**Task 5: Knowledge Distillation for GSNR Prediction**

**Objective:**  
In this assignment, you will learn and apply knowledge distillation techniques by training a model on a European dataset (teacher model) and then using this model to guide the training of a smaller model (student model) for predicting GSNR (General Signal-to-Noise Ratio) on the same European dataset. Do the same for USA dataset as well.

1. **Understanding Knowledge Distillation:**
   * Research and read about knowledge distillation, focusing on its principles, benefits, and common techniques. Look for reliable sources such as academic papers, textbooks, and reputable online articles.
   * Understand the concepts of hard and soft distillation in knowledge distillation.
2. **Data Preparation:**
   * Obtain the European dataset provided.
   * Preprocess the dataset as necessary (e.g., handling missing values, normalization).
3. **Model Training on European Dataset (Teacher Model):**
   * Choose a suitable model architecture for GSNR prediction (e.g., a large neural network).
   * Train the teacher model on the European dataset.
   * Evaluate the teacher model's performance on a validation set.
4. **Knowledge Distillation - Hard Distillation:**
   * Train a smaller student model using the ground truth labels from the European dataset.
   * Evaluate the performance of the student model on the European validation set.
5. **Knowledge Distillation - Soft Distillation:**
   * Train a smaller student model using the softened outputs (probability distributions) of the teacher model as well as the ground truth labels from the European dataset.
   * Experiment with different temperature settings for softening the teacher model's outputs.
   * Evaluate the performance of the student model on the European validation set.
6. **Comparison and Analysis (European Dataset):**
   * Compare the results of the hard distillation method and the soft distillation method.
   * Analyze computational efficiency and memory consumption of the teacher and student models.
   * Provide visualizations (e.g., training/validation loss curves, GSNR prediction accuracy) to support your analysis.
7. **Report:**
   * Include code snippets and relevant plots in your report.
   * Discuss the challenges faced and how you overcame them.
   * Summarize your findings and suggest potential improvements or future work.

**Deadline:** 21st July 2024