IAF SYMPOSIUM ON FUTURE SPACE ASTRONOMY AND SOLAR-SYSTEM SCIENCE MISSIONS (A7)
Science Goals and Drivers for Future Exoplanet, Space Astronomy, Physics, and Outer Solar System Science Missions (2)

Author: Mr. Suryoday Basak
University of Texas at Arlington, United States, suryodaybasak@gmail.com

CEESA: A HABITABILITY SCORE COMPUTATION APPROACH VALIDATED BY MACHINE LEARNING

Abstract

The rate of discovery of extrasolar planets is rapidly increasing. The idea of planets existing other than Earth which could possibly harbor life has intrigued and captured the imagination of scientists for centuries. In the last decade, thousands of planets were discovered in our own galaxy. This led to the inference that stars with planets are a usual occurrence, and the estimates of the actual number of planets exceeds the number of stars by orders of magnitude. Led by the NASA Kepler Mission, the planetary searches yielded nearly 7000 confirmed and yet to be confirmed exoplanets. The discovery and characterization of exoplanets require both extremely accurate instrumentation and sophisticated statistical methods in order to extract the weak planetary signals. We propose a new model, *Constant Elasticity Earth Similarity Approach* (CEESA) to quantify habitability and match the Earth-similarity approach to classification schemes of extrasolar planets via deep learning architectures. The data used for the analysis is taken from Planetary Habitability Laboratory's Habitable Exoplanets Catalog (HEC). The catalog exhibits significant imbalance, which are dealt with in a twofold manner for the task of classification: first, by iterative artificial balancing, and second, by artificial data augmentation. In either case, the efficacy of classification algorithms are analyzed and related to the structure and distribution of the data.

The CEESA model is based on the CES production function and is represented as follows:

$$Y = f(R, D, T_s, V_e, E) = (r.R^\rho + d.D^\rho + t.T_s^\rho + v.V_e^\rho + e.E^\rho)^{\frac{\eta}{\rho}} \quad (1)$$

where, R, D, T_s, V_e and E are the radius, density, surface temperature, escape velocity and eccentricity of an exoplanet respectively; r, d, t, v, and e are the corresponding coefficients of those parameters and lie in $(0, 1)$; Y is the target output; The range of ρ and ν values are in $(0, 1)$.

Both the CEESA and deep learning classification approaches produce a similar outcome. The complexity of the functional form presents a challenge in terms of finding the optimum habitability score. The habitability computation problem is rewritten as a constrained optimization problem and the difficulty is mitigated using a particle swarm optimization based approach. The classification method used here draws advantages from neural networks and fuzzy logic. The exoplanets are classified into mesoplanets, psychroplanets and non-habitable planets.

The accuracy of the classifier is above 95%. The strength of the proposed method lies in bringing together "earth similarity" and "planetary classification" approaches for a more meaningful interpretation of habitability.