



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VI Month of publication: June 2021

DOI: <https://doi.org/10.22214/ijraset.2021.35798>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com



Virtual Personal Assistant for Desktop Automation using Selenium

Tummala Sri Ranga Sai Krishna¹, Chinni Varshith Reddy², Mohammed Shafee-Ur-Rehman³, K. Vikram Reddy⁴
^{1,2,3,4}Computer Science and Engineering Department, Matrusri Engineering College

Abstract— In recent years, Virtual Personal Assistants(VPA) have worked with utmost efficacy sorting out queries and specific tasks posted by the individual users on the website by AI and Natural Language Processing . VPA developers develop functions to either scrape the query result from the Internet. The result data include copious formats from a simple definition in Wikipedia to complex calculations or recommendations. However, VPA's designed for desktops do not work as extensively as the VPA's featuring in the smart phones . They do not provide a complete automation of desktop websites due to continuous and frequent development. The current desktop personal assistant's can show you the top results of the query 'Biryani', but cannot order on behalf of you. In this study, we propose a Virtual Personal Assistant ARCHER for desktop automation using Selenium by using the specifications of the behavior data of websites.

Keywords— Desktop Automation, Selenium, Virtual Personal Assistant

I. INTRODUCTION

The advent of technology if observed in the past few decades made it evident that the whole objective is to develop a platform which provides a sophisticated experience to users. Automation is key to make such platforms. The usage of virtual assistants for automation observed an exponential rise in the past few years. Inorder to make the voice assistants congenial they require a means of input which is both fast and reliable at the same time. So, in usage of voice commands to input the data into the system has become popular.

Today we have a wide range of Virtual Personal Assistants being deployed in the market, however Google's Google Assistant and Amazon's Alexa for the smart devices powered by the Android and Apple iphone's SIRI rule the market. While mobile phones are getting sophisticated VPA's, desktop's are being neglected. Windows powered CORTANA is in no place comparable with ALEXA or SIRI.

The mere concept of VPA in desktops are constricted to showing results in search engines or file explorers. For example, with current vpa users can only get the web results of "order Biryani!" but can not actually get the virtual assistant to do it for them. The proposed project focuses on developing a Virtual Personal Assistant for desktops which provides complete web automation with help of Selenium. The proposed application saves a lot of time and effort for users by automating websites usage through chrome browser along with other standard features for easing the daily chores which include from playing music, marking dates to sending email.

II. LITERATURE REVIEW

Nowadays people are frequently using Virtual Personal Assistant (VPA) that has a natural language user interface, such as Amazon Echo, Apple Siri, Google Home, Microsoft Cortana, Softbank Pepper, Sharp RoBoHon, etc. Assistants can interact with people using voice and text to help them for finding web content, managing their daily routine, finding a route, purchasing right goods, helping them in social conversation etc. Consequently, the need for the much sophisticated VPA's arised [1]. In Tulshan, Amrita & Dhage, Sudhir. (2019)."

Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa", While all the four IVAs were capable of answering nearly 17.35% of daily based questions, only Google Assistant was found to be most efficient by answering 59.80% of the question. Contrary to the expectations, Siri brought a bit of disappointment by answering only 45%. Whereas Cortana could only cover one third of the questions. That was surely a terrible defeat[2]. This survey indicates the need of VPA for desktops which could work as efficiently as Google Assistant and can replace Cortana. In Gojare, S., Joshi, R. and Gaigaware, D., 2015." Analysis and design of selenium webdriver automation testing framework" to test the web application, they proposed an automation framework that can generate the customized test reports and also analyze the failures using screenshots of failed test cases. Tester can maintain all the data from the central place. This framework is very useful for dynamically changing web applications [3]. A widely used approach to detect emotion is through voice signals. Anagnostopoulos *et al.* [4] provide a survey of speech based emotion recognition. Our proposed application is a fusion of general VPA with the above discussed selenium methodology and emotion recognition to provide users an automated web experience with voice command.

III. PROPOSED APPROACH

The proposed application contains three main operations which are explored below:

A. Voice-Command

The input to the application will be the voice command which the system processes and produces the output in both the textual and audio formats. There are several approaches to implement this Natural Language Processing system, we used “pyttsx3” text-to-speech platform available in Python.

- 1) The audio is received through the interface using Microphone as source. The received audio must be adjusted for ambient noise to have an efficient recognition.
- 2) Google’s “speech recognition” library available in Python is used for feature extraction with an Acoustic model producing the text format of the received audio file.

