

EBC 5125: Data Science Applications

Spring/Summer 2021

Assignment 4: Intelligent Chatbot (Report)

Submitted by:

Juveriya Zameer (300116986)

## Table of Contents

[OBJECTIVE 3](#_bookmark0)

[ONTOLOGY BY USING PROTÉGÉ 3](#_bookmark1)

[COMPETENCY QUESTIONS 5](#_bookmark2)

[CHATBOT TRAINING USING DIALOG FLOW 5](#_bookmark3)

[TEST CASES 5](#_bookmark4)

[EVALUATION 5](#_bookmark5)

[CONCLUSION 28](#_bookmark6)

# OBJECTIVE

The overall objective of this assignment is to create an intelligent chatbot agent capable of answering questions from specific data science concepts. First, I have developed the ontology using the Protégé and I have selected the DSA class lectures as the basis of our material to do so. The next step involved preparing competency questions of six sets in order to create an agent which is capable of answering questions according to the intents made and also involved preparing test cases to check the quality while testing our chat-bot agent.

# ONTOLOGY BY USING PROTÉGÉ

The ontology for the topic is developed by using the tool Protégé. The base material for creating this ontology is derived from the class lectures of the EBC5125: DSA course. Protégé is an open source tool which is used to edit the ontology and a knowledge management system. It provides a graphic user interface to define the ontology. I can check if the models are consistent by the deductive classifiers. To create an ontology by using Protégé the first step is picking a domain, define the classes and class hierarchy, define the properties of the classes and create the instances and then create a new OWL ontology.

Before creating the ontology in I actually drew the domain it all its classes, subclasses and subclasses of subclasses with pen and paper. After analyzing it, I tried to portrait the ontology individual classes and top object properties in Protégé.

The following ontology for our basis material:

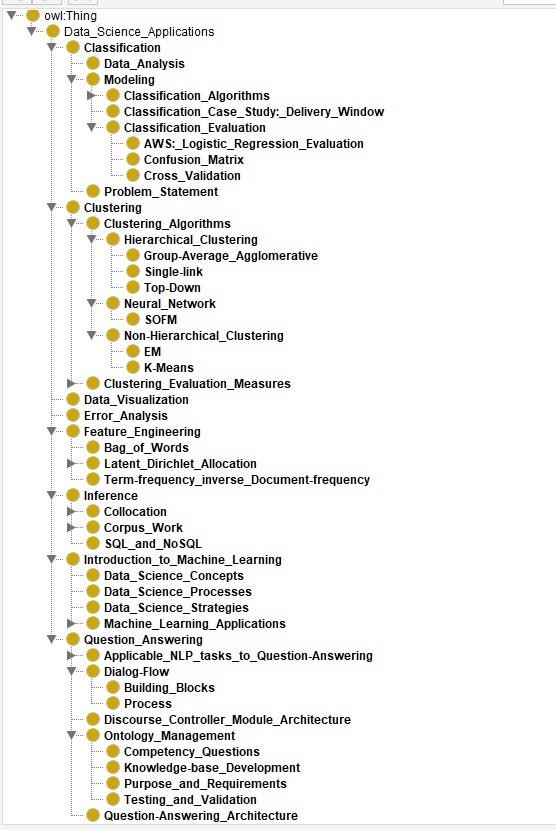


Figure: Individuals by Class (created in Protege)

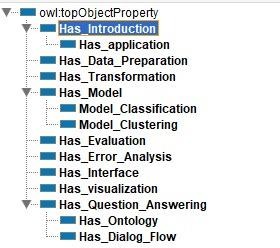


Figure: Object Properties (created in Protege)

# COMPETENCY QUESTIONS

I have created a set of six competency questions as mentioned in the problem statement. The competency questions are made according to the ontology created. The questions are mentioned below later in the report.

# CHATBOT TRAINING USING DIALOG FLOW

Dialog flow is owned by Google and is based on natural language conversation. It’s a platform which is used to create an application in a conversational manner for different company’s customers in various languages on multiple platforms. First step is to create an agent and then using the competency questions I have created different intents for different questions. I trained our agent on the basis of the competency questions and tested it with the test cases.

# TEST CASES

To check the quality and performance of the chat-bot agent I have made test cases every competency question. These test cases can be used to test the chat-bot agent after it is trained with the training phrases.

# EVALUATION

I have used the test cases to verify the output for the particular questions.

The following part of the report will contain some example of Competency Questions, the associated test cases and the instance of the chatbot when asked questions from the set of test cases for a particular competency question (intent):

Competency Question:

1. What is Data Science? Test Cases:

 What do you mean by Data Science?

