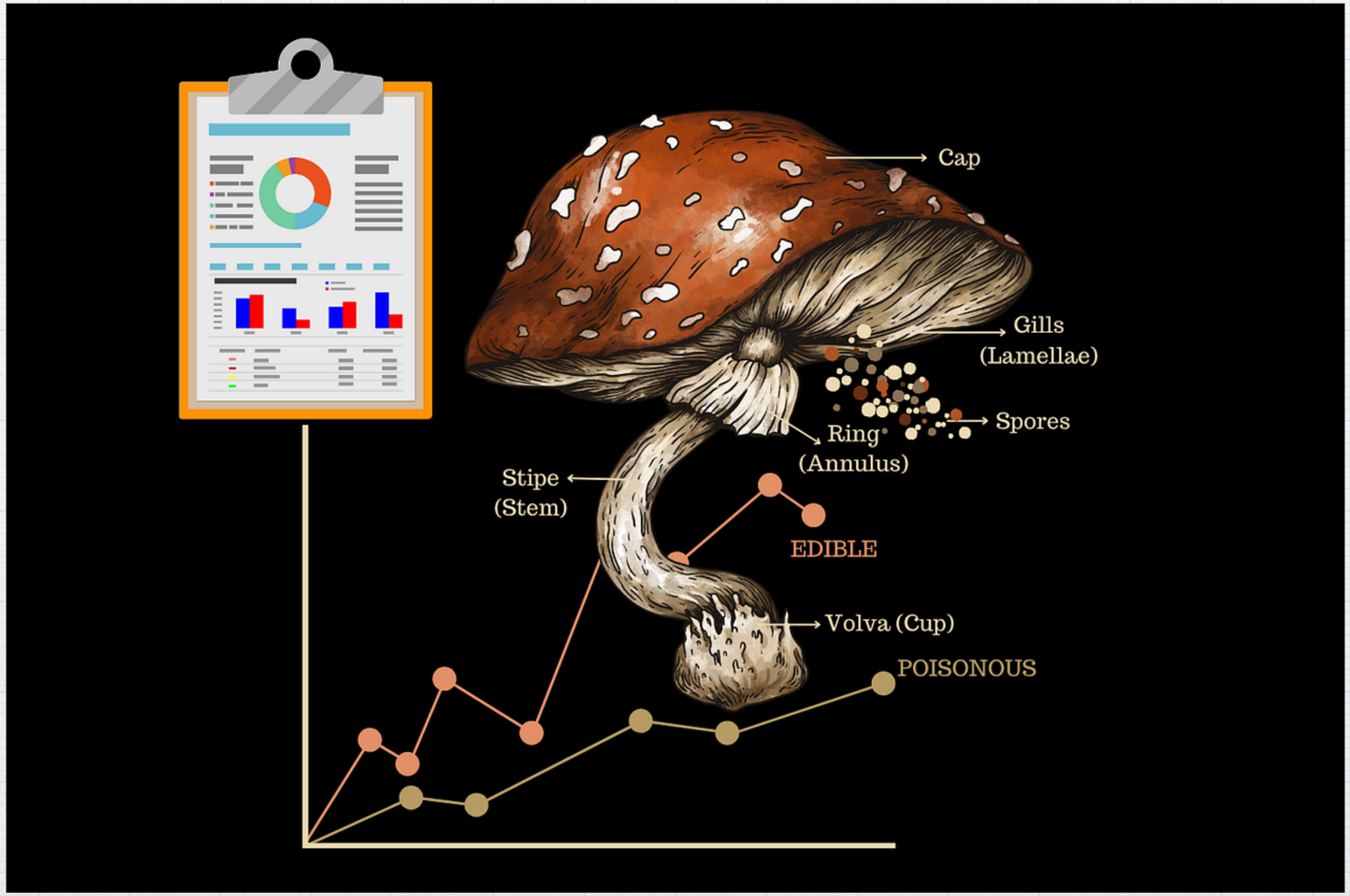
Good to eat or deadly poison?

