Prediksi Jamur: Klasifikasi beracun-dapat dimakan.

by KKN

Anggota:
Aulia Kemal Syah
Naufal Dzakirsyah
Ucok Kamal

Tujuan: Klasifikasi Jamur (Beracun / Aman).

Tantangan :

- Data kotor - Banyak null, simbol aneh ('%', '`'), dan inkonsistensi (Data acak).

- (Kritis-Keselamatan):* False Negative (memprediksi beracun sebagai aman) sangat berbahaya.

Alur kerja:

IDE & EDA

↓

Feature engineering

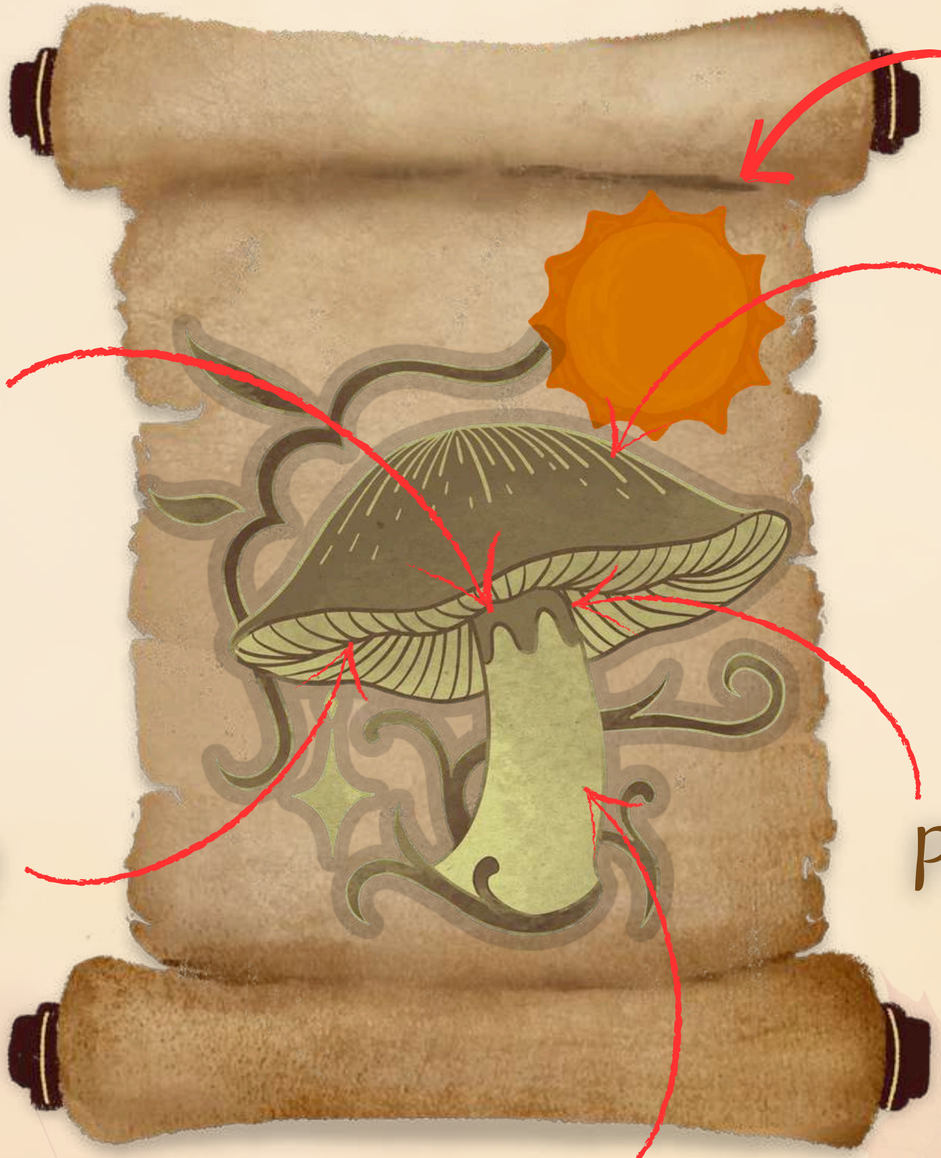
↓

Modeling

↓

Evaluasi

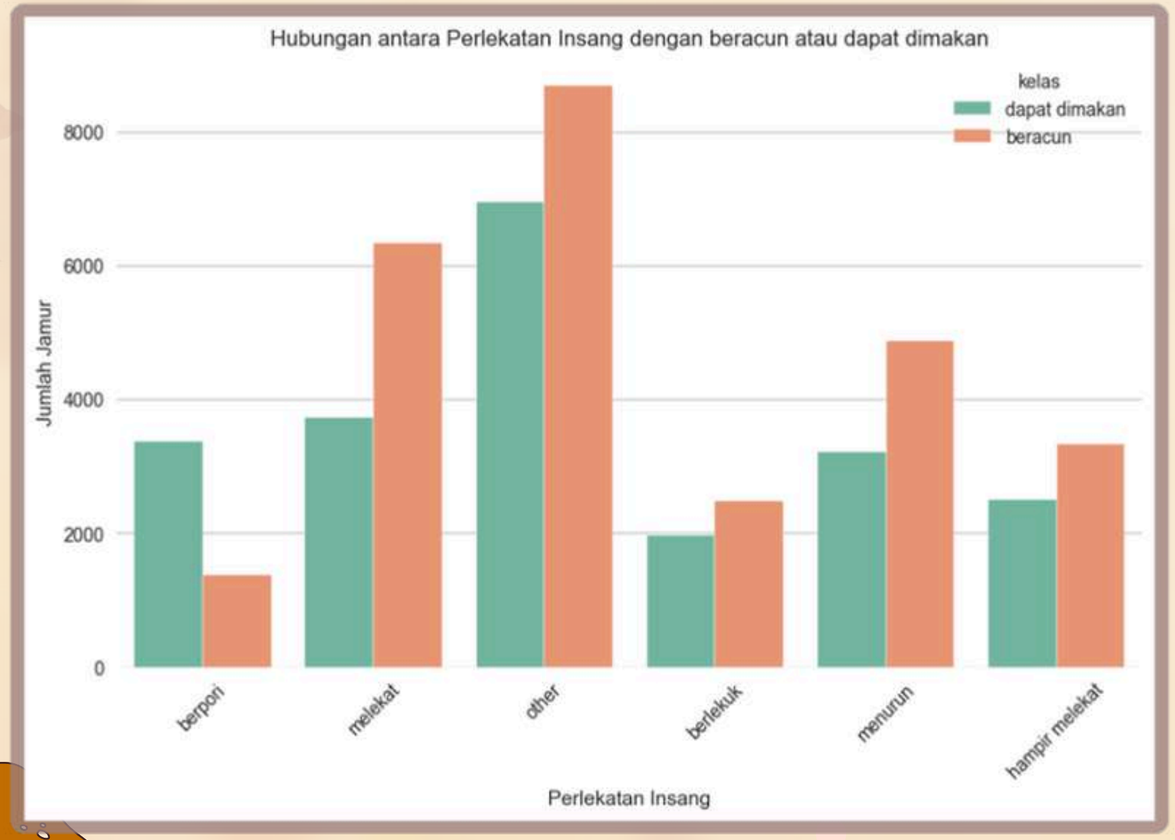
Fitur yang berpengaruh



Fitur yang berpengaruh

Warna insang

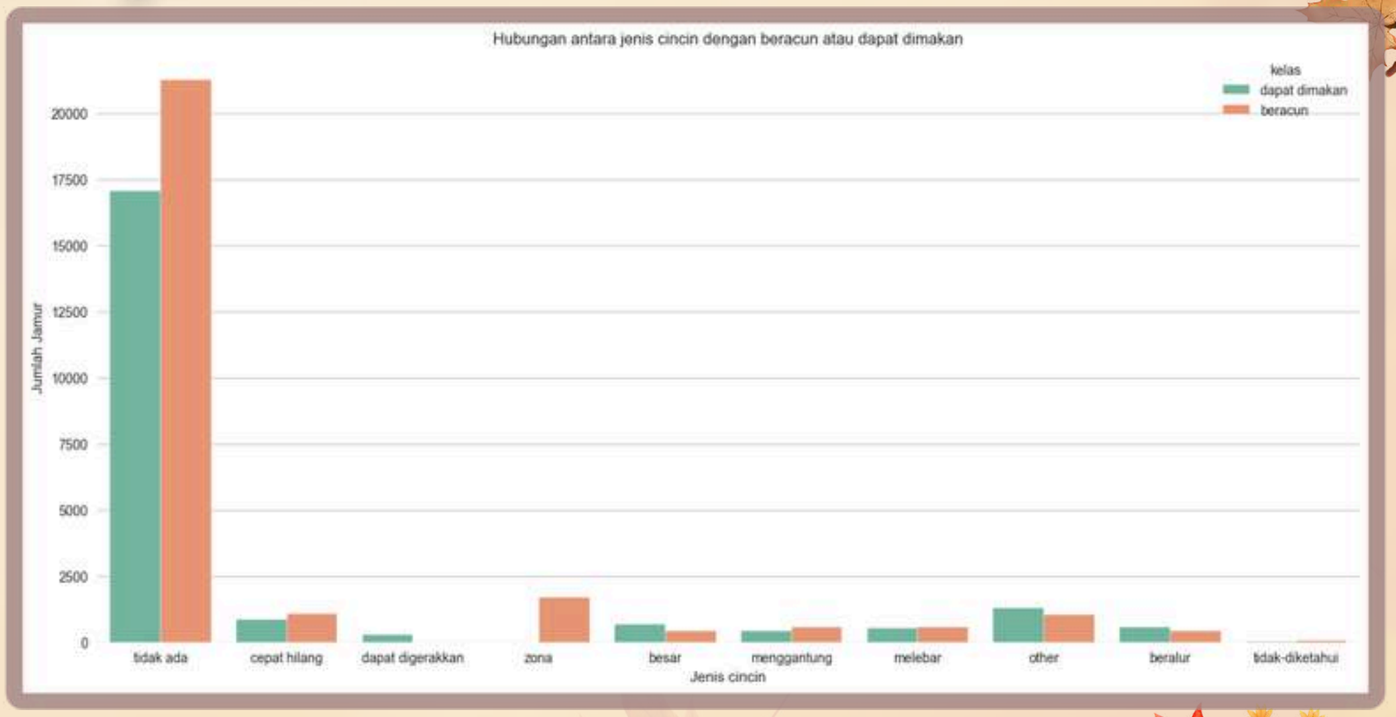
