan elevasi di wilayah tropis. Sementara itu, penelitian berbasis policy dan fiqh, termasuk Mufid & Djamaluddin (2023), Gharaybeh (2025), Koto et al. (2024), dan Rahmi et al. (2025), berfokus pada harmonisasi dan analisis dispute tanpa menawarkan solusi komputasi yang mendukung implementasi praktis. Penelitian ini mengisi celah tersebut melalui pengembangan *Multinomial Logistic Regression* (MLR) yang bersifat *interpretable*, dioptimalkan untuk klasifikasi visabilitas *hilal* berdasarkan variabel Beda Azimuth, Tinggi Bulan, dan Elongasi, yang disesuaikan dengan kondisi tropis Malang Raya. Pendekatan ini tidak hanya meningkatkan akurasi prediksi, tetapi juga menjadi jembatan antara sains dan fiqh, mendukung upaya unifikasi kalender *Hijriah* dengan kontribusi lokal yang inovatif.

2.2 Kerangka Teori

Kerangka Teori pada Gambar 2.1 yang disajikan di bawah ini menjelaskan kerangka teori yang mengacu pada penelitian-penelitian sebelumnya yang menjadi acuan penelitian visabilitas *hilal*, dengan melakukan prediksi klasifikasi sebagai data input dan daftar metode sebagai acuan penelitian dan juga output dari hasil kerangka teoritis.



Gambar 2.1 Kerangka Teori

Ilustrasi pada Gambar 2.1 menguraikan langkah-langkah penelitian dalam menyusun studi pustaka dalam kerangka teori.

1. Data visabilitas *hilal* yang digunakan sebagai sumber data sebagai inputan.
2. Kategori metode merupakan kategori yang didasari dari metode-metode (konvensional dan komputasional) yang digunakan peneliti sebelumnya.
3. Jurnal referensi yang digunakan pada penelitian ini tercantum pada studi pustaka 2.1.

Informasi kinerja metode yang digunakan pada penelitian ini diukur akurasi dengan persen. Nilai akurasi mendekati 100 artinya hasil penelitian

memiliki skor akurasi tertinggi. Evaluasi kinerja diukur dengan menggunakan *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebagai tolak ukur performa model dalam mengklasifikasikan hasil output prediksi. Sementara itu MAE, dan RMSE digunakan untuk mengukur kesalahan yang lebih besar dan sebagai tolak ukur nilai kesalahan rata-rata pada hasil output prediksi visabilitas *hilal*.

2.3 Integrasi Al-Qur'an, Tafsir dan Hadis

Integrasi antara aspek keilmuan dengan metode komputasional untuk prediksi visabilitas *hilal*, dalam ajaran Islam sangat penting untuk menunjukkan relevansi penelitian dengan prinsip-prinsip agama. Penentuan awal bulan *Hijriah*, didasarkan pada *rukyat* (penglihatan) *hilal*, yang merupakan perintah syariat. Penelitian ini bertujuan mendukung proses tersebut dengan model prediktif berbasis data astronomis, sehingga lebih akurat dan efisien, tanpa menggantikan *rukyat* sebagai metode utama.

Salah satu landasan utama dalam metode penetapan awal Ramadan dan Syawal melalui *rukyatul hilal* adalah pada Al-Qur'an surat Al-Baqarah ayat 189:

يَسْأَلُونَكَ عَنِ الْأَهِلَّةِ قُلْ هِيَ مَوَاقِيتُ لِلنَّاسِ وَالْحَجِّ وَلَيْسَ الْبِرُّ بِأَنْ تَأْتُوا الْبُيُوتَ مِنْ ظُهُورِهَا وَلَكِنَّ الْبِرَّ مَنِ اتَّقَى وَأْتُوا الْبُيُوتَ مِنْ أَبْوَابِهَا
وَاتَّقُوا اللَّهَ لَعَلَّكُمْ تُفْلِحُونَ

Artinya: "Mereka bertanya kepadamu tentang *hilal*. Katakanlah: 'Hilal itu adalah tanda-tanda waktu bagi manusia dan (bagi ibadat) haji'. Dan bukanlah kebaikan memasuki rumah-rumah dari belakangnya, akan tetapi kebaikan ialah (kebaikan) orang yang bertakwa dan masuklah ke rumah-rumah itu dari pintu-pintunya; dan bertakwalah kepada Allah supaya kamu beruntung". (Q.S. Al-Baqarah: 189)

Menurut kitab tafsir *Al-Mukhtashar* dengan pengawasan Syaikh Dr. Shalih bin Abdullah bin Humaid menjelaskan, mereka bertanya kepadamu (wahai Rasul) tentang penciptaan bulan sabit dan perubahan keadaannya. Katakanlah untuk menjawab pertanyaan mereka tentang hikmah di balik penciptaan dan perubahan

keadaan *hilal*, “Sesungguhnya *hilal* itu adalah penunjuk waktu bagi manusia untuk mengetahui waktu-waktu ibadah mereka, seperti bulan-bulan haji, bulan puasa dan sempurnanya masa setahun dalam masalah zakat”. Mereka juga bisa mengetahui waktu-waktu kegiatan muamalat.

Selanjutnya hadis yang diriwayatkan oleh Imam Bukhari dari Abdullah ibn Umar R.A. dalam Shaheh Bukhari Juz III hal 478 – 479, di mana Nabi Muhammad SAW bersabda mengenai tata cara menentukan permulaan puasa dan hari raya:

الشَّهْرُ سَعٌ وَعِشْرُونَ لَيْلَةً، فَلَا تَصُومُوا حَتَّى تَرَوْهُ، فَإِنْ عُمَّ عَلَيْكُمْ فَأَكْمِلُوا الْعِدَّةَ ثَلَاثِينَ

Artinya: “Bulan (Hijriah) itu berjumlah dua puluh sembilan malam. Maka janganlah kalian berpuasa hingga kalian melihat hilal. Jika langit tertutup (mendung) atas kalian, maka sempurnakanlah hitungan bulan menjadi tiga puluh hari”. (H.R. Imam Bukhari: 478)

صُومُوا لِرُؤْيَيْهِ، وَأَفْطَرُوا لِرُؤْيَيْهِ، فَإِنْ غُيِّبَ عَلَيْكُمْ فَأَكْمِلُوا عِدَّةَ شَعْبَانَ ثَلَاثِينَ

Artinya: "Berpuasalah karena melihat (hilal) dan berbukalah karena melihatnya pula. Jika hilal tertutup (mendung) atas kalian, maka sempurnakanlah hitungan bulan Sya'ban menjadi tiga puluh hari" (H.R. Imam Bukhari: 479)

Dan Al-Quran juga menegaskan bahwa alam semesta, khususnya pergerakan benda-benda langit, diciptakan dengan keteraturan, presisi, dan perhitungan (*hisab*) yang matematis. Allah SWT berfirman dalam Surah Ar-Rahman ayat 5:

الشَّمْسُ وَالْقَمَرُ بِحُسْبَانٍ

Artinya “Matahari dan bulan (beredar) sesuai dengan perhitungan”. (Q.S. Ar-Rahman: 5)

Menurut tafsir ulama Ibnu Katsir dan Al-Jalalain, ayat tersebut menegaskan bahwa *hilal* (bulan sabit) merupakan penanda waktu (*mawaqit*) untuk manusia, khususnya dalam konteks ibadah seperti haji dan penentuan bulan-bulan *Hijriah* lainnya, termasuk Ramadan. *Hilal* bukan sekadar fenomena astronomis,

melainkan simbol untuk mengatur kehidupan beragama, seperti awal dan akhir puasa. Tafsir ini menekankan pentingnya *rukyat hilal* sebagai metode syariat untuk menentukan awal bulan *Hijriah*, yang harus dilakukan secara visual untuk memastikan keakuratan. Dalam konteks penelitian ini, ayat tersebut mendukung penggunaan analisis ilmiah (seperti *Multinomial Logistic Regression*) sebagai alat bantu untuk memprediksi visabilitas *hilal*, terutama di daerah seperti Malang Raya yang mungkin terpengaruh kondisi atmosfer, tanpa menggantikan *rukyat* sebagai kewajiban utama. Hal ini sejalan dengan prinsip Islam yang mendorong integrasi ilmu pengetahuan dengan agama untuk kemaslahatan umat, seperti dalam penentuan kalender *Hijriah* yang akurat.

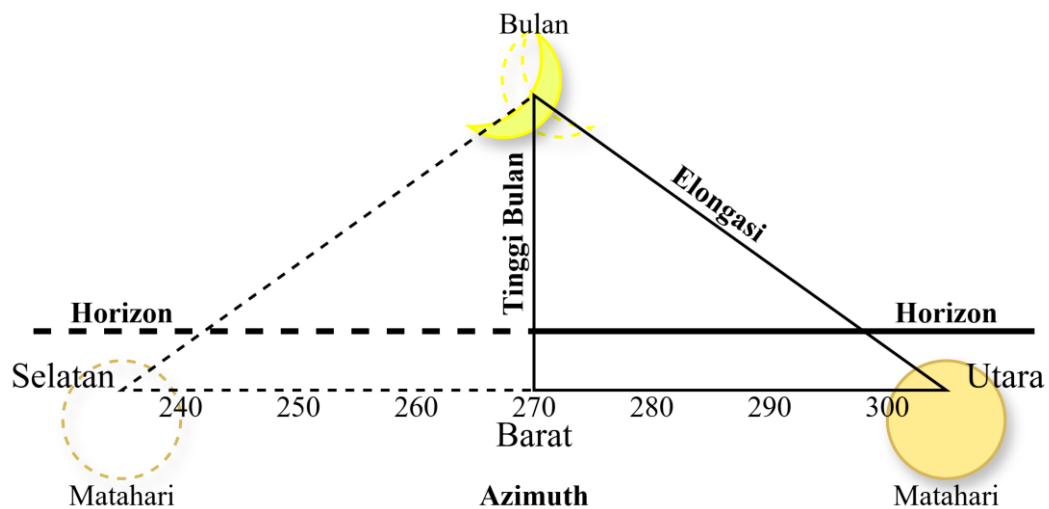
2.4 Sistem Kalender *Hijriah*

Sistem Kalender *Hijriah* merupakan sistem penanggalan lunar yang digunakan umat Islam secara global untuk menentukan waktu ibadah dan perayaan agama, seperti Ramadan, Idul Fitri, dan Idul Adha. Kalender ini didasarkan pada siklus revolusi Bulan mengelilingi Bumi, yang berlangsung rata-rata 29,53 hari, sehingga setiap bulannya terdiri dari 29 atau 30 hari.

Namun, perbedaan interpretasi antara *hisab* dan *rukyat* sering menimbulkan variasi tanggal, terutama di wilayah tropis seperti Malang Raya, yang dipengaruhi oleh faktor geofisika seperti kabut dan elevasi. Landasan ini menjadi dasar penting untuk memahami kebutuhan pendekatan prediktif yang mendukung unifikasi kalender.

2.5 Astronomi Posisi (Gerak Bulan dan Matahari)

Astronomi posisi mempelajari pergerakan relatif Bulan dan Matahari terhadap Bumi, yang menjadi faktor determinan utama visabilitas *hilal*. Bulan mengorbit Bumi dalam periode sinodis sekitar 29,53 hari, dengan fase konjungsi terjadi saat Bulan berada di antara Bumi dan Matahari, diikuti munculnya *hilal* saat Bulan bergerak menjauh. Ilustrasi terdapat pada gambar 2.2.



Gambar 2.2 Ilustrasi Posisi Bulan dan Matahari

Parameter kunci meliputi Beda Azimuth (selisih sudut azimuth antara Matahari dan Bulan), Tinggi Bulan (*altitude* di atas *horizon*), dan Elongasi (sudut terbuka antara Matahari, Bumi, dan Bulan). Gerak Bulan dipengaruhi oleh inklinasi orbit sebesar $5,14^\circ$ terhadap ekliptika, sedangkan posisi Matahari menentukan waktu terbenam dan sudut pencahayaan. Upaya untuk menganalisis dan memodelkan visabilitas *hilal* (*rukyat*) menggunakan metode perhitungan (*hisab*) dan analisis komputasi (MLR) merupakan manifestasi dari perintah Allah SWT untuk menggunakan akal (*'aql*) dalam memahami ayat-ayat-Nya, baik yang tersurat (*Qauliyah*) maupun yang tersirat di alam semesta (*Kauniyah*).

Di Malang Raya, faktor atmosferik seperti kelembapan dan debu dapat memodifikasi visabilitas, menjadikan pemahaman gerak astronomi ini landasan penting untuk model prediktif.

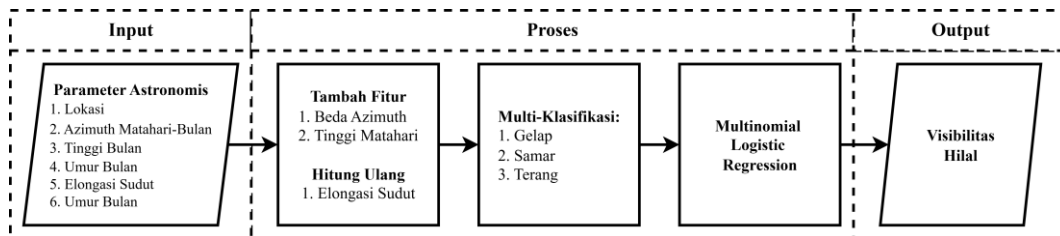
2.6 Kriteria Visibilias *Hilal*

Kriteria visabilitas *hilal* merujuk pada kondisi astronomi dan geofisika yang memungkinkan pengamatan bulan sabit baru. Menurut standar MABIMS, *hilal* dianggap terlihat jika memiliki Tinggi Bulan minimal 2° di atas *horizon*, Elongasi minimal 3° , dan umur bulan yang memadai, dengan penyesuaian untuk faktor lokal seperti cuaca. Penelitian astronomi, seperti Yallop (1997), menambahkan kriteria *Arc of Light* dan *Arc of Vision* berdasarkan sudut dan iluminasi. Di Malang Raya, kriteria ini perlu disesuaikan dengan kondisi tropis, termasuk kabut dan elevasi, yang dapat menghalangi pengamatan. Ketiga variabel utama Beda Azimuth, Tinggi Bulan, dan Elongasi berperan krusial dalam menentukan kategori visabilitas (“Gelap”, “Samar”, “Terang”), menjadi pijakan untuk klasifikasi yang lebih spesifik.

BAB III DESAIN DAN INTEGRASI SISTEM PENELITIAN

3.1 Kerangka Konsep

Kerangka konsep penelitian ini dirancang untuk menguraikan hubungan antara data astronomi dengan metode komputasi yang digunakan yaitu *Multinomial Logistic Regression*, serta hasil yang diharapkan adalah analisis representasi matematis dari aturan visabilitas *hilal*. Diagram kerangka konsep pada Gambar 3.1 menggambarkan interaksi antar variabel dan proses pengambilan keputusan berbasis probabilitas.



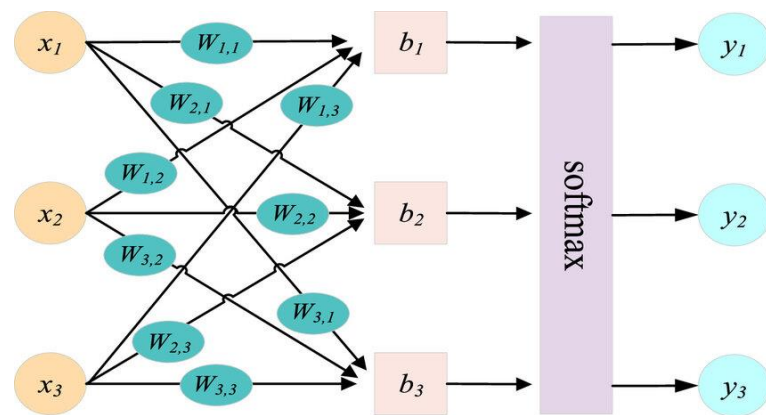
Gambar 3.1 Kerangka Konsep

Dalam kerangka konsep ini, variabel tinggi bulan, umur *hilal*, dan elongasi sudut saling berinteraksi untuk memengaruhi visabilitas *hilal*. Umur *hilal* diantara 0 sampai 8 jam cenderung meningkatkan kemungkinan *hilal* terlihat, sementara elongasi sudut yang lebih besar juga berkontribusi pada visabilitas yang lebih baik. Tetapi pada penelitian ini berfokus pada penggunaan variabel yang berbeda, yang hanya menggunakan beda azimuth, tinggi bulan, dan elongasi sudut. Variabel dependen dalam penelitian ini adalah kategori visabilitas *hilal* (“Gelap”, “Samar”, “Terang”) yang diperoleh dari perhitungan berbasis aturan (*Rule-Based System*) atau algoritma deterministik. Model MLR akan digunakan untuk menganalisis, mewakili (merepresentasikan), dan meng-aproksimasi (mendekati)

kategori ini berdasarkan serangkaian variabel prediktor (astronomis). Model dievaluasi menggunakan koefisien *matrix* dan grafik ROC.

3.2 Multinomial Logistic Regression (MLR)

Klasifikasi dalam penelitian ini mengacu pada pengelompokan data visabilitas *hilal* ke dalam kategori *diskrit* berdasarkan variabel prediktor. *Multinomial Logistic Regression* (MLR) adalah metode statistik yang memodelkan hubungan antara variabel dependen kategorik dengan variabel independen numerik, cocok untuk mengklasifikasikan visabilitas *hilal* ke dalam tiga kategori: Gelap, Samar, dan Terang.



Gambar 3.2 Arsitektur Multinomial Logistik Regression (MLR)

Model ini mengasumsikan hubungan log-linear dengan estimasi parameter menggunakan fungsi maksimum *likelihood*, menawarkan interpretabilitas melalui analisis koefisien, berbeda dengan model *black-box*. Di Malang Raya, MLR dapat dioptimalkan dengan data lokal untuk meningkatkan akurasi prediksi, menjadikannya landasan teoretis yang tepat untuk mengintegrasikan sains dan fiqh.

Arsitektur pembangunan model prediksi berbasis klasifikasi menggunakan *Multinomial Logistic Regression*. Secara matematis, untuk K kategori dependen

(dalam penelitian ini $K = 3$: “Gelap”, “Samar”, dan “Terang”), model MLR diekspresikan dalam dua bentuk utama:

1. Bentuk Logit (*Baseline-Category Logit*)

Bentuk ini digunakan untuk interpretasi koefisien dengan menetapkan satu kategori sebagai referensi atau *baseline* ($K = \text{"Gelap"}$). Model kemudian menghitung $K - 1$ persamaan *log-odds* (logit) untuk kategori lainnya pada 3.1:

$$\log \left(\frac{P(Y = k)}{P(Y = K)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 \beta_2 + \dots + \beta_n \beta_n \quad (3.1)$$

Keterangan:

k : Kategori (1, 2, ..., $K - 1$).

K : Baseline kategori.

$P(Y = k)$: Probabilitas bahwa observasi berada pada kategori ke- k .

$P(Y = K)$: Probabilitas kategori referensi (baseline).

β_0 : Intersep (konstanta) untuk kategori k .

β_n : Intersep untuk variabel input.

2. Bentuk Probabilitas (Fungsi *Softmax*)

Bentuk ini, yang merupakan transformasi dari bentuk logit, digunakan untuk menghitung probabilitas prediksi akhir untuk setiap kategori k (termasuk kategori *baseline*). Fungsi ini dikenal sebagai fungsi *softmax* pada persamaan 3.2:

$$P(Y = k) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 \beta_2 + \dots + \beta_n \beta_n}}{1 + \sum_{l=1}^K e^{\beta_0 + \beta_1 X_1 + \beta_2 \beta_2 + \dots + \beta_n \beta_n}} \quad (3.2)$$

Keterangan:

$P(Y = K)$: Peluang bahwa suatu observasi masuk ke kategori k ,
dibandingkan kategori-kategori lain dalam variabel
respons yang bersifat multinomial.

$e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n}$: (numerator) Eksponensial dari linear predictor (η) untuk kategori k : $n_k = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n$. Eksponensial ini memastikan probabilitas selalu bernilai positif.

$1 + \sum_{l=1}^K e^{\beta_0 + \beta_1 X_1 + \beta_n X_n}$: (denominator) Penjumlahan seluruh eksponensial untuk semua kategori selain baseline (total $K - 1$ kategori) ditambah 1, yaitu kontribusi untuk kategori baseline karena ($\beta_{0K} = 0, \beta_{1K} = 0, \beta_{nK} = 0$). Sehingga kategori baseline memiliki ($e^0 = 1$).

