BobaBot

...

An Interactive chatbot

Sanskriti Seernani, Matthew Chong

Outline

- Scope Comparison: The Evolution of Our Project
- Our Journey: Challenges Faced
- Conclusion/Summary
- Demo

The Evolution of Our Project From CI 102 to Now

- Initial Expectations
  - An interactive chatbot
  - Responds intelligently to user input
  - Voice recognition
  - Audio Response
  - A well formed GUI of bubble tea with a smiley face

- Meeting with Professors Burlick and Popyack (AI + Machine Learning)
  - Encouraged self directed research in the field
  - Possibility of using an API
  - Initial scope was far too advanced given our mere experience
  - Implementing voice recognition and an audio-visual aspect in the bot required complex knowledge that is often taught at the graduate level
  - Ideal to limit ourselves to implementing intelligent responses in textual form

Evolution of Project cntd..

- Multiple Machine Learning Techniques
  - Suggested by Dr. Burlick
  - Bag of words method
  - Nearest neighbor algorithm
  - Natural Language Processing

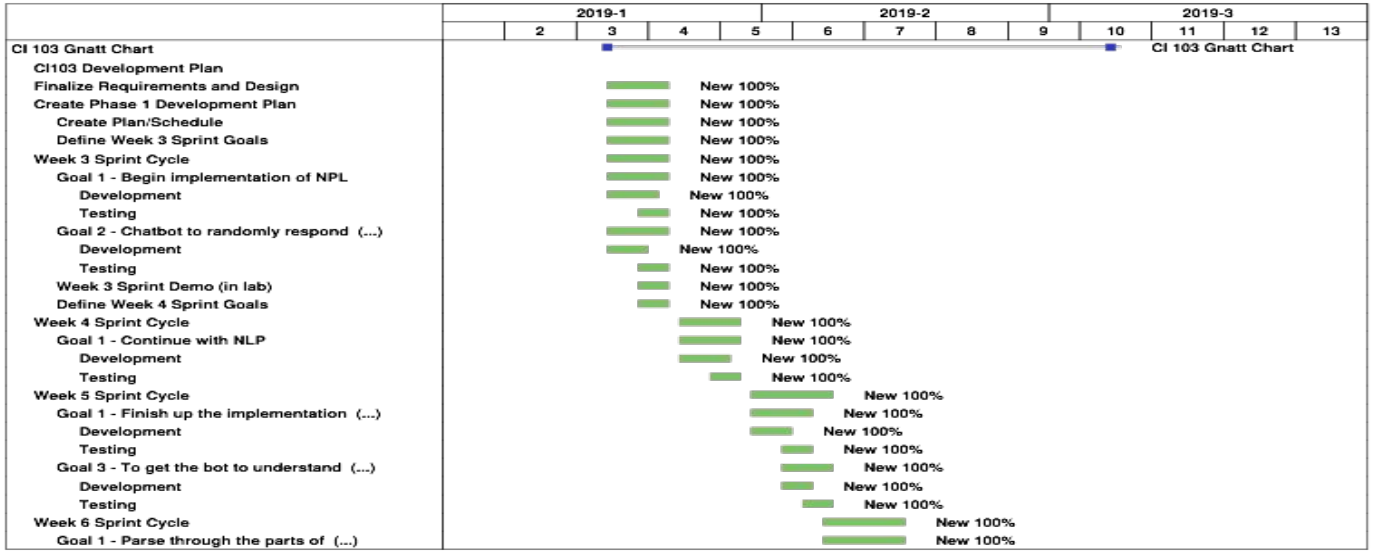
- Method Chosen for Implementation
  - Bag of words and nearest neighbor beyond our scope
  - Simple implementation of an array with possible user inputs
  - Get bot to match input to appropriate array element
  - Have another array store responses
  - Based on user input, pick a response from responses array
  - Print to screen
  - Intelligent handling of unfamiliar user inputs

