Measuring effectiveness of learning chatbot systems on Student’s learning outcome and memory retention

Suhni Abbasi¹, Hameedullah Kazi²

Lecturer, Information Technology Centre, Sindh Agriculture University, Hyderabad, Pakistan
Professor, Department of Electrical Engineering and Computer Sciences, Isra University, Hyderabad, Pakistan

ARTICLE INFO

ABSTRACT

Creating a learning environment in which students learn more effectively remains the great challenge from decades; different approaches are proposed, for example, Intelligent Tutoring Systems, Question Answering System and chatbot. All these approaches used natural language to achieve that goal. A comparison of these systems viz-a-viz student learning outcome and behavior is of eminent importance. To achieve this goal a chatbot system with knowledge base for Object-Oriented Programming Languages is developed and deployed. Case study was made to assess and evaluate the chatbot system for student learning methodology. Learning outcomes and Memory retention have been measured for the developed system. Comparisons were made between the results obtained using Google search engine and our chatbot system. The results indicate that learning through Chatbot have a significant impact on memory retention and Learning outcomes of the students.

Keywords: Chatbot, Question Answering System, memory retention, learning outcomes

Source of Support: Nil, Conflict of Interest: None declared.


This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license only allowing others to download our works and share them with others as long as they credit us, but they can’t change them in any way or use them commercially.

INTRODUCTION

The advent of human computer interface is to provide ease of communication between human and computer in user friendly way. User usually approaches Google, Yahoo and other information retrieval systems for finding the solutions of their problems but either they do not retrieve concise or relevant information, or
they retrieve documents or links to these documents instead of an appropriate answer of their problems. To address such problem the idea of natural language dialog system arises in which user’s questions in natural language and receives concise and appropriate answer Shawar and Atwell (2007).

Intelligent Tutoring Systems (ITSs) is computer based system that use Artificial Intelligence techniques to represent knowledge and offer support and guidance to learners in the process of learning struggling to master the learning and cognitive skills. In such systems when learners make any mistake they receive customized instructions or direct feedback on their problem solving task Corbett et al. (1997). Intelligent tutoring systems that use natural language have mostly been appropriate to be either tutor, pedagogical agents or chatbotsKerly et al. (2006).

A chatbot is a computer program designed to simulate communication with the user in natural language. The chatbot conversational approach can be integrated well with question answering system due to the fact that when user questions they should retrieve meaningful and concise answer Smith (2010). Benotti et al.(2014) reported that Chatbot perform as an efficient tool for increasing the Student’s retention and engagement in a classroom environment, as well as online competition in the field of Computer science mainly.

Kay (1997) describes that Presenting information in different ways may help students to think about their knowledge and also improve it. Further,He envisaged that all these goals can be achieved by chatbot suitably integrated into intelligent tutoring system. Constructing a dialogue-based human computer interface is quite challenging task in the same order of difficulty (quantitatively as well as qualitatively) as building an agent that can enhance their learning abilities.

Thus, determining the efficacy of the learning chatbots on the student’s behaviours like their cognitive load management, stress management, learning outcomes, memory retention etc. especially in the domain specificchatbots seems to be great challenge. This research pursues which and how behavior metrics are necessary to understand and consider for the student’s learning outcomes and memory retentions through the chatbots systems.

**Behavior Indicators**

Behavior is response generated by the system by either internal or external stimulus states of the agent. Further, it describes some activities to accomplish the job associated with behaviour Qiu and Hu (2008).

Kolb (1984) describes learning is a process whereby knowledge is created through the transformation of experience. Learning environment includes the systems and dynamics that facilitate and make possible student involvement Coates (2006). It is reasonable to assume that the learning environment will have an influence on how user behaves in the learning environment. Aside from the learning environments, the design, building and delivery of contents Coates et al. (2005) are also factors that can influence the user behavior in learning environment.

