Human Interactive Intelligent System for Managing Rural Libraries

Yesha Krishna. V[1], Dr. Priyadharshini. M[2*]

Computer Science Engineering, KPRIET, Coimbatore, India[1,2]

*Corresponding author

Abstract: Rural library information is maintained manually using hard bounded records which were replaced by software that stored records in form of database. It is also enabled with a wide variety of resources such as E-books, PDF and much more using the Internet. The objective of the proposed system is to provide an interactive system that could be incorporated with voice interaction. The voice interaction is carried out with a help of a tool that could build natural voice experiences and offer users a more intuitive way to interact to get their work done. The Voice User Interface models the real human assistant strategy by incorporating artificial intelligence techniques. In specific, Artificial Intelligence makes the library management system more realistic which can adapt to the scalable requirements of a learner. The proposed system is an interactive library system that uses Amazon's Echo device, which can help users with the regular operation of new book entry, new user entry, issue and return of books and to query for books on specific inputs such as title, author, publisher etc. through voice interaction without any manual assistance. The system's specific application is the prediction of books matching the combinations of user requirements. The natural language processing is used to provide support to the user and the artificial intelligence uses an algorithm that can help to suggest or to predict the resource available based on the user's choice.

Keywords: Artificial Intelligence, Alexa Skills set, Human-Computer Interaction, Lambda function, Natural Language Processing

I. INTRODUCTION

A Library is the most significant part in a student's life since helps in enhancing their knowledge, skills as well as their personality. Interactive libraries enable students to access a wide variety of resources from the library and can also guide them with their requirements. This aims to bring Artificial Intelligence into play along with Natural Language Processing to make the system intelligent and interactive. A system developed specifically for Worthington Libraries [1] using Amazon's Alexa Skills and Alexa Voice Service is implemented to make book entries and provides list of events currently going around. Likewise, the idea is to develop an Alexa Skill and enable it using Alexa Voice Service in rural libraries. This can help students to interact and utilize the library resources much effectively. The Alexa Skill should be developed to carry out all the functions along with a specific lambda function that has to be integrated into the Alexa Voice Service. The Alexa skill set is to be developed so as to enable prediction using prominent Long Short-Term Memory (LSTM) algorithm [2] has using Python or Node.js.

II. PROPOSED SYSTEM

(A) System Architecture

Fig. 1 System Architecture

The user enables the Echo device by calling out the wake word which is followed by the request known as the utterance. This utterance is converted into computer language commands and vice-versa by the Alexa Voice Service (AVS) that forwards the commands to the Alexa Skill Kit (ASK). The ASK determines the suitable skill using the skill invocation name in the utterance and the lambda function integrates the ASK with the AWS to process the query and sends back the reply which is converted back to speech and becomes audible to the user.

(B)   SYSTEM COMPONENTS

Amazon Echo Device
Amazon Echo [3] is a device that when connected to the internet becomes a voice-controlled intelligent personal assistant service. It has 360° room-filling sound and even can hear you from across the room in any direction. It can be awakened only by calling out the wake word such as "Alexa", "Echo" or "Computer" followed by the utterance which will be processed later. In the default mode, the device keeps listening to all speech, monitoring for the wake word to be spoken. The device supports voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, playing audio books, and providing weather, traffic and other real-time information. It can also control several smart devices in the field of Internet of Things (IoT).

Alexa Voice Service
Alexa Voice Service (AVS) [4] is Amazon's suite of services built around its voice-controlled Artificial Intelligence assistant for the home and other environments. It is comprised of interfaces that correspond to client-functionality, like speech recognition, audio playback, and volume control. Each interface contains various *skills*. Users can access these new abilities by asking questions or making requests to the AVS. Developers can build skills that provide users with many different types of abilities. It is comprised of tools, application program interfaces (APIs), code samples and documentation that enables a developer to add skills to the voice recognition capabilities available on Alexa. A developer must certify, publish and update skills, which are made available through the Alexa Skills Store in order to enable its usage.

Alexa Skills Kit
The Alexa Skills Kit [5] is a software development kit (SDK) that enables a developer to build skills, also called conversational applications, on the Amazon Alexa artificial intelligence assistant. It is comprised of tools, application program interfaces (APIs), code samples and documentation that enables a developer to add skills to the voice recognition capabilities available on Alexa. A developer must certify, publish and update skills, which are made available through the Alexa Skills Store in order to enable its usage.

Lambda Function
Amazon Web Service (AWS) Lambda Function [6] is a computation service that lets you run code without provisioning or managing servers. AWS Lambda function executes your code only when needed and scales automatically, from a few requests per day to thousands per second. It can be used to extend other AWS services with custom logic or to create your own back-end services which can operate at AWS scale, performance, and security. Every Lambda function includes the code as well as some associated configuration information that includes the function name and resource requirements. AWS Lambda supports codes written in one of the following languages Node.js, Java, C#, Go and Python. It enables other Alexa capabilities such as Amazon S3 bucket, an Amazon DynamoDB table, Amazon API Gateway and also to build server-less applications.

