Custom Speech Service using Natural Language Processing in Microsoft Azure

Presented by:
Lakkireddy Indhu

Guided by:
Mrs. Loyola Jasmine
Assistant Professor

Introduction

- Natural Language Processing (NLP) is the branch of artificial intelligence that deals with the interaction between computers and human language. It involves processing, understanding, and generating natural language text or speech. Azure, Microsoft's cloud computing platform, offers various tools and services for NLP.
- Azure Cognitive Services provides pre-built APIs for NLP tasks such as language detection, sentiment analysis, entity recognition, and key phrase extraction. These APIs can be easily integrated into applications to perform NLP tasks without the need for extensive coding or machine learning expertise.

Software Tool

Azure is a cloud computing platform and a suite of software services offered by Microsoft. It enables organizations to build, deploy, and manage applications and services in a secure and scalable environment.

Some of the key features of Azure

- Scalability: Azure allows you to easily scale your applications and services up or down to meet changing demands.
- Security: Azure provides a secure and compliant environment with built-in security features such as threat detection and data encryption.

Software Tool

- Flexibility: Azure supports a wide range of programming languages, frameworks, and tools, making it easy to integrate with your existing systems and workflows.

- Devops: Azure provides a complete Devops solution with tools for continuous integration and deployment, testing, and monitoring.

Software Tool used

- Azure Machine Learning is a cloud-based platform for building, training, and deploying machine learning models. It offers various algorithms and tools for NLP, such as text classification, language translation, and text generation. It also provides integration with other Azure services such as Cognitive Services and Azure Databricks.
- Azure Bot Service is a platform for building and deploying chatbots using natural language processing. It offers various tools for creating conversational experiences, including Language Understanding Intelligence Service (LUIS) for natural language understanding, and QnA Maker for creating and managing a knowledge base for frequently asked questions.

Reported Literature

- Custom speech recognition system for hearing-impaired individuals using deep learning techniques" by S. M. A. Bhuiyan et al. This paper proposes a custom speech recognition system using deep learning techniques for hearing-impaired individuals, and it demonstrates improved accuracy and performance over traditional speech recognition systems.
- "Customization and adaptation of automatic speech recognition for dialects and accents" by S. Saha et al. This paper describes a framework for customizing and adapting automatic speech recognition systems for dialects and accents, and it presents experimental results showing improved performance and accuracy.
- "End-to-end custom speech recognition using neural networks" by X. Liu et al. This paper proposes an end-to-end custom speech recognition system using neural networks, and it demonstrates improved accuracy and performance over traditional speech recognition systems.

Objective of project

- The main objective of CUSTOM SPEECH SERVICE in Azure is to provide organizations and individuals with a highly accurate and personalized speech recognition solution that can be tailored to their specific needs and requirements.
- Custom Speech Service in Azure is to provide organizations and individuals with a powerful and flexible speech recognition solution that can help them improve communication, increase efficiency, and gain a competitive advantage in their industry.

Timeline Proposal of a Project

- The circular of the Microproject was passed by 15-12-2023.
- Then we approval for the financial aid for the coursera by 19-12-2022.
- The financial is approved by the 29-01-2023.
- Then I start writing the assignments in the coursera
- I got the certificate on 13-01-2023.
- Then I started working project on 15-01-2023.
- Then I collected the data from the github.
- And working on the datasets.
- Then I test and train the data.
- Lastly I deployed my project on 21-02-2023.

Algorithm used

Azure Custom Speech Service uses a combination of machine learning algorithms to convert audio to text, including:

Recurrent neural networks (RNNs): RNNs are used for speech recognition, which involves decoding the audio features into words. The RNNs use a technique called sequence-to-sequence modeling to predict the sequence of words that corresponds to the audio features.

Work done in step by step description

- **Sign up for Azure:** The first step is to sign up for Azure if you don't have an account already. You can create an account on the Azure portal.
- **Create a Custom Speech resource:** Once you have an Azure account, you need to create a Custom Speech resource. You can do this by going to the Azure portal and selecting "Create a resource", then searching for "Custom Speech".
- **Configure your custom speech model:** Once you have created a Custom Speech resource, you can configure your custom speech model. This involves defining the language, vocabulary, recognition mode, and other parameters for your model.

