EXOPLANET EXPLORATION

DATS-6202
MACHINE LEARNING

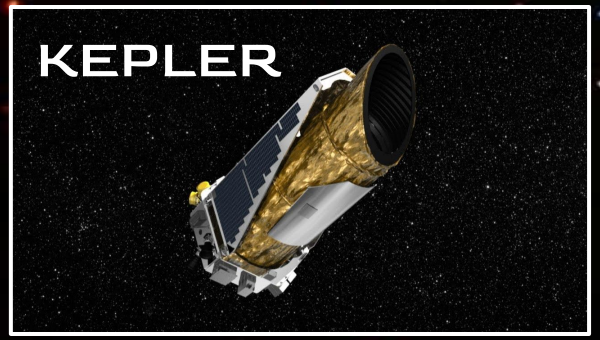
CHIEMEZIEM OGUAYO
NATE EHAT

ANALYSIS BACKGROUND

EXOPLANET DISCOVERIES [TOP OBSERVATORIES]

5,014
CONFIRMED EXOPLANETS

~32,500
PLANETARY SYSTEMS

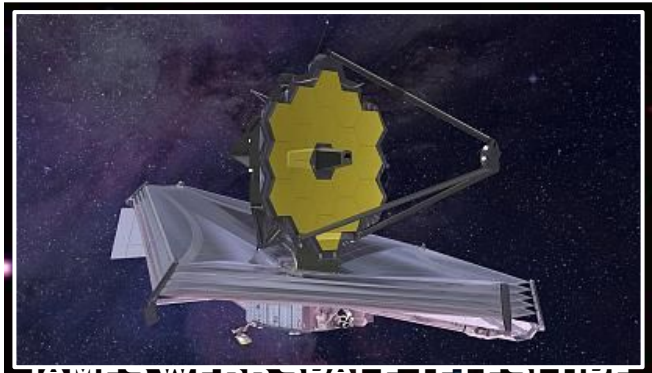


Kepler	2709
K2	537
La Silla Observatory	263
Transiting Exoplanet Survey Satellite (TESS)	205
Multiple Observatories	188
W. M. Keck Observatory	182
SuperWASP	113
OGLE	80
HATSouth	73
HATNet	67
Haute-Provence Observatory	64
Paranal Observatory	38
Anglo-Australian Telescope	37
Lick Observatory	35
CoRoT	33
SuperWASP-South	32
Okayama Astrophysical Observatory	31

PRIMARY TELESCOPE / MISSION DATA

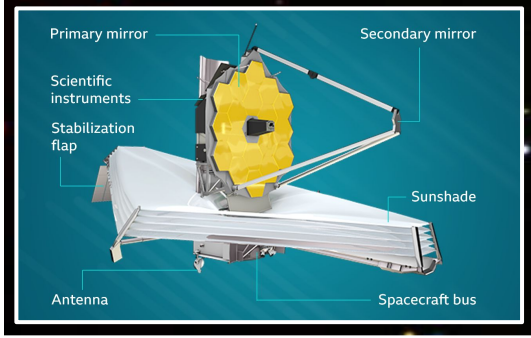
- Kepler (2009-2018)
  - First mission search for Earth-sized planets within habitable zones of nearby stars.
- K2 (2018-Present)
  - Continuation of Kepler's exoplanet discoveries, utilizing vast data archives.
- TESS (2018-Present) – Transiting Exoplanet Survey Satellite
  - Surveying the sky to find transiting exoplanets around brightest stars near Earth.
- James Webb Space Telescope (2021-Present)
  - Advanced instrument capabilities; most powerful telescope launched to date.

