



JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AI-Notes

¹Yashfeen Zari, ²Noah Sayed, ³Saad Evara, ⁴Wael Shaikh, ⁵Shaikh Mohd. Ashfaque

⁵Assistant Professor

¹⁻⁴Students, ¹⁻⁵Department of Computer Science Engineering,

¹Rizvi College of Engineering Mumbai, India.

Abstract: Artificial intelligence (AI) is a science that involves simulation of intelligent behaviors in machineries, like visual perception, decision making, speech recognition and so on. While the rate of progress in AI has been patchy and unpredictable, there have been significant advances. In this paper we discuss about one of the upcoming fields in artificial intelligence which is automatic speech recognition from neural signals. We aim to provide a method to quickly take notes, and securely store them all at one platform and help to scan your notes, textbooks, study material and convert it into digital text via OCR, for your ease, we also created a system which gives an audio output when asked any question pertaining to the notes, which helps in easy understanding. The user can insert their own inputs, which may be an image of digital text or a photo of text on a piece of paper and ask questions based on that text via Natural Language Processing.

Speech driven services like Siri and Google voice search are used by millions of people in their everyday lives. These speech interfaces allow natural interaction with electronic devices and enable faster communication. Brain-computer interface (BCI), also called as mind machine interface (MMI), direct neural interface (DNI) is a collaboration between a brain and an external component or device that facilitates brain signals to manipulate or control some activity of that device. This provides direct communication pathway between the brain and the component to be controlled. BCI's usage is limited. This is partly due to the unnatural paradigms that are used to give commands and texts to the BCI's like motion of hands etc. This is slow and unnatural. On the other hand, if speech was used as a paradigm, then it would be faster and more natural. But only the concept of thought to speech would enable severely disabled persons (i.e., locked-in syndrome, Lou Gehrig's syndrome) to communicate with the outside world.

Following are the languages used in this project: - HTML, CSS, JavaScript

Key Words: AI, NLP, Notes, OCR

1. INTRODUCTION

Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal. A subset of artificial intelligence is machine learning, which refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans. Deep learning techniques enable this automatic learning through the absorption of huge amounts of unstructured data such as text, images, or video.

Digital note taking is a combination of techniques that allow you to take and store your notes electronically. Most think of digital note taking as simply typing notes, but with digital handwriting apps and software, you can also handwrite, store, and share your digital notes.

By definition, note taking is nothing more than recording information captured from another source. While this can be done on paper, there are benefits to having access to your notes online including better organization, increased security, more flexibility and improved shareability.

2. RESEARCH METHODOLOGY

2.1 Surveying Existing System

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech computer or speech synthesizer, and can be implemented in software or hardware products. A text-to-speech (TTS) system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions into speech. The reverse process is speech recognition. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored speech units; a system that stores phones or diphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output. Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output.

The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood clearly. An intelligible text-to-speech program allows people with visual impairments or reading disabilities to listen to written words on a home computer. Many computer operating systems have included speech synthesizers since the early 1990s.

The Web Speech API enables you to incorporate voice data into web apps. The Web Speech API has two parts: SpeechSynthesis (Text-to-Speech), and SpeechRecognition (Asynchronous Speech Recognition.)

Text-to-speech (TTS) refers to the ability of computers to read text aloud. A TTS engine converts written text to a phonemic representation, then converts the phonemic representation to waveforms that can be output as sound. TTS engines with different languages, dialects and specialized vocabularies are available through third-party publishers.

Optical character recognition or optical character reader (OCR) is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example: from a television broadcast).

2.2 Main body

By studying the Survey on the Existing System, it was found that there is a need among the students to access the notes easily anywhere in the world from college /school or from home so you don't have to go for any other books and note taking apps.

Digital note taking is a combination of techniques that allow you to take and store your notes electronically. Most think of digital note taking as simply typing notes, but with digital handwriting apps and software, you can also handwrite, store, and share your digital notes.

By definition, note taking is nothing more than recording information captured from another source. While this can be done on paper, there are benefits to having access to your notes online including better organization, increased security, more flexibility and improved shareability.

2.3 Objective

The objective of the project is to quickly and securely scan and upload notes, textbooks, study material, etc., to be able to access all notes, study materials, etc., in an organized manner in a central place, to infer and understand the data via natural language processing and question answering, to quickly find answers and solutions from their text.

3. PROPOSED SYSTEM

It was found that there is a need among the students to access the notes easily anywhere in the world from college /school or from home so you don't have to go for any other books and note taking apps. We propose a system that enables users to quickly and securely scan and upload notes, textbooks, study material, etc., to be able to access all notes, study materials, etc., in an organized manner in a central place, to infer and understand the data via natural language processing and question answering, to quickly find answers and solutions from their text. The notes submitted by the user to the system, get saved to local storage and are fetched whenever the user revisits.