Machine Learning Techniques for Mushroom Classification



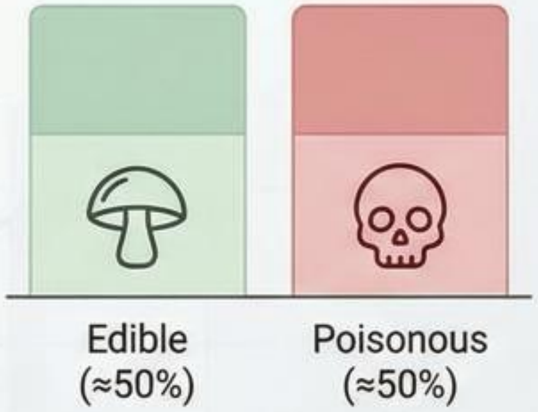
Introduction

- **The danger of mis-identifying mushrooms: edible vs poisonous varieties.**
- **Motivation: use data + ML to help classify mushrooms.**
- **Overview: dataset → modeling → evaluation → insights & recommendations.**

The Enhanced Data Set: Balanced, Categorical, and Field-Ready

Optimizing AI for Mushroom Safety through Data Structure

Dataset Balance & Size

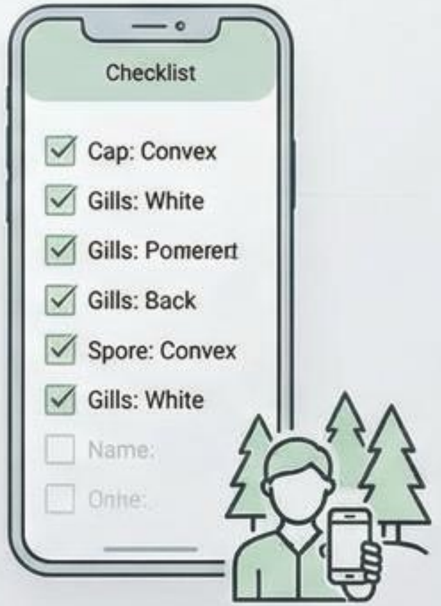


Prevents model bias towards either class

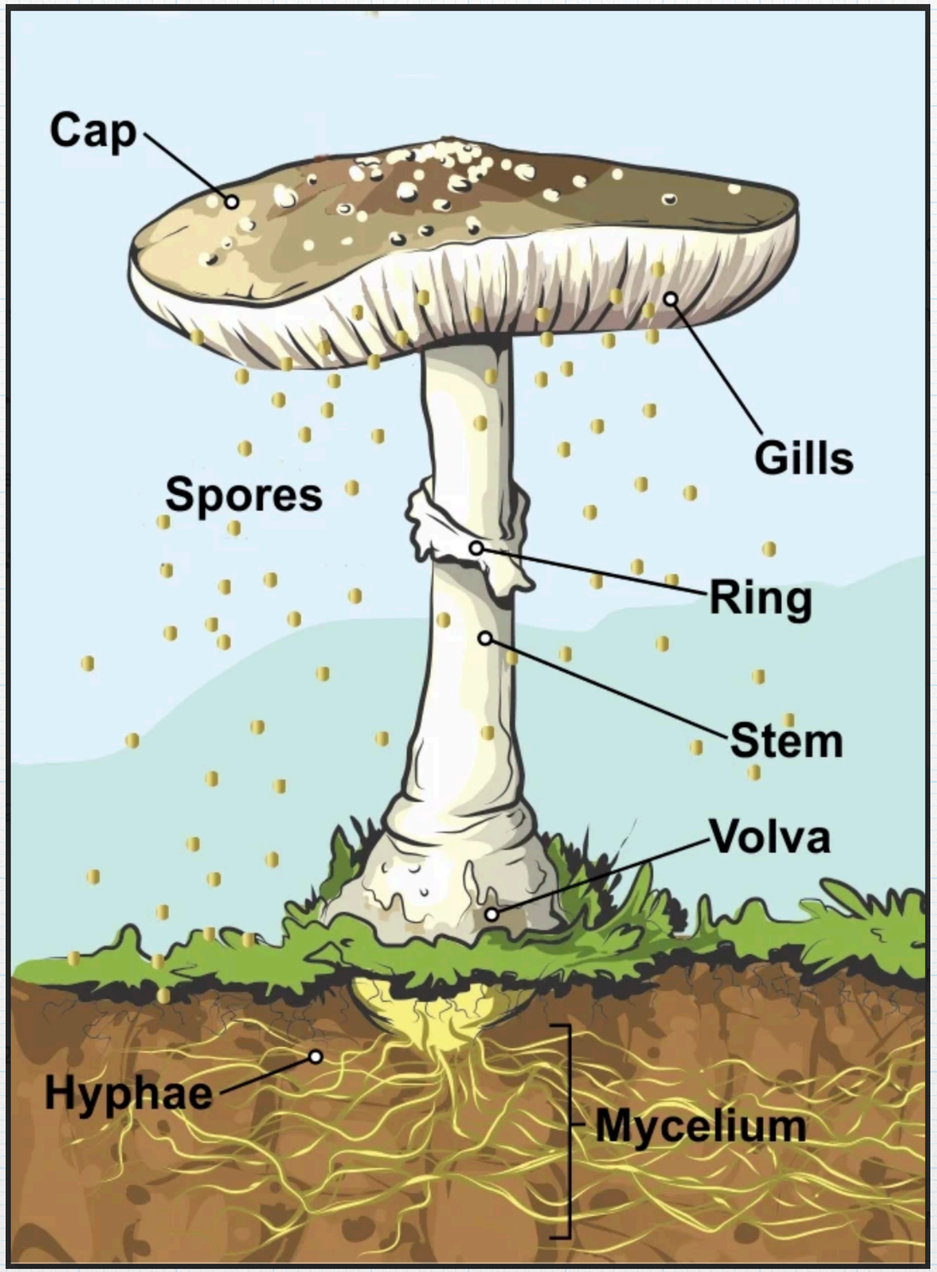
Categorical Descriptors

All 23 variables are discrete and categorical

Field Deployment Benefits



Simplifies input; No measuring tools required


Cap
Gills
Spores
Ring
Stem
Volva
Hyphae
Mycelium

Feature Analysis: Cramér's V Association Test

Primary Predictive Power (High Association)

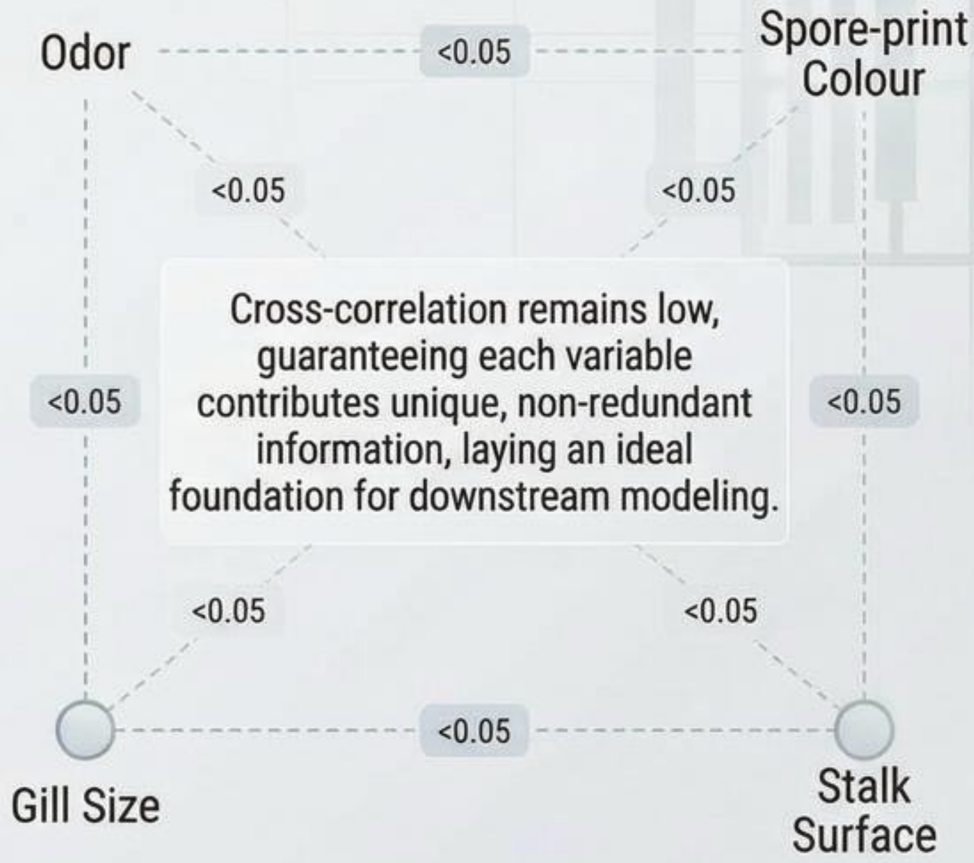
Target Label (Toxicity)