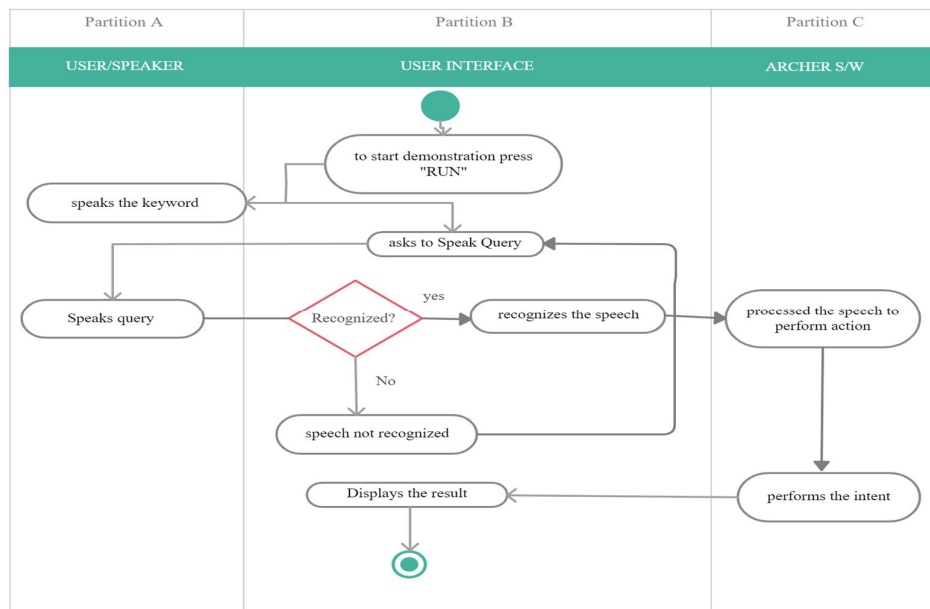


Fig. 1 An Activity diagram describing the voice command operation of the proposed application

B. Selenium based Automation

This is the important phase of the application ensuring the model’s idiosyncrasy with the existing models. In this phase we implement the selenium driver testing methodology to automate the web experience of users. In the proposed system, we used the “Selenium web driver” library available in Python to perform the automate commands. The implementation requires two key elements discussed below:

- 1) *Browser Driver*: Browser driver plays a key role as the Selenium Script creates an HTTP Request for each selenium command and sends it to the browser driver. An HTTP request is then sent to the server using Browser Driver. The steps are executed on the HTTP server. The execution status is sent to the HTTP server which is then captured by the automation script.
- 2) *X-path*: Even though the commands are sent the system should know where to perform the commands. That is Selenium driver sends the command to “click login button” to the browser, but how should the browser know what the login button is? X-path comes into play here. X-path is a unique address for behavioural data of the website which allows the browser to distinguish between the buttons, input fields, etc.

C. Recommendation based on the Emotional State of User

This phase is an experimental effort to provide a dynamic experience to the user. The proposed application ARCHER design allows it to work as Social Companion to the users. This is achieved by proposing recommendations of books, Songs and movies based on the mood of the user. Neural-network is used to achieve this feat. The modules of the Neural-network are as follows:

- 1) **Feature Extraction:** In pattern recognition and audio processing, feature extraction is a unique kind of dimension modification. If the input to the algorithm is colossal to process and is considered excessive, the input is transformed into a meager set of expression functions. Mel-Frequency cepstral(MFCC) is used in the proposed VPA ARCHER for feature extraction. The mel frequency cepstral coefficients (MFCCs) of a signal are a small set of features (usually about 10-20) which concisely describe the overall shape of a spectral envelope. In MIR, it is often used to describe timbre.
- 2) **librosa** is a python package for music and audio analysis It provides the building blocks necessary to create music information retrieval systems.
- 3) **MLPClassifier** stands for Multi-layer Perceptron classifier which in the name itself connects to a Neural Network. Unlike other classification algorithms such as Support Vectors or Naive Bayes Classifier, MLPClassifier relies on an underlying Neural Network to perform the task of classification.