4. Methodology

We provide a home page that greets the user with stress relieving memory game, along with an inspiring quote generator.



Fig 4.1 Home Page

The quote generator fetches quotes via an API.

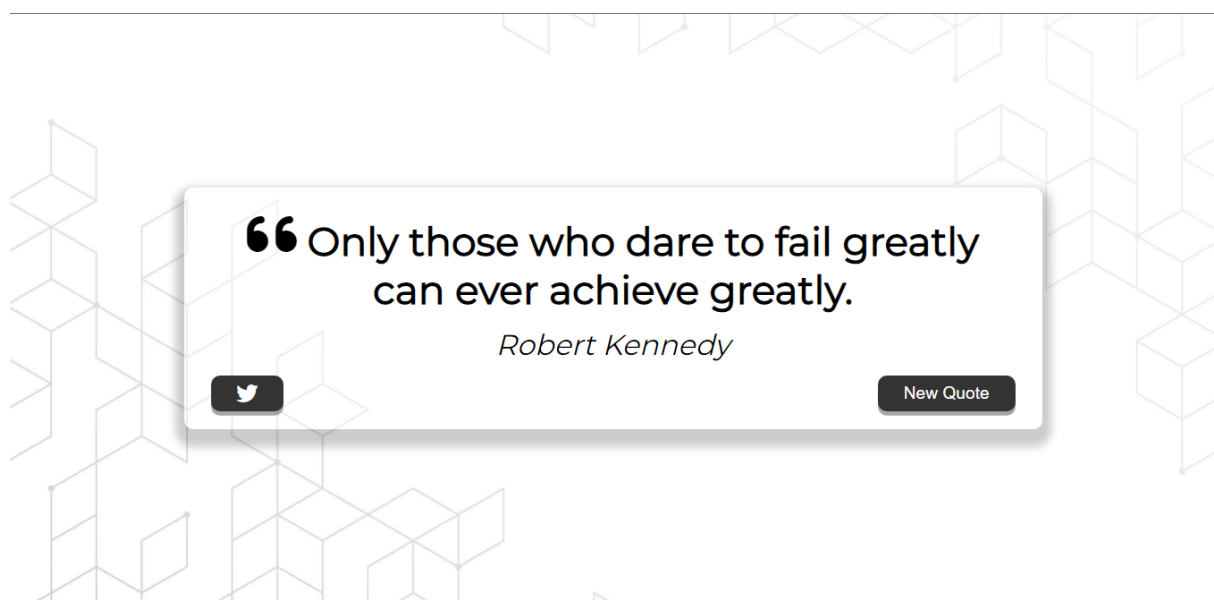


Fig 4.2 Quote Generator

The memory game tests the user's memory by having the user keep track of the game's state as the game gets progressively harder each turn.

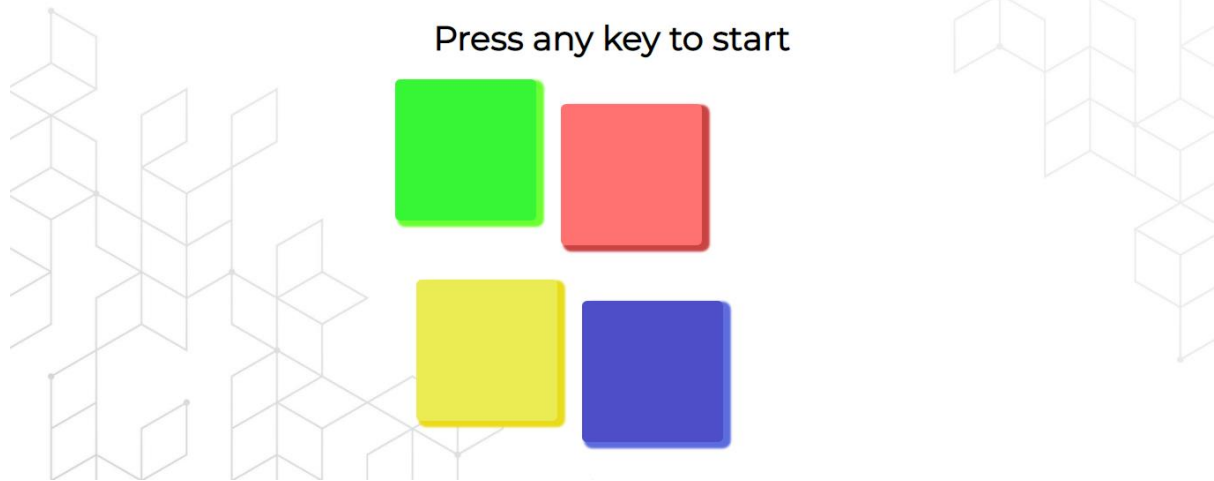


Fig 4.3 Memory Game

The notes page allows the user to keep track of various notes, topics, tasks, etc. It allows the user to move tasks/notes between New, In Progress, Complete and On Hold states.

