A Major Project on

SIGMA

"Personal Voice Assistant"

Submitted in Partial Fulfillment of the Requirements for

8th semester under

Pokhara University

Submitted by:

Sushant Tuladhar, 14738
Asim Giri, 14706

Submitted to:
Department of Software Engineering,NCIT

Department of Software Engineering
NEPAL COLLEGE OF
INFORMATION TECHNOLOGY
Balkumari, Lalitpur, Nepal

List of Figures

Table of Contents

1 Abstract:

SIGMA is a voice assistant which is built on Android platform. The main feature of Sigma is to provide service as a personal voice assistant to perform different tasks as requested in a smartphone. As the growing use of internet has increased in today's world data has increased massively too which is utilized by application SIGMA to filter out data and give the desired output of searched content with the application of machine learning.

In this project, we will be using different tools like Android Studio, Python, JSON, PHP, Firebase, MySQL, Java and so on. MySQL will be used as relational database to keep record of activity and different useful information whereas firebase will be used as cache database to handle real time operations.

Here, SIGMA will handle the voice command given by user by converting the speech to text and there are actually two different operations to check where there will be predefined operations to perform different tasks offline and find the desired output from massive data present in the internet through filtering of data by data mining which would be checked by breaking the voice command to natural language processing. The processed data set when giving output will convert the text to speech to deliver the searched result and notify user what they have requested could be accessed or not.

Keywords:

SIGMA, Voice Assistant, Natural Language Processing, Data Mining, Internet, Smartphone

2 Problem Statement

We have consider the various voice assistant applications and devices in the present market today like Siri (iOS), Cortana (Microsoft), Alexa (Amazon), Google Home etc. All these virtual assistant can really make our life comfortable and fun.

These applications are real time savers and people seem to get satisfied from using them. But one of the analysis done by Voice Labs, shows that most of these apps don't seem to get the reviews, which suggest that they are not so popular. People are using these apps but are not making its full use.

To develop this project we take into some of the consideration which is about the modern personal voice assistants which are present in the market. The modern technological giants Google, Microsoft, Yahoo, Apple, Huawei, and Samsung all are actually planning to develop a complete AI based software which could work as an assistant to help their customers to perform different tasks easily. The different kinds of voice assistant system which are prepared by other companies rather than this technological giants have made some static assistant which could only utilize predefined formats of data to given to give outputs. We want to create a machine learning system which utilizes the data from technological giants and filter them and produce the learning system so that we can create an intelligent agent which could give good results as output to satisfy maximum requests to accuracy.

Identification and manipulation of patterns of input data would be another problem to be faced. The other problems would be performing data engineering, cleaning, feature extraction/ selection etc. Creating and using of an independent Algorithm was another problem. Representation of input data in Machine Learning, Understanding the context and the intent of input data can be another issue. Other issues faced during its development process is finding the datasets for Machine Learning. Gaining access to useful datasets can also be challenging. As our project is based on the language specific (Nepali language) it is difficult to find the libraries for this specific language.

3 Project Objectives:

With the rapid development of voice interaction and speech recognition systems, the way how smartphones markets and their interactions in present market has also increased. The trend of giving commands to smartphones using voice as input to do different tasks daily can be of a greater importance for giving people a greater reliability to perform tasks easily.

The major objectives of this project are:

- Taking voice as input for natural language processing.
- Taking language specific commands (Nepali) from users.
- To discover patterns in your data and make predictions based on these data.
- Analyzing and Representation of input data.
- Performing actions/tasks based on these commands.
- To make interactions with users to make it more lively.
- The application must be laser focused on problem solving.
- Filtering out of the data from the massive data present.

4 Significance of study/project:

The major significance of this project is to help users to easily access their smartphones and perform different actions easily. SIGMA, a personal voice assistant is an assistant based on Android Platform. SIGMA has two categories, as one which does offline tasks to do different features as defined in the system to access different features of smartphones. The other does take input to search from internet domain filter it with the concept of data mining to search the filtered and required data so that it can be naturally processed back to perform different actions and retrieve results from different queries.

The main use of project is to focus on how to make it easy for users and giving higher accuracy of data which are being stored. Here, SIGMA follows the activity of user and checks what kind of data and what kind of operations are actually being done by user. The tracking of user activity is also done in offline system which is stored in local database where user's local activities are tracked to check how users are using assistant. The local activities will then be synced with the online system when internet is on. The more the data collected the more improvisation on assistant can be done. Backend server needs to do fully commit into the system otherwise reports can never be used. The server will also check from the cache server which is a replica of server made from real time database firebase.