Odor carries the highest coefficient, indicating strong predictive power.

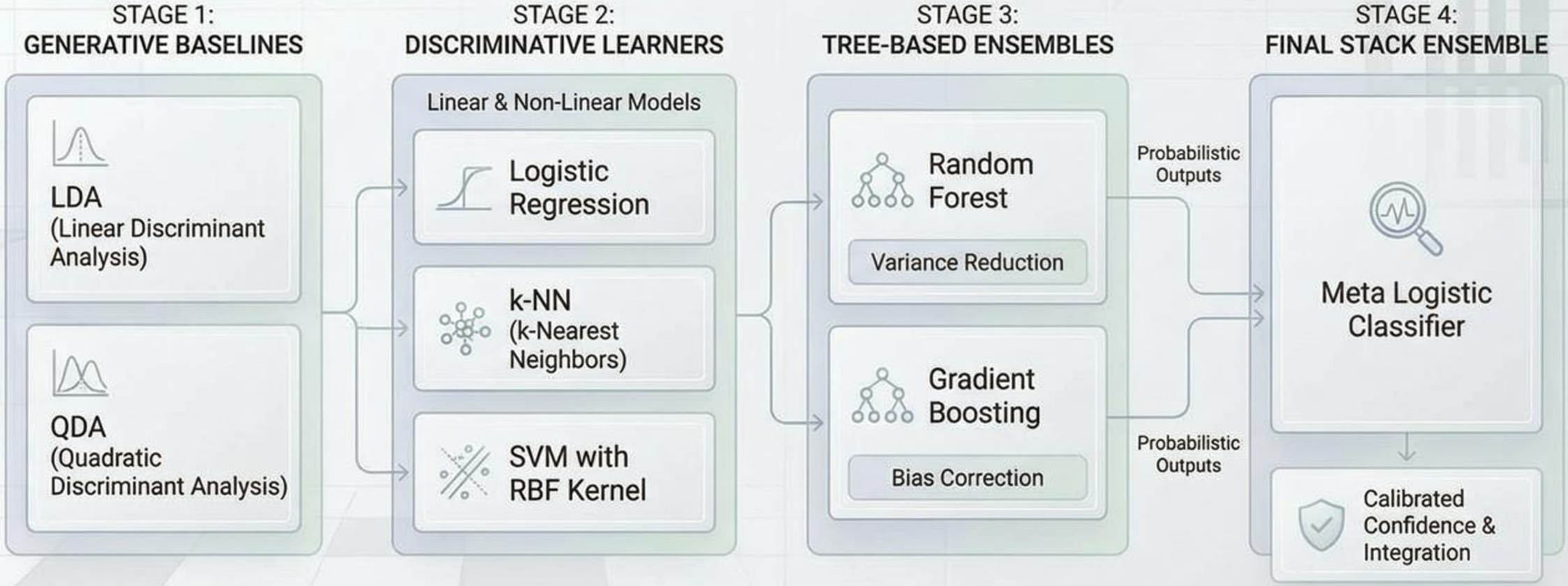
Weak Association (<0.1)

