An Automatic Assessment of Multilingual Recognition with GUI

Saranya.K[#], Rajasekar.S.S[*]

[#,*] *Assistant Professor/CSE/Bannari Amman Institute of Technology, Tamil Nadu, India*

***Abstract*— A great deal of negotiations throughout the past few years regarding the possibility of Text-to-Speech machines synthesize natural synthetic speech from texts. In the present paper, a Text-to-Speech synthesizer is developed that converts text into spoken word, by analyzing and processing it using google Text-To-Speech (gTTS) technology to convert this processed text into synthesized speech representation of the text. Similarly, Voice-to-Text synthesizer is developed that converts spoken words into text, by analyzing and processing it using speech_recognition as sr. This work is fully developed in Python language which is currently demanded in market using Python GUI Tkinter. This project increases the work opportunities for the physically challenged people and has increased their economy status.**

***Keywords*—google Text-To-Speech technology, speech_ recognition, GUI Tkinter.**

I. INTRODUCTION

An automated billing machine for deaf and dumb is enforced with few ideas of machine learning with Python programming tool. Python enables to work quicker and integrates the systems additional expeditiously. It uses English keywords whereas different languages use punctuations and has fewer grammar constructions than different languages. Python is a processor at run time by the interpreter.

The program execution is preferably done only before the compilation. Speech Recognition is the procedure of extracting essential information from input speech signal to make accurate decision about the corresponding text. Text to speech is the automatic conversion which configures the concept of speech recognition, speech analysis, speech synthesis, speech tuning, speech alteration etc. This concept can be implemented through Python.

Machine learning is a strategy to find out from specimen and practices, while not being specifically programmed. Rather than writing code it is focused to feed knowledge to generic algorithmic program, and it enables to build logic support of the information given. Machine leaning is a field with the computing knowledge. Applying AI, the requirement to make higher and intelligent machines is possible. There are three kinds of machine learning algorithms such as supervised Learning, unsupervised Learning and Reinforcement Learning.

Machine learning is associated with machine that additionally focuses on prediction-making through the utilization of computers. Inside the sector of knowledge analytics, machine learning can be a technique devises the advanced models and algorithms that lend themselves to prediction; in business utilization, this is often called prognostic analytics.

By victimization these two prospective the Deaf and Dumb Billing System is meant to beat the physically challenged people's complex by providing work opportunities even in skilled work places that requires the speaking and listening capability. For cross checking the calculated total worth in build calculator is embedded. Additionally for sales man reference the distinctive bill range is generated which helps the cashier to differentiate the orders of the customers.

1.1 Python Features

- *Easy to learn:* Python has few keywords from a straight forward structure and clearly outlined syntax.
- *Easy to read:* Python code is more clearly outlined and simple to handle.
- *A Board standard library:* Python's library is incredibly moveable and cross-platform compatible with UNIX, Windows and Macintosh.
- *Interactive mode:* Python supportive in associative interactive mode that permits interactive testing and debugging of snippets of code.
- *Portable:* It runs on a large type of hardware platforms and has a similar interface on all platforms.
- *Extendable:* It adds low-level modules to the Python interpreter .These modules change the programmers to feature or customize their tools to be a lot of economical.
- *Database:* Python provides interface to any or all the major business databases.
- *GUI Programming:* Python supports the GUI applications that can be created and ported to several system calls, libraries and windows system like Windows MFC, Macintosh and therefore the X-Window system of UNIX.

II. TOOLS DESCRIPTION

2.1 GUI Tkinter

The essential GUI toolbox is Tk, Python's default GUI. Tk can be accessed from its Python interface called Tkinter. Tk isn't the most recent and most noteworthy, nor does it have the most vigorous arrangement of GUI building blocks, yet it is genuinely easy to utilize, and can construct GUIs that keep running on generally stages. It is originally intended for the Tool Command Language (Tcl). Because of Tk's prominence, it has been ported to an assortment of other scripting language, including (Perl/Tk), (Ruby/Tk), and Python (Tkinter).The blend of Tk's GUI improvement movability and adaptability alongside the straightforwardness of a scripting language coordinated with the intensity of frameworks language gives the instruments to quickly to structure and actualize a wide assortment of business quality GUI applications.

2.2 Google Text-to-Speech (gTTS)

Text-to-speech systems are able to produce high quality speech, with great naturalness. This high quality is mostly outcomes due to large speech database used in building the synthetic signal. Most of those systems produce speech in a neutral style that limits their field of application and frequently prevents the user from utilizing them. TTS use to convert a text into speech that resembles, as closely as possible for a native speaker of the language who trying that text. TTS is technology by which a computer can speak to user and give the computed information. TTS system acquires text as input and then a computer algorithm called TTS engine analysis a text, pre-processes the text and synthesis the speech with some mathematical model. The TTS engine usually generates complete data in an audio format (.mp3) as the output even in a Multilanguage format.

