David Schulz

CS 4980 001

5/3/2022

**Prototyping**

Once I got my model in a state where it could give an answer, correct or not, obtained from the knowledge database, I ran it through the 3 categories of questions from the QuestionsCorpus that shared categories with the knowledge database: blocks, mobs, and items. I went through every Q/A combination and manually determined if each answer was correct and fully answered the question. As expected, the performance was not great. Out of 460 questions, only about 57 were fully correct, giving an accuracy of about 12%. Some of the correct answers are shown below.

Correct Answers:

1. Q: Can I mine a Diamond without a Pickaxe?
   1. A: Blocks of diamond can only be mined using an iron or diamond pickaxe.
2. Q: What is the most rare Ore in Minecraft?
   1. A: Emerald Ore is the rarest block in Minecraft.
3. Q: What has the highest blast resistance?
   1. A: End Stone has the highest blast resistance of any block that can be blown up easily.
4. Q: How do Pressure Plates help?
   1. A: A Pressure Plate is an item that can be used to activate certain features when it is stepped on.
5. Q: What else destroys Cobwebs besides Swords?
   1. A: Water, lava, and pistons will destroy the cobweb as well.

This was expected because the way the model works is it chooses 3 files that are most likely to have the answer based on the tf-idf tables for each category. It then reads every line of all 3 files and parses out all of the sentences. It then converts the question and every sentence from the 3 files to vectors and finds the cosine similarity between the question and every sentence. The most similar sentence is chosen as the answer. This approach creates a very naïve model.

As a result of this performance, I realized that the knowledge database still needed severe cleaning and standardized formatting. Some sentences weren’t information at all, but just scraped labels of images that were on the websites, so they were just a few words instead of complete sentences. I decided to manually clean the files about chickens and ask my own questions about them to see if that improved performance. Since all text in those files were now complete sentences with pure information, I could word the questions somewhat similarly to how they were worded to get full, correct answers more consistently. Some of those examples are shown below.

Chicken Questions:

1. Q: How can chickens be bred?
   1. A: Chickens can be bred using seeds.
2. Q: What do chickens drop when they are killed?
   1. A: Chickens drop 0-2 feathers and 1 raw chicken when they’re killed.
3. Q: What is the chance of a baby chicken to spawn from an egg?
   1. A: There’s a 1/8 chance of a baby chicken hatching from an egg.
4. Q: How often does a chicken lay one egg?
   1. A: A chicken will lay one egg every 5 to 10 minutes, unless it’s part of a chicken jockey.

The current performance of the model is not close to my vision statement’s desired performance at all. Very few questions actually get answered fully and correctly. With more dedicated time and some better resources, other approaches would create a much better model, like using an ontology for example.