# Capstone Project 2: Milestone Report

## Quora – Toxic question prediction

**Introduction:**

An existential problem for any major website today is how to handle toxic and divisive content. Quora wants to tackle this problem head-on to keep their platform a place where users can feel safe sharing their knowledge with the world.

[Quora](https://www.quora.com/) is a platform that empowers people to learn from each other. On Quora, people can ask questions and connect with others who contribute unique insights and quality answers. A key challenge is to weed out insincere questions -- those founded upon false premises, or that intend to make a statement rather than look for helpful answers.

I will develop models that identify and flag insincere questions. To date, Quora has employed both machine learning and manual review to address this problem. With my help, they can develop more scalable methods to detect toxic and misleading content.

An insincere question is defined as a question intended to make a statement rather than look for helpful answers. Some characteristics that can signify that a question is insincere:

* Has a non-neutral tone
  + Has an exaggerated tone to underscore a point about a group of people
  + Is rhetorical and meant to imply a statement about a group of people
* Is disparaging or inflammatory
  + Suggests a discriminatory idea against a protected class of people, or seeks confirmation of a stereotype
  + Makes disparaging attacks/insults against a specific person or group of people
  + Based on an outlandish premise about a group of people
  + Disparages against a characteristic that is not fixable and not measurable
* Isn't grounded in reality
  + Based on false information, or contains absurd assumptions
* Uses sexual content (incest, bestiality, pedophilia) for shock value, and not to seek genuine answers

The training data includes the question that was asked, and whether it was identified as insincere (target = 1). The ground-truth labels contain some amount of noise: they are not guaranteed to be perfect.

Note that the distribution of questions in the dataset should not be taken to be representative of the distribution of questions asked on Quora. This is, in part, because of the combination of sampling procedures and sanitization measures that have been applied to the final dataset.

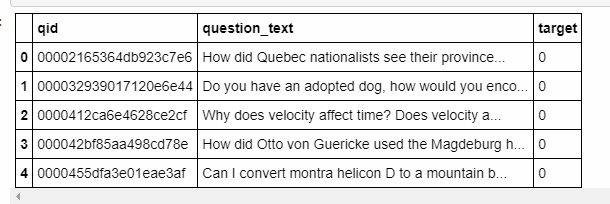
Data Wrangling

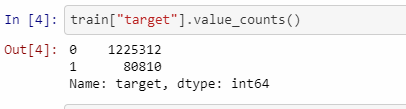
## **File descriptions**

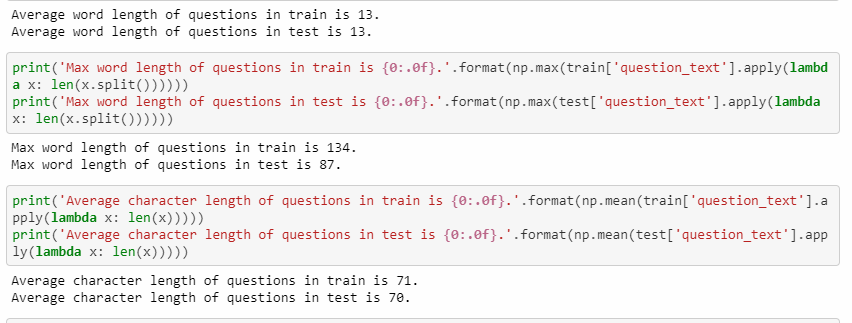
* train.csv - the training set
* test.csv - the test set
* sample\_submission.csv - A sample submission in the correct format
* enbeddings/ - (see below)

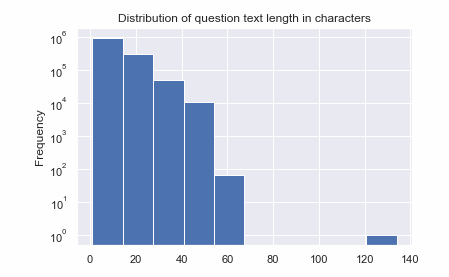
## **Data fields**

* qid - unique question identifier
* question\_text - Quora question text
* target - a question labeled "insincere" has a value of 1, otherwise 0

