NM SMART GRID CENTER Fall 2023 Externship

Protecting Smart Grids with Machine Learning: Eastern New Mexico University

Jacqueline Chavez

ENMU - Undergraduate Computer Science Major

The "Protecting Smart Grids with Machine Learning" research project was conducted during the Fall 2023 semester at Eastern New Mexico University (ENMU). The project was undertaken by Jacqueline Chavez, an accomplished computer science student from El Paso, Texas, in collaboration with her mentor, Dr. Edgar Eduardo Ceh-Varela, an Assistant Professor of Computer Science at ENMU. Dr. Ceh-Varela, originally from Yucatan, Mexico, holds a Ph.D. in Computer Science from New Mexico State University (NMSU) and actively engages in multiple research projects with his students as part of the EMMA lab.

Smart grids, integral to modernizing electrical power systems, are highly susceptible to cyber threats. Therefore, the imperative to strengthen these systems against potential cyber-attacks stresses the importance of research efforts in this domain.

Jacqueline's research focuses on cyber-attack detection, a facet often categorized as anomaly detection due to its emphasis on identifying oddities from normal system behavior. The principal objective of this research was to investigate the impact of dimensionality reduction on the prediction accuracy of anomaly detection models, a critical consideration for enhancing smart grid security.

The hypothesis establishes that simplifying the data through dimensionality reduction would expedite computations and enhance the efficacy of machine learning models. To this end, various dimensionality reduction techniques were employed, including Uniform Manifold Approximation and Projection (UMAP), Principal Component Analysis (PCA), and t-Distributed Stochastic Neighbor Embedding (t-SNE). Complementary to these techniques, two semi-supervised anomaly detection algorithms, Isolation Forest and One-Class Support Vector Machine (SVM), and two unsupervised methods, Local Outlier Factor and Elliptic Envelope, were employed.

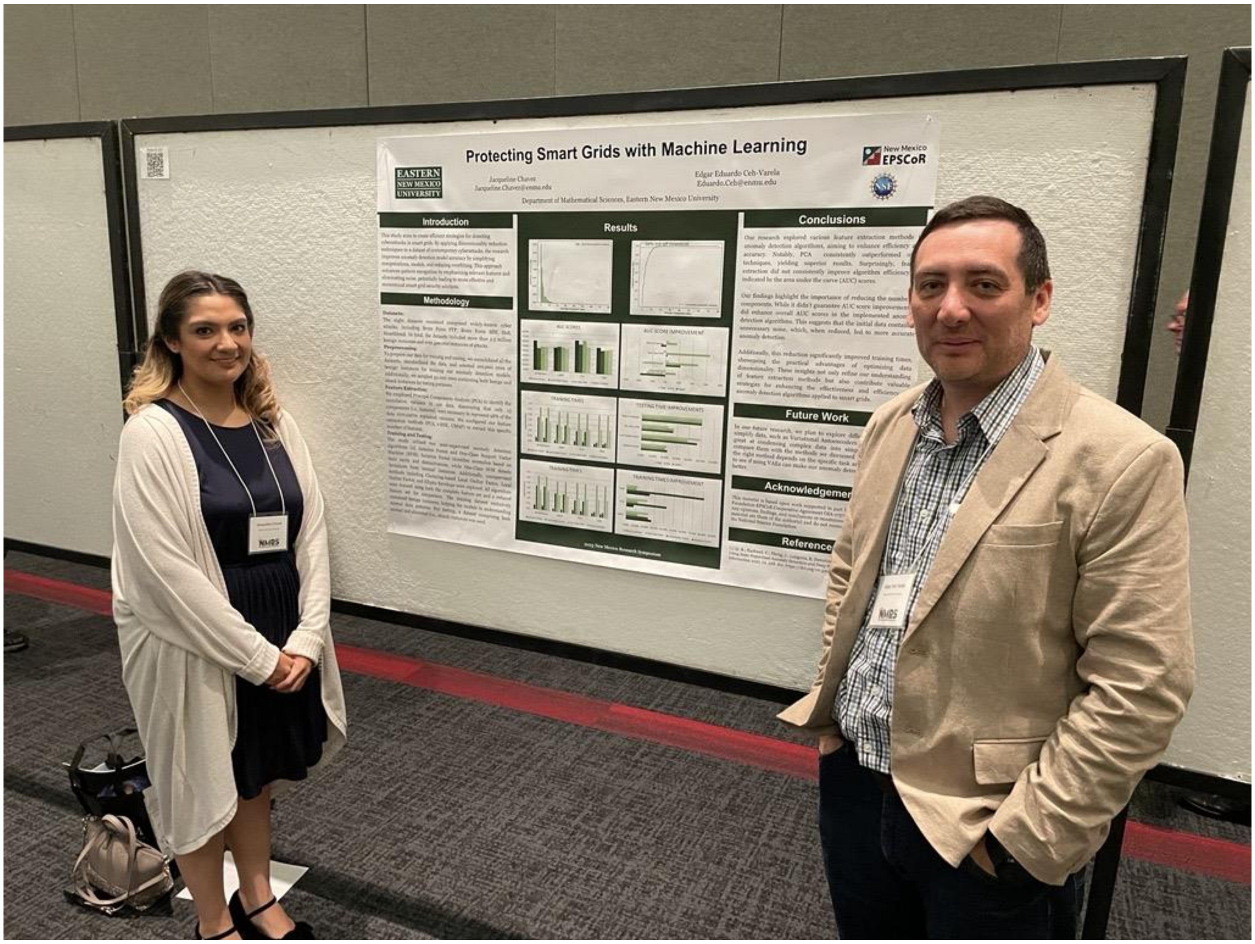
These techniques were comprehensively evaluated across eight widely utilized datasets concerning cyber-attacks. The combination of PCA with Isolation Forest outperformed other methods, as measured by the area under the curve (AUC) metric. This combination simplified computations and the models and reduced overfitting, improving anomaly detection accuracy.

The dimension reduction from 83 to 13, with a minimal 2% information loss, was achieved through dimensionality reduction. While enhancing the importance of identifying relevant features, this approach simultaneously mitigated implicit noise within the data. Beyond the immediate benefits to anomaly detection model accuracy, this method contributed to a notable improvement in training times for all tested models.

The implications of this research extend beyond anomaly detection. The insights garnered refine the understanding of feature extraction methods, offering valuable strategies for enhancing the effectiveness and efficiency of anomaly detection algorithms applied to smart grids. These results contribute to the academic knowledge in the field and present practical applications for more efficient smart grid security solutions.

Jacqueline's involvement in this research, her third in the field, signifies her dedication to computer science and showcases her ability to transcend theoretical knowledge into practical problem-solving. The skills acquired through these research works provide her with a competitive advantage in the dynamic field of computer science.

Looking forward, Jacqueline, upon her expected graduation in Spring 2024, aspires to pursue a Master's in Computer Science. Her commitment to contributing to the field through ongoing research highlights her dedication to advancing smart grid security and computer science.


Protecting Smart Grids with Machine Learning
EASTERN NEW MEXICO UNIVERSITY
Jacqueline Chavez
Jacqueline.Chavez@enmu.edu
Edgar Eduardo Ceh-Varela
Eduardo.Ceh@enmu.edu
Department of Mathematical Sciences, Eastern New Mexico University
New Mexico EPSCoR
Introduction
Methodology
Results
Conclusions
Future Work