Applied Data Mining

Third Assignment – December 6, 2024

Hello everyone. This is the third assignment for the Applied Data Mining course. The submission deadline is Wednesday, December 18th. I encourage all students to submit their work on time. If you have any questions regarding the exercises, please feel free to reach out.

Part 1:

1) Explain the problems that arise when sampling from the $q(z|x)$ distribution in Variational Autoencoders. Discuss how the sampling process can be rewritten to address these issues and provide an explanation of how this trick works.

2) How does the hyperparameter β affect the trade-off between reconstruction accuracy and disentanglement in a Variational Autoencoder?

3) Discuss the impact of "peephole connections" in LSTMs. How do they theoretically improve the LSTM's ability to maintain its cell state? Also, analyze situations where peephole connections may not be beneficial.

4) Explain the architecture of xLSTM and its functionality. Discuss the key advancements it introduces over the standard LSTM architecture and how these improvements address LSTM's limitations.

Part 2:

The **Pascal Heart Sound Dataset** has been gathered from two sources:

- **Dataset A**: Collected from the general public using the iStethoscope Pro iPhone app, with four classes: **Normal, Murmur, Extra Heart Sound, Artifact**.

- **Dataset B**: Collected from a clinic trial in hospitals using the DigiScope digital stethoscope, with three classes: **Normal, Murmur, Extrasystole.**

Your task is to classify the heart audio recordings into the five categories above by following the steps outlined below:

- You need to perform necessary **preprocessing** steps. This may include, but is not limited to: denoising audio signals to remove background noise, enhancing and resampling the dataset if needed, and converting audio to suitable formats like spectrograms or MFCCs for further analysis. Document all the preprocessing steps clearly in your notebook

- Build a CNN-RNN model to classify the heart sound recordings: Use a pre-trained CNN for feature extraction, then pass these features into an RNN for sequential processing. Finally, classify the heart sounds into relevant categories.

- To address the limitations of the baseline RNN's capacity for long-term dependencies, replace the simple RNN in your model with advanced variants, including**: LSTM** (Long Short-Term Memory), **Bi-LSTM** (Bidirectional LSTM), **xLSTM** (Extended Long Short-Term Memory), **GRU** (Gated Recurrent Unit) and train and evaluate each variant for performance comparison.

- Calculate appropriate metrics such as Accuracy, Precision, Recall, F1-Score, Confusion Matrix. Compare the performance of the baseline CNN-RNN model with the enhanced LSTM, Bi-LSTM, xLSTM and GRU models.

- Your final model should achieve at least 80% accuracy to be considered effective. If the accuracy falls below this threshold, improve preprocessing, model architecture, or hyperparameters to enhance performance.

- Write a detailed report in your notebook, including a brief explanation of each method used, the results of each model, and a clear comparison of their performance.

- **Bonus:** Enhance the LSTM-based model by incorporating a soft attention mechanism to focus on critical parts of the heart sound sequence. Compare the performance improvements over previous models without attention.

Plagiarism will not be tolerated. Homework submissions will be cross-checked against other students' submissions. Additionally, the use of AI to fully generate answers or code for assignments is strictly forbidden.