

Chatbot Project

System description:

Me and Zach were interested in neural networks tools, so we wanted to build a chatbot using machine learning and NLTK techniques.

The system uses MongoDB as the backend database, python as the environment, and TensorFlow to train models.

The requirements are:

- To create a bot that can have a simple conversation about a topic, which is football in our case.
- The user input should be understood, and a response should be returned from the chatbot.
- Understand the context in a conversation to reduce the number of responses to be more contextually relevant.
- Store separate user models

Process/ Design

As mentioned above, the system uses Mongo DB to store documents that have:

- A list of responses to return to the user.
- A list of patterns that the user might say that are used to build the model to predict probabilities.
- The name of the label to determine the correct classification.
- A context value of the intent used to guide or filter which response lists makes sense to return.

The system uses TensorFlow to build models from the database information. This will allow users to interact with the bot.

The chatbots model uses a prebuilt predictive model and then uses the backend DB to retrieve documents which have possible responses and context information.

The Process is simple

Capture user input -> process the input using NLTK and predict an intent category using the neural network model -> randomly pick a response from the document store in the DB.

Implementation

NLTK

1. Read the data from the chatbot intent collection.
2. Tokenize each pattern
3. Parse the data into array of stemmed words
4. Parse data array of intent classifications
5. Parse data array of documents with words for a pattern tagged with the classification name.

TensorFlow Model

After categorizing the data, we created a training set.

1. Create deep neural network
2. Evaluate the data set to create predictive model from the patterns in the DB
3. Convert the dataset into NumPy array
4. Partition the dataset into input data and possible outcomes for each input

Neural Network

1. Set shape and the number of layers
2. Used regression algorithm to fit the model
3. The predictive model is produced by TensorFlow using a Deep Neural Network, using the defined training data
4. Store the model by pickling the file along with its words, classes, and training data for the chatbot to use

Chatbot Structure

Process

1. Load the training data made from building the model
2. Create neural net with the same size and shape of the loaded model.
3. Load the predictive model into the neural network
4. Ask the user for input to have an interaction with the bot
5. For every input, classify which intent it belongs to and pick a random response for that intent

Evaluations of the chatbot

Strengths and Weaknesses:

