Can AI Tell us Where to Find Earth-II ?

Shane J. Robinson

shaner.mail@icloud.com

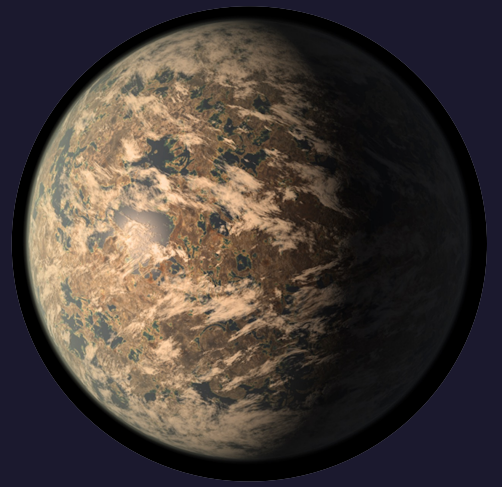
github.com/shanergit

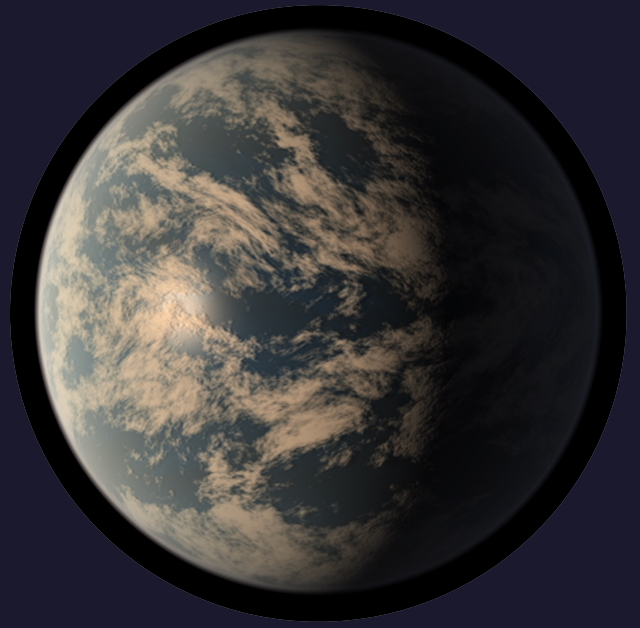
linkedin.com/in/page-shanejrobinson

shanejrobinson_

Choosing Where we Look

- NASA confirms over 5,000 exoplanets
- Likely hundreds of billions in our galaxy
- Distance from Earth?
- Likelihood of habitability?

