Data Science and Machine Learning for Power Systems

Course Description

This course explores the integration of data science and machine learning (ML) into modern electric power systems. Students will learn to model, analyze, and solve key problems in power systems using statistical and ML methods. Topics include load forecasting, renewable generation forecasting, energy disaggregation, electricity price forecasting, anomaly detection, and decision-focused learning for power system operations.

Prerequisites

- Power system fundamentals (basic operation, optimal power flow, etc.)
- Linear algebra, probability, and statistics
- Introductory machine learning or data science course
- Programming experience in Python (NumPy, Pandas, PyTorch/Scikit-learn)

Course Outline

Week 1–2: Introduction

- Overview of modern power systems and the role of data
- Types of power system data (e.g., AMI, SCADA, PMU, market)
- Intro to data science workflow in energy systems

Week 3–4: Load Forecasting

- Short-term and long-term load forecasting
- Time-series models: ARIMA, exponential smoothing
- ML models: MLP, LSTM, XGBoost
- Evaluation metrics and uncertainty quantification

Week 5: Renewable Generation Forecasting

- Wind and solar prediction challenges
- Weather data integration
- Probabilistic forecasting

Week 6: Electricity Price Forecasting

- Market structure and price formation
- Regression models, tree-based models, and LSTM for price prediction
- Congestion prediction and locational marginal prices (LMPs)

Week 7: Energy Disaggregation and Demand Response

- Non-intrusive load monitoring (NILM)
- Dictionary learning and sparse coding
- Demand response data analytics and segmentation

Week 8: Data-Driven Anomaly Detection and Cybersecurity

- PMU-based event detection
- Anomaly detection with autoencoders
- Cyber-physical attack detection

Week 9: Optimization and Machine Learning

- Optimization in power systems (unit commitment, OPF)
- Differentiable optimization layers
- Decision-focused learning (predict-then-optimize)

Week 10: Graph-Based Learning in Power Systems

- Power grids as graphs
- Graph signal processing
- Graph neural networks (GNNs) for topology and load prediction

Week 11: Reinforcement Learning for Grid Control

- Basics of reinforcement learning (RL)
- RL for frequency regulation, voltage control, and energy storage
- Safe RL and constrained RL

Week 12: Forecasting and Control under Uncertainty

- Uncertainty quantification: Bayesian methods, conformal prediction
- Scenario generation for planning and scheduling
- Robust and stochastic optimization

Week 13: Scalability and Real-World Applications

- Data platforms and cloud computing for power systems
- Real-world case studies: ERCOT, CAISO, NYISO, PJM
- Industry and policy implications of ML in power systems

Week 14–15: Final Projects and Presentations

- Project presentations
- Review and future trends: trustworthy AI, explainability, ML for grid resilience

Assessment

- Homework Assignments (30%)
- Midterm Exam (20%)

- Final Project (30%)
- Participation (20%)

Software and Tools

- Python (NumPy, Pandas, Scikit-learn, PyTorch, TensorFlow)
- Jupyter Notebooks
- Power system tools: MATPOWER, GridLAB-D, or PowerModels.jl (optional)
- Datasets: Pecan Street, UMass Smart*, ERCOT public data