2.3 Voice to Text

Speech Recognition is the procedure of extracting essential information from input speech signal to make accurate decision about the corresponding text. Computer follows human voice commands with the help of speech recognition mechanism and understands human languages i.e it acts as good interface human computer interaction. Generally, today's speech recognition technologies are designed for English language. So that illiterate rural communities or educationally under-privileged people are being kept away of computer technology. If the processing of computer technology in native language is made possible i.e if computer technologies can apprehend the native language then it will be easy to use computer technologies for illiterate people.

III. LITERATURE REVIEW

Abdulla[1], proposed a reference layout for speech recognition systems which communicate with human.Speech Recognition System (SRS) is otherwise called Automatic Speech Recognition (ASR) or PC discourse acknowledgment which is the way toward changing over a discourse flag to a grouping of words by methods for a calculation executed as a PC program. It has the capability of being an essential method of connection between people and computer

A. E. E.El Alf et al[2] has presented a Arabic sign language mobile chat application on Intelligent Arabic text to Arabic Sign Language Translation for Easy Deaf Communication. Focal points were this information based framework has tackled number of Arabic language issue, for example, equivalent words, inflectional, derivational, diacritical and plural, permitted finger spelling interpretation and drawbacks were this framework had not permitted to video handling, this framework had not made an interpretation of Arabic dialect content to Arabic communication through signing, it was hard to coordinate punctuation guidelines of Arabic language with Arabic sign language.

Butzberger, et al [3], proposed a Spontaneous speech effects in large vocabulary speech recognition application for speech analysis technique. Discourse information contains distinctive sorts of data that demonstrates a speaker's personality. This incorporates speaker's particular data because of vocal tract, excitation source and conduct include. The physical structure and measurement of vocal tract and excitation source are one of a kind for every speaker. The discourse examination manages stages with reasonable edge measure for portioning discourse motion for further investigation and extracting information.

H. Hon, et al. [4] proposed a simple rule that is able to predict the place of articulation with high accuracy. The rule relies on the frontness of the vowel following the artifacts, and is consistent with articulator principles.

Lakshmi Sahu et al[5],presented a Text analysis task for identifying words in the text. Text normalization includes Token Identification, which identifies the special symbols and numbers. Token to Words, which convert the identified tokens to words for which there is a well defined method of pronunciation.

Masatsune Tamura et al[6], proposed the framework for TTS, which is utilized for recognized the qualities of voice by changing HMM parameters of the speech units in the MLLR adjustment structure. To create speech with a self-assertively given focus on speaker's voice, The speaker autonomous models is adjusted, i.e., normal voice models, to the objective speaker. This method is presented for adjusting voice qualities and prosodic highlights of HMM-based TTS framework to a self-assertively given target speaker.

Poonam S.Shetake[7] , proposed a text-to-speech (TTS) convention transforms linguistic information stored as data or text into speech.TTS systems make it possible to access textual information over the telephone. The synthesizer produces speech signals of 16 bits, the sampling rate of which is determined by the sampling rate of the diphone database used. A text to speech (TTS) synthesizer is a system that can read text aloud automatically, which is extracted from Optical Character Recognition (OCR).

Reddy, D. R.[8],proposed a speech recognition technique by the direct analysis of the speech wave. This procedure utilizes best probabilistic model so it is most effective strategy for discourse acknowledgment. The most well known procedure of all qualities of information is Hidden Markov modeling. This method utilizes best probabilistic model so it is most proficient system for discourse acknowledgment.

Tapas Kumar Patra et al [9], proposed an educational learning tool application. This exploration has utilized model strategy. The framework was enhancing data access, as well as get as primary preferred standpoint and inconveniences of this exploration. It was not interpret Indian communication through signing in to the content, framework takes just straightforward English sentences as info, sign synthesis module utilizing an energized symbol has not been produced, and some language rules can't be connected to make an interpretation of English to ISL, not given clear thought regarding how framework functions appropriately.

IV. MOTIVATION

Nowadays, billing system is processed in online, which require internet but it is not applicable in remote areas. At the same time the physically challenged people are unable to find work opportunities in private sectors such as cashier in restaurant, stationary jobs and in places where automatic billing systems are being implemented.

