Intelligent (Task-Oriented) Conversation Assistant for Course Selection

Progress Report



Information Technology Capstone Project

COMP5703

Group Members

1. Shengyuan Sun (460257820)
2. Biying Wang (470067400)
3. Rui Chen (470111585)
4. Quan Chen (470199228)

Table of Contents

[Table of Contents i](#_Toc521413584)

[1. Progress Status 1](#_Toc521413585)

[2. Roles & Responsibilities 2](#_Toc521413586)

[3. Individual Achievements 2](#_Toc521413587)

[4. Group Collaboration 2](#_Toc521413588)

# Progress Status

|  |  |
| --- | --- |
| **Project Name** | CS17 Intelligence(Task-Oriented) Conversation assistant for course selection |
| **Project Start Date** | 06/ 3/ 2019 |
| **Project Manager** | Biying Wang |

|  |  |
| --- | --- |
| Project Description | The project focuses on using NLP and other technologies to build a dialog system to analyse and answer the questions that students ask in terms of course units on University of Sydney education system |

|  |  |  |
| --- | --- | --- |
| Project Status Report | # | Date: 15/ 3/ 2019 |

|  |  |  |
| --- | --- | --- |
| **Status Item** | **Status up to last week** | **Planned for this week** |
| **Major deliverables** |  | Initial project architecture and Initial Requirement list |
| **Planned delivery date** | 7/ 3/ 2019 | 13/ 3/ 2019 |
| **Major issues** | Forming team  Introduce each other  Discuss the general understand of the topic | Identify all interface and structure of the project  Understanding the special requirement of the project |
| **Major risks** | The misunderstand of the task.  Group members issue | Misunderstand the requirement  Designed an unachievable requirement |
| **External dependencies** | Slack Github | Slack Github |
| **Estimated effort (hr)** | 20hr | 20hr |
| **Recorded effort (hr)** | 30hr | 40hr |
| **Status (R, Y, G)** | Green | yellow |

# Roles & Responsibilities

Shengyuan Sun

Role: developer,

Responsibility:

* system architecture design
* intent classification methods review
* design top-level architecture

Biying Wang

Role: Project manager; Data collector

Responsibility:

* Check whether weekly deliverables is done by each member on time; make sure project progessing is successful
* Write meeting minutes as a recording
* Check meeting time, location and topic for group
* Collect information for dataset designing
* Design dataset structure
* Help and discuss with other team member for project work

Rui Chen

Role: Analyser and Developer

Responsibility:

* Collect testing corpus/dataset (MIT-Moive corpus)
* Analysing MIT-Moive corpus structure
* Reading the rule-based chatbot paper and analyse intent-slot model
* Discussing and analysing the top-level architecture of cassandra system
* Design a sample questions for interview
* Analysing the CUSP web structure for crawling
* Design the working flow graph for further co-working

Quan Chen

Role: Developer

Responsibility:

* Back-end database system building and data collecting
* Help to design the whole process of system (such as slot and intent)

# Individual Achievements

1. Wrote initial requirement list

After weekly discussion with Caren and Divad, The start point of our project is to define the requirement and initial project architecture.

Based on the blogs and literature review. First, the gerneal task-oriented chatbox need four part, NLU(nature language understanding), dialog state traking, policy learning and NLG(nature language generation).

A picture containing clock

Description automatically generated

In our project, the NLU is the first issue, so the popular methods for NLU are

* Attention-based recurrent neural network models
* Capsule Neural Networks
* Slot-Gated (Intent Atten.)

In initial version production, there no need dialogue state tracking and policy learning , then our team design a rule-based syetem and dialogue recoding system.

1. the idea of the rule based system: design all rules for the question
2. design SQL for searching the database
3. design all responses type

combine with the database ,the system can direct find infroamation.

Last part is how to generate the nature language for answer. Which we consider design a template answer table for the initial version of the project.

The link of the report : <https://drive.google.com/file/d/1OCqSokNG7o-Q-UGWZy32lelt-imOAJEZ/view?usp=sharing>

A screenshot of a cell phone

Description automatically generated

1. Review literature about Intenet classification

I find some method from the Github the link are:

|  |  |  |  |
| --- | --- | --- | --- |
| model | dataset | acc | link |
| Attention-based recurrent neural network models | ATIS | 98.43 | [Attention-based recurrent neural network models for joint intent detection and slot filling](https://arxiv.org/pdf/1609.01454.pdf) |
| Capsule Neural Networks | ATIS | 95.0 | [Joint Slot Filling and Intent Detection via Capsule Neural Networks](https://arxiv.org/pdf/1812.09471.pdf) |
| Slot-Gated (Intent Atten.) | ATIS | 94.1 | [Slot-Gated Modeling for Joint Slot Filling and Intent Prediction](http://www.aclweb.org/anthology/N18-2118) |

And I test the Attention-based recurrent neural network models with an other dataset which is MIT-Moive corpus my colab Link: <https://colab.research.google.com/drive/1gqZdEd_1a9FMOExkdUKG0jXfyNhw_YM1>

A screenshot of a social media post

Description automatically generated

1. Wrote a descirbtion for high-level work-flow diagram

As we discuss the structure of the Top-level structure ,Rui Chen start to draw the workflow diagram and I wrote a describtion of the each stage and some details of the work-flow diagram. The link : <https://drive.google.com/open?id=1J_Esx7r-8frP24f38g8hPllHhZnhp_6ATY_Fxw9M-zg>

A screenshot of a cell phone

Description automatically generated

# Group Collaboration

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated

A screenshot of a cell phone

Description automatically generated