5 Scope and Limitations:

The voice assistant application is just like a companion for our mobile devices. Right now voice assistant are taken as an extensions of voice-dialing and voice-search functions on phones but they're the next generation of consumer user interfaces. As tech companies have recognized the growing importance of voice assistant the scope and importance of project like ours has sky-rocketed.

Both voice assistant and NLP has existed for a long time but in recent years the advances in the machine learning, deep learning and neural networks has changed the face of the Voice assistant applications. At present you can provide thousands and millions of voice samples to these applications. Voice assistants will also continue to offer more individualized experiences as they get better at differentiating between voices. In the upcoming years the voice assistant will be able to uniquely identify multiple user's voice. To be clear, voice assistants are very useful and their application will continue to expand and become integrated into an increasing number of domains in our daily lives, Also, making the application language specific means even the people who are not familiar or those who have hard time interacting with today's top-notch and sophisticated mobile phones can easily interact giving certain commands to mobile phones in their own native language (i.e. Nepali).

This project can be of a great importance for people with visual impairments and those with physical problems (cannot interact with mobile phones in a normal way). Others may find it useful because talking is easier than typing and it is fun.

No project is perfect and there are also some limitations of this project, they are:

- Making voice assistant learn a new language means going through an elaborate development process.
- Translating sound into characters has always been a fundamental challenge.
- Lack of Accuracy and Misinterpretation.
- Visual limits of the voice assistant application unlike other applications.
- Not suitable for complex and multistep tasks.
- Background noise interference and difficulty in voice recognition if voice changes or changes in accents.

6 Literature Study:

For this project, we studied different types of existing systems which interacts with users with same kind of input methods to provide different services. The different existing systems are personal assistant developed by bigger technological giants like Siri (A voice assistant developed by Apple Inc.) , Voice Assistant, Alexa (A voice assistant developed by Amazon), Bixby (Voice Assistant which represents a major reboot for S Voice, Samsung's voice assistant app introduced in 2012 with the Galaxy S III) , Lyra (A voice assistant developed by Artificial Solutions which works for both Ios and Android), Cortana (A voice assistant developed for Windows 10 by Microsoft), Amazon Echo etc.

We show some of the popular voice assistant present in the market with their ratings according to users can be shown here:

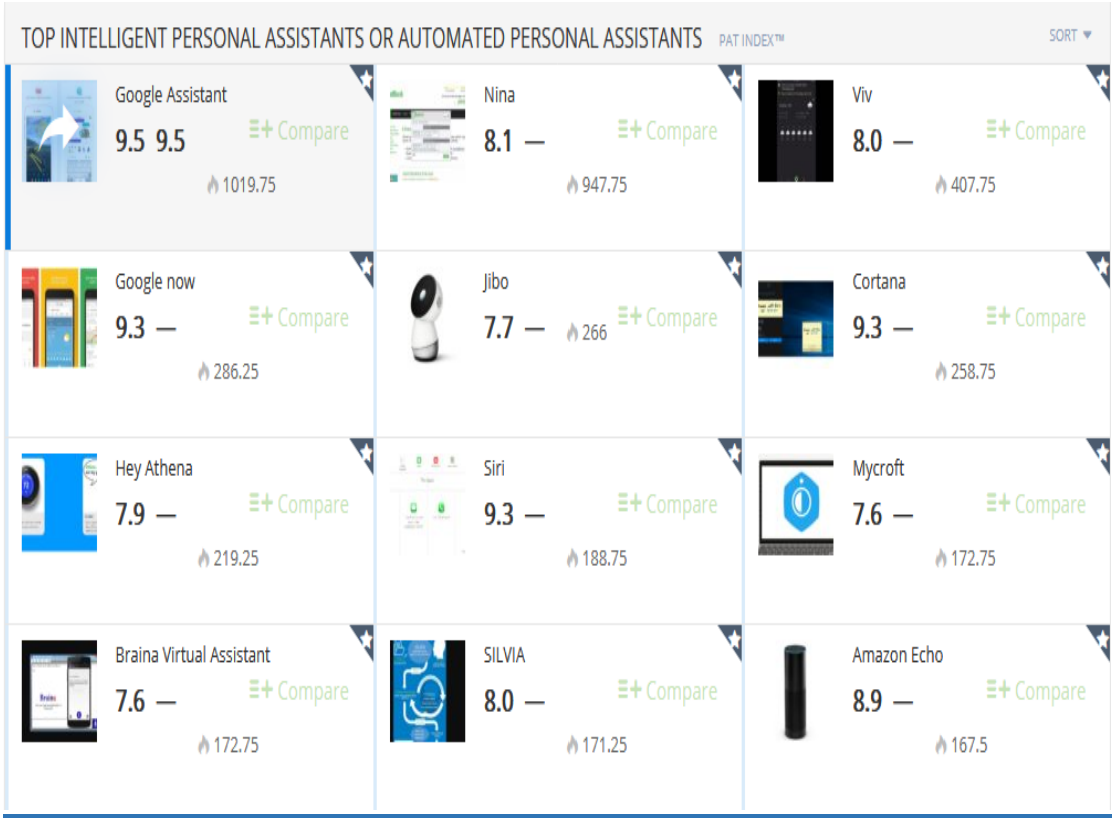


Figure 1: List of Voice Assistant with ratings

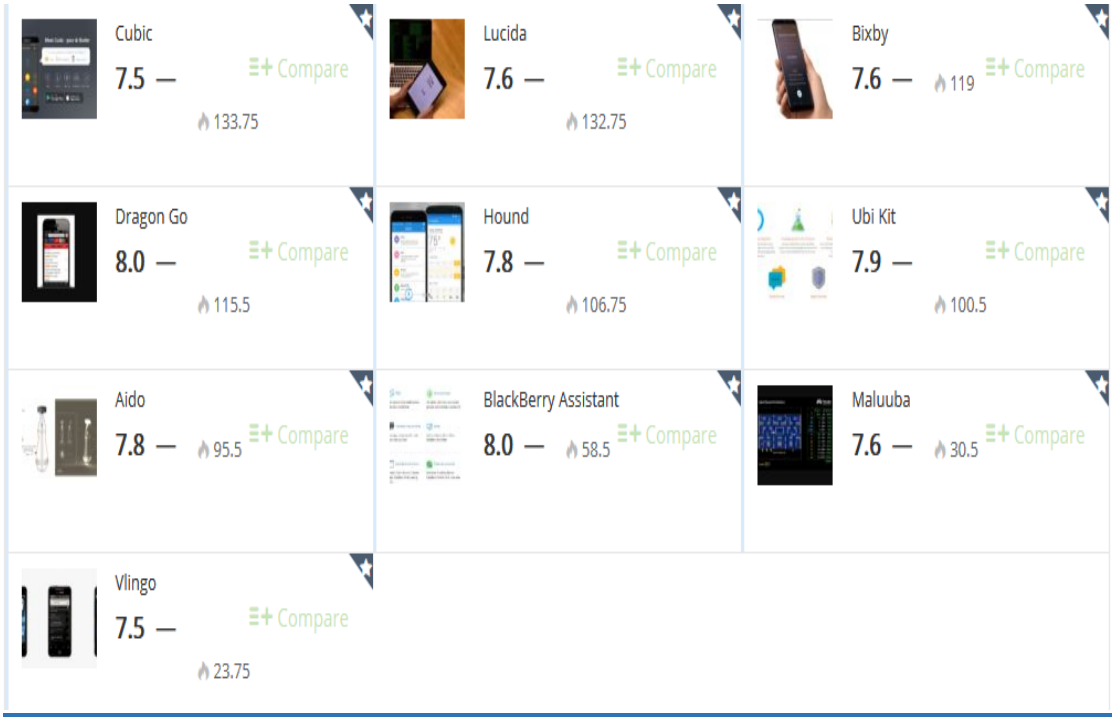


Figure 2: List of Voice Assistant with ratings

6.1 Voice Assistant and its past:

With the rise of machine learning and natural language processing skills, the technological era is now being transferred to voice assistants which could interpret different languages to perform different operations. The development of how voice assistant have progressed compared to its predecessors is an important part of study for this project. Here, we discuss about the past of voice assistant.

Back in 1784 Wolfgang von Kempelen creates the Acoustic-Mechanical speech Machine in Vienna. Kempelen's first experiment with speech synthesis involved only the most rudimentary elements of the vocal tract necessary to produce speech-like sounds. A kitchen bellows, used to stoke fires in wood-burning stoves, was invoked as a set of lungs to supply the airflow. A reed extracted from a common bagpipe was implemented as the glottis, the source of the raw fundamental sound in the vocal tract. The bell of a clarinet made for a sufficient mouth, despite its rigid form. This basic model was able to produce simple vowel sounds only, though some additional articulation was possible by positioning one's hand at the bell opening to obstruct airflow.

However, it was not until the 1950s that this line of inquiry would lead to genuine speech recognition. Up to this point, we see attempts at speech creation and recording, but not yet interpretation. Audrey, a machine created by Bell Labs, could understand the digits 0–9, with a 90% accuracy rate. Interestingly, this accuracy level was only recorded when its inventor spoke; it hovered between 70% and 80% when other people spoke to Audrey.

The IBM shoebox was a 1961 IBM computer that was able to perform mathematical functions and perform speech recognition. It recognized 16 spoken words and the digits 0 through 9. It was developed by William C.