PRIMARY TELESCOPE / MISSION DATA



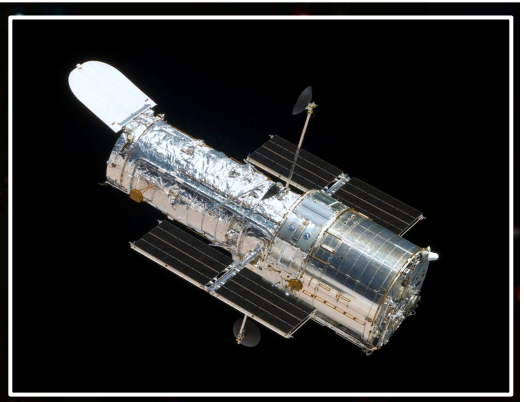
JAMES WEBB SPACE TELESCOPE

TESS [TRANSITING EXOPLANET SURVEY



SPECTROGRAPHY
TRANSIT / EMISSION

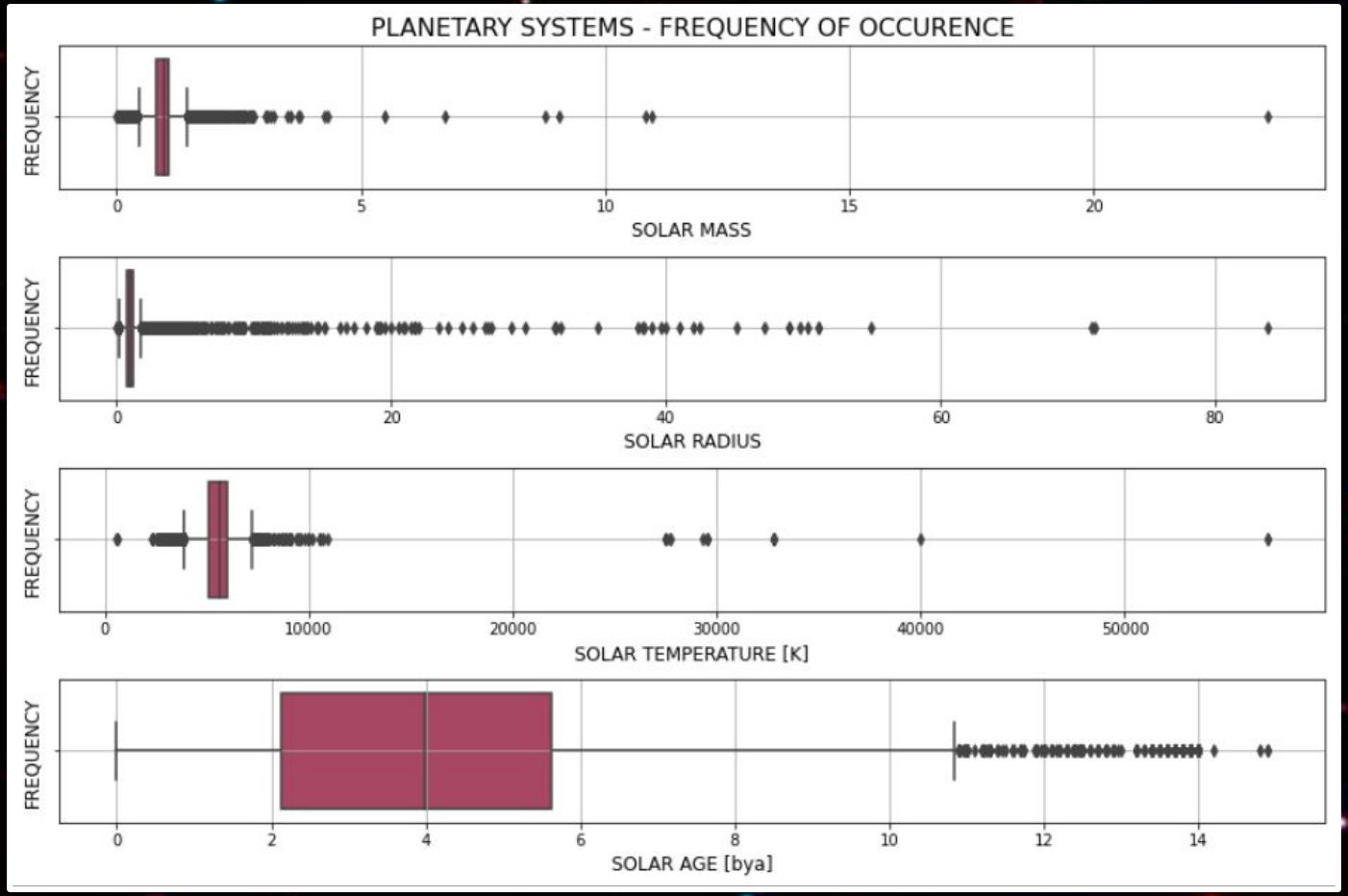
- Hubble Space Telescope [NASA]
  - Wide-Field Camera
- Spitzer Space Telescope [NASA]
  - Infrared Array Camera (IRAC)
  - Infrared Spectrograph (IRS)
- Hale Telescope [CalTech]
  - Wide-Field Infrared Camera (WIRC)
- HADES – HArps-n red Dwarf Exoplanet Survey



