

Predicting Exoplanet Discovery

*Flatiron Part Time Cohort
Instructor: Amber Yandow*

Author: Fennec Nightingale

THE METHOD

I used the
OSEMN process
& logistic
regression to
create my model
and predictions

THE GOAL

Showcase how well
machine learning
can predict whether
or not a star has an
exoplanet without
observing changes
in its light

THE DATA

Data collected from
the HYG and OEA
databases containing
information on ~4k
stars with planets and
over 100k without
collected from
various telescopes.

The results

99.32%

accuracy

99.57%

recall

99.14%

precision

accuracy

recall

precision

Recommendations

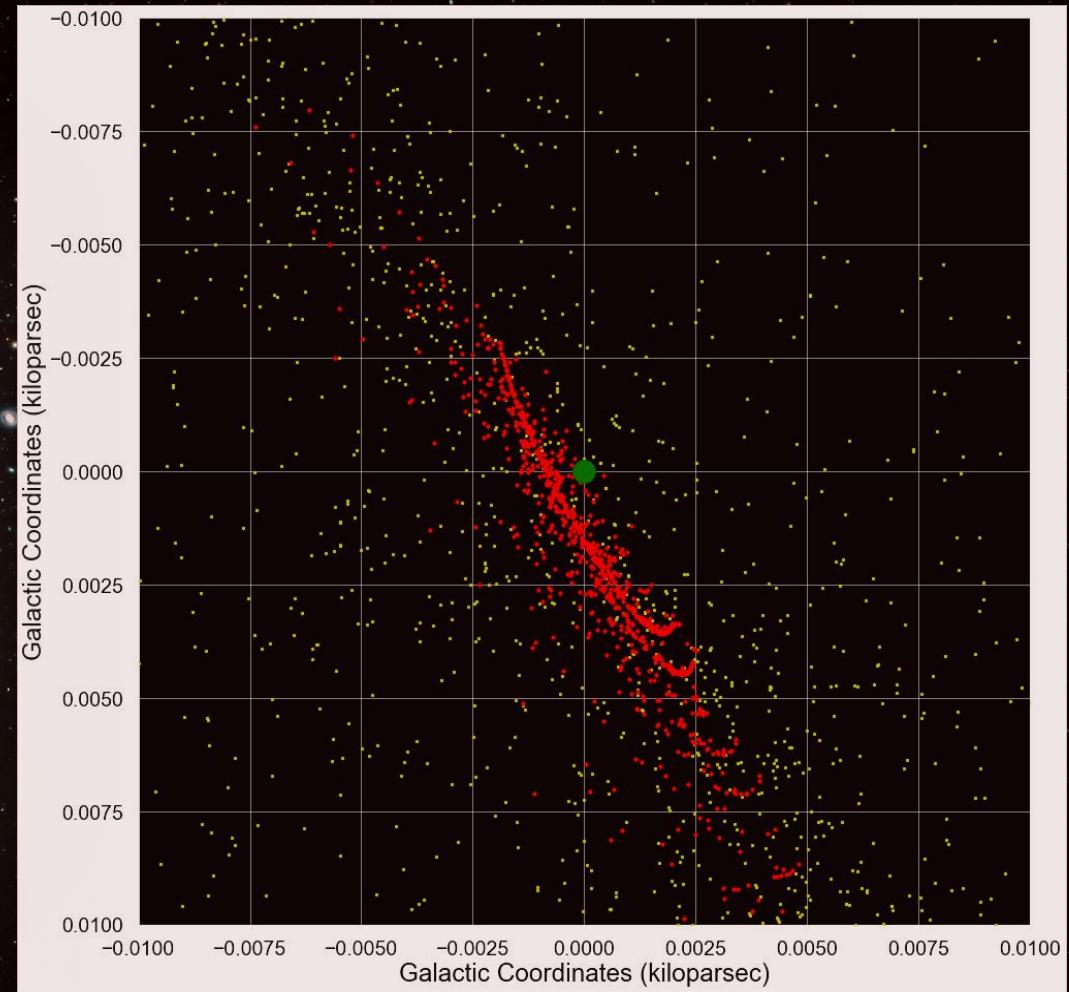
COMMUNICATIONS

Focus on our neighbors



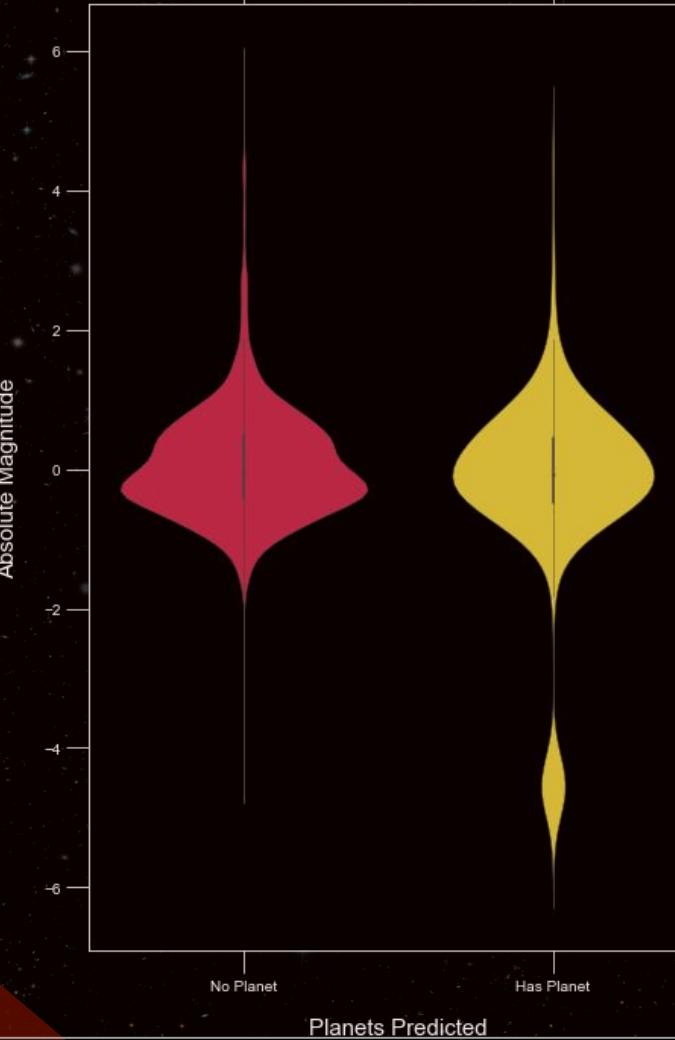
Use machine learning to
know where we should
start on the road to
discovering new
planets.

DATA SCIENCE
DISCOVERY

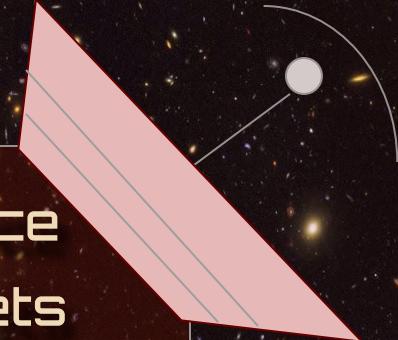


Check out brighter stars

Our brightest stars have been predicted to be about 5% more likely to have planets around them, and despite the challenges bright stars bring for us as well as the planets around them, we have still confirmed a number of them



Conclusion



While I certainly can't say we're in a place where we can completely confirm planets without direct observation, we've made good enough progress that I feel we can make a reasonable prediction to guide where we should be looking and speed up our confirmation process as well as cut down on time spent observing fluxuations in stars unlikely to have exoplanets.

Future work

1. Use kepler labelled time series data to train deep learning algorithms to detect exoplanets based on light fluxuations in observed stars.
2. Write something that is able to parse and accurately separate stellar types (as well as predict missing values) to test predictions made against more random data.
3. Use additional data from the Open Exoplanet Catalogue to predict features of planets around stars & predicted stars.
4. When more data is available, expand predictor to include multi-planetary predictions.



Thank you for your time.

Please check out my github for a more detailed analysis & access to the data.