Predicting Anime Ratings: Linear Regression vs. Neural Nets

Austin Chen

May 9, 2015

Just like movies, TV shows, and other forms of media, anime can be categorized into different genres: sci-fi, mystery, romance, etc. In addition, there are many common themes found in anime, some of which are quite unique to the medium (e.g. magical girls). My project leverages linear regression and neural networks as two different ways to predict anime ratings from these genres and themes. By drawing on hundreds of thousands of votes from users, I aim to quantify exactly what makes an anime "good".

1   Featurizing

For each anime, I included as features: the number of votes, the number of views, and a binary indicator for each of the genres and the 50 most common themes. These are used to regress against anime ratings, which are provided as averages across a 10-point scale.

Some example genres are action, adventure, and comedy; example themes are aliens, assassins, and bishounen.

All anime data was made available courtesy of Anime News Network.
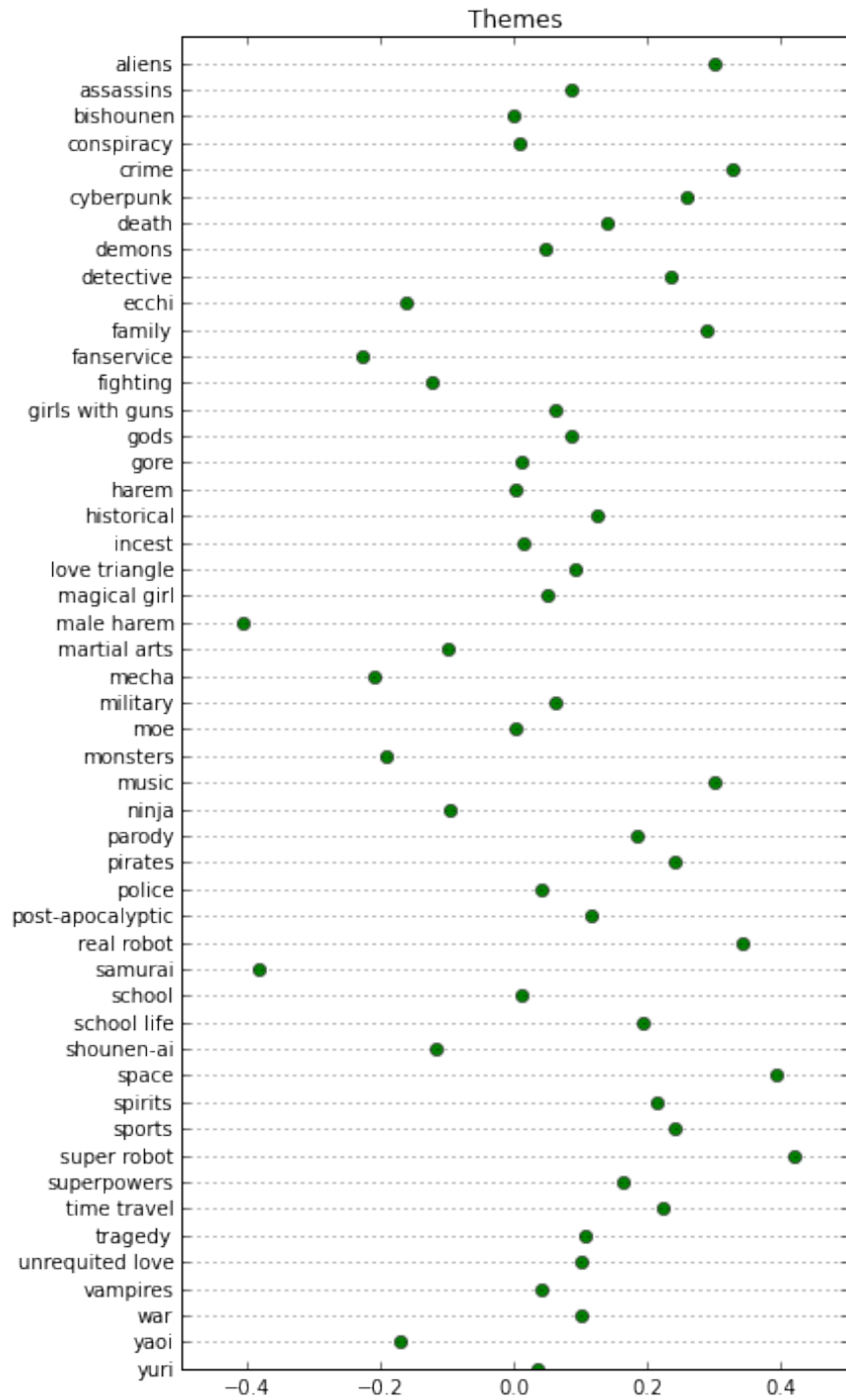
2   Linear Regression

Linear regression on the average user rating for each anime provides a baseline prediction strategy, with rather large errors (between -2 and 2 on a 10 point scale). However, it lends itself to clean interpretations for the various effects of different features. Each entry in the weight vector describes how much the presence of a particular genre or theme affects the rating.

