EV Charging Demand Forecasting

Advanced machine learning analysis of electric vehicle charging patterns across California stations from 2021-2024. This study leverages neural networks, decision trees, and SHAP interpretability frameworks to predict demand patterns and optimize grid infrastructure planning.

Dataset Overview

Data Characteristics

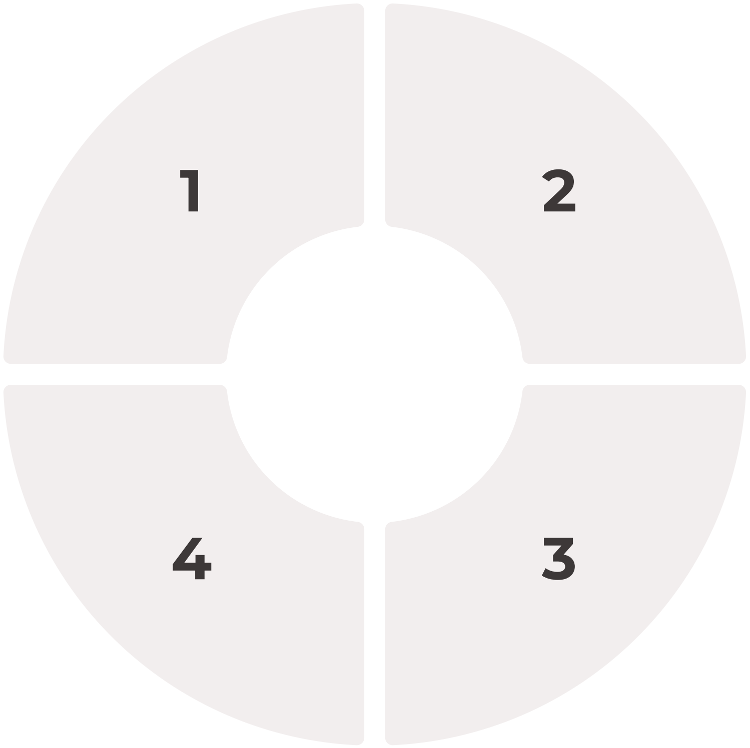
- 59,810 hourly observations across 3.5 years
- Two California charging stations combined
- 24 features capturing energy, pricing, infrastructure, and environmental metrics
- Temporal granularity: hourly measurements with seasonal patterns

Key Variables

- Target: EV Charging Demand (kW)
- Renewable energy production (solar + wind)
- Grid stability and availability metrics
- Weather conditions and temporal features

Seasonal Demand Patterns

EV charging demand exhibits distinct seasonal and daily fluctuations driven by renewable energy availability, grid conditions, and user behavior patterns.



Winter

Highest evening demand, moderate morning usage. Lower renewable production increases grid dependency.

Spring

Balanced demand across day periods. Improved renewable energy availability supports efficient charging.

Summer

Peak afternoon demand coinciding with solar production. Lowest evening renewable output.

Autumn

Stable nighttime demand with moderate fluctuations. Transitional renewable energy patterns.

Hourly Demand Distribution

Normalized charging demand reveals clear temporal patterns that inform infrastructure planning and load management strategies.

1. **Nighttime (0-6h)**

   Lowest demand period with normalized values 0.35-0.45. Ideal window for grid-friendly charging.

2. **Morning (6-12h)**

   Gradual increase to 0.50-0.55 as commuters begin charging. Solar production ramps up.

