# An Application of t-SNE to Exoplanets

Jess Vriesema

December 14, 2016

## Introduction

As we enter the golden age of planetary discovery, we will undoubtedly want to group or classify planets more precisely than we already do. Already, we have a diverse and loosely-defined set of classifications for exoplanets, such as “gas giants”, “hot jupiters”, “super-earths”, “sub-neptunes”, and so on. The exoplanet community has even come up with designations for groups of planets which have yet to be discovered, such as “water worlds” and “Cthonian planets”. In order for a designation to persist, however, it needs to be useful and sensible.

Machine learning can help us determine groupings of planets in a statistical sense, and this in turn can inform our physical basis for planet categorization. The challenge is that there are many features which may affect these classifications, and it is difficult to find relationships between these features in a multi-dimensional space. A dimensional-reduction technique is needed.

T-distributed stochastic neighbor embedding[1] (t-SNE) is well-suited to perform an initial exploration of dimension-reduction and help us visualize these groupings, so we use it here to study exoplanet classification.

## Methodology

### Data Source

There are several significant exoplanet datasets available (e.g. exoplanets.org, the Extrasolar Planets Encyclopaedia, and the NASA Exoplanet Archive). I chose to use the one from exoplanets.org because it had the most extensive complete set of data for the features I was interested in (it had roughly 1600 usable planets compared to several hundred in the other databases) and was relatively well-documented.

[1] van der Maaten, L.J.P.; Hinton, G.E. (Nov 2008). "Visualizing High-Dimensional Data Using t-SNE" (PDF). Journal of Machine Learning Research. 9: 25792605.

This dataset has 314 features, although on average 195 of them are missing/null. The remainder are primarily uncertainties and reference URLs to the papers from which the numbers came. A dozen or two of the values are physically meaningful. Of these, I chose to focus on the following thirteen features:

- $A$: the semi-major axis, in AU
- $AR$: the ratio of the semi-major axis to the stellar radius ($R_\star$)
- $BINARY$: a flag indicating whether or not the host star was a binary
- $DENS$: the density of the planet, in $\mathrm{g\,cm^{-3}}$
- $ECC$: the orbital eccentricity of the planet
- $FE$: the iron abundance (or metallicity) of the host star
- $LOGG$: the spectroscopic $\log g$ (surface gravity) of the host star
- $MASS$: the mass of the planet, in Jupiter masses ($M_J$)
- $MSTAR$: the mass of the host star ($M_\star$), in solar masses ($M_\odot$)
- $PER$: the orbital period of the planet, in days
- $R$: the radius of the planet, in Jupiter radii ($R_J$)
- $RSTAR$: the radius of the host star, in solar radii ($R_\odot$)
- $TEFF$: the effective temperature of the host star, in K

In addition to the exoplanets in the dataset, I included the eight planets from our solar system and found values for each of the thirteen features listed above. My purpose was initially to use these to give my audience reference points to understand the transformed data. Their inclusion can also be justified by acknowledging that they are also planets and there is no physical reason *not* to include them in an analysis that seeks to understand relationships between planets.

## Missing Data

About 93% of the planets did not include a density, and a few others had densities that were not realistic (e.g. $\rho < 0.05\mathrm{g\,cm^{-3}}$ or $\rho > 500\mathrm{g\,cm^{-3}}$), so I calculated the density from $MASS$ and $R$ for the planets without good densities.

I used the `pandas.DataFrame.dropna()` command to remove null rows and/or columns. This left me with 1576 exoplanets, plus the eight planets in our solar system, for a total of 1584 planets.

## Strategy and Data Preparation

I wanted to group the planets using relevant physical inputs, and then visualize it in a way to see where certain known planetary groups—namely, hot jupiters and Earth Analogues—are in the transformed space.

I therefore calculated five new features and two new labels in order to calculate three new labels:

- *PERIAPSE*: the planets periapse, in AU
- *APOAPSE*: the planets apoapse, in AU
- *LUMINOSITYSTAR*: the luminosity of the host star ($L_\star$), in W
- *HZ_RMIN*: the inner boundary of the system's habitable zone, in AU
- *HZ_RMAX*: the outer boundary of the system's habitable zone, in AU
- *INSIDE_HZ*: where in the system the planet lies (cases shown below)
  - (-2): $PERIAPSE < APOAPSE < HZ\_RMIN$
  - (-1): $PERIAPSE < HZ\_RMIN < APOAPSE$
  - (0): $HZ\_RMIN < PERIAPSE <= APOAPSE < HZ\_RMAX$
  - (+1): $PERIAPSE < HZ\_RMAX < APOAPSE$
  - (+2): $HZ\_RMAX < PERIAPSE < APOAPSE$
- *IS_HOT_JUPITER*: whether or not the planet is a hot jupiter
- *EARTH_MASS_COMPARISON*: modified signum comparison of the planet's mass to that of Earth:
  - (-1): $M_P < 0.5M_E$

- (0): $0.5M_E < M_P < 2M_E$
- (+1): $2M_E < M_P$

- *EARTH_SIZE_COMPARISON*: modified signum comparison of the planet's size to that of Earth:
  - (-1): $R_P < 0.8R_E$
  - (0): $0.8R_E < R_P < 1.9R_E$
  - (+1): $1.9R_E < R_P$
- *EARTH_ANALOGUE*: whether or not the planet is similar to Earth in terms of radius and mass

The habitable zone was calculated using the following equations, taken from this page, for the inner ($r_i$) and outer ($r_o$) radii as follows:

$$r_i = \sqrt{\frac{L_\star}{1.1}} \tag{1}$$

$$r_o = \sqrt{\frac{L_\star}{0.53}} \tag{2}$$

# Procedure

I ran a battery of "Trials" with different input features for the t-SNE analysis. For each Trial, I whitened the input features, performed the t-SNE analysis to transform the data down to two dimensions (for 2D visualization), and plotted the transformed data for each and every feature listed above with a colormap corresponding to that feature. These plots are included in a compressed archive file available at https://www.lpl.arizona.edu/ṽriesema/files/palafox/plots.zip.