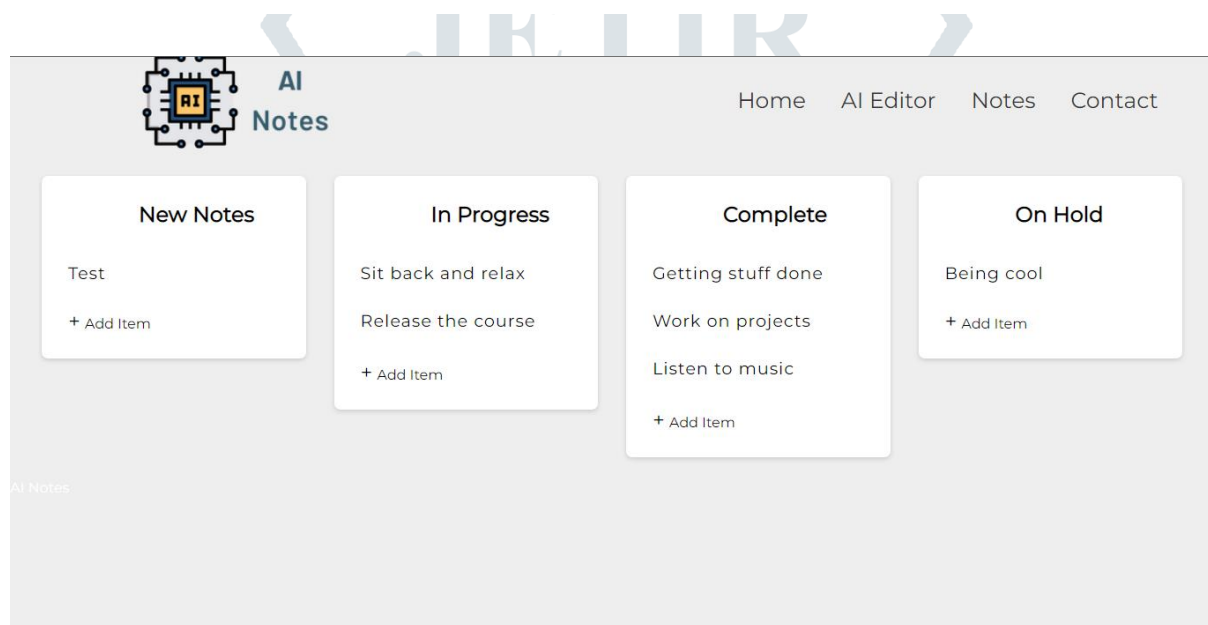


Fig 4.4 Notes Page

The editor page allows user to upload their notes via typing, copy-pasting or via OCR, which can extract text from photos of textbooks, screenshots of pdf/websites, etc. Whatever work the user does in the editor gets saved to the local storage and is retrieved when the user revisits, enabling the user to continue where they left off.