Spore-print colour, gill size, and stalk surface lag significantly, indicating weak predictive power.

Feature Interdependence (Low Cross-Correlation)



Experimental Machine Learning Pipeline: From Baselines to Ensemble Stacking



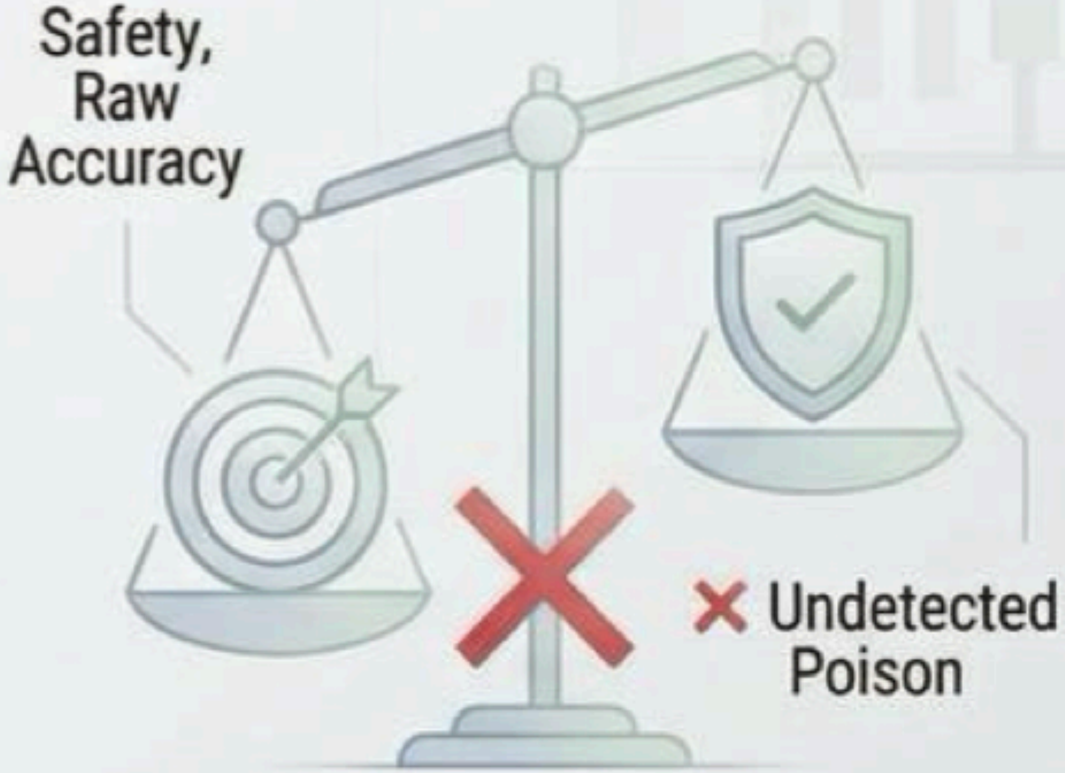
Key Performance Metrics

Accuracy	Accuracy morittification information specification	↗
ROC-AUC	ROC-AUC memic is base recont ROC-AUC	↗
F2-Score	F2-Score is oraxfully outoff probability to score	↗
False-Positive Rate	False-Positive rate is size and negatives	→
False-Negative Rate	False-negative rate is ensure False-negative rate	↗

Key Performance Metrics

Tracking comprehensive metrics to assess model reliability and risk.

