**TEXT TO SPEECH CONVERSION**

**SYNOPSIS REPORT**

*by*

**Alka Jaiswal (00704092021)**

**Anjali Rajesh Baldia (00804092021)**

**Anjali Gupta (00904092021)**

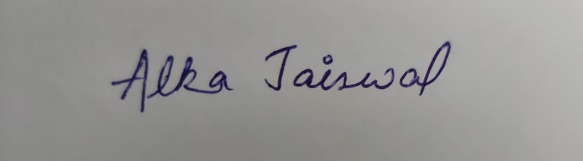
**Azra Parwez Ansari (01004092021)**

**Pooja Srivastava (03904092021)**



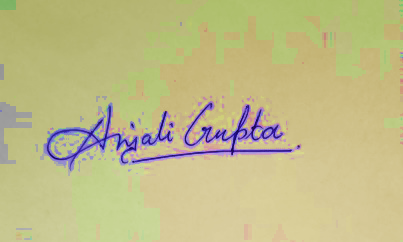
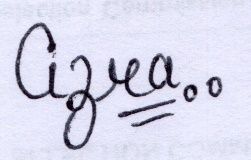
**STUDENT DECLARATION**

This is to declare that this report has been written by us. No part of the report is copied from other sources. All information included from other sources have been duly acknowledged. We are aware that if any part of the report is found to be copied, we shall take full responsibility for it.

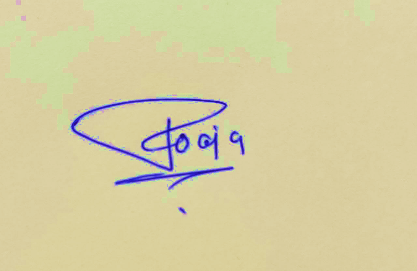
 Alka Jaiswal Anjali Rajesh Baldia

(00704092021) (00804092021)

Anjali Gupta Azra Parwez Ansari

(00904092021) (01004092021)



Pooja Srivastava

(03904092021)

Place: Delhi

Date: 23-11-2021

**TABLE OF CONTENTS**

**TITLE** **PAGE NO.**

1. **Background and objectives of the project assigned. 04**
   1. Background 04

1.1.1

* 1. Objective 04

1. **Description of Project 07**

2.1 Overall Description 07

2.1.1 Product Perspective 07

2.1.2 Product Feature and Developer Application 07

2.1.3 User Classes and Characteristics 07

2.1.4 The Operating Environment 07

2.1.5 Design and Implementation Constraints 08

2.2 System Features 08

2.2.1 Entered Text or Selected Text to Speech Conversion 08

2.3 External Interface Requirements 08

2.3.1 User Interfaces 08

2.3.2 Hardware Interfaces 08

2.3.3 Software Interfaces 08

2.4 Block Diagram 09

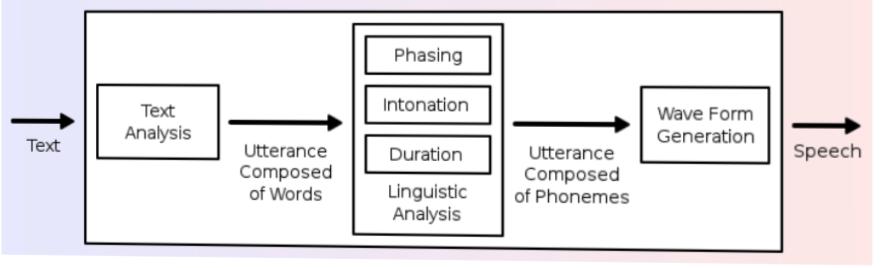
2.5 Flow Chart 10

2.6 Methodology of Work 10

1. **Description of Work Division in terms of Roles**
2. **Technologies and Framework to be used. 13**