Long Short-Term Memory
Long Short-Term Memory (LSTM) [7] is a Recurrent Neural Networks (RNN) and a very powerful algorithm which is used in search engines for retrieval of queries. This algorithm can be used in the system in order to retrieve the requested books from a database. An LSTM unit is consists of a cell, an input gate, an output gate and a forget gate. Every memory blocks are considered to be cells, the input gate is responsible for the addition of information to the current cell state, an output gate displays the output after computing the inputs in the cell and the forget gate is responsible for forgetting the information inside the cell state by denoting a binary values despite the fact that 1 means to keep the weight in memory and 0 means to forget it. These gates can be determined as a series of matrix operations and they also possess some weights. These networks are well-suited for classifying, processing and making predictions based on time series data and when LSTM units partially occur in RNN, they are able to solve the vanishing gradient problem, because LSTM units allow gradients to flow unchanged.
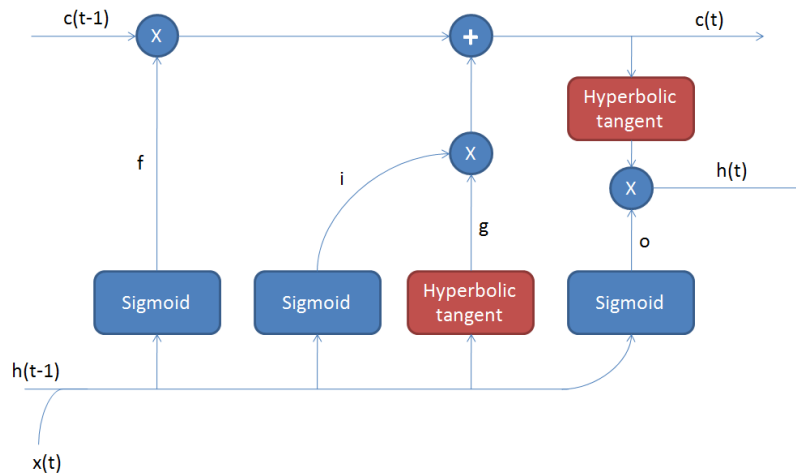
Fig. 2 Architecture of LSTM Cell

x(t) - input vector

i   - input gate's activation vector

o   - output gate's activation vector

f   - forget gate's activation vector

g   - new candidate activation vector

c(t) - cell state vector

h(t) - hidden state vector

LSTM uses the following functions:

The hyperbolic tangent function is used for dealing with the vanishing gradient problem since its second derivative sustain for a longer range before going to zero and determines the candidate values that can be added to the internal state.

Hyperbolic Tangent Function (Eq.1)

$$\tanh = \frac{2}{1 + e^{-2x}} - 1 \quad (Eq.1)$$

The sigmoid function is used in the 3 gates and since its value can be either 0 or 1, it can either allow information to flow completely or does not allow any information at all. The 'x' in both the activation function takes in a range of values.

Sigmoid Function (Eq.2)

$$f = \frac{1}{1 + e^{-x}} \quad (Eq.2)$$

One point to emphasize is tanh activation function is stronger than sigmoid function but they both have a vanishing gradient problem.

