

# History Repeats Itself

Renay Oshop, AyurAstro.com, January 7, 2020, Boulder, CO, USA

## Summary

Using natural language processing, a system of using artificial intelligence to understand text, the author constructs a predictor based on a simple, constrained, description of the solar system on the start dates of various world historical events (n=7819). This predictor is then statistically shown to be effective at characterizing similar events of the past. An online offering of this predictor for use of characterizing the future accompanies this paper.

## Introduction

Understandably, the future is a subject of shared fascination.

[Future history](#) is a term of ripe heritage. Typically, it is used to describe works of science fiction. What if future historical events could be predicted computationally, even if hazily?

This paper shows that the simple solar system astronomy of historical event start dates can map onto similar major world events of the past as numerically measured by textual analysis of one-sentence descriptors. By looking at the astronomy that accompanies future dates, those textual characterizations can easily be employed to see astronomy mappings onto historical events.

## Methods and Materials

The project makes heavy use of the [Mathematica](#) system of knowledge representation, computation, and analysis.

### Characterization of Historic Events Dataset

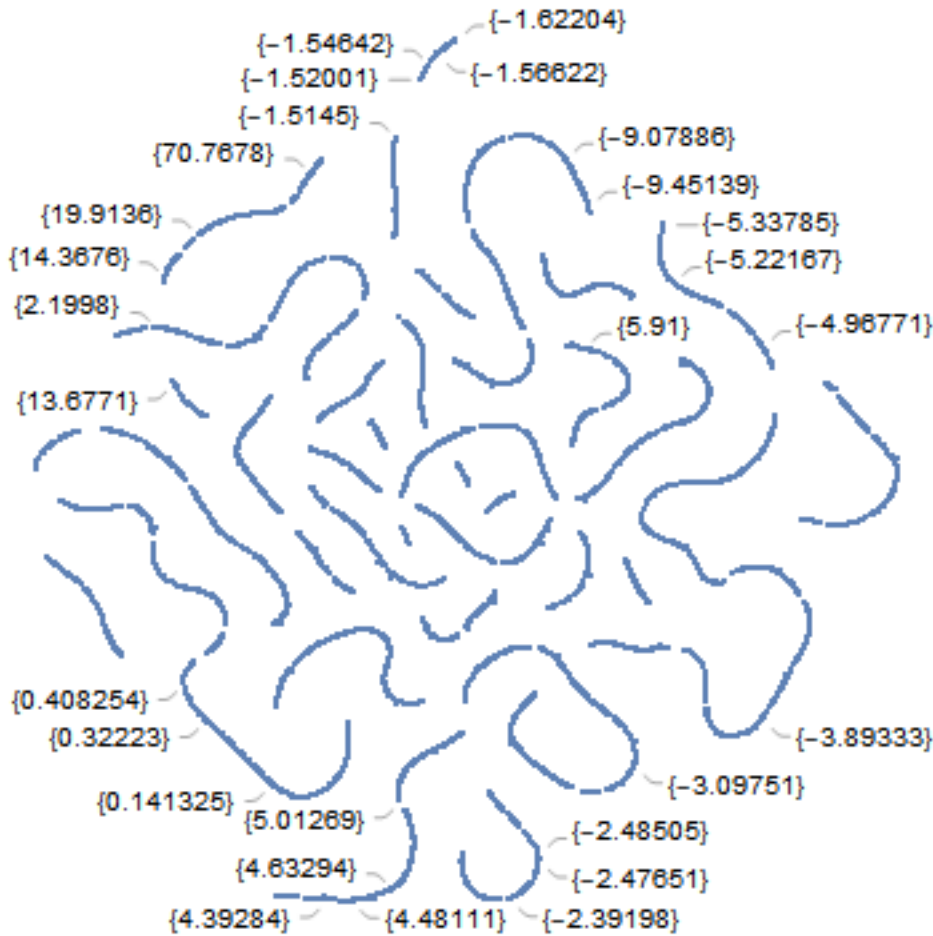
The first step is to seek the historical events. What counts as a true historical event? What counts as the start date? How does one even characterize a world event given differing cultural perceptions? For ancient events, how confident can one be for either the start date or the description?

These are all excellent questions that the author largely side stepped by employing [Mathematica's "Historical Events" feature](#). There were 7819 such events available thereby with start dates and one-line descriptions included. The sources from Mathematica for this historical event dataset are numerous. They can be found through [this mechanism](#).

The following is a set of descriptors for 400 or so example historic events out of the 7819.







The simple geocentric astronomy of the start date was mapped onto this unique feature space number for each event. The question then becomes: can the astronomy of an event's start date describe the feature space number and by extension, the textual description of the event?

A randomization created a training set of 6000 events, a validation set of the next 1000 events, and a test set of the following 819 events.

The simple astronomy considered here are right ascensions (i.e. Tropical degrees) of Sun, Moon, Mercury, Venus, Mars, Jupiter, Saturn, Uranus, Neptune, and Pluto. To be sure, more solar system planetary details (such as retrogressions) could have been included, as well as those of nodes, asteroids, etc. The focus here on historical events though asks that we consider the more slow-moving solar system entities, at least as is understood through astrology discourse.

