COMP433/6331
#0000

Course Project Proposal: EEG-to-Speech Kaggle Challenge

Anonymous COMP433/6331 submission

Paper ID 0000

1. Problem Statement and Application

We investigate linguistic neural decoding for restoring communication in people who lost speech (e.g., post-stroke, ALS). An interface captures brain activity related to speech; models decode it to text. We will (i) reproduce a working baseline/benchmark and (ii) explore improvements. Challenges include standing up a full pipeline, data handling, and understanding model components. Improvement avenues: data augmentation, architectural changes, alternative losses, and tokenization strategies. Results will be reported and discussed with attention to reproducibility.

2. Reading Material

EEG→Speech (aux phonemes): [2]. **ASR architectures overview (CTC/seq2seq):** [4]. **Transformer for EEG decoding:** [3]. **Host team's neuroprosthesis baseline context:** [1].

3. Possible Methodology

Benchmarks: run Kaggle-provided baselines without training to verify end-to-end I/O; targets include: Stanford-NPTL causal RNN (ensemble, TTA+5-gram) and UCD-NPL causal RNN (+5-gram). Add a submission script if missing. **Simple baselines:** random/mean/median; linear/logistic, kNN, SVM. **Improvements:** augmentation/preprocessing; GRU/Transformer variants; hyperparameter search; loss swaps. **Repo workflow:** clone and track benchmarks; swap components (e.g., optimizer) under version control. **Kaggle ops:** submit dummy CSV (done), submit loaded baselines, then trained models.

4. Metric Evaluation

Primary: *word error rate* (WER)—edit distance at word level (subs/ins/del). Deliverable: CSV predictions for 1,450 test sentences. During training: track loss and learning curves; post-hoc: compare WER across baselines/improvements and visualize trends.

COMP433/6331
#0000

Supplement: Gantt Chart (1 page)

Phase / Task	**Description**	**Responsible**	W1	W2	W3	W4	W5	W6	W7	W8	**Milestone**
Project Setup	Choose competition, scope, roles	All	✓								Selected competition
Proposal Draft	Write and format proposal	All (Lead: Ion)	✓	✓							Submission-ready draft
Literature Review	Collect and summarize key papers	Kirill		✓	✓	✓					Curated reading list
Environment Setup	Kaggle/Colab/Codespaces; clone baselines	David		✓	✓						Reproducible env
Benchmark Exploration	Run/analyze baseline models	All			✓	✓					Initial submission
Baseline Validation	Train/validate baselines	Elion & Ion				✓	✓				Leaderboard score
Model Improvement	Arch changes, tuning, losses	All					✓	✓	✓		Improved model
Evaluation	WER and qualitative outputs	Kirill & David						✓	✓		Eval report
Final Report	Paper & slides	Elion & Ion							✓	✓	Final submission

References

[1] Nicholas S. Card, Maitreyee Wairagkar, Carrina Iacobacci, Xianda Hou, Tyler Singer-Clark, Francis R. Willett, Erin M. Kunz, Chaofei Fan, Maryam Vahdati Nia, Darrel R. Deo, Aparna Srinivasan, Eun Young Choi, Matthew F. Glasser, Leigh R. Hochberg, Jaimie M. Henderson, Kiarash Shahlaie, David M. Brandman, and Sergey D. Stavisky. An accurate and rapidly calibrating speech neuroprosthesis. *medRxiv*, page 2023.12.26.23300110, 2024. 1

[2] Jihwan Lee, Tiantian Feng, Aditya Kommineni, Sudarsana Reddy Kadiri, and Shrikanth Narayanan. Enhancing Listened Speech Decoding from EEG via Parallel Phoneme Sequence Prediction, 2025. arXiv:2501.04844 [eess]. 1

[3] Young-Eun Lee and Seo-Hyun Lee. EEG-Transformer: Self-attention from Transformer Architecture for Decoding EEG of Imagined Speech, 2021. arXiv:2112.09239 [cs]. 1

[4] Ilias Papastratis. Speech Recognition: a review of the different deep learning approaches, 2021. 1