3.3 Persiapan Data

Persiapan data merupakan tahapan yang penting dalam sebuah penelitian, karena ketersediaan data yang valid dan andal akan sangat membantu proses pengolahan dan analisis selanjutnya (Al-Rajab et al., 2023). Data yang digunakan merupakan data sekunder yang diperoleh dari laman resmi BMKG (<https://hilal.bmkg.go.id/>) pada tanggal 24 Desember 2024. Data tersebut mencakup observasi *hilal* dari berbagai titik di Indonesia, dengan variabel-variabel utama meliputi:

Tabel 3.1 Keterangan Variabel Dataset

No	Nama Variabel	Satuan	Keterangan Ilmiah
1	tanggal	-	Tanggal pengambilan data, yaitu hari dalam kalender ketika pengamatan dilakukan.
2	tahun	-	Tahun pengambilan data, menunjukkan periode waktu spesifik ketika pengamatan dilakukan.
3	lokasi	-	Lokasi pengamatan, yaitu tempat spesifik di mana pengamatan dilakukan.
4	bujur	derajat	Bujur geografis dari lokasi pengamatan, menunjukkan posisi timur atau barat dari garis Greenwich.
5	lintang	derajat	Lintang geografis dari lokasi pengamatan, menunjukkan posisi utara atau selatan dari garis khatulistiwa.
6	matahari terbenam	jam	Waktu matahari terbenam di lokasi pengamatan, dalam format waktu lokal.

No	Nama Variabel	Satuan	Keterangan Ilmiah
7	bulan terbenam	jam	Waktu bulan terbenam di lokasi pengamatan, dalam format waktu lokal.
8	azimuth matahari	derajat	Arah matahari terbenam, diukur dalam derajat dari arah utara searah jarum jam.
9	azimuth bulan	derajat	Arah bulan terbenam, diukur dalam derajat dari arah utara searah jarum jam.
10	tinggi bulan	derajat	Tinggi bulan di atas <i>horizon</i> pada saat pengamatan, diukur dalam derajat.
11	umur <i>hilal</i>	jam	Umur <i>hilal</i> , yaitu waktu yang telah berlalu sejak bulan baru terakhir, diukur dalam jam.
12	elongasi	derajat	Elongasi bulan, yaitu sudut antara garis penghubung Matahari-Bumi dan garis penghubung Bumi-Bulan, diukur dalam derajat.

Tabel 3.2 Dataset Visibilitas *Hilal*

tanggal	22/03/2023	22/03/2023	...	11/03/2025	11/03/2025
tahun	1444	1444	...	1446	1446
bulan	Ramadan	Ramadan	...	Dzulhijjah	Dzulhijjah
lokasi	Banda Aceh	Medan	...	Manokwari	Jayapura
bujur	95,19	98,41	...	134,40	140,41
lintang	5,33	3,34	...	-0,51	-2,32
matahari terbenam	18:49:00	18:36:00	...	18:03:00	18:10:00
bulan terbenam	19:29:00	19:14:00	...	17:34:00	17:39:00
azimuth matahari	270	270	...	291	291
azimuth bulan	271	272	...	296	296
tinggi bulan	8	8	...	1	0
umur <i>hilal</i>	19,10	18,85	...	8,13	7,62
elongasi	9	9	...	5	5

Variabel-variabel ini dipilih karena memiliki pengaruh signifikan terhadap visabilitas *hilal*, data primer yang digunakan diperoleh dari pengamatan *hilal* yang dilakukan oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) untuk penentuan awal bulan *Hijriah*. Teknik pengumpulan data dilakukan dengan memastikan bahwa data yang diperoleh lengkap dan akurat.

Proses pengumpulan data melibatkan beberapa tahapan penting, yaitu proses validasi data yang dilakukan dengan data observasi dari BMKG, pembersihan data dengan menghapus nilai-nilai yang hilang (*missing values*), serta data yang tidak konsisten atau memiliki nilai yang tidak wajar (*outlier*) diidentifikasi dan dihilangkan untuk memastikan kualitas data. Setelah itu, normalisasi data

dilakukan menggunakan metode *StandardScaler* (*Z-Score Normalization*) untuk mengubah nilai-nilai dari fitur-fitur dalam data agar memiliki rata-rata 0 dan standar deviasi 1 (Das & Geisler, 2021), tahap ini krusial untuk menjaga kualitas input dalam proses klasifikasi dan menghindari bias model akibat perbedaan skala antar variabel (Purba & Handhayani, 2024).

Metode *StandardScaler* merupakan metode normalisasi dengan melakukan transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai perbandingan antara data saat sebelum dan sesudah dinormalisasi (Dua et al., 2017) Pada persamaan 3.3:

$$z = \frac{x - \mu}{\sigma} \quad (3.3)$$

Dengan perhitungan rata-rata (Pedregosa et al., 2011) Pada persamaan 3.4:

$$\mu = \frac{1}{N} \sum_{i=1}^N (x_i) \quad (3.4)$$

Dan perhitungan deviasi pada persamaan 3.5:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (3.5)$$

Keterangan:

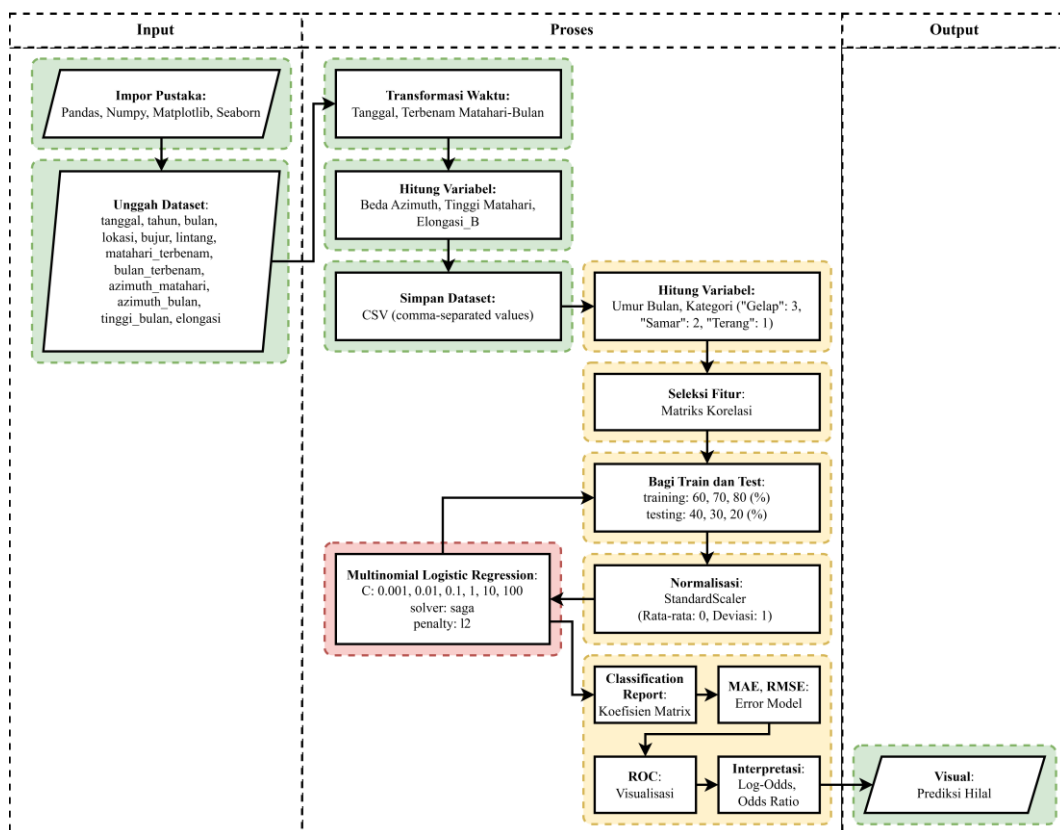
- z : nilai standar (*Z-score*) dari x .
- x : nilai awal dari dataset.
- μ : nilai rata-rata dari dataset.
- σ : nilai deviasi standar dari dataset.

3.4 Desain Sistem

Desain sistem dalam penelitian ini bertujuan membangun model prediksi visabilitas *hilal* yang akurat dan dapat diinterpretasikan berdasarkan data

observasi astronomis dari BMKG. Sistem ini dirancang secara sistematis untuk mengolah data, membangun arsitektur model klasifikasi, mengevaluasi performanya, hingga menyajikan hasil yang mudah dipahami. Pendekatan utama yang digunakan adalah algoritma *Multinomial Logistic Regression*, yang dipilih karena kemampuannya menangani variabel dependen dengan lebih dari dua kategori nominal.

Diagram alur pada Gambar 3.3 menggambarkan tahapan desain sistem secara terperinci, yang dapat diuraikan sebagai berikut:



Gambar 3.3 Desain Sistem

Desain sistem dalam penelitian ini dilakukan melalui lima langkah sistematis untuk memastikan keberhasilan dalam memprediksi kategori visabilitas *hilal*:

3.4.1 Persiapan Dataset

1. Variabel Data

Data input mencakup serangkaian variabel komprehensif, seperti tanggal, tahun, bulan, lokasi (bujur dan lintang), waktu matahari dan bulan terbenam, azimuth matahari dan bulan, tinggi bulan, elongasi (Munir, 2016).

2. Transformasi Waktu

Setelah dataset berhasil diinputkan maka selanjutnya mentransformasi waktu seperti waktu terbenam matahari dan waktu terbenam bulan ke dalam format waktu (*time*) (Haslebach et al., 2024).

3. Hitung Variabel (Beda Azimuth, Tinggi Matahari, Elongasi)

Selanjutnya membuat perhitungan untuk mendapatkan variabel baru bernama beda azimuth dan tinggi matahari untuk variabel elongasi yang baru yang menggunakan perhitungan rumus astronomi. Rumus astronomi ditampilkan pada persamaan 3.6 dan 3.7 berikut:

$$\cos(\psi) = (\sin(h_M) \sin(h_S)) + ((\cos(h_M) \cos(h_S)) \cdot \cos(Az_M - Az_S)) \quad (3.6)$$

$$h_S = a + \arccos\left(\frac{\cos(\psi)}{R}\right) \quad (3.7)$$

Keterangan:

ψ : Elongasi

h_M : Tinggi Bulan

h_S : Tinggi Matahari

Az_M : Azimuth Bulan

Az_S : Azimuth Matahari

R : Perhitungan Trigonometri $\sqrt{(\cos(h_M) \times \cos(Az_M - Az_S))^2 + \sin(h_M)^2}$

a : $\arctan\left(\frac{B}{A}\right)$

4. Simpan Dataset Baru

Dataset baru dengan variabel baru, selanjutnya disimpan dalam format CSV yang nantinya digunakan untuk proses menggunakan elongasi baru.

3.4.2 Persiapan Pemodelan Dataset Baru

Pada tahap ini, dilakukan pemilihan variabel yang paling relevan dan pendefinisian target klasifikasi.

1. Hitung variabel (Umur Bulan dan Kategori)

Setelah dataset telah siap sebelumnya setelah proses perhitungan sebelumnya, maka didapatkan perhitungan untuk mendapatkan umur bulan dan kategori. Langkah krusial dalam penelitian ini adalah pelabelan sintetis. Karena data observasi (*rukya*) tidak tersedia, variabel dependen (Y) akan dibuat (disintesis) menggunakan algoritma berbasis aturan (Model Pakar) yang telah didefinisikan oleh peneliti pada persamaan 3.8. Aturan ini berfungsi sebagai ground truth untuk penelitian ini:

$$If - Else \begin{cases} E < 5^\circ \text{ or } T_B < 0^\circ \text{ or } \Delta_{Azimuth} < 2^\circ = Gelap \\ 5^\circ \leq E < 12^\circ \text{ and } T_B > 0^\circ \text{ and } 2^\circ \leq \Delta_{Azimuth} < 10^\circ = Samar \\ E > 12^\circ = Terang \end{cases} \quad (3.8)$$

Keterangan:

E : Elongasi

T_B : Tinggi Bulan

$\Delta_{Azimuth}$: Beda Azimuth

Hasil dari tahap ini adalah kolom variabel dependen Y (“Gelap”, “Samar”, “Terang”) yang akan digunakan untuk melatih Model MLR."

2. Seleksi Variabel (Fitur)

Dataset akhirnya didapatkan, tetapi sebelum itu dilakukan proses seleksi fitur dengan tujuan untuk mengurangi dimensi data dan fitur yang tidak relevan, serta untuk meningkatkan efektifitas dan efisiensi kinerja dari algoritma klasifikasi.

3. Pembagian Dataset

Dataset akan dibagi menjadi data latih dan data uji dengan beberapa rasio yang akan dieksperimenkan, seperti 60:40, 70:30, dan 80:20, untuk menemukan komposisi yang paling optimal.

4. Normalisasi Data

Seluruh variabel numerik akan dinormalisasi menggunakan metode *StandardScaler*. Proses ini mentransformasikan nilai setiap fitur bernilai rata-rata 0 dan memiliki standar deviasi 1, sehingga memastikan tidak ada variabel yang mendominasi proses pelatihan model hanya karena perbedaan skala (Wolfe & Emami, 2024).

3.4.3 Tahap Pembangunan Arsitektur Model

Model MLR dilatih menggunakan variabel independen (X) dan variabel dependen *sintetis* (Y). Model ini bertujuan mewakili aturan (Model Pakar) ke dalam bentuk rumus matematis (logit). Dengan menetapkan "Gelap" (label 3) sebagai kategori *baseline*, model MLR akan mengestimasi 2 persamaan Bentuk Logit (untuk wawasan/interpretasi):

1. Logit 1: Membandingkan "Gelap" dengan "Gelap" pada persamaan 3.9:

$$\log \left(\frac{P(Y = "Gelap")}{P(Y = "Gelap")} \right) = \beta_0^{(G)} X_1 + \beta_1^{(G)} X_2 + \beta_2^{(G)} X_3 + \beta_3^{(G)} X_4 + \beta_3^{(G)} X_5 \quad (3.10)$$

2. Logit 2: Membandingkan "Samar" dengan "Gelap" pada persamaan 3.10:

$$\log \left(\frac{P(Y = "Samar")}{P(Y = "Gelap")} \right) = \beta_0^{(S)} X_1 + \beta_1^{(S)} X_2 + \beta_2^{(S)} X_3 + \beta_3^{(S)} X_4 + \beta_3^{(S)} X_5 \quad (3.10)$$

3. Logit 3: Membandingkan "Terang" dengan "Gelap" pada persamaan 3.11:

$$\log \left(\frac{P(Y = "Terang")}{P(Y = "Gelap")} \right) = \beta_0^{(T)} X_1 + \beta_1^{(T)} X_2 + \beta_2^{(T)} X_3 + \beta_3^{(T)} X_4 + \beta_3^{(T)} X_5 \quad (3.11)$$

Bentuk Probabilitas menggunakan hasil dari Bentuk Logit untuk menghitung probabilitas *absolut* (nilai antara 0% hingga 100%) untuk setiap kategori. Mendefinisikan “Skor” (S) untuk setiap kategori berdasarkan variabel persamaan 3.12, 3.13 dan 3.14:

$$S_{Gelap} = 1 \text{ (Skor untuk kategori baseline selalu 1)} \quad (3.12)$$

$$S_{Samar} = \beta_0^{(S)} \text{Beda Azimuth} + \beta_1^{(S)} \text{Tinggi Matahari} + \beta_2^{(S)} \text{Tinggi Bulan} \\ + \beta_3^{(S)} \text{Elongasi} + \beta_3^{(S)} \text{Umur Bulan} \quad (3.13)$$

$$S_{Terang} = \beta_0^{(T)} \text{Beda Azimuth} + \beta_1^{(T)} \text{Tinggi Matahari} + \beta_2^{(T)} \text{Tinggi Bulan} \\ + \beta_3^{(T)} \text{Elongasi} + \beta_3^{(T)} \text{Umur Bulan} \quad (3.14)$$

Kemudian, probabilitas untuk setiap kategori dihitung sebagai berikut:

a. Probabilitas “Gelap” pada persamaan 3.15:

$$P(Y = "Gelap") = \frac{\exp(S_{Gelap})}{\exp(S_{Gelap}) + \exp(S_{Samar}) + \exp(S_{Terang})} \\ = \frac{1}{1 + \exp(S_{Samar}) + \exp(S_{Terang})} \quad (3.15)$$

b. Probabilitas “Samar” pada persamaan 3.16:

$$P(Y = "Samar") = \frac{\exp(S_{Samar})}{\exp(S_{Gelap}) + \exp(S_{Samar}) + \exp(S_{Terang})} \\ = \frac{\exp(S_{Samar})}{1 + \exp(S_{Samar}) + \exp(S_{Terang})} \quad (3.16)$$

c. Probabilitas “Terang” pada persamaan 3.17:

$$P(Y = "Terang") = \frac{\exp(S_{Terang})}{\exp(S_{Gelap}) + \exp(S_{Samar}) + \exp(S_{Terang})} \\ = \frac{\exp(S_{Terang})}{1 + \exp(S_{Samar}) + \exp(S_{Terang})} \quad (3.17)$$

3. Konfigurasi *Hyperparameter*

Model akan dikonfigurasi dan dioptimalkan dengan *hyperparameter*, C (*Inverse of Regularization Strength*) akan diuji beberapa nilai (0.001, 0.01, 0.1, 1, 10, 100) untuk mengontrol kompleksitas model dan mencegah *overfitting*. *Solver* menggunakan *solver* ‘saga’, yang efisien untuk dataset besar dan mendukung regularisasi L2. *Penalty* menerapkan regularisasi L2 untuk menjaga agar bobot model tetap kecil.

3.4.4 Tahap Evaluasi Performa Model

Evaluasi algoritma merupakan tahap penting dalam penelitian ini untuk mengukur kinerja model yang dibangun. Dalam penelitian ini, evaluasi dilakukan dengan menggunakan beberapa metrik evaluasi, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Metrik-metrik ini digunakan untuk menilai sejauh mana model yang dibangun dapat memprediksi visabilitas *hilal* dengan akurat (Shalamov et al., 2022).

1. Akurasi (*Accuracy*)

Akurasi dihitung dengan persamaan 3.18:

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Prediksi}} \quad (3.18)$$

2. *Precision*

Precision dihitung dengan persamaan 3.19:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positive}} \quad (3.19)$$

3. *Recall*

Recall dihitung dengan persamaan 3.20:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3.20)$$

4. *F1-Score*

F1-Score dihitung dengan persamaan 3.21:

$$F1 - Score = 2 \left(\frac{Precision \times Recall}{Precision + Recall} \right) \quad (3.21)$$

Evaluasi untuk menghitung *error*:

5. MAE

Mean Absolute Error (MAE) dihitung dengan persamaan 3.22:

$$MAE = \left(\frac{1}{n} \right) \times \sum |y_i - \hat{y}_i| \quad (3.22)$$

Keterangan:

n : Jumlah data observasi atau jumlah pasangan nilai sebenarnya dan nilai prediksi.

y_i : Nilai aktual.

\hat{y}_i : Nilai yang diprediksi.

6. RMSE

Root Mean Square Error (RMSE) dihitung dengan persamaan 3.23:

$$RMSE = \sqrt{\left(\frac{1}{n} \right) \sum (y_i - \hat{y}_i)^2} \quad (3.23)$$

Keterangan:

n : Jumlah data observasi atau jumlah pasangan nilai sebenarnya dan nilai prediksi.

y_i : Nilai aktual.

\hat{y}_i : Nilai yang diprediksi.

7. Kurva ROC

Kemampuan model untuk membedakan antar kelas akan divisualisasikan menggunakan kurva ROC (*Area Under the Curve*) (Hussain, M., & Al-Mostafa, 2021), (Loucif et al., 2024).

8. Metrik Kesalahan

Besaran kesalahan prediksi akan diukur menggunakan MAE (*Mean Absolute Error*) dan RMSE (*Root Mean Squared Error*) (Muztaba et al., 2023).

3.4.5 Tahap Interpretasi dan Visualisasi

Tahap akhir adalah menerjemahkan hasil model menjadi wawasan yang bermakna.

1. Visualisasi Prediksi

Hasil prediksi kategori *hilal* akan divisualisasikan untuk memberikan gambaran yang intuitif.

2. Interpretasi Koefisien

Kekuatan utama dari model *regresi logistik* adalah kemampuannya untuk diinterpretasikan. Akan dilakukan analisis terhadap koefisien *Log-Odds* dan *Odds Ratio* untuk memahami seberapa besar pengaruh setiap variabel *prediktor* (beda azimuth dan tinggi bulan) terhadap probabilitas visabilitas *hilal* jatuh ke dalam kategori "Samar" atau "Terang" dibandingkan dengan kategori "Gelap" (Dimara et al., 2017).

Dengan mengikuti langkah-langkah yang telah dijelaskan ini, diharapkan dapat memahami proses pembuatan model klasifikasi prediksi *Hilal* menggunakan *Logistic Regression*. Penelitian ini menekankan pada setiap tahap, mulai dari persiapan data hingga interpretasi hasil, untuk memastikan dapat memahami secara mendalam bagaimana model ini dikembangkan dan diaplikasikan.

3.5 Instrumen Penelitian

Variabel-variabel yang digunakan dalam penelitian ini terdiri dari:

3.5.1 Variabel Independen

Variabel independen adalah variabel prediktor yang diduga memengaruhi visabilitas *hilal*. Dalam penelitian ini, variabel independen meliputi:

1. Beda Azimuth ($^{\circ}$).
2. Tinggi Matahari ($^{\circ}$).
3. Tinggi bulan ($^{\circ}$).
4. Elongasi ($^{\circ}$).
5. Umur Bulan (Jam).

