EasyChair Preprint

Talk-to-Code

Jay Rabadiya, Meet Bhalodiya, Aditya Das and Kamesh Patel

April 5, 2020

Talk-to-Code

Jay Rabadiya[1], Meet Bhalodiya[2], Aditya Das[3] and Kamesh Patel[4]
Dept of Information & Communication Technology, Parul University, Vadodara

Abstract. Speech is probably is the most efficient and natural way to communicate with each other. This system provide coding through speech. Using these users can code easily. Talk-to-Code command-based applications will make life easier due to the fact that people will get easy and fast access to information. Therefore, the popularity of automatic speech-based coding system has been greatly increased. The work of speech recognition further helps in establishing an easy way communication between interactive response system and users/ customers i.e. as a part of post processing of the speech recognizing process, we can accomplish some computational task with such a system making voice input as a trigger to do some task within the system.

Keywords: Speech Recognition, Machine Learning, Speech to Text, Natural Languages Processing

1. Introduction

Speech Recognition, it is the ability of the machine or program to evaluate word, idiom or a sentence in spoken expression and convert those words into a machine readable format. The more sophisticated software has the ability to obtain natural language as well. Speech recollection works using algorithms through acoustic and language modeling. In addition, acoustic modeling represents the link between linguistic units of speech and audio signals; whereas language modeling matches sound with string to help categorize between words that sound similar. Additionally, Hidden Markov Models are used as well to make materialistic patterns in a speech to enhance accuracy with the system. Furthermore, it is seen that a person working on a computer cannot work or type for longer duration because if they can then there will be an issue of back or wrist pain that will be pernicious for the human body, but it can be avoided easily by switching from typing to speaking whenever needed. Voice code, "Advanced Voice Control Platform", in these articles, writer faced some issues because of working on computer for longer duration, so they came up with the idea that, by taking advantage of our brain's natural aptitude for language we can control our computer more efficiently and naturally, so they created a plugin for UNIX based system. Nuzhat Atiqua Nafis, Md. Safaet Hossain, "Speech to Text Conversion in Real-Time", In this paper, they have used a method to design a text-to-speech version module by the use of MATLAB. That method is simple to implement and involves the much lesser use of memory spaces.

2. Proposed System

The prime objective of the project being proposed is to design and build a system that a basic user can interact so that she/he can make use of voice commands to deal with system i.e. making a system that has capability of recognizing the isolated speaker words and process the request to forward the given task.
• User can code easily and faster based on voice.
• To understand the basics of speech processing.
• To get knowledge on various speech recognition approaches.
• To get insights on speech responsive application development
• To Save the Time of user-system

This is a machine learning based Speech-to-Text software that will automatically write the converted text into the python programming language syntax. It based on the Natural Language Processing associated with the Speech Recognition API for recognizing the speech from microphone. System Flow Diagram The hardware requirements of our system are: 1. Computer System 2. Microphone Google Speech API For Speech recognition from Microphone we are using google Speech API Cloud Speech-to-Text Google Cloud Speech-to-Text enables developers to convert audio to text by applying powerful neural network models in an easy-to-use API. The API recognizes 120 languages and variants to support your global user base using Google's machine learning technology. Machine learning Machine learning is defined as the field of study that provides computers with the ability to learn from input data without being