Fitur yang berpengaruh



Perlekatan insang

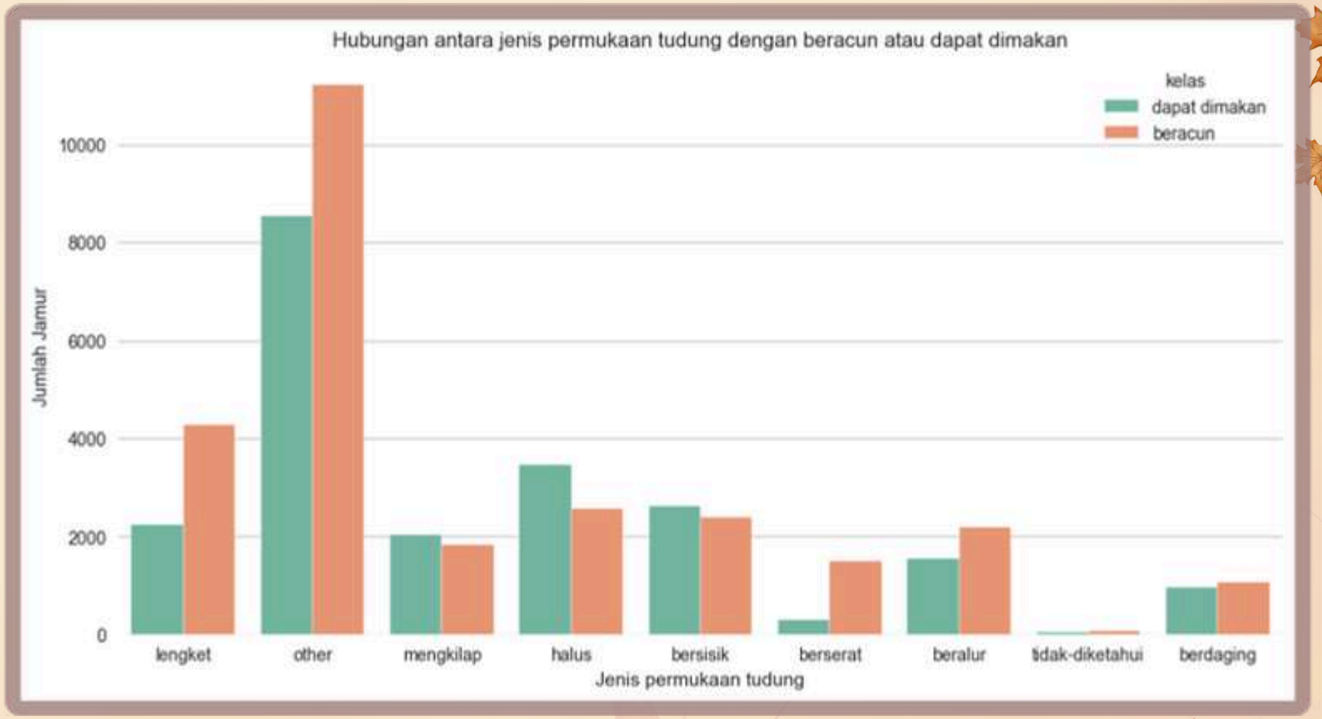
Fitur yang berpengaruh

Cincin Jamur



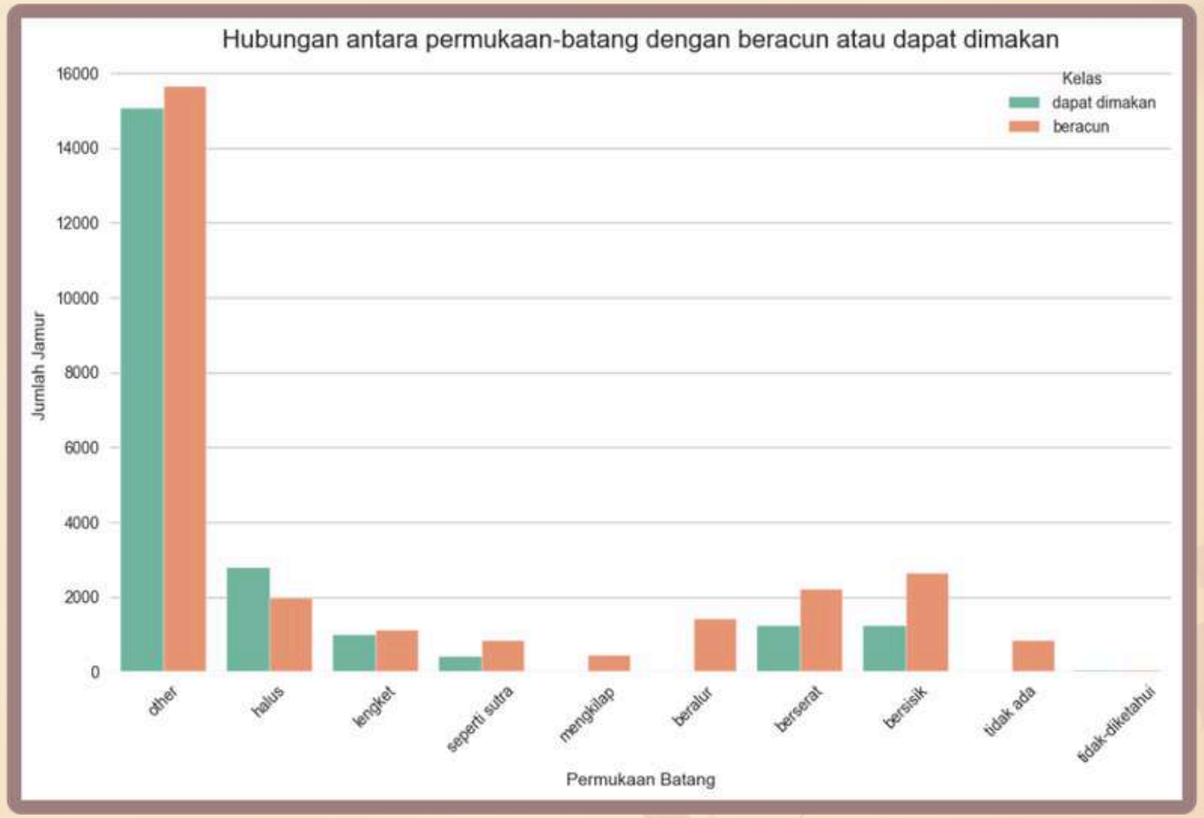
Fitur yang berpengaruh

Tudung Jamur



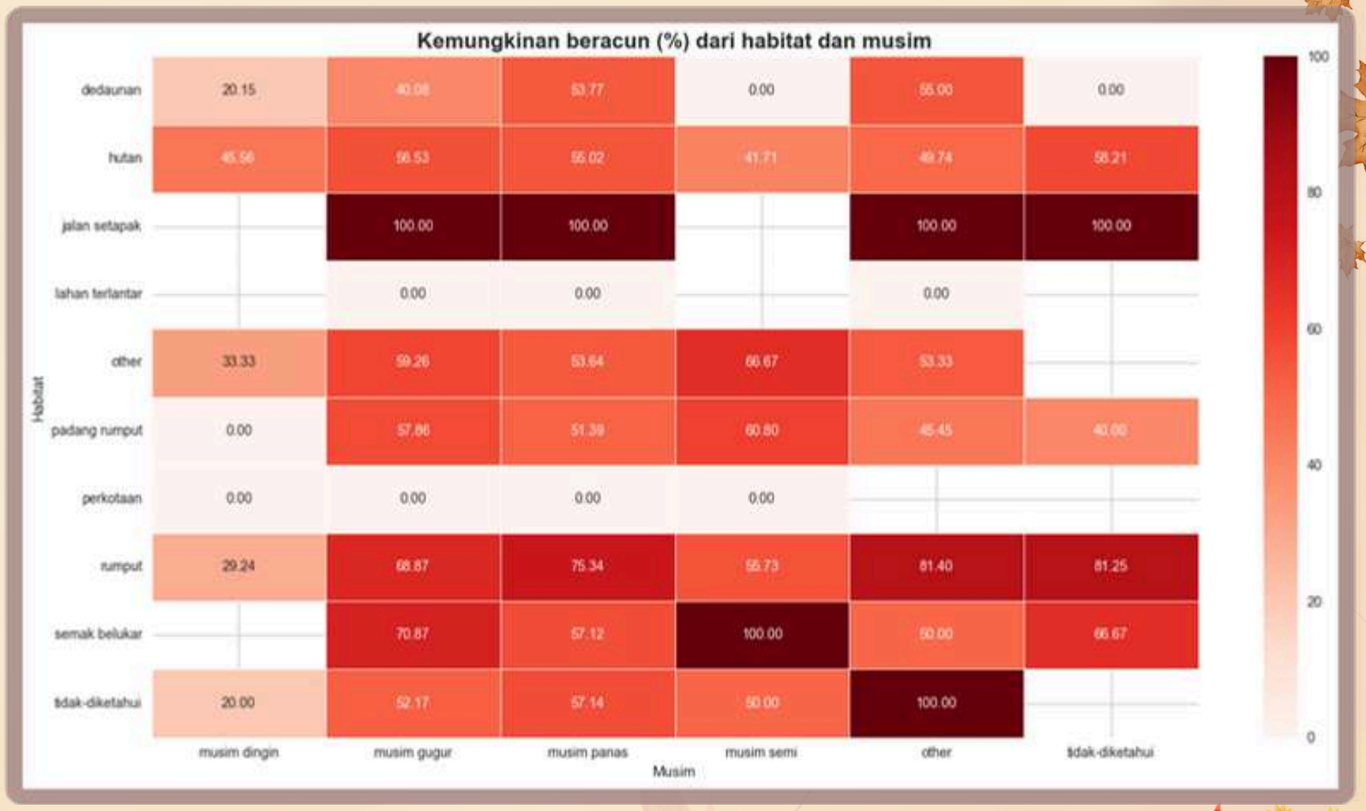
Fitur yang berpengaruh

Batang Jamur



Fitur yang berpengaruh

Musim-Habitat



Fitur yang kurang berpengaruh

- memar-atau-getah (99.0% hilang)
- jenis-selaput (93.9% hilang)
- warna-cetakan-spora (88.8% hilang)
- warna-selaput (87.1% hilang)
- akar-batang (85.4% hilang)

```
==================================================