IBM Tangora, released in the mid-1980s and named after Albert Tangora, then the world's fastest typist, could adjust to the speaker's voice. It still required slow, clear speech and no background noise, but its use of hidden Markov models allowed for increased flexibility through data clustering and the prediction of upcoming phonemes based on recent patterns. Although it required 20 minutes of training data (in the form of recorded speech) from each user, Tangora could recognize up to 20,000 English words and some full sentences.

It was only in 1997 that the world's first "continuous speech recognizer" (ie. one no longer had to pause between each word) was released, in the form of Dragon's NaturallySpeaking software. Capable of understanding 100 words per minute, it is still in use today (albeit in an upgraded form) and is favored by doctors for notation purposes.

This culminated in the launch of the Google Voice Search app for iPhone in 2008. Driven by huge volumes of training data, the Voice Search app showed remarkable improvements on the accuracy levels of previous speech recognition technologies. Google built on this to introduce elements of personalization into its voice search results, and used this data to develop its Hummingbird algorithm, arriving at a much more nuanced understanding of language in use. These strands have been tied together in the Google Assistant, which is now resident on almost 50% of all smartphones.

It was Siri, Apple's entry into the voice recognition market, that first captured the public's imagination, however. As the result of decades of research, this AI-powered digital assistant brought a touch of humanity to the sterile world of speech recognition.

After Siri, Microsoft launched Cortana, Amazon launched Alexa, and the wheels were set in motion for the current battle for supremacy among the tech giants' respective speech recognition platforms.