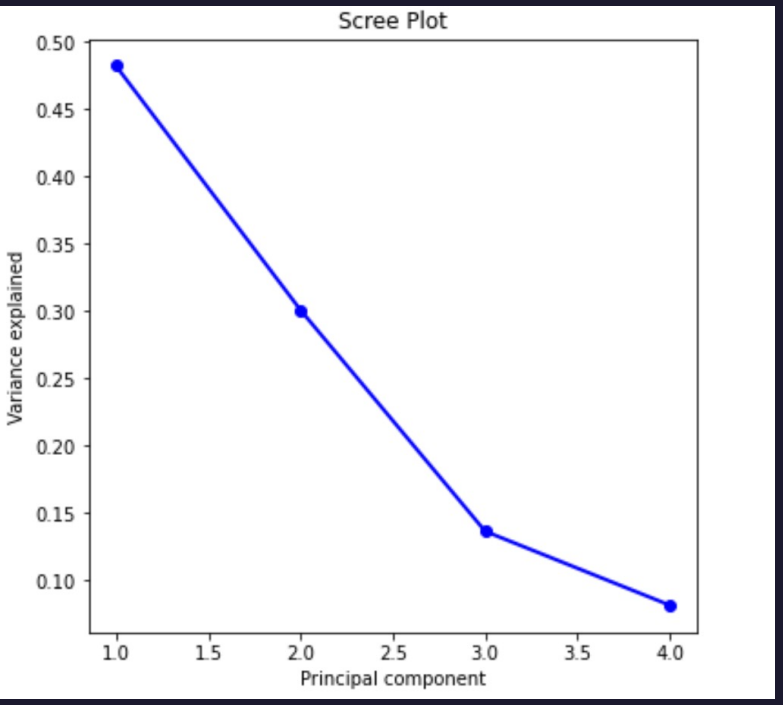


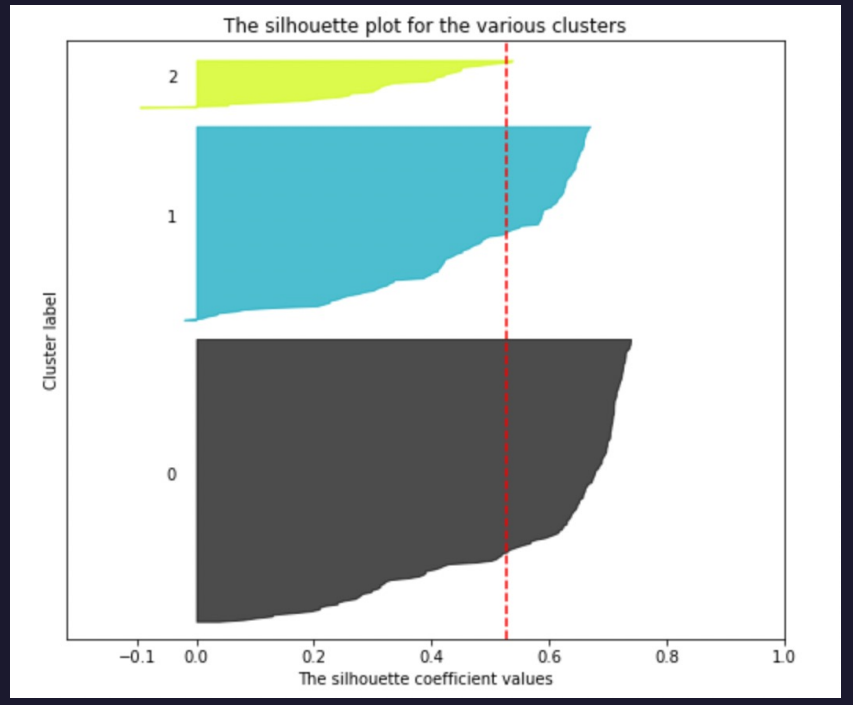


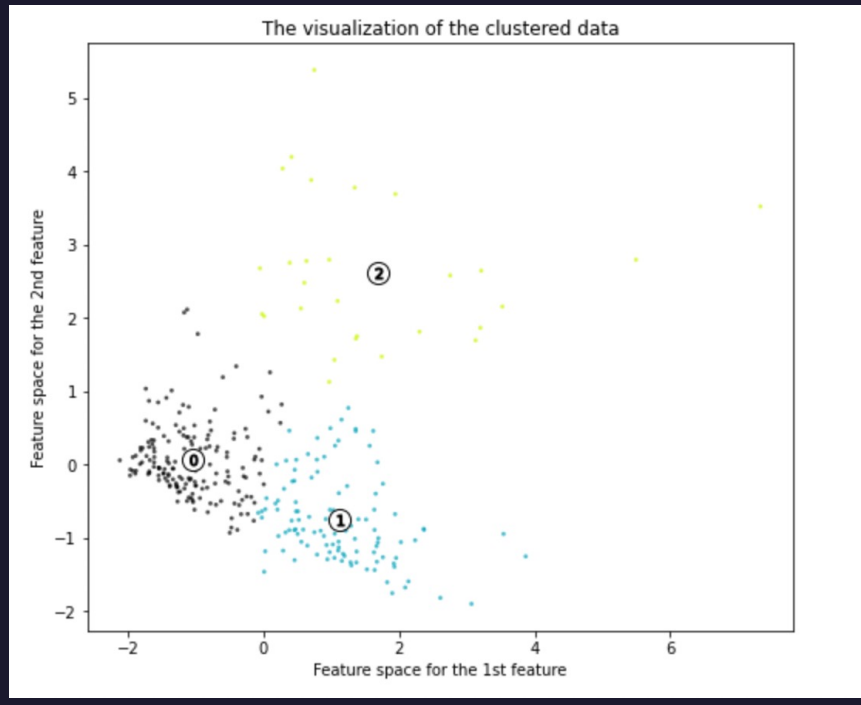
Data | Feature Engineering

Planet Name	Planet Radius	Planet Mass	Orbital Eccentricity	Equilibrium Temperature
WASP-89 b	11.65	1875.197	0.193	1120
Wolf 503 b	2.04	6.26	0.41	790
XO-7 b	15.39	225.34	0.038	1743
Earth	1	1	0.0167	255

