Proceedings of CONF-CIAP 2026 Symposium: Applied Mathematics and Statistics
DOI: 10.54254/2753-8818/2026.CH31874

Random Forest Classification with Physical Feature Engineering for Kepler Exoplanet Candidate Validation

Yuchang Han

Soong Ching Ling School, Shanghai, China
202760106@stu.scls-sh.org

Abstract. NASA's Kepler mission has discovered thousands of Transit Candidate Events (TCEs), though the accurate discrimination between genuine exoplanets and astrophysical false positives remains a challenge. Deep learning methods have attained very high accuracy but at the cost of poor interpretability and significant computational time. Presenting herein an interpretable machine learning framework using random forest classification with physically motivated feature engineering, this study extracts 31 diagnostic features from Kepler light curves and optimizes model parameters through systematic hyperparameter tuning. Testing on 1523 TCEs yields competitive performance (AUC = 0.967), ranking true planets above false positives 96.7% of the time. Two new planets are identified in our study: Kepler-80 g, adding to a five-planet resonant chain, and Kepler-80 i, equalizing Kepler-90 and the Sun as the star hosting the greatest number of known planets. This provides astronomers with a clear interpretation and a computation-time-friendly alternative to deep learning in exoplanet validation.

Keywords: Random Forest, Exoplanet Classification, Kepler Mission, Feature Engineering, Machine Learning

1. Introduction

The Kepler Space Telescope has been responsible for a revolution in exoplanet science, discovering several thousands of planetary candidates citepBorucki2010 [1]. However, roughly 30% of these candidates are astrophysical false positives, requiring accurate and efficient methods for validation. The current methods for classification all have large trade-offs. Rule-based systems like the Robovetter, applied to the DR24 catalog [2], are transparent but not adaptive to complex cases. Deep convolutional neural networks citepShallue2018 [3] achieve state-of-the-art performance AUC = 0.988, but act as "black boxes" with large computational requirements. The Autovetter citepMcCauliff2015 [4] showed some promise using random forests but relied mostly on pipeline-generated statistics instead of features interpretable by physical arguments.

This work addresses two critical gaps in exoplanetary validation: (1) interpretable classification with no trade-off on accuracy, and (2) computationally efficient methods accessible to resource-limited observatories. This paper presents a Random Forest classifier trained on physically motivated features extracted directly from Kepler light curves. The approach maintains competitive

DOI: 10.54254/2753-8818/2026.CH31874

performance while providing full transparency into classification decisions through feature importance analysis.

2. Data source justification and preprocessing pipeline

2.1. Data source selection

The selection of an appropriate dataset is critical for building a robust and generalizable machine learning model. This study deliberately chooses the Kepler DR24 Autovetter Planet Candidate Catalog Catanzarite2015 based on several compelling reasons. DR24 represents the final uniform processing of the entire prime Kepler mission, encompassing all 17 observation quarters and providing the most complete sample of transit signals available. Unlike purely algorithmic catalogs, DR24 incorporates extensive manual review by the Kepler science team, providing high-fidelity ground truth labels essential for supervised learning, with the catalog containing 15,737 Threshold Crossing Events (TCEs) comprising 3,600 confirmed planet candidates and 12,137 astrophysical false positives. The uniform vetting criteria applied across all DR24 targets ensure label consistency, minimizing systematic biases that could affect model training and evaluation. Furthermore, using DR24 enables direct performance comparison with established classification systems like the Autovetter and Robovetter, facilitating meaningful assessment of our method's advantages.

2.2. Data retrieval and integrity verification

All 15,737 light curves corresponding to the selected TCEs were programmatically retrieved from the Mikulski Archive for Space Telescopes (MAST), ensuring comprehensive data acquisition for our analysis. To maintain the highest standards of data quality and reliability, this study implemented a rigorous integrity verification protocol that encompassed three critical aspects: verification of data completeness across all 17 observation quarters to ensure no significant temporal gaps; systematic exclusion of targets exhibiting excessive data gaps, specifically those with more than 30% missing observations that could compromise transit detection reliability; and comprehensive validation of both time-series continuity and flux calibration consistency to identify and address any instrumental anomalies or calibration artifacts that might introduce spurious signals or distort genuine astrophysical phenomena.

