

Chatbot for HR Department Using AIML and LSA

Ajaykumar Kushwaha¹, Rishikesh Nanaware², Damini Pandare³

¹Student, Dept. of Information Technology, Atharva College of Engineering, Malad, Mumbai, Maharashtra

²Student, Dept. of Information Technology, Atharva College of Engineering, Malad, Mumbai, Maharashtra

³Student, Dept. of Information Technology, Atharva College of Engineering, Malad, Mumbai, Maharashtra

Abstract - Chatbots are programs that impersonate human discussion utilizing Artificial Intelligence (AI). Today's Chatbots are becoming smarter, more responsive and more useful. Chatbot technology at their core uses Natural Language Processing (NLP). Softwares used by the HR department for Payroll management, Human Resources Management (HRM) are very tedious to use as they require visiting various pages, sections and forms even for benign details. A chatbot can simplify this task by letting user communicate in Natural language and get precise information with just one query. In, this paper we have proposed a chatbot which can give a response in different interface elements like tables, graphs, etc. It can understand the meaning of different queries searching for same response and gives response according to that.

Key Words: Chatbot, NLP, AIML, LSA, Human resource Management, Stemming.

1.INTRODUCTION

A Chatbot is a new tool, introduced in the market that simplifies user interaction with computers. In banking, sales, services it has made tremendous growth in the past. As per IBM - 265 billion client demands are recorded every year and organizations spent almost \$1.3 trillion to address them. Utilizing chatbots could assist spare with increasing to 30% of this expense^[1]. Microsoft CEO Satya Nadella believes that the evolution of user behaviour with computing will naturally involve the use of chatbots as the underlying basis for communication between user and information^[2]. According to Suresh Iyer, CIO of Blue Star, Sales executives spend a lot of time assembling and processing information by themselves which can be done by chatbot and chatbot may also be used as a guide^[3]. Various management tools available in the market can help manage human resource but mastering these tools and drawing insights is time-consuming and tedious.

In this paper, we are proposing an intelligent chatbot that is not only easy to use but also helps the user derive faster business insights. The bot can also respond to common greetings. The chatbot is built using Artificial Intelligence Markup Language (AIML) and Latent Semantic Analysis (LSA) and the required data is fetched from the database of the company.

In the proposed chatbot, there are two types of responses, template-based and analysis-based. Template-based responses are usually provided for greetings and predefined queries whereas analysis-based responses are provided for complex inputs. Template-based queries are handled by AIML solely for other responses LSA, stemming, and other NLP tools are additionally used.

Artificial Intelligence Markup Language(AIML)^[4]

AIML has four major tags, namely <aiml>, <category>, <pattern> and <template>.The AIML tag defines the start and the end of AIML document. The category tag defines a unit of knowledge and contains pattern and template tag. The pattern tag defines user input and may have wild card characters like star(*) and underscore(_).The template tag represents the response of the bot to the user input. AIML lacks the capability of performing mathematical operations and needs updating regularly. ^[5]

Code:

```
<category>
  <pattern>
    Do you know *
  </pattern>
  <template>
    Sorry ,I have no Idea about <star/>
  </template>
</category>
```

Output:

```
// User : Do you know Ajay
//Bot : Sorry ,I have no Idea about Ajay
```

Latent Semantic Analysis^[6]

Latent Semantic Analysis(LSA) practices 'bag of words' model, it learns latent topics by performing a matrix decomposition on the document-term matrix using Singular value decomposition (SVD). This can help in classifying texts and concluding the topic of the text. The initial step in LSA is the term frequency-inverse document frequency (TF-IDF). It reflects how important a word is to a document in a collection or corpus. The Next and Final

step is to use SVD to perform dimensionality reduction on the TF-IDF vectors.