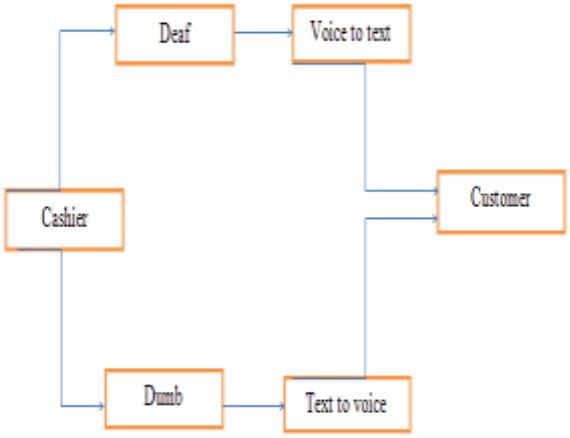

Fig 4.1 Overview Architecture

V. PROPOSED SYSTEM

The project is all about Billing System for Deaf and Dumb. Nowadays manual billing consumes more time. This GUI is created for user friendly interaction. The ways of implementing this software are given as follows:

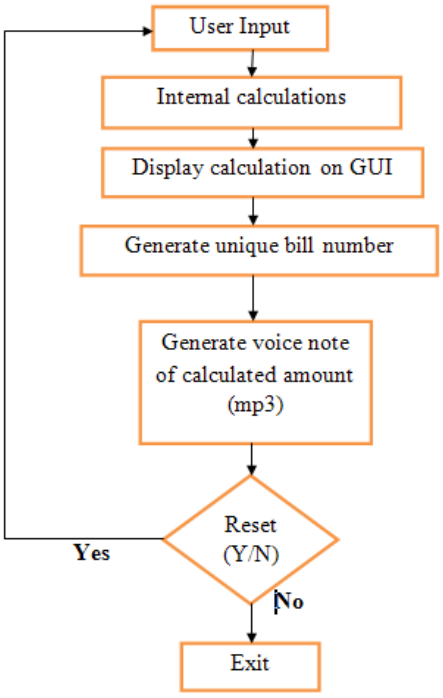

Fig 5.1 Flow chart

5.1 User Input

In this half, A range of menu has to be inserted for required things by the client for every item like carpets, fabric and blinds and this module even have choices for home delivery. If home delivery option is given then corresponding charge is also intersalary with total value supported amount of things else it may not be included in total calculation.

Fig 5.2 User Input Fields

5.1.1 Pre-Defined fields

It consists of two predefined fields like

- Sales Reference – Here random range can generate that is exclusive tosspot specific customers order.
- Order of date – The date of ordering the things is shown mechanically supported the system date.

5.2 Internal Calculation

During this half, cost of carpets, value of cloth, value of blinds, paid tax, sub-total, total cost are going to be generated mechanically by clicking the total button. First, it calculates the worth of everything supported in the figure of amount and show the whole price as well as GST that is calculated internally. At the equivalent time the calculator even be obtainable on the left facet of the screen for cross checking.

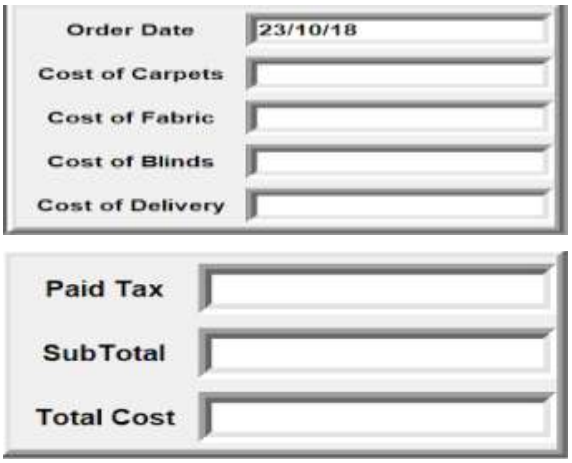

Fig 5.3 Internal Calculation Fields

5.3 Buttons

Total cost, sales reference, Reset and Exit are the buttons that is provided using the user interface.

- Total Price – Total buttons provides the calculation of the bill.
- Sales References – On click, generates random variety of number on the sales reference field.
- Reset – Resets all the values.
- Exit – Closes the window.

Fig 5.4 Buttons

5.4 Text to Voice

During this module, the given text is reborn into voice for communication with client with efficiently. The full value is spelled out explicitly just in case of the cashier is physically challenged that is unable to talk and even the opposite details like balance enquiry or any queries relating to purchase or exchange of product and others information are efficiently exchanged between the client and the cashier.

5.5 Voice to Text

In this module the given voice note is converted into textual equivalent in order that the cashier who is unable to listen to be able to respond the client on seeing the matter illustration of the purchasers voice note.