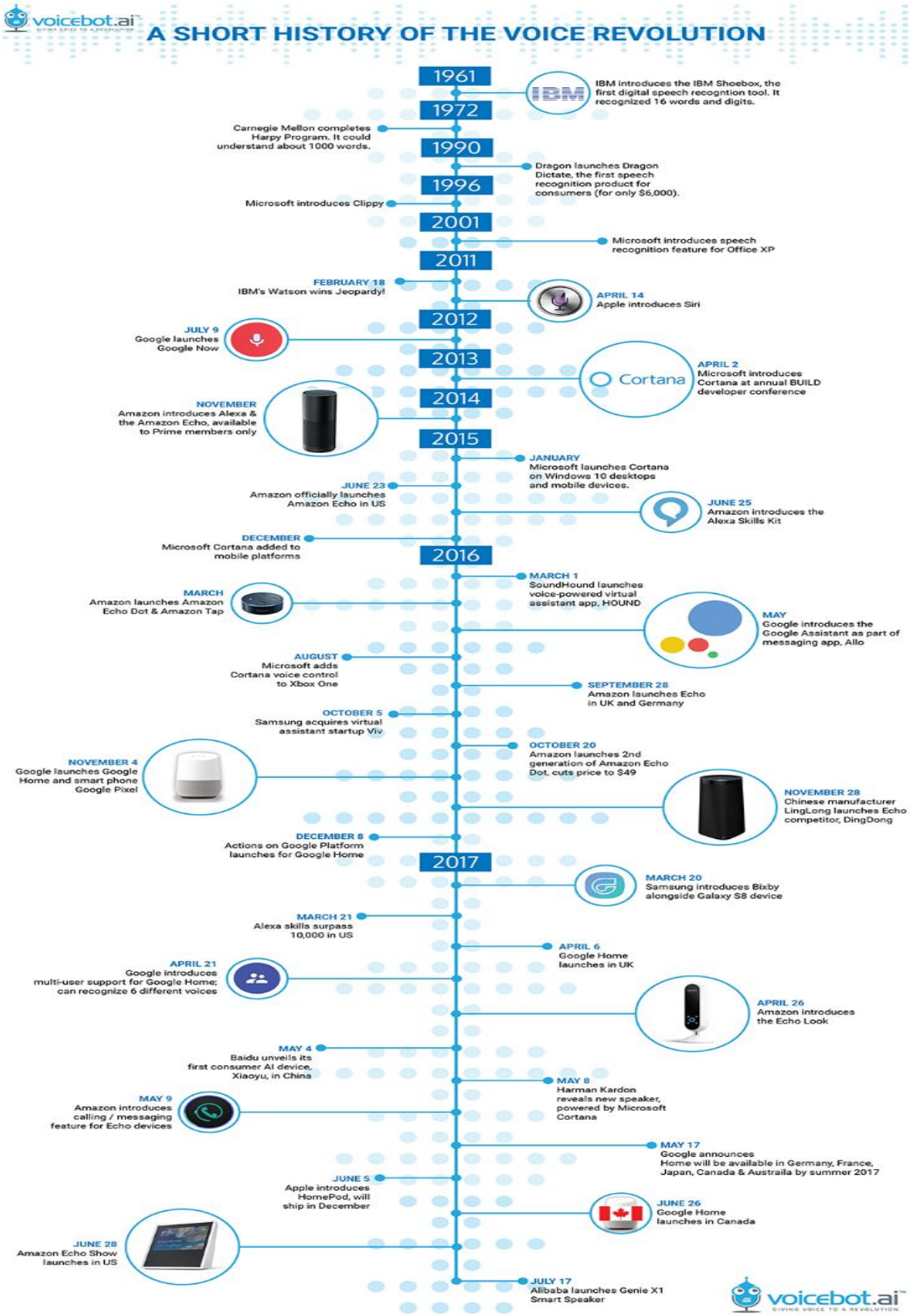


Figure 3: History of Voice Assistant

7 Proposed Methodology:

If we talk about the current methodologies for developing mobile applications they are mostly based on the API (applications programming interfaces) offered by the native platform. Hence most of the solutions to development of an app are characterized by a low portability and reusability.

Mobile app development is not an easy job and every app developer develops an app with a high hope of success. We must develop an app in such a way that it ensures a minimal development cost and maximum output. For this we must have a clear cut methodology.

Among the various approaches we have selected Agile methodology/approach. Agile development is quick and iterative. Here the tasks are divided into short phases of work. We can quickly adapt plans and ways of doing things. Agile methodology is way more adaptive and helps us to create apps that are seamless and easy to work upon. Agile methodology have proven to be quite efficient and helpful in mobile app development. Agile development makes the app more stable with few errors thus increasing the quality.

7.1 Type of Agile Development (Overview of Scrum):

For this specific project we have used the Agile Scrum which is a type of Agile development framework. In this technique the product owner creates a product backlog (like a wish list of tasks). The important part is that the product backlog must be prioritized. That's the job of the Product Owner. Each scrum is managed by an individual project manager called scrum-master. The Scrum team conducts a sprint planning session.

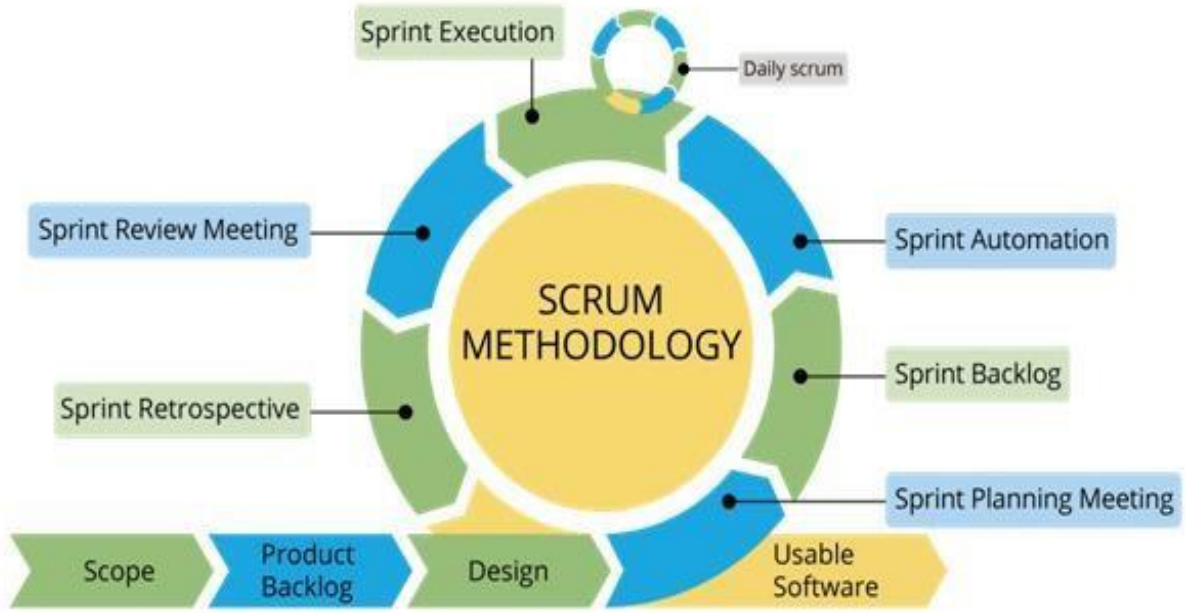


Figure 4: Scrum Methodology

Some of the reasons why we have selected Agile Scrum as a development approach:

- It follows a team based approach.
- It allows us to make changes, add new features and keep evolving with the changing trends.
- Testing can be performed at each stages.
- Customer satisfaction by continuous and rapid delivery of app modules.
- It also ensures effective use of time and money.
- Large projects are divided into small sprints.
- Massive involvement of users/customers in the overall development of the project.
- The individual effort of each team member is visible during daily scrum meetings.