Safety-First Prioritization



Safety-First Prioritization

Explicitly penalizing missed toxic cases to ensure safety dominates final selection.

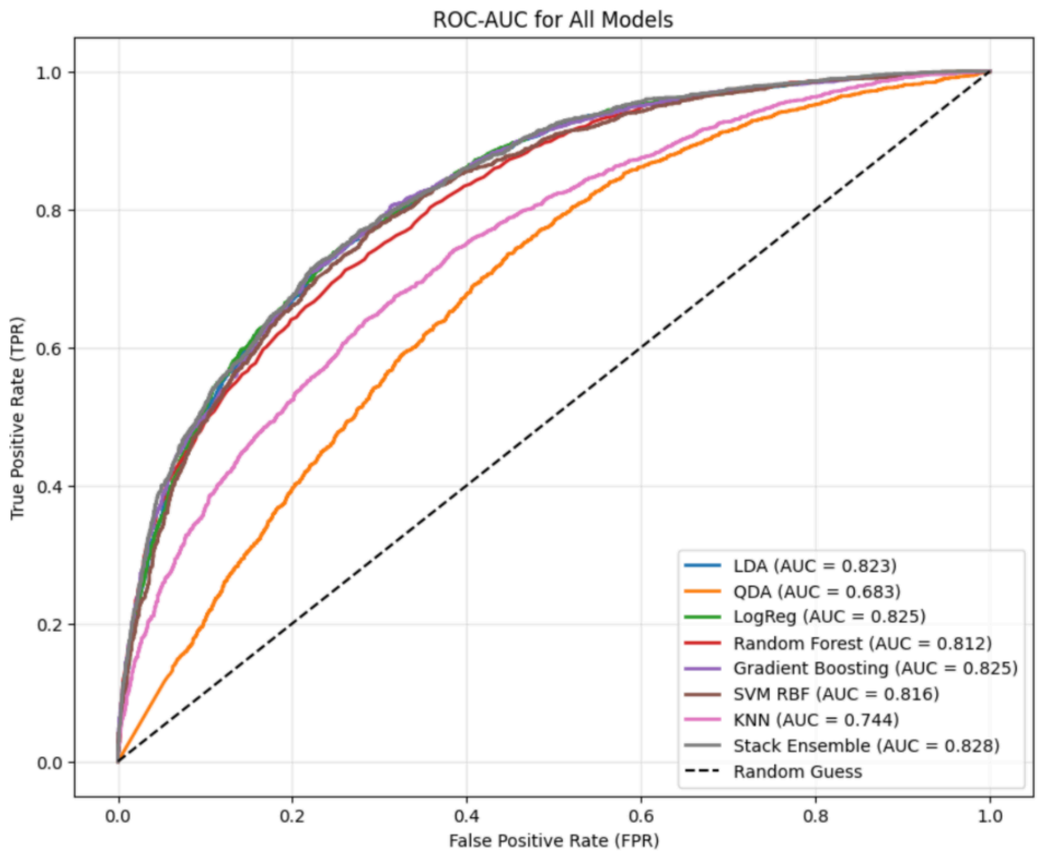


Table 2: All model performance summary

Model	F2	Accuracy	TPR	FNR	FPR	TNR
LDA	0.80	0.71	0.88	0.12	0.43	0.57
QDA	0.77	0.61	0.86	0.14	0.60	0.40
Random Forest	0.81	0.68	0.90	0.10	0.50	0.51
Boosting	0.82	0.70	0.89	0.10	0.46	0.54
Logit Reg.	0.82	0.71	0.89	0.12	0.44	0.56
SVM	0.76	0.73	0.79	0.21	0.32	0.68
KNN	0.79	0.60	0.89	0.11	0.64	0.36
Stack Ensemble	0.82	0.71	0.89	0.11	0.45	0.55

Stacked Model Performance Analysis & Recommendation

Evaluation of F2-score, Accuracy, and Error Rates for Mushroom Safety

F2-score
(Recall-Correctness Trade-off)

0.82 → **Best Trade-off**

Balances recall of poisonous species and overall correctness.

