LIFE BEYOND EARTH:

EXPLORING MACHINE LEARNING AND EARTH SIMILARITY INDEX (ESI) FOR EXOPLANET HABITABILITY ESTIMATION

Kishor Baniya[1,2], Vlad Veksler[1], Yangxing Yang[1], Suresh Bhattarai[2]

Caldwell University[1], Nepal Astronomical Society (NASO)[2]

MOTIVATION

- High volume of real-time data makes conventional analysis challenging. [1]
- Earth Similarity Index (ESI) offers a compact metric to rank potentially habitable worlds, but its empirical robustness is underexplored. [2]
- Understanding the PHL-adopted weightage of habitability parameters for ESI [3] with a data-driven model could unlock a new avenue.

METHOD

- Trained a Random Forest classifier in Python using Habitable World Catalog (HWC)[4] and a synthetic non-habitable exoplanetary database. (Total 800 records.)
- Queried the NASA Exoplanet Archive [5] using Astroquery [6] to classify habitability of real-time discoveries.
- Compared the feature importance rankings from the ML model to the PHL's theoretical weights used in the ESI.
- Analyzed convergence and divergence between ML predictions and ESI scores, using verified planetary assessments for validation.

ANALYSIS & RESULTS

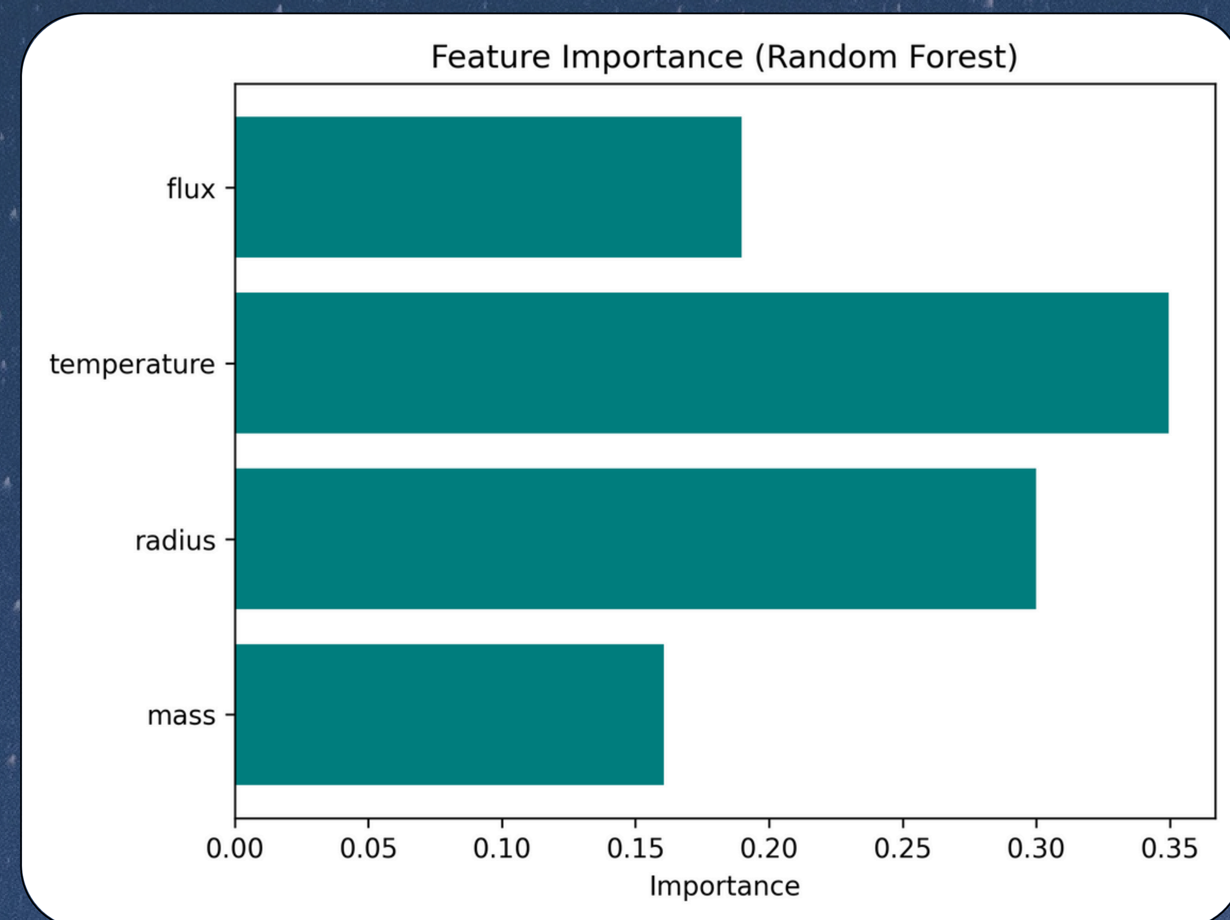

Figure 1: Feature Importance on Habitability Determination

name	mass	radius	temperature	flux	habitable?	ESI(%)
K2-3 d	2.20	1.458	305.2	1.44	1	45.63
TRAPPIST-1 f	0.68	1.045	219	0.382	1	34.82
TRAPPIST-1 g	1.34	1.127	198.6	0.258	1	26.29
Kepler-22 b	9.10	2.1	279	1.013	1	12.6
LP 890-9 c	25.30	1.367	272	0.906	1	12.47
LHS 1140 b	5.60	1.73	226	0.43	1	11.39
LHS 1140 b	6.65	1.43	230	0.46	1	10.98
PH2 b	87.08	9.33708	295.1	1.2	1	10.95
LHS 1140 b	6.98	1.727	235	0.503	1	10.5
TOI-2134 c	41.89	7.27	306	1.4	1	9.88
K2-18 b	8.63	2.61	254.9	1.005	1	9.52
Kepler-1704 b	1322.17	11.9488	253.8	0.342	1	8.15
Kepler-553 c	2129.45	11.5789	251	0.589	1	7.89
Kepler-1661 b	17.00	3.87	243	0.88	1	7.16
Kepler-1654 b	158.92	9.18017	206	0.3	1	4.26
EPIC 248847494 b	4131.79	12.442	183	0.19	1	2.87
WD 1856+534 b	4386.05	10.4	163	0.181	1	1.92

Figure 2: List of potentially habitable exoplanets (with their respective ESI score)

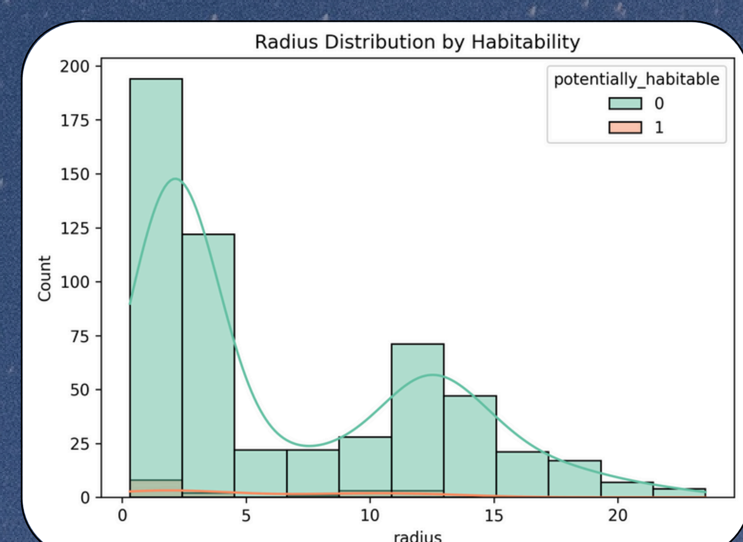

Figure 3: Occurrence of Habitable (1) or non-Habitable (0) planets by planetary radius.

