Integrating AI into Imbalanced data learning

Yifan Hsu, Assistant Professor of Statistics, Department of Mathematics, Texas AM University - Kingsville

The abstract: Imbalanced data is a prevalent challenge in machine learning, particularly in applications such as fraud detection, medical diagnosis, and anomaly detection. Traditional machine learning algorithms often struggle with skewed class distributions, leading to biased models that favor majority classes. To address this issue, generating synthetic data algorithms have been developed. This presentation will introducing key approaches such as Synthetic Minority Over-sampling Technique (SMOTE), Random Over-Sampling Examples (ROSE), and explore the integration of artificial intelligence (AI) with imbalanced data learning as Autoencoders, and Generative Adversarial Networks (GANs). Experimental results demonstrate that integrating these methods improves classification performance, leading to more robust and fair machine learning models.