Prochaska and Clemente (1992) introduce the model which describe how people can develop behavior and attitudes in learning environment and how the learning takes places. The five phases of the model are described below

- Acknowledgement of the new behavior or attitude,
- Recognition of one’s similarity with the observation,
- Identification of the outcomes,
- Remembrance of the behavior or an attitude,
• Reproduction of the behavior or an attitude,
• Reinforcement of the model.

MEMORY RETENTION

Mayer (2001) stated that the retention is the power to recognize or recall past learning events. The interaction with the chatbot system is done in Natural language dialogue, therefore, its interface is supposed to need improvement in the procedure of communication and some other features, such as usability, user satisfaction, learnability, and memory retention. Lindsey et al. (2014) reported that human memory is not completely perfect, for long term retention the regular review is very much important. Students in the academic environment are always facing problems to retain the increasing amount of material, as well as master in crucially new knowledge and skills. In the Chatbot Systems involvement of the users is more active during learning; user’s reflection and self-explanation may well encourage, many parallel channels as well as frontal communication is allowed in chatbot systems. The ability to retain the answers for the longer time can be achieved by adapting these systems Jan et al. (2003).

LEARNING OUTCOMES

Dunlosky et al. (2013) reported that students can achieve their learning goals more effectively if the cognitive and educational psychologists develop and evaluate easy-to-use learning techniques. In distance education the use of chatbot or a computer software that is self-directed, interactive, anthropomorphized, and aimed at enhancing the educational goals and outcomes play an important role. The Purpose of designing such computer software is focused by numerous hypothetical frameworks strained from different disciplines. Fosnot (1996) highlights exchange and communication of knowledge in natural language is effective way learning elaboration. The approach of using the chatbot in the learning environment have been identified as a learning tool with the ability to enhance learner interest, memory retention, and knowledge transference Hadwin et al. (2005).

Johnson et al. (2000) argues that the chatbot have very great potential to enhance the communication abilities between students and computers and it also stimulate the students to engage with computer. Additionally it is supposed that traditionally improving learning outcomes and learning experiences tedious and were very challenging in distance learning. Reiners et al. (2014) suggested that developing the chatbot with the specialist resource with strong educational scenario have significant impact on the learning of students.

EXISTING CHATBOT SYSTEM

ELIZA

Eliza is the first and well known chatbot of the world; it was developed by Joseph Weizenbaum in 1966. Eliza simulated a therapist role in clinical treatment. ELIZA search for simple keyword in the user request and then tries to define minimal context where it found the key word. ELIZA also choose appropriate transformation rule to, adjust the user request. To summaries, Eliza works by turning the user sentences around Angeli et al. (2001).
ALICE

ALICE (Artificial Linguistic Internet Computer Entity) was first developed by Dr. Wallace in 1995 continuously improved by means of the time. Brain or knowledge base of ALICE is stored in AIML format, AIML (or Artificial Intelligence Mark-up Language) an XML specification for programming Chatbot software.

Currently, ALICE corpus size is more than 10K AIML categories and these are continuously increasing with respect to time, but all the categories are hard-coded, so adoption of new domains and languages is restricted Shawar and Atwell (2007).

FAQChat

FAQChat was first developed to chat about the UNIX OS but later on it was modified by the School of computing at the University of Leeds. The information retrieved from the FAQChat is similar to those which are found from Google search. The User interface was developed on the basis of such correspondence. The Interface requires input query and on the basis of that query it will produce two responses, one response is produced from the FAQchat, and the other one is retrieved from Google after filtering it to the FAQ of SoCShawar et al. (2005).

FAQchat will give the response by selecting most significant words as token, and attempts to match longest possible pattern. FAQChat is language independent, so its response are not generated by using any morphological tools or by analysing the meaning of words Shawar et al. (2005).

The working of FAQchat is summarized in the following steps:

- Filtering process is applied to the database of questions and answers to remove any unnecessary tags and then questions and answers are extracted from the whole database.
- The FAQ database gives questions and answers. After this, a list of links is constructed, containing the links from FAQ to web pages containing responses.
- All the words in questions are recorded in the dictionary with their frequencies of occurrence. From each question, the first and second most-significant words are extracted.
- In this step, the AIML categories are created. The input could either match a complete FAQ question or 1st or 2nd most-significant words in the question. If only one match is found, then response is a direct answer else in case of multiple matches, FAQchat returns links as a reply.

