**BA865: Individual Homework Assignment #2**

Submit this file, edited to include your answers to conceptual questions, along with a Collab notebook (i.e., an .ipynb file) via Blackboard. Leave the rubric blank (we will use the Rubric for our grading).

|  |  |  |
| --- | --- | --- |
| **Criteria** | **Comments** | **Grade** |
| **Comments / Formatting (3)**   * Clean code, intuitive variable naming conventions, code-re-use, etc. * Comments to clarify what purpose your code blocks serve. * Clean notebook formatting, with subsections, e.g., #\*\*TEXT\*\* headers, sub-headers. |  |  |
| **Code & Conceptual Implementation (7)**   * Code executes as is without any errors. * Data imported and pre-processed. * Properly works with pre-trained Keras model. * Properly implements recovery of embeddings. * Properly implements embedding similarity calculations. * Makes suggestions about things to explore going forward that may improve performance. |  |  |
| **Conceptual Questions (5)**   * Responses to conceptual questions (detailed below) are clear, coherent, and correct. |  |  |
| **Total** | | /15 |

**Context & Instructions**

Populate the provided .ipynb notebook in Google Colab to implement a neural network that achieves speaker identification in our spoken digit dataset. As part of your solution, record yourself speaking a short sentence (save the .wav file to your computer and submit it with your notebook). You should produce a well-formatted, well-commented notebook in Google Colab, implementing all the necessary logic. You will submit this notebook as your assignment solution, your recording, and a copy of this document answering the conceptual questions below. Be sure to stick to the methods we have learned and discussed in class!

**Conceptual Questions**

1. (1 point) Leveraging the concepts and ideas we learned in this class, what would you consider persuasive evidence that two voice recordings originated from the same speaker (assume you have a ground-truth label for only one of the recordings)?
2. (1 point) Suppose a tech company or government tried to use an approach like the one you have implemented for real-world voice recognition. What key challenge do you see (name at least 1 and explain)?

1. (1.5 points) An alternative way to use your new model for speaker identification would be to recover the prediction for any new audio sample. What tradeoffs do you see between the prediction-based approach and the approach we implemented in the assignment (name at least 1 and explain)?
2. (1.5 points) Various audio libraries can be used to extract numeric features of audio signals, such as pitch. Conceptually, how might you examine the influence of voice pitch on your model’s speaker identity predictions?

**Peer Support**

Indicate here what other students you spoke to when you were working on this project to brainstorm solutions or work through problems:

**Generative AI Use**

Briefly explain any AI large-language models (LLMs) you used when working on the assignment and how you used them here. Bear in mind that I reserve the right to ask you to explain an aspect of your code to me orally if I see something oddly sophisticated (i.e., something we have covered in class). You will receive a point deduction if you cannot explain orally what the code is doing.