4.1 Technology 13

4.2 Framework 13

1. **SWOT Analysis achieved in project.** **15**
2. **BACKGROUND AND OBJECTIVES OF THE PROJECT**
   1. **Background :**Text - to - speech conversion software project is windows-based application that reads a text file to the user. The software reads a text file and associated pronunciations in its Text - to - speech conversion software project is windows-based application that reads a text .The software reads a text file and associated pronunciations in its temporary database. The program then reads an entire word to the user. The software can be effectively used to help read the text document for the user so that the user does not constantly need to look at the screen and read the entire document.
      1. **Text to speech converter** is a recent software project that allows even the visually challenged to read and understand various documents. The blinds cannot read a document, so this software can be an assistant to them who would read out those documents for them. It can also be a great help for those who cannot speak. The person can simply type what he/she wants to say and the software would give a voice to them by speaking what they wanted to say. So, this software is not just an advancement towards the future development but also a boon for those who cannot speak and see. Temporary database. The program then reads an entire word to the user. The software can be effectively used to help read the text document for the user so that the user does not constantly need to look at the screen and read the entire document.
      2. **About language:** The language used for the project text – to speech conversion is python. Python is a high-Level, interpreted, interactive and object-oriented scripting language. Python is designed to Be highly readable. It uses English keywords frequently where as other languages use Punctuation, and it has fewer syntactical constructions than other languages. Python is Interpreted – Python is processed at runtime by the interpreter. You do not Need to compile your program before executing it. This is similar to PERL and PHP.

* 1. **Objective :** Among the many definition that could be given of text – to – speech, that Describes it as a way of having computer audibly communicate information to the user is probably the most relevant within the context of this statement. In Situations where visual feedback is inadequate or even impossible, audible feedback may be an essential feature; in many situations it may just add extra Value to a product. Generally, text – to – speech provides a very valuable and flexible alternative for digital – audio recordings where recordings are too expensive. Disk storage is insufficient to store the recordings. The application does not know ahead of time what it will need to Speak. The information varies too much to record and store all the alternatives.

**Project object:** The term “Text – to – Speech” or TTS for short, refers to the process by which Plain text is converted into digital audio and the Python IDLE - Python’s Integrated Development and Learning Environment For writing programs. IDLE has two main window types, the Shell window and the Editor window. It is possible to have multiple editor windows simultaneously. The shell is used for the construction of the project. Tkinter library – Tkinter is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit, and is Python's de facto standard GUI. Operating system used is Windows1Tkinter library – Tkinter is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit, and is Python's de facto standard GUI. Operating system used is Windows10.

**1.3 Motivation** :Among the many definition that could be given of text – to – speech, that describes it as a way of having computer audibly communicate information to the user is probably the most relevant within the context of this statement. In situations where visual feedback is inadequate or even impossible, audible feedback may be an essential feature; in many situations it may just add extra value to a product. Generally, text – to – speech provides a very valuable and flexible alternative for digital – audio recordings where : • Recordings are too expensive. • Disk storage is insufficient to store the recordings. • The application does not know ahead of time what it will need to speak. • The information varies too much to record and store all the alternatives.

**1.4 Outcomes & Goals:** This project may offer text-to-speech (TTS) solutions for websites, mobile apps, e-books, e-learning material, documents, conversational customer experience and transport experience systems, media, robotics, embedded devices, self-service applications, IoT, and more.

1. **DESCRIPTION OF PROJECT**

**2.1 Overall Description**

**2.1.1 Project Perspective:-**

Text – to – speech program that lets you type – in any English text and then plays it as an audio stream.

Instantly convert desired text to audio.

Converting files into speech format. File can be txt, pdf.

Supported language: English.

**2.1.2Project Feature and developer application:-**

Different implementations of text – to – speech system exist. This section discusses some of the concepts on which these systems are built. Generally, a text – to – speech system can be broken down into three parts: a linguistic, a phonetic and an acoustic part. First, an ordinary text is input to the system. A linguistic module converts this text into a phonetic representation. From this representation, the phonetic processing module calculates the speech parameters. Finally, an acoustic module uses these parameters to generate a synthetic speech signal*.*

**2.1.3 User Classes and Characteristics :-**

Using this project user can listen his entered text or selected text. He or she can listen given input file text which can be txt, pdf format.

User can listen the entered text in the interactive mode.

**2.1.4 The Operating Environment :-**

Software Requirements used are Windows XP and any other latest editions, Python Technologies.

Hardware Requirements used are P4processor, 512MB of Main Memory (RAM) and 40GB hard disk and base memory.