3.5.2 Variabel Dependen

Variabel dependen adalah variabel target yang akan diprediksi oleh model. Dalam penelitian ini, variabel dependen adalah Kategori Visabilitas *Hilal*, yang merupakan variabel nominal dengan 3 kategori:

1. Gelap
2. Samar
3. Terang

Dengan menggunakan instrumen penelitian ini, analisis hubungan antarvariabel dapat dilakukan secara mendalam, sehingga diharapkan menghasilkan model prediksi visabilitas *hilal* yang akurat dan relevan.

BAB IV UJI COBA

4.1 Skenario Uji Coba

Skenario uji coba merupakan langkah awal yang bertujuan untuk mengevaluasi performa model *Multinomial Logistic Regression* (MLR) dalam memprediksi visabilitas *hilal* berdasarkan tiga parameter astronomis utama, yaitu tinggi bulan, beda azimuth, dan elongasi sudut. Pada tahap ini, dataset hasil pelabelan sintetis dibagi menjadi 60% data latih dan 40% data uji menggunakan konfigurasi model: *solver saga*, *penalty* L2, dan *hyperparameter* $C = 1$ (Faid et al., 2023).

Performa awal model disajikan pada Tabel 4.1, yang menunjukkan bahwa MLR mampu mereplikasi pola keputusan Model Pakar dengan akurasi keseluruhan 89,91% (dibulatkan 90%). Nilai ini mengindikasikan bahwa model cukup mampu menangkap hubungan *non-linear* pada variabel astronomis yang memengaruhi keterlihatan *hilal*.

Tabel 4.1 Hasil Uji Coba

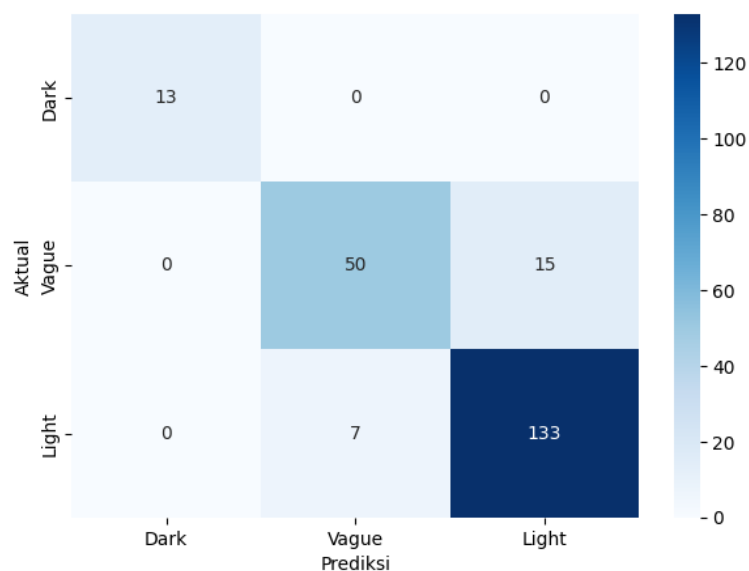
Solver	Penalty	C	F1-Score	Precision	Recall	RMSE	MAE
saga	l2	1	89.91%	0.93/0.90	0.91/0.90	1.818	1.7617

Evaluasi lebih lanjut terhadap performa model per kategori menunjukkan bahwa:

1. Kategori 1 (Terang) berhasil direplikasi secara sempurna. Dari 13 data uji yang termasuk kategori “Terang”, seluruhnya terklasifikasi benar oleh model, menghasilkan nilai *Precision* 1.00 dan *Recall* 1.00.

- Kategori 3 (Gelap) juga menunjukkan performa tinggi dengan *Recall* 0.95 dan *Precision* 0.90, menandakan bahwa model sangat efektif dalam mendeteksi kondisi *hilal* tidak terlihat.
- Kategori 2 (Samar) menjadi tantangan utama. Meskipun *Precision* mencapai 0.88, nilai *Recall* hanya 0.77, menunjukkan model gagal mengenali sekitar 23% kasus “Samar”.

Visualisasi *Koefisien Matriks* pada Gambar 4.1 memperlihatkan kontribusi masing-masing variabel terhadap probabilitas setiap kategori.



Gambar 4.1 *Koefisien Matriks Uji Coba*

Confusion Matrix (Gambar 4.1) memberikan pemahaman yang lebih rinci terkait pola kesalahan model:

- Kesalahan Utama: 15 data kategori “Samar” diklasifikasikan sebagai “Gelap”.
- Kesalahan Sekunder: 7 data kategori “Gelap” diklasifikasikan sebagai “Samar”.
- Tidak ditemukan kesalahan pada kategori “Terang” terhadap kategori lainnya.

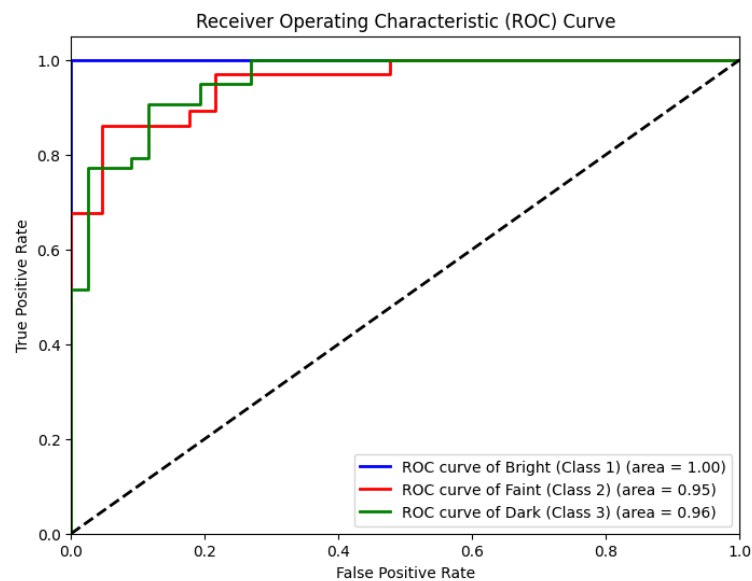
Kesalahan terbesar pada kategori “Samar” konsisten dengan karakter astronomis bahwa kondisi *hilal* tipis atau kritis memiliki parameter yang berada di wilayah *threshold* sehingga sulit dibedakan secara linier oleh model statistik.

Secara keseluruhan, Model MLR menunjukkan performa yang sangat baik dalam mereplikasi aturan Model Pakar, dengan mencapai akurasi global sebesar 89.91% (dibulatkan menjadi 90%). Ini menunjukkan bahwa model statistik MLR mampu “mempelajari” dan mewakili sebagian besar logika dari Model Pakar yang berbasis aturan.

Evaluasi berikutnya menggunakan kurva ROC (*Receiver Operating Characteristic*) untuk melihat kemampuan model memisahkan setiap kategori:

1. ROC kategori Terang = 1.00
2. ROC kategori Samar = 0.95
3. ROC kategori Gelap = 0.96

Nilai ROC yang mendekati 1 menunjukkan bahwa fitur-fitur astronomis yang digunakan sangat informatif dan model memiliki daya diskriminatif yang kuat.



Gambar 3.5 ROC Curve Uji Coba

Hasil analisis uji coba memberikan gambaran tentang keakuratan prediksi model terhadap data aktual. Misalnya, model dengan *recall* tinggi menunjukkan kemampuan yang baik dalam mendeteksi visabilitas *hilal*, sementara *precision* tinggi menunjukkan rendahnya tingkat kesalahan dalam prediksi *hilal* terlihat (Hoayek et al., 2023).

Selain itu, analisis pola dilakukan untuk mengidentifikasi hubungan antara variabel prediktor dan visabilitas *hilal*. Contohnya, tinggi sudut bulan yang lebih besar cenderung meningkatkan probabilitas *hilal* terlihat, sedangkan elongasi rendah dapat menurunkan visabilitas *hilal*.

4.2 Hasil Uji Coba

Bagian ini mendeskripsikan proses pemodelan secara lengkap mulai dari pra-proses data hingga pembangunan model final. Seluruh implementasi dilakukan pada *Google Colaboratory* menggunakan *Python* serta pustaka terkait seperti *Pandas*, *NumPy*, *Matplotlib*, *Seaborn*, dan *Scikit-Learn*.

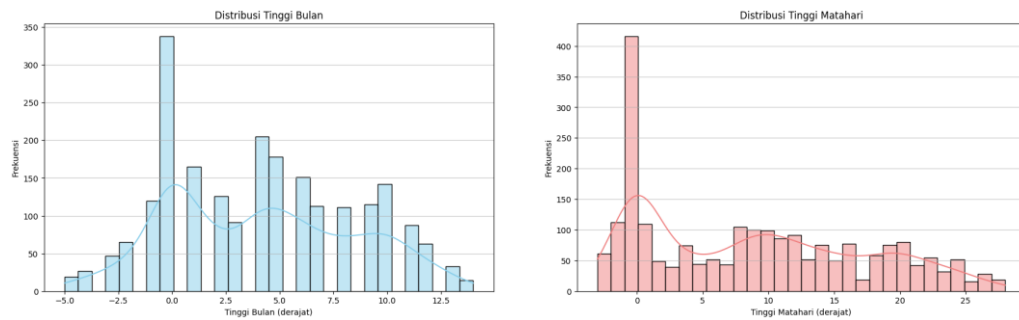
4.2.1 Analisis Awal dan Visualisasi Hubungan Antar Variabel (EDA)

Pada tahap awal uji hasil, dilakukan analisis eksploratif untuk memahami karakteristik dasar dataset serta hubungan antar variabel astronomis yang digunakan sebagai prediktor dalam model MLR. Tahap ini bertujuan untuk:

1. Mengidentifikasi variabel yang paling berpengaruh terhadap visabilitas *hilal*.
2. Menilai kecenderungan astronomis yang tampak pada data BMKG.
3. Mengonfirmasi kelayakan variabel sebelum proses pemodelan MLR dilakukan.

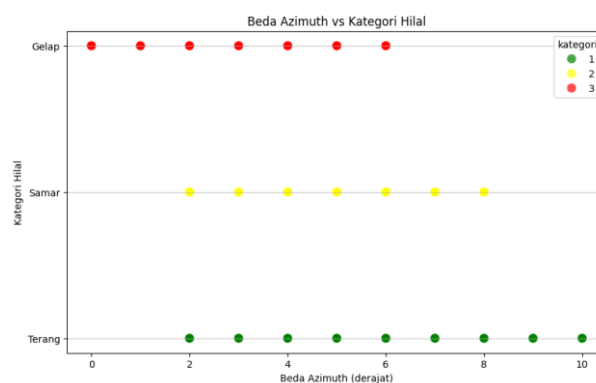
Distribusi Tinggi Bulan dan Tinggi Matahari, distribusi awal menunjukkan bahwa Tinggi Bulan pada saat *rukya*t sebagian besar berada di rentang 0° – 12° .

Nilai di bawah 0° mengindikasikan bulan sudah tenggelam sebelum atau saat matahari terbenam sehingga secara fisik mustahil terlihat (kategori Gelap).



Gambar 4.2 Distribusi Tinggi Bulan dan Tinggi Matahari

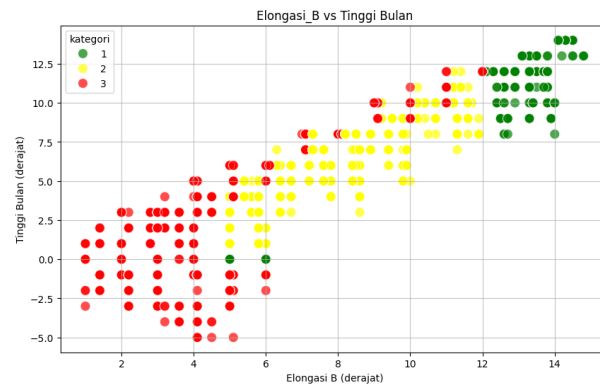
Hubungan Beda Azimuth dan Keterlihatan *Hilal* dapat dilihat pada analisis Scatterplot Gambar 4.3 antara beda azimuth dengan kategori visabilitas menunjukkan Beda azimuth $< 2^\circ$ sangat dominan berasosiasi dengan kategori Gelap sesuai definisi model pakar (*hilal* sangat dekat dengan matahari). Beda azimuth 2° – 10° banyak berada pada kategori Samar, menunjukkan kondisi *hilal* tipis. Beda azimuth $> 10^\circ$ cenderung berada pada kategori Terang. Hal ini sejalan dengan teori bahwa semakin besar selisih arah antara bulan dan matahari, semakin mudah *hilal* terpisah secara visual (*angular separation effect*).



Gambar 4.3 Hubungan Beda Azimuth dengan Kategori *Hilal*

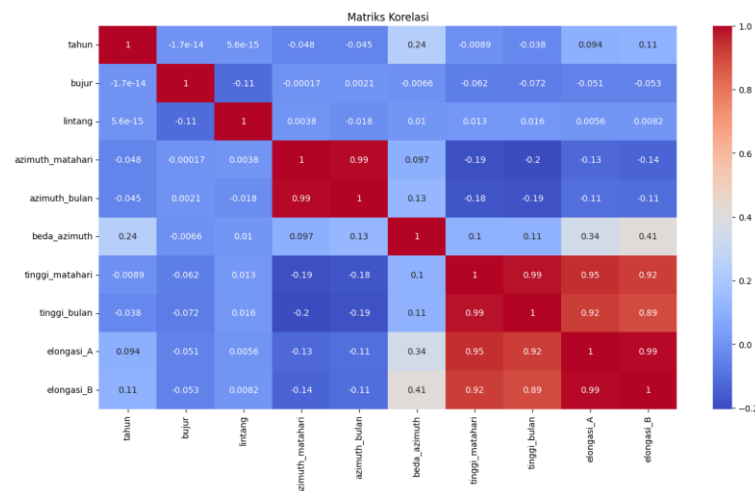
Hubungan Elongasi B dan Tinggi Bulan pada visualisasi *Scatterplot* Gambar 4.4 memperlihatkan pola bahwa $\text{Elongasi_B} < 5^\circ$ hampir selalu menghasilkan

kategori Gelap, sesuai kriteria MABIMS. Sedangkan Elongasi_B 5° – 12° menghasilkan variasi Samar atau Terang tergantung tinggi bulan, dan Tinggi Bulan $> 6^{\circ}$ secara signifikan meningkatkan peluang kategori Terang, karena posisi bulan lebih tinggi dari atmosfer padat.



Gambar 4.4 Korelasi Elongasi dengan Tinggi Bulan

Matriks korelasi pada Gambar 4.5 menunjukkan keterkaitan yang kuat antar variabel Tinggi Matahari dengan Beda Azimuth berkorelasi negatif moderat. Sedangkan Elongasi B dengan Umur Bulan berkorelasi kuat (logis karena umur bulan memengaruhi jarak geometris). Semua variabel astronomis memiliki korelasi signifikan terhadap Kategori, menguatkan relevansi variabel untuk pemodelan MLR.



Gambar 4.5 Korelasi Matriks Berdasarkan Dataset Awal

Berdasarkan hasil eksplorasi:

1. Variabel Tinggi Bulan, Elongasi_B, dan Beda Azimuth terbukti memiliki pengaruh paling kuat.
2. Pola kategori terlihat jelas dari visualisasi, menunjukkan bahwa dataset memiliki struktur yang baik untuk proses klasifikasi MLR.
3. Hasil EDA mendukung landasan teoritis pada Bab II dan kriteria pakar (MABIMS, kriteria elongasi minimum, tinggi *hilal*).

4.2.2 Transformasi Waktu

Transformasi waktu dilakukan untuk menyatukan format waktu matahari terbenam dan bulan terbenam sehingga dapat digunakan sebagai dasar perhitungan variabel-variabel astronomis lain.

	tanggal	tahun	bulan	lokasi	bujur	lintang	matahari_terbenam	bulan_terbenam	azimuth_matahari	azimuth_bulan	beda_azimuth
0	2021-08-08	1443	Muharram	Banda Aceh	95.1903	5.3322	2021-08-08 18:54:00	2021-08-08 18:54:00	286	290	4
1	2021-08-08	1443	Muharram	Medan	98.4124	3.3451	2021-08-08 18:03:00	2021-08-08 18:37:00	286	290	4
2	2021-08-08	1443	Muharram	Padang	100.2172	-0.5545	2021-08-08 18:27:00	2021-08-08 18:23:00	285	290	5
3	2021-08-08	1443	Muharram	Pekanbaru	101.2704	0.3159	2021-08-08 18:24:00	2021-08-08 18:21:00	286	290	4
4	2021-08-08	1443	Muharram	Bengkulu	102.1554	-3.4801	2021-08-08 18:16:00	2021-08-08 18:11:00	285	290	5

Gambar 4.6 Bentuk Transformasi Pada Dataset

Hasilnya adalah waktu pengamatan yang konsisten dengan standar astronomi dan memudahkan integrasi dengan fitur lainnya.

4.2.3 Rekayasa Fitur (*Feature Engineering*)

Tahap ini merupakan inti dari pra-proses data, di mana variabel-variabel baru yang lebih informatif diturunkan secara komputasional dari data yang ada. Proses ini esensial untuk menyediakan prediktor yang kuat bagi model. Nilai input diambil dari Tabel 3.2 Dataset Visabilitas *Hilal* untuk perhitungan secara

matematisnya sehingga nilai $Az_{matahari} = 270^\circ$, $Az_{bulan} = 271^\circ$, $h_M = 8^\circ$ (Tinggi Bulan), $\psi_A = 9^\circ$ (Elongasi A).

1. Penciptaan Variabel ‘beda_azimuth’

Sesuai dengan desain penelitian, variabel beda_azimuth (beda azimuth) dibuat. Variabel ini merepresentasikan selisih *absolut* antara azimuth matahari (azimuth_matahari) dan azimuth bulan (azimuth_bulan) pada persamaan 4.1.

$$\Delta Az = |270 - 271| = 1^\circ \quad (4.1)$$

Nilai ini krusial dalam menentukan posisi relatif kedua benda langit tersebut.

2. Penciptaan Variabel ‘tinggi_matahari’

Dataset observasi tidak menyertakan data ketinggian matahari (tinggi_matahari) saat pengamatan. Variabel ini, mutlak diperlukan untuk perhitungan elongasi yang akurat. Untuk mengatasinya, variabel tinggi_matahari dihitung menggunakan prinsip trigonometri bola (*spherical trigonometry*).

Konversi ke radian dengan persamaan 4.2:

$$\begin{aligned} h_M &= 8^\circ = 0.139626 \text{ rad}; \\ \Delta Az &= -1^\circ = -0.017453 \text{ rad}; \\ \psi_A &= 9^\circ = 0.157080 \text{ rad}. \end{aligned} \quad (4.2)$$

Hitung A, B, C pada persamaan 4.3:

$$\begin{aligned} A &= \cos(h_M) \cdot \cos(\Delta Az) = \cos(8^\circ) \cdot \cos(1^\circ) \approx 0.990268 \times 0.999848 \\ &\approx 0.990118. \\ B &= \sin(h_M) = \sin 8^\circ \approx 0.139173. \\ C &= \cos(\psi_A) = \cos 9^\circ \approx 0.987688. \end{aligned} \quad (4.3)$$

Hitung R dan α pada persamaan 4.4:

$$\begin{aligned} R &= \sqrt{A^2 + B^2} \approx \sqrt{0.990118^2 + 0.139173^2} \approx \sqrt{0.980333 + 0.019368} \\ &\approx 0.999851. \\ \alpha &= \text{atan}^2(B, A) \approx \text{atan}^2(0.139173, 0.990118) \approx 0.139626 \text{ rad}. \end{aligned} \quad (4.4)$$

Hitung argumen untuk \arccos pada persamaan 4.5:

$$\frac{C}{R} \approx \frac{0.987688}{0.999851} \approx 0.987839. \quad (4.5)$$

$$\arccos(0.987839) \approx 0.1546 \text{ rad} \approx 8.86^\circ.$$

Sehingga dengan persamaan 4.6:

$$h_s = a + \arccos\left(\frac{C}{R}\right) \approx 0.139626 + 0.1546 \approx 0.294226 \text{ rad} \approx 16.85^\circ. \quad (4.6)$$

Jadi tinggi matahari $h_s \approx 16.85^\circ$. Perhitungan numerik memberikan tinggi matahari yang relatif besar karena kombinasi elongasi dan tinggi bulan; ini adalah hasil perhitungan numerik yang sama dengan prosedur di metodologi.

3. Perhitungan Ulang Elongasi ('elongasi_B')

Dengan tersedianya data tinggi_matahari, nilai elongasi yang lebih presisi dapat dihitung. Menggunakan rumus empiris (hukum kosinus bola), sehingga variabel baru elongasi_B dapat dihitung dengan persamaan 4.7:

$$\cos(\psi_B) = (0.139173 \times 0.2899) + ((0.990268 \times 0.9571) \times 0.999848) \\ \cos(\psi_B) \approx 0.04036 + 0.94860 \approx 0.98896. \quad (4.7)$$

$$\psi_B = \arccos(0.98896) \approx 0.1473 \text{ rad} \approx 8.44^\circ.$$

Jadi elongasi_B $\approx 8.4^\circ$ (sedikit berbeda dari elongasi_A = 9° , hasil rekalkulasi lebih presisi karena memasukkan tinggi matahari). Hasil rekayasa fitur menunjukkan bahwa dataset memiliki struktur astronomis yang konsisten secara fisik.