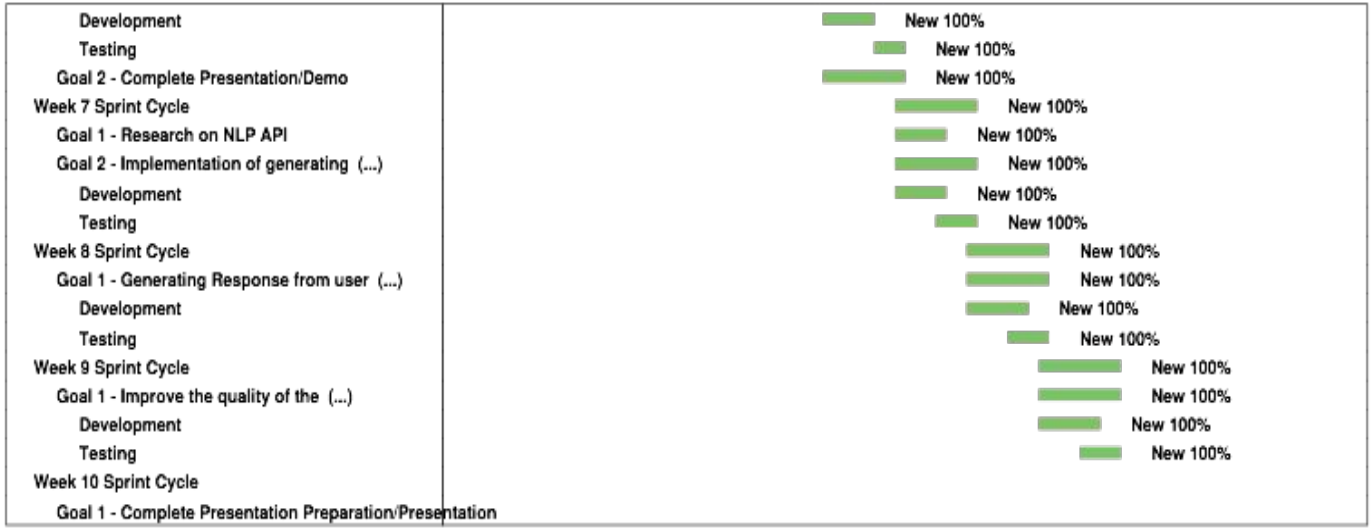
Evolution of Project cntd..

- Further Development
  - Looked into Google API based on Costa's suggestion
  - Found GPT-2 to be a more suitable choice
  - Integrated the API into our code
  - GPT-2 takes in user input
  - Uses trained sample models given to bot to generate a response
  - Functions like a ML Bot

Gnatt Chart

CI 103 Gnatt Chart




Development
Testing
Goal 2 - Complete Presentation/Demo
Week 7 Sprint Cycle
Goal 1 - Research on NLP API
Goal 2 - Implementation of generating (...)
Development
Testing
Week 8 Sprint Cycle
Goal 1 - Generating Response from user (...)
Development
Testing
Week 9 Sprint Cycle
Goal 1 - Improve the quality of the (...)
Development
Testing
Week 10 Sprint Cycle
Goal 1 - Complete Presentation Preparation/Presentation
New 100%

Summary of Project Activities

- Initially, we...
  - Parsed through user input to get different parts of speech
  - Tried to create our own bot that would take in parts of speech and generate responses
  - Due to time constraints, might not be possible to generate coherent responses
- Researched different API (recommendations from Costa)
  - Looked at Google NLP API
  - Looked at GPT-2 open AI
- Ended up picking GPT-2 based off the needs for the project

Information about the GPT-2 Model

- Natural language processing tasks are approached with supervised learning on specific databases
- Language models begin to learn these tasks without any explicit supervision when trained on a new dataset of millions of webpages called WebText.
- Learn to perform tasks from their naturally occurring demonstrations
- When given new input, it will go through the different dataset it was trained
    - The bot will generate a response based off the dataset

Challenges Faced

- Further learning of the nltk library and learning GPT-2
- Thinking about how to parse through the tuple of POS and pick out important aspects of the sentence
- Completing the project on time/within project scope

Conclusion

- We were able to generate a response though the use of GPT-2
- Project ended up not being what we envisioned, but we are still proud of what we accomplished