



Internship Assignment for NLP(Voice AI)

Creation of for Speech to Text Model



Objective

The primary objective of this internship assignment is to evaluate your proficiency in natural language processing by developing a highly accurate transcription model for the Marathi language. This entails creating a robust model that can accurately transcribe spoken Marathi content. You will be provided with a test dataset, and you are encouraged to use open source datasets to enhance the model's performance and training.

Data

Please download the assignment folder from [here](#).

Audio files and corresponding transcriptions are stored in a designated folder.

- The provided folder contains audio files and a text file that includes the names of the audio files along with their respective transcriptions.

Output

- 1. Metric Selection:** Carefully choose and justify the use of appropriate metrics that optimise the model's performance.
- 2. Model Development:** Implement a state-of-the-art natural language processing model for transcription. You can leverage any existing architecture.
- 3. Code Implementation:** Provide clear and well-commented code that demonstrates the implementation of the transcription model. Ensure that your code is organised, efficient, and easy to follow.
- 4. Comprehensive Documentation:** Accompany your code with comprehensive documentation that guides readers through the transcription model implementation process. Include detailed explanations of the model selection, data preprocessing techniques, training methodology, and hyperparameter tuning. Provide instructions for replicating your results.
- 5. Testing and Evaluation:** Evaluate the model's performance using the designated test dataset. Report the results using the chosen metrics, and discuss the model's strengths and limitations.



Note

You are permitted to utilise any open-source data for training purposes, excluding the data we have provided.

Submissions

Code format

The code submission shall be a .py script that contains the following function.

```
def get_transcription(filename: str):  
    transcription = model_output  
    return transcription
```

The above functions shall have the following behaviour:

1. The functions shall only accept arguments of string data types. If any other data type is passed, it should raise an exception.
2. The functions should return the transcription of the audio as a string.

Model

Additionally, you are required to share the model construction script and the model file to facilitate the utilization of the transcription model.

Submission guidelines

1. Your code and submission will be subject to automated checks, unit tests and plagiarism checks. Please follow the below guidelines so that it is not rejected.
2. Your code will also be automatically checked for code quality as well. Please follow PEP8 guidelines.
3. Share a writeup discussing the methods and techniques you've used in creating your solution.
4. Share the submission as a script (.py) file and include proper documentation that supports your solution.
5. Zip your code and documentation with the following naming convention and upload to the appropriate slot in the google forms <yourname>_nlp_assignment.zip



Solution Guidelines

1. Feel free to use any NLP and text processing techniques, libraries, or external data.
Please provide due credits to those.
2. Try Large Language Models