Stemming

Stemming is a process of reducing a word to its root form i.e. lemma, eliminating the suffix and prefix. It is an important step in natural language processing. Porter's Stemmer is one of the most popular stemming methods proposed in 1980. [7]

Eg:

- EED -> EE
(Agreed -> Agree)
- TION -> T
(Introduction -> Introduce)

2. LITERATURE REVIEW

1. Chatbot for University Related FAQs [8]

Chatbots are programs that mimic human conversation using Artificial Intelligence (AI). Chatbots are being used for e-commerce, health-related, tourism-related, hotels, payment websites, and applications. This paper proposed a chatbot that is designed for students who have queries at the university. Chatbot uses AIML (Artificial Intelligence Markup Language) and LSA (Latent Semantic Analysis) technologies. Template-based queries such as welcome greetings will be responded using AIML, as AIML stores responses to these queries and other questions will be answered anytime using LSA.

User discussion starts with welcome greetings and general questions. When a user submits a query it is tested whether it has response stored in AIML format.

- User post the query on chatbot
- The user's query is matched with the predefined format stored in the AIML file by the developer.
- Pattern matching is performed between the user entered query and knowledge (pattern).
- If the pattern is matched then stored string is given as response, otherwise, it is handled by LSA.

2. Automated Thai-FAQ Chatbot using RNN-LSTM [9]

Users interact with businesses using emails and live chatbots. Though emails look formal in a business environment, users prefer to use chatbots as they are easy to use. Business needs to keep admins who can reply to customers' queries. This is a time-consuming process as admins need to type responses to each query manually and the customer has to wait for the response. Also, this is costly as admins need to be paid by businesses. Hence chatbots seem very efficient replacement to emails and give

interactive user experience. In this paper, the proposed chatbot uses a Recurrent Neural Network (RNN) in the form of Long Short-Term Memory (LSTM) for text classification. The proposed chatbot automatically responds to customers Frequently Asked Questions (FAQs).

Preparing data

Developers had used 2,636 pairs of questions and answers. These were then manually classified into 80 classes according to the number of FAQ types and numbered them.

Classification Model

This model is a neural network that learns to categories the questions. It includes three layers.

- Embedding layer is the first layer where words are mapped to vectors of real numbers that learn representation for predefined fixed-sized vocabulary from a corpus of text.
- The second layer is the long short-term memory (LSTM) layer. It is a particular kind of recurrent neural network (RNN). It is capable of learning sequential data such as text and video. With the help of LSTM, RNN can remember inputs for a long period.
- The third layer is the Dense layer (Output layer) with a softmax activation function.

The softmax activation operates in the output layer and represents a categorical distribution over class labels and obtains the probabilities of each inputs belonging to a label. Because of the softmax activation function, we have to encode the label of questions to one-hot format for the learning process of the model.

3. A Pilot Study Integrating an AI-driven Chatbot in an Introductory Programming Course [10]

In this research, the developer team developed an intelligent chatbot interface for a computer programming course. The chatbot initially had a very limited dataset. It would be advanced by user interaction with the chatbot. This chatbot is capable of evolving with the needs of the student. This paper discusses, how the chatbot was developed and integrated into the course, how the dataset was developed, the usage during the pilot study and future scope for the chatbot. The future scope includes improving the user interface, improving error rate and much more. This paper has also detailed the mechanism to handle issues like false-positive responses. The database for chatbot was populated with MATLAB functions. 21 unique users interacted with chatbot out of which six users interacted with the chatbot for more than one session and others interacted only once.

3. PROPOSED MODEL

In this work, we have developed an interactive Chatbot for the HR department to find answers to their queries. User will interact in natural language with the bot. Our system employs AIML and LSA technologies.

The collection of FAQs are preprocessed using tokenization, stemming and stopword removal.