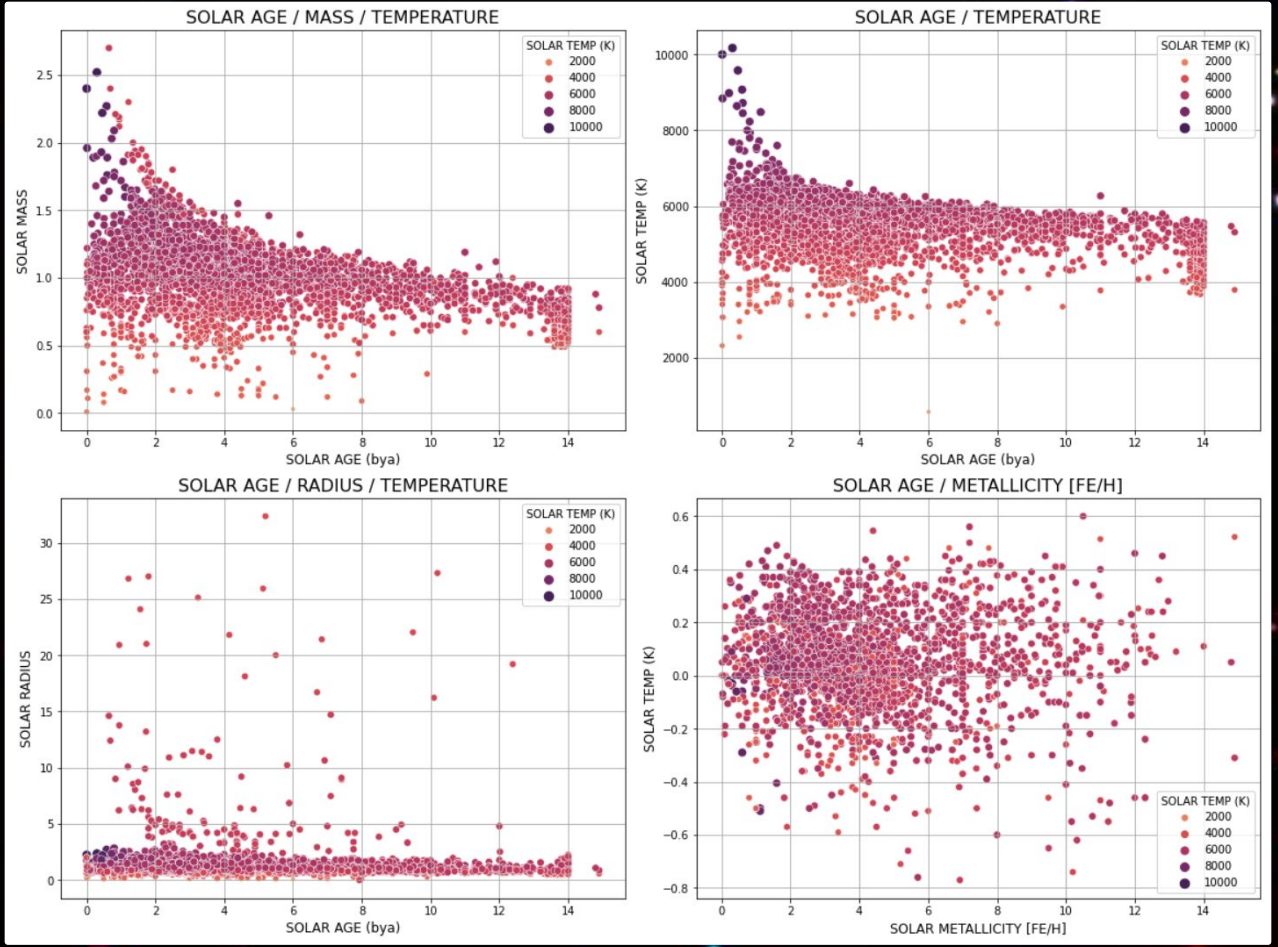
DIRECT IMAGING / MICROLENSING

- DIRECT IMAGING
  - NRM, NACO, NICS, IRCS, SPHERE, SIMON
- MICRO-LENSING
  - Various Telescopic Instruments
- PLANNED MISSIONS
  - Nancy Grace Roman Space Telescope [2027]
  - ARIEL - European Space Agency [2029]