explicitly programmed to do so. Among the two types of learning method we are using Supervised Learning as it has dataset containing Input as well as corresponding output.
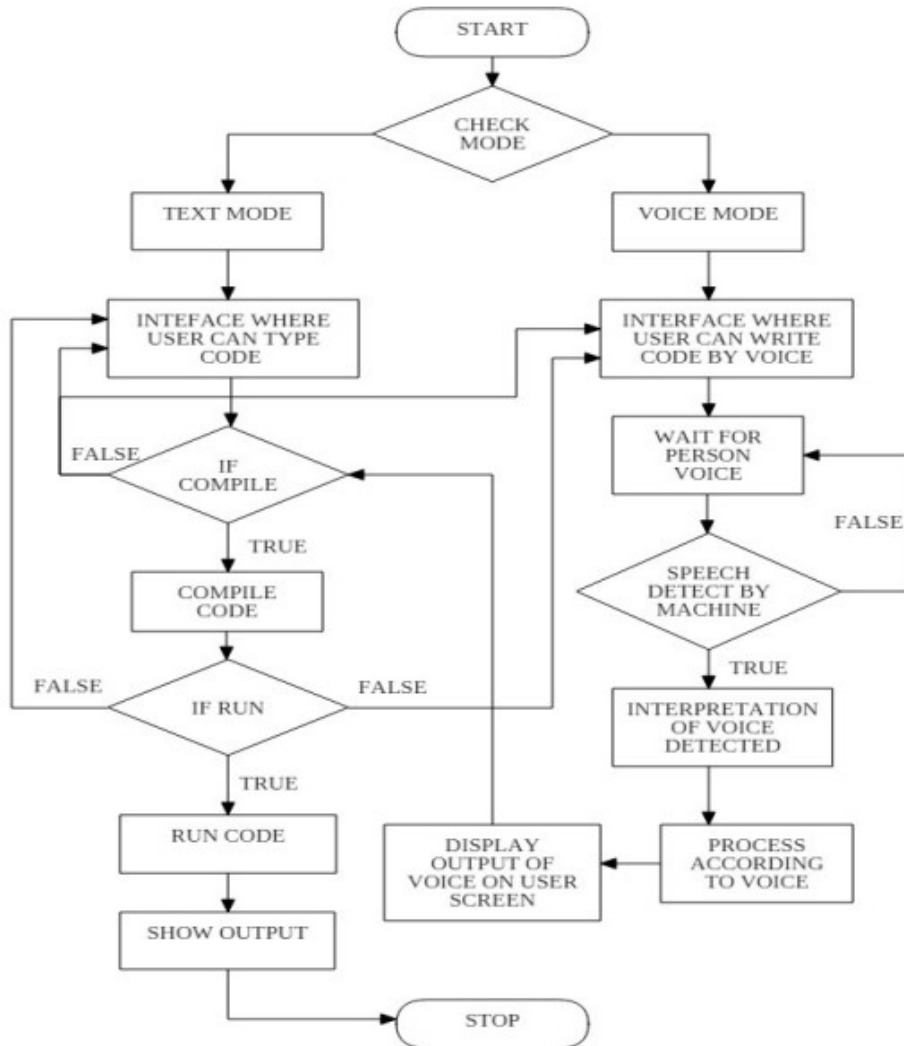
Flowchart

Fig. 1. Flowchart

3. Problem definition

Traditionally there is no application available in the market which can develop code in python using user voice.so we decide to create some interesting application to make coding in python using voice. This will be interesting for student to Perform Programs in Python based on Voice. This will save time of student to write code and to remember code.

4. Literature Survey

In this chapter I have given my critical evaluation & summary of all research papers that I read related to my project. Critical Evaluation of Research Paper Speech Recognition Using Neural Networks: A Systematic Review This paper provided a thorough analysis of speech recognition by extracting the evolution technique used. Speech signals can provide us with different kinds of information such as: • Speech recognition, which gives information about the content of speech signals. • Speaker recognition that carries information about the speaker identity. • Emotion recognition, which delivers information about the speaker's emotional state. • Health recognition, which offers information on the patient's health status. • Language recognition, that yields information of the spoken language. • Accent recognition, which produces information about the speaker accent. • Age recognition that supplies information about the speaker age. • Gender recognition, which carries information about the speaker gender. Recently data has become very easily obtained through numerous numbers of open sources. Extracting knowledge from data is considered the real challenge. With the use of computers and smart software that can perform numerous computations and calculations in seconds, the process of analyzing data has become easier. Moreover, learning from obtained data is also essential as adapting with new inputs is a highly important process that ensures the continuous development of any smart system. For that reason, a lot of attention has been given on the field of machine learning over the past years. Machine learning is defined as the field of study that provides computers with the ability to learn from input data without being explicitly programmed to do so. The learning process is done iteratively from analyzed data and new input data This iterative aspect allows computers to identify hidden insight sand repeated patterns and use these findings to adapt when exposed to a new data The different types of data used in this learning process can vary from observations and examples to instructions. The gained knowledge will help in producing reliable and repeated results. Thus, we can describe machine learning as a method that learns from past experiences and uses gained knowledge to do better in the future. Five main techniques of machine learning exists which are supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning and finally deep learning.