7.2 Implementation of Scrum Framework:

To implement the Scrum framework in our project first we have to consider ourselves as a Product Owner. We will have to create the ‘Product Backlog’ and also have to prioritize the tasks. During the development process, any team members can add anything to the product backlog, but only Product Owner has the right to prioritize it.

The next step will be to determine the ‘Sprint’ (timeframe), within this timeframe we (team) must complete set of tasks from the ‘Product Backlog’. Also, each Sprint ends with a review where the team reviews their work and discusses way to improve the next Sprint.

We must also decide the team’s Sprint budget. This will provide us the rough idea on the available numbers of hours for people in the Sprint. We must try to deduct an approximate time that team members will not be able to spend working on the Sprint.

The next step will be to create a collaboration hub. It will act as a place where the team will discuss and give feedback on one another. These steps would surely help us to boost our product development success rate in several ways.

Nothing is perfect, and Scrum methodology is no exception. But proper planning and smart decision making can help us get past these disadvantages with Scrum methodology. For example, the Scrum master can be hold the responsibilities with definite goals, so that there is no compromise on quality and no excuse for failure.

8. Task done so far:

Different tasks were performed for the making of this software SIGMA and they are described with different iterations which were performed when following the SDLC (Software Development Life Cycle). The different tasks which were done during different stages of SDLC are defined below:

8.1. Requirement Analysis:

Requirement analysis involves the step of collecting the requirements of the software and how it should behave and perform when it goes live. The different steps which were performed during this stage are as described below:

1. Study of different datasets which were available and how feasible they can be when used with different programming language and their output time were also calculated.
2. Study of how different assistants performed and how they handle different operations given to them.
3. Preparation of SRS was done for the database schema before implementation of it in real life.
4. Significance of how the real program should actually behave and how they can be bug free.
5. Dividing of how tasks can be performed in a team.
6. Sprint Planning Meeting importance which was done in different interval to how we can enhance different features and what are the problems that are faced during building of process.

8.2 Designing:

Designing involves different steps and procedures which were followed for building this project. The different designing steps which were followed for making this software are as below:

1. The schema design which were discussed in the requirement analysis steps and during different Scrum Planning Meeting was actually designed.

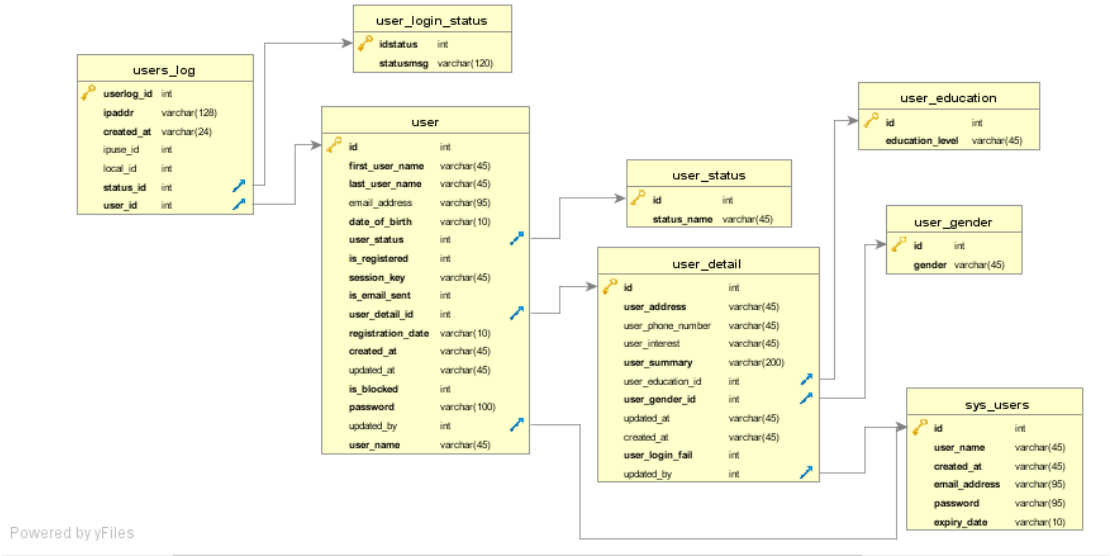
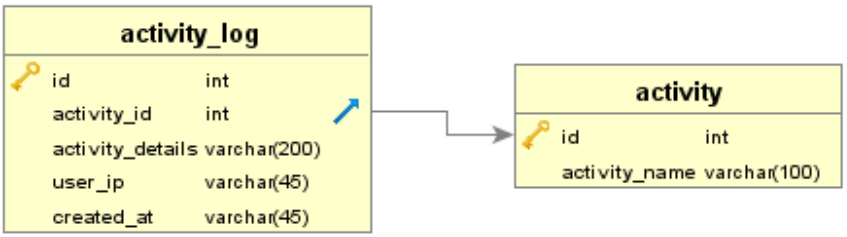


Figure 8.2.1.1: Database Schema Design 1



les

Figure 8.2.1.2: Database Schema Design 2

2. Different helper class for Android Database Connection, Remote Database Connection and different layouts which needs to be developed was made different design models.