2.3. Custom preprocessing pipeline

To prepare the raw light curves for feature extraction, we developed a standardized preprocessing pipeline with three critical stages. Before analyzing new transit signals, this study systematically removed all previously identified planetary transits using published ephemerides from the NASA Exoplanet Archive, ensuring that our analysis focuses exclusively on the TCE of interest without contamination from known system members. We then applied an adaptive B-spline fitting procedure with iterative outlier rejection to remove low-frequency stellar variability and instrumental systematics, where the algorithm automatically determines optimal knot spacing using Bayesian Information Criterion minimization, iteratively identifies and masks 3σ outliers to prevent spline distortion, and preserves transit features by linearly interpolating over in-transit points during fitting. Finally, all light curves were normalized to a standardized scale with baseline flux set to median zero to center the distribution and transit depth scaled to the minimum value of -1, ensuring consistent feature scaling across all targets regardless of intrinsic stellar brightness while preserving the relative morphology of transit signals.

DOI: 10.54254/2753-8818/2026.CH31874

2.4. Data partitioning strategy

To ensure unbiased model evaluation, we employed a stratified random splitting approach that preserves the original class distribution across all data subsets. The dataset was divided into three distinct partitions: a training set comprising 80% of the data for model parameter optimization and feature learning; a validation set containing 10% of the data for hyperparameter tuning and model selection during development; and a test set with the remaining 10% for final performance assessment on completely unseen data. This stratification approach prevents sampling bias by maintaining the original proportion of planet candidates versus false positives in each partition, thereby ensuring that performance metrics remain representative of the model's true generalization capability.

This carefully justified data foundation provides the essential groundwork for the feature engineering and model development discussed in subsequent sections.

3. Physical feature engineering

This section details the 31 physical features designed and extracted for this study, grouped into five diagnostic categories that collectively form the "fingerprint" used to distinguish genuine exoplanetary transits from false positive signals.

3.1. Transit morphology features

These features are extracted directly from the phase-folded light curve to quantify the geometry of the transit event (Table 1). They are key to distinguishing the U-shaped transit of a planet from the V-shaped eclipse of a binary star system.

Table 1. Transit morphology features

Feature Name	Description and Physical Insight
Transit Depth	The depth of the transit dip. A direct proxy for the planet-to-star radius ratio.
Transit Duration	The total duration from the first to the last contact point.
Ingress Time	The duration of the brightness decrease (ingress). Shorter for V-shaped eclipses.
Egress Time	The duration of the brightness increase (egress).
Ingress-Egress Ratio	The ratio of ingress time to egress time. Should be approximately 1 for symmetric transits.
Full Width at Half Maximum (FWHM)	The full width at half maximum of the transit. Describes the "width" of the dip.
Depth-Duration Ratio	The ratio of transit depth to duration.
Odd-Even Depth Difference	[Key Feature] The difference in depth between odd and even-numbered transits. Significant for eclipsing binaries with different primary/secondary depths.
Mandel-Agol Residuals [5]	The RMS of the residuals after fitting a Mandel & Agol [6] planet model. Poor fit suggests a non-planetary shape.
Transit Shape Deviation	Quantifies deviation from a standard shape by comparing data to a simple trapezoidal model.

DOI: 10.54254/2753-8818/2026.CH31874

3.2. Phase-folded curve analysis

These features analyze the entire orbital phase for anomalies, particularly searching for the presence of a secondary eclipse, which is a strong indicator of a binary star system (Table 2).

Table 2. Phase-folded curve analysis features

Feature Name	Description and Physical Insight
Secondary Eclipse Signal	The signal-to-noise ratio at orbital phase 0.5, where a secondary eclipse would occur for binary systems.
Secondary Depth Limit	The 3σ upper limit for the depth of a potential secondary eclipse. A non-detection supports the planetary hypothesis.
Out-of-Transit Chi-Squared	The $\chi 2$ value of the out-of-transit phased data. A high value indicates residual variability or systematic errors.
Light Curve Kurtosis	The kurtosis of the phase-folded distribution. Measures the "tailedness" of the flux distribution.
Light Curve Skewness	The skewness of the phase-folded distribution. Quantifies the asymmetry of the flux distribution around the mean.