**2.1.5 Design and Implementation Constraints :-**

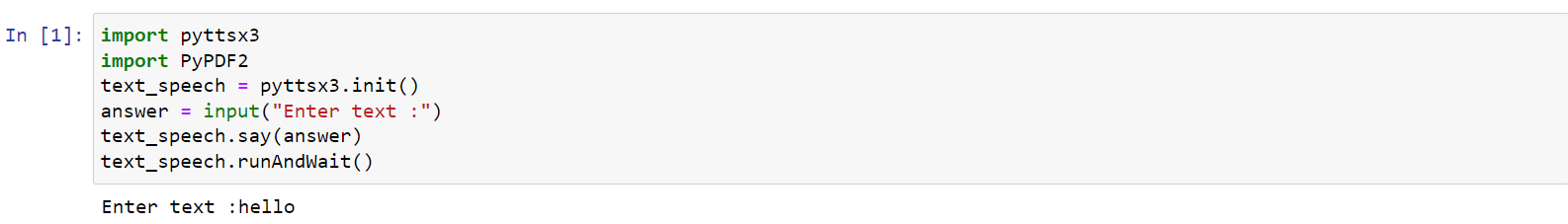
Design constraints developers. All modules are coded thoroughly based on requirements. The software is designed in such a way that the user can easily interact with the screen. Software is designed in such a way that it can be extended into real time business.

**2.2 System Features**

In this project, we have 2 modules:

**2.2.1 Entered Text or Selected Text to Speech conversion**

In this module user has to enter some text and it can listen the speech by clicking the Convert button present at the bottom. User can listen selected text or entered text. In this module we have to design GUI which provides text area to enter text. This Module opens up when we click on Interactive Mode in the main menu.



**2.2.2 Text File and PDF file conversion Module**

In this module, user can input text file as input for converting text into speech. In this module functionalities are:

* Getting path of input file
* Open the file
* Reading file
* Read text passed to speech module

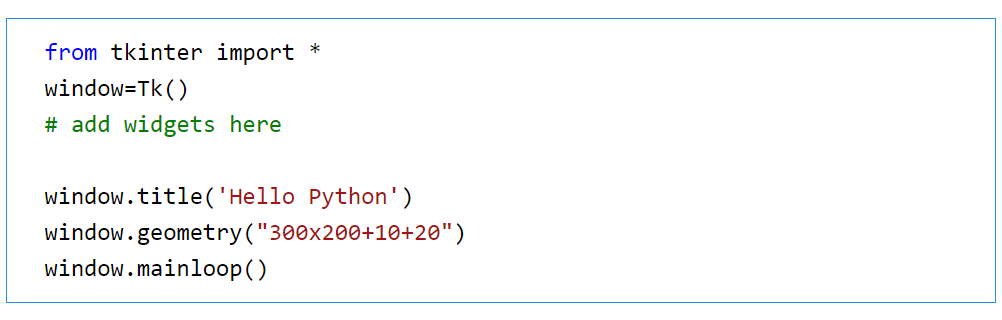
**2.6 Methodology of Work:-**

**Different libraries used in the project are:**

**Tkinter -** Tkinter is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit,and is Python's *de facto* standard GUI. Tkinter is included with standard Linux, Microsoft Windows and Mac OS X installs of Python. The name *Tkinter* comes from *Tk interface*. Tkinter was written by Fredrik Lundh. Tkinter is free software released under a Python license. As with most other modern Tk bindings, Tkinter is implemented as a Python wrapper around a complete Tcl interpreter embedded in the Python interpreter. Tkinter calls are translated into Tcl commands which are fed to this embedded interpreter, thus making it possible to mix Python and Tcl in a single application. Python 2.7 and Python 3.1 incorporate the "themed Tk" ("ttk") functionality of Tk 8.5. This allows Tk widgets to be easily themed to look like the native desktop environment in which the application is running, thereby addressing a long-standing criticism of Tk (and hence of Tkinter). There are several popular GUI library alternatives available, such as wxPython, PyQt (PySide), Pygame, Pyglet, and PyGTK

Creating a GUI application using Tkinter is an easy task. All you need to do is perform the following steps −

* Import the *Tkinter* module.
* Create the GUI application main window.
* Add one or more of the above-mentioned widgets to the GUI application.
* Enter the main event loop to take action against each event triggered by the user.

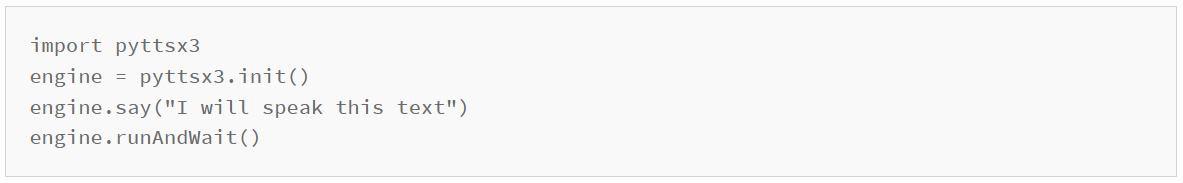


**Pyttsx3 –**Pyttsx is a good text to speech conversion library in python but it was written only in python2 until now ! Even some fair amount of googling didn’t help much to get tts library compatible with pyton3.

