The Astronomical Journal, 166:28 (20pp), 2023 July https://doi.org/10.3847/1538-3881/acd344


# Multiplicity Boost of Transit Signal Classifiers: Validation of 69 New Exoplanets using the Multiplicity Boost of ExoMiner

Hamed Valizadegan[1,2], Miguel J. S. Martinho[1,2], Jon M. Jenkins[2], Douglas A. Caldwell[2,3], Joseph D. Twicken[2,3], and Stephen T. Bryson[2]
[1] Universities Space Research Association (USRA), Mountain View, CA 94043, USA; hamed.valizadegan@nasa.gov
[2] NASA Ames Research Center (NASA ARC), Moffett Field, CA 94035, USA
[3] The SETI Institute, Mountain View, CA 94043, USA
*Received 2022 December 9; revised 2023 March 21; accepted 2023 April 27; published 2023 June 26*

## Abstract

Most existing exoplanets are discovered using validation techniques rather than being confirmed by complementary observations. These techniques generate a score that is typically the probability of the transit signal being an exoplanet ($y(x)$ = exoplanet) given some information related to that signal (represented by $x$). Except for the validation technique in Rowe et al. (2014), which uses multiplicity information to generate these probability scores, the existing validation techniques ignore the multiplicity boost information. In this work, we introduce a framework with the following premise: given an existing transit-signal vetter (classifier), improve its performance using multiplicity information. We apply this framework to several existing classifiers, which include `vespa`, `Robovetter`, `AstroNet`, `ExoNet`, `GPC` and `RFC`, and `ExoMiner`, to support our claim that this framework is able to improve the performance of a given classifier. We then use the proposed multiplicity boost framework for `ExoMiner` V1.2, which addresses some of the shortcomings of the original `ExoMiner` classifier, and validate 69 new exoplanets for systems with multiple Kepler Objects of Interests from the Kepler catalog.

*Unified Astronomy Thesaurus concepts:* Exoplanet astronomy (486); Exoplanet catalogs (488); Exoplanets (498); Exoplanet detection methods (489); Exoplanet systems (484); Convolutional neural networks (1938); Neural networks (1933)

*Supporting material:* machine-readable tables

## 1. Introduction

Since the traditional approach for the confirmation of new exoplanets, which requires complementary observations, is not possible or practical for all candidates due to the increase in the number of candidates and their specifics (e.g., small planets around faint stars), the focus of the discovery of new exoplanets has been shifting from manual follow-up studies to mass validation using automated processes. These automated processes include statistical techniques (Rowe et al. 2014; Morton et al. 2016) and machine learning approaches (Shallue & Vanderburg 2018; Armstrong et al. 2021; Valizadegan et al.2022).

The statistical approaches include `vespa` (Morton et al. 2016), a generative (Bayesian) approach (Vapnik 1998) to calculate the posterior probability of a transit signal for classes of interest (planet candidate or different false positive scenarios), and multiplicity boost (Rowe et al. 2014), a frequentist approach that calculates the probability of a specific planet and False Positive (FP) configuration (i.e., a specific number of planets and FPs) in a multi-planet system. The generative approach to classification models the prior and likelihood in order to obtain the posterior probability. This is in contrast to the discriminative approach, which was introduced by Vapnik (1998) and models the posterior probability directly without the need to model priors and likelihood. The compelling justification of Vapnik (1998) to propose the discriminative approach was that, "one should solve the (classification) problem directly and never solve a more general problem as an intermediate step (such as modeling $p(X|y)$)."

The machine learning approaches that have been applied to transiting planet data mainly consist of discriminative approaches (Vapnik 1998) to calculate the posterior probability, which include the following methods: (1) Autovetter (Jenkins et al. 2014; McCauliff et al. 2015), which uses a random forest to classify transit signals summarized in the form of diagnostic metrics represented by scalar values; (2) the first Deep Neural Network (DNN) model to classify transit signals, `AstroNet` (Shallue & Vanderburg 2018), which has been used to validate two new exoplanets; (3) the machine classifiers that were developed in Armstrong et al. (2021) and used to validate 50 new exoplanets; and (4) a recently developed DNN model called `ExoMiner` (Valizadegan et al. 2022) that validated 301 new exoplanets. These statistical and machine learning models validate new exoplanets when the model's confidence about the Planet Candidate (PC) disposition of a transit signal is sufficiently high (typically >0.99).

Most existing validation models rely on the posterior probability of a transit signal alone for validation, i.e., they do not take the configuration of a target star and whether there are existing known planets or FPs in that system into account. However, knowledge of existing planets and FPs for a given system should affect our confidence about the validation of a new unknown signal. More importantly, it is strongly believed that the candidates in multi-planet systems are highly likely planets (Latham et al. 2011; Lissauer et al. 2012; Christiansen 2022). Blindly applying a single threshold 0.99 on the posterior for two different unknown signals, one around a target star with multiple existing exoplanets and another around

a target star with known FPs, is not an optimal solution. One might want to adjust this threshold or equivalently update the score of the classifier based on the multiplicity information. The latter is the focus of this work.

We propose to incorporate multi-planet configuration information to update the posterior probability of a given model in order to more accurately validate new exoplanets in the Kepler data (Borucki et al. 2010). Our framework is agnostic to the type of the base classifier and can be applied to any model, whether statistical or machine learning. We show that incorporating such configuration information to an existing classifier improves its performance. By applying this method to a new version of the `ExoMiner` model, we validate 69 new exoplanets in the Kepler Objects of Interest (KOI) list from the Q1-Q17 DR25 KOI catalog (Thompson et al. 2018). These new exoplanets are related to transit signals around target stars whose specific configuration increases the likelihood of these signals being associated with transiting exoplanets.

The new version of `ExoMiner`, which we call `ExoMiner` V1.2, addresses some of the shortcomings of the original model (Valizadegan et al. 2022). `ExoMiner` V1.2 receives multiple new inputs that are useful for the correct classification of FPs due to eclipsing binaries and background sources.

## 2. Background

### 2.1. Setup

In this paper, we assume there is a classifier that is able to provide a score for any given transit signal. We also assume that such a classifier does not have access to the system configuration of the target star, i.e., it does not use the dispositions of other transit signals detected for the target star. Existing transit-signal classifiers in this category include `vespa` (Morton et al. 2016), `Autovetter` (Jenkins et al. 2014; McCauliff et al. 2015), `Robovetter` (Coughlin 2017), `AstroNet` (Shallue & Vanderburg 2018), `ExoNet` (Ansdell et al. 2018), `GPC`, `RFC`, and other classifiers introduced in Armstrong et al. (2016), and `ExoMiner` (Valizadegan et al.2022).

The objective of this work is to utilize the knowledge of the configuration of a target star in order to improve the performance of a given transit-signal classifier. In our discussion, we distinguish between two sources of information: (1) the multiplicity information of the system defined by the number of Confirmed Planets (CPs), FPs, and unknown (i.e., unclassified) transit signals for that system; and (2) the transit information of a signal. We denote the multiplicity and transit information of signal $x$ by multi($x$) and transit($x$), respectively.

Let us denote by $s_f(x)$ the score of a classifier $f$ operating only on transit information transit(x).[4] The multiplicity information multi($x$) is represented by a triple $N_{\text{FPs}}(x)$, $N_{\text{CPs}}(x)$, and $N_{\text{UNKs}}(x)$, which are the total number of FPs (Bryson et al. 2017), the total number of CPs, and the total number of unknown objects of interest (Thompson et al. 2018), respectively, for the star hosting signal $x$, i.e., multi($x$) = [$N_{\text{FPs}}(x)$, $N_{\text{CPs}}(x)$, $N_{\text{UNKs}}(x)$]. We aim to build a model $g$ that combines the multiplicity and transit information in order to build a more accurate classifier.

In our discussion, we denote an exoplanet in the output from a model by $y = 1$ and an FP by $y = 0$.

[4] $f$ does not have access to multi($x$).

### 2.2. Preliminary Work

Before developing the multiplicity boost framework of this work, which we will discuss in Section 3.1, we report the results of applying the multiplicity boost framework of Lissauer et al. (2012, 2014) and Rowe et al. (2014) in order to boost the performance of an existing classifier. This approach proved to be unsatisfactory, which motivated the development of our data-driven approach (introduced in Section 3.1).

Lissauer et al. (2012, 2014) and Rowe et al. (2014) introduced a framework to compute the probability of a transit signal being associated with a planet using multiplicity boost. Their framework assumes that "false positives are randomly distributed among the targets" and that "there is no correlation between the probability of a target to host one or more detectable planets and to display false positives." Their working data set was the Q1-Q8 KOI catalog, for which few gold standard labels were known at that point. Therefore, to test their model, Lissauer et al. (2014) and Rowe et al. (2014) used planet candidates and identified FPs to build two data sets: one that included low-depth Eclipsing Binaries (EB; depth <2%) and another that did not. After counting the total number of targets in the sample with $i$ planet candidates for different values of $i$ ($i = 1, 2, 3, \ldots$), they estimated two parameters in a data-driven approach: (1) the fidelity of the sample of single-planet candidates, $P_1$; and (2) the number of existing FP planet candidates present among the multi-planet systems, $n_{fm}$. To this end, they estimated $P_1$ as the number of single-planet candidates that are actual planets by excluding identified FPs from the population of singles, and $n_{fm}$ using the number of known observations for different scenarios (e.g., number of targets with two and three FPs, one planet + one FP, one planet + two FPs). They showed that this approach provided predictions of these different scenarios that are in close agreement to the observed values.

By denoting $p_{\text{multi}}(y = 1|x)$ as the probability of exoplanet using the multiplicity information alone, one can utilize the multiplication strategy introduced in Armstrong et al. (2021) and adapted by Valizadegan et al. (2022) in order to combine the probability score of the classifier $f$, i.e., $p_f(y = 1|x) = s_f(x)$, with $p_{\text{multi}}(y = 1|x)$ as follows:

$$p_{f,\text{multi}}(y = 1|x) = \frac{p_f(y = 1|x)p_{\text{multi}}(y = 1|x)}{\sum_y p_f(y|x)p_{\text{multi}}(y|x)}. \quad (1)$$

For this to work, we need a machine classifier $f$ that is able to generate meaningful probability scores, i.e., $s_f(x) = p_f(y = 1|x)$.

To calculate $p_{\text{multi}}(y = 1|x)$, we applied the statistical framework of Lissauer et al. (2012, 2014) and Rowe et al. (2014) to the more recent Kepler Q1-Q17 DR25 KOI catalog (Thompson et al. 2018) by performing similar stellar and KOI cuts to our initial target star population. The stellar catalog contains a total of 104,159 targets ($n_t = 104{,}159$) that pass the following constraints: 3 hr transit duration combined differential photometric precision (CDPP, Christiansen et al. 2012; Jenkins 2020) smaller than 160 ppm; not evolved as defined by Huber et al. (2014); not binary per the analysis of Berger et al. (2018); with a renormalized unit weight error (RUWE, See Section 5) less than 1.2; and with an observation duty cycle greater than 25%. For those Kepler Q1-Q17 DR25 KOIs that are already confirmed as planets or certified as FPs (Bryson et al. 2017), we use their corresponding label. For those not listed as confirmed or certified as FP, we used the Exoplanet Archive disposition in the Cumulative KOI table (i.e., candidate or FP).

**Table 1**
Input Target Counts for Our Experiments for the DR25 Data Set

| | $n_t$ | $n_1$ | $n_2$ | $n_3$ | $n_4$ | $n_5$ | $n_6$ | $n_7$ |
|---|---|---|---|---|---|---|---|---|
| Without EBs | 104159 | 3215 | 383 | 127 | 47 | 17 | 4 | 1 |
| With EBs | 104159 | 3714 | 406 | 128 | 47 | 17 | 3 | 2 |

**Note.** $n_t$ is the number of targets from which the sample is drawn and $n_i$, $i \in \{1,\ldots,7\}$ are the number of targets with exactly $i$ KOIs. With and without EBs refer to the cut for low-depth EBs as per Rowe et al. (2014).

**Table 2**
Comparison Between the Number of Predicted and Identified FPs in Multis using the Statistical Framework of Rowe et al. (2014) for the DR25 Data Set

| Scenarios | With EBs | | Without EBs | |
|---|---|---|---|---|
| | Predicted | Observed | Predicted | Observed |
| 2 FPs | 41.20 | 34 | 22.74 | 17 |
| 3 FPs | 0.27 | 0 | 0.11 | 0 |
| 1 planet + 1 FP | 40.45 | 30 | 30.13 | 23 |
| 1 planet + 2 FPs | 0.72 | 1 | 0.39 | 1 |
| $\geqslant$2 planets + 1 FP | 13.29 | 7 | 9.68 | 6 |
| $\geqslant$2 planets + 2 FPs | 0.24 | 0 | 0.12 | 0 |
| $n_{fm}$ | 96.3 | 72 | 63.25 | 47 |

**Note.** $n_{fm}$ is the expected total number of identified FP candidates in multis and is the sum of the rows. We followed Rowe et al. (2014) by running two tests: one with low-depth EBs and one without.

Table 1 shows the input candidate abundances used to run the statistical framework (similar to Table 2 in Lissauer et al. 2014). Using ordinary least squares, we minimized the squared error between the input value for $n_{fm}$ and the predicted value based on the estimates of expected number of FPs for the different scenarios. We used the values that were estimated in Lissauer et al. (2014) as initial values for $n_{fm}$ in our optimization process. Table 2 reports the results of this experiment for two data sets: one that includes low-depth EBs and one that does not (similar to Table 4 in Lissauer et al. 2014). Assuming that the number of observations provides a good measure of mean and the observations follow a Poisson distribution, the method in Lissauer et al. (2014) generates reasonable results (i.e., difference between Predicted and Observed $<3\sigma$) for the Q1-Q17 DR25 data.

However, the probability of observing 72 or fewer FPs for the "with EBs" case when the expected number is 96.3 is 0.59%. Similarly, the probability of observing 47 or fewer FPs for the "without EBs" case when the expected number is 63.25 is 2.0%. These results indicate that the model is not a particularly accurate representation of the DR25 data set.

Therefore, instead of using the multiplicity boost model from Lissauer et al. (2014) and Rowe et al. (2014), we develop a fully machine learning model to boost the probability scores of a KOI using the information available for the other KOIs of the host star. A machine learning model is more flexible in that: (1) it does not require the base classifier $f$ to generate probability scores, i.e., the generated scores can have any range; and (2) we do not have to use the multiplicity rule introduced in Equation (1). Instead, we can directly learn the mapping from inputs that consist of multiplicity information and classifier scores to the output label (exoplanet or FP). This will be the focus of the following section.

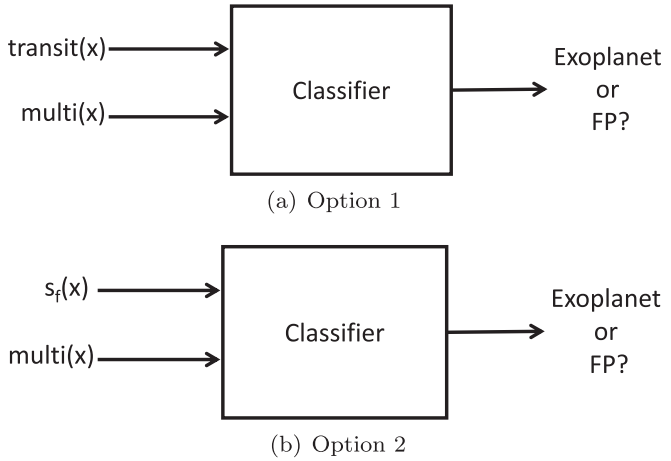


**Figure 1.** Two machine learning models to use multiplicity boost information.

## 3. Multiplicity Boost

Equation (1) provides a way to compute the posterior probability using two initial probabilities, $p_{\text{multi}}$ and $p_f$, that are computed using two sources of information, multi($x$) and transit ($x$), respectively. However, as we discussed in Section 2.2, the results obtained from the multiplicity boost model of Rowe et al. (2014) are not a particularly accurate representation of the DR25 data set. Therefore, we propose to directly learn the posterior probability of a transit signal being a planet using the combination of multiplicity information multi($x$) and transit information transit($x$) in a fully data-driven way. This will be the focus of this section.