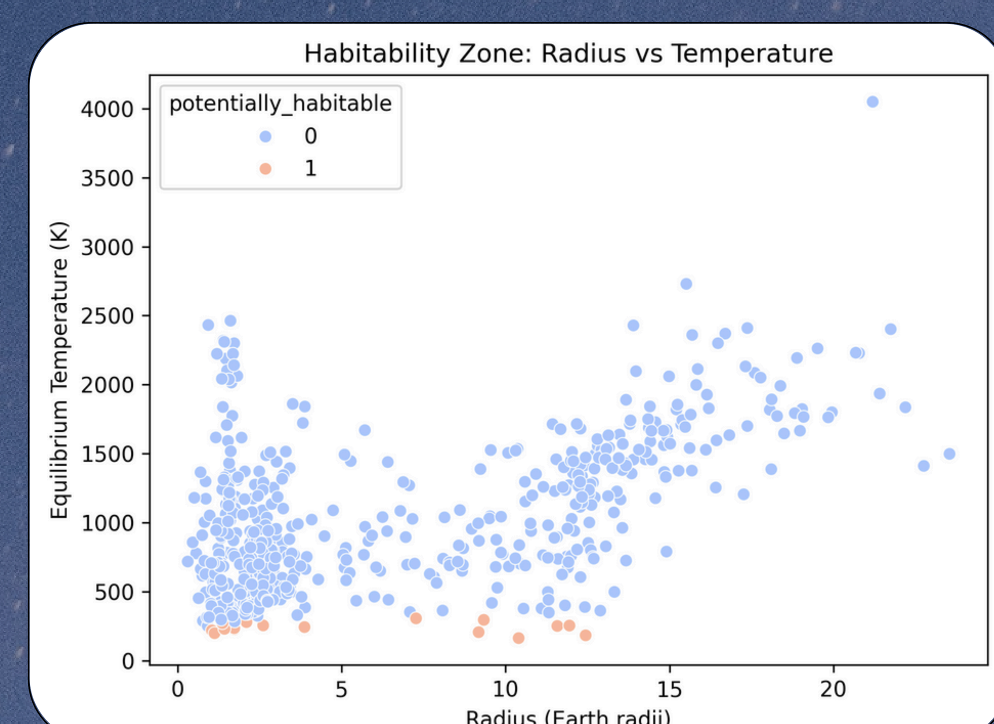

Figure 4: Radius by Temperature plot in the occurrence of habitability.

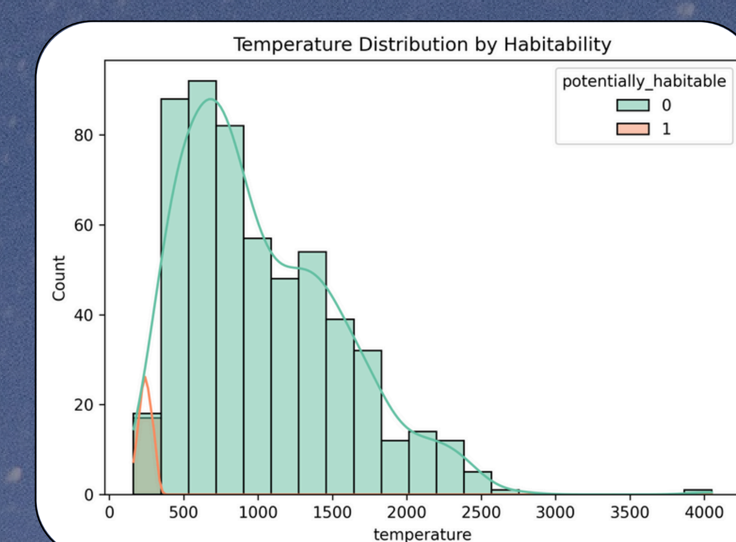

Figure 5: Occurrence of Habitable (1) or non-Habitable (0) planets by planetary equilibrium temperature.

KEY FINDINGS

P lanetary Property	Reference Value	Weight Exponent
Mean Radius	1.0 Eu	0.57
Bulk Density	1.0 Eu	1.07
Escape velocity	1.0 Eu	0.70
Surface Temperature	288 K	5.58

Figure 6: PHL-adopted metric for ESI calculation. [3]

- Interestingly, in the feature importance results (Figure 1), surface temperature and planetary radius emerged as the most influential features, matching PHL's theoretical weights on habitability parameters.
- A well-defined cluster of potentially habitable planets lies within 0.5–2.5 Earth radii and 200–400 K, providing a refined target envelope.
- However, not all exoplanets flagged as potentially habitable (Figure 2) are confirmed candidates—e.g., K2-3d, a likely tidally locked planet with a high H-He composition, shows high ESI approval.
- While machine learning complements theoretical models, reliance solely on current static parameters may limit it from reaching its potential. A more dynamic approach, such as data imputation for incomplete or evolving datasets, may therefore enhance future habitability assessments.

REFERENCES

[1] *Shallue, C. J. & Vanderburg, A. (2018), AJ 155, 94.*
[2] *Brandl, C. (2015), Exoplanet Habitability and the ESI, ResearchGate.*
[3] *Planetary Habitability Laboratory, UPR (2025), ESI*
[4] *Méndez, A. (2023), Habitable World Catalog (HWC), PHL, UPR-Arecibo. phl.upr.edu/hwc*
[5] *NASA Exoplanet Archive (2025), Caltech/IPAC Exoplanet Science Institute.*
[6] *Ginsburg, A. et al. (2019), astroquery, ASCL*

GitHub

Linked in