There is however , one library **gTTS** which works perfectly in python3 but it needs internet connection to work since it relies on google to get the audio data.ButPyttsx is completely offline and works seamlessly and has multiple tts-engine support. The codes in these repos are slightly modified version of the pyttsx module of python 2.x and is a clone from westonpace’s repo. The purpose of creating this repo is to help those who want to have an offline tts lib for Python3 and don’t want to port it from python2 to python3 themselves.

**Installation:**

**Usage :**



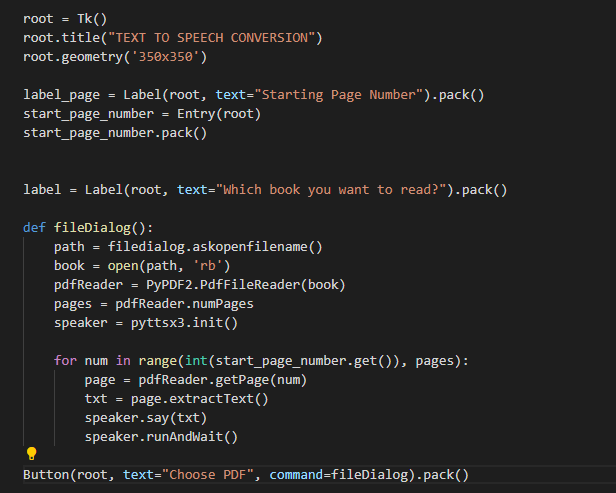
**PyPDF2:-** The **PyPDF2** package is a pure-Python PDF library that you can use for splitting, merging, cropping and transforming pages in your PDFs. According to the PyPDF2 website, you can also use PyPDF2 to add data, viewing options and passwords to the PDFs too. Finally, you can use PyPDF2 to extract text and metadata from your PDFs.PyPDF2 is actually a fork of the original pyPdf which was written by Mathiew Fenniak and released in 2005. However, the original pyPdf's last release was in 2014. A company called Phaseit, Inc spoke with Mathieu and ended up sponsoring PyPDF2 as a fork of pyPdf.

**Installation:**

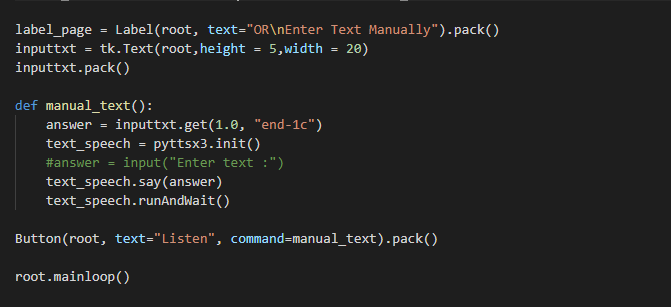
PyPDF2 is a pure Python package, so you can install it using pip (assuming pip is in your system's path):



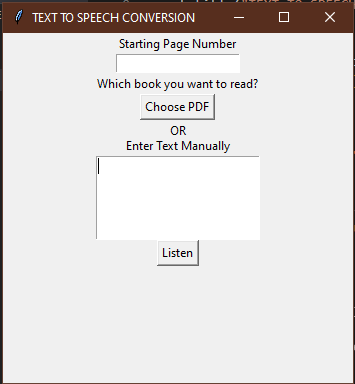
**Extracting Text of PDF Pages:**



**Extracting text entered manually:**



**OUTPUT SCREEN**



**2.4 BLOCK DIAGRAM**



**2.5 FLOW CHART**



1. **DESCRIPTION OF WORK DIVISION IN TERMS OF ROLES**

**ALKA JAISWAL** collected information from various links, research papers, PDF’s available on the internet. She has described the SWOT analysis of the project.

**ANJALI GUPTA** searched about the background and Objectives of the project . Also collected important information from different links ,PDF's available on the Internet .Such as which can be used the technologies available text and audio files that can be used , working of the models or projects .

**AZRA PARWEZ ANSARI** has acquired information about the technology and different modules used in our project like the best technology available and what modules to be implemented so that it can be effective for our project. She has also contributed to analysing the overall code.