### 3.1. Proposed Methodology

To directly learn the posterior probability of a transit signal being a planet using the combination of multiplicity information multi($x$) and transit information transit($x$), we propose to train a classifier that receives the multiplicity and transit information as inputs and generates a posterior probability, as depicted in Figure 1(a).

In this work, rather than training a single classifier to use both the transit and multiplicity information directly, we train a classifier that receives the score provided by an existing transit classifier $f$, i.e., $s_f(x)$, and multiplicity information multi($x$) to return the probability of exoplanet, as depicted in Figure 1(b). This approach has multiple advantages, including:

1. Transit-signal classifiers often work on Threshold Crossing Events (TCEs) but multiplicity information is only meaningful for FP and planetary candidate KOIs. By separating the model that uses transit information from the one that uses multiplicity information, we make the best use of each source of information.
2. This provides a clear design that can be applied on the results of any existing classifier without defining a new classifier. As an example, incorporating multiplicity information into `Robovetter` (Coughlin et al. 2016) could be very difficult because it is an expert system classifier that leverages manually incorporated domain knowledge in the form of if-then conditions to check different types of diagnostic values and transit fit values


**Figure 2.** The machine learning model to boost the performance of a given classifier using multiplicity information.

to classify a TCE. Adapting it to accept multiplicity boost information is not straightforward.

3. The multiplicity information does not interact with the transit information (e.g., diagnostic metrics, and plots) directly. Thus, there is no clear advantage in combining them as the input of a single classifier.

To combine the score of an existing classifier with the multiplicity information of the system, we build a multiplicity classifier $g$ that receives $s_f(x)$, $N_{\rm FPs}(x)$, $N_{\rm CPs}(x)$, and $N_{\rm UNKs}(x)$ as inputs and generates a score $s_g(x)$. To train $g$, we built a training set that consists of input/output pairs in which the input is a quadruple, $[s_f(x), N_{\rm FPs}(x), N_{\rm CPs}(x), N_{\rm UNKs}](x)]$, and the output is whether $x$ is a planet ($y = 1$) or an FP ($y = 0$). For the remainder of this paper, we remove $(x)$ and use $[s_f, N_{\rm FPs}, N_{\rm CPs}, N_{\rm UNKs}]]$ to simplify the notation.

We can train any machine classifier for $g$. In this work, we train a logistic regression model (Cox 1958; Bishop 2006) for two reasons: (1) it generates meaningful probability scores that do not need calibration and can be immediately used for validation; and (2) it is especially useful when input variables include discrete ordinal variables (e.g., counts). A diagram of this system is given in Figure 2. We provide a brief review of logistic regression below.

Under rather general assumptions (Bishop 2006), the posterior probability of class $y = 1$ can be written as a logistic sigmoid function operating over the linear function of input $x$, i.e.,

$$p_g(y = 1|x) = \sigma(w'x) = \frac{1}{1 + \exp(-w'x)} \tag{2}$$

where $'$ indicates the transpose operator. In statistics, this model is called logistic regression. Unlike its name, this model is designed for classification problems (Cox 1958) and not regression. Optimizing this model with regard to the linear parameter, $w$, is achieved by maximum likelihood over the observed data. Our observed data includes pairs of $\langle x_i, y_i\rangle$, $i \in \{1,\ldots,n\}$, where $x_i$ is a quadruple $[s_f, N_{\rm FPs}, N_{\rm CPs}, N_{\rm UNKs}]_i$ (the inputs to the model in Figure 2), and $y_i$ is its label, i.e., exoplanet or FP (right-hand side of Figure 2). By denoting $p_g(y|x_i)$ by $g_i$ and having $\boldsymbol{g} = (g_1,..,g_n)$, the likelihood function can be written as:

$$p_g(\boldsymbol{g}|w) = \prod_{i=1}^{N} g_i^{y_i}(1 - g_i)^{1-y_i}, \tag{3}$$

which can be maximized over the observed data to learn $w$.

For the multiplicity boost framework of this work, which has four input data elements, $w'x$ in Equation (2) can be expanded as follows:

$$w'x = w_{s_f} \times s_f + w_{N_{\rm FPs}} \times N_{\rm FPs} + w_{N_{\rm CPs}} \times N_{\rm CPs} + w_{N_{\rm UNKs}} \times N_{\rm UNKs}. \tag{4}$$

The weights in logistic regression have a nice interpretation. Note that:

$$\log\left(\frac{p_g(y = 1|x)}{1 - p_g(y = 1|x)}\right) = w'x. \tag{5}$$

Given that $w = [w_{s_f}, w_{N_{\rm FPs}}, w_{N_{\rm CPs}}, w_{N_{\rm UNKs}}]$ and $x = [s_f, N_{\rm FPs}, N_{\rm CPs}, N_{\rm UNKs}]$, this exponential relationship provides an interpretation for the weights. When the value of an input changes, the odds will change exponentially proportional with the change. As an example, the odds are multiplied by $\exp(w_{N_{\rm FPs}})$ when we increase $N_{\rm FPs}$, the total number of FPs, by 1. After training the logistic regression and knowing the values of $w$, we know how each input affects the output.

### 3.2. Constructing the Training Set

To train the multiplicity boost framework described in the previous section, we need to have (1) a base classifier that generates a score for a given transit signal and (2) access to multiplicity information of the target stars in the data set. We assume that a transit-signal classifier is already trained and provided. Here, we discuss how to construct the multiplicity information. Assume that we have access to a catalog of stars and their number of known planets, FPs, and other detected transits that are not annotated yet. Without loss of generality and to simplify the discussion, we focus on Kepler data to describe the details of how we construct our training set. However, the same approach can be applied to data of any other survey.

In the Kepler catalog, each star is associated with multiple labeled (exoplanets and FPs) and unlabeled (unknowns) KOIs. Given that our objective is to to train the logistic regression model depicted in Figure 2, we need to build input/output pairs $\langle[s_f, N_{\rm FPs}, N_{\rm CPs}, N_{\rm UNKs}], y(x)\rangle$ where $x$ is a transit signal and $y$ is the corresponding label (exoplanet or FP).

A given annotated KOI provides one single input/output pair that can be used to generate more pairs in order to create a uniform data set. To this end, we construct additional input/output pairs from a single KOI by combinatorially assuming that each subset of annotated KOIs (CPs or FPs) for the target star could be unknown. To explain this, suppose there is exoplanet B on a target star with one FP, two other CPs (hence, three planets in total), and one unknown KOI. The original pair for this KOI is $\langle[s_f(B), 1, 2, 1], y = 1\rangle$. We then build the following pairs out of this single KOI by turning each exoplanet or FP to unknown:

1. Original: $\langle[s_f(B), 1, 2, 1], y = 1\rangle$
2. Changing one CP to unknown: $\langle[s_f(B), 1, 1, 2], y = 1\rangle$
3. Changing two CPs to unknown: $\langle[s_f(B), 1, 0, 3], y = 1\rangle$
4. Changing one FP to unknown: $\langle[s_f(B), 0, 2, 2], y = 1\rangle$
5. Changing one FP and one CP to unknown: $\langle[s_f(B), 0, 1, 3], y = 1\rangle$
6. Changing one FP and two CPs to unknown: $\langle[s_f(B), 0, 0, 4], y = 1\rangle$.

For a KOI on a target star that has $N_{\rm FPs}$ FPs and $N_{\rm CPs}$ CPs (excluding that KOI), we build a total of $(N_{\rm FPs} + 1) \times (N_{\rm CPs} + 1)$ new instances. The idea here is that the logistic

regression model in Figure 2 needs to learn the mapping, even when a labeled KOI is assumed unknown. We aim to provide all combinations of inputs to the model so that it can generalize better to unseen situations that can have any subset of the KOIs labeled or not.

This procedure, however, repeats the same KOIs multiple times, which leads to more emphasis on KOIs for systems with more known companion KOIs. To address this issue, we reweight each generated sample by the number of examples that we generated from the original KOI, i.e., $[(N_{\mathrm{FPs}}+1)\times(N_{\mathrm{CPs}}+1)]^{-1}$. For instance, in the above example, we reweight each generated sample by $1/6$. Utilizing weights for samples in machine learning models, including logistic regression, is a standard practice in imbalanced cases. Even if a classifier does not accept such weights, one can simply repeat the samples to incorporate their appropriate weights (e.g., with probabilistic sampling with replacement).

## 4. Proof of Concept

As a proof of concept prior to using the proposed framework to validate new exoplanets, we show that the methodology introduced in Section 3 can improve the performance of any given classifier.

### 4.1. Base Classifiers

To show that the methodology introduced in this work can improve the performance of any given classifier, we apply this methodology to multiple classifiers described in Valizadegan et al. (2022), as summarized below:

1. *vespa:* We use `vespa` FP probabilities provided in the Kepler Q1-Q25 DR25 False Positive Probability table described in Morton (2012) and Morton et al. (2016).
2. *Robovetter:* We use the `Robovetter` scores (Coughlin et al. 2016) for the Kepler Q1-Q17 DR25 TCE catalog.
3. *AstroNet:* In Valizadegan et al. (2022), the authors used the `AstroNet` code available on GitHub (Shallue & Vanderburg 2018), preprocessed the data, and trained the model using the same setup and DNN architecture as provided in Shallue & Vanderburg (2018). In this work, we use the scores reported in Valizadegan et al. (2022).
4. *ExoNet:* The original code of `ExoNet` is not available. We use the scores of `ExoNet` reported in Valizadegan et al. (2022).
5. *Random Forest Classifier (RFC):* This is one of the classifiers introduced in Armstrong et al. (2021). We use the scores provided by the authors.
6. *Gaussian Process Classifier (GPC):* This is one of the classifiers that was introduced in Armstrong et al. (2021). We use the scores provided by the authors, similar to `RFC`.
7. *ExoMiner:* We use the scores reported for the `ExoMiner` classifier in Valizadegan et al. (2022).

For each of the aforementioned classifiers, we use two sets of classifiers' scores as inputs to our logistic regression model: (1) the raw output of the classifier (Table 4 in Valizadegan et al. 2022); and (2) the posterior probability of the classifier after the application of prior probability of different scenarios, as discussed in Morton et al. (2016), Armstrong et al. (2021), and Valizadegan et al. (2022) (see also Table 11 in Valizadegan et al. 2022). The only exception is `vespa`, for which computing the posterior probability is part of the model. Thus, we only use the posterior probability as an input to the logistic regression model. As we will discuss in Section 4.3, our extensive study shows that our framework is agnostic to the source of the scores and generates similar performance boosts when applied directly to the outputs of the original classifiers.

### 4.2. Evaluation Metrics

We compare the performance using the following metrics:

1. *Accuracy:* This is the fraction of correctly classified cases. When the data set is imbalanced in terms of the percentage of positive examples, this metric is not particularly informative. To see this, note that a classifier that classifies all transits as FPs has an accuracy of 90% if 90% of examples are FPs. However, accuracy provides some insights when studied in conjunction with other metrics.
2. *Precision:* Also called positive predictive value, this is the fraction of transit signals classified as planets that are indeed true planets, i.e.,
$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \tag{6}$$
3. *Recall:* Also called the true positive rate, this is the fraction of planets correctly classified as such, i.e.,
$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \tag{7}$$
4. *Precision-Recall (PR) curve:* The PR curve summarizes the trade-off between precision and recall by varying the threshold used to convert the classifier's score into a label. PR area under the curve (AUC) is the total area under the PR curve. An ideal classifier would have an PR AUC of 1.
5. *Receiver Operating Characteristic (ROC) curve:* The ROC curve summarizes the trade-off between the true positive rate (recall) and false positive rate (fall-out) when varying the threshold used to convert the classifier score into a label. ROC AUC is the total area under the ROC.

### 4.3. Performance Boost of Existing Classifiers

To test and compare the performance of different base classifiers (with and without priors), we use the same labels and the 10-fold cross-validation scheme that was reported in Valizadegan et al. (2022). The performance of these classifiers before and after the application of our multiplicity boost methodology is reported in Table 3. As can be seen from the results, the multiplicity information improves the performance of all of the classifiers across all evaluation metrics, with the exception of PR AUC for some models. The reason why the multiplicity boost information reduces the performance in terms of PR AUC in some cases is because logistic regression was not designed to optimize PR AUC directly. Overall, the proposed method is effective and independent of the type of the employed base classifier.

### 4.4. Model Interpretation

To study the behavior of our machine learning multiplicity boost model, we report the weight values of the logistic regression trained for different classifiers in Table 4. As

**Table 3**
The Multiplicity Model Introduced in This Paper Can Be Applied to Any Given Classifier to Boost its Performance

| Raw Scores | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Original (Table 4, Valizadegan et al. 2022) | | | | Boosted Using Our Multiplicity Approach | | | |
| Model/Metric | Precision and Recall | Accuracy | PR AUC | ROC AUC | Precision and Recall | Accuracy | PR AUC | ROC AUC |
| Robovetter | 0.951 and 0.975 | 0.994 | 0.958 | 0.994 | 0.955 and 0.98 | 0.995 | 0.975 | 0.999 |
| AstroNet | 0.861 and 0.885 | 0.981 | 0.925 | 0.993 | 0.875 and 0.921 | 0.984 | 0.959 | 0.996 |
| ExoNet | 0.925 and 0.864 | 0.985 | 0.956 | 0.995 | 0.922 and 0.933 | 0.989 | 0.976 | 0.998 |
| GPC | 0.921 and 0.964 | 0.991 | 0.982 | 0.998 | 0.92 and 0.975 | 0.992 | 0.986 | 0.998 |
| RFC | 0.929 and 0.955 | 0.991 | 0.979 | 0.998 | 0.928 and 0.973 | 0.992 | 0.961 | 0.998 |
| ExoMiner | 0.968 and 0.974 | 0.996 | 0.995 | 1.0 | 0.971 and 0.977 | 0.996 | 0.993 | 1.0 |
| Posterior Probabilities | | | | | | | | |
| | Original (Table 11, Valizadegan et al. 2022) | | | | Boosted Using Our Multiplicity Approach | | | |
| Model/Metric | Precision and Recall | Accuracy | PR AUC | ROC AUC | Precision and Recall | Accuracy | PR AUC | ROC AUC |
| vespa | 0.666 and 0.968 | 0.801 | 0.865 | 0.928 | 0.749 and 0.973 | 0.865 | 0.895 | 0.943 |
| Robovetter | 0.947 and 0.979 | 0.994 | 0.964 | 0.997 | 0.954 and 0.981 | 0.995 | 0.975 | 0.999 |
| AstroNet | 0.920 and 0.920 | 0.988 | 0.968 | 0.996 | 0.921 and 0.955 | 0.990 | 0.980 | 0.998 |
| ExoNet | 0.933 and 0.902 | 0.988 | 0.97 | 0.997 | 0.933 and 0.953 | 0.991 | 0.983 | 0.998 |
| GPC | 0.910 and 0.977 | 0.991 | 0.987 | 0.998 | 0.911 and 0.983 | 0.992 | 0.986 | 0.998 |
| RFC | 0.942 and 0.964 | 0.993 | 0.983 | 0.999 | 0.942 and 0.978 | 0.994 | 0.965 | 0.999 |
| ExoMiner | 0.971 and 0.978 | 0.996 | 0.996 | 0.999 | 0.974 and 0.983 | 0.997 | 0.993 | 1.000 |

**Note.** The top part of this table is the result of our multiplicity boost framework when the raw classifiers' scores are used as inputs to the logistic regression. The bottom portion of the table shows the performance of our multiplicity boost when the posterior probabilities of the classifiers are used as inputs to the logistic regression. As can be seen, our multiplicity boost framework is agnostic to how the original probability scores are obtained and can improve the scores independently using the multiplicity information.

