Addressing Data-Centric AI Challenges in Machine Learning with Advanced Sample-Selection Techniques

Deep learning has achieved significant success across various domains, primarily due to the availability of large datasets and increased computational power. However, the quality of training data remains a crucial factor affecting model performance. As a result, the approach of Data-Centric AI, which emphasizes high-quality data and efficient training processes, is gaining increasing importance.

The goal of our project is to use data sample selection methods to address important issues related to Data-Centric AI, such as robust training with noisy labels, training resilient to adversarial attacks, and explaining the training process. These issues have not yet been thoroughly investigated in the literature, making our project pioneering in nature.

We have chosen this research topic because optimizing data usage is crucial due to the high computational costs associated with training deep learning models. Selecting appropriate data samples for training can significantly enhance model robustness and optimize resource utilization. Methods such as Importance Sampling, Selective Backprop, and the Streaming Approach have already demonstrated that effective sample selection can improve model performance and stability.

Our research aims to develop new sample selection techniques that will increase model robustness against label noise and adversarial attacks, enhance the training process, and create frameworks for explaining the training processes. The project results will contribute to the development of trustworthy and reliable AI systems, in line with the European Union's strategic goals in "Explainable and Robust AI."

The preliminary results of our research are promising. We have developed an innovative streaming approach with a Base-Values Mechanism, which significantly improves the neural network training process. Our original method, called Persistent Entropy-Based (PEB), uses Base-Values to indicate the importance of individual training elements. In experiments on the EMNIST dataset, this method achieved higher accuracy on test data compared to traditional methods, confirming the effectiveness of our approach.

The project has great potential to introduce significant innovations in the field of AI, contributing to the creation of more robust and understandable machine learning models that will better serve society in various applications.