                MISSING VALUES ANALYSIS
==================================================
memar-atau-getah        | 48357 missing (99.0%)
jenis-selaput           | 45896 missing (93.9%)
warna-cetakan-spora     | 43365 missing (88.8%)
warna-selaput           | 42541 missing (87.1%)
akar-batang             | 41735 missing (85.4%)
permukaan-batang        | 30285 missing (62.0%)
jarak-insang            | 19900 missing (40.7%)
permukaan-tudung        | 14766 missing (30.2%)
perlekatan-insang       | 7875 missing (16.1%)
jenis-cincin            | 1941 missing (4.0%)
diameter-tudung         | 1940 missing (4.0%)
bentuk-tudung           | 1572 missing (3.2%)
tinggi-batang           |  380 missing (0.8%)
lebar-batang            |  366 missing (0.7%)
```

Feature Engineering

1

Membersihkan data anomali dan mengatasi skewness (distribusi miring) menggunakan Transformasi Log.

2

Menciptakan fitur baru (interaksi, binning, agregat) untuk menemukan pola yang lebih kompleks.

3

Mengatasi NaN numerik dengan KNN Imputer dan NaN kategorikal dengan Label Khusus (-1).

4

Menggunakan LabelEncoder untuk mengubah feature kategorikal menjadi angka agar dapat dimanfaatkan model secara optimal

Model Selection

Model Performance Comparison:

	Category	Accuracy	F1-Score	Recall	Precision	Overall_Score	Training_Time	F1_Rank	Recall_Rank
LightGBM	Advanced	0.9979	0.9976	0.9968	0.9984	0.9974	15.31s	1	2
XGBoost	Advanced	0.9974	0.9971	0.9961	0.9982	0.9969	3.74s	2	3
CatBoost	Advanced	0.9966	0.9962	0.9949	0.9975	0.9959	83.96s	3	4
KNN	Baseline	0.9860	0.9843	0.9846	0.9839	0.9843	nans	4	5
Decision Tree	Baseline	0.9833	0.9813	0.9811	0.9814	0.9812	nans	5	6
Naive Bayes	Baseline	0.4610	0.6228	1.0000	0.4523	0.7396	nans	7	1
Logistic Regression	Baseline	0.6910	0.6468	0.6355	0.6584	0.6446	nans	6	7

Cross Validation:

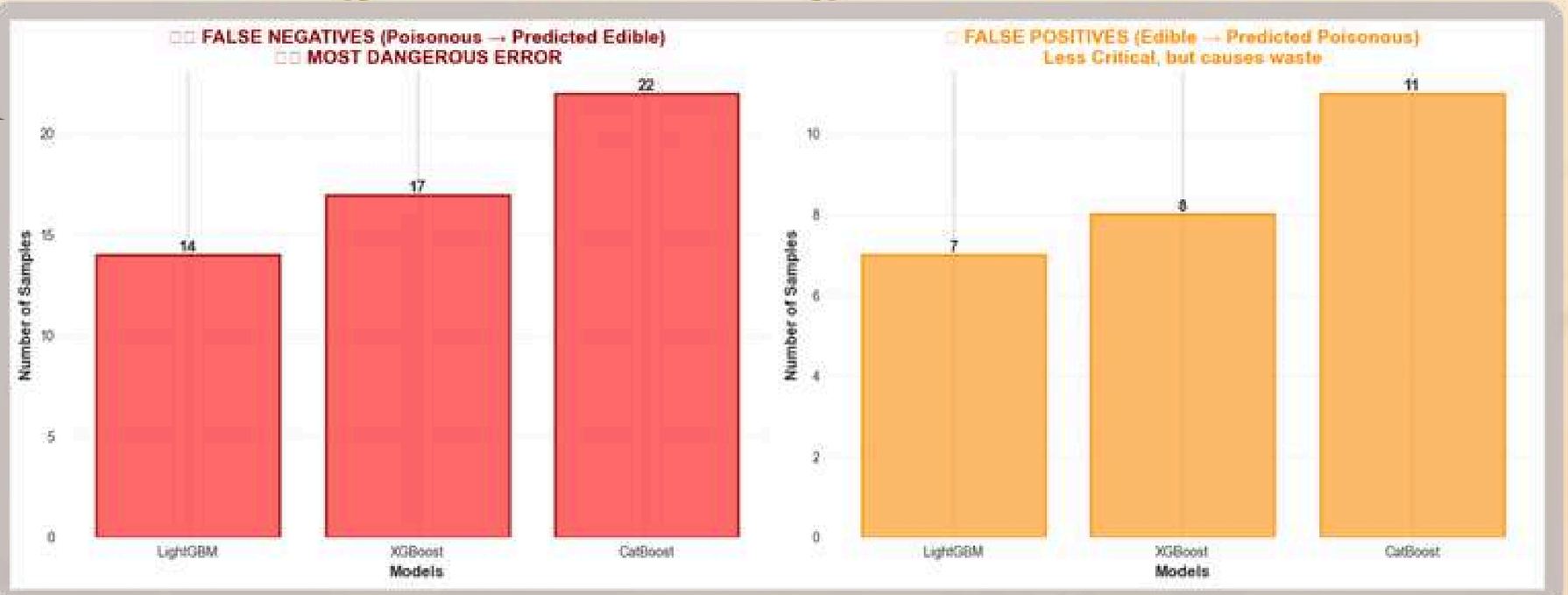
Model	Single F1	CV F1	Diff	Status
LightGBM	0.9976	0.9972	0.0004	✅ Excellent
CatBoost	0.9962	0.9962	0.0000	✅ Excellent

Pemenang: LightGBM (LGBM)
Skor Kunci: F1-Score CV 0.9972
Validasi: Unggul signifikan (p-value 0.0049) & sangat stabil (Std Dev 0.0004).

Error Analysis

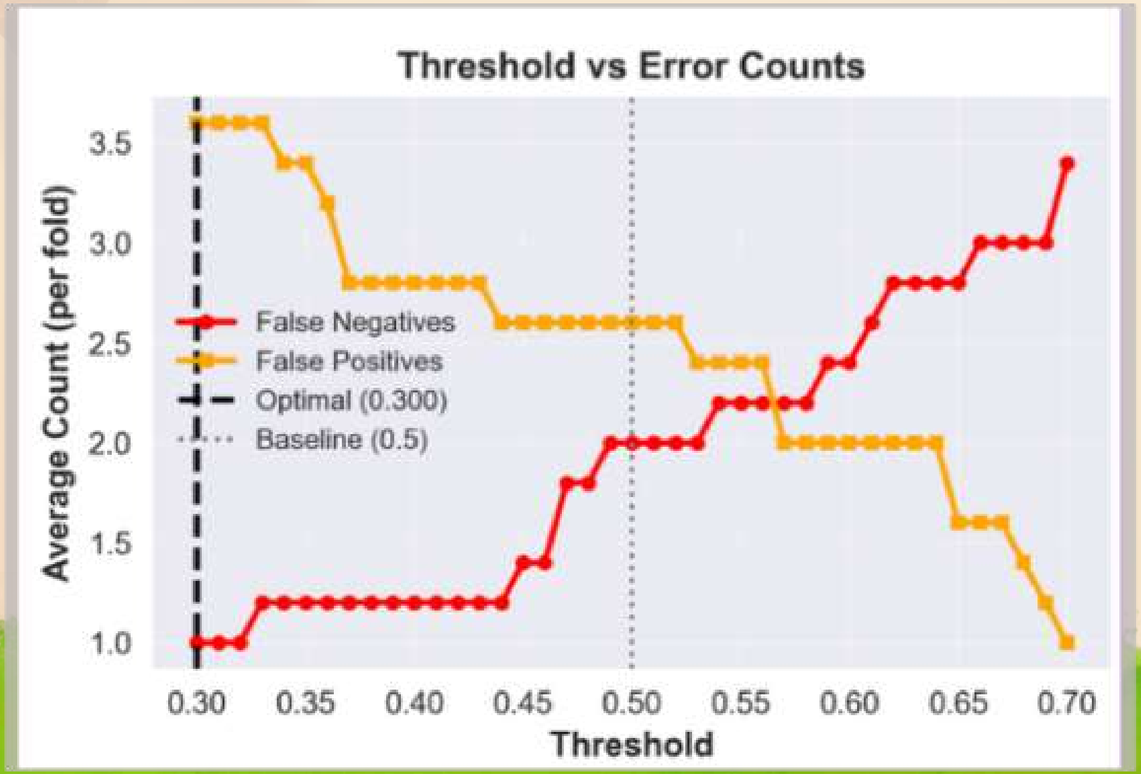
Sebelum optimalisasi threshold

```
LightGBM:
   False Negative Rate: 0.32% (14/4349)
   False Positive Rate: 0.13% (7/5422)

XGBoost:
   False Negative Rate: 0.39% (17/4349)
   False Positive Rate: 0.15% (8/5422)

CatBoost:
   False Negative Rate: 0.51% (22/4349)
   False Positive Rate: 0.20% (11/5422)
```