**Table 4**
Logistic Regression Weights for Different Classifiers

| | Original (Table 4, Valizadegan et al. 2022) | | | | Posterior Probabilities (Table 11, Valizadegan et al. 2022) | | | |
|---|---|---|---|---|---|---|---|---|
| Model/Parameter | $w_{sf}$ | $w_{N_{\rm FPs}}$ | $w_{N_{\rm CPs}}$ | $w_{N_{\rm UNKs}}$ | $w_{sf}$ | $w_{N_{\rm FPs}}$ | $w_{N_{\rm CPs}}$ | $w_{N_{\rm UNKs}}$ |
| vespa | | | | | $6.20 \pm 0.05$ | $-1.93 \pm 0.05$ | $4.63 \pm 0.07$ | $2.60 \pm 0.02$ |
| Robovetter | $7.79 \pm 0.03$ | $-2.49 \pm 0.07$ | $3.27 \pm 0.06$ | $2.06 \pm 0.03$ | $7.90 \pm 0.03$ | $-2.24 \pm 0.09$ | $3.24 \pm 0.06$ | $2.07 \pm 0.04$ |
| AstroNet | $6.54 \pm 0.02$ | $-2.77 \pm 0.06$ | $3.54 \pm 0.04$ | $1.93 \pm 0.02$ | $8.02 \pm 0.03$ | $-2.02 \pm 0.08$ | $3.33 \pm 0.04$ | $1.96 \pm 0.02$ |
| ExoNet | $8.98 \pm 0.03$ | $-2.51 \pm 0.09$ | $3.25 \pm 0.05$ | $1.73 \pm 0.02$ | $9.64 \pm 0.03$ | $-2.73 \pm 0.10$ | $3.06 \pm 0.05$ | $1.78 \pm 0.02$ |
| GPC | $10.81 \pm 0.04$ | $-1.63 \pm 0.03$ | $1.76 \pm 0.05$ | $1.00 \pm 0.03$ | $10.46 \pm 0.03$ | $-2.16 \pm 0.04$ | $1.68 \pm 0.05$ | $1.03 \pm 0.03$ |
| RFC | $9.34 \pm 0.03$ | $-2.36 \pm 0.16$ | $2.39 \pm 0.04$ | $1.47 \pm 0.02$ | $9.76 \pm 0.04$ | $-3.00 \pm 0.11$ | $2.13 \pm 0.04$ | $1.36 \pm 0.03$ |
| ExoMiner | $9.04 \pm 0.03$ | $-1.10 \pm 0.10$ | $2.10 \pm 0.05$ | $1.46 \pm 0.02$ | $9.35 \pm 0.03$ | $-1.46 \pm 0.11$ | $2.04 \pm 0.06$ | $1.54 \pm 0.03$ |

mentioned in Section 3.1, the weight directly demonstrates the effect of the corresponding input parameter on the odds of the event, i.e., the transiting signal is an exoplanet. The weight for the number of FPs, i.e., $N_{\rm FPs}$ is negative, while the weight for all other input parameters, i.e., $S_f$, $N_{\rm CPs}$, and $N_{\rm UNKs}$, are positive for all base classifiers. This shows that the model learned from the data to decrease the odds for systems that already have FPs and increase the odds for systems that have existing CPs or unknown KOIs. The amount of change in the odds, compared to the original odds generated by the original classifier, depends on a number of factors, which include: (1) the accuracy of the base classifier, (2) the precision/recall trade-off of the base classifier, and (3) the regions where the classifier makes more mistakes (e.g., single or multi-planet systems). In general, however, the weight for the original classifier, $s_f$, is higher for more accurate classifiers, as can be seen by comparing the numbers in Tables 3 and 4. Moreover, the weights for multiplicity information, i.e., $N_{\rm FPs}$, $N_{\rm CPs}$, and $N_{\rm UNKs}$, are generally smaller than the weight for $s_f$. This happens because: (1) the base classifier uses transit data, which is more informative than the multiplicity data; and (2) the range of the base classifier score is only between 0 and 1, which is smaller than that of the other input parameters.

By studying the values of individual weights, we can also better understand the importance of different input parameters for the multiplicity model. For example, the weight for $N_{\rm CPs}$ when vespa is used as the base classifier is 4.629, which means that the odds of an exoplanet will increase by a factor of $\exp(4.629) = 102.41$ for any addition of a CP to the system. This number is $\exp(2.043) = 7.71$ for posterior from ExoMiner. The collective behavior can be summarized in mapping plots, where the $x$-axis is the score of the original classifier, the $y$-axis is the score generated by the multiplicity boost framework, and the different curves show the mapping for specific multiplicity scenarios [$N_{\rm FPs}$, $N_{\rm CPs}$, $N_{\rm UNKs}$]. Figure 3 shows such a mapping for the posterior probability of ExoMiner. For each multiplicity scenario [$N_{\rm FPs}(x)$, $N_{\rm CPs}(x)$, $N_{\rm UNKs}(x)$], one solid line represents the mapping for that scenario. We note a few observations here:

1. When the total number of existing exoplanets or unknown KOIs increases for a star, the chances that a new KOI is an exoplanet increases. In the extreme case when there are four existing exoplanets and no FPs, the

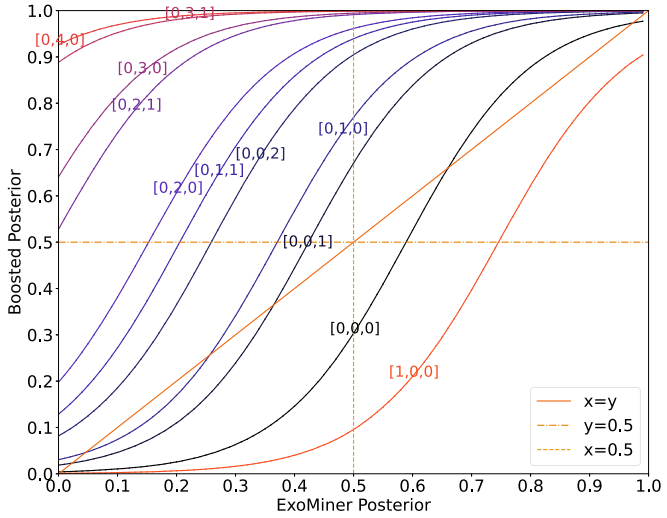


**Figure 3.** The mapping between `ExoMiner` scores and the boosted scores using the logistic regression classifier. Each colored curve represents the mapping for a different multiplicity scenario [$N_{\rm FPs}(x)$, $N_{\rm CPs}(x)$, $N_{\rm UNKs}(x)$]. Scenarios are only plotted if they have more than 10 counts in the Kepler Q1-Q17 DR25 KOI catalog. The larger the number of existing CPs ($N_{\rm CPs}(x)$), the more the logistic regression classifier favors classification as an exoplanet.

multiplicity boost predicts with high confidence that the next KOI is an exoplanet independent of the score of the original classifier. As we will discuss in Section 6, this could go wrong for very rare situations where a background object is detected on a multi-planet system. We will introduce mitigating solutions in Section 5.4 to avoid validating such cases.

2. When the total number of FPs increases, the probability that a new KOI is an exoplanet decreases. However, the prediction of the model for systems with more than two existing FPs should be taken with caution because there are no training data with more than three FPs and the multiplicity classifier extrapolates in that region.
3. When the KOI being classified is the only KOI for a target star (blue line for [0, 0, 0]), the multiplicity boost framework is conservative and lowers the scores when the original scores are less than 0.57. This reflects the larger abundance of single FP KOIs. In the Kepler catalog, there are only 1315 exoplanets out of 4721 single KOI stars. Also note that our multiplicity boost framework does not generate scores larger than 0.99 for any single KOI, which implies that it does not validate single KOIs. Basically, no existing exoplanets in single-planet systems can be validated using this framework. This is an acceptable property because this framework is built to validate exoplanets for multi-planet systems.

Given that this is a multiplicity boost framework, one might expect to see that the mapping for Scenario [0, 0, 0] lies on the diagonal. However, this is not the case because: (1) the model learns a mapping between input quadruple [$s_f$, $N_{\rm FPs}$, $N_{\rm CPs}$, $N_{\rm UNKs}$] and the binary output variable (planet or FP) to perform well for all scenarios and given that the model is more confident for KOIs in multi-planet systems, it adjusts its confidence of the cases for single KOIs, respectively; and (2) the specific shape of the mapping is dictated by the logistic regression assumption that the log odds for the signal being an exoplanet is a linear combination of the input variables [$s_f$, $N_{\rm FPs}$, $N_{\rm CPs}$, $N_{\rm UNKs}$]. The combination of these two properties leads to the specific form of the mapping curves, including one for Scenario [0, 0, 0].

## 5. ExoMiner V1.2

In this section, we introduce an improved version of the `ExoMiner` classifier described in Valizadegan et al. (2022). This new model, which we call `ExoMiner` V1.2, is depicted in Figure 4. Compared to the original `ExoMiner`, the new version includes the following changes, itemized by the branch to which they apply:

1. Transit-View Centroid Branch: We added two new scalar values to the transit-view centroid branch: (1) target star magnitude, $K_p$, required to identify saturated stars for which the centroid diagnostic test is invalid. Instead of using the original value, we converted it to a binary value to represent whether a target star is saturated or not based on a conservative threshold $\text{thr}_{K_p} = 12$; and (2) Renormalized Unit Weight Error (RUWE, Lindegren et al. 2021) that gives information about the quality of the transit source, i.e., larger values can indicate that the source is not a single star or that we are in the presence of significantly crowded fields. This leads to an array of $1 \times 7$ centroid scalars in this branch. The other five features that are already used in the original `ExoMiner` model are the flux-weighted centroid motion detection statistic, centroid offset to the target star according to the Kepler Input Catalog (Brown et al. 2011), centroid offset to the out-of-transit (OOT) centroid position, and finally the respective uncertainties for these two centroid offset estimates.
2. Transit-View Odd and Even Branch: We added one OOT variability scalar for each view (odd and even). This is related to a problem identified in the original `ExoMiner` model that prevented it from effectively learning from the odd and even data. OOT variability scalar describes the variability in the odd/even view and is computed as the median absolute deviation (as a robust estimate of the standard deviation) for the out-of-transit cadences in the odd/even phase folded time series normalized by the square root of the number of cadences; a larger standard deviation or a lower number of cadences make this feature larger, providing information regarding the general variability of these two views.
3. We updated the architecture of the model in Valizadegan et al. (2022) by adding fully connected (FC) layers at the end of the stellar parameters and data validation (DV) diagnostic branches to make sure the contribution of each branch is learned through training instead of being dictated by the architecture. Similar to the other convolutional branches, these FC layers have four units each.

These changes are highlighted by the red-dotted rectangles in Figure 4.

To train `ExoMiner` V1.2, we built a data set similar to Valizadegan et al. (2022) using the Kepler Q1-Q17 DR25 TCE table (Twicken et al. 2016) and the most up-to-date planet catalog, as follows: we first removed all 1498 rogue TCEs from the list. Rogue TCEs are three-transit TCEs that were generated by a bug in the Kepler pipeline (Coughlin et al. 2016). For the planet category, we used the TCEs that are listed as CPs in the Cumulative KOI catalog. For the FP class, we used: (1) TCEs in

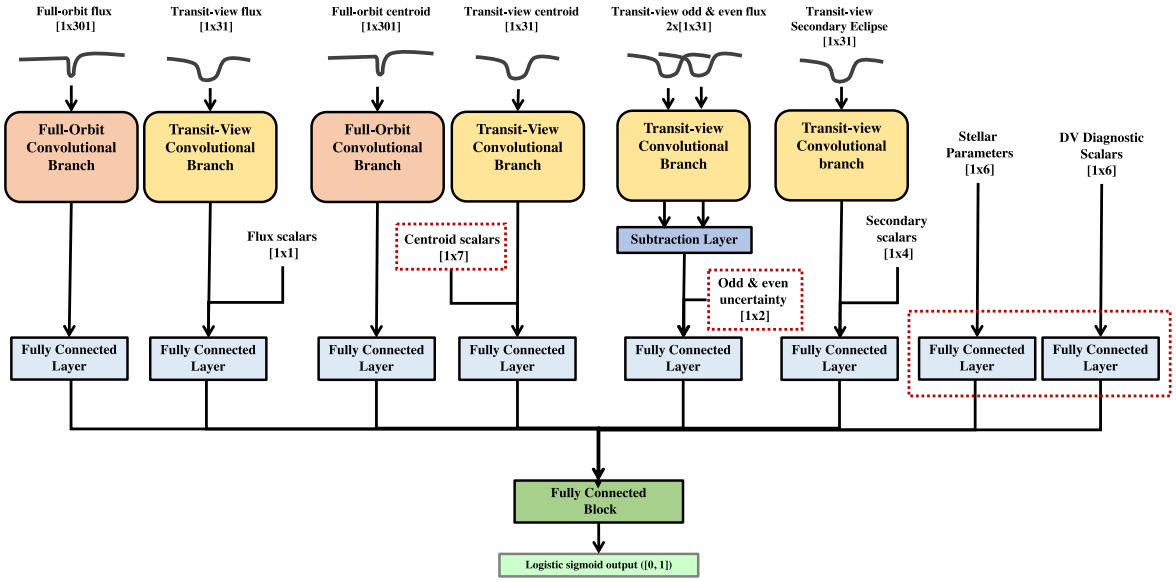


**Figure 4.** An adapted version of the classifier in Valizadegan et al. (2022). The red-dotted rectangles show the differences between ExoMiner and ExoMiner V1.2.

the Q1-Q17 DR25 list that are Certified FPs (CFPs) or certified false alarms (CFAs) in the Kepler Certified False Positive table (Bryson et al. 2017) and (2) TCEs vetted as Non-Transiting Phenomena (NTP) by Robovetter (any TCE from Q1-Q17 DR25 TCE catalog that is not in the Cumulative KOI catalog). This resulted in a total of 30,957 TCEs that consisted of 2643 CPs and 28,314 FPs. Note that this data set is almost identical to the one used in Valizadegan et al. (2022); there are only more exoplanets in this data set as the result of recent validations (Armstrong et al. 2021; Valizadegan et al.2022).

To study the behavior of different models when the training/test split changes, we perform a 10-fold cross-validation (CV), i.e., we split the data into 10 folds, each time we take one fold for test and the other nine folds for training/validation (eight folds for training and one fold for validation). Also, similar to Valizadegan et al. (2022), we split TCEs by their respective target stars instead of by TCEs to remove dependency between training and test sets.

### 5.1. Data Set for the Multiplicity Model

Out of 30,957 Q1-Q17 DR25 TCEs obtained to train ExoMinerV1.2, 8054 TCEs are associated with KOIs in the Cumulative KOI table (Thompson et al. 2018). To build the training set for the multiplicity boost, we used TCEs associated with CP or CFP[5] KOIs, which resulted in a total of 2643 planets and 3538 FPs. For each target, we counted the total number of KOIs, number of planets, FPs, and remaining unclassified (UNK) KOIs. This led to a total of 6181 KOIs, whose scenarios and counts are presented in Table 5 . We split this table into three sections: (1) the top section represents scenarios that have at least one FP, (2) the middle section represents the single-planet systems, and (3) the bottom section represents scenarios that do not have any FP. We would like to emphasize that the numbers of FPs, planets, and unknowns for each KOI include only the other KOIs in the given system. Before providing some insights and to better explain these numbers, we discuss some specific interesting scenarios below that have only a few FP and CP counts. Note that each scenario is represented by $[N_{\mathrm{FPs}}, N_{\mathrm{CPs}}, N_{\mathrm{UNKs}}]$.

**Table 5**
Multiplicity Scenarios with Their Total Number of Exoplanets and FP Counts

| Scenarios | Confirmed Exoplanets | Certified FPs | Total |
|---|---|---|---|
| [3, 0, 0] | 0 | 4 | 4 |
| [1, 0, 0] | 6 | 90 | 96 |
| [1, 0, 1] | 0 | 2 | 2 |
| [1, 0, 2] | 1 | 0 | 1 |
| [1, 1, 0] | 8 | 0 | 8 |
| [1, 2, 2] | 3 | 0 | 3 |
| [1, 4, 0] | 5 | 0 | 5 |
| [0, 0, 0] | 1315 | 3406 | 4721 |
| [0, 0, 1] | 121 | 23 | 144 |
| [0, 0, 2] | 23 | 0 | 23 |
| [0, 0, 3] | 1 | 0 | 1 |
| [0, 1, 0] | 510 | 6 | 516 |
| [0, 1, 1] | 68 | 0 | 68 |
| [0, 1, 2] | 6 | 1 | 7 |
| [0, 1, 3] | 2 | 0 | 2 |
| [0, 2, 0] | 294 | 4 | 298 |
| [0, 2, 1] | 30 | 0 | 30 |
| [0, 2, 2] | 6 | 0 | 6 |
| [0, 3, 0] | 136 | 0 | 136 |
| [0, 3, 1] | 20 | 0 | 20 |
| [0, 3, 2] | 0 | 1 | 1 |
| [0, 4, 0] | 75 | 0 | 75 |
| [0, 5, 0] | 6 | 1 | 7 |
| [0, 6, 0] | 7 | 0 | 7 |
| Total | 2643 | 3538 | 6181 |

**Note.** A scenario is represented by a tuple $[N_{\mathrm{FPs}}(x), N_{\mathrm{CPs}}(x), N_{\mathrm{UNKs}}(x)]$ accounting for all the KOIs in the system other than the CP or CFP under consideration.