Accuracy
(Overall Correctness)

0.70

Represents overall model correctness.

False-Negative Rate
(Critical Safety Metric)

11% **Lowest Among Contenders**

Crucial for minimizing missed poisonous species.

False-Positive Rate
(Safety "Insurance Premium")

45%

Deemed an acceptable cost for public safety.

Key Insight: Prioritizing Public Safety

With the lowest false-negative rate and the best trade-off (F2-score), the stacked model prioritizes minimizing the risk of severe outcomes. The 45% false-positive rate is acceptable as a safety precaution.

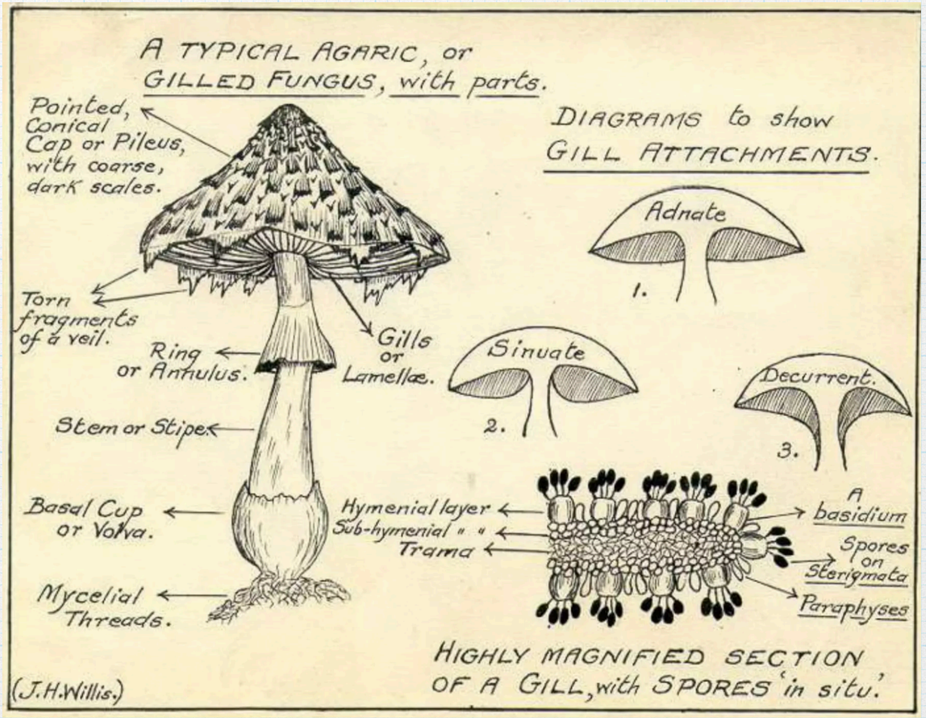
Recommended Production Classifier

Feature Importance in Mushroom Safety Analysis: Key Predictors Ranked

Based on Performance Drop Analysis and Robust Across Random Seeds

Recommendation & Future Improvements

- **For better reliability: ensure dataset covers wide variety of species — avoid focusing on just few types.**
- **Consider combining traditional feature-based ML with image-based deep learning (when images available) for robust classification.**
- **Use explainability methods (e.g. feature importance, SHAP, model-agnostic analysis) to understand and trust model decisions.**
- **Always treat predictions as "aids", not definitive — misclassifying a poisonous mushroom can be fatal.**



Conclusion

- **ML models — especially ensemble methods — can effectively classify mushrooms as edible or poisonous with high accuracy.**
- **Feature-based classification (using physical characteristics) is feasible and interpretable.**
- **Combining good data, sound ML practices, and interpretability makes such classification safer and more reliable.**
- **However, model predictions should not replace expert judgement — use as supportive tool only.**