8.3 Coding/ Implementation:

This steps involves the changing of design models into implementation. The different steps which were done during this implementation process are described below:

8.3.1 Development of Front end:

The development of front end application was done through JAVA programming which was done in Android Studio. Different design models which were created was implemented in this model.

8.3.2 Development of Middleware:

The development of Middleware was done by implementing Python as the middleware for performing different operations of Machine Learning and different other tasks.

8.3.3 Development of Backend:

The development of Backend was done by implementing PHP as the backend server language. The backend interacts with the middleware and performs different operations.

8.4. Testing:

The testing involves different methodologies which were done during performing this project. The different testing methodologies which were used are as shown below:

8.4.1 Unit Testing:

As all the models were built independently they are done unit testing for each models which were created in front-end, middleware and back-end. As all the models passed the unit testing. Each class and their interactions were checked.

8.4.2 Model Testing:

Each model of front end, back end and middleware were tested independently and passed the test.

8.5 Documentation:

The documentation is also one of the important process which needs to be done in the project. As the project was moving forward so the documentation was also done in the incremental manner. With changes done and implementation of changes which was passed in the development stage through testing were documented in the proper manner. The proper documentation gave us guidance and easiness for proper implementation.

9 Results and Discussion:

The development of front end was done using Java programming in Android Studio. The monitoring of how an application performed was done in the Android Studio. The testing of different independent classes has been done. The results of testing showed positive results. The monitoring of network was also done through inbuilt functions of Android Studio.

The middleware consists of python modules which performs different activities as well as different machine learning algorithms in the server end. Python acts as middleware for performing different functions when called by PHP modules.

The backend consists of PHP modules which when getting request from an android device performs different activities. The users have to login to access the different information present in server. User validation and verification are also performed. The user sends the voice message in the server which is then processed in the server.

10 Task Remaining:

Developing an application of any kind is not an easy tasks. It requires a lot of analysis, research of the existing application, requirement gathering from the end users etc. SIGMA was no exception to this. Applications which perform direct interactions with the users in a simplistic way as possible (through voice in this case) are much more difficult as there are many levels of development tasks to be considered and much more exceptions and errors to be considered.

As no project/application ever developed is hundred percentage accurate. There is always some sort of loop holes, unfinished tasks or some functionality which the application may not be able to perform, or contain defects. Similar is the case with SIGMA.

These are some of the lists of tasks remaining in SIGMA which always can be worked out later are:

- Identifications and manipulation of patterns of input data.
- SIGMA is unable to perform data engineering.
- Cannot fully represent the input data in Machine Learning.
- Cannot fully understand the context and the intent of input data.
- Integration of different modules needs to be done.
- Feature extractions as well as its selections.

These are some of the tasks that are remaining in SIGMA or which SIGMA is unable to function properly. In the further iterations of SIGMA we will be working on these remaining tasks and see what we will be able to accomplish.

11 Performance Analysis and Validation:

Performance analysis is a widely accepted method for measuring service quality that is well known of its simplicity and stress-free application. The concept of IPA (Importance-performance analysis) was introduced in 1977 by Martilla and James. Essentially, the idea of IPA comes from the theory of customer satisfaction as a function of the expectation on important attribute and the judgment of attribute performance toward customer satisfaction is linear and symmetric. Thus, IPA focuses on the gap between the customer expectation on the importance and judgment on the performance of specific attribute of service consumed.

Well the factors which were checked for the performance analysis and validation of software are as mentioned in the below points:

1. The importance of each attribute of the system were checked and their performance and interaction were also monitored.
2. The judgment of the performance of the system was done with the help of team mates and some of the external person who helped to review the system.
3. The validation of the data components which were given were checked through the data sets which were available and rendering time of the data were calculated.
4. Verification and validation of user which are allowed to connect to the system are classified through the email verification link which is sent to client using the system. Once user has been verified there is no need to do the verification. Validation of the system to correctly process the data was checked. The performance time for sending email and response to client was also calculated.

12 Deliverables/Output:

SIGMA will be able to cover voice commands, voice searching, and voice-activated device control, letting us to do things like send messages, check appointments and so on the Android devices, just like Apple's Siri on iPhone or iPad.

SIGMA will be able to assist users with various tasks like Requesting information, similarly to inputting a search query, performing various tasks like playing music.