1. Scenario [0, 0, 3]: There is only one system, K01082, with four KOIs of which three are unknown. The known KOI of this system, K01082.03, is a CP. This generates one count of CP for Scenario [0, 0, 3].

[5] Note that we did not include CFAs in our analysis because multiplicity boost is based on the statistics of astrophysical FPs and planet multiplicity, whereas FAs do not follow these statistics.

**Table 6**
Information Related to FP Cases for Scenario [0,0,1]

| KOI/Data | KOI Flags | | | | Scores | |
|---|---|---|---|---|---|---|
| | Not Transit-like | Stellar Eclipse | Centroid Offset | Ephemeris Match | ExoMiner V1.2 Posterior | Multiplicity |
| K06137.02 | 0 | 0 | 1 | 0 | 0.497 | 0.732 |
| K00126.01 | 0 | 1 | 0 | 0 | 0.480 | 0.699 |
| K01232.01 | 0 | 1 | 0 | 0 | 0.143 | 0.067 |
| K05449.01 | 0 | 1 | 0 | 0 | 0.132 | 0.064 |
| K04388.01 | 0 | 1 | 0 | 0 | 0.113 | 0.056 |
| K03641.01 | 0 | 1 | 0 | 0 | 0.012 | 0.024 |
| K07124.01 | 0 | 1 | 0 | 0 | 0.003 | 0.018 |
| K07544.02 | 0 | 1 | 0 | 0 | 0.003 | 0.022 |
| K02882.02 | 0 | 1 | 0 | 0 | 0.000 | 0.022 |
| K00379.01 | 0 | 0 | 1 | 0 | 0.000 | 0.021 |
| K06464.01 | 0 | 1 | 0 | 0 | 0.000 | 0.020 |
| K01957.02 | 0 | 1 | 0 | 0 | 0.000 | 0.018 |
| K02159.02 | 0 | 0 | 1 | 0 | 0.000 | 0.018 |
| K06751.02 | 0 | 1 | 1 | 0 | 0.000 | 0.017 |
| K03087.01 | 0 | 0 | 1 | 0 | 0.000 | 0.018 |
| K02184.01 | 0 | 0 | 1 | 0 | 0.000 | 0.020 |
| K02050.01 | 0 | 1 | 1 | 1 | 0.000 | 0.017 |
| K02404.02 | 1 | 0 | 1 | 1 | 0.000 | 0.017 |
| K04323.02 | 1 | 0 | 1 | 1 | 0.000 | 0.022 |
| K01562.01 | 0 | 0 | 0 | 1 | 0.000 | 0.022 |
| K01731.02 | 0 | 0 | 1 | 0 | 0.000 | 0.019 |
| K03230.02 | 1 | 0 | 0 | 1 | 0.000 | 0.019 |
| K04881.02 | 0 | 0 | 0 | 1 | 0.000 | 0.019 |

2. Scenario [1,2,2] and [0,3,2]: The system that generates these scenarios is K02433. There are a total of six KOIs for this system in DR25, which includes three CPs, one CFP, and two unknown KOIs. This results in three CPs for Scenario [1,2,2] and one FP for Scenario [0,3,2].
3. Scenario [1,4,0]: K00082 has five CPs and one CFP, which results in five CPs for this scenario and one FP for Scenario [0, 5, 0], discussed below.
4. Scenario [0, 5, 0]: There are six exoplanets for K00157 that results in six CP for scenario [0, 5, 0]. There is also K00082 that has five exoplanets and one FP, resulting in a single FP for [0, 5, 0].

In general the statistics provided in Table 5 confirm that when there is at least one FP, it is highly likely that the unknown KOI (the KOI in question) is also an FP. When there is no FP, it is highly likely that the unknown KOI is a planet. We provide some insights regarding multiplicity scenarios with at least one CP or unknown KOI but no FPs (i.e., $[x, y, z]$, $x = 0$, $y > 0$ or $z > 0$), and those that have at least 1 FP count (third section in Table 5):

1. Scenario [0, 0, 1]: This scenario has 121 CP and 23 FP counts (first row of the third section of Table 5), the most among scenarios with at least one CP or one unknown KOI. The list of the KOIs for this scenario is provided in Table 6. Interestingly, ExoMiner V1.2 classifies all the FP KOIs correctly and gives very low scores to most of them. None of them get close to the validation threshold of 0.99 after the application of the multiplicity boost.
2. Scenario [0, 1, 0]: There are six FPs for this scenario: K00199.02, K01944.02, K02362.01, K03685.01, K03936.01, and K03954.01. These KOIs have at least one Robovetter FP flag on in the Cumulative KOI table indicating that they failed one or more FP diagnostic tests. Our ExoMiner V1.2 was able to correctly classify these cases with high confidence with scores 2.61E-10, 0.001, 0.407, 0.009, 1.88E-11, and 1.03E-05, respectively. The scores for these KOIs get boosted to 0.03, 0.030, 0.752, 0.040, 0.027, and 0.028, respectively. None of them gets close to the validation threshold 0.99.
3. Scenario [0, 1, 2]: The only FP for this scenario is K03741.01 with the "Stellar Eclipse Flag" on in the Cumulative KOI table. ExoMiner V1.2 gives a very low score of 1.10E-05 to this KOI, indicating that our classifier correctly classified this KOI with high confidence. The multiplicity boost framework boosts its score to 0.381.
4. Scenario [0, 2, 0]: There are four FPs for this scenario: K01196.01, K01806.01, K00672.03, and K01792.02. All four KOIs are flagged by at least one FP indicator in the Cumulative KOI table. ExoMiner V1.2 was also able to successfully capture the FP pattern in the data classifying them with low scores of 0.072, 4.69E-04, 2.76E-06, and 2.02E-4, respectively. The multiplicity boost framework boosts their scores to 0.201, 0.207, 0.182, and 0.136, respectively, which are all well below the validation threshold.
5. Scenario [0, 3, 2]: The only KOI for this scenario is K02433.05, which is an FP. This KOI is flagged by two FP flags in the Cumulative KOI table and scored 1.55E-05 by ExoMiner V1.2. Given that this KOI is around a system with three CPs and two other unknown KOIs, its score gets boosted to 0.987 by the multiplicity boost framework, which is very close to the validation threshold.
6. Scenario [0, 5, 0]: Note that the only FP KOI for scenario [0, 5, 0] is the sixth KOI of K00082 with five existing exoplanets. As mentioned in Valizadegan et al. (2022), this is incorrectly certified as FP in the Certified False Positive table. ExoMiner V1.2 gives a very high score of 0.997 to this KOI, which is boosted to almost 1.0 by the multiplicity boost framework.

**Table 7**
Performance of ExoMiner V1.2 and Its Performance Boost Using the Multiplicity Framework

| | Original | | | | Boosted Using Our Multiplicity Approach | | | |
|---|---|---|---|---|---|---|---|---|
| Model/Metric | Precision and Recall | Accuracy | PR AUC | ROC AUC | Precision and Recall | Accuracy | PR AUC | ROC AUC |
| ExoMiner V1.2 | 0.978 and 0.980 | 0.996 | 0.997 | 0.999 | 0.979 and 0.984 | 0.997 | 0.995 | 1.000 |

**Table 8**
Logistic Regression Weights for ExoMiner V1.2

| Model/Parameter | $w_{sf}$ | $w_{N_{\text{FPs}}}$ | $w_{N_{\text{CPs}}}$ | $w_{N_{\text{UNKs}}}$ |
|---|---|---|---|---|
| ExoMiner V1.2 | $9.728 \pm 0.03$ | $-1.71 \pm 0.10$ | $2.044 \pm 0.07$ | $1.596 \pm 0.03$ |

Note that we did not correct this label noise in the data because there are potentially other sources of label noise. As a matter of the fact we are aware of five CPs that have been demoted to FP.[6] These are Kepler-486 b Díaz et al. (2014), Kepler-492 b (Díaz et al. 2013), Kepler-699 b (Niraula et al. 2022), Kepler-840 (Niraula et al. 2022), and Kepler-854 b (Niraula et al. 2022). All these CPs are single-planet systems which do not affect the learning process.[7] However, we would like to emphasize that machine learning models are generally robust to a significant amount of label noise. A machine learning model does not rely on the statistics for each scenario individually but rather aims to find a mapping that optimizes its objective function, in this case Equation (3).

From the 6181 initial scenario sets reported in Table 10, we generated 8458 new scenarios and their weights using the procedure described in Section 3.2.

### 5.2. Performance Study

Table 7 summarizes the performance result of applying ExoMiner V1.2 posterior to the most recent data set discussed in the previous subsection. It also reports the results after application of the multiplicity boost framework proposed in this work. As can be seen, our multiplicity boost framework also improves the performance of this new model. We also report in Table 8 the weights learned by the multiplicity boost framework for each input parameter. Compared to the original classifier (Table 4), ExoMiner V1.2 scores obtains higher weight because it is more accurate.

To provide insights into how the multiplicity boost improves the performance on specific scenarios, in Table 10 we provide the total number of predicted FP for each scenario for ExoMiner V1.2 and the multiplicity boost framework. As can be seen, multiplicity information helps improve the performance for individual scenarios.

We also report in Table 9 the scores of ExoMiner V1.2 and the multiplicity logistic regression model on all 6181 CP and CFP KOIs discussed in Section 5.1. The multiplicity boost approach gives a score >0.99 to 1281 KOIs, including one FP: K00082.06. This FP KOI is the sixth KOI in a system with five existing exoplanets. As mentioned in Valizadegan et al. (2022), this is incorrectly certified as FP in the Certified False Positive table. Given that there is a total of 1328 planets in multi-planet systems in Kepler Q1-Q17 DR25 data, the precision and recall value at the validation threshold of 0.99 are 1.000 and 0.965, respectively. The equivalent precision and recall values for the ExoMiner V1.2 posterior without multiplicity boost are 1.0 and 0.836. Thus, the multiplicity boost framework significantly improves the recall at the validation threshold of 0.99.

**Table 9**
The Original Scores of ExoMiner V1.2 and the Boosted Scores After the Application of Multiplicity Boost Framework on 6181 known KOIs

| Column | Description |
|---|---|
| KIC | KIC ID |
| TCE | TCE planet number |
| KOI name | KOI name |
| Period (days) | TCE period |
| Radius (Re) | planet radius |
| tce_max_mult_ev | TCE MES |
| ruwe | RUWE value |
| positional prob | positional probability |
| label | 0 for CFP, 1 for CP |
| ExoMiner score | ExoMiner V1.2 score + priors |
| Boosted ExoMiner Score | score assigned by multiplicity |

(This table is available in its entirety in machine-readable form.)

To show how the known KOIs are mapped by our multiplicity boost framework, we plot the score mapping of the CPs and CFPs in the training set for the logistic regression model in Figure 5. We only show lines for tuples $[N_{\text{FPs}}(x),\ N_{\text{CPs}}(x),\ N_{\text{UNKs}}(x)]$ with more than 10 points in the known KOI data. The mapping learned by the multiplicity boost classifier follows the general pattern of data seen in Table 10, i.e., the probability of CPs increases when the scenario has more existing CPs and unknown KOIs. Our multiplicity boost classifier also captures this pattern by increasing the scores for KOIs for target stars with already CPs and unknown KOIs. For single KOIs ([0, 0, 0] curve), whose scenario includes the most points in this figure, even though the boosted score can change relative to the score of the original classifier ($x$-axis), it is always less than 0.99 and does not allow validation of any still unconfirmed single KOIs.

### 5.3. Stability of New Planets

To ensure that the newly validated planets are stable, similar to Dietrich & Apai (2020), we use the dynamical stability parameter $\Delta$ (Fabrycky et al. 2014), defined as:

$$\Delta = \frac{2(a_2 - a_1)}{a_2 + a_1}\left(\frac{3M^*}{m_1 + m_2}\right)^{1/3}, \tag{8}$$

where $a_1$ is the inner planet semimajor axis, $a_2$ is the outer semimajor axis, $M^*$ is the stellar mass, and $m_1$ and $m_2$ are the

[6] https://exoplanetarchive.ipac.caltech.edu/docs/removed_targets.html
[7] They have a negligible effect on the relative ratio of FPs for Scenario [0,0,0].

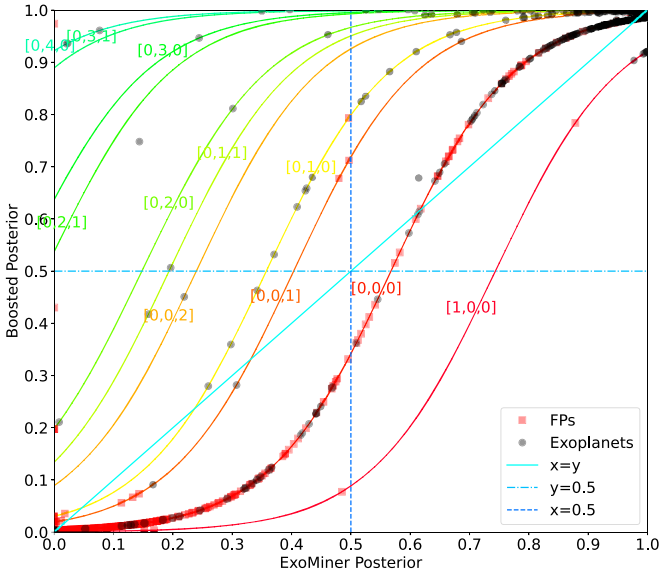


**Figure 5.** The score mapping of the existing exoplanets and FPs in the training data. We only show lines for tuples $[N_{\text{FPs}}(x),\ N_{\text{CPs}}(x),\ N_{\text{UNKs}}(x)]$ with more than 10 points in the known KOI data for clarity. The lines for tuples with less than ten points exhibit similar patterns.

**Table 10**
The Performance of Multiplicity Boost Framework Compared to ExoMiner V1.2 on Individual Scenarios in Terms of FP Counts

| Scenarios | Observed | ExoMiner V1.2 | Multiplicity Boost |
|---|---|---|---|
| [3, 0, 0] | 4 | 4 | 4 |
| [1, 0, 0] | 90 | 89 | 89 |
| [1, 0, 1] | 2 | 2 | 2 |
| [1, 0, 2] | 0 | 0 | 0 |
| [1, 1, 0] | 0 | 0 | 0 |
| [1, 2, 2] | 0 | 0 | 0 |
| [1, 4, 0] | 0 | 0 | 0 |
| [0, 0, 0] | 3406 | 3387 | 3395 |
| [0, 0, 1] | 23 | 25 | 23 |
| [0, 0, 2] | 0 | 1 | 1 |
| [0, 0, 3] | 0 | 0 | 0 |
| [0, 1, 0] | 6 | 14 | 9 |
| [0, 1, 1] | 0 | 2 | 2 |
| [0, 1, 2] | 1 | 2 | 1 |
| [0, 1, 3] | 0 | 0 | 0 |
| [0, 2, 0] | 4 | 7 | 5 |
| [0, 2, 1] | 0 | 0 | 0 |
| [0, 2, 2] | 0 | 0 | 0 |
| [0, 3, 0] | 0 | 1 | 0 |
| [0, 3, 1] | 0 | 0 | 0 |
| [0, 3, 2] | 1 | 1 | 0 |
| [0, 4, 0] | 0 | 4 | 0 |
| [0, 5, 0] | 1 | 0 | 0 |
| [0, 6, 0] | 0 | 1 | 0 |
| Total | 3538 | 3540 | 3531 |

planet masses. Similar to Dietrich & Apai (2020), we use $\Delta > 8$ for stable systems.