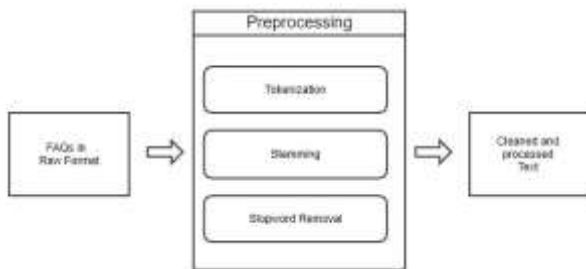


Fig-1: preprocessing FAQs

The user will enter a query using natural language and thereafter AIML will match the query with AIML script to provide a proper response. The system will switch to LSA, if AIML Script is unable to answer the query, as giving all possible user input is not feasible. LSA has an assumption that words with related meaning will occur in a similar piece of text.

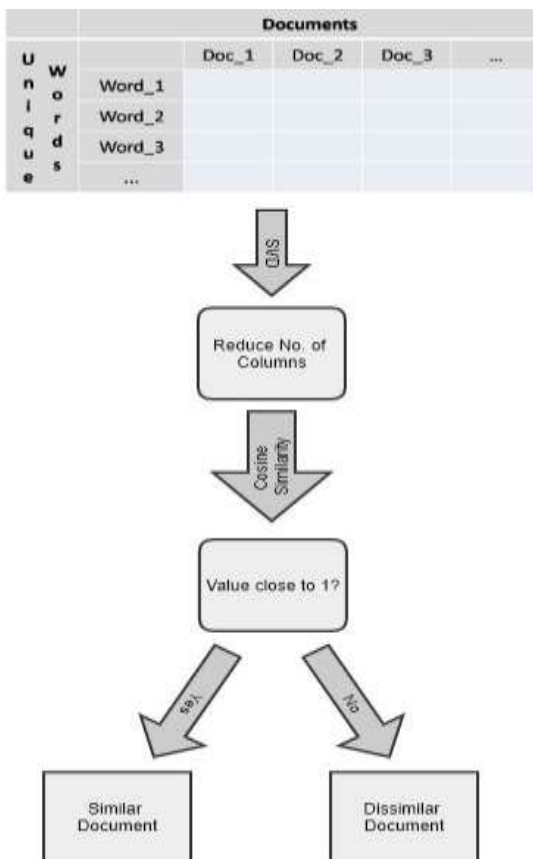


Fig-2: LSA computation

A matrix will be generated containing "Words counts per document". This matrix will comprise of the row representing a unique word and columns representing each document. We will now apply a mathematical technique called singular value decomposition (SVD), which is used to reduce the number of rows while preserving similarity. Documents are compared by taking a cosine similarity i.e. dot product between the normalizations of two vectors. Values close to 1 indicate similar document and values close to 0 indicate dissimilar document.