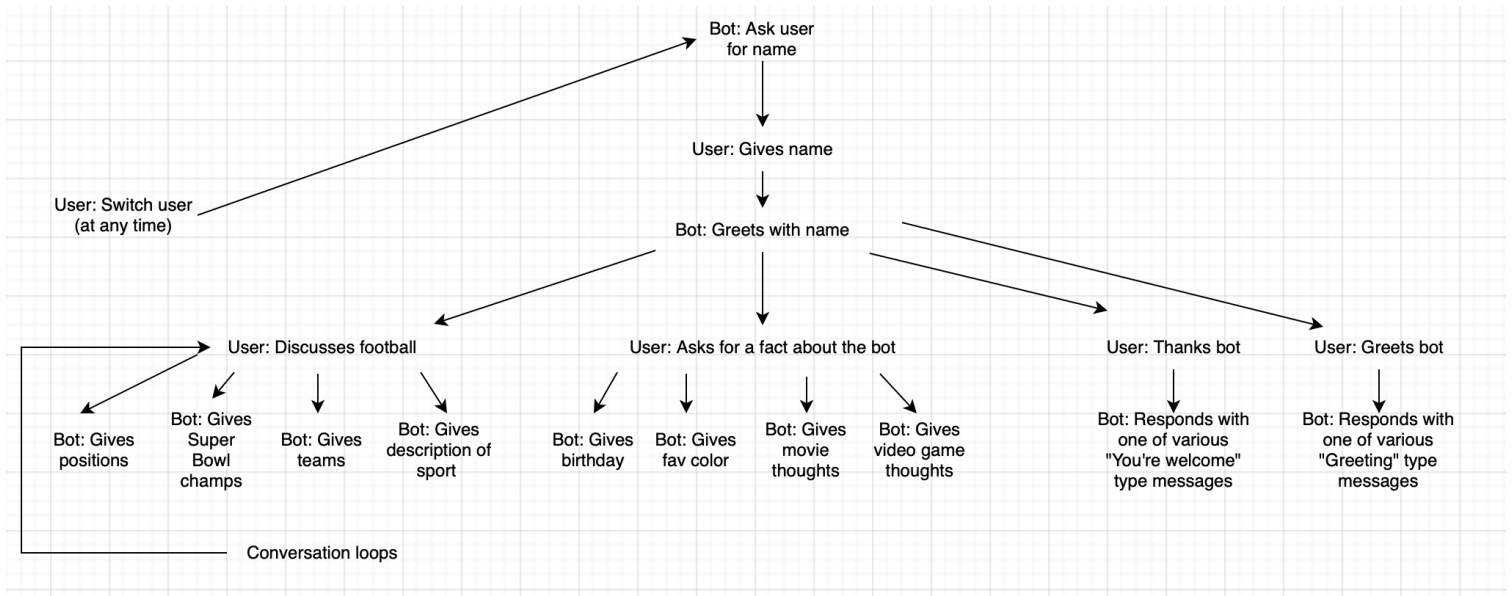
Strengths

- MongoDB implementation to maintain chatbot data.
- Powerful Neural Network training to best analyze user input and give correct output.
- Various conversation paths, including more than just our main topic of football.
- Built using TensorFlow machine learning.
- Uses NLTK tokenization and other NLTK libraries for better NLP of user input.
- Ability to change between users on the fly without restarting bot.
- User models written to and maintained in individual files.

Weaknesses

- Database for responses and user input patterns is static, must be updated by developers for new information.
- TensorFlow environment can be difficult to setup if no prior experience.
- Uses random selection for user output based on user input category.
- User model does not use former input for decision making.
- User input stores name and basic conversation, not certain user preferences.

Dialog Tree



Link: <https://app.diagrams.net/#G1dyeVkobcqmcfYBFdDaq3jxfjqJf8tg6>

Sample Dialog

```
Type "quit" to exit.
type 'switch to change users
This is a chatbot that can talk to you about American football.
ChatBot: What is your name?
User: Nebil
ChatBot: Hello, Nebil
Nebil: football
Chatbot: OK. Lets talk about Football.
Nebil: who won the super bowl
Chatbot: Tom Brady won his seventh and final Super Bowl with his new team in his home stadium, besting the defending champion Chiefs
Nebil: thank you
Chatbot: Think nothing of it
Nebil: switch
ChatBot: Hello new user, what is your name
newUser: Zach
ChatBot: Hello, Zach
Zach: what is your favorite team
Chatbot: I like the Cowboys
Zach: what is football
Chatbot: I love to play football!
Zach: what is football
Chatbot: One of the most popular competitive sports in the United States of America is American football.
Zach: switch
ChatBot: Hello new user, what is your name
newUser: John
ChatBot: Hello, John
John: tell me a fact
Chatbot: My favorite color is purple.
John: Tell me more about your self
```

