Problem:

An important 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.

In this capstone we develop models that identify and flag insincere questions. A question is assumed to be "insincere" if it is intended to make a statement rather than look for helpful answers.

Some of the characteristics we consider for insincere questions are:

- Having a non-neutral tone
- Being disparaging or inflammatory
- Being based on false information or absurd assumptions
- Uses sexual content
- (incest, bestiality, pedophilia) for shock value, and not to seek genuine answers

Data:

Data is from Kaggle's dataset: https://www.kaggle.com/c/quora-insincere-questions-classification/data.

Training data includes the question that was asked, and whether it was identified as insincere (target = 1) with following fields:

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

Scope:

Investigate and compare different methods including naïve Bayes, logistic regression and CNN.