(C)   SYSTEM IMPLEMENTATION

The system is tested with a single alexa echo plus device in the college library and few other rural libraries and found to be working with few limitations such as difficulty in recognizing phonetics and multi-lingual.
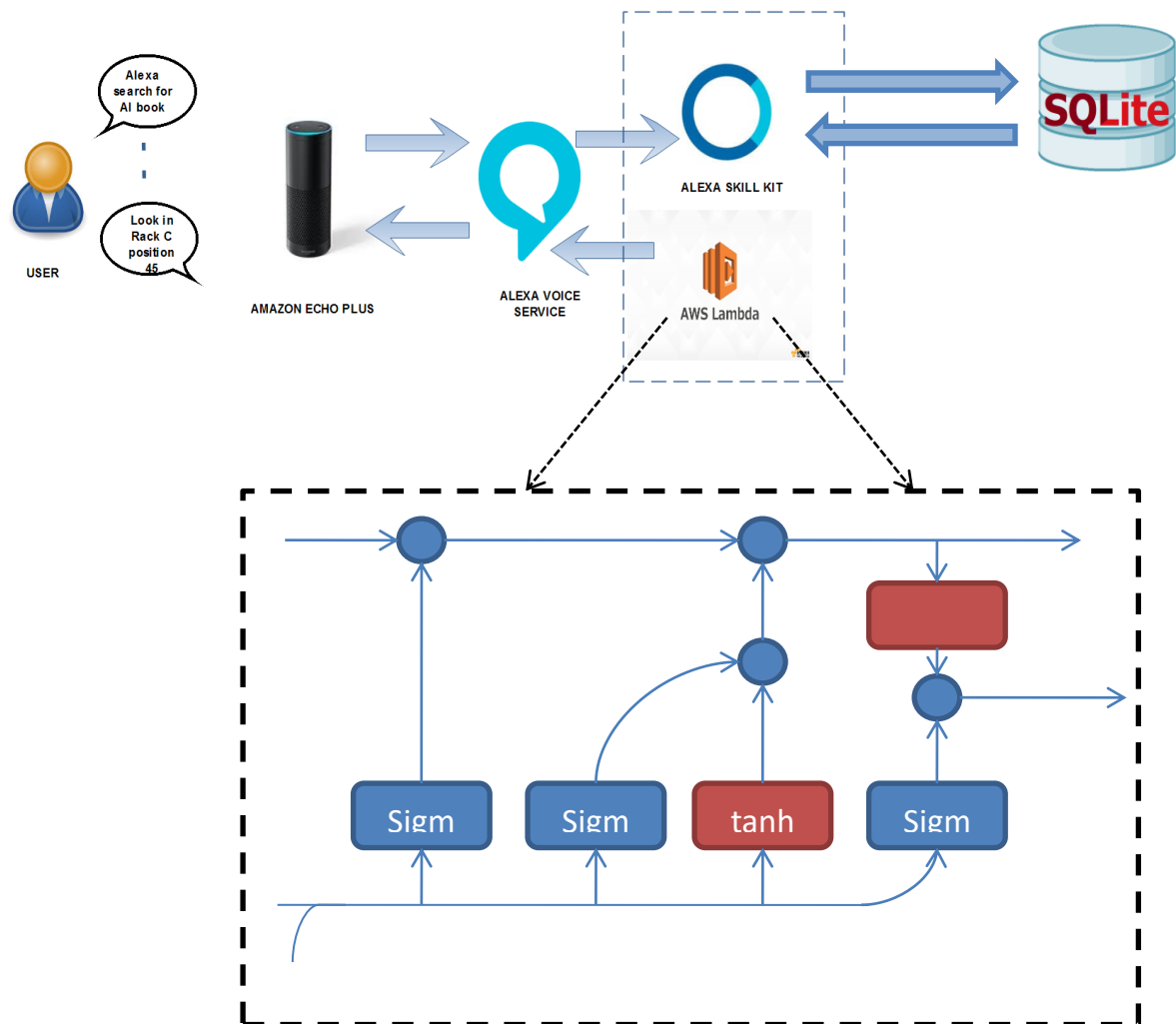
Fig. 3 System Implementation

CONCLUSION

The system implemented helps in automating the whole library management process. In case of huge volume of books and resources this could help the libraries in their process effectively. This interactive system can save human resources to be present at the counter and the artificial intelligent systems could help the users to find resources of their need and choice. The system though poses many advantages, holds few limitations such that user must be able to communicate in English in order to query. These limitations could be considered worthwhile for future extensions so as to be able to take in every random utterance and must not stick the predefined utterance and a particular language. This system could also be extended so as to interact with the e-resources and respond to the users with offline and online contents.

REFERENCES

[1]. https://www.worthingtonlibraries.org/visit/services/voice-assistants
[2]. Filippova, Katja & Alfonseca, Enrique & A. Colmenares, Carlos & Kaiser, Lukasz & Vinyals, Oriol. (2015). Sentence Compression by Deletion with LSTMs. 360-368. 10.18653/v1/D15-1042.
[3]. https://thewirecutter.com/reviews/what-is-alexa-what-is-the-amazon-echo-and-should-you-get-one/
[4]. https://whatis.techtarget.com/definition/Alexa-Voice-Services-AVS
[5]. https://searchaws.techtarget.com/definition/Alexa-Skills-Kit
[6]. https://docs.aws.amazon.com/lambda/latest/dg/welcome.html
[7]. http://colah.github.io/posts/2015-08-Understanding-LSTMs/
[8]. A. Sartin, Maicon & Rodrigues da Silva, Alexandre. (2013). Approximation of hyperbolic tangent activation function using hybrid methods. 2013 8th International Workshop on Reconfigurable and Communication-Centric Systems-on-Chip, ReCoSoC 2013. 1-6. 10.1109/ ReCoSoC.2013.6581545.
[9]. Jamel, Thamer & Mohammed, Ban. (2012). Implementation of a Sigmoid Activation Function For Neural Network Using FPGA.