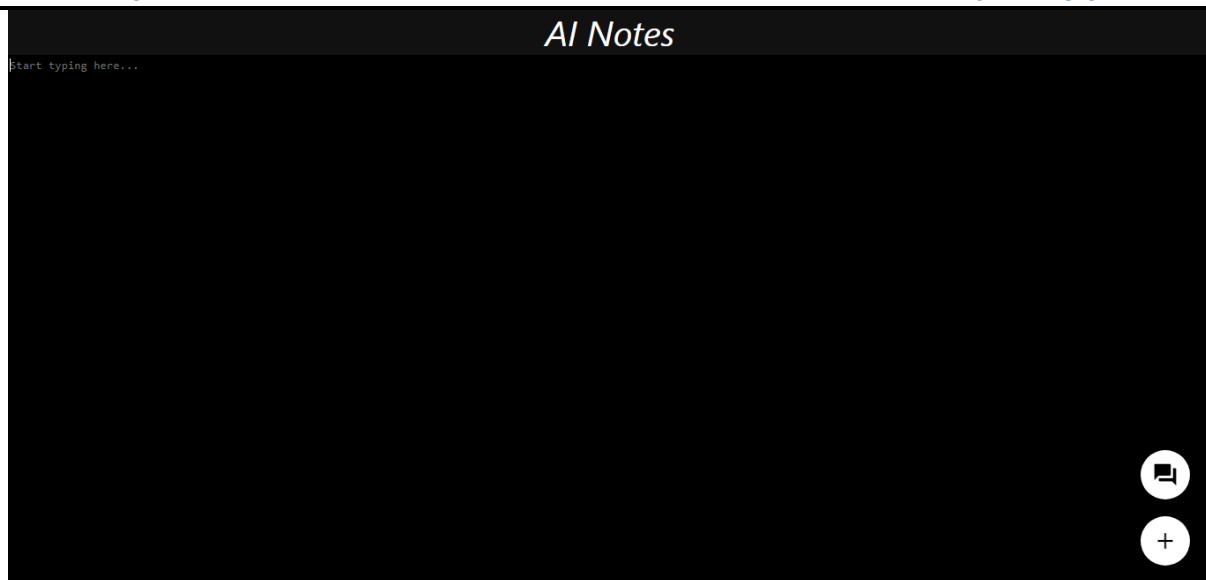


Fig 4.5 Editor Page

Once the user has uploaded their notes to the editor, they can then get answers and insights from the notes by querying the NLP Question Answering AI. The user simply has to provide a question in natural language and the AI will infer and understand the given data via NLP and provide an answer to the user's query by outputting it to the editor as well as by speaking it out loud using a TTS (Text to Speech) Engine.

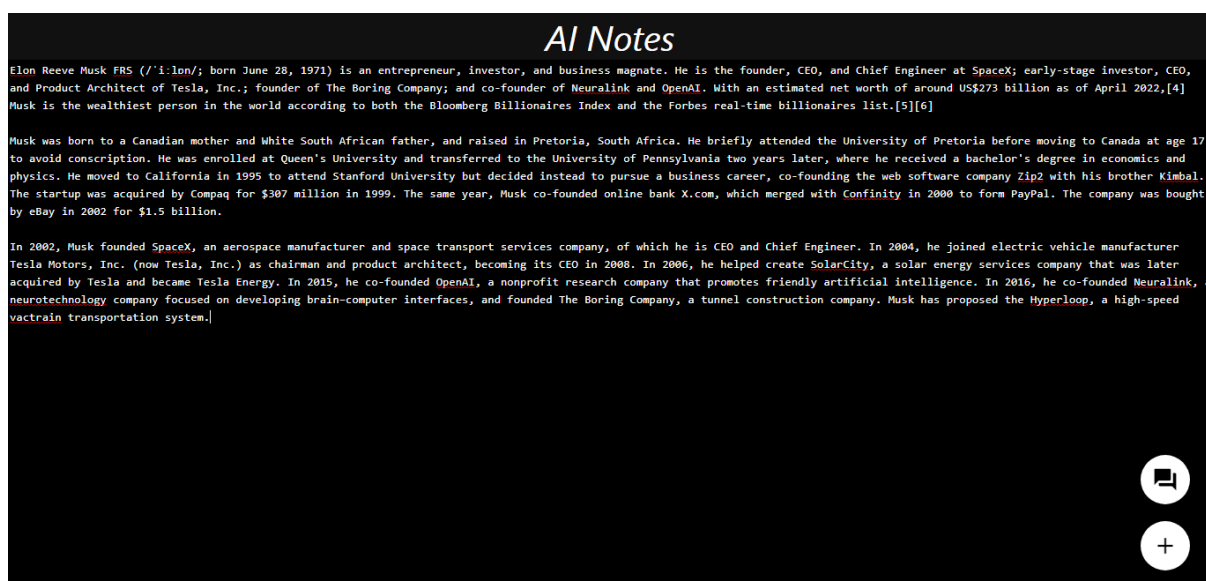


Fig 4.6 Editor Page with OCR scanned text

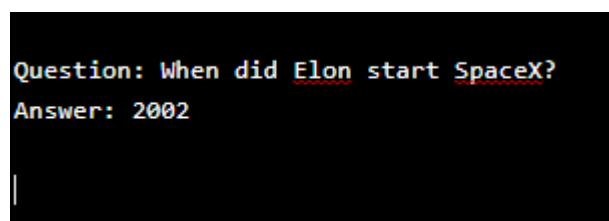


Fig 4.7 Question answering via AI

5. Conclusion

This project helps you in quickly taking your notes, and securely storing them all at one platform and helps to scan your notes, textbooks, study material and converts it in a digital text via OCR, for your ease, and gives an audio output when asked any question pertaining to the notes, which helps in easy understanding. The user can insert their own inputs, which may be an image of digital text or a photo of text on a piece of paper and ask questions based on that text via Natural Language Processing.

6. References

AI:

<https://blog.tensorflow.org/2018/04/a-gentle-introduction-to-tensorflowjs.html>

<https://github.com/tensorflow/tfjs>

<https://www.google.com/amp/s/www.geeksforgeeks.org/tensorflow-js-introduction/amp/>

TensorFlow.js: Machine Learning for the Web and Beyond

[<https://arxiv.org/abs/1901.05350>]

Speech Synthesis:

https://en.m.wikipedia.org/wiki/Speech_synthesis

https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API

OCR:

https://en.m.wikipedia.org/wiki/Optical_character_recognition