Genres that are highly rated include drama, thriller, slice of life, and mystery, and highly rated themes included superpowers, space, and real robots. Meanwhile, erotica and horror are heavily penalized as genres, as are the themes of male harem and samurai.
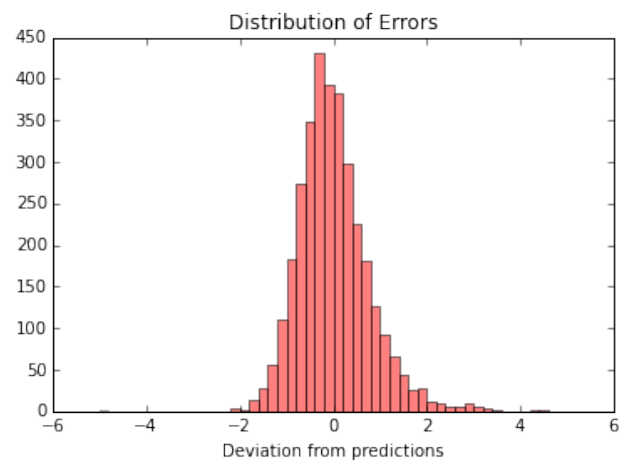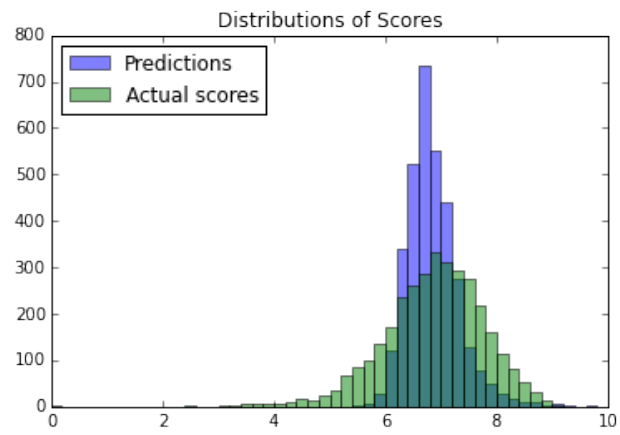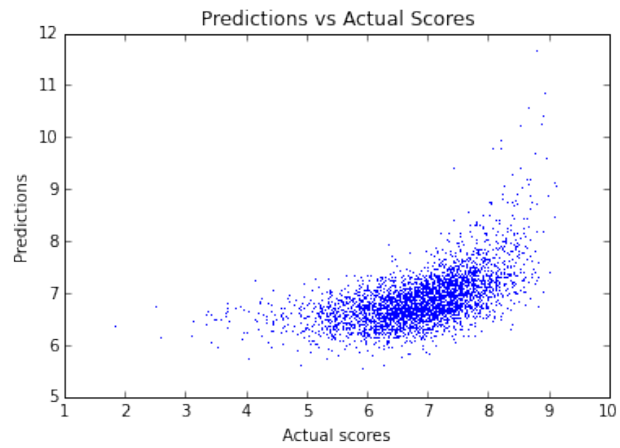
Interestingly, while anime ratings are positively correlated with the number of votes, they are negatively correlated with the number of views, leading to the tentative conclusion is that popular anime is bad.

2.1 Linear Regression Weights

Other

Genres

Themes
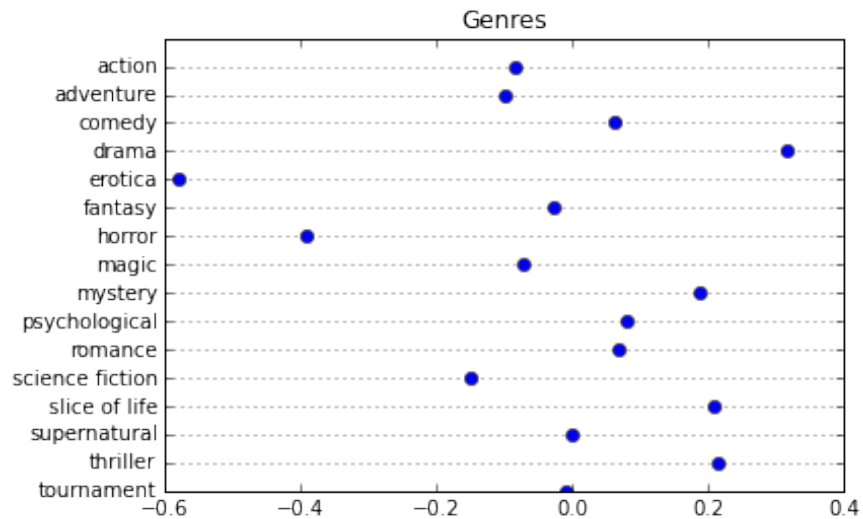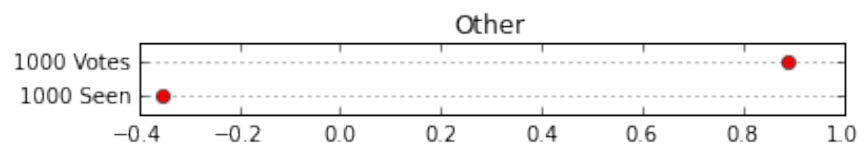
2.2 Linear Regression Analysis

