ScaLaDE (Scalable Labeling for Data Enrichment)

NASPI Work Group Meeting
DOE Digitizing Utilities Prize

Bruno P. Leao (bruno.leao@siemens.com): Siemens Technology US
Olga Lositsky (Siemens Technology US)
Prof. Jianhui Wang (SMU)
Han Guo (SMU)

September 2023
Problem Definition

Opportunities

Continuous Performance Monitoring
• Focus on time-frequency patterns and how they change over time
• Opportunities
  • Automatic identification of new patterns for investigation
  • Automatic tracking of patterns over time
  • Automatic identification of region affected by phenomena
  • Automatic identification of similar patterns in historical data
  • Visualization in order to properly communicate information to stakeholders
Proposed Solution

Data enrichment that enables identification and comparison of patterns over space and time
ScaLaDE Data Enrichment Concept
Technical Concept
Implementation

Processing Pipeline

0: Data Quality
1: Pre-processing
2: Model Training
3: Model Selection
4: Application

- **Data**: 2 data streams. 18 months of data.
- **Adaptations and prioritizations** performed due to the constraints in time, resources and computational power
AE Training Results

Full Spectrum

input
output

Frequency Filtered

input
output
Embedding Results
Sample Applications

- Distance / Similarity
- Clustering
- Anomaly Detection

Cluster visualizations and graphs illustrating the applications.
Conclusion and Future Steps

Identification of incipient phenomena → Reliable power delivery

• Achievements
  • Successful implementation of proposed data enrichment concept
  • Exploration of sample applications
• Future Work
  • Performance Improvement
  • Application and Impact
  • Integration into Operation
Contact

Bruno Paes Leao
Siemens Corporation, Technology
755 College Road East
Princeton, NJ 08540
USA

E-Mail: bruno.leao@siemens.com
Web: https://www.siemens.com/research