# Trial A

I wanted to get a feel of the data by including all relevant fields. I included the following: A, BINARY, ECC, DENS, FE, MASS, MSTAR, PER, R, RSTAR, TEFF, INSIDE_HZ and IS_HOT_JUPITER. The results of this are included in an attached file.

This trial is problematic for several reasons. The first thing one notices is that Venus and Mercury are far from the other planets, but that the other planets are clustered together. This is nonsensical, as Venus and Earth have more in common, for instance, than Earth and Neptune. The second thing worthy of note is that the hot jupiters are all by themselves. This is problematic because they ought to be a *little* closer to Jupiter, and it seems unlikely that they would be grouped so distinctly from all other planets. The same problem occurs with planets in binary star systems. Lastly, it does a horrible job at classifying Earth Analogues, and Earth is not even near this loose cluster.

I explain the faults of Trial A as primarily being due to including several of the labels—which are nonphysical and intended only for visualization—as inputs to the t-SNE routine. This artificially increased their distance from all other points because of how I classified them *a priori*, rather than grouping them based on underlying physical similarities. I believe the similarities between most of the planets in the solar system is due to them sharing Sol as their host star and having a large number of input features be properties of the host star, which these planets share. For example, the reason Earth is closer to Jupiter than to the Earth Analogues is because Earth Analogues have similar mass and radii (two features), but Earth and Jupiter share the same star (four features).

# Trial D

The goal of Trial D was to remove the artificial labels that plagued Trial A. I also removed FE because it seemed to have no effect on anything else (indeed, it had a negligible-to-low correlation with all other features). Lastly, since A is essentially redundant with PER, I removed PER to avoid giving the fundamental physical quantity (A xor PER) double weight in the t-SNE analysis. This left me with the following inputs to t-SNE: A, BINARY, ECC, DENS, MASS, MSTAR, R and RSTAR.

Trial D was successful in grouping the four terrestrial planets together, Jupiter and Saturn together, and Uranus and Neptune together. The t-SNE transform primarily grouped planets by R in the horizontal direction and MSTAR in the vertical direction. It does a fine job of grouping Earth Analogues together, and correctly places Earth and Venus near the center of all of them.

It also has a separate regime for hot jupiters. Unfortunately, there are a number of non-hot jupiters mixed in with them, so this transformation does not satisfactorily distinguish between hot jupiters and non-hot jupiters.

Another fault of Trial D is that it still grouped all planets in a binary star system tightly together, and far away from other planets. If there is any physical basis for this, I am unaware of it. More than likely, the BINARY flag introduced additional distance in the t-SNE analysis in the same way that IS_HOT_JUPITER did in Trial A. Of course, planets in binary star systems would be expected to be different in some respects, but one would not expect this behavior to be adequately captured by the input parameters I used.

# Trial F

The goal of Trial F was to determine what kind of groupings could be done using just stellar parameters (A, FE, MSTAR, RSTAR, LUMINOSITYSTAR, FE) plus the planets' semi-major axis.

This Trial is not worthy of note other than making two observations. First, the planets in our solar system are all grouped together, which is to be expected because they share the same host star. The second observation is that there is essentially no structure whatsoever for any feature plotted as the color axis. This implies that the properties of the host star do not significantly determine the kinds of planets which form around them. I believe it also implies that the stellar types in the dataset were somewhat well-sampled, as there are no significant gaps in the transformation.

# Trial G

My purpose in running Trial G was to see how the planets would group if I used only used features describing the planets themselves, and not features that describe their host star. To this end, I used A, ECC, DENS, MASS and R as inputs to t-SNE.

This Trial performed exceptionally well. This was the only Trial to achieve the characteristic "spindly" grouping seen in many other t-SNE plots for other applications. Trial G correctly grouped Earth with Venus, Mars with Mercury, Neptune with Uranus, and Saturn with Jupiter.

The plot which colors PER shows an unusual pattern: the thin, darker (shorter period), denser groupings have lighter (longer period) planets coming out from them in a direction pointing towards the center of the plot. I take this to reflect the fact that the dataset contains many more shorter period planets than longer period planets. This is likely because shorter period planets are easier to detect—an observational or sample bias in the data.

Another success of Trial G is that it groups planets with eccentric orbits closer together. Hot jupiters are similar to Mercury and Mars in this regard, and this is apparent in the ECC plot.

Finally, Trial G does a superior job of grouping Earth Analogues and grouping hot jupiters.

# Conclusions

None of the Trials were able to group planets that were likely to be in the habitable zone. Each Trial spread these planets out throughout the t-SNE space This is likely because the presence of an atmosphere strongly influences whether or not a planet can be habitable. It is also likely because the metric I used for whether a planet is in the habitable zone is a very outdated and overly simplistic approach.

I observed that artificial labels such as IS_HOT_JUPITER were tremendously useful for plotting the data and identifying clusters in t-SNE space, but gave misleading distance to the transformation when used as inputs to t-SNE. These labels would likely be very useful when training a neural network, however.

I also observed that choosing physically-motivated features as inputs to t-SNE improved the results. Furthermore, planetary characteristics were much more interesting and useful than stellar characteristics.