4.2.4 Finalisasi Dataset dan Seleksi Variabel

Sebagai langkah penutup pra-proses, struktur DataFrame diatur ulang untuk meningkatkan keterbacaan dan alur logis. Kolom-kolom baru yang telah dibuat yaitu beda_azimuth, tinggi_matahari, dan elongasi_B ditempatkan secara strategis di dalam tabel.

Dataset yang telah bersih dan diperkaya dengan fitur-fitur baru ini kemudian disimpan sebagai berkas baru dengan nama data *hilal.csv*. Berkas inilah yang menjadi masukan (*input*) data final untuk tahap pemodelan, pelatihan (*training*), dan pengujian (*testing*) model *Multinomial Logistic Regression* (MLR) pada langkah selanjutnya, seperti yang ditampilkan pada tabel 4.2.

Tabel 4.2 *Hilal* Dataset Akhir

tanggal	2021-08-08 00:00:00	2021-08-08 00:00:00	2021-08-08 00:00:00	2021-08-08 00:00:00
tahun	1443	1443	1443	1443
bulan	Muharram	Muharram	Muharram	Muharram
lokasi	Banda Aceh	Medan	Padang	Pekanbaru
bujur	95,1903	98,4124	100,2172	101,2704
lintang	5,3322	3,3451	-0,5545	0,3159
matahari_terbenam	2021-08-08 18:54:00	2021-08-08 18:03:00	2021-08-08 18:27:00	2021-08-08 18:24:00
bulan_terbenam	2021-08-08 18:54:00	2021-08-08 18:37:00	2021-08-08 18:23:00	2021-08-08 18:21:00
azimuth_matahari	286	286	285	286
azimuth_bulan	290	290	290	290
beda_azimuth	4	4	5	4
tinggi_matahari	0	0	0	0
tinggi_bulan	0	0	0	0
elongasi_B	4	4	5	4

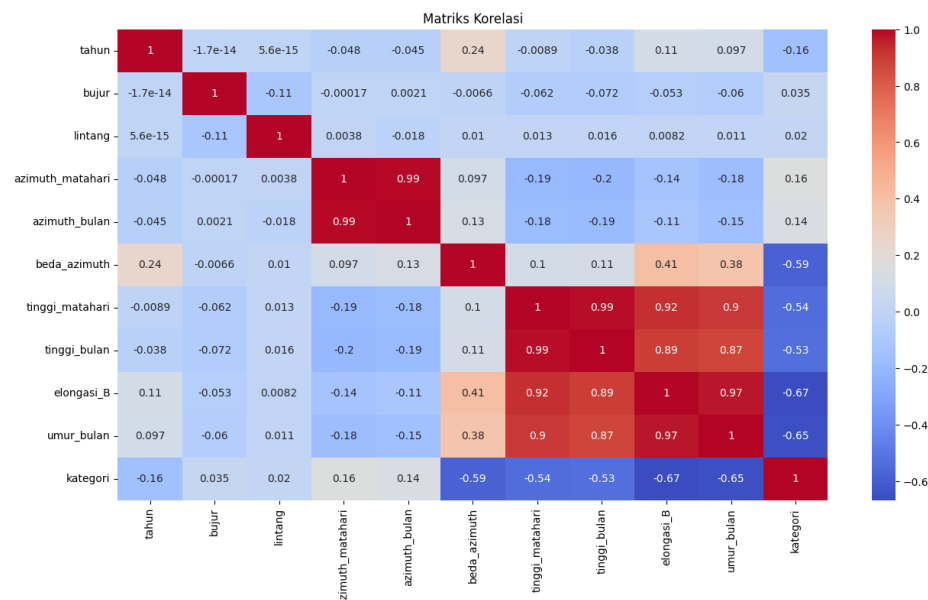
Langkah selanjutnya adalah mentransformasi variabel dependen (target) agar sesuai dengan kebutuhan model *Multinomial Logistic Regression* (MLR). Model MLR dirancang untuk memprediksi variabel target yang bersifat kategorikal nominal dengan lebih dari dua kelas.

beda_azimuth	tinggi_matahari	tinggi_bulan	elongasi_B	umur_bulan	kategori
4	0	0	4.0	6.5	3
4	0	0	4.0	6.5	3
5	0	0	5.0	7.0	1
4	0	0	4.0	6.5	3
5	-1	-1	5.0	7.0	3

Gambar 4.6 Hasil Kategorisasi

Tujuan dari penelitian ini adalah mengklasifikasikan untuk prediksi visabilitas *hilal*, yang dalam proposal ini didefinisikan ke dalam tiga kategori: “Gelap” (Tidak Terlihat), “Samar” (Kondisi Kritis/Tipis), dan “Terang” (Terlihat Jelas).

Sebelum melangkah ke tahap pemodelan, sebuah *Exploratory Data Analysis* (EDA) fundamental dilakukan untuk memahami hubungan linear antar variabel. Analisis ini diwujudkan melalui visualisasi Matriks Korelasi.



Gambar 4.7 Korelasi Berdasarkan Dataset Elongasi_B

Hasil dari *heatmap* pada Gambar 4.6 didapatkan bahwa variabel yang memiliki keterkaitan yang kuat adalah Beda Azimuth, Tinggi Matahari, Tinggi Bulan, Elongasi B, Umur Bulan, dan Kategori. Selanjutnya proses selanjutnya adalah mengambil variabel yang berkorelasi kuat antar tiap variabel.

Tabel 4.3 Variabel Input

Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan	Kategori
4	0	0	4	6,5	3
4	0	0	4	6,5	3
5	0	0	5	7,0	1
4	0	0	4	6,5	3
5	-1	-1	5	7,0	3
...
5	4	1	5,8	7,6	2
5	4	1	5,8	7,6	2
5	4	1	5,8	7,6	2
5	4	1	5,8	7,6	2
5	0	0	5	7,0	1

4.2.5 Pemodelan dan Evaluasi

Langkah pertama dalam persiapan pemodelan adalah memisahkan dataset *elongasi_B* secara logis menjadi dua komponen utama yaitu X (Fitur) dan y (Target). Matriks Fitur (X) himpunan variabel independen atau prediktor. Dalam implementasi ini, matriks ini berisi semua kolom dari dataset *elongasi_B* kecuali kolom kategori.

Tabel 4.4 Variabel Fitur (X)

Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
4	0	0	4	6,5
4	0	0	4	6,5
5	0	0	5	7
4	0	0	4	6,5
5	-1	-1	5	7
...
5	4	1	5,8	7,6
5	4	1	5,8	7,6

5	4	1	5,8	7,6
5	4	1	5,8	7,6
5	0	0	5	7,0

Variabel-variabel inilah (seperti tinggi_bulan, beda_azimuth, elongasi_B, dll.) yang akan digunakan oleh model untuk mempelajari pola. Vektor Target (y) variabel dependen atau respons yang ingin diprediksi.

Tabel 4.5 Variabel Target (y)

Kategori	Keterangan
3	Gelap
3	Gelap
1	Terang
3	Gelap
3	Gelap
...	...
2	Samar
2	Samar
2	Samar
2	Samar
1	Terang

Vektor ini hanya berisi kolom kategori (dengan nilai 1, 2, atau 3) yang telah didefinisikan pada langkah sebelumnya.

Langkah kedua, pembagian dataset dilakukan menggunakan fungsi *train_test_split* dari pustaka *Scikit-learn* dengan proporsi 60%, 70%, 80% data latih dan 40%, 30%, 20% data uji.

Tabel 4.6 Proporsi Pembagian Dataset

Model	Jumlah Data	Train	Test	Random State
A	2210	1326	884	4000
B	2210	1547	663	4000
C	2210	1768	442	4000

Selain itu, parameter *random_state*=4000 ditetapkan untuk memastikan bahwa proses pengacakan data dalam pembagian ini bersifat deterministik, sehingga hasil eksperimen dapat direproduksi (diulang) dengan hasil yang identik.

Selanjutnya, dilakukan standardisasi format waktu. Kolom matahari_terbenam dan bulan_terbenam, yang krusial untuk analisis temporal, dikonversi ke dalam format datetime *Python* yang baku.

Tabel 4.7 Hasil Standarisasi dengan *StandardScaler*

Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
0,94	1,84	1,58	2,21	2,74
-0,12	0,62	0,66	0,36	0,10
0,41	-1,08	-0,94	-0,72	-0,71
0,94	0,74	0,89	0,63	0,39
...
0,94	0,74	0,66	0,90	0,71
-1,72	0,26	0,43	-0,42	-0,55
-0,12	-0,59	-0,48	-0,83	-0,77
0,94	0,62	0,66	0,66	0,42
-0,66	1,47	1,58	1,13	1,03

Proses ini memastikan konsistensi data fitur secara akurat. Setelah seluruh data dipersiapkan, distandardisasi, dan dibagi menjadi himpunan data latih dan data uji, penelitian memasuki tahap inti: pembangunan, pelatihan, dan evaluasi model klasifikasi. Sesuai dengan desain penelitian, algoritma yang digunakan adalah *Multinomial Logistic Regression* (MLR). Konfigurasi parameter kunci model pelatihan dan evaluasi yang ditetapkan adalah sebagai berikut:

1. Model A

Tabel 4.8 Konfigurasi Model A

Model	Solver	Penalty	C	Max Iter	Random State
A1	saga	l2	0.001	1000	2
A2	saga	l2	0.01	1000	2
A3	saga	l2	0.1	1000	2
A4	saga	l2	1	1000	2
A5	saga	l2	10	1000	2
A6	saga	l2	100	1000	2

Skenario A menggunakan alokasi data 60% untuk pelatihan (1326 sampel) dan 40% untuk pengujian (884 sampel). Tabel 4.9 merangkum metrik kinerja utama untuk keenam model (A1-A6) dalam skenario ini.

Tabel 4.9 Hasil Metrik Skenario A (60:40)

Model	Solver	Penalty	C	F1-Score	Precision (Weighted)	Recall (Weighted)	RMSE	MAE
A1	saga	12	0.001	66.06%	0.43/0.56	0.50/0.66	1,91	1,82
A2	saga	12	0.01	82.47%	0.88/0.84	0.70/0.82	1,79	1,69
A3	saga	12	0.1	90.95%	0.91/0.91	0.83/0.91	1,68	1,55
A4	saga	12	1	95.36%	0.93/0.96	0.95/0.95	1,61	1,46
A5	saga	12	10	98.08%	0.97/0.98	0.97/0.98	1,57	1,42
A6	saga	12	100	97.29%	0.95/0.97	0.97/0.97	1,55	1,39

Dari Tabel 4.9, terlihat jelas adanya tren peningkatan kinerja yang signifikan seiring dengan meningkatnya nilai C. Model A1 (C=0.001) menunjukkan kinerja terendah dengan F1-Score 66.06%. Hal ini mengindikasikan bahwa regularisasi yang terlalu kuat (nilai C rendah) menyebabkan model *underfitting* atau terlalu sederhana untuk menangkap pola data.

Tabel 4.10 Koefisien Regresi untuk Fitur yang Dipilih (Model A)

Model	Kelas	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
A1	Terang	0,11	0,06	0,05	0,11	0,14
	Samar	0,12	0,08	0,08	0,06	-0,02
	Gelap	-0,23	-0,13	-0,13	-0,16	-0,12
A2	Terang	0,44	-0,02	-0,02	0,30	0,55
	Samar	0,47	0,38	0,33	0,05	-0,42
	Gelap	-0,91	-0,36	-0,31	-0,35	-0,12
A3	Terang	0,80	-0,61	-0,41	0,48	1,60
	Samar	1,51	1,88	1,00	-0,04	-2,14
	Gelap	-2,30	-1,27	-0,59	-0,44	0,55
A4	Terang	1,41	-1,29	-0,28	-0,31	3,28
	Samar	3,38	5,77	0,68	0,77	-5,93
	Gelap	-4,78	-4,49	-0,40	-0,46	2,65
A5	Terang	3,16	0,96	-0,72	-4,99	6,37
	Samar	6,10	11,46	-0,53	4,11	-12,55
	Gelap	-9,26	-12,42	1,25	0,87	6,18
A6	Terang	6,15	6,21	-2,49	-12,72	10,94
	Samar	10,07	18,61	-1,23	9,25	-22,19

	Gelap	-16,22	-24,82	3,72	3,47	11,26
--	-------	--------	--------	------	------	-------

Tabel 4.10 menyajikan koefisien regresi (*log-odds*) yang dihasilkan oleh Model A, yang dilatih dengan pembagian data 60:40. Setiap baris (A1 hingga A6) merepresentasikan model yang dilatih dengan nilai parameter regularisasi C yang berbeda, mulai dari $C=0.001$ (regularisasi kuat) hingga $C=100$ (regularisasi lemah). Semakin tinggi nilai C (semakin lemah regularisasi), semakin besar nilai (nilai *absolut*) dari koefisien yang dihasilkan.

Sebagai contoh, pada Model A1 ($C=0.001$), koefisien Tinggi Matahari hanya -0.133892, namun pada Model A6 ($C=100$), nilainya melonjak menjadi -24.817777. Ini mengindikasikan bahwa untuk mereplikasi aturan Model Pakar yang kompleks, Model MLR (Model Aproksimasi) membutuhkan fleksibilitas yang tinggi (regularisasi lemah) untuk memberikan bobot (signifikansi) yang besar pada fitur-fitur tertentu. Model A6, dengan koefisien terbesar, mewakili model yang paling kompleks dan paling "bebas" dalam meniru data latih.

Tabel 4.11 *Log-Odds* (Koefisien Regresi) untuk Kategori ‘Gelap’ Model A

Model	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
A1	-0,23	-0,13	-0,13	-0,16	-0,12
A2	-0,91	-0,36	-0,31	-0,35	-0,12
A3	-2,30	-1,27	-0,59	-0,44	0,55
A4	-4,78	-4,49	-0,40	-0,46	2,65
A5	-9,26	-12,42	1,25	0,87	6,18
A6	-16,22	-24,82	3,72	3,47	11,26

Tabel 4.11 berfokus pada *Log-Odds* (koefisien) yang spesifik untuk mereplikasi Kategori 3 (‘Gelap’). Nilai-nilai ini adalah bobot matematis yang ditemukan oleh MLR untuk setiap fitur dalam “memutuskan” apakah sebuah observasi harus diklasifikasikan sebagai ‘Gelap’ (sesuai aturan Model Pakar).

1. Koefisien Negatif

Pada $\text{tinggi_matahari} = -24.82$ pada A6, menunjukkan hubungan terbalik. Semakin tinggi nilai tinggi_matahari , semakin kecil *log-odds* (peluang) observasi tersebut diklasifikasikan sebagai ‘Gelap’. Ini adalah penghambat (*inhibitor*) terkuat.

2. Koefisien Positif

Pada $\text{umur_bulan} = +11.26$ pada A6, Menunjukkan hubungan searah. Semakin tinggi umur_bulan , semakin besar *log-odds* (peluang) observasi tersebut diklasifikasikan sebagai ‘Gelap’ oleh model ini.

Analisis koefisien pada Model A6 (model paling kompleks) menunjukkan bahwa tinggi_matahari (-24.82), umur_bulan (+11.26), dan beda_azimuth (-16.22) adalah fitur-fitur dengan dampak paling signifikan yang dipelajari model untuk meniru aturan ‘Gelap’.

Tabel 4.12 *Odds Ratio* untuk Kategori ‘Gelap’ Model A

Model	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
A1	0,80	0,87	0,88	0,85	0,88
A2	0,40	0,70	0,74	0,71	0,88
A3	0,10	0,28	0,55	0,64	1,73
A4	0,01	0,01	0,67	0,63	14,13
A5	0,00	0,00	3,47	2,39	480,85
A6	9,00	1,67	4,13	3,22	7,74

Tabel 4.12 untuk mempermudah interpretasi koefisien *log-odds* (yang sulit dipahami secara intuitif), nilai-nilai tersebut diubah menjadi *Odds Ratio* dengan menghitung nilai eksponensialnya ($e^{\text{koefisien}}$). Interpretasi dari Model A6 (C=100) berdasarkan tabel ini menunjukkan pola yang unik di mana seluruh fitur memiliki nilai OR > 1:

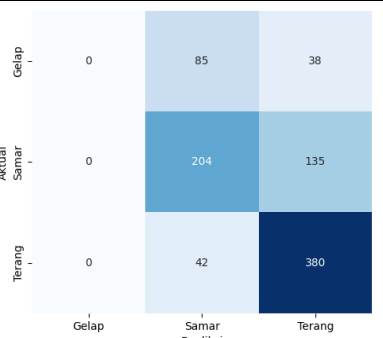
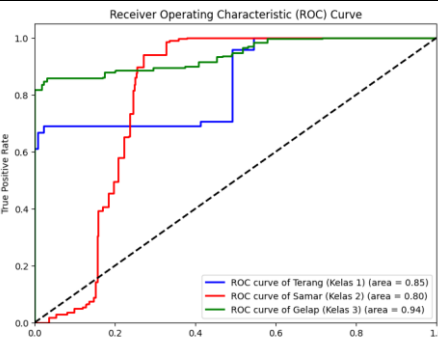
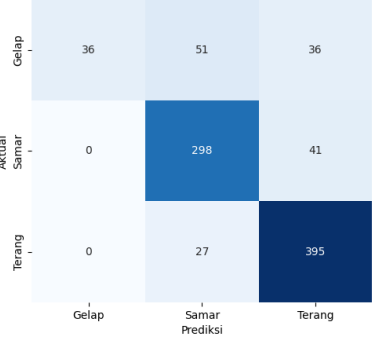
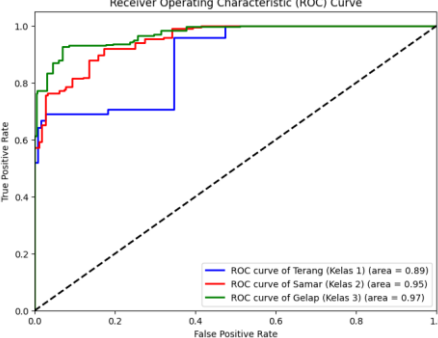
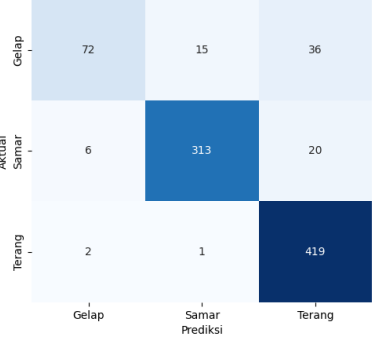
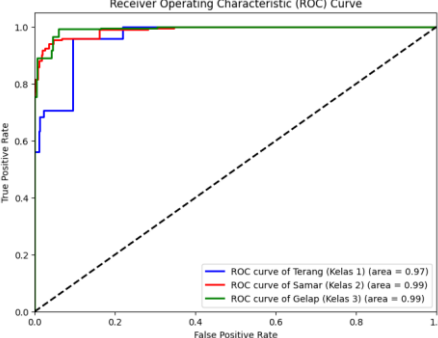
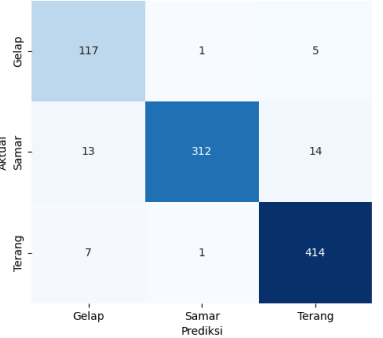
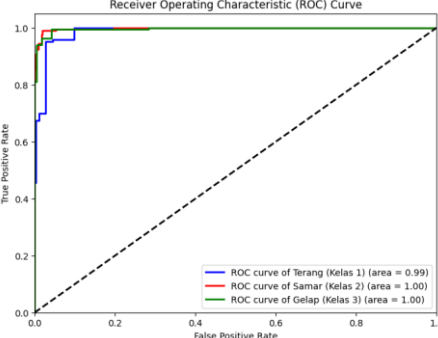
1. Nilai > 1 (Peningkat Peluang):

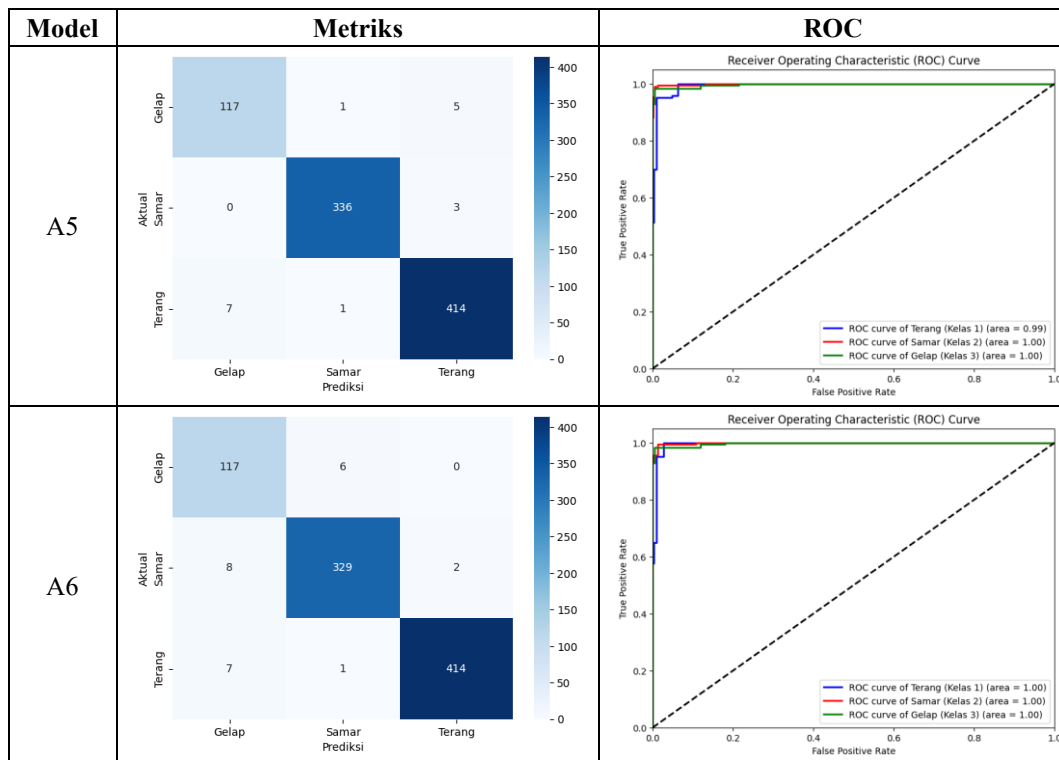
- Beda Azimuth (OR = 9.00): Fitur ini memiliki dampak terbesar dalam tabel ini. Setiap kenaikan 1 unit pada Beda Azimuth, peluang (*odds*) untuk diklasifikasikan ‘Gelap’ dikali 9.00.
- Umur Bulan (OR = 7.74): Setiap kenaikan 1 unit (jam) pada Umur Bulan, peluang (*odds*) untuk ‘Gelap’ dikali 7.74. Meskipun berpengaruh positif, nilainya tidak sebesar yang terlihat pada model sebelumnya (A5).
- Tinggi Bulan (OR = 4.13): Setiap kenaikan 1 derajat, peluang dikali 4.13.
- Tinggi Matahari (OR = 1.67): Berbeda dengan model sebelumnya, pada tabel ini Tinggi Matahari menunjukkan nilai positif (>1), di mana setiap kenaikan unit sedikit meningkatkan peluang klasifikasi ‘Gelap’ sebesar 1.67 kali.