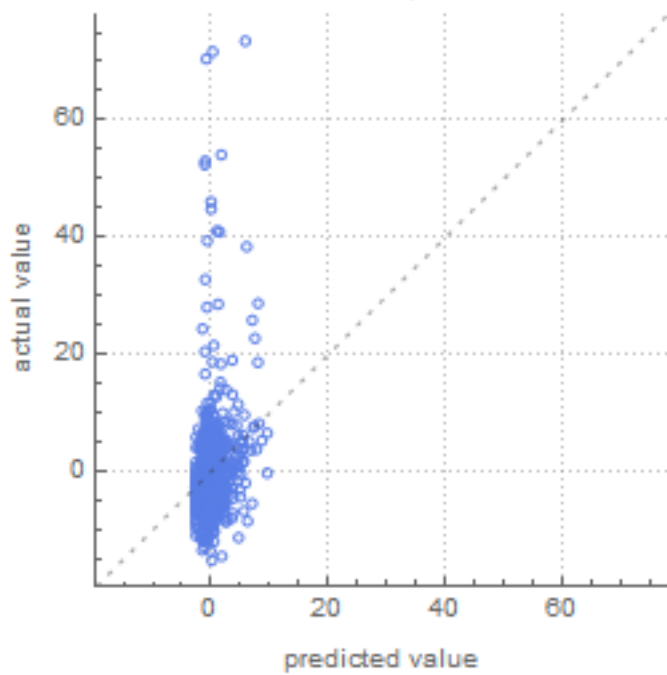
A method using gradient-boosted trees was automatically selected for predicting the BERT encoder's value for a text description based only on the set of astronomy placements for the start of an event.

## Results

A 3.25 percent reduction in standard deviation from baseline was achieved in the test set with some linearity emerging of actual value plotted against predicted value. (R-squared is 0.0475.)

### Predictor Measurements

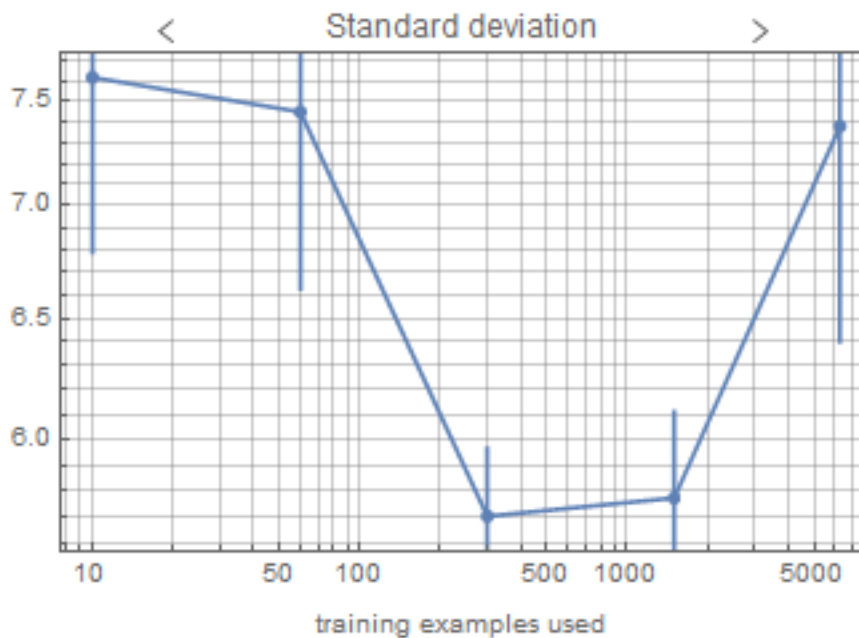
Number of test examples	<b>819</b>
Standard deviation	<b>8.52 ± 0.78</b>
Standard deviation baseline	<b>8.73 ± 0.79</b>
Mean cross entropy	<b>3.56 ± 0.093</b>
Single evaluation time	<b>6.44 ms/example</b>
Batch evaluation speed	<b>1.57 examples/ms</b>



Moreover, when the full set of 7819 historical events were evaluated at once, the internal testing showed a lower entropy (loss).

### Predictor information

Data type	Mixed (number: 10)
Standard deviation	5.70 ± 0.26
Method	GradientBoostedTrees
Single evaluation time	5.62 ms/example
Batch evaluation speed	14.7 examples/ms
Loss	3.31 ± 0.017
Model memory	790. kB
Training examples used	7819 examples
Training time	41.3 s



It is this latter predictor function, based on the full events dataset, that is offered for use at AyurAstro.com.

### Conclusion

To be sure, more astronomical features can be tested. As well, a more inclusive dataset could be used.

This relatively short study nonetheless offers a demonstration of the use of astronomical features in predicting context and vocabulary for describing historical events whether in the past, present, or future.