Dissertation Proposal

1. Project Summary

The project aims to enhance the accuracy of automatic speech recognition (ASR) systems through the application of deep learning techniques. ASR is an essential component of various applications, including voice assistants, transcription services, and speech-to-text systems. Despite advancements, ASR systems still face challenges in accurately recognizing speech due to factors like background noise, accents, and variations in speech patterns. This research proposes the utilization of deep learning techniques, such as convolution neural networks (CNNs) and recurrent neural networks (RNNs), to improve ASR accuracy. By training these models on large speech datasets, the project aims to develop a robust ASR system capable of accurately transcribing speech across diverse conditions and contexts

https://theaisummer.com/speech-recognition/

2. Research Area

The research area focuses on automatic speech recognition (ASR) and deep learning techniques. It explores the application of CNNs and RNNs in ASR systems, investigating their effectiveness in improving recognition accuracy. Relevant research papers, such as "Listen, Attend and Spell" by William Chan et al. and "Deep Speech 2: End-to-End Speech Recognition in English and Mandarin" by Dario Amodei et al., provide valuable insights into the field of ASR and deep learning techniques.

https://ieeexplore.ieee.org/abstract/document/7472621/
http://proceedings.mlr.press/v48/amodei16.html?ref=https://codemonkey.link

3. Research Questions

- How can deep learning techniques, specifically CNNs and RNNs, be applied to enhance automatic speech recognition accuracy?
- What are the key factors affecting ASR accuracy, and how can deep learning models address these challenges?
- How does the performance of the proposed deep learning-based ASR system compare to traditional ASR approaches?

4. Aim and Objectives:

Aim: The aim of this research is to enhance the accuracy of automatic speech recognition systems using deep learning techniques.

Objectives:

- Investigate the current state-of-the-art in automatic speech recognition and deep learning techniques.
- Design and implement a deep learning-based ASR system using CNNs and RNNs.
- Train and optimize the deep learning model on large speech datasets to improve recognition accuracy.
- Evaluate and compare the performance of the proposed ASR system with traditional ASR approaches.
- Identify potential limitations and challenges in implementing deep learning-based ASR systems.

5. Expected Practical Element Output:

The expected practical element output includes:

- Development of a deep learning-based ASR system using CNNs and RNNs.
- Collection and preprocessing of a large speech dataset for training and evaluation.
- Training the ASR model on the dataset and optimizing its performance.
- Evaluation of the ASR system's accuracy and performance using appropriate metrics.
- Comparison of the proposed deep learning-based ASR system with existing ASR approaches.
- Documentation of the implemented system, experimental results, and analysis of findings.

6. Required Resources

- Access to a suitable computing environment with GPUs for training deep learning models.
- Large speech dataset for training and evaluation.
- Programming tools and libraries for deep learning, such as TensorFlow or PyTorch.
- Text-to-speech and speech recognition APIs for data acquisition and system evaluation.

https://www.kaggle.com/datasets/mozillaorg/common-voice
https://asr.science.energy.gov/about/resources-for-newly-funded-investigators

7. Prerequisite Knowledge/ Skills Required:

- Strong understanding of deep learning concepts and techniques, particularly CNNs and RNNs.
- Proficiency in programming, preferably in Python, and experience with deep learning frameworks.

- Familiarity with speech recognition algorithms and techniques.
- Knowledge of data preprocessing and feature extraction methods for speech signals.
- Understanding of evaluation metrics for assessing ASR system performance.

https://realpython.com/python-speech-recognition/

8. Project Plan

- Literature review and research methodology development
- Data acquisition and preprocessing
- Design and implementation of deep learning-based ASR system
- Training and optimization of the ASR model
- Evaluation and analysis of the ASR system's performance
- Comparison with existing ASR approaches
- Documentation and report writing

1. Literature review and research methodology development:

   During this phase, I will conduct an extensive literature review to understand the current state-of-the-art in automatic speech recognition (ASR) and deep learning techniques. I will gather relevant research papers, analyze their methodologies, and identify gaps in the existing approaches. This phase also involves developing a research methodology that outlines the steps and techniques I will employ to achieve the objectives of my study.