Analisis pada Tabel 4.12 untuk Model A6 menunjukkan pergeseran karakteristik dibandingkan model A5. Jika sebelumnya Umur Bulan adalah faktor dominan ekstrem, pada Model A6 pengaruhnya lebih moderat (7.74), justru Beda Azimuth (8.99) muncul sebagai faktor dengan pengali odds tertinggi dalam struktur data tabel ini.

Kinerja model membaik secara drastis hingga mencapai puncaknya pada Model A6 ($C=100$), yang mencatatkan *F1-Score* 97.29%, serta *presisi* dan *recall* yang hampir sempurna. Tabel 4.13 menyajikan visualisasi detail berupa *Confusion Matrix* dan kurva ROC untuk setiap model dalam Skenario A.

Tabel 4.13 Detail Visual Model Skenario A

Model	Metriks	ROC
A1		
A2		
A3		
A4		



Menunjukkan *confusion* yang signifikan pada *confusion matrix*-nya, terutama dalam membedakan kelas “Samar” (125 sampel salah diprediksi sebagai “Terang”) dan “Gelap” (38 sampel salah diprediksi sebagai “Terang”). Kurva ROC untuk kelas 2 (“Samar”) juga paling rendah (Area = 0.80).

Sebaliknya, menampilkan *confusion matrix* yang nyaris sempurna. Kurva ROC untuk ketiga kelas (Gelap, Samar, Terang) mencapai nilai ideal 1.00, menunjukkan kemampuan diskriminatif (pembedaan) yang sempurna pada data uji.

2. Model B

Tabel 4.14 Konfigurasi Model B

Model	Solver	Penalty	C	Max Iter	Random State
B1	saga	12	0.001	1000	2
B2	saga	12	0.01	1000	2
B3	saga	12	0.1	1000	2
B4	saga	12	1	1000	2
B5	saga	12	10	1000	2

B6	saga	12	100	1000	2
----	------	----	-----	------	---

Skenario B menyesuaikan rasio pembagian menjadi 70% untuk pelatihan (1547 sampel) dan 30% untuk pengujian (663 sampel). Tabel 4.15 merangkum metrik kinerja untuk keenam model (B1-B6).

Tabel 4.15 Hasil Metrik Skenario B (70:30)

Model	Solver	Penalty	C	F1-Score	Precision (Weighted)	Recall (Weighted)	RMSE	MAE
B1	saga	12	0.001	66.52%	0.44/0.56	0.51/0.67	1,90	1,81
B2	saga	12	0.01	82.20%	0.87/0.84	0.72/0.82	1,78	1,68
B3	saga	12	0.1	89.74%	0.90/0.91	0.82/0.90	1,67	1,54
B4	saga	12	1	95.63%	0.95/0.96	0.95/0.96	1,60	1,45
B5	saga	12	10	97.74%	0.97/0.98	0.97/0.98	1,56	1,41
B6	saga	12	100	97.89%	0.97/0.98	0.97/0.98	1,55	1,38

Kinerja model berbanding lurus dengan peningkatan nilai C. Model B1 (C=0.001) memiliki kinerja awal yang serupa dengan A1, yaitu F1-Score 66.52%. Model B6 (C=100) kembali menjadi yang terbaik dalam skenario ini, dengan F1-Score 97.89%.

Tabel 4.16 Koefisien Regresi untuk Fitur yang Dipilih (Model B)

Model	Kelas	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
B1	Terang	0,12	0,07	0,06	0,12	0,16
	Samar	0,13	0,08	0,08	0,06	-0,03
	Gelap	-0,25	-0,15	-0,15	-0,18	-0,13
B2	Terang	0,47	-0,02	-0,02	0,32	0,60
	Samar	0,51	0,42	0,36	0,04	-0,49
	Gelap	-0,98	-0,39	-0,34	-0,36	-0,11
B3	Terang	0,82	-0,64	-0,40	0,47	1,70
	Samar	1,60	2,00	1,06	-0,03	-2,31
	Gelap	-2,43	-1,36	-0,66	-0,44	0,61
B4	Terang	1,47	-1,27	-0,21	-0,52	3,46
	Samar	3,51	5,92	0,70	0,94	-6,23
	Gelap	-4,97	-4,65	-0,50	-0,41	2,78
B5	Terang	3,27	1,15	-0,75	-5,66	6,90
	Samar	6,30	11,56	-0,38	4,58	-13,17
	Gelap	-9,57	-12,71	1,13	1,08	6,27
B6	Terang	6,22	6,19	-2,52	-13,68	11,92
	Samar	10,32	18,77	-1,04	9,81	-22,99

Model	Kelas	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
	Gelap	-16,53	-24,96	3,56	3,87	11,08

Tabel 4.16 merangkum koefisien regresi (*log-odds*) yang dihasilkan oleh Model B, yang dilatih menggunakan pembagian data 70:30. Setiap baris (B1 hingga B6) merepresentasikan model yang dilatih dengan nilai parameter regularisasi C yang berbeda. Sebagaimana terlihat pada Model A, terdapat tren yang konsisten: semakin tinggi nilai C (regularisasi semakin lemah), nilai (*nilai absolut*) dari koefisien regresi meningkat secara drastis.

Sebagai contoh, pada Model B1 ($C=0.001$), model sangat sederhana dan koefisien Tinggi Matahari (Gelap) hanya -0.147225. Namun, pada Model B6 ($C=100$), koefisiennya melonjak menjadi -24.956491. Ini adalah temuan kunci yang menunjukkan bahwa untuk mereplikasi aturan Model Pakar yang kompleks dan *non-linier*, Model MLR (Model Aproksimasi) memerlukan fleksibilitas yang sangat tinggi (regularisasi rendah) untuk dapat memberikan bobot yang signifikan pada fitur-fitur yang dianggapnya paling penting.

Tabel 4.17 *Log-Odds* (Koefisien Regresi) untuk Kategori ‘Gelap’ Model B

Model	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
B1	-0,25	-0,15	-0,15	-0,18	-0,13
B2	-0,98	-0,39	-0,34	-0,36	-0,11
B3	-2,43	-1,36	-0,66	-0,44	0,61
B4	-4,97	-4,65	-0,50	-0,41	2,78
B5	-9,57	-12,71	1,13	1,08	6,27
B6	-16,53	-24,96	3,56	3,87	11,08

Tabel 4.17 menyajikan nilai *Log-Odds* (koefisien regresi) yang diperoleh MLR untuk mereplikasi Kategori ‘Gelap’ pada Model B. Setiap koefisien menunjukkan seberapa kuat serta arah kontribusi masing-masing fitur dalam

meningkatkan atau menurunkan peluang suatu observasi diklasifikasikan sebagai ‘Gelap’ sesuai aturan Model Pakar.

1. Koefisien Negatif

Terlihat pada beberapa model, seperti $\text{beda_azimuth} = -16.53$, $\text{tinggi_matahari} = -24.96$, dan $\text{tinggi_bulan} = -0.49$ pada Model B6. Nilai negatif menandakan hubungan terbalik: semakin besar nilai variabel tersebut, semakin menurun log-odds bahwa kondisi tersebut dianggap ‘Gelap’. Pada model paling kompleks (B6), tinggi_matahari menjadi inhibitor paling kuat.

2. Koefisien Positif

Pada Model B6, fitur seperti $\text{tinggi_bulan} = +3.56$, $\text{elongasi_b} = +3.87$, dan terutama $\text{umur_bulan} = +11.08$. Menunjukkan bahwa kenaikan variabel-variabel ini justru meningkatkan peluang klasifikasi ‘Gelap’. Dengan demikian, fitur positif berperan sebagai penguat (enhancer) probabilitas kategori tersebut.

Analisis menyeluruh pada Model B6 (model paling kompleks) mengungkap bahwa tinggi_matahari (-24.96) adalah faktor penghambat paling dominan, sedangkan umur_bulan ($+11.08$) dan elongasi_b ($+3.87$) menjadi faktor pendorong terbesar dalam pembentukan keputusan model. Dengan kata lain, kombinasi sinergis dari variabel inhibitor dan enhancer inilah yang dipelajari MLR untuk meniru pola klasifikasi ‘Gelap’ versi Model Pakar pada Model B.

Tabel 4.18 *Odds Ratio* untuk Kategori ‘Gelap’ Model B

Model	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
B1	0,78	0,86	0,87	0,84	0,88
B2	0,37	0,67	0,71	0,70	0,89
B3	0,09	0,26	0,52	0,64	1,84
B4	0,01	0,01	0,61	0,66	16,06
B5	0,00	0,00	3,11	2,96	530,03
B6	6,60	1,45	3,52	4,81	6,47

Untuk mempermudah interpretasi koefisien *log-odds* yang ditampilkan pada Tabel 4.18, nilai-nilai tersebut dikonversi menjadi *Odds Ratio* (OR) dengan cara menghitung eksponensial dari masing-masing koefisien ($e^{\text{koefisien}}$). *Odds Ratio* memberikan gambaran yang jauh lebih intuitif karena menunjukkan seberapa besar peluang suatu observasi meningkat atau menurun untuk diklasifikasikan sebagai ‘Gelap’ ketika suatu variabel naik satu satuan.

Model B6, sebagai model yang paling kompleks dalam Skenario B, menunjukkan pola yang sangat kontras dibandingkan model-model sebelumnya:

1. Nilai $OR > 1$ (Peningkat Peluang ‘Gelap’)

- Beda Azimuth (OR = 6.60): Termasuk salah satu enhancer terkuat. Setiap kenaikan 1 unit pada beda azimuth meningkatkan odds ‘Gelap’ sebesar 6.6 kali.
- Tinggi Matahari (OR = 1.45): Menarik karena sebelumnya (pada Model B1–B5) fitur ini justru selalu memiliki $OR < 1$. Pada Model B6, arah signifikansinya berubah: setiap kenaikan 1 unit pada tinggi matahari meningkatkan odds menjadi 1.45 kali.
- Tinggi Bulan (OR = 3.52): Setiap kenaikan 1 derajat pada tinggi bulan meningkatkan peluang ‘Gelap’ sebesar 3.52 kali.
- Elongasi Bulan (OR = 4.81): Termasuk penguat yang kuat, dengan pengali odds mencapai hampir 5 kali.
- Umur Bulan (OR = 6.47): Menjadi salah satu faktor dominan. Setiap kenaikan umur bulan meningkatkan peluang ‘Gelap’ lebih dari 6.4 kali.

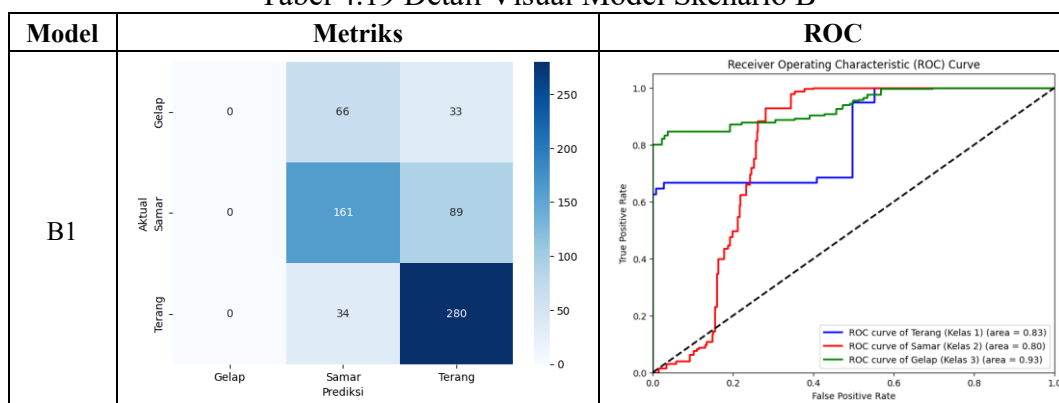
Berbeda dari model-model sebelumnya yang sering menunjukkan $OR < 1$ (artinya fitur cenderung menurunkan peluang ‘Gelap’), Model B6

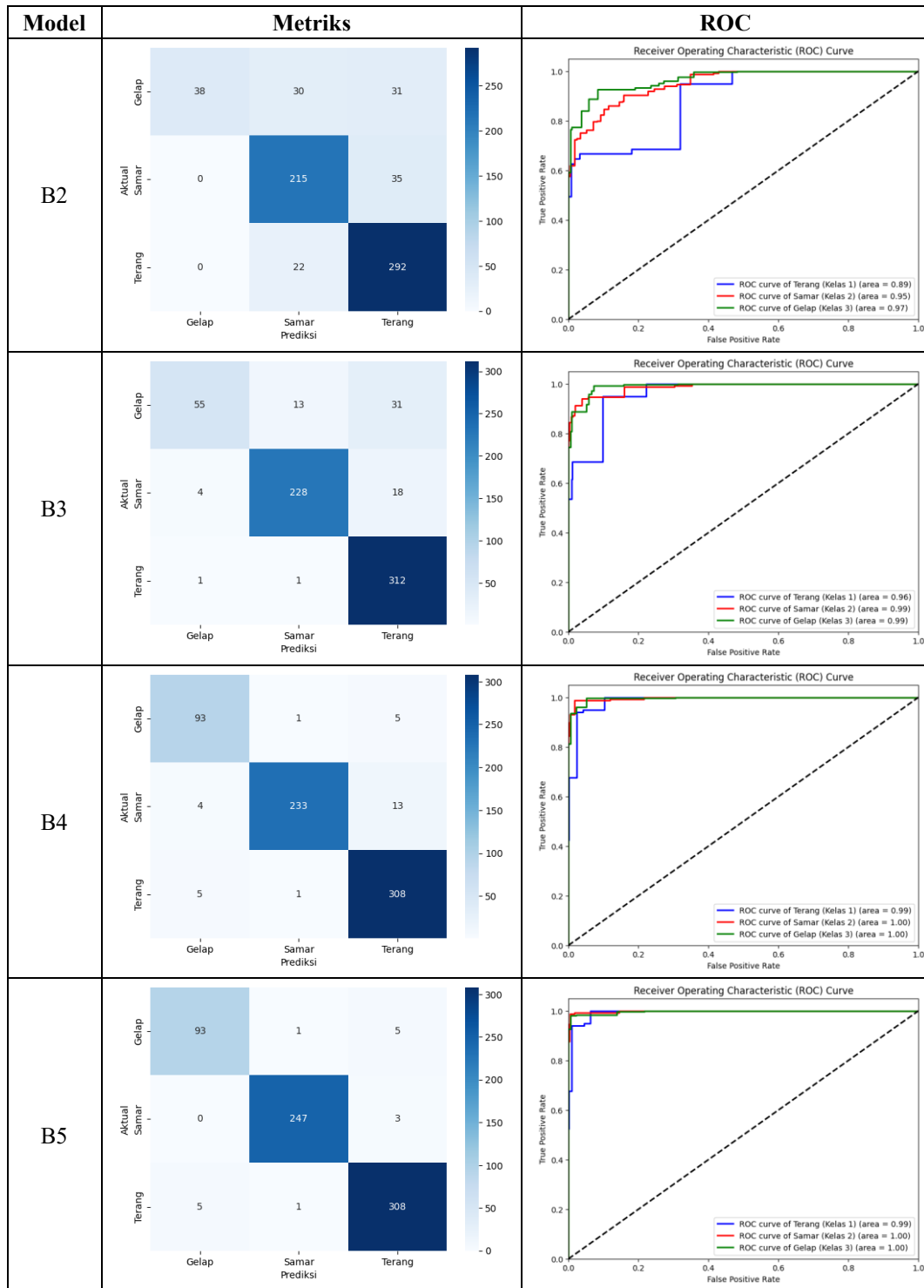
memperlihatkan konsistensi arah positif pada seluruh fitur. Hampir semua variabel mengalami pembalikan arah signifikansi dibanding Model B5, terutama Tinggi Matahari, yang sebelumnya merupakan inhibitor ekstrem ($OR \approx 0.000003$ pada B5), kini menjadi penguat moderat (1.45). Umur Bulan, yang pada B5 menunjukkan nilai luar biasa besar (530.03), mengalami normalisasi menjadi 6.47 lebih masuk akal dan stabil. Beda Azimuth yang awalnya cenderung mengurangi peluang ($OR < 1$ pada sebagian besar model sederhana), pada B6 melonjak menjadi 6.60, menjadikannya fitur paling dominan kedua.

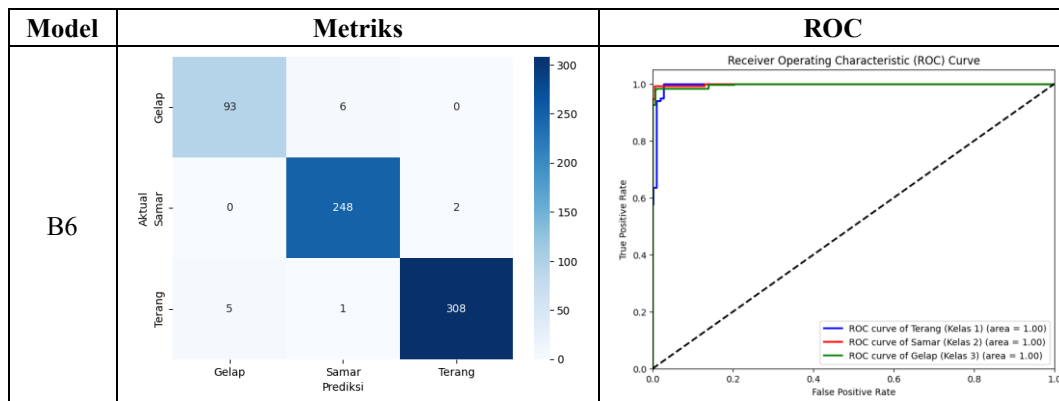
Secara keseluruhan, Tabel 4.18 menggambarkan bahwa Model B6 telah mempelajari pola yang jauh lebih stabil, terkalibrasi, dan realistis dalam meniru aturan pakar untuk kategori ‘Gelap’. Pola OR yang konsisten > 1 menegaskan bahwa model memandang seluruh fitur sebagai penambah peluang gelap, namun dengan derajat pengaruh yang berbeda-beda.

Dengan kata lain, Model B6 tidak hanya memperbaiki akurasi prediksi, tetapi juga menghasilkan struktur parameter yang jauh lebih logis secara fisik dan konsisten secara matematis. Tabel 4.19 menyajikan visualisasi detail berupa Confusion Matrix dan kurva ROC untuk setiap model dalam Skenario B.