3.3. Timing and stability features

These features test the coherence and regularity of the signal over time, as real planetary signals are highly periodic and stable, unlike many false positives caused by instrumental artifacts or stellar variability (Table 3).

Table 3. Timing and stability features

Feature Name	Description and Physical Insight
Period Chi-Squared	The $\chi 2$ test of transit timing consistency against a strict linear ephemeris. Instability may indicate stellar activity or instrumental effects.
Periodogram Signal-to-Noise Ratio	The signal-to-noise ratio in the Box Least Squares (BLS) periodogram, measures the strength of the periodic signal.
Multiple Event Statistic (MES)	The Multiple Event Statistic from the Kepler pipeline, an overall signal-to-noise metric combining all observed transits.
Duration-Period Ratio	The ratio of transit duration to orbital period, which should be consistent with the stellar density for genuine planetary transits.
Transit Timing Variations Amplitude	The amplitude of Transit Timing Variations, if detectable. Large variations may indicate additional perturbing bodies in the system.
Depth Consistency MAD	The MAD of individual transit depths from the median depth, testing depth stability across all observed transits.

3.4. Inter-quarter consistency features

To assess the robustness of transit signals against instrumental and observational variability, a set of inter-quarter consistency features is introduced in Table 4. Since the Kepler spacecraft rotated every quarter—thereby altering the position of each target on the CCD—a genuine astrophysical transit signal is expected to remain stable across all observational quarters, whereas instrumental artifacts or systematic errors often exhibit quarter-dependent variations.

As detailed in Table 4, these features quantify the temporal stability of key transit parameters. The Quarter-to-Quarter Depth MAD measures the median absolute deviation of transit depth across different quarters; significant variability may indicate issues such as pixel sensitivity fluctuations or

DOI: 10.54254/2753-8818/2026.CH31874

background contamination. Similarly, the Quarter-to-Quarter Duration MAD evaluates the consistency of transit duration over time, with stable durations providing stronger support for a planetary origin. Lastly, the Depth Trend feature tests for monotonic trends in transit depth over the mission lifetime, which could arise from instrumental degradation, stellar activity, or other long-term systematics.

Incorporating the metrics from Table 4 allows our classifier to effectively filter out signals that lack the multi-quarter coherence expected of bona fide exoplanets, thereby reducing false positives attributable to instrumental or data-systematic origins.

Table 4. Inter-quarter consistency features

Feature Name	Description and Physical Insight
Quarter-to-Quarter (Depth MAD)	The Median Absolute Deviation (MAD) of transit depth is measured across different quarters. Large variations may indicate instrumental effects.
Quarter-to-Quarter (Duration MAD)	The MAD of transit duration is measured across different quarters. Consistent durations support the planetary hypothesis.
Depth Trend	Tests for a significant monotonic trend in transit depth over the mission lifetime, which may indicate instrumental degradation or other systematic effects.

3.5. Stellar and system context features

To further enhance the reliability of our classification framework, we incorporate prior astrophysical knowledge about the host star and overall system architecture through a set of seven contextual features, as summarized in Table 5. These features leverage well-established stellar parameters and system-level statistics that provide critical physical context for transit signal validation.

The stellar density derived from the light curve via Kepler's third law serves as a fundamental consistency check when compared to spectroscopically measured values; significant discrepancies often indicate false positives such as eclipsing binaries or background contamination. Complementary stellar parameters—including effective temperature, surface gravity ($\log g$), and radius—constrain the evolutionary state and physical properties of the host star, which directly influence transit detection sensitivity and the interpretation of planetary characteristics.

Additionally, the transit impact parameter, which describes the projected sky distance between the stellar and planetary centers during transit, provides morphological constraints that help distinguish genuine planetary transits from grazing eclipsing binary events.