These are some of the lists of deliverables/outputs that SIGMA is intended to provide or fairly represent:

- Voice-enabled, voice control, voice interaction, voice queries.
- Natural Language interactions i.e. Commands, Results.
- Find information. Weather, Traffic, News.
- Answer questions.
- Make Recommendations.
- Support for multiple users on single device. Example, Google Assistant, Voice Match etc.
- Personalization, Adaptation, Responses and actions.

13 Project Task and Time Schedule:

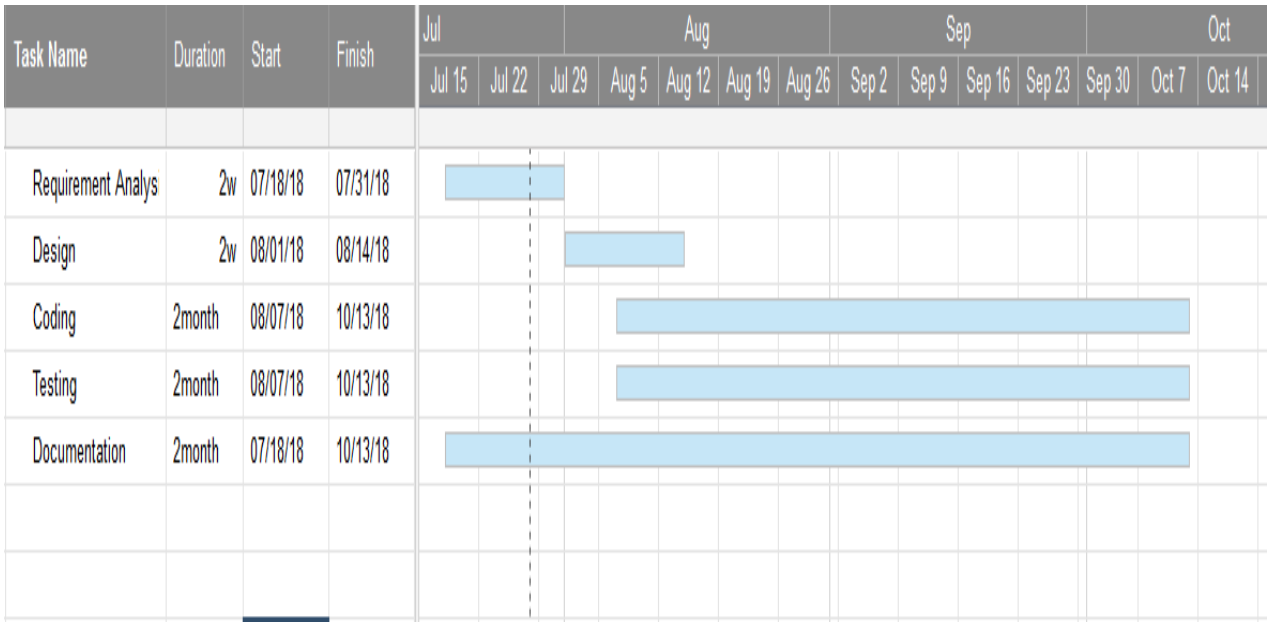


Figure 5: Overall Process in Gantt chart

The chart above shows the overall tasks and its time schedule. Since we followed the Scrum framework our first task was gathering the requirements. We gathered the various requirements based on our project. Upon gathering the requirements first we considered ourselves as a product owner. We try to explain what a user wants in the application, what the expectations of the users or what a user want to see in such an application.

The next step was to make mere concept of the project into its structural form. The designing process involved various design phases along with prototyping. We focused on designing the smaller units of the overall product. Since the design of the product gives us an idea of how the product will look upon its completion.

After visualizing/structuring the product we are developing (designing) it we have to make it functional, making it functional requires us to code. Most of the development process involved coding the product.

The next phase was testing. But as we followed the Scrum framework, testing the process was crucial in almost every phases of development. Testing helped us to

know how functional our application actually is. It also helped us to know the various defects of the application, which we tried to counter with.

Documenting the project performed is as crucial as any other any tasks. Documentation of the project performed starts early in the life cycle/development phase. We tried to document every work performed along with the important information, descriptions of the project.

14 Bibliography/References:

- Voice Assistant and its past :
  https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf
  Access Date: 9th July, 2018
- Information about different libraries:
  https://pypi.org/project/SpeechRecognition/
  Access Date: 10th July, 2018
- Challenges faced during making Voice Assistant
  https://softwareengineering.stackexchange.com/questions/207926/writing-my-own-voice-recognition-code
  Access Date: 11th July, 2018
- Natural Language Processing Basics and how it can be used in Python
  https://pythonprogramming.net/tokenizing-words-sentences-nltk-tutorial/
  Access Date: 15th July, 2018