5. Conclusion and Future Work

The main purpose of this study is to develop code using voice in python Programming. We conclude that using this system user can perform programming very flexibly and easier way. It will also save Coding Time of User. We will add more python modules and different programming languages.

References

1. D. Ardagna, M. Comuzzi, E. Mussi, B. Pernici, and P. Plebani, "PAWS: A Framework for Executing Adaptive Web-Service Processes," IEEE Software, vol. 24, no. 6, pp. 39–46, 2007.
2. L. Baresi, S. Guinea, R. Kazhamiakin, and M. Pistore, "An Integrated Approach for the Run-Time Monitoring of BPEL Orchestrations," in Proceedings of the 1st European Conference Towards a Service-Based Internet (Servicewave08), 2008.
3. L. Bodenstaff, A. Wombacher, M. Reichert, and M. Jaeger, "Monitoring Dependencies for SLAs: The MoDe4SLA Approach," in Proceedings of the 2008 IEEE International Conference on Services Computing (SCC'08), 2008.
4. D. Shukla and B. Schmidt, Essential Windows Workflow Foundation. Microsoft .Net Development Series, 2006.
5. J. Skene, D. D. Lamanna, and W. Emmerich, "Precise service level agreements," in Proceedings of the 26th International Conference on Software Engineering (ICSE'04), 2004.
6. Davis, D. and Parashar, M. Latency Performance of SOAP Implementations. In Proceedings of the IEEE Cluster Computing and the GRID 2002 (CCGRID'02), Berlin, Germany, IEEE, 2002
7. Gregor Hohpe, Bobby Woolf, Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions, Addison-Wesley Professional (October 10, 2003), ISBN-13: 978-0321200686
8. Kohlhoff, C. and Steele, R. Evaluating SOAP for High Performance Business Applications: Real–Time Trading Systems. In Proceedings of the WWW2003, Budapest, Hungary, 2003
9. Menascé, D. A. 2002. QoS Issues in Web Services. IEEE Internet Computing 6, 6 (Nov. 2002), 72-75.

10. 10.Menascé, D. A and Almeida, V. A., Capacity Planning for Web Services: metrics, models and methods, Prentice Hall, 2001, ISBN 0-13-065903-7.
11. G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, pp. 504-507, 2006.
12. J. V. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," Journal of Clinical Epidemiology, vol. 49, pp. 1225-1231, 1996.
13. R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," Proceedings of the 23rd international conference on Machine learning, 2006, pp. 161-168.
14. R. Caruana, N. Karampatziakis, and A. Yessenalina, "An empirical evaluation of supervised learning in high dimensions," International conference on Machine learning, 2008, pp. 96-103.
15. J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 1/2015
16. Hasegawa-Johnson M. Multivariate-state hidden Markov models for simultaneous transcription of phones and formants. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP); 2000. Istanbul, Turkey: 2000:1323e26. http://www.isle.illinois.edu/pubs/2000/hasegawa-johnson00icassp.pdf (accessed 20 Jul 2011).
17. "Programming Language Trends - O'Reilly Radar". Radar.oreilly.com. 2 August 2006.
18. Rowles, C., Huang. X., and Aumann, G. "Natural Language Understanding and Speech Recognition: Exploring the Connections, Pr~c, 3rd lust. Conf. p_r! S~eech Science and Technoloav SST-90, Melbourne, Nov. 1990, pp. 374-379.