Error Analysis

Setelah optimalisasi threshold (0,32)




```
Step 5.2: Calculate Test Metrics (Optimal vs Baseline)

📊 TEST SET COMPARISON:
Metric              Baseline (0.5)      Optimal (0.300)     Change
----------------------------------------------------------------------
F1-Score            0.997584            0.997702            +0.000118
Recall              0.996781            0.998390            +0.001610
Precision           0.998388            0.997015            -0.001373
Accuracy            0.997851            0.997953            +0.000102
----------------------------------------------------------------------
False Negatives     14                  7                   -7
False Positives     7                   13                  +6
True Negatives      5415                5409                -6
True Positives      4335                4342                +7
```

Follow-up Decision Pre-Tuning

- Uji Coba: Tidak cocok menggunakan hyperparameter tuning
- Hasil: F1 sudah 99.7%. Risiko Overfitting lebih besar dari potensi gain.
- Keputusan: SKIP (Model default sudah sangat optimal)

```
BASELINE ANALYSIS SUMMARY
==================================================

Current Performance:
  • F1-Score: 0.99717
  • Recall: 0.99700
  • Precision: 0.99726

95% Confidence Interval:
  • F1: [0.99707, 0.99727]

Theoretical Ceiling:
  • F1: 0.80731

Potential Gain:
  • ΔF1: -0.18986 (-19.040%)

TUNING DECISION:
  □ NOT RECOMMENDED

  SKIP tuning. Focus on error analysis & feature engineering instead.
```

Follow-up Decision Pre-Tuning

```
FEATURE ENGINEERING SUMMARY
==========================================

DATASET STATISTICS:
------------------------------------------
Original Features:          38
New Features Created:       16
New Features Accepted:       0
Final Feature Count:        38
------------------------------------------
Train Samples:            48,855
Test Samples:             12,214

PERFORMANCE METRICS:
------------------------------------------
Baseline CV F1:              nan
New CV F1:                0.997471
Improvement:                 +nan
Improvement %:              +nan%
------------------------------------------
p-value:                     nan
Significance:             NO

DECISION:
------------------------------------------
Status:                   REJECTED
Reason:                   Insufficient
                          improvement

NEXT PHASE:
------------------------------------------
Sub-Bab 7.5: Final Model Training
Expected F1 Target:       ≥ 0.997471

==========================================
```

- Uji Coba: FE Ronde 2 (Menambah 16 Fitur Baru)
- Hasil: Performa CV turun & tidak stabil.
- Keputusan: FE Ronde 2 DITOLAK (Lebih banyak fitur tidak selalu lebih baik)