VI. RESULT

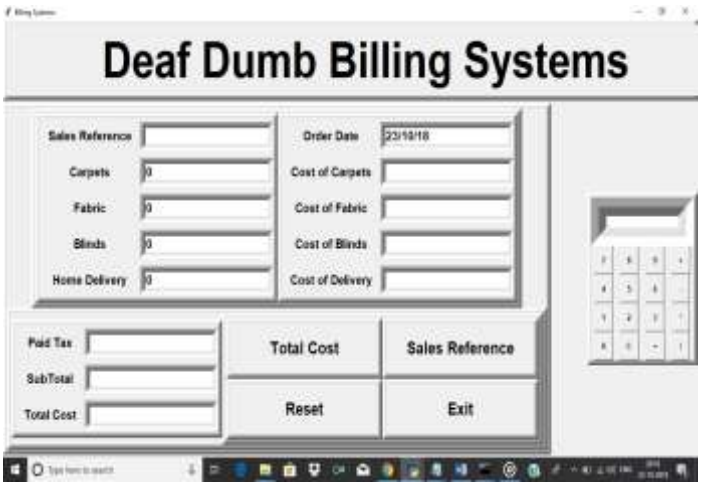

Fig 6.1 GUI

Fig 6.2 In Build Calculator

Fig 6.3 User Input

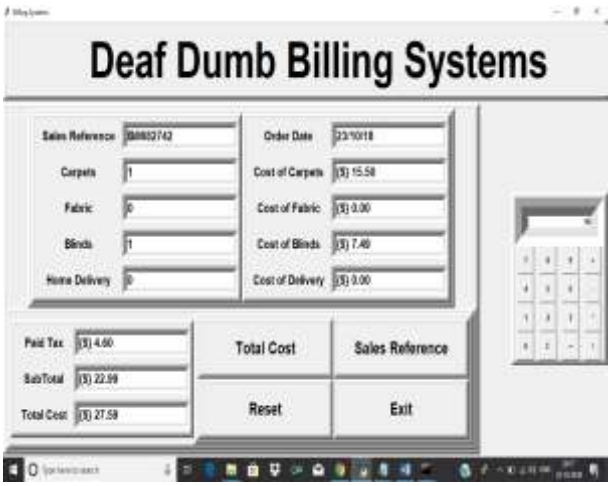

Fig 6.4 Generated Bill

VII. CONCLUSIONS

As we have used Python Tkinter as a GUI it provides numerous controls like buttons, labels, and text boxes so on to create a user friendly application. Therefore this project is very helpful for able individuals and helps in increasing the economic life style by providing work opportunities. Because the calculations are performed mechanically the time taken for the bill generation of every client is greatly reduced and it minimizes the waiting time for the client. There is no go to recruit additional dedicated person or instrumentation to handle this application.

REFERENCES

[1]. Abdulla, W. H., Chow, D., & Sin, G. (2003, October). Cross-words reference template for DTW-based speech recognition systems. In TENCON 2003.Conference on Convergent Technologies for the Asia-Pacific Region (Vol. 4, pp. 1576-1579). IEEE.

[2]. Alfi.A, M. Basuony and S. Atawy, "Intelligent Arabic text to Arabic Sign Language Translation for Easy Deaf Communication",International Journal of Computer Applications, vol. 92, no. 8, pp. 22-29, 2014.

[3]. Butzberger, J., Murveit, H., Shriberg, E., & Price, P. (1992, February). Spontaneous speech effects in large vocabulary speech recognition applications.In Proceedings of the workshop on Speech and Natural Language (pp. 339-343).Association for Computational Linguistics.

[4]. Hon .H, A. Acero, X. Huang, J. Liu, and M. Plumpe,"Automatic Generation of Synthesis Units for TrainableText-to-Speech Systems".

[5]. Lakshmi Sahu and Avinash Dhole,"Hindi &Telugu text to-Speech Synthesis(TTS) and inter-language text Conversion",International Journal of Scientific and Research Publications, Volume 2, Issue 4, April 2012 1 ISSN2250-315.

[6]. Masatsune Tamura, Takashi Masuko, Keiichi Tokuda, Takao Kobayashi," Text-To-Speech Synthesis With Arbitrary Speaker's Voice From Average Voice", Interdisciplinary Graduate School of Science and Engineering, Euro speech 2001 Scandinavia.

[7]. Poonam.S.Shetake, s.A.Patil, p. M Jadhav,"Review Of Text To Speech Conversion Methods",Proceedings of 10th IRF International Conference, 01st June-2014, Pune, India, ISBN: 978-93-84209-23-0.

[8]. Reddy, D. R. (1966). Approach to computer speech recognition by direct analysis of the speech wave.The Journal of the Acoustical Society of America, 40(5), 1273-1273.

[9]. Suryawanshi .S.D, Mrs. R. R. Itkarkar, Mr. D. T.Mane, "High Quality Text to Speech Synthesizer using Phonetic Integration" , International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 3, Issue 2, February 2014.

[10]. Tapas Kumar Patra, Biplab Patra, Puspanjali Mohapatra," Text to Speech Conversion with Phonematic Concatenation",International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 2 Issue 5 (September 2012).