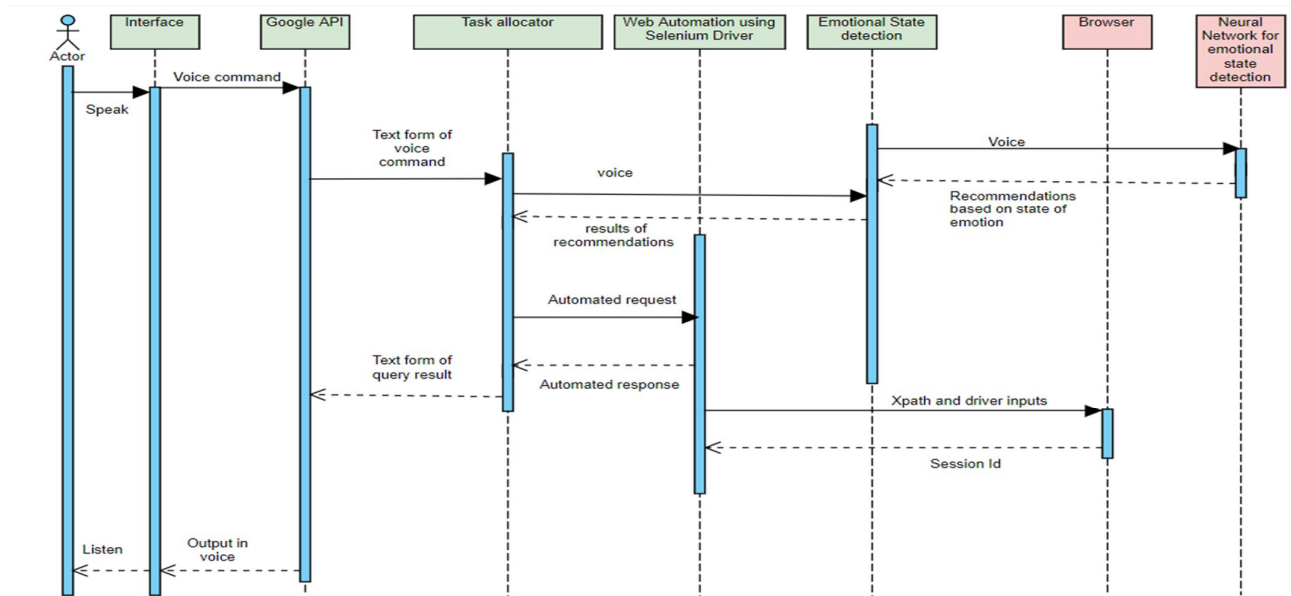


Fig. 2 Sequence diagram of the proposed Virtual Personal Assistant ARCHER

D. Neural Network Working

- 1) Step-1: Take Audio Input from Interface
- 2) Step-2: Convert the Audio Input into sinusoidal signals using Short Time Fourier Transform(STFT)
- 3) Step-3: Feature Extraction Using Mel-Frequency cepstral(MFCC)
- 4) Step-4: MLPClassifier (Multi-layer Perceptron classifier) model is used to map the features extracted from the MFCC to that of features in the Dataset.
- 5) Step-5: Train on that pattern
- 6) Step-6: Predict the behaviour

$$y = \varphi\left(\sum_{i=1}^n w_i x_i + b\right) = \varphi(\mathbf{w}^T \mathbf{x} + b)$$

Fig. 2 Formula of MLP Classifier where \mathbf{w} denotes the vector of weights, \mathbf{x} is the vector of inputs, \mathbf{b} is the bias and φ is the non-linear activation function

IV. CONCLUSIONS

The proposed Virtual Personal Assistant (VPA) concentrates on providing desktop users a complete dynamic browsing experience using selenium methodology for web automation and Neural-network based on Multi-Layer Perceptron classifier for detecting the emotions of the user and recommending accordingly. This proposal system enables the user to actually control the web application activities such as login,logout, buying,etc using voice commands and also provide distress support based on the emotio of the user.



REFERENCES

- [1] Ahmed, S., Qaasar, M., Sholikah, R. W., & Morimoto, Y. (2018, March). Early Dementia Detection through Conversations to Virtual Personal Assistant. In *2018 AAAI Spring Symposium Series*.
- [2] Tulshan, Amrita & Dhage, Sudhir. (2019). Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa: 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers. 10.1007/978-981-13-5758-9_17.
- [3] Gojare, S., Joshi, R., & Gaigaware, D. (2015). Analysis and design of selenium webdriver automation testing framework. *Procedia Computer Science*, 50, 341-346.
- [4] AnagnostopoulosC.-N.Iliou.Giannoukos.: ' Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011', *Artif. Intell. Rev.*, 2015, 43, (2), pp. 155–177
- [5] Sheikhan, M., Bejani, M. & Gharavian, D. Modular neural-SVM scheme for speech emotion recognition using ANOVA feature selection method. *Neural Comput & Applic* 23, 215–227 (2013).



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)