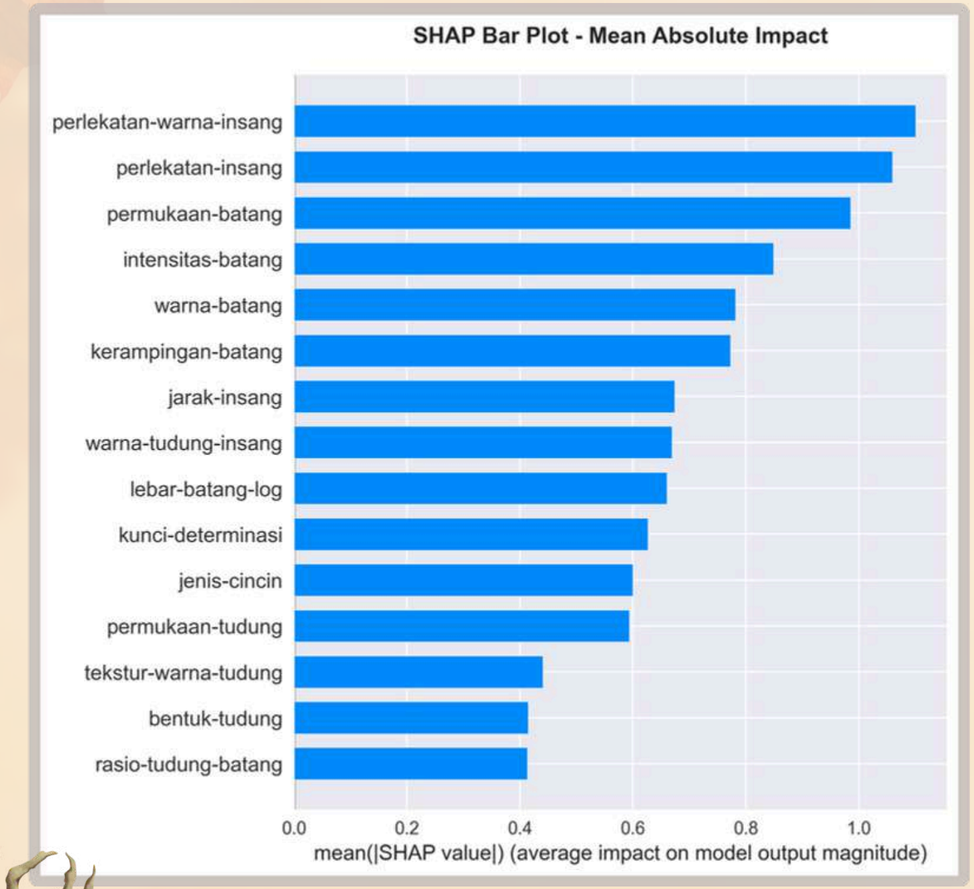
Summarize

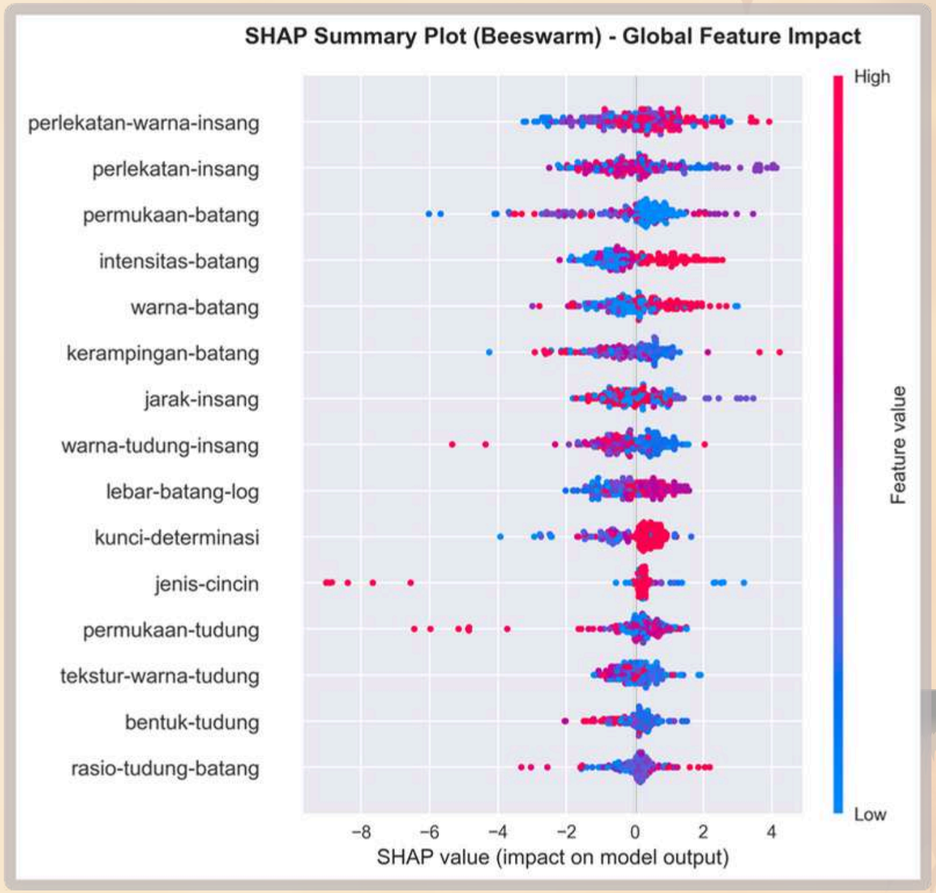
```
FINAL MODEL SUMMARY
==================================================

Model: LightGBM
Threshold: 0.300 (from 7.3)
Features: 38
Training Samples: 48,855

CV PERFORMANCE (5-Fold):
--------------------------------------------------
F1-Score:  nan (±nan)
Recall:    nan
Precision: nan

EXPECTED TEST PERFORMANCE:
--------------------------------------------------
Best Case:  nan
Expected:   nan
Worst Case: nan

IMPROVEMENTS:
--------------------------------------------------
vs 7.2: +nan (+nan%)
vs 7.4: +nan (+nan%)

OVERFITTING CHECK:
--------------------------------------------------
Train-CV Gap: +nan
Status: □ High

TEST PREDICTIONS:
--------------------------------------------------
Total: 12,214
Safe (0): 6772 (55.4%)
Poisonous (1): 5442 (44.6%)

Note: Submission file will be created in Sub-Bab 9.0
```

- Model Final: LightGBM
- Fitur Final: 38 Fitur (FE Ronde 2 ditolak)
- Kontribusi Kunci: Threshold 0.32 (Mengurangi False Negative 50%)
- Pelajaran Utama: Model yang lebih aman (safe) lebih penting daripada model yang sedikit lebih akurat (accurate).

Phase	Threshold	F1-Score	Features	Notes
7.2 Baseline	0.5	0.997584	38	Initial baseline
7.3 Threshold Opt	0.3	0.997702	38	Cost-optimized threshold
7.4 Feature Eng	0.5	NaN	38	FE rejected (p=0.70)
7.5 Final	0.3	NaN	38	Final with 38 features

Feature Importance





Thank You