Analysis For TF-IDF Homework

Analysis 1: What answers get confused with each other most easily? What kinds of mistakes does this guesser make?

For question that answer is confused with guesses, its answer has a higher probability to be a specific name, e.g. name for a human or a place, that has no or few synonyms.

- Answer Examples: Mass_(music), Byzantine_Empire, Russian_Empire, Sandman, Los_Angeles, Lord_Byron, Matthew_Arnold, S-waves, Density, Georgia_(country), Luminosity, Apollo

By counting the number of answers' synonyms of three classes of results (close, hit and miss), it can be found that confused answer is more concentrate around 0, while answers in the other two classes tend to have more synonyms.
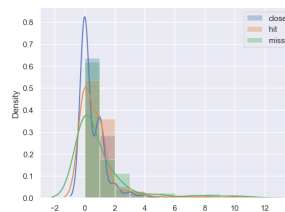

Figure 1 Count of Synonyms for Close, Hit and Miss Results

According to the question, answer and guesses in miss category, it is found that the guesser is overfitting the training dataset. The f1 score on training is 0.97, which is much higher than F1 score on testing dataset (0.51). From cases study, it is found that the guesser prefers to select guesses which words appear in the original question (see Figure 2). Except this, the guesser also fails to find answers based on multiple given information. Except for constraint of psychologist, there are other constraints for the person. But the guesser wrongly found the other psychologist (see Figure 3).


```
guess: Georg_Simmel
answer: Federal_Republic_of_Germany
    id: 102343
  text: A sociologist from this modern-day country inspired Joseph Schumpeter
        by first using the term "creative destruction" in War and Capitalism,
        and argued that the ability of America's two party system to co-opt
        third party movements helped answer the question "Why is there no
        Socialism in the United States?" A sociologist from this modern-day
        country argued that society's taste in clothing is a social creation
        that has the function of "class distinction" in his essay "Fashion."
        That man from here stated that a blasé outlook characterizes people
        who developed a "protective organ" in a book examining the
        psychological effects of the city. That man from this modern country
        argued that money functions as an "impersonal agent" that bridges the
        middle distance in another book. The Metropolis and Mental Life and
        The Philosophy of Money are from, for 10 points, what home country of
        Werner Sombart and Georg Simmel?
```

Figure 2 Preference in Repeat of the Words in Question


```
guess: Abraham_Maslow
answer: Albert_Ellis
    id: 102175
  text: With Milton Blum, this thinker wrote a book using a "tough shit"
        approach about business called Executive Leadership. He coined the
        terms "awfulizing," and "catastrophizing" to describe behavior that
        should be fought with shame-attack exercises. He called the harmful
        act of imposing absolute values on the world "musturbation." Many of
        his ideas, such as the distinction between "discomfort disturbances"
        and "ego disturbances," were developed while he writing Sex Without
        Guilt and The Art and Science of Love. This thinker developed a model
        of causation in which an activating event induces a belief, which
        itself is responsible for a human feeling – that is the "ABC" model.
        Like Aaron Beck, this psychologist developed an individualist approach
        that seeks to correct harmful behavior by converting socially
        destructive tendencies to constructive ones. For 10 points, name this
        psychologist who developed rational emotive behavioral therapy.
```

Figure 3 Unable to Handle Multiple Constraints

Analysis 2: Where does your buzzer make mistakes? How might you further improve the buzzer?

According to the coefficient of logistic regression classifier, the buzzer mainly relies on the confidence of tf-idf guesser. This means when guesser make a wrong guess but with a high confidence, the buzzer cannot figure it out. For aggressive category of the evaluation result, while facing the guess that has the same words with the question, the buzzer is more likely to give a higher buzzing probability.

To further improve the buzzer, a different tf-idf guesser could be trained to generate a difference confidence score for this question & guesser. For example, using the Wikipedia page and doc as the data source to train an additional guesser.

Besides, a feature presents whether the word from guess is in the question, especially when guess is name and other specific definition could be added.