NATURAL LANGUAGE ASSISTANT

NLA serves users in finding notebooks according to their needs via natural language conversational dialogue. At each turn NLA provides incremental feedback for its understanding of user’s requirements about a product. This iterative process of user’s query refinement allows the system to find and recommend product that best matches with the user’s requirements and constraints Chai et al. (2001).

The system’s architecture is composed of three major components. The Presentation Manager is used as an interpreter for user’s questions and also generates system responses. The user’s current requirements are input to the Dialogue Manager and on the basis of these requirements; the DM formulates action plans for Action Manager. The AM performs backend operations such as, database access. The DM then creates responses based on the results returned by AM and the discourse history and sends it to PM which responsible for displaying it to the user.
NLA is unable to handle meta-level queries such as “What is DVD?” or “How could I add memory to this model.”

**Methodology**

**Participants**

Seventy-two 2nd year students of Information Technology Centre, Sindh Agriculture University, Tandojam participated in this study. The participants of this study have been exposed to the questions from the Object Oriented Programming languages domain. As this study measures the effect on memory retention and learning outcomes while using the chatbot system. The participants were randomly divided into two equal groups, one who used Google search engine and the other one used the Chatbot system to find the solution of their problems.

**Experiment Design**

For designing the knowledge base for the selected domains i.e. Object Oriented Programming languages; MySql Database Server is used. For producing the responses of the student’s natural language queries, the GUI is connected to the knowledge Base via MySqlConnector.net.

For developing the knowledge base total, 5000 question samples were collected from all the students of 2nd year’s students of the Information Technology Centre then those questions were divided into categories depending on the patterns of question. The identified question patterns in the survey were “What,” “Why,” “Perform Operation,” “How,” “Advantage or Disadvantage,” “Application” and “Who”. The classification of questions and percentage of each category collected in the survey is shown in Fig. 1.

![Fig. 1 Percent of Questions collected from students](image-url)

The queries which are asked by the users are in natural language and in order to match these queries against a knowledge base the Artificial Intelligence Markup Language (AIML) is used. User Queries are matched against patterns and responses are generated in the form of Key Word from the Templates of the AIML file. AIML is very suitable for this
purpose as it was initially designed for conversational chatbot. The various forms of these types of questions are also implemented in the AIML files. The partial implementation of “What *” Query is shown in Fig. 2.

**SYSTEM ARCHITECTURE**

The system architecture is designed on the principles and techniques of chatbots and natural language processing. The question-answer responses in existing chatbots are hard-coded. But it is a cumbersome task to feed enough categories manually. An alternate to overcome this problem is to create an automated chatbot which extracts user responses from some knowledge source such as, database. Such chatbots are known as information retrieval chatbots. An information retrieval chatbot is the best alternative for traditional information retrieval systems such as Google as they provide direct answers to users' natural language questions, unlike traditional information retrieval systems. The proposed solution that is an information retrieval chatbot reduces the time to browse and search through the documents to find the required answer. The architecture of the OOPL comprises of three main components: the User Interface, the AIML parser and AIML database. The system works as shown in Fig. 3.
DATA ANALYSIS AND RESULTS

To measure the students learning and memory retention through the chatbot it is essential to record individual performance. Appropriate computer-based pre-tests are required to understand factors that are useful to measure the learning abilities. To achieve this several surveys were conducted to measure memory retentions and students learning abilities. Students were divided into two groups, one who used Google search engine and the other ones used the Chatbot to find the solution of their problems. A total of 36 samples for each test were collected and evaluated. We adopted two Surveys, one for to measure memory retention in terms of students remembers the responses which they receive from Google or Chatbot and other is to measure the Learning Outcomes based on the quality of responses. The results obtained from both measures were evaluated from the expert faculty members of Information technology centre, Sindh Agriculture University, Tandojam. Below steps will briefly explained about the data analysis procedure for both measures will be explained briefly:

For measuring memory retention, students were randomly divided into two groups (n=36); one who used Chatbot to find the solution of their problem and other group used Google Search Engine. The task was assigned to each student of the group, and three-time frames were set. One-time frame was set to 5 minutes, 2 hours and 1 day, and according to pre-defined time responses were checked that either student remember the responses they receive from both systems or not. The average counts of the responses as remember or not remember the exact answers from the OOPLChatbot and Google is shown in Fig 4.
The results show that there is a significant difference in the responses received from Chatbot and Google search Engine. Results indicate high count for remembering the responses of ChatBot for all time’s frames i.e. After 5 minutes, 2 hours and 1 day.

![Results of Memory Retention](image)

**Fig. 4. Average count results of Memory Retention**

For learning outcomes, pre and post were conducted of their pedagogical agent. We measured learning outcomes by randomly dividing students into two groups (n=36); one who used Chatbot to find the solution of their problem and other group used Google Search Engine. Two-sample tests were conducted; one test was conducted before using any system to perceive the knowledge about the subject’s students wants to learn, and another test was conducted after using Chatbot or Google Search Engine. Five different tasks were randomly assigned in Pre-Test and Post-test to each student of the group. The time to find the solution of their problem was ten minutes for each test. The mean and standard deviation of all tests are shown in table and figure below (Table 1 and 2, Fig.5)

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRE-TEST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google</td>
<td>36</td>
<td>24.05</td>
<td>9.7657</td>
<td>1.62761</td>
</tr>
<tr>
<td>Chatbot</td>
<td>36</td>
<td>28.88</td>
<td>6.9930</td>
<td>1.16549</td>
</tr>
</tbody>
</table>

*Table 1. Pre-Test for the learning outcomes responses from chatbot and google search Engine*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>POST-TEST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google</td>
<td>36</td>
<td>37.55</td>
<td>8.3408</td>
<td>1.39013</td>
</tr>
<tr>
<td>Chatbot</td>
<td>36</td>
<td>54.77</td>
<td>8.3225</td>
<td>1.38708</td>
</tr>
</tbody>
</table>

*Table 2. Post-Test for the learning outcomes responses from chatbot and google search Engine*

<table>
<thead>
<tr>
<th>Learning Outcomes</th>
<th>Pre-Test</th>
<th>Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>24.05(9.76)</td>
<td>37.55(8.34)</td>
</tr>
<tr>
<td>ChatBot</td>
<td>28.88(6.99)</td>
<td>54.77(8.32)</td>
</tr>
</tbody>
</table>

*Fig. 5. Mean Score and standard deviation of Learning Outcomes*
The results show that Learning Outcomes measurements the post-test scores from Google and the Chatbot were significantly higher than the pre-test scores from these sessions (p < 0.001, Mann-Whitney). Thus, learning was achieved by users in both groups. However, learning outcomes in the Chatbot group was significantly higher than that of the Google group (p < 0.001, Mann-Whitney).

**CONCLUSION**

In this paper for measuring the various behaviours which effect on learning of students through chatbot system, a chatbot prototype was designed and assessed by the student. From the literature review, it is found that chatbot system is the efficient tool to not only measure but enhance the learning of the students. The designed prototype was evaluated for memory retention and learning outcomes. From the overall evaluations, it is concluded that students learning in terms of memory Retention and learning outcomes by using the chatbot system is significantly high in terms of quality as well as quantity as compares to the learning through the conventional search engines. Further, it is concluded that problems with students misspelling and false leads, quality of questions from the student can improve the overall learning through chatbot system.

**REFERENCES**


AJASE!!!
“Speedy publication service, Online archives, Paperless, web-based peer review system, Open access policy, Indexing in world known citation databases, Global circulation, Broad international readership and authorship, Online submission system, Minimum publication charge”