**POOJA SRIVASTAVA** handled the description part of the project as in explain what all modules and library used in the project like Tkinter,Pyttsx3,PyPDF2,their installation & how they are working in our project. At the end she describes the whole project through a Flowchart and Block Diagram.

**ANJALI RAJESH BALDIA** has managed the coding part of the project. Also, the formatting, tracking & analysing the different parts of the project were done by her.

1. **TECHNOLOGIES AND FRAMEWORK TO BE USED**
   1. **Text-to-speech (TTS):** It is a technology used to assist in reading digital text in either a smartphone, computer, or tablet aloud and can be used everywhere, including home and school. It is commonly known as read-aloud technology. The technology is easy to use and only require basic knowledge. The technology works with just a click on the button where it captures the words on a digital device and converts the text to audio. Text-to-speech technology was first developed to help kids with disabilities, especially those who were struggling to read. Still, today it helps all kids in improving their editing and writing ability and also increases their general focus while reading. Today, the technology is used by millions of people, including the big companies, freelancers to help them read the documents.

The technology read-aloud all kinds of documents, including the online web pages, pages document and word. So, you don’t have to worry about the nature of your file, it is sorted with TTS. TTS technology is very efficient, accurate and uses computer-generated realistic voices that can sound like adults or children. This is why it is valuable for everyone. The reading can be controlled depending on your speed. Text-to-Speech technology has different types and the one you choose to use should depend on your digital device. The models include the web-based tools, chrome tools, Built-in-to-text-speech, Text-to-speech software programs and Text-to-speech apps. Each of these tools works differently but in almost all the digital devices.

* 1. **Python:** Python is a very popular general-purpose programming language which was created by Guido van Rossum, and released in 1991. It is open-source and you can freely use & distribute Python, even for commercial use. It is very popular for web development and you can build almost anything like mobile apps, web apps, tools, data analytics, machine learning etc. It is designed to be simple and easy like english language. It's much easier to read and write Python programs compared to other languages like C++, Java, C#. It's is highly productive and efficient which makes it a very popular programming language.

## Features of this tool

* Simple & Minimalist Design, Lightweight, Easy & Fast
* Supported Python version - Python3.8
* Interactive program execution which makes the user to give program inputs at real time
* Helpful for beginners to learn and practice Python
* Dark & Light theme options and customizable code editor with more themes
* Undo and Redo Options in Code Editor

**4.3 Jupyter lab:** The Jupyter Notebook is an incredibly powerful tool for interactively developing and presenting data science projects. First, thought: **what is a “notebook”?**

A notebook integrates code and its output into a single document that combines visualizations, narrative text, mathematical equations, and other rich media. In other words: it's a single document where you can run code, display the output, and also add explanations, formulas, charts, and make your work more transparent, understandable, repeatable, and shareable.

Using Notebooks is now a major part of the data science workflow at companies across the globe. If your goal is to work with data, using a Notebook will speed up your workflow and make it easier to communicate and share your results.

Best of all, as part of the open source [Project Jupyter](https://jupyter.org/), Jupyter Notebooks are completely free. You can download the software [on its own](https://jupyter.org/install), or as part of the [Anaconda data science toolkit](https://www.anaconda.com/products/individual).

1. **SWOT ANALYSIS ACHIEVED IN PROJECT**



|  |  |
| --- | --- |
| **Strength**  Speech-enabling service minimizes human agent workload, provides personalized services, accelerates throughput, and reduces operational costs.  A single TTS voice across multiple contact points supports consistent, emotional branding.  With TTS technology that is web- or cloud-based on a SaaS (Software as a Service) platform, online content can quickly and easily be speech enabled, and maintenance is minimal. | **Weakness**  It is clearly a synthetic voice or conversational AI without emotion or intonation; this is not a narrative or a show  The written text is emotionless and often can't pronounce the "name" correctly. In producing female or young children's voice characters, text to speech often has problems because women's voices have a tone that is almost twice as high as men's voices, and in the case of children it is even up to three times higher |
| **Opportunities**  Text to speech technology market is witnessing increasing revenue growth owing to growing traction of the software globally  It can further improve digital health technology in healthcare by speech-enabling websites, health apps, nurse call systems in hospitals, and portable health devices synthetic voices that are very humanlike. | **Threats**  It is difficult to build a perfect system.  The flat voice and the limitation to convey emotion is the main reason why using text to speech for video applications, video game dialogues, or audiobooks is not the best choice. When performance is needed or emotions need to be conveyed, the human narrator is an irreplaceable choice. |