A Report for Artificial Intelligence Tutorial (CS 401)

Under the Guidance of

Dr. Partha Pakray,
Associate Professor

by **Shibam Debnath (2012181)**

PROJECT I

OBJECTIVE :

To build a Search for NIT Silchar website using Apache Nutch and Tomcat.

ABSTRACT :

A search engine uses crawling to gather web data and indexing to organise it for quick retrieval. Apache Nutch is an open-source tool for effective web crawling. Tomcat is an app server handling Java-based web apps. Combining Nutch and Tomcat makes a search engine.

ARCHITECTURE :

1. **Apache Nutch :**
   - Web crawler for data collection.
   - Responsible for indexing web content.
2. **Tomcat :**
   - Application server for executing the search engine.
   - Manages user queries and facilitates data retrieval.
3. **Integration :**
   - Nutch and Tomcat collaborate seamlessly.
   - Nutch's indexed data is processed and executed through Tomcat.

STEPS INVOLVED :

1. **Install Java**: We need Java installed on your machine

2. **Install Cygwin**: We need Cygwin to run the shell commands

3. **Install Tomcat**: We need Tomcat running on your machine. So we downloaded the tomcat.

4. **Download Nutch**: Download the release and extract it on your hard disk in a directory that does not contain a space in it (e.g., c:\\nutch-0.9).

5. **Create a URL seed list**: Create an empty text file in your nutch directory and add the URLs of the sites you want to crawl. For example, if you want to crawl the NITS website, you would add `http://www.nits.ac.in/` to this file.

6. **Configure Regular Expression Filters**: Add your URLs to the `regex-urlfilter.txt` file[15]. An entry could look like this: `+^http://([a-z0-9]*\\.)*nits.ac.in/`.

7. **Crawl the website**: Start the crawl:

8. **Index the crawled data**: After crawling the website, we need to index the crawled data.

9. **Start a Cygwin bash shell** : Navigate to the Tomcat bin directory. For example, if your Tomcat is installed in c:/Tomcat8:

10. **Start the Tomcat server** : Type in startup.sh and then hit Enter

11. **Open this url and start searching** : In your favourite web browser, goto

You can now p `http://localhost:8081/nutch-0.9/` ıe Nutch Search Web Application

OUTPUT SAMPLES :

Figure 1.1 : Actual Search Engine

Figure 1.2 : Search sample

PROJECT II

OBJECTIVE :

To build a movie recommendation chatbot in Assamese language using the RASA.

ABSTRACT :

AI chatbots handle tasks like ordering, reservations, and more. RASA is an open-source solution for conversational AI. It communicates, learns, and aims for top-notch customer interaction. Using RASA to create an AI Chatbot for Assamese.

ARCHITECTURE :

RASA has two main components :

1. RASA Core : Keeps a tracker for each session, i.e. for each user.
2. RASA NLU : Assists the chatbots in understanding what users are saying, extracts entities from the bot in the form of structured data, and classifies intents.

STEPS INVOLVED :

1. **Installation** : Install RASA open source using the set of commands shown in the output.
2. **Setting up the environment** : To get the project started, run rasa init. Three crucial files are produced by this: stories.yml, domain.yml, and nlu.yml.
3. **Intents** : Update nlu.yml with the required intentions.
4. **Slots**: In domain.yml, define slots for information storage.
5. **Responses** : To direct the bot's responses, add responses to domain.yml for each purpose.
6. **Stories** : To specify how the bot reacts to user inputs, set up the conversation flow in stories.yml.
7. **Validation** : Use the following methods to confirm the model's contents before training: **Validate $ rasa data**
8. **Training** : Use the following command to train the data: The fixed-model-name contact_bot is **$ rasa train**.
9. **Testing** : Use the **$ rasa shell** to have a conversation and test the chatbot.

OUTPUT SAMPLE :

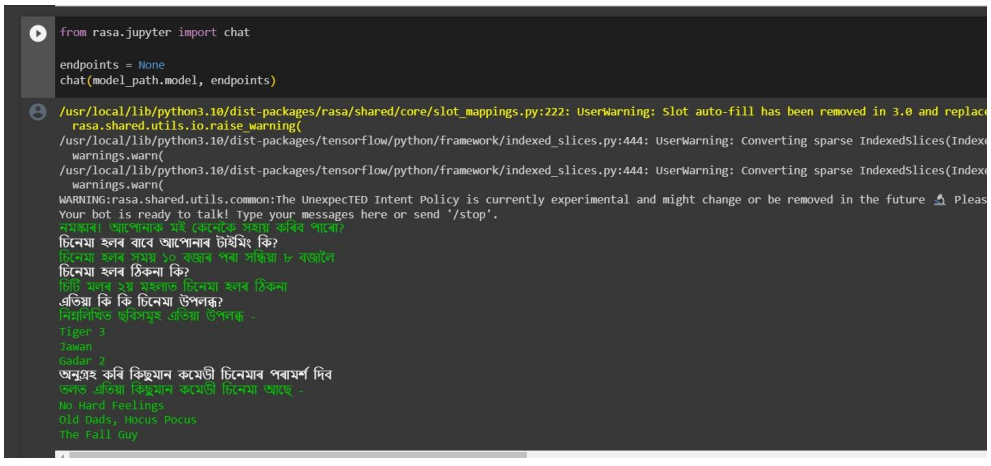

Figure 2.1 : Chatbot sample