Capstone Project: Fake Poem Detection

1. The Objectives

We want to build a deep neural network that creates poetry. Afterwards will train a classifier to be able to distinguish our fake poetry from real poetry. The underlying idea could be expanded to creating an algorithm that is able to detect fake news.

2. The Data

We obtained a dataset from Kaggle which includes 3310 unique poems. The data will be split into a training set to train our neural network and a separate set whose beginning words we will use to initiate creating fake poems. Afterwards we will create a test and training set from the rest of our data and feed it into our classifier.

<https://www.kaggle.com/johnhallman/complete-poetryfoundationorg-dataset>

Final Classification Algorithm

3. The Algorithm

We plan on using a Deep Neural Network structure with LSTM layers. We will use Tensorflow and Keras in order to make implementation as easy as possible.

The Second algorithm will be a binary classifier whose architecture is yet to be determined.