Work done in step by step description

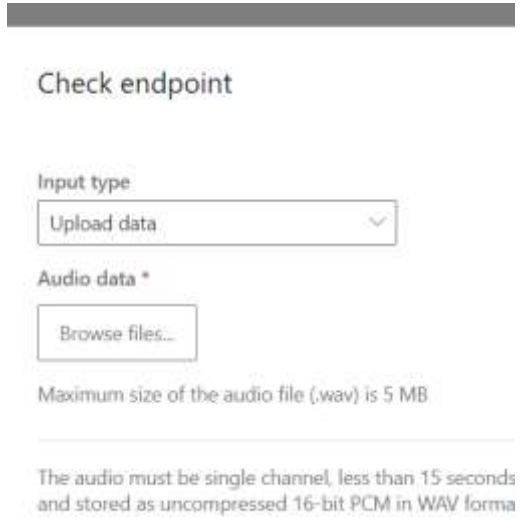
- **Upload data and train your model:** After configuring my custom speech model, you need to upload data and train your model. I upload audio data and corresponding transcripts to the Custom Speech Portal, and then train the model using the portal's training tools.
- **Evaluate and refine my model:** Once my model is trained, I need to evaluate its performance and refine it as necessary. I can use testing tools provided in the Custom Speech Portal to evaluate your model's performance and fine-tune the model as necessary.
- **As the clouds change every few months, there will be changes in the portals. In the recent times there is no need to upload and train the model using dataset, they are preloaded.**

Work done in step by step description

- **Integrate my model with your application or service:** After you have evaluated and refined your custom speech model, you can integrate it with your application or service. This may involve using Azure's Speech SDK to enable speech recognition in your application.
- **Monitor and maintain my model:** Finally, it is important to monitor and maintain your custom speech model over time. This may involve tracking performance metrics, making adjustments as necessary, and retraining the model periodically to ensure continued accuracy and performance.

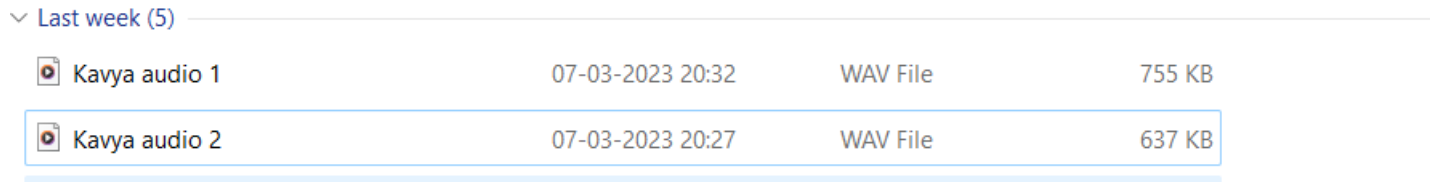
Results and Discussion

Here I Deploy my project.



Results and Discussion

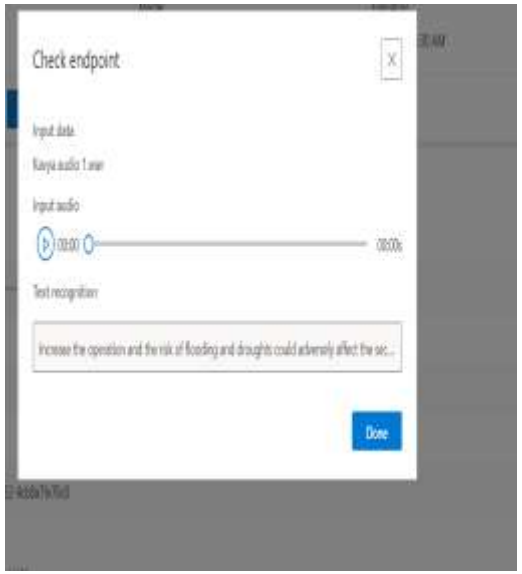
By the browse file I uploaded my own audio file.



Results and Discussion

Then we get the output like this.

It recognition the audio and converting in into text



Summary

Custom Speech Service in Azure offers a powerful solution for businesses and developers seeking to build accurate and customized speech recognition capabilities into their applications and services. By providing access to powerful machine learning tools and cloud-based computing resources, Custom Speech Service enables users to create highly accurate speech recognition models that are customized to their specific needs and use cases.

References

Speech Recognition with Microsoft Azure" - Microsoft Developer Network: This article provides an overview of speech recognition with Azure, including a discussion of the different types of speech recognition offered by Azure.

Speech to Text – Audio to Text Translation | Microsoft Azure

"Building a Custom Speech Recognition Model in Azure" - DZone: This article provides a step-by-step guide to building a custom speech recognition model in Azure, including information on how to collect and label training data, and how to use Azure's Custom Speech Portal to train and evaluate the model.

What is the Speech service? - Azure Cognitive Services

Thank you