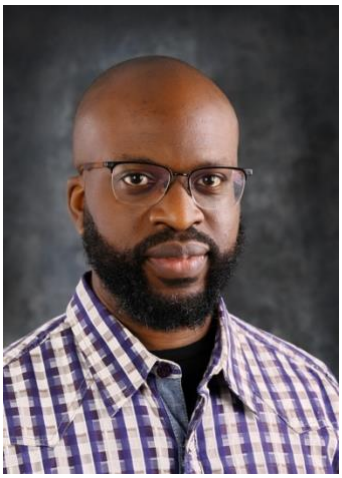
Cal Poly Pomona

Colloquium Series
Department of
Mathematics & Statistics

Dr. Kagba Suaray
Cal State Long Beach



Leveraging Baseball Data to Enhance At-Bat Outcome Prediction in Softball Via Transfer Learning

Abstract: This study explores the application of machine learning and domain adaptation techniques to enhance predictive analytics in women's softball, a sport with limited publicly available data compared to baseball. By leveraging Statcast baseball data as a source domain and LSU softball play-by-play data as the target domain, the project investigates whether transfer learning can improve the prediction of key offensive outcomes—including hits, strikeouts, and walks. Several machine learning models were evaluated, including logistic regression, decision trees, random forests, and domain adaptation approaches, with performance assessed using F1-score. Results indicate that domain adaptation techniques often outperform traditional models in imbalanced softball datasets, particularly for outcomes such as walks and strikeouts. We'll conclude the talk with a brief overview of graduate programs and certificates in statistics, data science and machine learning at CSU Long Beach.

March 18, 2026, 1:05-1:50 pm, room 4-2-314

Join remotely via Zoom: https://cpp.zoom.us/j/84908036425