To use the above stability criterion, we need the planet mass, $M_p$, which is generally not available. To estimate the planet mass, we use the approach in Fabrycky et al. (2014) as:

$$M_p = M_\oplus \left(\frac{R_p}{R_\oplus}\right)^\alpha \tag{9}$$

where $\alpha = 2.06$ for planet radius $R_p > R_\oplus$ and $\alpha = 3$ for $R_p \leqslant R_\oplus$.

As a conservative approach, we also assume that all unknown KOIs are confirmed planets when we calculate stability, i.e., the new planets should be stable when unknown KOIs are assumed to be planets.

### 5.4. Vetoing Criteria

Given that the constructed training set for this study does not include FAs, we remove KOIs that have MES <10.5 as a conservative precaution to avoid noise-induced FAs. This threshold is similar to the one used by existing validation work (Rowe et al. 2014; Morton et al. 2016; Armstrong et al. 2021; Valizadegan et al. 2022). It is actually based on the reliability analysis performed in Thompson et al. (2018). In particular, Thompson et al. (2018) showed that the reliability meets the 0.99 threshold for MES >10 and period <200 days. For period >200 days, reliability meets the 0.99 threshold when MES >20 and is 0.97 when 10 < MES <20.

Because the multiplicity boost only works if the signal is known to be very close to the target, we veto KOIs from validation if (1) RUWE >1.2 (Lindegren et al. 2021) or (2) positional probability <0.99 (Bryson & Morton 2017) to prevent validation of KOIs for blended sources. This is because the multiplicity boost increases the score of the classifier significantly for scenarios with multiple planets and unknown KOIs.

In addition to these vetoes, we make sure that ExoMiner V1.2 score >0.5 to prevent validating planets that ExoMiner V1.2 labels as FP. As another level of reliability, we remove any KOIs that have any FP flag set in the Cumulative KOI catalog.

### 5.5. Newly Validated Exoplanets

We applied ExoMiner V1.2 and the multiplicity classifier trained in Section 5.2 to the 1570 unknown KOIs in the Kepler DR25 data set. Table 11 reports the ExoMiner V1.2 and the multiplicity boost scores of all these 1570 unknown KOIs. This resulted in a total of 208 KOIs with boosted scores >0.99. Three out of these 208 KOIs were already confirmed by previous works but were not included in our 6181 known KOIs, which are K01831.03 (Kepler-324 d, Jontof-Hutter et al. 2021), K00089.02 (Kepler-462 c, Masuda & Tamayo 2020), and K01783.02 (Kepler-1662c, Vissapragada et al. 2020). Out of these three, K01831.03 has a positional probability <0.99.

Out of the remaining 205 KOIs, only K02926.05 has an original ExoMiner score <0.5. K02926 has five previously confirmed planets and K02926.05 is the last KOI for this system in DR25. In addition to the original score veto, K02926.05 has a positional probability <0.99. Regarding the KOIs' FP flags, there is only one KOI out of these 205 KOIs, K00408.05, that has the "Not Transit-Like Flag" set. The original ExoMiner score of this KOI was 0.575, which was boosted to 0.999 given that there are four other CPs for this system. K00408.05 also has a low MES of 7.9.

After removing KOIs with MES <10.5, RUWE >1.2, or positional probability <0.99, we validate a total of 69 new

**Table 11**
The Original Scores of ExoMiner V1.2 and the Boosted Scores After the Application of Multiplicity Boost Framework on 1570 Unknown KOIs

| Column | Description |
|---|---|
| KIC | KIC ID |
| TCE | TCE planet number |
| KOI name | KOI name |
| Period (days) | TCE period |
| Radius (Re) | planet radius |
| tce_max_mult_ev | TCE MES |
| ruwe | RUWE value |
| positional prob | positional probability |
| ExoMiner score | ExoMiner V1.2 score + priors |
| Boosted ExoMiner Score | score assigned by multiplicity |

(This table is available in its entirety in machine-readable form.)

**Table 12**
Multiplicity Scenarios and Their Total Number of Cases for 1570 Unknown Kepler Q1-Q17 DR25 KOIs Are Reported in Columns 1 and 2, Respectively

| Scenario | Counts | Score >0.99 | Not Validated due to MES <10.5 only | Validated Exoplanets |
|---|---|---|---|---|
| [0, 0, 0] | 1142 | 0 | 0 | 0 |
| [0, 0, 1] | 153 | 66 | 19 | 27 |
| [0, 0, 2] | 12 | 4 | 3 | 1 |
| [0, 0, 3] | 12 | 11 | 2 | 5 |
| [1, 0, 0] | 18 | 0 | 0 | 0 |
| [2, 0, 0] | 1 | 0 | 0 | 0 |
| [0, 1, 0] | 118 | 59 | 31 | 15 |
| [0, 1, 1] | 45 | 24 | 6 | 8 |
| [0, 1, 2] | 3 | 3 | 1 | 2 |
| [1, 1, 1] | 2 | 1 | 0 | 1 |
| [0, 2, 0] | 34 | 16 | 8 | 7 |
| [0, 2, 1] | 6 | 5 | 2 | 1 |
| [0, 2, 2] | 3 | 3 | 0 | 0 |
| [0, 3, 0] | 10 | 5 | 2 | 2 |
| [0, 3, 1] | 4 | 2 | 0 | 0 |
| [1, 3, 1] | 2 | 2 | 2 | 0 |
| [0, 4, 0] | 5 | 4 | 1 | 0 |
| Total | 1570 | 205 | 77 | 69 |

**Note.** The total number of KOIs with score >0.99, the number of KOIs that could be validated if their multiple event statistics (MES) >10.5, and validated exoplanets per each scenario are reported in Columns 3, 4, and 5, respectively.

exoplanets, whose scenario counts presented in Table 12. All 69 of these planets satisfy the stability criteria discussed in Section 5.3. Among these validated planets, we have two that have periods >200 days: K06103.01, whose MES is 35.4, and K01608.03, whose MES is 12.5. The former meets the reliability threshold of 0.99 (Check Section 5.4). The latter has a reliability of 0.97 based on Thompson et al. (2018). However, given that the reliability in Thompson et al. (2018) is calculated for single-planet systems and given that K01608 has already two confirmed planets, the reliability of 0.97 is reasonable.

There are 77 KOIs that did not get validated simply because their MES <10.5. Note that out of 1570 unknown KOIs, 943 have MES <10.5. Thus, the majority of remaining unknown KOIs are in the low-MES region. This is due to the fact that almost all previous validation techniques avoid validating planets in this region.

In general, the percentage of newly validated exoplanets increases for scenarios as the number of CPs or unknown KOIs increase. The exceptions to this rule are mainly due to our vetoing conditions, which prevent the validation of new exoplanets. For example, there are five unknown KOIs around target stars that already have four existing exoplanets. Although none of these five unknown KOIs are validated, four of these five unknown KOIs have probability scores higher than 0.99 but fail at least one of the veto conditions.

Table 13 provides the list and properties of the 69 newly validated exoplanets. It also reports the probability score of ExoMiner V1.2 after applying the prior and the probability score after using the logistic regression multiplicity boost approach. Note that the original score of ExoMiner V1.2 for eight of these validated exoplanets is already larger than 0.99, so they could be validated even without multiplicity information. Figure 6 shows the score mapping done by our multiplicity boost model for these new exoplanets. We also report the list of KOIs with boosted score >0.99 that only failed the MES condition in Table 14 and those that did not pass other vetoes in Table 15.

All of the newly validated planets are planet candidates in the Q1-Q17 DR25 KOI catalog according to the Exoplanet Archive disposition. We also manually examined the newly validated exoplanets to make sure there is nothing in the data to indicate that they should not be validated. Among the KOIs that do not have any existing CP, we discuss the following newly validated exoplanets, which are especially interesting. In several cases, the KOI failed the standard difference image centroiding test relative to the centroid of the OOT image but passed the difference image centroiding test relative to the KIC catalog position of the target star. In these cases, the OOT centroid erroneously converged to a nearby star in the postage stamp rather than to the target star itself. The OOT centroid offset diagnostic result is incorrect when the OOT centroid does not locate the position of the target star (Bryson et al. 2013; Twicken et al. 2018). Also note that none of these are flagged for "Centroid Offset Flag" or for any other FP flags in the Cumulative KOI catalog.

1. K01358.01, K01358.02, K01358.03, and K01358.04: We validate all four KOIs of K01358 (represented by black circles on line [0,0,3] in Figure 6). There is no FP or other KOI for this system. The minimum score of 0.635 for the KOIs in this system is assigned to K01358.01 by ExoMiner V1.2. This score is boosted to 0.993 by the multiplicity boost information. All four candidates failed the difference image centroiding test relative to the out-of-transit centroid because the out-of-transit image centroid locked onto a much brighter ($Kp = 13.5$) star 6″ to the north of the $Kp = 15.5$ target star. These four candidates also failed the centroid shift test due to the presence of the brighter nearby star. However, all four passed the KIC offset difference image centroiding test, i.e., all four difference image centroids were consistent with the catalog position of the target star.
2. K03145.01: There are a total of three TCEs in DR25 for this system, of which two are KOIs. Neither of these two KOIs are CPs or FPs, and both appear to have difference image centroid positions relative to the catalog position of the host star consistent with being on target. The ExoMiner V1.2 score for this KOI was 0.986, which

**Table 13**
List of 69 Newly Validated Exoplanets, Sorted by TCE KIC

| Number | TCE KIC | KOI Name | Period (days) | Radius (Re) | MES | Pos. Prob. | RUWE | ExoMiner V1.2 | Multiplicity Score | Kepler Name |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1717722.2 | K03145.01 | 4.54 | 1.82 | 13.8 | 1.000 | 0.96 | 0.986 | 0.997 | Kepler-1977 b |
| 2 | 2832589.1 | K01942.01 | 10.85 | 3.01 | 34.5 | 1.000 | 1.03 | 0.891 | 0.991 | Kepler-1978 b |
| 3 | 3229150.1 | K02150.01 | 18.51 | 4.32 | 22.5 | 1.000 | 1.08 | 0.965 | 0.996 | Kepler-1979 b |
| 4 | 3338885.2 | K01845.02 | 5.06 | 5.64 | 30.2 | 1.000 | 1.06 | 0.981 | 0.998 | Kepler-975 c |
| 5 | 3559860.1 | K03440.01 | 33.03 | 2.85 | 11.1 | 1.000 | 1.05 | 0.937 | 0.994 | Kepler-1980 b |
| 6 | 3561464.1 | K03398.02 | 35.80 | 8.96 | 15.3 | 1.000 | 0.99 | 0.895 | 0.999 | Kepler-1487 c |
| 7 | 3634051.1 | K06103.01 | 453.54 | 9.41 | 35.4 | 1.000 | 0.98 | 0.881 | 0.991 | Kepler-1981 b |
| 8 | 4077526.4 | K01336.04 | 4.46 | 1.69 | 11.4 | 1.000 | 1.00 | 0.861 | 1.000 | Kepler-58 e |
| 9 | 4157325.3 | K01860.03 | 3.08 | 1.73 | 22.4 | 1.000 | 0.95 | 0.927 | 1.000 | Kepler-416 d |
| 10 | 4173026.2 | K02172.02 | 116.58 | 4.06 | 15.0 | 1.000 | 0.99 | 0.990 | 0.998 | Kepler-1801 c |
| 11 | 4458082.2 | K02303.02 | 8.93 | 1.74 | 12.1 | 0.995 | 1.06 | 0.955 | 0.997 | Kepler-1181 c |
| 12 | 4548098.1 | K04157.01 | 3.82 | 0.87 | 13.4 | 1.000 | 0.95 | 0.991 | 0.997 | Kepler-1982 b |
| 13 | 4665571.1 | K02393.02 | 0.77 | 1.46 | 21.0 | 1.000 | 1.10 | 0.994 | 0.998 | Kepler-1834 c |
| 14 | 4770365.3 | K01475.03 | 4.73 | 1.60 | 10.9 | 1.000 | 1.03 | 0.899 | 0.999 | Kepler-1669 d |
| 15 | 4857058.1 | K03061.01 | 7.33 | 1.64 | 11.4 | 1.000 | 0.98 | 0.993 | 0.997 | Kepler-1983 b |
| 16 | 5531953.2 | K01681.02 | 1.99 | 0.81 | 12.9 | 1.000 | −1.00 | 0.610 | 0.994 | Kepler-1984 b |
| 17 | 5942808.2 | K02250.02 | 0.63 | 1.65 | 23.5 | 1.000 | 1.02 | 0.987 | 0.998 | Kepler-1814 c |
| 18 | 6697605.1 | K02851.01 | 3.42 | 2.22 | 16.8 | 1.000 | −1.00 | 0.986 | 0.997 | Kepler-1985 b |
| 19 | 7102227.3 | K01360.03 | 0.76 | 0.89 | 11.5 | 1.000 | 1.02 | 0.961 | 1.000 | Kepler-290 d |
| 20 | 7202957.1 | K02687.01 | 1.72 | 0.83 | 50.0 | 1.000 | 0.89 | 0.905 | 0.995 | Kepler-1869 c |
| 21 | 7285757.1 | K03271.01 | 19.55 | 2.95 | 11.1 | 1.000 | 1.01 | 0.951 | 0.995 | Kepler-1986 b |
| 22 | 7376983.1 | K01358.01 | 5.64 | 3.27 | 51.4 | 1.000 | 1.01 | 0.636 | 0.996 | Kepler-1987 d |
| 23 | 7376983.2 | K01358.02 | 8.74 | 2.45 | 25.0 | 1.000 | 1.01 | 0.803 | 0.999 | Kepler-1987 e |
| 24 | 7376983.3 | K01358.03 | 3.65 | 1.70 | 15.5 | 1.000 | 1.01 | 0.964 | 1.000 | Kepler-1987 c |
| 25 | 7376983.4 | K01358.04 | 2.35 | 1.51 | 14.6 | 1.000 | 1.01 | 0.960 | 1.000 | Kepler-1987 b |
| 26 | 7841925.2 | K01499.03 | 6.21 | 1.06 | 11.0 | 1.000 | 0.91 | 0.890 | 0.999 | Kepler-865 c |
| 27 | 7869917.2 | K01525.02 | 11.81 | 3.16 | 12.1 | 0.998 | 0.94 | 0.964 | 0.997 | Kepler-880 c |
| 28 | 7939330.1 | K01581.01 | 29.54 | 3.11 | 31.4 | 1.000 | 1.00 | 0.988 | 0.998 | Kepler-896 c |
| 29 | 7983117.1 | K03214.01 | 11.49 | 1.55 | 15.8 | 1.000 | −1.00 | 0.987 | 0.997 | Kepler-1988 b |
| 30 | 8162789.1 | K00521.01 | 10.16 | 4.76 | 102.0 | 1.000 | 0.98 | 0.903 | 0.992 | Kepler-1989 b |
| 31 | 8456679.1 | K00102.01 | 1.74 | 3.91 | 432.9 | 1.000 | 0.88 | 0.986 | 0.997 | Kepler-1990 b |
| 32 | 8456679.2 | K00102.02 | 4.07 | 1.19 | 22.5 | 1.000 | 0.88 | 0.998 | 0.997 | Kepler-1990 c |
| 33 | 8780959.3 | K03741.04 | 9.63 | 2.38 | 19.7 | 1.000 | 0.96 | 0.991 | 0.997 | Kepler-1518 c |
| 34 | 8804283.1 | K01276.01 | 22.79 | 2.91 | 40.1 | 1.000 | 0.93 | 0.977 | 0.996 | Kepler-1991 c |
| 35 | 8804283.2 | K01276.02 | 13.26 | 1.85 | 14.6 | 1.000 | 0.93 | 0.989 | 0.997 | Kepler-1991 b |
| 36 | 8827575.2 | K03052.02 | 15.61 | 1.03 | 11.3 | 1.000 | 1.07 | 0.953 | 0.995 | Kepler-1992 b |
| 37 | 8890924.1 | K04269.01 | 26.66 | 4.43 | 11.7 | 1.000 | 1.00 | 0.914 | 0.993 | Kepler-1993 b |
| 38 | 9071593.2 | K02257.02 | 59.28 | 1.88 | 10.7 | 1.000 | 1.09 | 0.878 | 0.994 | Kepler-1162 c |
| 39 | 9100953.2 | K04500.02 | 44.99 | 2.21 | 10.7 | 1.000 | 1.03 | 0.980 | 1.000 | Kepler-1610 c |
| 40 | 9117416.2 | K03425.02 | 3.16 | 1.69 | 13.8 | 1.000 | 1.14 | 0.987 | 1.000 | Kepler-1921 c |
| 41 | 9205938.2 | K02162.02 | 199.67 | 1.68 | 11.1 | 1.000 | 0.91 | 0.922 | 0.996 | Kepler-1126 c |
| 42 | 9285265.2 | K03410.02 | 61.57 | 3.13 | 10.9 | 1.000 | 1.11 | 0.958 | 0.997 | Kepler-1491 c |
| 43 | 9602613.1 | K02612.01 | 4.61 | 0.62 | 11.4 | 1.000 | 0.96 | 0.968 | 0.996 | Kepler-1994 b |
| 44 | 9634821.1 | K02037.01 | 73.76 | 4.62 | 29.9 | 1.000 | 1.12 | 0.830 | 0.997 | Kepler-1995 b |
| 45 | 9636135.2 | K01498.02 | 2.42 | 1.77 | 13.3 | 1.000 | 0.90 | 0.966 | 0.997 | Kepler-864 c |
| 46 | 9729691.2 | K01751.02 | 21.00 | 5.16 | 27.8 | 1.000 | 0.97 | 0.873 | 0.993 | Kepler-949 c |
| 47 | 9758089.1 | K01871.01 | 92.73 | 2.81 | 31.7 | 1.000 | 1.08 | 0.955 | 0.995 | Kepler-1996c |