- The NASA Exoplanet Archive provides our dataset
- From domain knowledge, four planetary features are chosen
- Principal Component Analysis (PCA)

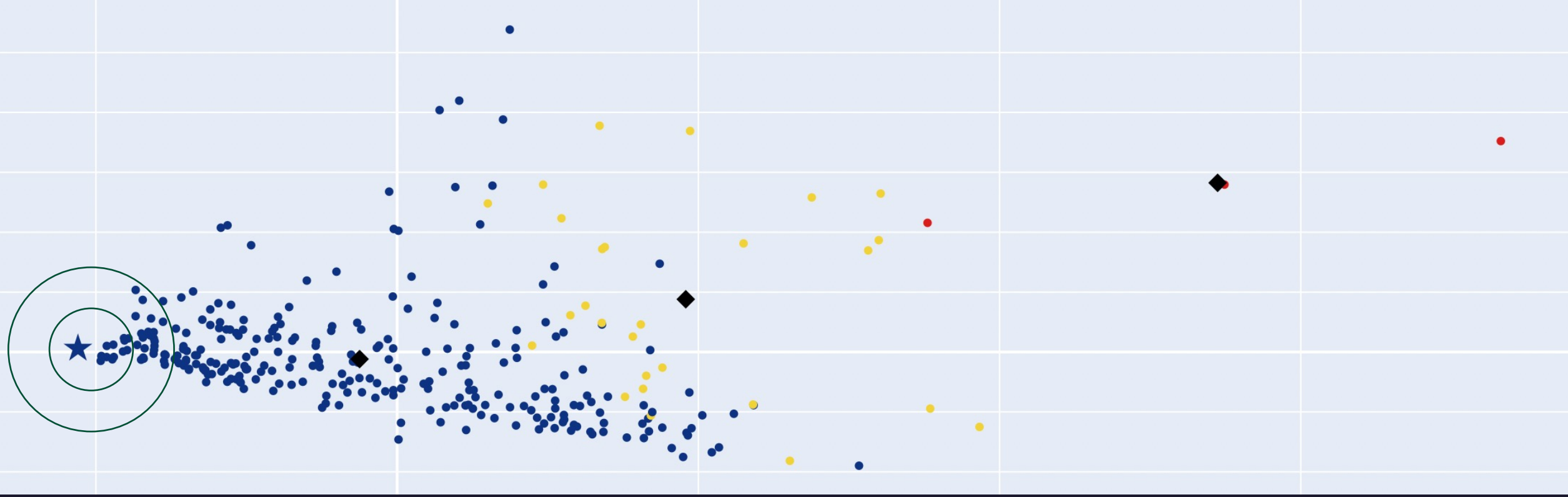






Scree | Silhouette

- Scree: principal components 1 & 2 explain 80% of variance
- Silhouette plots inform the number of clusters for algorithm
- Tested number of clusters was 2 - 6
- 3-Cluster score was highest, with .528



K-means Clusters

- Like-exoplanets grouped and plotted
- Blue star represents Earth
- Black diamonds are cluster centroids
- Exoplanets within a certain radius of Earth are best candidates?

Summary | Future Work

SUMMARY

- Thousands of confirmed exoplanets, billions more
- Unsupervised machine learning with K-means
- Distinct clusters were produced, Earth in one
- Focus on candidates plotted proximally to Earth?

FUTURE WORK

- Investigate the planets plotted nearest to Earth
- Adjust feature engineering with data collection
- Compare K-means with other algorithms

Connect

- Shane J. Robinson
- shaner.mail@icloud.com

- github.com/shanergit/exoplanet_kmeans
- linkedin.com/in/page-shanejrobinson
- shanejrobinson_