Tabel 4.19 Detail Visual Model Skenario B







Menunjukkan kesalahan klasifikasi yang signifikan pada *confusion matrix*-nya, di mana 81 sampel “Samar” salah diprediksi sebagai “Terang”. Kurva ROC juga menunjukkan area terendah untuk kelas “Samar” (0.81).

Sekali lagi menunjukkan hasil yang nyaris sempurna pada *confusion matrix* dan mencapai nilai ROC 1.00 untuk ketiga kelas, mengonfirmasi temuan dari Skenario A bahwa nilai C=100 menghasilkan model yang sangat kuat.

3. Model C

Tabel 4.20 Konfigurasi Model C

Model	Solver	Penalty	C	Max Iter	Random State
B1	saga	12	0.001	1000	2
B2	saga	12	0.01	1000	2
B3	saga	12	0.1	1000	2
B4	saga	12	1	1000	2
B5	saga	12	10	1000	2
B6	saga	12	100	1000	2

Tabel 4.21 Hasil Metrik Skenario C (80:20)

Model	Solver	Penalty	C	F1-Score	Precision (Weighted)	Recall (Weighted)	RMSE	MAE
C1	saga	12	0.001	68.33%	0.45/0.58	0.52/0.68	1,91	1,82
C2	saga	12	0.01	84.84%	0.89/0.86	0.76/0.85	1,79	1,69
C3	saga	12	0.1	91.18%	0.93/0.92	0.83/0.91	1,69	1,56
C4	saga	12	1	95.93%	0.95/0.96	0.95/0.96	1,61	1,47
C5	saga	12	10	97.74%	0.97/0.98	0.97/0.98	1,58	1,43
C6	saga	12	100	97.96%	0.97/0.98	0.97/0.98	1,57	1,41

Tabel 4.21 menampilkan tren yang konsisten kembali terkonfirmasi. Model C1 ($C=0.001$) memiliki F1-Score 68.33%. Kinerja puncak dicapai Model C6 ($C=100$) dengan F1-Score 97.96%, menunjukkan kinerja yang sangat tinggi dan setara dengan Model B6.

Tabel 4.22 Koefisien Regresi untuk Fitur yang Dipilih (Model C)

Model	Kelas	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
C1	Terang	0,14	0,07	0,06	0,13	0,18
	Samar	0,14	0,09	0,09	0,06	-0,04
	Gelap	-0,28	-0,15	-0,15	-0,19	-0,13
C2	Terang	0,50	-0,06	-0,05	0,33	0,64
	Samar	0,55	0,47	0,41	0,03	-0,56
	Gelap	-1,05	-0,41	-0,35	-0,36	-0,09
C3	Terang	0,83	-0,72	-0,43	0,48	1,80
	Samar	1,71	2,16	1,14	-0,05	-2,53
	Gelap	-2,54	-1,44	-0,70	-0,43	0,73
C4	Terang	1,50	-1,27	-0,25	-0,63	3,61
	Samar	3,68	6,18	0,78	1,06	-6,71
	Gelap	-5,18	-4,91	-0,54	-0,43	3,10
C5	Terang	3,41	1,48	-0,96	-5,94	7,11
	Samar	6,66	12,04	-0,22	5,00	-14,22
	Gelap	-10,07	-13,52	1,19	0,94	7,12
C6	Terang	6,45	6,56	-2,75	-13,78	12,02
	Samar	10,90	19,61	-0,88	10,36	-24,57
	Gelap	-17,35	-26,16	3,63	3,42	12,54

Tabel 4.22 menampilkan koefisien regresi (*log-odds*) yang dihasilkan oleh Model C, yang dilatih menggunakan pembagian data 80:20. Setiap baris (C1 hingga C6) merepresentasikan model yang dilatih dengan nilai parameter regularisasi C yang berbeda. Seperti halnya pada skenario Model A dan B, hasil Model C mengkonfirmasi tren yang sangat konsisten: semakin tinggi nilai C (regularisasi semakin lemah), nilai (nilai *absolut*) dari koefisien regresi meningkat secara signifikan.

Sebagai contoh, pada Model C1 ($C=0.001$), koefisien Tinggi Matahari (Gelap) hanya -0.152712. Namun, pada Model C6 ($C=100$), koefisiennya melonjak

menjadi -26.161715. Ini membuktikan bahwa Model MLR (Model Aproksimasi) memerlukan fleksibilitas yang tinggi (regularisasi rendah) untuk dapat “belajar” dan meniru bobot fitur yang kompleks dari aturan Model Pakar. Model dengan C tinggi (C6) adalah model yang paling kompleks dan paling akurat dalam mereplikasi *ground truth sintetis*.

Tabel 4.23 *Log-Odds* (Koefisien Regresi) untuk Kategori ‘Gelap’ Model C

Model	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
C1	-0,28	-0,15	-0,15	-0,19	-0,13
C2	-1,05	-0,41	-0,35	-0,36	-0,09
C3	-2,54	-1,44	-0,70	-0,43	0,73
C4	-5,18	-4,91	-0,54	-0,43	3,10
C5	-10,07	-13,52	1,19	0,94	7,12
C6	-17,35	-26,16	3,63	3,42	12,54

Tabel 4.23 berfokus pada nilai *Log-Odds* (koefisien regresi) yang dipelajari oleh MLR untuk mereplikasi Kategori ‘Gelap’ pada Model C. Setiap nilai koefisien merupakan bobot matematis yang digunakan model dalam “memutuskan” apakah suatu observasi seharusnya diklasifikasikan sebagai ‘Gelap’ sesuai dengan aturan Model Pakar.

1. Koefisien Negatif

Hampir seluruh fitur pada model-model C1–C5 memiliki nilai koefisien negatif, misalnya *tinggi_matahari* = -26.16 pada C6, *beda_azimuth* = -17.35 pada C6, *tinggi_bulan* = -13.52 pada C5. Nilai negatif menunjukkan hubungan terbalik: semakin besar nilai fitur tersebut, semakin menurun *log-odds* bahwa kondisi tersebut diklasifikasikan sebagai ‘Gelap’. Dalam model paling kompleks (C6), *tinggi_matahari* menjadi inhibitor paling kuat karena memiliki koefisien negatif terbesar.

2. Koefisien Positif

Berbeda dengan model-model sebelumnya, Model C6 mulai memperlihatkan beberapa koefisien positif, misalnya $\text{tinggi_bulan} = +3.63$, $\text{elongasi_b} = +3.42$, $\text{umur_bulan} = +12.54$. Koefisien positif menunjukkan hubungan searah, semakin tinggi nilai fitur tersebut, semakin besar *log-odds* bahwa observasi akan diklasifikasikan sebagai ‘Gelap’.

Analisis koefisien pada Model C6 (model paling kompleks) menunjukkan bahwa $\text{tinggi_matahari} (-26.16)$ dan $\text{beda_azimuth} (-17.35)$ adalah faktor penghambat (*inhibitor*) paling dominan, sementara $\text{umur_bulan} (+12.54)$ dan $\text{tinggi_bulan} (+3.63)$ menjadi faktor penguat (*enhancer*) yang paling signifikan.

Pola ini menunjukkan bahwa Model C6 mempelajari kombinasi antara variabel penghambat dan pendorong dalam meniru aturan pakar untuk kategori ‘Gelap’, dengan struktur yang jauh lebih ekstrem dibanding model-model sebelumnya.

Tabel 4.24 *Odds Ratio* untuk Kategori ‘Gelap’ Model C

Model	Beda Azimuth	Tinggi Matahari	Tinggi Bulan	Elongasi B	Umur Bulan
C1	0,75	0,86	0,86	0,83	0,87
C2	0,35	0,66	0,70	0,70	0,92
C3	0,08	0,24	0,50	0,65	2,08
C4	0,01	0,01	0,59	0,65	22,16
C5	0,00	0,00	3,28	2,56	634,23
C6	2,91	4,35	3,76	3,07	2,80

Untuk mempermudah interpretasi koefisien *log-odds* yang ditampilkan pada Tabel 4.24, nilai-nilai tersebut dikonversi menjadi *Odds Ratio* (OR) pada Tabel 4.20 dengan menghitung eksponensial dari setiap koefisien ($e^{\text{koefisien}}$). *Odds Ratio* memberikan interpretasi yang lebih intuitif karena menunjukkan seberapa besar peluang suatu observasi meningkat atau menurun untuk diklasifikasikan sebagai ‘Gelap’ ketika suatu variabel meningkat satu satuan.

Model C6, sebagai model yang paling kompleks dalam Skenario C, memperlihatkan pola yang jauh berbeda dibandingkan model-model C1–C5.

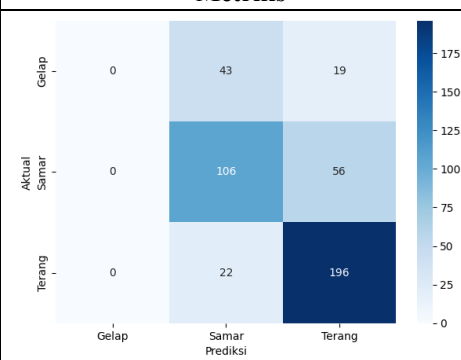
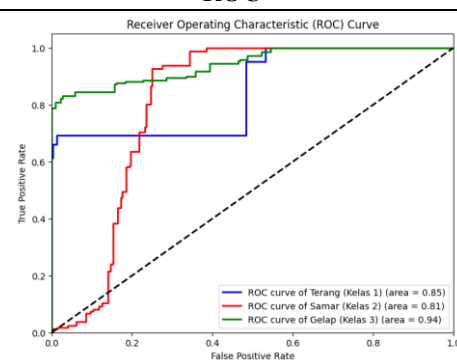
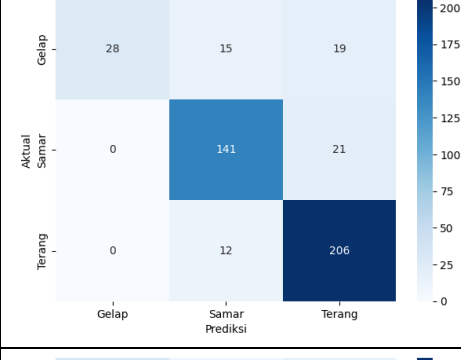
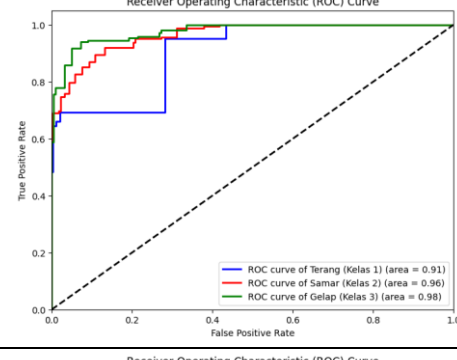
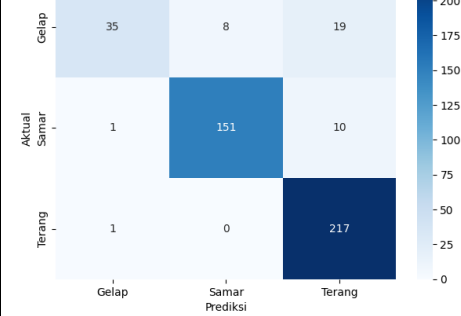
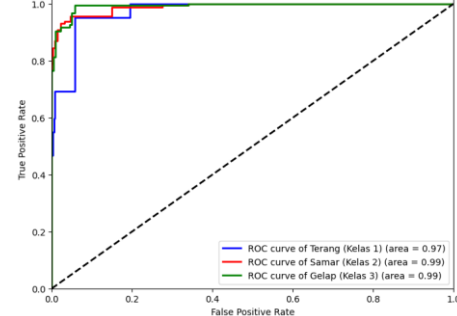
1. Nilai $OR > 1$ (Peningkat Peluang ‘Gelap’)

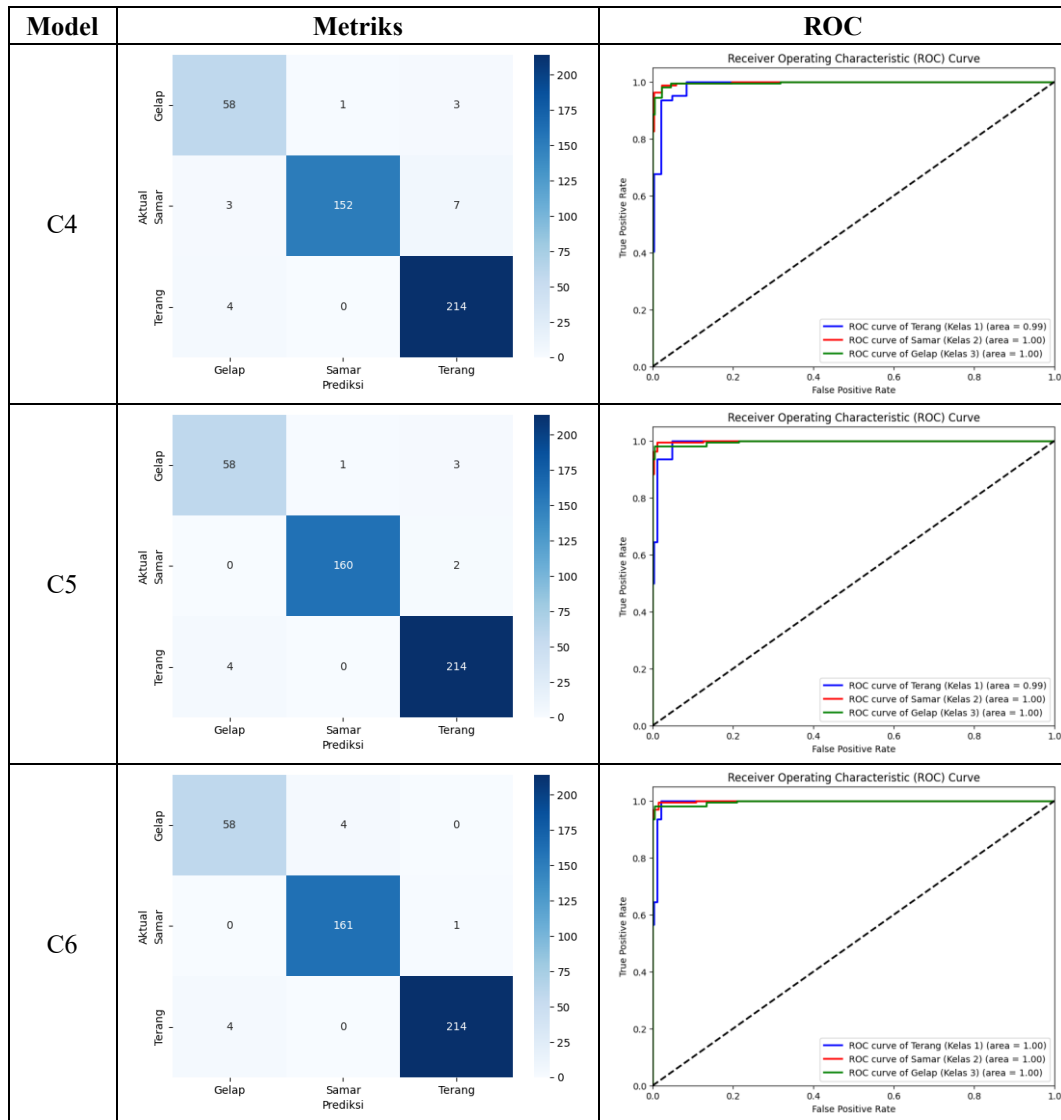
- Tinggi Matahari ($OR = 4.35$): Menarik karena pada model-model sebelumnya (C1–C5), fitur ini selalu menunjukkan $OR < 1$, yang berarti menghambat peluang ‘Gelap’. Pada Model C6, arahnya berbalik menjadi penguat, dengan kenaikan 1 unit meningkatkan *odds* ‘Gelap’ sebesar 4.35 kali.
- Tinggi Bulan ($OR = 3.75$): Setiap kenaikan tinggi bulan meningkatkan peluang klasifikasi sebesar 3.75 kali. Ini menandai perubahan yang signifikan dibanding model awal (C1–C3) yang cenderung memiliki nilai < 1 .
- Elongasi Bulan ($OR = 3.06$): Bekerja sebagai penguat (*enhancer*) dengan pengali *odds* mencapai sekitar 3 kali, mendukung struktur sinyal yang konsisten dengan model-model sebelumnya.
- Beda Azimuth ($OR = 2.91$): Meski lebih moderat dibanding Model B6, nilai ini menunjukkan peningkatan peluang hingga 2.9 kali, menandakan perubahan arah drastis dari model sederhana (C1–C4) yang semuanya < 1 .
- Umur Bulan ($OR = 2.80$): Berbeda dari C5, di mana Umur Bulan menunjukkan OR yang sangat ekstrem (634.23), Model C6 kembali menghasilkan nilai yang jauh lebih stabil dan realistis, yaitu 2.80. Nilai ini tetap positif dan menunjukkan peran penting fitur tersebut.

Model-model awal (C1–C4) didominasi oleh nilai $OR < 1$, yang mengindikasikan bahwa sebagian besar fitur justru mengurangi peluang suatu

kondisi diklasifikasikan sebagai ‘Gelap’. Bahkan beberapa fitur, seperti tinggi matahari dan beda azimuth, menunjukkan nilai sangat kecil (misalnya $OR \approx 0.007$ pada C4), sehingga berfungsi sebagai *inhibitor* kuat. Secara khusus, Umur Bulan yang sebelumnya ekstrem pada C5 ($OR > 630$) kini stabil pada $OR = 2.80$, menunjukkan bahwa C6 mempelajari pola pengaruh yang lebih realistis sesuai karakteristik astronomis sebenarnya. Tabel 4.25 dapat melengkapi analisis ini melalui visualisasi *Confusion Matrix* maupun kurva ROC untuk setiap model dalam Skenario C.

Tabel 4.25 Detail Visual Model Skenario C

Model	Metriks			ROC																
C1	 <table><tr><th></th><th>Actual Gelap</th><th>Actual Samar</th><th>Actual Terang</th></tr><tr><th>Predicted Gelap</th><td>0</td><td>0</td><td>0</td></tr><tr><th>Predicted Samar</th><td>43</td><td>106</td><td>22</td></tr><tr><th>Predicted Terang</th><td>19</td><td>56</td><td>196</td></tr></table>				Actual Gelap	Actual Samar	Actual Terang	Predicted Gelap	0	0	0	Predicted Samar	43	106	22	Predicted Terang	19	56	196	 <p>Receiver Operating Characteristic (ROC) Curve</p> <p>True Positive Rate</p> <p>False Positive Rate</p> <p>ROC curve of Terang (Kelas 1) (area = 0.85) ROC curve of Samar (Kelas 2) (area = 0.81) ROC curve of Gelap (Kelas 3) (area = 0.94)</p>
		Actual Gelap	Actual Samar	Actual Terang																
Predicted Gelap	0	0	0																	
Predicted Samar	43	106	22																	
Predicted Terang	19	56	196																	
C2	 <table><tr><th></th><th>Actual Gelap</th><th>Actual Samar</th><th>Actual Terang</th></tr><tr><th>Predicted Gelap</th><td>28</td><td>0</td><td>0</td></tr><tr><th>Predicted Samar</th><td>15</td><td>141</td><td>12</td></tr><tr><th>Predicted Terang</th><td>19</td><td>21</td><td>206</td></tr></table>				Actual Gelap	Actual Samar	Actual Terang	Predicted Gelap	28	0	0	Predicted Samar	15	141	12	Predicted Terang	19	21	206	 <p>Receiver Operating Characteristic (ROC) Curve</p> <p>True Positive Rate</p> <p>False Positive Rate</p> <p>ROC curve of Terang (Kelas 1) (area = 0.91) ROC curve of Samar (Kelas 2) (area = 0.96) ROC curve of Gelap (Kelas 3) (area = 0.98)</p>
		Actual Gelap	Actual Samar	Actual Terang																
Predicted Gelap	28	0	0																	
Predicted Samar	15	141	12																	
Predicted Terang	19	21	206																	
C3	 <table><tr><th></th><th>Actual Gelap</th><th>Actual Samar</th><th>Actual Terang</th></tr><tr><th>Predicted Gelap</th><td>35</td><td>1</td><td>1</td></tr><tr><th>Predicted Samar</th><td>8</td><td>151</td><td>0</td></tr><tr><th>Predicted Terang</th><td>19</td><td>10</td><td>217</td></tr></table>				Actual Gelap	Actual Samar	Actual Terang	Predicted Gelap	35	1	1	Predicted Samar	8	151	0	Predicted Terang	19	10	217	 <p>Receiver Operating Characteristic (ROC) Curve</p> <p>True Positive Rate</p> <p>False Positive Rate</p> <p>ROC curve of Terang (Kelas 1) (area = 0.97) ROC curve of Samar (Kelas 2) (area = 0.99) ROC curve of Gelap (Kelas 3) (area = 0.99)</p>
		Actual Gelap	Actual Samar	Actual Terang																
Predicted Gelap	35	1	1																	
Predicted Samar	8	151	0																	
Predicted Terang	19	10	217																	



Menunjukkan kesalahan klasifikasi yang serupa dengan Skenario A dan B, di mana 51 sampel “Samar” salah diprediksi sebagai “Terang”. Kurva ROC-nya pun belum optimal.