**Table 13**
(Continued)

| Number | TCE KIC | KOI Name | Period (days) | Radius (Re) | MES | Pos. Prob. | RUWE | ExoMiner V1.2 | Multiplicity Score | Kepler Name |
|---|---|---|---|---|---|---|---|---|---|---|
| 48 | 9758089.2 | K01871.02 | 32.38 | 2.29 | 26.3 | 1.000 | 1.08 | 0.948 | 0.995 | Kepler-1996 b |
| 49 | 9785921.1 | K03372.01 | 26.76 | 2.44 | 11.1 | 0.993 | 1.02 | 0.975 | 0.996 | Kepler-1997 b |
| 50 | 9839821.2 | K02012.02 | 180.93 | 2.37 | 12.0 | 0.999 | 0.97 | 0.917 | 0.996 | Kepler-1052 c |
| 51 | 9896018.2 | K02579.02 | 3.60 | 1.72 | 12.8 | 1.000 | 1.00 | 0.964 | 0.999 | Kepler-1859 c |
| 52 | 10055126.3 | K01608.03 | 232.05 | 2.07 | 12.5 | 1.000 | 1.12 | 0.909 | 0.999 | Kepler-311 d |
| 53 | 10141900.1 | K01082.01 | 6.50 | 1.97 | 15.9 | 1.000 | 0.91 | 0.991 | 1.000 | Kepler-763 d |
| 54 | 10141900.3 | K01082.02 | 4.10 | 1.64 | 11.0 | 1.000 | 0.91 | 0.988 | 1.000 | Kepler-763 c |
| 55 | 10189546.1 | K00427.01 | 24.61 | 4.55 | 53.9 | 1.000 | 1.15 | 0.969 | 1.000 | Kepler-549 d |
| 56 | 10265898.2 | K00732.03 | 5.25 | 1.52 | 10.6 | 1.000 | 1.01 | 0.749 | 0.995 | Kepler-656 c |
| 57 | 10350571.1 | K01175.02 | 17.16 | 1.77 | 13.0 | 1.000 | 1.01 | 0.996 | 1.000 | Kepler-784 c |
| 58 | 10387742.1 | K02583.01 | 3.03 | 0.69 | 12.9 | 1.000 | 1.01 | 0.968 | 0.996 | Kepler-1998 b |
| 59 | 10460984.3 | K00474.03 | 94.89 | 4.00 | 19.5 | 1.000 | 1.10 | 0.976 | 1.000 | Kepler-164 e |
| 60 | 10843431.1 | K07378.01 | 8.74 | 3.55 | 12.1 | 1.000 | 1.00 | 0.884 | 0.991 | Kepler-1999 b |
| 61 | 10973664.2 | K00601.02 | 11.68 | 4.38 | 43.0 | 1.000 | 1.03 | 0.943 | 1.000 | Kepler-618 d |
| 62 | 11122894.2 | K01426.03 | 150.02 | 7.12 | 92.4 | 1.000 | 0.94 | 0.754 | 0.997 | Kepler-297 d |
| 63 | 11450414.3 | K01992.03 | 85.52 | 2.63 | 11.2 | 1.000 | 1.02 | 0.947 | 1.000 | Kepler-347 d |
| 64 | 11618601.2 | K03022.02 | 5.05 | 1.59 | 10.6 | 0.999 | 1.05 | 0.879 | 0.994 | Kepler-1894 c |
| 65 | 11752906.1 | K00253.01 | 6.38 | 3.32 | 55.2 | 1.000 | 1.01 | 0.980 | 0.996 | Kepler-2000 b |
| 66 | 11752906.2 | K00253.02 | 20.62 | 2.20 | 12.3 | 1.000 | 1.01 | 0.989 | 0.997 | Kepler-2000 c |
| 67 | 11810124.2 | K03344.01 | 11.60 | 2.24 | 11.7 | 1.000 | 1.14 | 0.989 | 1.000 | Kepler-1471 c |
| 68 | 12061969.1 | K02061.01 | 14.09 | 4.08 | 15.9 | 1.000 | 0.94 | 0.930 | 0.994 | Kepler-2001 c |
| 69 | 12061969.2 | K02061.02 | 1.09 | 1.19 | 10.5 | 1.000 | 0.94 | 0.973 | 0.996 | Kepler-2001 b |

(This table is available in machine-readable form.)

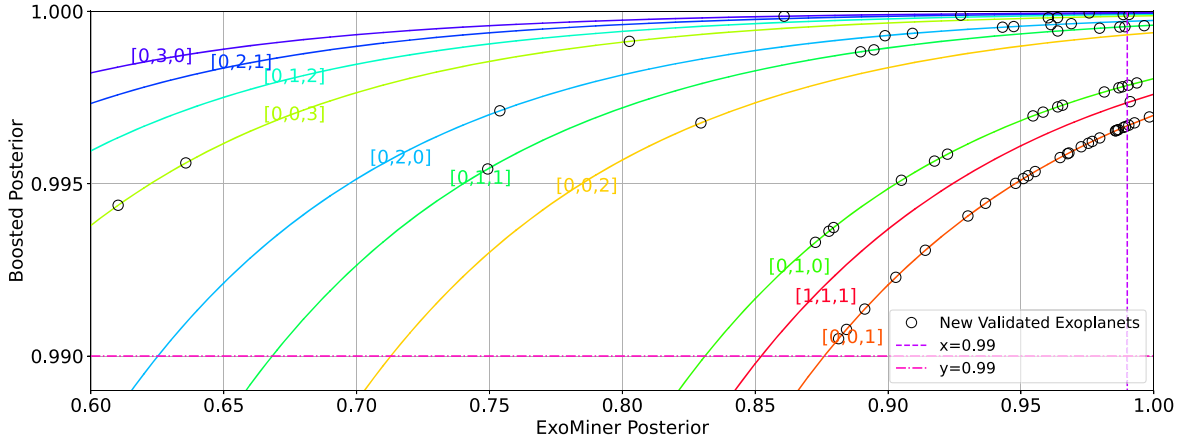


**Figure 6.** The newly validated exoplanets are shown using black circles. The mapping lines between ExoMiner scores and the boosted scores for each tuple $[N_{\mathrm{FPs}}(x),\ N_{\mathrm{CPs}}(x),\ N_{\mathrm{UNKs}}(x)]$ setting are displayed using solid lines. Lines are only plotted for tuples that are represented in the unknown KOIs.

was boosted to 0.996 by the multiplicity boost information. The third (non-KOI) TCE corresponds to a 2% deep (diluted) single transit or eclipse on a neighboring star $\sim 6''$ away.

3. K00102.01 and K00102.02: There are two TCEs/KOIs in this system. We validate both of them. This is a double star system with a separation of $\sim 2\farcs8$ and a delta-magnitude of 1.5. In addition, there is a brighter star (KIC 8456687, $Kp = 10$) $16''$ south of the target star. Both KOIs fail the difference image centroid test relative to the OOT images due to the presence of this bright star. However, the difference image centroids relative to the KIC position for K00102.01 place the signal on the target star, ruling out the fainter companion at more than $30\sigma$. Although the difference images for K00102.02 are much noisier due to the lower strength signal, the difference images themselves indicate the transits fall on target. Kepler photometry of KIC 8456687 shows no transit or eclipse-like signal at the four day period of K00102.02, ruling out the flux contamination from that target as the source. The Kepler follow-up observing program[8] carried out extensive observations on this early KOI and found no evidence of any additional companion brighter than a delta-magnitude of +7 at $3''$ separation or greater. This indicates that a background source for this signal would require a stellar eclipse. However, the ExoMiner V1.2 score for K00102.02 of 0.998 strongly favors a planetary transit. The ExoMiner V1.2 scores for K00102.01 and K00102.02 were 0.986 and 0.998, which were boosted to 0.996 and 0.997, respectively.
4. K01276.01 and K01276.02: There are two TCEs/KOIs for this system. We validate both. There is a brighter star (KIC 8804292, $Kp = 13.3$) $\sim 10''$ to the southeast of the target star that impacts the OOT centroids; however, the difference image centroids are consistent with being on target relative to the catalog star position. The ExoMiner V1.2 scores for these KOIs were 0.977 and 0.989, which were boosted to 0.996 and 0.997, respectively.
5. K02612.01: There are two TCEs/KOIs for this system. We validate one. The target is a bright star, marginally saturated at $Kp = 11.8$. The observed centroid offsets are along the brightest pixel column, which is consistent with the saturation bleed, and speckle and adaptive optics imaging from Kepler follow-up observations find no nearby companions. The ExoMiner V1.2 score for this KOI was 0.968, which was boosted to 0.996.
6. K01871.01 and K01871.02: There are two TCEs/KOIs for this system. We validate both. This is a double star consisting of the target star ($Kp = 14.9$) and a companion (KIC 9758087, $Kp = 14.6$) $\sim 5''$ to the northwest. The difference image centroids for both K01871.01 and K01871.02 are on the KIC position of the target star and rule out the companion at more than $10\sigma$. The ExoMiner V1.2 scores for K01871.01 and K01871.02 were 0.955 and 0.947, respectively, which were both boosted to 0.995.
7. K00253.01 and K00253.02: This is a double star system of two $Kp = 15$ stars with a separation of $5''$. The second star (KIC 11752908) has a FP K02651.01 that triggered off of the signal from K00253.01. The difference image centroids for both K00253.01 and K00253.02 are within $\sim 0\farcs5$ of the KIC position of the target star and rule out the companion at more than $10\sigma$. The ExoMiner V1.2 scores for these KOIs were 0.980 and 0.989, which were boosted to 0.996 and 0.997, respectively.
8. K02037.01: There are three TCEs/KOIs for this system. While there is a nearby brighter star (KIC 9634819, $Kp = 14.2$) $10''$ to the north, the difference image centroids for all three KOIs are on the KIC position of the target star and rule out the companion. The ExoMiner V1.2 score for this KOI was 0.830, which was boosted to 0.996 by multiplicity information given that there are two other unknown KOIs for this system (yellow line on Figure 6).

As discussed in Section 5.1, K01082 is the only system with three unknown KOIs and one CP. We validate K01082.01 and K01082.2 in this work. The only remaining KOI for K01082 not validated by this work is K01082.04, which has a score >0.99, but we decline to validate because of its low-MES value (MES = 7.65). K01082.04 is listed in Table 14. Another interesting new validated planet is K03741.04. There are four TCEs/KOIs for K03741, which include one CP, one FP, and two unknown KOIs. This is the only planet that we validate that has a host star with an FP (black circle on red line for scenario [1,1,1] in Figure 6).

[8] https://exofop.ipac.caltech.edu/

**Table 14**
List of low-MES KOIs (MES <10.5) with Boosted ExoMiner V1.2 Score >0.99, RUWE <1.2, and Positional Probability >0.99