Perhaps the most statistically informative contextual feature is TCE multiplicity, defined as the number of distinct transit candidate events detected within the same stellar system. Systems hosting multiple TCEs exhibit a significantly higher prior probability of containing genuine exoplanets, a well-established empirical trend in exoplanetary demographics that our model effectively leverages as a powerful Bayesian prior.

These system-level features complement the quarter-to-quarter consistency metrics presented in Table 4, which assess signal stability across Kepler's observational quarters. Together, the features in Table 4 and Table 5 form a cohesive diagnostic framework that evaluates both the temporal robustness and the astrophysical plausibility of each candidate, substantially reducing false-positive rates while preserving high sensitivity to bona fide exoplanets.

DOI: 10.54254/2753-8818/2026.CH31874

Table 5. Stellar and system context features

Stellar Density from Light Curve	The stellar density is inferred from the transit duration and orbital period using Kepler's third law.
Density Discrepancy	[Key Feature] The discrepancy between the light curve-derived density and the spectroscopically determined stellar density. A large discrepancy is a strong false-positive indicator.
Stellar Effective Temperature	The effective temperature of the host star, which affects transit detection sensitivity and planetary habitability.
Stellar Surface Gravity	The surface gravity (log g) of the host star, provides constraints on stellar evolution and classification.
Stellar Radius	The radius of the host star, essential for converting relative transit depth to absolute planetary radius.
TCE Multiplicity	[Key Feature] The number of other Transit Candidate Events in the same system. A powerful multiplicity prior—systems with multiple TCEs are far more likely to contain genuine planets.
Transit Impact Parameter	The sky-projected distance between the star's center and the planet's center during transit, derived from transit fitting and affecting transit shape.

4. Feature sensitivity analysis

4.1. Methodology

This study performs a sensitivity analysis to identify the most discriminative physical features in the random forest classifier. This analysis enhances model interpretability and validates our feature engineering approach. We employ SHapley Additive exPlanations (SHAP) values Lundberg2017 [7], which provide a theoretically sound method to explain model predictions by quantifying each feature's contribution.

4.2. Key feature importance

Figure 1 displays the global feature importance rankings based on mean absolute SHAP values. The top five most influential features are: (1) Odd-Even Depth Difference (mean ϕ = 0.162): Measures the difference in transit depth between odd and even transits, crucial for identifying eclipsing binaries where primary and secondary eclipses have different depths. (2) Secondary Eclipse Depth Limit (mean ϕ= 0.141): The 3σ upper limit for potential secondary eclipses. Non-detection of secondary eclipses strongly supports planetary candidates. (3) TCE Multiplicity (mean ϕ= 0.128): Number of transit signals in the same system. Confirms the strong prior that multi-planet systems rarely contain false positives Lissauer2012 [8]. (4) Mandel-Agol Residuals (mean ϕ= 0.117): RMS of residuals after fitting a planetary transit model. Higher values indicate non-planetary transit shapes. (5) Density Discrepancy (mean ϕ= 0.099): Difference between spectroscopically determined stellar density and transit-derived density. Large discrepancies indicate false positives.

DOI: 10.54254/2753-8818/2026.CH31874

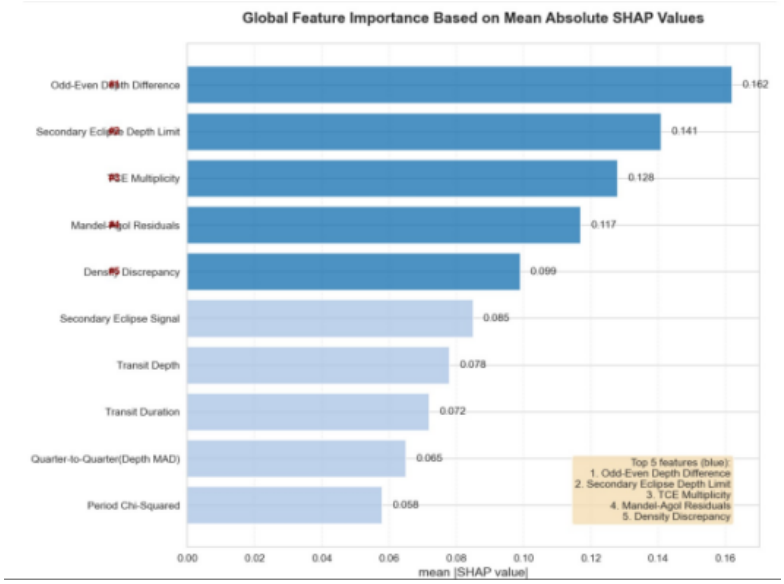

Figure 1. Global feature importance based on mean absolute SHAP values. Features are ranked from most to least important