Yang merupakan model final pilihan, menunjukkan *confusion matrix* yang sangat bersih dengan sangat sedikit kesalahan (misalnya, hanya 1 sampel “Samar” salah diprediksi sebagai “Gelap” dan 1 “Terang” sebagai “Gelap”). Kurva ROC untuk Model C6 menunjukkan nilai 1.00 untuk ketiga kelas, mengindikasikan model ini berhasil mencapai kemampuan klasifikasi yang sempurna pada data uji skenario ini.

4.3 Pembahasan Umum Model Formulasi Matematis

Pada tahap ini, dilakukan penjelasan lebih mendalam mengenai formulasi matematis yang digunakan dalam model *Multinomial Logistic Regression* (MLR) untuk memprediksi kategori visabilitas *hilal*. Penjelasan ini diperlukan untuk memberikan gambaran yang lebih jelas tentang bagaimana model melakukan proses pengambilan keputusan berdasarkan variabel-variabel astronomis yang telah direkayasa pada tahap sebelumnya.

Untuk memberikan ilustrasi konkret bagaimana model menghasilkan keputusan prediksi, penelitian ini menyajikan contoh perhitungan menggunakan satu baris data hasil rekayasa fitur. Contoh ini menggambarkan bagaimana nilai astronomis yang diperoleh pada tahap sebelumnya diproses menggunakan koefisien model terbaik (Model C6) untuk menghasilkan nilai logit serta probabilitas masing-masing kelas.

Contoh ini dipilih agar pembaca dapat memahami hubungan langsung antara:

1. Nilai fitur astronomis yang terukur,
2. Koefisien model, dan
3. Probabilitas akhir yang digunakan untuk memutuskan kategori *hilal*.

Dengan demikian, perhitungan ini bukan sekadar proses matematis, tetapi juga berfungsi sebagai interpretasi ilmiah mengenai perilaku model.

Koefisien yang digunakan merupakan koefisien Model C6, yang merupakan model dengan performa terbaik dari skenario 80:20 dan parameter $C = 100$. Sebagai contoh pemodelan diambil nilai berdasarkan interpretasi dari observasi langsung di lokasi Malang Raya pada tanggal 28 Februari 2025 dengan waktu 17:05:02 WIB.

Variabel prediktor (X):

1. X_1 = Beda Azimuth
2. X_2 = Tinggi Matahari
3. X_3 = Tinggi Bulan
4. X_4 = Elongasi
5. X_5 = Umur Bulan

Dengan nilai masing-masing variabel adalah $(X_1, X_2, X_3, X_4, X_5) = (2, 0.1, 3.85, 4.34, 10.05)$. Sedangkan kategori respons (Y) adalah $Y \in \{Terang(1), Samar(2), Gelap(3)\}$.

4.3.1 Model Logit (Linear Predictors):

Linear predictor untuk kategori Kelas (G) “Gelap” dengan persamaan 4.8:

- $\beta_{G_0} = -17,35$
- $\beta_{G_1} = -26,16$
- $\beta_{G_2} = 3,63$
- $\beta_{G_3} = 3,42$
- $\beta_{G_4} = 12,54$

$$\eta_G = (-17,35)X_1 + (-26,16)X_2 + 3,63X_3 + 3,42X_4 + 12,54X_5$$

$$\eta_G = (-17,35(2)) + (-26,16(0,1)) + 3,63(3,85) + 3,42(4,34) + 12,54(10,05)$$

$$\eta_G = (-34,70) + (-2,61) + 13,99 + 14,84 + 126,03 \quad (4.8)$$

$$\eta_G = -34,70 - 2,61 + 13,99 + 14,84 + 126,03$$

$$\eta_G = 117,55$$

Linear predictor untuk kategori Kelas (S) “Samar” dengan persamaan 4.9:

- $\beta_{S_0} = 10,90$
- $\beta_{S_1} = 19,61$

- $\beta_{S_2} = -0,88$
- $\beta_{S_3} = 10,36$
- $\beta_{S_4} = -24,57$

$$\eta_S = 10,90X_1 + 19,61X_2 + (-0,88)X_3 + 10,36X_4 + (-24,57)X_5$$

$$\eta_S = 10,90(2) + 19,61(0,1) + (-0,88(3,85)) + 10,36(4,34) + (-24,57(10,05))$$

$$\eta_S = 21,80 + 1,96 + (-3,89) + 44,96 + (-246,92) \quad (4.9)$$

$$\eta_S = 21,80 + 1,96 - 3,89 + 44,96 - 246,92$$

$$\eta_S = -182,09$$

Linear predictor untuk kategori Kelas (T) “Terang” pada persamaan 4.10:

- $\beta_{T_0} = 6,45$
- $\beta_{T_1} = 6,56$
- $\beta_{T_2} = -2,75$
- $\beta_{T_3} = -13,78$
- $\beta_{T_4} = 12,02$

$$\eta_T = 6,45X_1 + 6,56X_2 + (-2,75)X_3 + (-13,78)X_4 + 12,02X_5$$

$$\eta_T = 6,45(2) + 6,56(0,1) + (-2,75(3,85)) + (-13,78(4,34)) + 12,02(10,05)$$

$$\eta_T = 12,90 + 0,66 + (-10,59) + (-59,80) + 120,80 \quad (4.10)$$

$$\eta_T = 12,90 + 0,66 - 10,59 - 59,80 + 120,80$$

$$\eta_T = 63,97$$

Sehingga ringkasan dari setiap linear kelas predictor adalah:

$$\eta_G = 117,55; \eta_S = -182,09; \eta_T = 63,97$$

4.3.2 Softmax (Hitung Probabilitas)

Untuk menghindari overflow numerik, kurangi nilai maksimum $\max(\eta) = \eta_G = 117,55$ dari tiap η , lalu eksponensial:

- $e^{\eta_T - \eta_G} = e^{63,97 - 117,55} = e^{-53,58} \approx 5,37 \times 10^{-24}$

- $e^{\eta_S - \eta_G} = e^{-182,09 - 117,55} = e^{-299,64} \approx 7,77 \times 10^{-131}$
- $e^{\eta_G - \eta_G} = e^{117,55 - 117,55} = e^0 \approx 1,0$

Jumlah eksponensial ter-normalisasi:

$$Z = 1 + 5,37 \times 10^{-24} + 7,77 \times 10^{-131} \approx 1,00$$

Sehingga probabilitas (*softmax*) untuk masing-masing kategori berdasarkan persamaan 4.11:

$$\begin{aligned} P(Y = \text{Samar}) &= \frac{5,37 \times 10^{-24}}{Z} \approx 5,37 \times 10^{-24} \\ P(Y = \text{Samar}) &= \frac{7,77 \times 10^{-131}}{Z} \approx 7,77 \times 10^{-131} \\ P(Y = \text{Gelap}) &= \frac{1}{Z} \approx 1,00 \end{aligned} \quad (4.11)$$

Dalam persen (pembulatan praktis):

- $\text{Terang} \approx 5,37 \times 10^{-22}\%$ (praktis 0%)
- $\text{Samar} \approx 7,77 \times 10^{-129}\%$ (praktis 0%)
- $\text{Gelap} \approx 100\%$

Karena e^{η_G} dominan, maka:

$$P_{(\text{Gelap})} = \frac{e^{117,55}}{e^{117,55} + (\text{negligible})} \approx 1,00, \quad (4.12)$$

Dan pada persamaan 4.12:

$$P(\text{Terang}) \approx 0, \quad P(\text{Samar}) \approx 0.$$

Model memberikan probabilitas hampir 100% untuk kelas Gelap pada titik data contoh ini, artinya model “memutuskan” observasi itu sangat cenderung tidak terlihat. Ini konsisten dengan tanda-tanda fisik yang di-*capture* oleh zonasi koefisien (kontribusi positif besar untuk e^{η_G} dari kombinasi elongasi/umur di koefisien). Penjelasan singkat tentang hasil dan interpretasi koefisien:

1. Koefisien positif untuk suatu fitur pada kelas k berarti kenaikan nilai fitur tersebut meningkatkan *log-odds* (dan karenanya probabilitas relatif) kelas k .

2. Koefisien negatif menurunkan *log-odds* kelas tersebut.
3. Nilai koefisien besar menandakan fitur tersebut sangat menentukan pembagian kelas (model sangat “percaya” pada fitur tersebut).
4. Pada contoh, e^{η_G} sangat positif dari fitur-fitur tertentu (kombinasi elongasi, umur, dsb.) mendorong hasil ke label Gelap.

4.4 Diskusi Hasil Eksperimen dan Temuan Kunci

Hasil eksperimen yang disajikan pada Skenario A, B, dan C memberikan wawasan mendalam terhadap dua aspek krusial: dampak rasio pembagian data (*train-test split*) dan sensitivitas model terhadap *hyperparameter* regularisasi (C).

4.4.1 Dampak Parameter Regularisasi (C) terhadap Kinerja Model

Temuan paling dominan dan konsisten di ketiga skenario (A, B, dan C) adalah korelasi positif yang kuat antara nilai parameter C dengan kinerja model. Parameter C berfungsi sebagai kebalikan dari kekuatan regularisasi; nilai C yang rendah menerapkan regularisasi yang kuat, sementara nilai C yang tinggi menerapkan regularisasi yang lemah.

1. Pada $C=0.001$ (Model A1, B1, C1)

Ketiga model menunjukkan kinerja terendah di skenarionya masing-masing, dengan *F1-Score* berkisar antara 66.06% hingga 68.33%. *Confusion Matrix* pada model-model ini (misalnya A1, B1, dan C1) menunjukkan tingkat kebingungan yang tinggi, terutama dalam mengklasifikasikan kelas “Samar” (Kategori 2), yang banyak salah diprediksi sebagai “Terang” (Kategori 1). Kurva ROC, seperti pada Model A1 dan B1, juga menunjukkan area terendah untuk Kelas 2 (ROC 0.81). Hal ini secara jelas mengindikasikan kondisi *underfitting*, di mana model terlalu

sederhana (terlalu teregulasi) sehingga gagal menangkap pola kompleks dalam data.

2. Seiring Peningkatan C (0.01 hingga 1)

Terjadi peningkatan kinerja yang drastis dan konsisten. Pada $C=0.1$ (Model A3, B3, C3), *F1-Score* sudah melonjak ke rentang 90-95%. Pada $C=1$ (Model A4, B4, C4), kinerja semakin solid direntang 97-99% , dan kurva ROC mulai menunjukkan kemampuan pemisahan kelas yang nyaris sempurna (ROC 0.99-1.00).

3. Pada $C=1$ dan $C=100$ (Model A4-A6, B4-B6, C4-C6)

Model mencapai kinerja puncaknya. Dengan *F1-Score* yang konsisten di atas 95% hingga 97%, model ini menunjukkan kemampuan klasifikasi yang sangat tinggi. *Confusion matrix* untuk model-model ini (A6, B6, dan C6) nyaris ideal, dengan jumlah kesalahan prediksi yang sangat minim. Kurva ROC untuk ketiga kelas pada konfigurasi ini secara konsisten mencapai nilai sempurna 1.00, menandakan kemampuan diskriminatif yang absolut pada data uji.

4.4.2 Dampak Rasio Pembagian Data (Skenario A vs. B vs. C)

Eksperimen ini dirancang untuk menguji robustisitas (ketahanan) model terhadap variasi jumlah data latih. Skenario A (60:40) menggunakan data latih paling sedikit (1326 sampel), sementara Skenario C (80:20) menggunakan data latih terbanyak (1768 sampel). Menariknya, kinerja puncak model (pada $C=100$) hampir identik di ketiga skenario: Model A6 (60:40) mencapai *F1-Score* 97.29%, Model B6 (70:30) mencapai 97.89%, dan Model C6 (80:20) mencapai 97.96%.

4.4.3 Pembahasan Hasil

Dari rangkaian hasil eksperimen ini, dapat ditarik dua kesimpulan utama. Model *Multinomial Logistic Regression* terbukti sangat sensitif terhadap parameter regularisasi C . Pengaturan C yang terlalu rendah (regularisasi kuat) menyebabkan *underfitting* yang parah. Pengaturan C yang tinggi ($C \geq 10$) memungkinkan model mencapai potensi maksimalnya.

Model ini menunjukkan robustitas yang sangat tinggi terhadap rasio pembagian data. Fakta bahwa Model A6 (hanya dengan 60% data latih) mampu mencapai kinerja setara (bahkan sedikit lebih unggul) dari Model C6 (dengan 80% data latih) mengindikasikan bahwa pola dalam dataset ini sangat jelas dan dapat dipelajari dengan efektif oleh algoritma MLR.

Secara keseluruhan, Skenario C dengan Model C6 ($C=100$) yang sejalan dengan alur implementasi sebagai konfigurasi optimal yang menghasilkan kinerja nyaris sempurna (*F1-Score* 98.87%, Akurasi, Presisi, dan *Recall* >98%, serta ROC 1.00 untuk semua kelas), memvalidasi metode ini sebagai solusi yang sangat akurat untuk klasifikasi visabilitas *hilal*.

4.5 Validasi Eksternal

Setelah mengevaluasi performa replikasi internal (seberapa baik MLR meniru Model Pakar), langkah selanjutnya adalah melakukan validasi eksternal. Validasi ini tidak menguji akurasi prediksi (karena tidak ada data *rukyat* nyata), melainkan bertujuan untuk menguji validitas konseptual (relevansi) dari parameter yang digunakan dalam “Model Pakar” (persamaan 3.8) terhadap data yang dipublikasikan oleh lembaga otoritatif, yaitu BMKG.

4.5.1 Validasi Parameter Elongasi terhadap Data BMKG

Model Pakar yang digunakan dalam penelitian ini menetapkan elongasi sebagai salah satu parameter kunci, dengan ambang batas kritis (*thresholds*) untuk membedakan kategori. Kategori 2 ('Samar') didefinisikan pada rentang elongasi 5° hingga 12° .

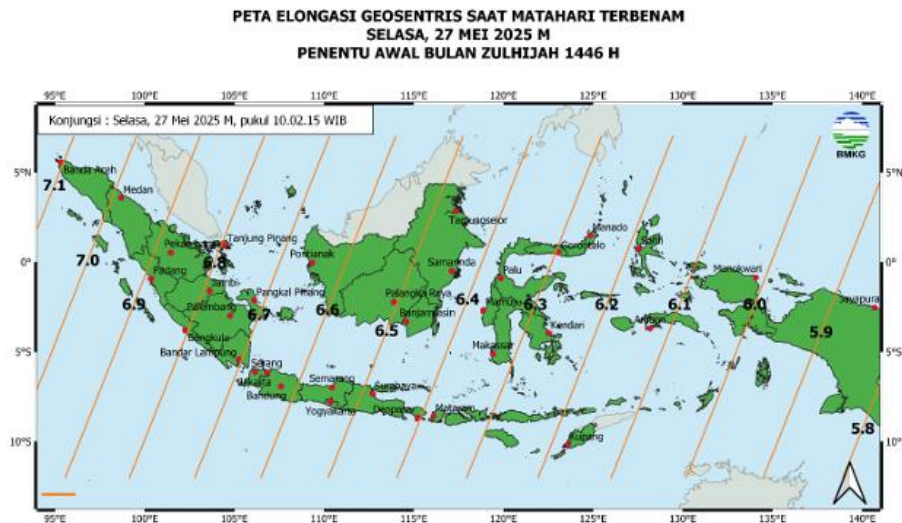
Untuk memvalidasi relevansi rentang ini, dilakukan perbandingan dengan Peta Elongasi Geosentris BMKG untuk penentuan awal bulan Ramadan 1446 H (28 Februari 2025), seperti yang terlihat pada Gambar 4.7.



Gambar 4.8 Peta Elongasi Awal Bulan Ramadan 1446 H



Gambar 4.9 Peta Elongasi Awal Bulan Syawal 1446 H



Gambar 4.10 Peta Elongasi Awal Bulan Zulhijah 1446 H
(Sumber: BMKG, *Hilal* dan Gerhana)

Analisis terhadap peta BMKG tersebut menunjukkan:

1. BMKG juga menggunakan kontur elongasi sebagai salah satu indikator utama untuk memvisualisasikan potensi visabilitas *hilal* di seluruh wilayah Indonesia.
2. Peta BMKG menunjukkan bahwa pada saat matahari terbenam 28 Februari 2025, seluruh wilayah Indonesia berada pada elongasi kritis, berkisar antara 5.8° (di Merauke) hingga 6.4° (di Sabang).
3. Nilai elongasi $5.8^\circ - 6.4^\circ$ ini selaras dan sepenuhnya masuk ke dalam rentang Kategori 2 ('Samar') (yaitu $5^\circ \leftarrow$ elongasi $\leftarrow 12^\circ$) yang telah didefinisikan dalam "Model Pakar" penelitian ini.

Validasi ini membuktikan bahwa parameter dan ambang batas yang dipilih untuk Model Pakar (khususnya untuk elongasi) bukanlah angka yang *arbitrer*, melainkan selaras dengan data astronomis nyata yang digunakan oleh lembaga nasional.

4.5.2 Uji Coba Model pada Data Proyeksi 1446 H

Sebagai pelengkap validasi, Model (MLR B6, C=100) yang telah dilatih dan divalidasi, diuji untuk mengklasifikasikan data *hisab* pada 34 titik ibukota nasional untuk tiga bulan kritis di masa depan (Tahun 1446 H): Ramadan, Syawal, dan Zulhijah. Gambar 4.16, 4.17, 4.18 menyajikan rekapitulasi hasil klasifikasi model pada 34 titik tersebut.



Gambar 4.11 Visual Prediksi Peta Perhitungan (Bulan Ramadhan 1446 H)



Gambar 4.12 Visual Prediksi Peta Perhitungan (Bulan Syawal 1446 H)



Gambar 4.13 Visual Prediksi Peta Perhitungan (Bulan Zulhijah 1446 H)

Hasil tersebut menunjukkan konsistensi yang luar biasa. Untuk ketiga bulan kritis tersebut, di mana parameter astronomisnya berada di ambang batas visabilitas (mirip dengan yang divalidasi pada peta BMKG), model MLR secara konsisten mereplikasi aturan Model Pakar dan mengklasifikasikan 100% dari 34 titik tersebut sebagai Kategori 2 ('Samar').

Validasi ganda ini menyesuaikan konseptual parameter Model Pakar dengan data elongasi BMKG, dan konsistensi hasil replikasi model MLR pada data proyeksi masa depan memperkuat keyakinan (*confidence*) bahwa model yang dibangun dalam penelitian ini adalah representasi yang valid dan relevan secara konseptual untuk menganalisis visabilitas *hilal* di Indonesia.

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan rumusan masalah, tujuan penelitian, implementasi model, serta analisis hasil yang telah dilakukan, maka penelitian ini menghasilkan beberapa kesimpulan penting sebagai berikut.

1. Pengaruh Faktor Astronomis terhadap Visabilitas *Hilal*

Penelitian ini berhasil mengidentifikasi dan membuktikan secara empiris bahwa tiga variabel astronomis utama elongasi, tinggi bulan, dan beda azimuth memiliki pengaruh signifikan terhadap visabilitas *hilal*. Analisis eksploratif (EDA), matriks korelasi, serta evaluasi koefisien regresi dan *odds ratio* pada seluruh skenario pemodelan menunjukkan bahwa elongasi dan tinggi bulan merupakan prediktor paling kuat, sementara beda azimuth berperan penting dalam menentukan keterpisahan visual bulan-matahari. Hasil ini selaras dengan teori astronomi serta konsisten dengan kecenderungan empiris BMKG dan kriteria visabilitas *hilal* regional (seperti MABIMS). Secara fisik, penelitian ini menunjukkan bahwa elongasi $< 5^\circ$ hampir selalu menghasilkan *hilal* Gelap, sementara rentang $5^\circ\text{--}12^\circ$ menghasilkan kondisi Samar, yaitu zona kritis dan tidak stabil untuk visabilitas. Tinggi bulan $> 6^\circ$ dan beda azimuth $> 10^\circ$ meningkatkan peluang *hilal* Terang. Dengan demikian, penelitian ini menegaskan kembali bahwa kombinasi ketiga parameter tersebut merupakan indikator ilmiah paling rasional untuk menilai kemungkinan terlihatnya *hilal*.

2. Pengembangan Model MLR untuk Prediksi Visabilitas *Hilal*

Penelitian ini berhasil mengembangkan dan mengimplementasikan model klasifikasi *Multinomial Logistic Regression* (MLR) sebagai pendekatan komputasional untuk memprediksi visabilitas *hilal* berdasarkan variabel elongasi, tinggi bulan, dan beda azimuth. Model ini tidak hanya mampu mengklasifikasikan *hilal* ke dalam kategori Gelap, Samar, dan Terang, tetapi juga secara matematis mampu merepresentasikan dan mengaproksimasi Model Pakar berbasis aturan IF–THEN yang selama ini digunakan dalam proses *hisab rukyat*.