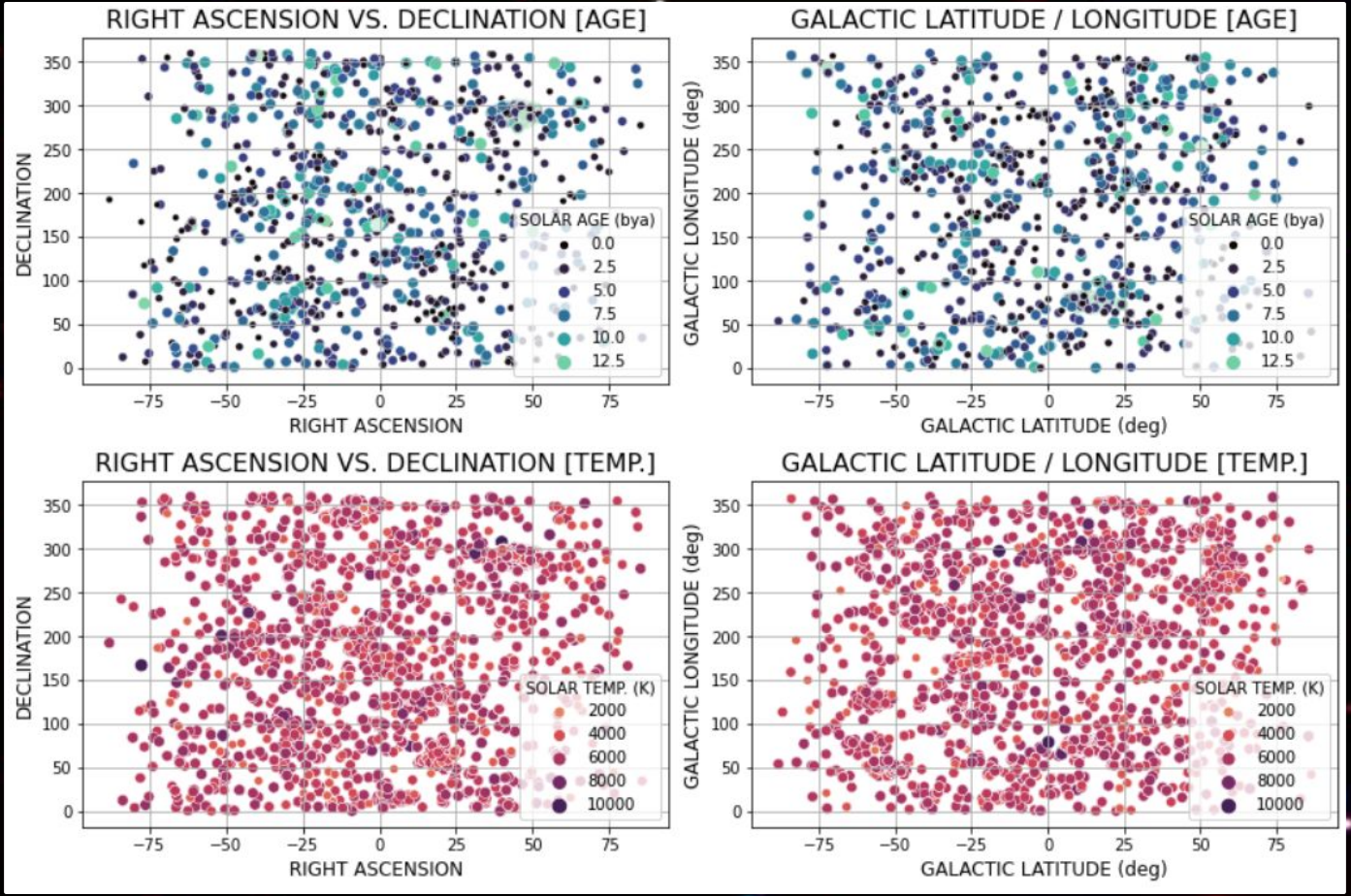
PLANETARY SYSTEMS



STAR SYSTEMS



GALACTIC LOCATION



IDEOLOGY / METHODOLOGY

MODELING PIPELINE

1) Scraping / Mining — CalTech Exoplanet Archive database primary source of data for extraction and exploration
2) Pre-Processing — Handling uncommon features, handling identifiers, scaling the data, imputing null values
3) Model Predictions — Building model and assessing models' ability to predict exoplanet confirmation
4) Model Evaluation — Evaluating best models performance

