

Contents

1 Quora Insincere Questions Classification	1
1.1 Problem Statement	1
1.2 Submission	1
1.3 Articles	2

1 Quora Insincere Questions Classification

- Competition Link
- Difficulty: Medium-Hard
- Problems: NLP, text analysis
- ML Problem Type: Classification
- Possible ML Approaches: Naive-Bayes (TF-IDF), SVM, CNN

1.1 Problem Statement

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 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.

In this competition, Kagglers 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 your help, they can develop more scalable methods to detect toxic and misleading content.

Here’s your chance to combat online trolls at scale. Help Quora uphold their policy of “Be Nice, Be Respectful” and continue to be a place for sharing and growing the world’s knowledge.

1.2 Submission

For each qid in the test set, you must predict whether the corresponding question_text is insincere (1) or not (0). Predictions should only be the

integers 0 or 1. The file should contain a header and have the following format:

```
qid,prediction
0000163e3ea7c7a74cd7,0
00002bd4fb5d505b9161,0
00007756b4a147d2b0b3,0
```

1.3 Articles

- Deep Learning for NLP: An Overview of Recent Trends
-