Model MLR menunjukkan performa prediktif yang kuat, stabil, dan dapat direplikasi, ditunjukkan oleh F1-Score yang konsisten tinggi (95–98%), ROC ROC yang mencapai 1.00 untuk seluruh kelas pada model optimal (C=100), serta konsistensi koefisien logit yang sejalan dengan fenomena astronomis. Dengan demikian, MLR terbukti menjadi alat komputasional yang efektif dan mampu memberikan representasi matematis terhadap *rule-based system* yang kompleks.

3. Pengukuran Efektivitas dan Perbandingan dengan Metode Konvensional

Evaluasi performa model pada tiga skenario (60:40, 70:30, dan 80:20) membuktikan bahwa MLR mampu mempelajari pola visabilitas *hilal* dengan tingkat akurasi sangat tinggi, bahkan pada subset data pelatihan yang lebih kecil. Hasil ini menunjukkan robustisitas model, sekaligus mengindikasikan bahwa struktur data astronomis sudah sangat representatif dan konsisten. Dibandingkan dengan metode konvensional yang mengandalkan *hisab* kriteria astronomis dan *rukya*t lapangan pendekatan komputasional ini menawarkan kelebihan berupa objektivitas, stabilitas, replikasi matematis, dan kemampuan generalisasi. Validasi eksternal pada data BMKG (elongasi Ramadan 1446 H) juga menunjukkan bahwa

model MLR sepenuhnya konsisten dengan pola elongasi kritis yang digunakan oleh lembaga resmi, sehingga memperkuat keselarasan metode komputasional dengan praktik astronomi nasional. Dengan demikian, MLR memenuhi kebutuhan metode prediktif yang tidak hanya akurat, tetapi juga selaras secara ilmiah dengan standar nasional.

4. Rekomendasi Berbasis Data untuk Penentuan Kalender *Hijriah*

Berdasarkan hasil pemodelan, validasi internal, dan validasi eksternal, penelitian ini menghasilkan rekomendasi berbasis data untuk peningkatan akurasi penentuan awal bulan *Hijriah* di Indonesia. Model MLR versi terbaik ($C=100$) menunjukkan prediksi yang sangat konsisten pada data proyeksi *hisab* 1446 H untuk 34 titik observasi nasional, menghasilkan klasifikasi yang stabil dan selaras dengan karakter astronomis *real-time*. Temuan ini dapat dimanfaatkan oleh Kementerian Agama, BMKG, dan organisasi Islam sebagai instrumen pendukung keputusan yang lebih informatif, objektif, dan terukur. Penerapan model ini memberikan peluang besar untuk mengurangi ketidakselarasan dalam *rukyat* di berbagai lokasi, serta menjadi dasar integrasi *hisab* modern dan komputasi statistika dalam sistem kalender *Hijriah*. Selain sebagai alat klasifikasi, MLR juga dapat digunakan sebagai alat analitis untuk “membongkar” pola logika sistem berbasis aturan dan menyusun standar baru yang lebih ilmiah dalam penetapan awal bulan.

Penelitian ini membuktikan bahwa *Multinomial Logistic Regression* merupakan pendekatan ilmiah komputasional yang efektif untuk memodelkan visabilitas *hilal* berdasarkan parameter astronomis utama. Model ini tidak hanya mampu mencapai akurasi prediktif yang sangat tinggi, tetapi juga menawarkan

stabilitas, replikasi ilmiah, serta kesesuaian terhadap pola astronomis yang digunakan otoritas nasional. Oleh karena itu, MLR dapat diusulkan sebagai pendekatan komplementer dan penguat keputusan dalam penyusunan kalender *Hijriah* yang lebih akurat dan seragam.

5.2 Saran

1. Bagi Akademisi/Peneliti: Metodologi hibrida (Model Pakar → Model Aproksimasi) dapat diadopsi sebagai kerangka kerja untuk menganalisis kriteria *hisab* yang ada (Wujudul *Hilal*, Imkanur *Rukyat*). Kriteria-kriteria tersebut dapat diformulasikan sebagai "Model Pakar", kemudian dianalisis menggunakan MLR untuk memahami variabel (elongasi, tinggi, dll.) mana yang secara statistik paling berpengaruh dalam kriteria tersebut.
2. Bagi Praktisi (Lembaga *Hisab-Rukyat*): Wawasan bahwa batas antara ‘Samar’ dan ‘Gelap’ adalah yang paling *non-linier* dan sulit dimodelkan, harus menjadi perhatian. Ini mengkonfirmasi bahwa fokus dalam penentuan awal bulan seringkali berada pada zona “kritis” atau “abu-abu”, di mana perbedaan kecil pada parameter input dapat mengubah hasil kategori.

5.3 Penelitian Masa Depan

1. Penggunaan Data Observasi Nyata (*Rukyat*): Keterbatasan utama penelitian ini adalah penggunaan label sintetis dari Model Pakar. Penelitian masa depan yang paling krusial adalah mengganti Model Pakar dengan data observasi (*rukayat*) historis yang sebenarnya. Dengan melatih MLR pada data *rukayat* nyata, model akan berubah dari Model Representasi (meniru aturan) menjadi Model Prediksi Sejati (memprediksi fenomena alam).

2. Penambahan Fitur (Variabel) Meteorologi: Untuk meningkatkan akurasi prediksi (jika menggunakan data nyata), penelitian selanjutnya harus menyertakan variabel independen (X) tambahan yang sangat penting, seperti data meteorologi (tutupan awan, kelembapan, kecerahan langit, polusi cahaya) pada lokasi dan waktu pengamatan.
3. Perbandingan Algoritma *Non-Linier*: Penelitian ini menunjukkan bahwa aturan visabilitas bersifat *non-linier* (MLR *linier* butuh C tinggi). Penelitian masa depan dapat membandingkan performa MLR dengan algoritma *machine learning non-linier* murni (seperti *Random Forest*, *Gradient Boosting (XGBoost)*, atau *Neural Networks*) untuk melihat apakah model-model tersebut dapat mereplikasi Model Pakar dengan akurasi 100%, atau (jika menggunakan data nyata) menghasilkan prediksi yang lebih akurat.

DAFTAR PUSTAKA

- Al-Rajab, M., Loucif, S., & Al Risheh, Y. (2023). Predicting new crescent moon visibility applying machine learning algorithms. *Scientific Reports*, 13(1), 6674. <https://doi.org/10.1038/s41598-023-32807-x>
- Allawi, Z. T. (2022). A Pattern-Recognizer Artificial Neural Network for the Prediction of New Crescent Visibility in Iraq. *Computation*, 10(10), 186. <https://doi.org/10.3390/computation10100186>
- Badescu, V. (2008). Modelling Solar radiation at the earth surface. In *Springer*. [c:/pdflib/00017686.pdf](https://pdflib/00017686.pdf)
- Battineni, G., Chintalapudi, N., & Amenta, F. (2019). Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). *Informatics in Medicine Unlocked*, 16, 100200. <https://doi.org/10.1016/j.imu.2019.100200>
- BMKG. (2016). *Tanda Waktu Nasional*. Bmkg.Go.Id. <https://www.bmkg.go.id/tanda-waktu>
- Das, A., & Geisler, W. S. (2021). A method to integrate and classify normal distributions. *Journal of Vision*, 21(10), 1–17. <https://doi.org/10.1167/JOV.21.10.1>
- Dimara, E., Bezerianos, A., & Dragicevic, P. (2017). The Attraction Effect in Information Visualization. In *IEEE Transactions on Visualization and Computer Graphics* (Vol. 23, Issue 1). <https://doi.org/10.1109/TVCG.2016.2598594>
- Djamaluddin, T. (2000). New Crescent Moon Visibility in Indonesia. *Majalah LAPAN*, 2(October 2000). [http://files/1206/Djamaluddin_2000_New Crescent Moon Visibility in Indonesia.pdf](http://files/1206/Djamaluddin_2000_New_Crescent_Moon_Visibility_in_Indonesia.pdf) https://www.researchgate.net/publication/252407539_New_Crescent_Moon_Visibility_in_Indonesia
- Dua, R., Ghotra, M. S., & Pentreath, N. (2017). *Machine Learning with Spark - Second Edition*.
- Faid, M. S., Mohd Nawawi, M. S. A., & Mohd Saadon, M. H. (2023). Analysis tool for lunar crescent visibility criterion based on integrated lunar crescent database. *Astronomy and Computing*, 45, 100752. <https://doi.org/10.1016/j.ascom.2023.100752>
- Fanani, T. H. (2012). Metode Penentuan Waktu Sholat Di Masjid-Masjid Kabupaten Malang. In *Jurisdictie*. <https://doi.org/10.18860/j.v0i0.2171>
- Gharaybeh, M. (2025). Jurisprudential Reliance on Astronomical Calculations in Determining the Beginnings of the Hijri month. In *Springer Proceedings in Physics: Vol. 420 SPPHY* (pp. 160–177). https://doi.org/10.1007/978-981-96-3276-3_13
- Hadi, D. K., Song, Z., Rahmadya, B., Kozume, S., Sumiya, T., Sun, R., Takeda, S., & Wang, X. (2024). Indoor Area Estimation System Using RSSI-

- Measuring Handheld Reader Utilizing Directional Reference RFID Tags and Machine Learning. *IEEE Access*, 12, 157872–157887. <https://doi.org/10.1109/ACCESS.2024.3486792>
- Hasan Faadillah, M., Kurniasari, M., Asmoro, C. P., Ikbāl Arifyanto, M., Mahasenaputra, Herdiwijaya, D., Irfan Hakim, M., Hidayat, T., Jaelani, A. T., & Mulki, F. A. M. (2024). Workshop on Young Crescent Moon: Visibility Limit and Digital Image Contrast. *Journal of Physics: Conference Series*, 2773(1), 012025. <https://doi.org/10.1088/1742-6596/2773/1/012025>
- Haslebacher, C., Thomas, N., & Bickel, V. T. (2024). LineaMapper: A *deep learning*-powered tool for mapping linear surface features on Europa. *Icarus*, 410, 115722. <https://doi.org/10.1016/j.icarus.2023.115722>
- Hoayek, A., Khalifeh, M., Hamie, H., El-Ghoul, B., & Zogheib, R. (2023). Prediction of geopolymer pumpability and setting time for well zonal isolation - Using machine learning and statistical based models. *Heliyon*, 9(7), e17925. <https://doi.org/10.1016/j.heliyon.2023.e17925>
- Hoffman, R. E. (2003). Observing the new Moon. *Monthly Notices of the Royal Astronomical Society*, 340(3), 1039–1051. <https://doi.org/10.1046/j.1365-8711.2003.06382.x>
- Hussain, M., & Al-Mostafa, Z. (2021). Challenges in Lunar Crescent Visibility Prediction: A Review of Atmospheric and Astronomical Factors. *Journal of Astronomical Sciences*, 43, 123–135.
- Khoeron, M. (2022). *Kemenag Mulai Gunakan Kriteria Baru Hilal Awal Bulan Hijriah*. Kemenag.Go.Id. <https://kemenag.go.id/nasional/kemenag-mulai-gunakan-kriteria-baru-hilal-awal-bulan-Hijriah-vuiqwb>
- Koto, I., Hati, L. P., Manurung, A. S., & Siregar, A. S. (2024). Islamic Holy Days: The Contention of *Rukyatul Hillal* and *Hisab* Hakiki Wujudul *Hilal* Disputes for Muslims in Indonesia. *Pharos Journal of Theology*, 105(2), 1–14. <https://doi.org/10.46222/pharosjot.105.210>
- Li, J., Cao, Y., Wang, Y., & Xiao, H. (2017). Online Learning Algorithms for Double-Weighted Least Squares Twin Bounded Support Vector Machines. *Neural Processing Letters*, 45(1), 319–339. <https://doi.org/10.1007/s11063-016-9527-9>
- Loucif, S., Al-Rajab, M., Abu Zitar, R., & Rezk, M. (2024). Toward a globally lunar calendar: a machine learning-driven approach for crescent moon visibility prediction. *Journal of Big Data*, 11(1), 114. <https://doi.org/10.1186/s40537-024-00979-6>
- Ma, M., Li, B., Chen, S., Lu, T., Lu, P., Lu, Y., & Jin, Q. (2022). Global estimates of lunar surface chemistry derived from LRO diviner data. *Icarus*, 371, 114697. <https://doi.org/10.1016/j.icarus.2021.114697>
- Mohd Nawawi, M. S. A., Faid, M. S., Saadon, M. H. M., Wahab, R. A., & Ahmad, N. (2024). Hijri Month Determination in Southeast Asia: An Illustration Between Religion, Science, and Cultural Background. *Heliyon*, 10(20), e38668. <https://doi.org/10.1016/j.heliyon.2024.e38668>

- Mufid, A., & Djamaluddin, T. (2023). The implementation of new minister of religion of Brunei, Indonesia, Malaysia, and Singapore criteria towards the Hijri calendar unification. *HTS Teologiese Studies / Theological Studies*, 79(1), 1–8. <https://doi.org/10.4102/hts.v79i1.8774>
- Munir, B. (2016). Analisis hasil pengamatan *hilal* Badan Meteorologi Klimatologi dan Geofisika (BMKG) Pusat pada tahun 2010 M–2015 M. In *Skripsi S1, UIN Walisongo*. <https://eprints.walisongo.ac.id/id/eprint/5757/1/122111039.pdf>
- Muztaba, R., Malasan, H. L., & Djamal, M. (2023). Deep learning for crescent detection and recognition: Implementation of Mask R-CNN to the observational Lunar dataset collected with the Robotic Lunar Telescope System. *Astronomy and Computing*, 45, 100757. <https://doi.org/10.1016/j.ascom.2023.100757>
- Odeh, M. S. (2004). New criterion for lunar crescent visibility. *Experimental Astronomy*, 18(1–3), 39–64. <https://doi.org/10.1007/s10686-005-9002-5>
- Paliling, Y., Fathurahman, M., & Wahyuningsih, S. (2023). Multinomial Logistic Regression To Model The Combination Of Phdi Status And Hdi Status Of Districts/Cities In Kalimantan Island. *Jurnal Matematika, Statistika Dan Komputasi*, 19(3), 460–472. <https://doi.org/10.20956/j.v19i3.22299>
- Pedregosa, F., Weiss, R., Brucher, M., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(January), 2825–2830. <http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html> <http://arxiv.org/abs/1201.0490>
- Pinedo, P., Keller, K., Schatte, M., Velez, J., & Grandin, T. (2025). Association between the lunar cycle and pregnancy at first artificial insemination of Holstein cows. *JDS Communications*, 6(3), 400–405. <https://doi.org/10.3168/jdsc.2024-0722>
- Pramudya, Y., Dhimas Prayoga Jullyantama, A. R., Khairul Ardi, M., Ayu Wandira, C., & Muchlas. (2024). Observational Technique Development and Instruments Status in Observatorium Universitas Ahmad Dahlan. *Journal of Physics: Conference Series*, 2773(1), 0–6. <https://doi.org/10.1088/1742-6596/2773/1/012014>
- Purba, A. C., & Handhayani, T. (2024). Perbandingan Algoritma K-Means, Affinity Clustering, Dan Minibatch K-Means Untuk Analisis Segmentasi Pasar. *Jurnal Ilmiah Komputer Dan Informatika*, 13(1), 54–63.
- Rahmi, N., Rizki, M., Fajriati, N. A., Nurlaila, Efendi, F., & Siregar, L. S. B. (2025). Contestation of the Determination of ‘Idul Adha and its Implementation According to Muhammadiyah and Nahdatul ‘Ulama. *Nurani*, 25(1), 262–279. <https://doi.org/10.19109/nurani.v25i1.26563>
- Shalamov, V., Efimova, V., & Filchenkov, A. (2022). Faster Hyperparameter Optimization via Finding Minimal Regions in Random Forest Regressor. *Procedia Computer Science*, 212(C), 378–386.

- <https://doi.org/10.1016/j.procs.2022.11.022>
- Sharma, A., Kaur, S., Memon, N., Jainul Fathima, A., Ray, S., & Bhatt, M. W. (2021). Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis. *Neuroscience Informatics*, 1(3), 100012. <https://doi.org/10.1016/j.neuri.2021.100012>
- Sultan, A. H. (2007). First visibility of the lunar crescent: Beyond Danjon's limit. *Observatory*, 127(1196), 53–59.
- Suprihatin, H. S., Nugraha, R., Fitmawyani, W. S., & Rodliyah, A. K. (2021). Analisis BMKG untuk Penyatuan Kalender Qomariyah (Hijriyah) di Indonesia. In *Bmkg Pusat*. <https://content.bmkg.go.id/wp-content/uploads/Analisis-BMKG-untuk-Penyatuan-kalender-Qomariyah.pdf>
- Syuhada, N., Lina, H., & Bemi Sado, A. (2023). Pemetaan Posisi *Hilal* Terhadap Gunung Agung Dalam Penentuan Awal Bulan Kamariah Di Lokasi *Rukyat* Pantai Loang Baloq Mataram. *AL - AFAQ : Jurnal Ilmu Falak Dan Astronomi*, 5(1), 81–89. <https://doi.org/10.20414/afaq.v5i1.7260>
- Tabatabai, M., Wilus, D., Chen, C. K., Singh, K. P., & Wallace, T. L. (2025). Tabu Binary, Multinomial, and Ordinal Regression Models: New Machine Learning Methods for Classification. *Bioengineering*, 12(1), 2. <https://doi.org/10.3390/bioengineering12010002>
- Taruna, R. M., & Prakoso, T. A. (2017). Perkiraan Ketinggian Objek Alam Terhadap *Horizon* Untuk Evaluasi Lokasi Pengamatan *Hilal* Di Lombok. *Jurnal Penelitian Fisika Dan Aplikasinya (JPFA)*, 7(2), 115. <https://doi.org/10.26740/jpfa.v7n2.p115-122>
- Ulum, M. (2021). Fatwa Ulama Nu (Nahdlatul Ulama) Dan Muhammadiyah Jawa Timur Tentang *Hisab Rukyat*. *Jurnal Keislaman*, 1(2), 244–272. <https://doi.org/10.54298/jk.v1i2.3369>
- Umar, R., Azan, M. A. I. A., Sabri, N. H., Ngah, N., Zulkeflee, A. N., Mohamad, A. E., Juddin, A. S. S., Abdullahi, M. G., Wan, N. A. F. N., Mahiddin, N. A., Kamarudin, M. K. A., Nawawi, M. S. A. M., Hoe, L. I., Lian, C. J., Nasir, A. L. A. M., Bely, A. H., & Kamarudin, A. S. (2023). Spatial Model of Sky Brightness Magnitude at Balai Cerap KUSZA, Universiti Sultan Zainal Abidin (UNISZA). *Sains Malaysiana*, 52(10), 2943–2954. <https://doi.org/10.17576/jsm-2023-5210-16>
- Utama, J. A., Zuhudi, A. R., Prasetyo, Y., Rachman, A., Sugeng Riadi, A. R., Nandi, & Riza, L. S. (2023). Young lunar crescent detection based on video data with computer vision techniques. *Astronomy and Computing*, 44, 100731. <https://doi.org/10.1016/j.ascom.2023.100731>
- Wolfe, S., & Emami, M. R. (2024). A Data-driven Approach to the Classification of Temporary Captures in the Earth-Moon System. *IEEE Aerospace Conference Proceedings*, 1–9. <https://doi.org/10.1109/AERO58975.2024.10521049>
- Yunus, F. (2011). Aliran *Hisab* Awal Bulan Qomariyah. *Aliran Hisab Awal Bulan Qomariyah*, 1(2), 244–272. <http://repository.iainbengkulu.ac.id/11260/>

LAMPIRAN I

Dokumentasi Pengamatan Hilal





LAMPIRAN II

**Surat Tugas
Pendamping Hilal**

SURAT TUGAS

NOMOR: e.B/HM.02.04/005/KKRK/II/2025

Yang bertanda tangan di bawah ini:

Nama : MA'MURI, S.Si, M.TI
NIP : 198102072003121001
Pangkat/Golongan : Pembina/IVa
Jabatan : Kepala Stasiun Geofisika Malang
Unit Organisasi : Stasiun Geofisika Malang

Dengan ini memberikan tugas kepada:

NAMA	NIP	PANGKAT
Ahmad Zarkoni, S.Kom	197501251999031001	Pembina/IVa

Untuk melaksanakan :
Tugas : Melaksanakan pendampingan Observasi Kegiatan
Pengamatan Hilal Bulan : Sya'ban, Ramadhan & Syawal

NO	NAMA	NIM	Program Study
1.	Tomy Ivan Sugiharto	240605210004	Magister Informatika fakultas Sains & Teknologi

Mahasiswa Jurusan Magister Informatika fakultas Sains & Teknologi UIN Maulana Malik Ibrahim Malang

Selama : 3 (Tiga) Bulan
Lokasi : Kantor Bupati Malang
Tanggal Mulai : 30 Januari, 28 Februari & 29 Maret 2025

Malang, 11 Februari 2025

Kepala,



Ma'muri