MODEL

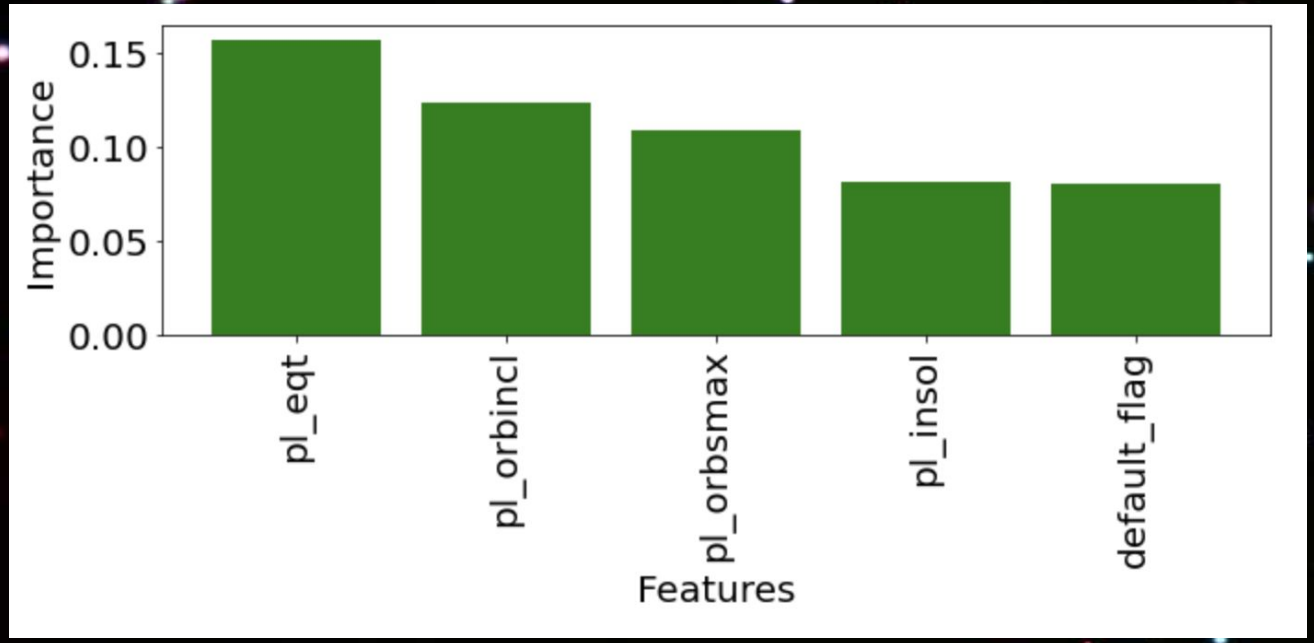
MODEL RESULTS

	best_score	best_param	best_estimator
0	0.908633	{'model__min_samples_leaf': 1, 'model__min_sam...	((DecisionTreeClassifier(max_features='auto', ...
1	0.902407	{'model__alpha': 1e-05, 'model__learning_rate_...	(MLPClassifier(alpha=1e-05, early_stopping=Tru...
2	0.745209	{'model__C': 0.001, 'model__tol': 1e-05}	(LogisticRegression(C=0.001, class_weight='bal...

MODEL EVALUATION

	Precision	Recall	F1-score	AUC
0	[0.9769830028328612, 0.8531026908292147]	[0.8375834851244688, 0.9795081967213115]	[0.90192873488068, 0.9119459935427062]	0.908546

Model Evaluation



Equilibrium Temperature, Inclination, Orbit Semi-Major axis, Insolation Flux, Default Parameter Set

CONCLUSIONS

CONCLUSIONS / TAKEAWAYS

- **DATA MINING:**
  - High accessibility of exoplanet data – yet many gaps

- **MODELING PIPELINE:**
  - Classification is most logical application of ML as applied to data set under investigation, though image recognition of galaxies may offer greater insights in the future

- **FUTURE RESEARCH:**
  - Supplemental NASA data (James Webb Telescope / TESS)
  - Image recognition - opportunity to deploy neural network architecture, utilizing multi-lens / light telescopic imagery

APPENDIX

SOURCES / CITATIONS

- **CalTech Exoplanet Archive**
  - https://exoplanetarchive.ipac.caltech.edu/index.html
  - https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=PS
  - https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html
- **NASA Publications**
  - https://exoplanets.nasa.gov/news/1702/cosmic-milestone-nasa-confirms-5000-exoplanets/
  - https://www.scientificamerican.com/article/5-000-exoplanets-nasa-confirms-a-cosmic-milestone/