Predictions vs Actual Scores

Distributions of Scores


Distribution of Errors

3 Filtering

While plotting predicted scores vs actual scores, I noticed that an abnormally large number of anime were being predicted as 6.0. This anomaly was caused by unpopular/niche anime not having any appropriate features to distinguish them from others. I controlled for this anomaly by filtering out anime with very few (less than 50) votes.
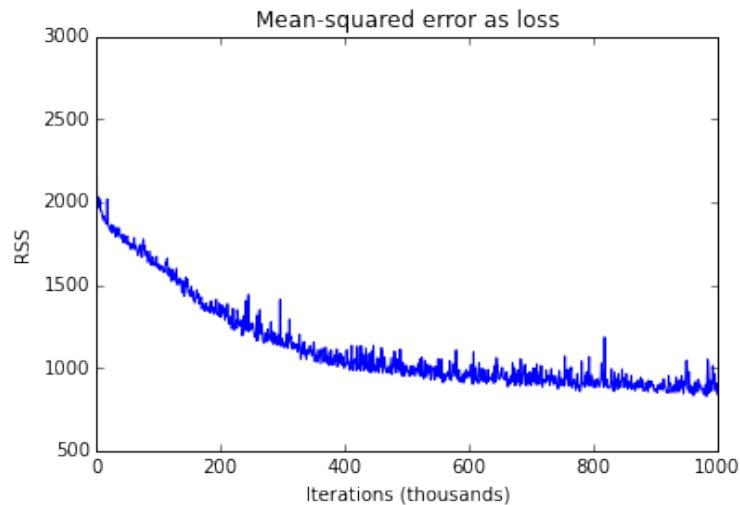
Before
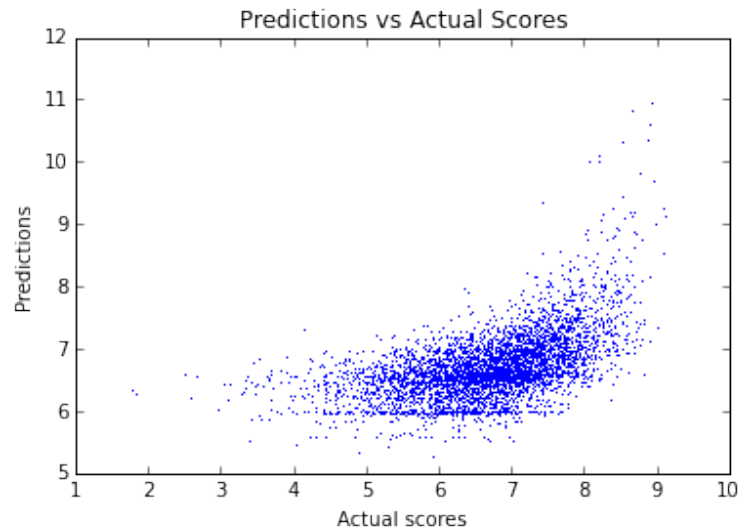
After
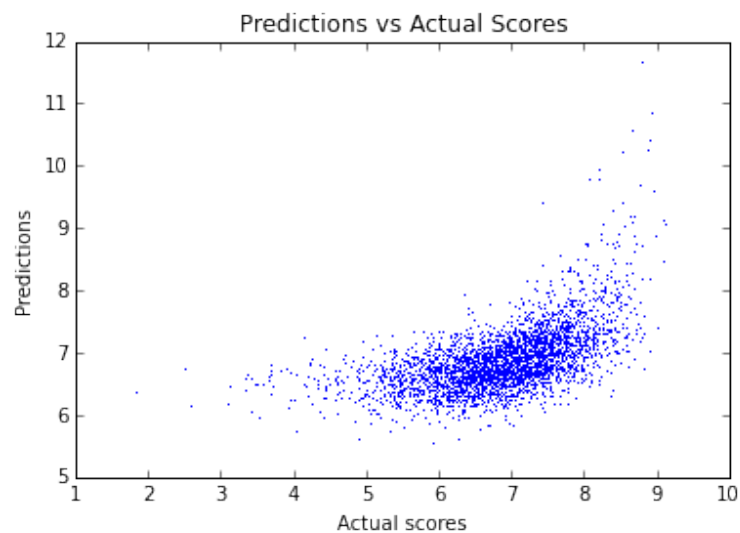
4 Neural Nets

Neural networks provide better predictions than linear regression. I used a neural net with 1 hidden layer of 40 nodes with an activation of tanh, and 1 output node with no activation function. After one million training iterations, it had a residual sum of squares of 841.9, compared to 1914.2 for linear regression. The distribution of scores predicted by the neural net much more closely matches the distribution of actual scores, and the errors are more concentrated (standard deviation of 0.50 as opposed to 0.75). Unfortunately, the neural net obscures insights into the effects of particular features upon anime ratings.

Some hyperparameters I experimented with to adjust neural net performance included the number of nodes in the hidden layer as well as the step size for the stochastic gradient descent.

4.1 Neural Net Analysis

Predictions vs Actual Scores

Distributions of Scores

Distribution of Errors