Appendix for the knowledge base

Here is the some of the knowledge base and its format.

```
{
  "_id" : ObjectId("5a160efe21b6d52b1bd58ce5"),
  "patterns" : [
    "Hi",
    "How are you",
    "Is anyone there",
    "Hello",
    "Good day"
  ],
  "responses" : [
    "Hello, thanks for talking to me.",
    "Hi there, how can I help?",
    "Hello",
    "Hey there."
  ],
  "name" : "greeting",
  "contextSet" : ""
}
{
  "_id" : ObjectId("5a160f7421b6d52b1bd58ce6"),
  "patterns" : [
    "Bye",
    "See you later",
    "Goodbye",
    "Good Bye",
    "Adios"
  ],
  "responses" : [
    "See you later, thanks for visiting",
    "Have a nice day",
    "Bye! Come back again soon.",
    "Adios",
    "Goodbye",
    "I would love to chat with you again sometime. Goodbye for now."
  ],
  "name" : "goodbye",
  "contextSet" : ""
}
{
  "_id" : ObjectId("5a16132321b6d52b1bd58ce7"),
  "patterns" : [
    "Tell me about yourself",
    "What is a fact about yourself?",
    "I'd like to know more about you.",
    "Give me a fun fact."
  ],
  "responses" : [
    "My birthday is November 12th, 2022",
```



```

        "My favorite color is purple.",
        "I love movies, especially horror!",
        "I play video games, I am basically one myself!"
    ],
    "name" : "conversation",
    "contextSet" : ""
}
{
    "_id" : ObjectId("5a17600021b6d52b1bd58ce8"),
    "patterns" : [
        "Thanks",
        "Thank you",
        "Thats helpful, Thank you very much",
        "Thanks a bunch"
    ],
    "responses" : [
        "You are welcome",
        "You are welcome. Anything else I can help you with?",
        "Its no bother",
        "Dont mention it",
        "No problem",
        "Happy to help!",
        "Any time!",
        "Its my pleasure",
        "Think nothing of it",
        "Sure",
        "Sure thing"
    ],
    "name" : "thanks",
    "contextSet" : ""
}
{
    "_id" : ObjectId("5a17662021b6d52b1bd58ceb"),
    "patterns" : [
        "what is football",
        "Can we talk about football",
        "Is football a sport.",
        "Explain American football to me."
    ],
    "responses" : [
        "American football also known as gridiron, is a team sport played by
two teams of eleven players on a rectangular field with goalposts at each
end.",
        "One of the most popular competitive sports in the United States of
America is American football.",
        "I love to play football!",
        "I used to play running back, I can talk football all day!"
    ],
    "name" : "football",
    "contextSet" : "knowledge"
}
{
    "_id" : ObjectId("5a176be221b6d52b1bd58cee"),
    "patterns" : [
        "what is your favorite team",
        "Do you like any particular team",

```

```

        "who do you support",
        "what football team do you like"
    ],
    "responses" : [
        "I like the Cowboys",
        "No, I like them all",
        "Lets go Cowboyssssssss!"
    ],
    "name" : "team",
    "contextSet" : ""
}
{
    "_id" : ObjectId("5a1b4dd721b6d52b1bd58d04"),
    "patterns" : [
        "Who won the super bowl?",
        "team",
        "who do you support",
        "Explain American football to me."
    ],
    "responses" : [
        "The Super Bowl is the annual final playoff game of the National
Football League (NFL) to determine the league champion.",
        "The New England Patriots began the year strong, beating the Atlanta
Falcons .",
        "Tom Brady won his seventh and final Super Bowl with his new team in
his home stadium, besting the defending champion Chiefs"
    ],
    "name" : "superBowl",
    "contextSet" : ""
}
{
    "_id" : ObjectId("5a1b4df721b6d52b1bd58d05"),
    "patterns" : [
        "Football"
    ],
    "responses" : [
        "OK. Lets talk about Football."
    ],
    "name" : "FootballSubjectAnswer",
    "contextFilter" : "knowledge"
}
{
    "_id" : ObjectId("5a17600021b6d52b1bd58ce5"),
    "patterns" : [
        "Did anyone win today?",
        "I wonder who won",
        "Did you watch the game?",
        "Do plays today"
    ],
    "responses" : [
        "The Cowboys play.",
        "You know the Cowboys."
    ],
    "name" : "win",
    "contextSet" : "knowledge"
}

```

Appendix for sample user models

The program stores user model in separate files as shown below:

File one (Nebil_model.txt)

```
Name: Nebil
  userResponse: football
  Response: OK. Lets talk about Football.

Name: Nebil
  userResponse: who won the super bowl
  Response: The Super Bowl is the annual final playoff game of the National Football
League (NFL) to determine the league champion.

Name: Nebil
  userResponse: thank you
  Response: Think nothing of it

Name: Zach
  userResponse: what is your favorite team
  Response: No, I like them all
```

File two (Zach_model.txt)

```
Name: Zach
  userResponse: what is football
  Response: One of the most popular competitive sports in the United States of America
is American football.

Name: Zach
  userResponse: what is football
  Response: I love to play football!
```

File three (John_model.txt)

```
Name: John
  userResponse: tell me a fact
  Response: My favorite color is purple.

Name: John
  userResponse: Tell me more about your self
  Response: Adios

Name: John
  userResponse: tell me more
  Response: I would love to chat with you again sometime. Goodbye for now.

Name: John
  userResponse: bye
  Response: See you later, thanks for visiting
```