| Number | TCE KIC | KOI Name | Period (days) | Radius (Re) | MES | Pos. Prob. | RUWE | ExoMiner V1.2 | Multiplicity Score |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 4860678.2 | K01602.02 | 3.03 | 1.83 | 10.4 | 1.000 | 0.94 | 0.960 | 0.997 |
| 2 | 6209677.2 | K01750.02 | 2.54 | 1.01 | 10.3 | 1.000 | 1.08 | 0.854 | 0.992 |
| 3 | 10717241.2 | K00430.02 | 9.34 | 0.98 | 10.2 | 1.000 | 1.05 | 0.831 | 0.990 |
| 4 | 5903749.1 | K03029.01 | 18.98 | 2.49 | 10.2 | 1.000 | −1.00 | 0.996 | 0.997 |
| 5 | 4164922.2 | K03864.02 | 18.26 | 1.07 | 10.2 | 1.000 | 1.04 | 0.999 | 0.998 |
| 6 | 7449554.2 | K02357.02 | 15.90 | 4.82 | 10.2 | 1.000 | 1.11 | 0.994 | 0.997 |
| 7 | 3230805.1 | K03068.01 | 3.92 | 1.00 | 10.2 | 1.000 | 0.94 | 0.986 | 0.997 |
| 8 | 7692093.1 | K03337.01 | 11.17 | 1.66 | 10.2 | 1.000 | 1.02 | 0.997 | 0.997 |
| 9 | 8874090.2 | K01404.02 | 18.91 | 1.19 | 10.1 | 1.000 | 1.06 | 0.997 | 0.998 |
| 10 | 9002538.2 | K03196.02 | 6.88 | 0.81 | 10.1 | 1.000 | 0.95 | 0.996 | 0.997 |
| 11 | 8760040.2 | K02963.02 | 7.35 | 1.54 | 10.1 | 1.000 | 1.07 | 0.991 | 0.998 |
| 12 | 9489524.4 | K02029.04 | 4.79 | 0.78 | 9.9 | 1.000 | 1.08 | 0.998 | 1.000 |
| 13 | 5511659.1 | K04541.01 | 8.53 | 2.84 | 9.9 | 1.000 | 0.94 | 0.987 | 0.997 |
| 14 | 8644365.2 | K03384.01 | 10.55 | 1.30 | 9.9 | 1.000 | 1.08 | 0.889 | 0.994 |
| 15 | 9836563.1 | K04421.01 | 4.73 | 0.73 | 9.8 | 1.000 | 0.96 | 0.883 | 0.991 |
| 16 | 5621333.2 | K03341.02 | 11.55 | 1.66 | 9.8 | 1.000 | 1.04 | 0.950 | 0.995 |
| 17 | 6436505.1 | K06707.01 | 43.54 | 1.82 | 9.8 | 1.000 | 1.02 | 0.993 | 0.997 |
| 18 | 6871071.4 | K02220.04 | 7.66 | 1.83 | 9.8 | 1.000 | 0.94 | 0.998 | 1.000 |
| 19 | 9093086.1 | K06191.01 | 9.70 | 1.67 | 9.7 | 1.000 | 0.99 | 0.953 | 0.995 |
| 20 | 4472818.2 | K03878.02 | 15.36 | 1.20 | 9.7 | 1.000 | 0.89 | 0.889 | 0.994 |
| 21 | 7289317.2 | K02450.02 | 7.19 | 1.32 | 9.7 | 1.000 | 0.96 | 0.996 | 0.998 |
| 22 | 4645174.2 | K03437.02 | 34.75 | 2.53 | 9.6 | 1.000 | 1.04 | 0.962 | 0.997 |
| 23 | 7825899.3 | K00896.03 | 28.87 | 1.70 | 9.6 | 1.000 | 0.92 | 0.994 | 1.000 |
| 24 | 6705026.2 | K03374.02 | 34.11 | 2.09 | 9.5 | 1.000 | 1.03 | 0.895 | 0.995 |
| 25 | 7289338.1 | K03420.02 | 12.50 | 2.73 | 9.5 | 1.000 | 0.94 | 0.990 | 0.997 |
| 26 | 4770617.3 | K02243.03 | 31.45 | 3.42 | 9.4 | 1.000 | 0.99 | 0.787 | 0.998 |
| 27 | 9967009.1 | K03462.01 | 12.43 | 2.01 | 9.4 | 1.000 | 0.98 | 0.938 | 0.994 |
| 28 | 8168187.2 | K02209.02 | 35.50 | 2.10 | 9.4 | 1.000 | 1.03 | 0.886 | 0.994 |
| 29 | 2854181.2 | K02232.02 | 12.84 | 1.95 | 9.3 | 1.000 | 0.97 | 0.989 | 0.998 |
| 30 | 9896018.3 | K02579.03 | 10.30 | 1.86 | 9.3 | 1.000 | 1.00 | 0.996 | 1.000 |
| 31 | 5003670.2 | K04524.02 | 4.53 | 1.23 | 9.3 | 1.000 | 0.86 | 0.887 | 0.999 |
| 32 | 11968463.4 | K02433.04 | 27.90 | 2.11 | 9.2 | 0.999 | 0.94 | 0.982 | 1.000 |
| 33 | 7106173.1 | K03083.01 | 10.18 | 0.92 | 9.2 | 1.000 | 0.98 | 0.871 | 0.998 |
| 34 | 12256520.2 | K02264.02 | 7.25 | 1.15 | 9.2 | 1.000 | 0.87 | 0.857 | 0.992 |
| 35 | 5769810.1 | K04913.02 | 8.97 | 3.39 | 9.2 | 0.999 | 1.12 | 0.900 | 0.992 |
| 36 | 6508221.3 | K00416.03 | 9.75 | 1.54 | 9.2 | 1.000 | 0.95 | 0.792 | 0.998 |
| 37 | 11918099.2 | K00780.02 | 7.24 | 1.81 | 9.2 | 1.000 | 0.96 | 0.979 | 0.998 |
| 38 | 4735826.2 | K03184.03 | 4.02 | 0.67 | 9.1 | 1.000 | 0.90 | 0.974 | 0.996 |
| 39 | 9655711.1 | K06209.01 | 12.71 | 2.29 | 9.1 | 1.000 | 0.97 | 0.942 | 0.995 |
| 40 | 7285757.2 | K03271.02 | 7.42 | 1.26 | 9.0 | 1.000 | 1.01 | 0.992 | 0.997 |
| 41 | 5177859.2 | K04246.02 | 8.76 | 1.36 | 9.0 | 1.000 | 1.10 | 0.985 | 0.998 |
| 42 | 1432789.2 | K00992.02 | 4.58 | 1.64 | 8.9 | 1.000 | 1.00 | 0.903 | 0.995 |
| 43 | 10290666.2 | K00332.02 | 6.87 | 0.88 | 8.8 | 1.000 | 1.14 | 0.977 | 0.998 |
| 44 | 4157325.4 | K01860.04 | 24.84 | 1.57 | 8.8 | 1.000 | 0.95 | 0.575 | 0.997 |
| 45 | 6543893.3 | K01627.03 | 3.81 | 3.28 | 8.8 | 1.000 | 1.03 | 0.811 | 0.998 |
| 46 | 7047922.2 | K01899.02 | 10.52 | 1.19 | 8.7 | 1.000 | 1.13 | 0.947 | 0.997 |
| 47 | 10000941.2 | K04146.02 | 2.57 | 0.69 | 8.7 | 1.000 | 1.07 | 0.942 | 0.997 |
| 48 | 1996180.2 | K02534.02 | 5.42 | 1.35 | 8.7 | 1.000 | 1.07 | 0.866 | 0.993 |
| 49 | 7107802.2 | K02420.02 | 5.47 | 1.19 | 8.6 | 1.000 | 1.05 | 0.979 | 0.998 |
| 50 | 5903749.2 | K03029.02 | 6.35 | 1.42 | 8.6 | 1.000 | −1.00 | 0.916 | 0.993 |
| 51 | 9100953.3 | K04500.03 | 14.75 | 2.01 | 8.6 | 1.000 | 1.03 | 0.953 | 0.999 |
| 52 | 12266636.2 | K01522.02 | 12.65 | 1.14 | 8.5 | 1.000 | 0.98 | 0.914 | 0.996 |
| 53 | 7663405.2 | K01519.02 | 57.13 | 1.93 | 8.5 | 1.000 | 1.02 | 0.975 | 0.998 |
| 54 | 9872283.2 | K01815.02 | 1.75 | 1.56 | 8.4 | 1.000 | 1.10 | 0.844 | 0.991 |
| 55 | 7512982.2 | K01480.02 | 7.00 | 1.47 | 8.4 | 1.000 | 0.99 | 0.975 | 0.998 |
| 56 | 6276791.2 | K04477.01 | 5.30 | 1.28 | 8.4 | 1.000 | 1.04 | 0.872 | 0.993 |
| 57 | 8613535.4 | K02263.03 | 15.59 | 1.15 | 8.4 | 1.000 | 0.98 | 0.740 | 0.997 |
| 58 | 9602613.2 | K02612.02 | 7.57 | 0.61 | 8.3 | 1.000 | 0.96 | 0.895 | 0.992 |
| 59 | 7021681.2 | K00255.02 | 13.60 | 0.79 | 8.3 | 1.000 | 1.03 | 0.965 | 0.999 |
| 60 | 10265898.3 | K00732.02 | 3.30 | 1.44 | 8.2 | 1.000 | 1.01 | 0.966 | 0.999 |
| 61 | 3645438.3 | K04385.03 | 17.37 | 1.83 | 8.2 | 1.000 | 1.05 | 0.827 | 0.999 |
| 62 | 7289338.2 | K03420.01 | 5.77 | 2.10 | 8.2 | 1.000 | 0.94 | 0.950 | 0.995 |
| 63 | 6221385.3 | K06145.03 | 7.31 | 2.41 | 8.2 | 1.000 | 1.08 | 0.947 | 1.000 |
| 64 | 8216763.1 | K04838.01 | 13.30 | 0.95 | 8.1 | 1.000 | 1.00 | 0.808 | 0.996 |

**Table 14**
(Continued)

| Number | TCE KIC | KOI Name | Period (days) | Radius (Re) | MES | Pos. Prob. | RUWE | ExoMiner V1.2 | Multiplicity Score |
|---|---|---|---|---|---|---|---|---|---|
| 65 | 7673841.3 | K02585.03 | 7.88 | 0.90 | 8.1 | 1.000 | 0.94 | 0.824 | 0.999 |
| 66 | 8216763.2 | K04838.03 | 24.07 | 0.97 | 8.1 | 1.000 | 1.00 | 0.961 | 0.999 |
| 67 | 6103377.2 | K03004.02 | 7.04 | 2.81 | 8.1 | 1.000 | 1.06 | 0.974 | 0.997 |
| 68 | 5531953.3 | K01681.04 | 21.91 | 1.05 | 8.0 | 1.000 | −1.00 | 0.919 | 1.000 |
| 69 | 5351250.5 | K00408.05 | 93.80 | 2.48 | 7.9 | 1.000 | 1.01 | 0.576 | 1.000 |
| 70 | 8226050.2 | K01910.02 | 18.38 | 1.36 | 7.8 | 1.000 | 1.02 | 0.874 | 0.993 |
| 71 | 5602588.2 | K02369.03 | 7.23 | 1.55 | 7.7 | 1.000 | 0.96 | 0.948 | 0.997 |
| 72 | 6716545.2 | K02906.03 | 21.94 | 1.37 | 7.7 | 1.000 | 1.05 | 0.908 | 0.999 |
| 73 | 7765528.2 | K01840.02 | 9.39 | 1.94 | 7.7 | 1.000 | 0.96 | 0.875 | 0.993 |
| 74 | 10141900.4 | K01082.04 | 9.66 | 1.77 | 7.7 | 1.000 | 0.91 | 0.559 | 0.994 |
| 75 | 5531953.4 | K01681.03 | 3.53 | 0.88 | 7.6 | 1.000 | −1.00 | 0.977 | 1.000 |
| 76 | 10028792.4 | K01574.04 | 8.98 | 1.84 | 7.4 | 1.000 | 0.90 | 0.896 | 1.000 |
| 77 | 11968463.6 | K02433.07 | 86.43 | 2.98 | 7.2 | 1.000 | 0.94 | 0.901 | 1.000 |

**Note.** This list is sorted by MES value.

(This table is available in machine-readable form.)

Figure 7 displays a scatterplot of the planetary radius versus orbital period for the previously confirmed and validated exoplanets from the Kepler, K2, and TESS missions along with the 69 new Kepler exoplanets that were validated by our multiplicity framework. The distribution of the new exoplanets is consistent with that of the Kepler sample, with periods ranging from ∼0.6 day to over 450 days, and with planetary radii as small as 0.6 $R_\oplus$ to as large as 9.5 $R_\oplus$. Figure 8 shows a scatterplot of the planetary radius versus the energy received by each planet from the Kepler, K2, and TESS missions, and the new validated planets. As with the previous parametric plot, the distribution of the radius versus received energy of the new sample generally follows that of the Kepler sample.

## 6. Caveats

Caveats related to the base classifier (ExoMiner V1.2 in our case), or the multiplicity boost framework can affect the performance of the model and lead to erroneous validations. We have discussed the caveats related to ExoMiner in Valizadegan et al. (2022). Here, we discuss those related to the multiplicity boost framework.

The shape of the multiplicity mapping is affected by the form of the underlying classifier model used for multiplicity boost, i.e., logistic regression here, and is largely dependent on the data set that we use for training the model. The logistic regression makes certain assumptions that might not fully hold for the multiplicity data:

1. Logistic regression requires that the independent variables are linearly related to the log odds. This assumption basically dictates the shape of the mapping between the original score and multiplicity score (Figure 3). Having a different assumption about the log odds will change the shape of the mapping, even though the general behavior will stay the same.

   There are also caveats directly related to the multiplicity boost independent of the multiplicity classifier that we use. This is mainly due to the fact that the multiplicity boost assumes that the candidates are not (1) FPs due to background objects and (2) FAs, as we discuss below.
2. FP due to background objects: The multiplicity boost approach will fail for FPs due to background objects when there are multiple existing planets and unknown KOIs. One example of this situation is K02433.05. Even though K02433.05 is a background FP, the multiplicity approach boosts its very low score to 0.987 because there are three CPs and two unknown KOIs for this system. Fortunately, this does not pass the validation threshold. However, there could exist similar KOIs that pass the validation threshold. Our vetoing conditions mitigate this problem for most stars, although positional probability is not computed for all targets. For K02433.05, not only are there two FP flags, "Stellar Eclipse Flag" and "Centroid Offset Flag" set, but it also has a low MES of 9.09.

   Calculating the rate of misclassification for background FPs can provide great insights. However, this is not easy because we do not have the gold standard labels for background FPs. To provide an estimate, we use the KOI flag for centroid shift, i.e., "Centroid Offset Flag," as the indicator for background FPs. There are a total of 1742 KOIs with this flag set. Out of these 1742 KOIs, ExoMiner V1.2 correctly classifies 1731 KOIs as FPs (only 11 mistakes), resulting in a recall of 99.4% for background FPs. Only one KOI out of these 11 misclassified cases, K00082.06, is in a multi-planet system. As mentioned in Valizadegan et al. (2022) and reported in this manuscript, K00082.06 is incorrectly certified as FP in the Certified False Positive table. After applying the multiplicity boost model, a total of 14 KOIs, including the 11 KOIs originally misclassified by ExoMiner V1.2, are misclassified. Except K00082.06, none of these 14 KOIs pass the validation threshold of 0.99 after the application of multiplicity boost. We also would like to mention that for a background FP to be incorrectly validated by our model, three conditions need to be satisfied: (1) ExoMiner V1.2 gives a high enough score to that KOI (as we explained above, this is very rare); (2) that KOI should be around a multi-planet system with a high enough ExoMiner score to get boosted above 0.99; and (3) all stability and vetoing conditions introduced in Sections 5.3 and 5.4 fail. This can happen in practice but it is highly unlikely.
3. FAs: Given that the multiplicity boost works for planets and FPs, it might boost the score of an FA for scenarios that have multiple planets and unknown KOIs. We do not