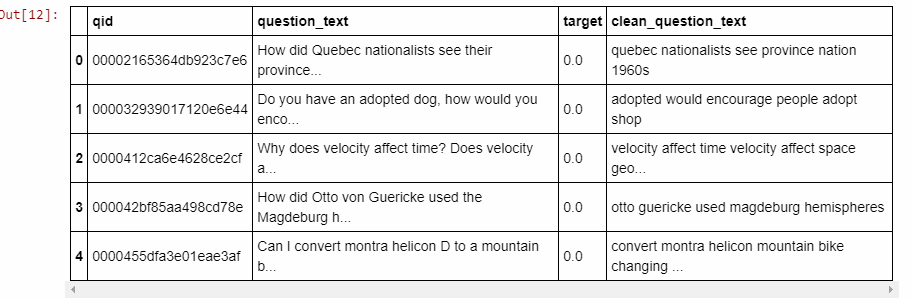




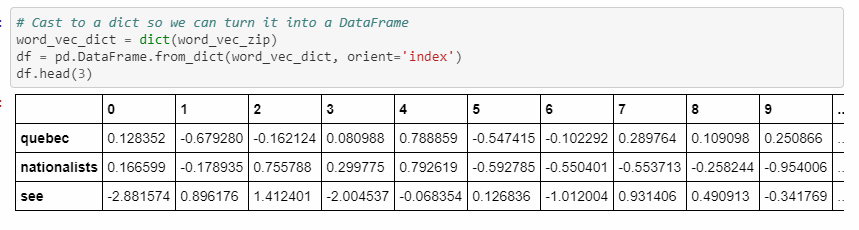


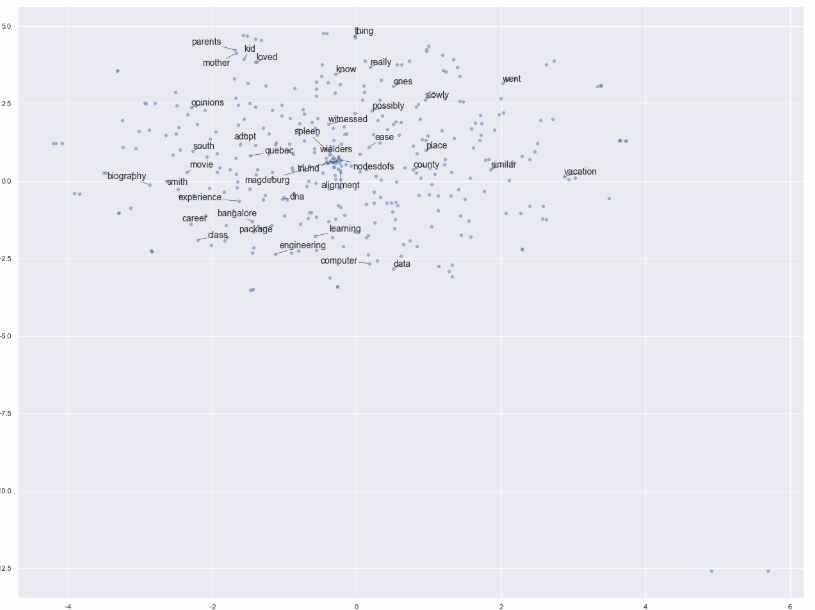


* 1. Utilized NLTK to clean sentence, removed stop words.



* 1. Utilized Word2Vec to visualize word embedding.





Utilize Deep Learning

1. Single Bidrectional LSTM, Single BatchNormaization, Double Dense

Test loss: 0.12538052813606687

Test accuracy: 0.952421111732375

1. # Added Bidirectional GRU layer, Single Batchnormalization, Double Dense

Test loss: 0.12459266796889229

Test accuracy: 0.9527122808404213

1. Different kernel initializer - RandomNormal for Conv1D layers

Test loss: 0.12350578806582761

Test accuracy: 0.9535475069661562

1. Single GRU, Double batchnormalizations. Double dense

Test loss: 0.12390515017195507

Test accuracy: 0.9533154997088309

1. Double Bidirectional GRU layers, Double Batchnormalization, Double Dense

Test loss: 0.11734020953069065

Test accuracy: 0.9550532340651615

1. Decrease units to 64 and build model 5 again

Test loss: 0.12014513911110418

Test accuracy: 0.954277169789871

1. One GRU, LSTM, Single Batch

Test loss: 0.22964484459516035

Test accuracy: 0.9390354530292473

1. One GRU, LSTM, Single Batch

Test loss: 0.22964484459516035

Test accuracy: 0.9390354530292473

**Conclusion & Next Steps**

1. Best model seen with double bidirectional GRU layers, with double batch normalization and double dense layers for regularization.
2. Consider different layers for regularization
3. Train on a system with more processing power because deep learning took too long to train.
   1. Consider researching multiprocessing with deep learning training.