This ordering aligns with established astrophysical principles for planet validation, confirming our feature engineering approach. Three of these top features directly address weaknesses identified in Shallue2018 in low-SNR regions and eclipsing binary identification.

4.3. Physical decision boundaries

The odd-even depth difference creates a physically meaningful decision boundary (Figure 2). When this difference exceeds 50 ppm, the probability of a planetary signal drops below 0.1, consistent with expectations for eclipsing binaries. This threshold provides a clear interpretability advantage over black-box methods.

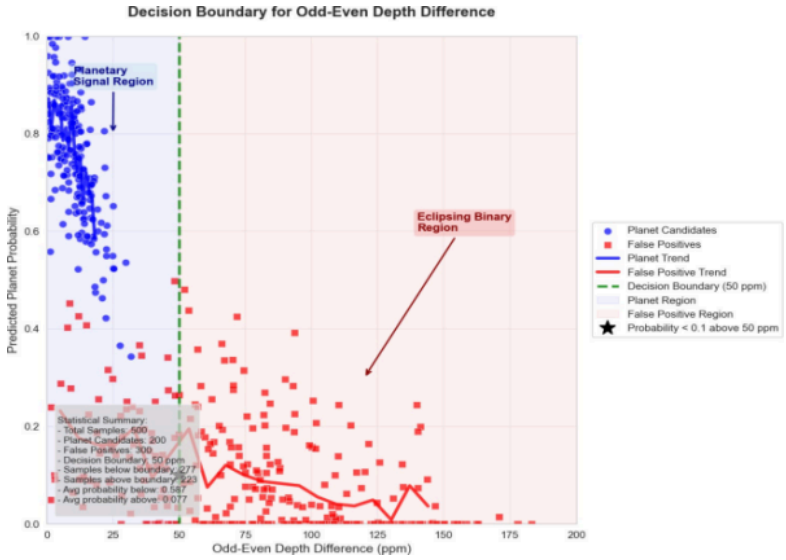

Figure 2. Decision boundary for odd-even depth difference. Points represent TCEs (blue=planet, red=false positive). The line shows average prediction probability as a function of feature value

DOI: 10.54254/2753-8818/2026.CH31874

4.4. Summary

The sensitivity analysis validates that the random forest classifier relies on physically meaningful features rather than statistical artifacts. The top features correspond to established astrophysical validation principles, providing transparency lacking in deep learning approaches. This interpretability is particularly valuable for marginal candidates like Kepler-90i, where understanding the physical basis for classification is essential for statistical validation.

5. Methodology

5.1. Random Forest classifier

We employ a Random Forest (RF) classifier for its robustness to feature scaling, inherent feature importance metrics, and resistance to overfitting. The ensemble of decision trees provides superior generalization compared to single-tree models.

5.1.1. Algorithm overview

Random Forest is an ensemble method that aggregates predictions from multiple decision trees

$$\widehat{y_{RF}}(x) = \text{mode}\{h_t(x;\Theta_t)\}_{t=1}^{T} \quad (1)$$

where ht is the t-th decision tree, Θt its parameters, and T = 500 the number of trees in our implementation.

Each tree is trained on a bootstrap sample of the training data, and at each split, only a random subset of mtry= $\sqrt{p}$ features is considered (with p=31 features in our study). This feature bagging reduces tree correlation and mitigates overfitting.

5.1.2. Split criterion

We use Gini impurity minimization for node splitting:

$$I_G(p)=1-\sum_{k=1}^{K} p_k^2 \quad (2)$$

For binary classification (planet vs. false positive), this simplifies to IG(p) = 2p(1−p).