**Table 15**
List of KOIs with Boosted ExoMiner V1.2 Score >0.99 That Did Not Pass Positional Probability or RUWE Tests

| Number | TCE KIC | KOI Name | Period (days) | Radius (Re) | MES | Pos. Prob. | RUWE | ExoMiner V1.2 | Multiplicity Score |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 10662202.3 | K00750.03 | 14.52 | 1.55 | 7.5 | 1.000 | 1.36 | 0.914 | 0.999 |
| 2 | 10397751.3 | K02859.05 | 5.43 | 0.76 | 9.4 | 1.000 | 1.65 | 0.949 | 1.000 |
| 3 | 10397751.4 | K02859.04 | 2.91 | 0.53 | 8.6 | 1.000 | 1.65 | 0.988 | 1.000 |
| 4 | 8280511.4 | K01151.04 | 17.45 | 0.87 | 8.4 | 1.000 | 2.01 | 0.948 | 1.000 |
| 5 | 8280511.5 | K01151.05 | 21.72 | 0.92 | 7.8 | 1.000 | 2.01 | 0.871 | 1.000 |
| 6 | 10875007.2 | K04149.02 | 14.71 | 1.68 | 10.6 | 1.000 | 2.05 | 0.916 | 0.993 |
| 7 | 8261920.1 | K02174.01 | 6.69 | 2.66 | 17.9 | 1.000 | 2.82 | 0.975 | 1.000 |
| 8 | 5956656.2 | K01053.02 | 46.25 | 2.48 | 8.8 | 1.000 | 3.33 | 0.984 | 0.998 |
| 9 | 5856571.2 | K01839.02 | 80.41 | 4.41 | 12.9 | 1.000 | 3.85 | 0.929 | 0.996 |
| 10 | 3529290.1 | K03340.02 | 13.73 | 1.08 | 11.4 | 1.000 | 4.09 | 0.919 | 0.993 |
| 11 | 5629353.1 | K06132.01 | 33.32 | 13.31 | 59.2 | 1.000 | 4.39 | 0.787 | 0.997 |
| 12 | 5542466.3 | K01590.03 | 4.75 | 1.63 | 9.6 | 1.000 | 5.07 | 0.988 | 1.000 |
| 13 | 4832837.2 | K00605.02 | 5.07 | 0.73 | 9.5 | 1.000 | 10.81 | 0.973 | 0.997 |
| 14 | 4736569.2 | K01996.02 | 7.07 | 1.09 | 9.1 | 1.000 | 10.88 | 0.969 | 0.997 |
| 15 | 11566064.1 | K00353.01 | 152.11 | 11.67 | 82.5 | 1.000 | 1.42 | 0.833 | 0.998 |
| 16 | 5384713.1 | K03444.02 | 60.33 | 5.25 | 59.2 | 1.000 | 1.95 | 0.959 | 1.000 |
| 17 | 11125797.1 | K03371.01 | 58.13 | 1.55 | 10.8 | 1.000 | 2.90 | 0.964 | 0.997 |
| 18 | 9413156.1 | K04700.01 | 3.83 | 1.32 | 9.3 | 0.981 | 0.90 | 0.903 | 0.992 |
| 19 | 4385148.1 | K02942.01 | 13.84 | 2.40 | 14.5 | 0.975 | 0.92 | 0.982 | 0.998 |
| 20 | 5031857.2 | K01573.02 | 7.14 | 1.31 | 13.6 | 0.974 | 1.08 | 0.954 | 0.997 |
| 21 | 5080636.2 | K01843.02 | 6.36 | 0.69 | 9.9 | 0.962 | 1.19 | 0.851 | 0.992 |
| 22 | 7100673.5 | K04032.05 | 7.24 | 0.88 | 9.0 | 0.961 | 1.19 | 0.986 | 1.000 |
| 23 | 8581240.1 | K03111.01 | 10.77 | 1.05 | 11.9 | 0.959 | 1.07 | 0.990 | 0.997 |
| 24 | 4466677.2 | K01338.02 | 42.04 | 1.73 | 12.1 | 0.950 | 0.96 | 0.934 | 0.999 |
| 25 | 4478168.2 | K00626.02 | 8.03 | 1.13 | 9.9 | 0.941 | 0.90 | 0.985 | 0.998 |
| 26 | 10028792.3 | K01574.03 | 5.83 | 1.79 | 9.3 | 0.940 | 0.90 | 0.750 | 0.999 |
| 27 | 9002538.1 | K03196.01 | 4.96 | 0.71 | 11.0 | 0.938 | 0.95 | 0.967 | 0.996 |
| 28 | 6026737.1 | K02949.01 | 10.17 | 1.15 | 10.0 | 0.925 | 0.89 | 0.997 | 0.997 |
| 29 | 10122538.5 | K02926.05 | 75.73 | 3.55 | 11.4 | 0.917 | 1.06 | 0.356 | 0.997 |
| 30 | 7449554.1 | K02357.01 | 2.42 | 2.64 | 17.8 | 0.908 | 1.11 | 0.985 | 0.997 |
| 31 | 5003670.3 | K04524.03 | 3.34 | 1.08 | 8.4 | 0.887 | 0.86 | 0.833 | 0.998 |
| 32 | 5621333.1 | K03341.01 | 27.10 | 2.30 | 14.7 | 0.865 | 1.04 | 0.994 | 0.997 |
| 33 | 7256914.2 | K04136.02 | 4.03 | 1.20 | 11.3 | 0.856 | 1.02 | 0.971 | 0.996 |
| 34 | 4851530.1 | K01884.01 | 23.08 | 2.07 | 17.6 | 0.846 | 1.21 | 0.969 | 0.996 |
| 35 | 4548011.2 | K04288.02 | 9.09 | 0.92 | 10.0 | 0.839 | 1.08 | 0.813 | 0.998 |
| 36 | 7375348.2 | K00266.02 | 47.74 | 1.79 | 14.1 | 0.795 | 0.93 | 0.991 | 0.998 |
| 37 | 12785320.2 | K00298.02 | 57.38 | 1.70 | 17.4 | 0.794 | 0.98 | 0.994 | 0.998 |
| 38 | 5211199.2 | K02158.02 | 6.68 | 2.12 | 10.1 | 0.777 | 0.98 | 0.996 | 0.998 |
| 39 | 11253711.1 | K01972.01 | 17.79 | 3.48 | 31.7 | 0.749 | 6.49 | 0.949 | 0.997 |
| 40 | 6527078.1 | K04657.01 | 7.58 | 0.64 | 10.2 | 0.740 | 0.92 | 0.985 | 0.997 |
| 41 | 5972334.3 | K00191.03 | 0.71 | 1.26 | 21.1 | 0.740 | 0.91 | 0.899 | 1.000 |
| 42 | 7673192.5 | K02722.05 | 16.53 | 1.24 | 7.8 | 0.710 | 1.05 | 0.984 | 1.000 |
| 43 | 5384713.2 | K03444.03 | 2.64 | 0.64 | 11.4 | 0.653 | 1.95 | 0.973 | 1.000 |
| 44 | 11135694.1 | K04896.02 | 49.54 | 1.84 | 9.3 | 0.643 | 1.04 | 0.983 | 0.996 |
| 45 | 6467363.1 | K02840.01 | 3.68 | 0.90 | 15.4 | 0.621 | 1.03 | 0.992 | 0.997 |
| 46 | 10471621.1 | K02554.02 | 10.27 | 1.07 | 13.8 | 0.618 | −1.00 | 0.992 | 0.997 |
| 47 | 5384713.3 | K03444.01 | 12.67 | 0.84 | 10.1 | 0.614 | 1.95 | 0.994 | 1.000 |
| 48 | 10397751.5 | K02859.03 | 4.29 | 0.61 | 7.4 | 0.589 | 1.65 | 0.557 | 0.999 |
| 49 | 8838950.1 | K02421.01 | 2.27 | 0.62 | 13.1 | 0.585 | 3.57 | 0.892 | 0.991 |
| 50 | 11030475.2 | K02248.03 | 0.76 | 1.26 | 16.8 | 0.579 | −1.00 | 0.921 | 0.999 |
| 51 | 6268648.4 | K01613.03 | 20.61 | 0.95 | 8.1 | 0.540 | 45.45 | 0.810 | 0.997 |
| 52 | 5531953.1 | K01681.01 | 6.94 | 1.18 | 18.4 | 0.540 | −1.00 | 0.997 | 1.000 |
| 53 | 6527078.2 | K04657.02 | 10.43 | 0.74 | 8.9 | 0.520 | 0.92 | 0.969 | 0.996 |
| 54 | 4770174.1 | K02971.01 | 6.10 | 1.58 | 14.3 | 0.514 | 1.04 | 0.948 | 0.995 |
| 55 | 4851530.2 | K01884.02 | 4.78 | 1.89 | 14.3 | 0.461 | 1.21 | 0.987 | 0.997 |
| 56 | 11967788.2 | K04021.02 | 4.93 | 1.42 | 14.0 | 0.416 | 20.55 | 0.987 | 0.997 |
| 57 | 10875007.1 | K04149.01 | 9.55 | 1.72 | 13.9 | 0.350 | 2.05 | 0.988 | 0.997 |
| 58 | 8008067.3 | K00316.03 | 7.31 | 2.18 | 33.3 | 0.268 | 1.07 | 0.739 | 0.997 |
| 59 | 5542466.1 | K01590.01 | 12.89 | 2.36 | 18.7 | 0.099 | 5.07 | 0.980 | 1.000 |

**Note.** This list is sorted by positional probability.

(This table is available in machine-readable form.)

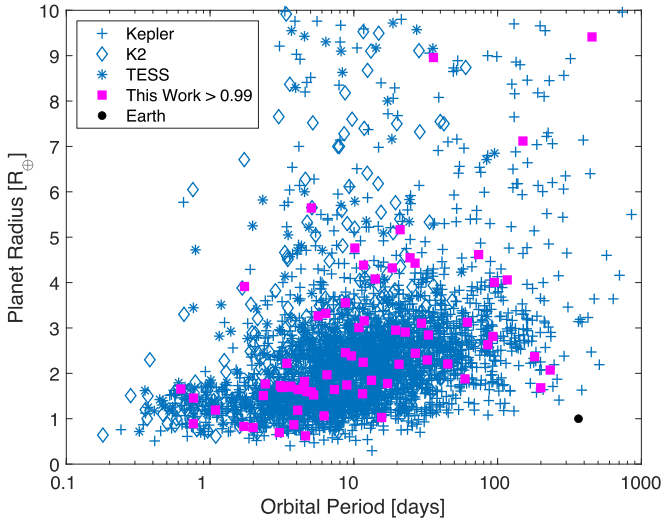


**Figure 7.** Planet radius vs. orbital period for confirmed transiting planets and validated planets in this paper. The CPs are indicated by the discovery source with Kepler indicated by "+," K2 by a diamond "◊," and TESS by "*." `ExoMiner` V1.2-boosted validated planets are indicated by a magenta square. For reference, Earth is indicated by a black circle.

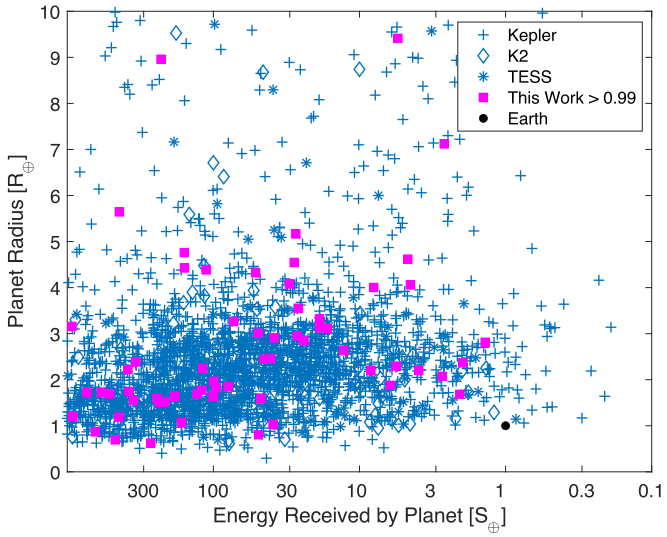


**Figure 8.** Planet radius vs. energy received by planet for confirmed transiting planets and the validated planets in this paper. The symbols are the same as in Figure 7.

have any examples of this in the labeled KOIs but, as we discussed in Section 5.5, the unknown K00408.05 with multiplicity score >0.99 has the "Not Transit-Like Flag" on. It is also a KOI with MES <10.5. So our vetoing conditions such as MES >10.5 and KOI FP flags also help for cases in this category.

To provide some insights regarding the chances of validating such observations, we would like to note that `ExoMiner` V1.2 is highly accurate when it comes to the NTP instances. Out of 24474 NTPs in the Kepler data, there are only five with `ExoMiner V1.2` score >0.5. This results in a 99.98% recall for the NTP set. It is also highly accurate for FA objects in the KOI tables. In the CFP table of Bryson et al. (2017), there are a total of 302 FAs out of which `ExoMiner` V1.2 gives a score >0.5 to only two of them: K03226.01 with a score 0.84 and K02768.03 with a score 0.97, both of which are on systems with a single KOI (note that there is only one KOI for K02768 in Q1-Q17 DR25). Interestingly, K02768.03 is a transiting planet that has been confirmed in Q1-Q16 and incorrectly labeled as FA.

KOI 3226.01 was dispositioned as an FA by the Kepler False Positive Working Group based on Q1-Q16 data processing, and using that data this KOI violated the FA criteria by having large oscillations at the same period and with a similar shape as the transit signal Bryson et al. (2017). However, these oscillations diminished in the final DR25 processing, so that they no longer violate the false alarm criterion. Therefore, we think of this KOI as an unclear case.

Overall, to have a validated FA KOI, it needs to satisfy the following three conditions: (1) `ExoMiner` V1.2 gives a high score to that FA (this is very rare, as we explained above); (2) the FA KOI should be around a system with enough other planets to be boosted to the validation threshold of 0.99 by the multiplicity boost framework; and (3) it needs to pass the vetoing and stability conditions explained in Sections 5.3 and 5.4. Having a FA that meets all these conditions is possible but unlikely.

4. Application to TESS: The same approach can also be beneficial when considering TESS exoplanet results. While the TESS pixel scale is ∼25 times bigger in area than that for Kepler (21″ versus 4″ on a side), the average impact of background blends in TESS data and Kepler data are roughly comparable. This is due to the fact that TESS's exoplanet targets are ∼5 mag brighter than Kepler's, meaning that there are ∼32 times fewer background stars at a comparable delta-magnitude. Of course, this average estimate breaks down in regions of high crowding (e.g., the Galactic plane or Galactic bulge), or in the cases of stars dimmer than Tmag ∼ 11, and results from these scenarios would require more careful checking and attention to other tests that provide information on background contamination, such as centroid and difference image analysis.

## 7. Conclusions

We introduced a new multiplicity boost framework that can be applied to any existing transit-signal classifier to boost its performance. Our framework does not require an existing classifier to retrained or redesigned, and can be applied to the output scores of a given classifier to improve its performance using the multiplicity information. We applied our framework to multiple state-of-the-art transit-signal classifiers to demonstrate that multiplicity information improves their performance. Furthermore, we applied it to an improved version of `ExoMiner` and validated 69 new exoplanets for systems with more than one candidate.

This work is dedicated to the Women, Life, Freedom movement in Iran. We would like to symbolically name four of the newly validated exoplanet, i.e., K01358.01, K01358.02, K01358.03, and K01358.04, to Zan, Zendegi, Azadi,[9] and Iran, respectively.

H.V. and M.M. are supported through NASA NAMS contract NNA16BD14C, TESS GI Cycle 4 contract 80NSSC22K0184, and NASA ROSES XRP proposal 22-XRP22_2-0173. D.C. and J.T. are supported through NASA Cooperative Agreement

[9] Persian words for woman, life, and freedom, respectively.

80NSSC21M0079. We would like to thank many people who directly or indirectly contributed to this work. This paper includes data collected by the Kepler mission and obtained from the MAST data archive at the Space Telescope Science Institute (STScI). Funding for the Kepler mission was provided by the NASA Science Mission Directorate. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center for the production of the Kepler SOC data products and for training our deep learning model, `ExoMiner` V1.2. This research has made use of the Exoplanet Follow-up Observation Program (ExoFOP; DOI:10.26134/ExoFOP5) website, which is operated by the California Institute of Technology, under contract with NASA under the Exoplanet Exploration Program.

## ORCID iDs

Hamed Valizadegan https://orcid.org/0000-0001-6732-0840
Miguel J. S. Martinho https://orcid.org/0000-0002-2188-0807
Jon M. Jenkins https://orcid.org/0000-0002-4715-9460
Douglas A. Caldwell https://orcid.org/0000-0003-1963-9616
Joseph D. Twicken https://orcid.org/0000-0002-6778-7552
Stephen T. Bryson https://orcid.org/0000-0003-0081-1797

## References

Ansdell, M., Ioannou, Y., Osborn, H. P., et al. 2018, ApJL, 869, L7
Armstrong, D. J., Gamper, J., & Damoulas, T. 2021, MNRAS, 504, 5327
Armstrong, D. J., Pollacco, D., & Santerne, A. 2016, MNRAS, 465, 2634
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Bishop, C. M. 2006, Pattern Recognition and Machine Learning (Berlin: Springer)
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, AJ, 142, 112
Bryson, S. T., Abdul-Masih, M., Batalha, N., et al. 2017, Kepler Science Document KSCI-19093-003
Bryson, S. T., Jenkins, J. M., Gilliland, R. L., et al. 2013, PASP, 125, 889
Bryson, S. T., & Morton, T. D. 2017, Kepler Science Document KSCI-19108-001, 16
Christiansen, J. 2022, in IAU Focus Meeting 10: Synergy of Small Telescopes and Large Surveys for Solar System and Exoplanetary Bodies Research (Busan: IAU)
Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, PASP, 124, 1279
Coughlin, J. L. 2017, KSCI-19114-001
Coughlin, J. L., Mullally, F., Thompson, S. E., et al. 2016, ApJS, 224, 12
Cox, D. R. 1958, J. R. Stat. Soc., B, 20, 215
Díaz, R. F., Damiani, C., Deleuil, M., et al. 2013, A&A, 551, L9
Díaz, R. F., Montagnier, G., Leconte, J., et al. 2014, A&A, 572, A109
Dietrich, J., & Apai, D. 2020, AJ, 160, 107
Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, ApJS, 211, 2
Jenkins, J. M. 2020, KSCI-19081-003
Jenkins, J. M., McCauliff, S., Burke, C., et al. 2014, in IAU Symp. 293, Formation, Detection, and Characterization of Extrasolar Habitable Planets, ed. N. Haghighipour (Cambridge: Cambridge Univ. Press), 94
Jontof-Hutter, D., Wolfgang, A., Ford, E. B., et al. 2021, AJ, 161, 246
Latham, D. W., Rowe, J. F., Quinn, S. N., et al. 2011, ApJL, 732, L24
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021, A&A, 649, A2
Lissauer, J. J., Marcy, G. W., Bryson, S. T., et al. 2014, ApJ, 784, 44
Lissauer, J. J., Marcy, G. W., Rowe, J. F., et al. 2012, ApJ, 750, 112
Masuda, K., & Tamayo, D. 2020, AJ, 160, 224
McCauliff, S. D., Jenkins, J. M., Catanzarite, J., et al. 2015, ApJ, 806, 6
Morton, T. D. 2012, ApJ, 761, 6
Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86
Niraula, P., Shporer, A., Wong, I., & de Wit, J. 2022, AJ, 163, 172
Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, ApJ, 784, 45
Shallue, C. J., & Vanderburg, A. 2018, AJ, 155, 94
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2018, PASP, 130, 064502
Twicken, J. D., Jenkins, J. M., Seader, S. E., et al. 2016, AJ, 152, 158
Valizadegan, H., Martinho, M. J. S., Wilkens, L. S., et al. 2022, ApJ, 926, 120
Vapnik, V. N. 1998, Statistical Learning Theory (New York: Wiley-Interscience), https://www.wiley.com/en-us/Statistical+Learning+Theory-p-9780471030034
Vissapragada, S., Jontof-Hutter, D., Shporer, A., et al. 2020, AJ, 159, 108