 Define Data Science.

 What is data science in simple words?

 Describe what’s Data Science?

 How do you define data science?

 What is data science for beginners? Chatbot:

|  |  |
| --- | --- |
|  |  |

1. What are the applications of Data Science? Test Cases:

 Where is data science applied?

 What is the purpose of data science?

 What is the importance of Data Science?

 How do you implement Data Science in real Life?

 What are the different types of processes in Data Science? Chatbot:

|  |  |
| --- | --- |
|  |  |

1. What is the difference between Machine Learning and Data Science? Test Cases:

 Distinguish Data Science and Machine Learning?

 How is Machine Learning different from Data Science?

Chatbot:

|  |  |
| --- | --- |
|  |  |

3. What is Feature engineering in the context of Data Science? Test Cases:

 What is feature engineering in data science?

 Define Feature Engineering.

 What do you mean by Feature Engineering?

 Explain Feature Engineering.

 What the hell is Feature Engineering?

Chatbot:

|  |  |
| --- | --- |
|  |  |

#### Competency Question:

What are the different types of feature engineering?

Test Cases:

 Branches of Feature Engineering.

 Different categories of feature engineering.

 What are the various types of feature engineering?

 Kinds of Feature extraction.

 List the various types of feature extraction methods in data science?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is a Bag of Words? Test Cases:

 What do you mean by Bag of words?

 Bow

 Is bag of words a transformation technique?

 What is BOW used for?

 Define Bag of Words.

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is LDA? Test Cases:

 What do you mean by LDA?

 Latent Dirichlet Allocation

 Define LDA

 LDA stands for?

 Is LDA used for transformation?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is TF-IDF? Test Cases:

 term frequency and inverse document frequency?

 tf-idf technique.

 What do you mean by Tf-idf?

 What is tf-idf used for?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is word embedding? Test Cases:

 what do you mean by word embedding?

 use of word embedding.

 meaning of word embedding?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is classification in the context of data science? Test Cases:

 What do you mean by Classification in machine learning?

 Why is classification needed?

 What is the motive of classification?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What are the different kinds of classification models ? Test Cases:

 Types of classification model

 techniques of classification modeling

Chatbot:

|  |  |
| --- | --- |
|  |  |

#### Competency Question:

What is clustering?

Test Cases:

 what is the meaning of clustering

 ii)give me some clue of clustering

Chatbot:

|  |  |
| --- | --- |
|  |  |

#### Competency Question:

What are the different kinds of clustering model? Test Cases:

 types of clustering models

 techniques of clustering

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is an evaluation technique? Test Cases:

 why I use evaluation techniques

 examine models in data science

Chatbots:

|  |  |
| --- | --- |
|  |  |

What are the evaluation techniques used in classification?

Test Cases:

 how can I test our supervised data model

 suggest me evaluation technique of classification model

Chatbot:

|  |  |
| --- | --- |
|  |  |

What are the evaluation techniques used in clustering? Test Cases:

 evaluation of unsupervised data

 determine the evaluation score of clustering

Chatbot:

|  |  |
| --- | --- |
|  |  |

#### Competency Question:

What is the question answering architecture in data science?

Test Cases:

 what do you mean by question answering

 Why question answering is used?

Chatbot:

|  |  |
| --- | --- |
|  |  |

#### Competency Questions:

What is a dialog flow?

Test Cases:

 Create a dialog flow

 the process of making a dialog flow

Chatbot:

|  |  |
| --- | --- |
|  |  |

What are the components of dialog flow? Test Cases:

 Definition of parts of dialog flow

 Describe components of dialog flow?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is ontology?

Test Cases:

 what actually ontology is?

 Explain the process of ontology.

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is a recommender system?

Test Cases:

 How recommender system works?

 What are the basics of recommender system?

Chatbot:

|  |  |
| --- | --- |
|  |  |

What is sentiment analysis in NLP?

Test Case:

 What sentiment analysis stands for?

 What I analyze in sentiment analysis?

Chatbot:

|  |  |
| --- | --- |
|  |  |

# CONCLUSION

Protégé tool has helped to get a clear view about the classes, subclasses and the hierarchy which is required for creating a chat-bot,using this tool I created a skeleton for the chat-bot. The competency questions helped us create an exact intent in the chat-bot with the right set of training phrases. After creating the agent for the chat-bot I have been able to understand what kind of training phrases should be used to train the chat-bot in order to get the required output. A set of well-defined training phrases has given us the required output for various questions. I have verified whether the chat-bot gives the right output for a particular question or multiple questions with the same meaning using the test cases that were created for every intent made in the chat-bot.