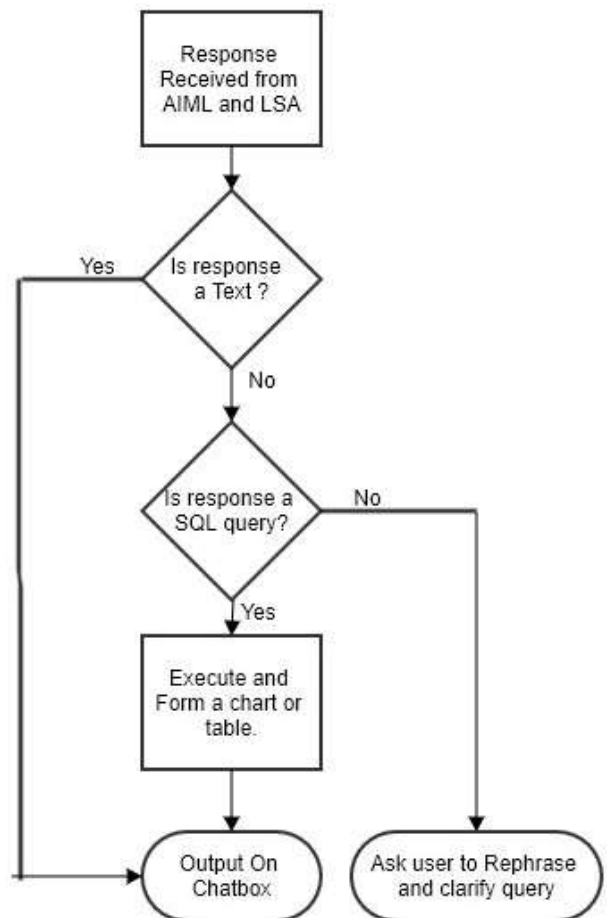


Fig-3: Formulating Response

In response to the user query, AIML and LSA will further produce either a plain-Text or SQL query as a response. If a plain text is produced it is output as it is, else if SQL query is returned, it is executed on employee database of the organization and output is displayed in an appropriate chart or tabular format, else for irrelevant queries we further ask for clarifications.

4. RESULT

The chatbot will be able to help the HR department get quick responses to their queries. It will help them manage Employee details, Employee salary, Employee attendance and many more. The results are displayed in text, tabular

and chart format. Snapshots of the proposed system are displayed in Fig, 2, Fig.3, Fig.4 respectively.



Fig-4: Employee Details



Fig-5: Employee Expertise

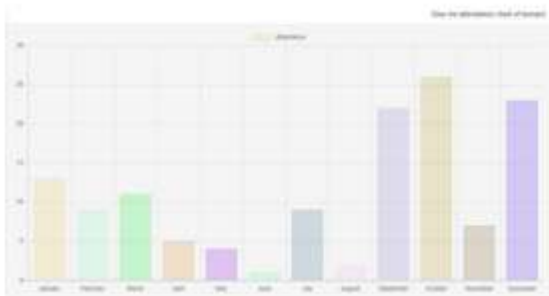


Fig-6: Employee Attendance Chart

5. CONCLUSION

This project will simplify the work of the HR department by obtaining accurate answers to their questions. The Chatbot will flag the user queries that had no response from the bot. Those queries can be further analyzed to enhance AIML patterns. Therefore, the chatbot will be able to continuously develop and learn from the data which is new to the system. As future work, we can provide the user with more interactive charts, buttons, emoticon, embedded maps and other GUI elements. We may also apply analysis such as Sentiment Analysis, Summarization and Topic modelling to make the bot more friendly and conversational, hence building an informal experience that mimics a human friend.

6. REFERENCES

[1] “How chatbots can help reduce customer service costs by 30%”, <https://www.ibm.com/blogs/watson/2017/10/how-chatbots-reduce-customer-service-costs-by-30-percent/>

[2] “Microsoft CEO Satya Nadella says chatbots will revolutionize computing”, <https://www.onmsft.com/news/microsoft-ceo-satya-nadella-says-chatbots-will-revolutionize-computing>

[3] “How chatbots are changing India Inc”, <https://cio.economictimes.indiatimes.com/news/next-gen-technologies/how-chatbots-are-changing-india-inc/74604260>

[4] “Artificial Intelligence Markup Language”, <http://www.aiml.foundation/>

[5] “Intelligent Chatting Service Using AIML”, Saqib G, Faizan, Ghatte N, IEEE International Conference on Current Trends toward Converging Technologies, 2018

[6] “Latent Semantic Analysis with python”, <https://www.datacamp.com/community/tutorials/covering-hidden-topics-python>

[7] “The Porter Stemming Algorithm”, <https://tartarus.org/martin/PorterStemmer>

[8] “Chatbot for University Related FAQs”, Bhavika R. Ranoliya, Nidhi Raghuvanshi and Sanjay Singh, International Conference on Advances In Computing Communications and Informatics 2017

[9] “Automated Thai-FAQ Chatbot using RNN-LSTM”, Panitan Muangkammuen, Narong Intiruk, Kanda Runapongsa Saikaew, 22nd International Computer Science and Engineering Conference 2018

[10] “A Pilot Study Integrating an AI-driven Chatbot in an Introductory Programming Course”, Matthew Verleger, James Pembridge, IEEE Frontiers in Education Conference (FIE), 2018