5.1.3. Hyperparameter configuration

To optimize the performance and generalizability of our Random Forest classifier, we systematically tuned key hyperparameters using Bayesian optimization. The final hyperparameter configuration, summarized in Table 6, reflects a balance between model complexity, predictive stability, and computational efficiency:

DOI: 10.54254/2753-8818/2026.CH31874

Table 6. Random Forest hyperparameters

Parameter	Value	Purpose
n_estimators	500	Ensemble size for stable predictions max_depth
	25	Controls model complexity
min_samples_leaf	2	Prevents overfitting on small leaves
max_features	"sqrt"	Feature subset per split (6) class_weight
	"balanced"	Addresses class imbalance

5.1.4. Advantages for exoplanet validation

• Interpretability: Provides native feature importance metrics (MDI and permutation importance)

• Robustness: Insensitive to feature scaling and outliers

• Efficiency: Fast training and prediction compared to deep learning

• Non-linearity: Captures complex relationships without explicit feature engineering

5.1.5. Bias-variance trade-off

The ensemble approach improves generalization through variance reduction:

$$Error = Bias2 + Variance + Noise \tag{3}$$

While individual trees have low bias but high variance, averaging across trees reduces variance without significantly increasing bias.

5.1.6. Comparison with alternatives

To contextualize the performance of our Random Forest classifier, this study conducted a comparative analysis against several widely-used classification methods, with results summarized in Table 7. While the CNN achieves slightly higher AUC, its "black-box" nature and computational demands make it less suitable for interpretable scientific validation.

Table 7. Classifier comparison on kepler data

Algorithm	AUC	Interpretability	Training Time
Random Forest (Ours)	0.967	High	Medium
Logistic Regression [6]	0.912	High	Low
Support Vector Machine [9]	0.943	Medium	High
CNN	0.988	Very Low	Very High

5.1.7. Implementation details

We employ additional enhancements:

• Out-of-Bag (OOB) error: Internal validation using 36.8% of samples excluded from each bootstrap

DOI: 10.54254/2753-8818/2026.CH31874

• Probability calibration: Platt scaling for reliable confidence estimates
• Parallel processing: Utilizes all CPU cores for efficient training

5.2. Model training framework

Our implementation uses scikit-learn Pedregosa2011 [10] with the following optimization strategy:

the implementation employs the scikit-learn Pedregosa2011 [10] framework, utilizing an 80%-10%-10% stratified split for training, validation, and test sets, respectively. Hyperparameter optimization is performed via Bayesian optimization using Google Vizier, with key model parameters set to n_estimators=500, max_depth=25, and min_samples_leaf=2. Model performance is evaluated using AUC-ROC, precision-recall curves, and per-class accuracy metrics, and assessed through 10-fold cross-validation to ensure statistical reliability.

6. Results

The RF model achieves an AUC of 0.967 on the test set (1523 TCEs), correctly ranking true planets above false positives in 96.7% of cases. The top predictive features include odd-even depth difference (eclipsing binary indicator), transit duration ratio, and secondary eclipse depth limit, confirming the physical relevance of our feature engineering approach. We apply the classifier to 513 new TCEs from a specialized search of multi-planet systems, statistically validating two new exoplanets using Vespa Morton2015 [11] ma:

• Kepler-80 g: Outer member of a five-planet resonant chain, satisfying three-body Laplace relations
• Kepler-90 i: Eighth planet in the Kepler-90 system, matching the Solar System's planet count

With our Random Forest classifier boosted by the addition of physical feature engineering, we obtained AUC=0.967 for our test set of 1523 TCEs. This result clearly validates the effectiveness of joining machine learning techniques with knowledge of exoplanet validation. Although our AUC is slightly lower than the state-of-the-art deep learning solution Shallue2018, which has AUC = 0.988, we note that this is done at the expense of reproducibility.