   Timeline: 20th June 2023 - 3rd July 2023

2. Data acquisition and preprocessing:

   In this stage, I will acquire a suitable speech dataset for training and evaluation purposes. I may collect pre-existing speech datasets or record my own environmental speech data. The acquired dataset will undergo preprocessing, including cleaning the data, removing any artifacts or silences, and dividing longer audio clips into smaller segments for efficient processing. I will also perform necessary conversions and normalization to ensure consistency in the dataset.

   Timeline: 4th July 2023 - 17th July 2023

3. Design and implementation of deep learning-based ASR system:

   During this phase, I will design and implement a deep learning-based ASR system using convolution neural networks (CNNs) and recurrent neural networks (RNNs). I will define the network architecture, including the layers, connections, and activation functions. The implementation will involve using deep learning frameworks such as TensorFlow or PyTorch to build and train the ASR model on the preprocessed speech dataset.

Timeline: 18th July 2023 - 7th August 2023

3. Training and optimization of the ASR model:

   In this stage, I will train the deep learning-based ASR model on the prepared speech dataset. Training will involve adjusting the model's parameters and optimizing its performance to maximize recognition accuracy. I will experiment with different hyperparameters, learning rates, and regularization techniques to fine-tune the model. This phase may require multiple iterations to achieve the desired performance.

   Timeline: 8th August 2023 - 21st August 2023

4. Evaluation and analysis of the ASR system's performance:

   After training the ASR model, I will evaluate its performance using appropriate evaluation metrics. I will test the model on a separate evaluation dataset to assess its accuracy and effectiveness in recognizing speech. The analysis will involve comparing the model's performance with traditional ASR approaches and identifying strengths, weaknesses, and potential areas of improvement.

   Timeline: 22nd August 2023 - 4th September 2023

5. Comparison with existing ASR approaches:

   During this phase, I will compare my deep learning-based ASR system with existing ASR approaches. I will review and analyze relevant research papers and commercial ASR systems to assess the advantages and limitations of my proposed approach. The comparison will provide insights into the effectiveness and novelty of my system.

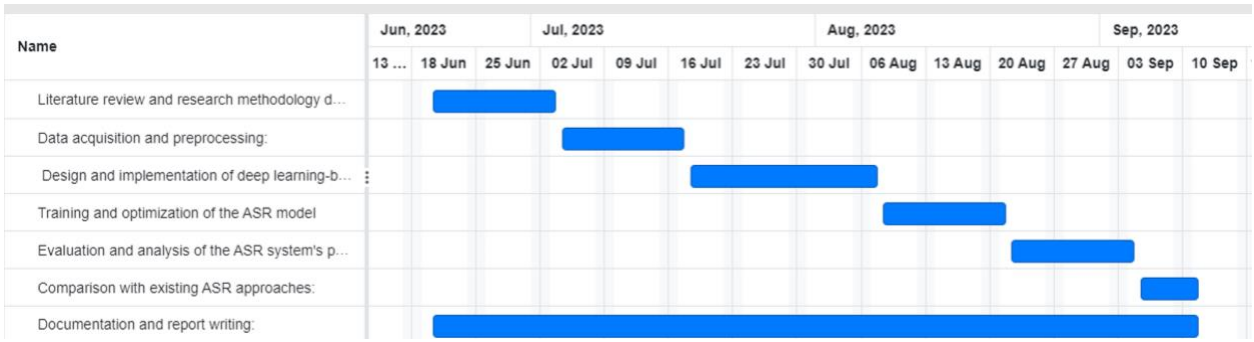
   Timeline: 5th September 2023 - 11th September 2023

6. Documentation and report writing:

   Throughout the project, I will document the process, methodologies, experimental results, and analysis. This phase involves compiling all the information into a comprehensive report. The report will include an introduction, literature review, methodology description, results and analysis, conclusion, and future work. I will ensure proper referencing and formatting to produce a high-quality dissertation.

   Timeline: 11th September 2023.

9. Gantt Chart



It is important to note that the project plan and Gantt chart provided are tentative and subject to change based on the actual progress and specific requirements of the research. Regular monitoring and adjustments will be made to ensure the successful completion of the dissertation within the given timeframe.