ChatBot : A Modern Way of Interaction

S ARUN PRASAD
CSE DEPT
IIIT BANGALORE
Bangalore, India
s.arunprasad@iiitb.org

***Abstract*—This paper discusses the design and working and implementation of a chatbot system using opensource tools.The tool used builds a local NLU (Natural Language Understanding) to create Ask-Tell Model , which is used as the brain of the chatbot.**

***Index Terms*—Natural Language Understanding , RASA-NLU.**

I. Introduction

A chatbot is computer program that used to stimulate a conversation or interact with humans via chats or voice messages.A chatbot can be a program , a software or a mobile app.Its primarily deployed by companies on their websites or apps to interact and answer queries to the users. Its an automatic answering machine.

A. PROS

•*Fast and Accurate*
• *Can provide effective assistance for business needs*
• *Cuts the cost of employing a person to answer queries*

B. CONS

•*Makes mistakes sometimes*
• *Technical requirements to deploy and maintain them.*
• *Needs Internet to work for online models.*

II. Chatbot Design

Chatbots can be implemented in a number of ways , from the simplest of them being ,just hardcoding the question and answers , to a highly complex system using AI ,ML and NLP which requires lot of training and testing. A good working bot is a program that constantly learns all time or periodically even after being deployed.This is possible with recent advancement in ML and AI field. The latest algorithms help or enable bots to be intelligent enough to learn all time.This requires logging user chat data history and storing it to train and provide update to the bot later.The days a lot of frameworks have been developed to help us build chatbots , these come with pre-trained algorithms and models enabling us to build application without having us to worry much about the algorithms or internal implementation.They come as Bot Development Frameworks and Bot Platforms.

A. Bot Development Frameworks

A set of predefined functions ,classes and libraries.Also provides access for inbuilt pre-trained ML models and algorithms. Used to build bots from scratch.
Example : Microsoft Bot framework, Wit.ai , API.ai etc.

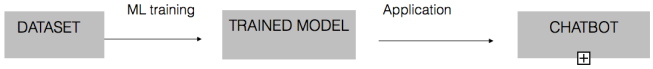
B. Bot Platforms

These are for non-technical people and beginners who want to build and deploy a bot with coding knowledge. The are online platforms mostly , provides a way to train and test bots and deploy it where it can interact with users.
Example : ChatFuel ,Motion.ai etc

III. Dataset

DataSet is the input or data used to train your chatbot initially. Just like teaching someone how to speak or answer.

To state otherwise a Dataset is just a meaningful conversation related to the task we are trying to solve , you can either create them manually or download and use a relevant dataset for the problem. To create a general chatbot one of the popularly used data set (open) to train models is reddit website comments and sub-comments.Pairing comment and sub-comment are a classical dataset for training a bot , this gives us tons of meaningful data for training them. This also has lot of irrelevant but that is only a minority.



Overview
Link To DataSet

IV. The Algorithm and Model

A. NLP and NLU Concept

1) NLP: Natural-Language Processing is the area of computer science , that deals with the math and algorithm of understanding large amount of natural data. This is the base for computers to understand large amount of natural data.This area has lot of application in speech recognition,natural-language understanding , Machine Learning.

2) NLU: Natural-Language is the subset of the above NLP , in artificial intelligence.This deals with machine perception and comprehenshion.This has lot of applications on AI , ML problems.This is classified as an AI-Hard.Further this is used in intent classification and entity extraction. This is used in building the brain of the chatbot system.

B. Intent Classification and Entity extraction

1) Entity Extraction: Entity extraction can be defined as the task of extracting useful information from the natural language and categorizing them into one of the pre-defined catogories. From the chatbot system point of view we pre-define or hard code a lot of entities or catorgories , this is done under the assumption that for a local chatbot system whose inputs are contrained , it is easy to apply entity extraction and provide answers based on it.

2) Intent Classification : One of the priority task for chatbots is to do intent classification on the input. This is a very complex task.This is basically catogorizing the input into several catogories , this is done to reduce the universal set of replies to quite a few section , so it is easy find the appropriate and best reply for the input.

Some methods used for intent classification.

- $PatternMatching$
- $Machinelearning$
- $DeepLearning - NeuralNetworks$

V. RASA-NLU

Having read briefly about some the background neceesities ,math and algorithm , we now go into the tools.The open source tool i am describing is RASA-NLU , this is a tool which build NLU for the given dataset , intents and entities have to be pre-defined and this builds the trained model for the given input data set.So for a relatively Ask-Tell model for buisness needs , the intents and entites are relatively few and hence they need to be hardcoded in JSON Format.

Here are a few links

Docs

JSON Generator Tool

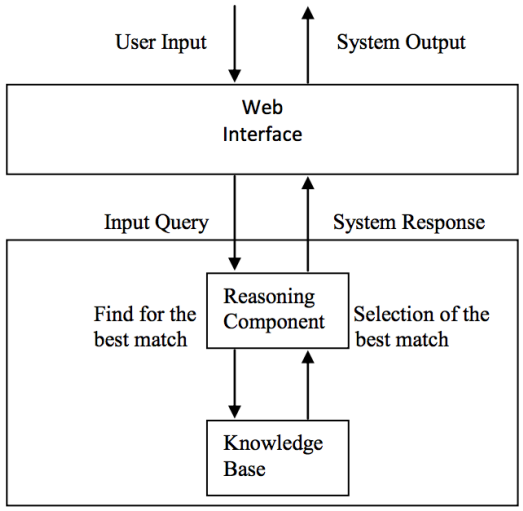
VI. Building a practical bot

In this section we will go about a practical way to implement a chatbot system. We will use a lot of open source tool to build it .We go for a popular Ask-Tell model.

1) Training a model: We are going to use RASA-NLU model , but we need to train the model according the need of the chatbot. For this we have to get the frequent question and appropriate answers as a training data.The more data we produce the more accurate it is going to be.In this dataset we need to also specify the intent and entity in each of the input sentence.Also we need to hardcode answers for each intent-entity.For this we can use the tool(JSON Generator Tool) to generate the JSON with input , intent and entity to feed it to the model.

2) Create Feedback System: Now we need to create a feed back system , assume we used a model to extract entities and intents , what do we do after that? We need to provide reply based on the intents and entities ,to this we create a system , which takes entity and intent as argument and provide a reply predefined by us. This can be simply done in python , using a dictionary , we need to initialize the dictionary once and we can retrieve the replies for each intent and entities in O(1).A code has been provided in the github link in references.

3) Create the chatbot system: Now we have data set , trained model and feedback system. Now the chatbot is nothing but the linking of the above into a complete working code. We use a python code or program , the reads the trained model and intialized the feedback system. We need to get input and pass it to the above and return the feedback as output.This completes out chatbot program. A python program for this along with complete implementation is given in the code in references section.This can be used as backend program for online chatbots.



Structure

Acknowledgment

A complete working chatbot for admission related queries for IIIT Banglore has been implemented using Python , Flask as Frontend , using RASA-NLU and python program as backend. It uses a listner to get queries and write back outputs.

References

[1]GitHub Repo for Project as reference

[2] An Intelligent Behaviour Shown by Chatbot System ,Vibhor Sharma, Monika Goyal , Drishti Malik

[3]Chatbots Greetings to Human-Computer Communication