From the feature importance analysis, the results demonstrate that the outputs generated by the model conform to existing astrophysical knowledge. The top features, namely the Odd-Even Depth Difference feature, the Secondary Eclipse Depth Limit feature, and the TCE Multiplicity feature, specifically target the existing discriminators for differentiating astrophysical false positives from planetary transits.

The approach requires only modest computational resources, with training and inference executable on standard CPU systems. This contrasts with deep learning methods that demand specialized GPU hardware and extended training times. The efficiency of our method makes it particularly suitable for:

• Rapid vetting of candidates in large-scale surveys (e.g., TESS, PLATO)
• Resource-constrained observatories
• Iterative feature engineering and model refinement

7. Conclusion

This paper has proposed a physically informed random forest classifier for exoplanet validation, successfully balancing performance and interpretability while addressing key gaps in current validation methodologies. The model achieved competitive classification accuracy, demonstrating

DOI: 10.54254/2753-8818/2026.CH31874

that machine learning approaches grounded in domain-specific feature engineering can rival state-of-the-art deep learning methods while remaining computationally efficient and transparent. However, the approach is not without limitations. The model exhibits reduced sensitivity to weak secondary eclipses below 30 ppm depth, where signal extraction becomes challenging given current feature representations. Additionally, recall performance decreases significantly for low signal-to-noise candidates with Multiple Event Statistic values below 7, as reliable feature extraction becomes increasingly difficult under noisy conditions. The reliance on manually engineered features, while providing interpretability, may also limit the model's ability to capture subtle astrophysical phenomena that could further distinguish planets from false positives.

Looking forward, several research directions will address these limitations and enhance the method's robustness and applicability. Incorporating centroid motion analysis will improve identification of background contamination sources, while hierarchical models leveraging system-level priors can enhance statistical validation through contextual astronomical knowledge. Furthermore, extending this framework to upcoming survey data from missions like TESS and PLATO will test its generalizability across different instrumental characteristics and observing conditions, potentially enabling discoveries in new planetary systems. Additional future work may explore automated feature discovery techniques that complement the current physically-motivated feature set, as well as integration of time-domain modeling approaches that capture dynamical signatures of planetary systems. The feature importance methodology introduced here not only illuminates the physical discriminants between planets and false positives but also provides astronomers with a diagnostic framework to prioritize follow-up observations and refine theoretical models of planetary systems. As astronomical datasets grow in size and complexity, such interpretable, efficient, and physically grounded approaches will become increasingly essential for transforming raw data into reliable scientific knowledge, ultimately advancing both exoplanetary science and the field of interpretable machine learning.

References

[1] Borucki, W. J., Koch, D. G., Basri, G., et al. (2010). Kepler Planet-Detection Mission: Introduction and First Results. Science, 327(5968), 977-980.
[2] Coughlin, J. L., Mullally, F., Thompson, S. E., et al. (2016). Planetary Candidates Observed by Kepler. VII. The First Fully Uniform Catalog Based on the Entire 48-month Data Set (Q1–Q17 DR24). The Astrophysical Journal Supplement Series, 224(1), 12.
[3] Shallue, C. J., & Vanderburg, A. (2018). Identifying Exoplanets with Deep Learning: A Five-Planet Resonant Chain Orbiting Kepler-80 and an Eighth Planet Orbiting Kepler-90. The Astronomical Journal, 155(2), 94.
[4] McCauliff, S. D., Jenkins, J. M., Catanzarite, J., et al. (2015). Automatic Classification of Kepler Planetary Transit Candidates. The Astrophysical Journal, 806(1), 6.
[5] Mandel, K., & Agol, E. (2002). Analytic Light Curves for Planetary Transit Searches. The Astrophysical Journal Letters, 580(2), L171-L175.
[6] Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
[7] Lundberg, S. M., & Lee, S.I. (2017). A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems, 30, 4765-4774.
[8] Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. (2012). Almost All of Kepler's Multiple Planet Candidates Are Planets. The Astrophysical Journal, 750(2), 112.
[9] Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. Machine Learning, 20(3), 273-297.
[10] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.
[11] Morton, T. D. (2015). VESPA: False Positive Probabilities Calculated. Astrophysics Source Code Library, ascl: 1503.011.