Task
Speech Recognition Model Development Using the SpeechData Dataset

Objective:
The objective of this task is to develop an automatic speech recognition (ASR) model using the SpeechData dataset (https://github.com/isaacwiafe/speech_data_ghana_ug). You need to pre-process the dataset, develop, and train a deep learning model for speech recognition in either Akan, Ewe, Dagbani, Dagaaare or Ikposo, and evaluate its performance using appropriate metrics. This task aims to assess your understanding of ASR techniques and your ability to apply them to real-world datasets.

Task Description:
Dataset Description:
- The SpeechData dataset consists of audio recordings of spoken sentences in five Ghanaian languages and accents.
- The audio data was collected by our research team specifically for this task
- There are 1000 hours of audio in each language and it is accompanied by 100 hours of transcript for a selection of the audios.
- The dataset is available in a structured format, with separate folders for audio files and their corresponding transcripts.

Pre-processing:
- Divide the dataset into training, validation, and testing sets.
- Perform audio pre-processing techniques such as normalization, feature extraction (e.g., MFCC), and data augmentation (if necessary).
- Convert text transcripts into numerical representations suitable for training the ASR model.

Model Development:
- Design a deep learning architecture for speech recognition using frameworks such as TensorFlow or PyTorch.
- Experiment with different architectures, including but not limited to Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Transformer models.
- Implement techniques such as attention mechanisms to improve model performance.

Training:
- Train the ASR model using the training set.
- Tune hyperparameters such as learning rate, batch size, and regularization techniques to optimize model performance.
- Monitor training progress and visualize training/validation loss and accuracy over epochs.

Evaluation:
- Evaluate the trained model on the validation set using appropriate evaluation metrics such as Word Error Rate (WER), Character Error Rate (CER), and Accuracy.
- Fine-tune the model based on validation performance if necessary.

Testing:
- Evaluate the final trained model on the testing set to assess its generalization performance.
- Calculate and report the final performance metrics on the testing set.

Documentation and Reporting:
- Prepare a report documenting the entire process of dataset pre-processing, model development, training, evaluation, and testing.

- Include descriptions of the chosen model architecture, hyperparameters, training methodology, evaluation metrics, and results analysis.
- Present visualizations (e.g., loss curves, confusion matrices) to aid in understanding the model's behavior.
- Discuss insights gained from the project, challenges encountered, and potential avenues for future improvements.

Submission Guidelines:
- Submit your report along with the source code (Python scripts or Jupyter notebooks) used for dataset pre-processing, model development, training, and evaluation to Github and provide the link to: https://link.webropolsurveys.com/S/E45355851898104A
- Ensure that the code is well-commented and organized for clarity and reproducibility.
- Submit by April 12, 2024.