3. **Afternoon (12-18h)**

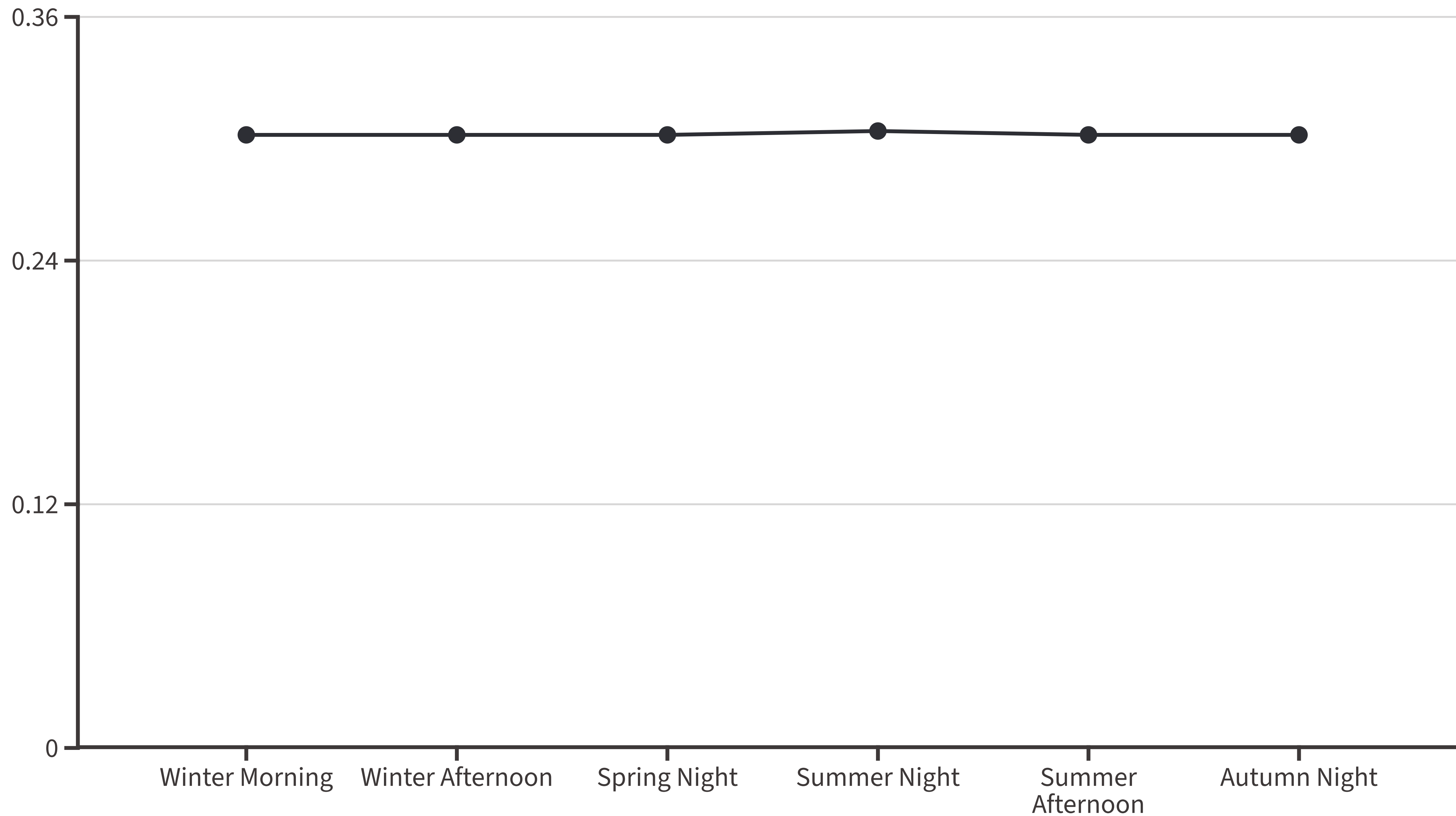
   Peak demand window reaching 0.60-0.65. Maximum solar availability enables renewable charging.

4. **Evening (18-24h)**

   Secondary peak at 0.55-0.60 as users return home. Grid pressure intensifies as solar declines.

Renewable Energy Production

Total renewable energy production (solar + wind combined) demonstrates seasonal variability that directly influences charging sustainability and grid impact.



Nighttime consistently produces the highest renewable energy across seasons, particularly in summer and spring. This counterintuitive pattern suggests strong wind generation compensates for solar absence, creating optimal conditions for overnight charging strategies.

Weather Impact on Charging Efficiency

Weather conditions significantly influence EV charging efficiency, with normalized efficiency values ranging from 0.45 to 0.53 across different meteorological scenarios.

Sunny Conditions

Highest efficiency at 0.53 normalized. Optimal temperature and solar production create ideal charging environment.

Cloudy Weather

Moderate efficiency at 0.50 normalized. Reduced solar but stable temperatures maintain performance.

Partly Cloudy

Efficiency at 0.49 normalized. Variable conditions create moderate charging performance.

Rainy Periods

Efficiency at 0.48 normalized. Lower temperatures and humidity impact charging systems.

Clear Skies

Lowest efficiency at 0.46 normalized. Extreme temperatures may reduce battery performance.

SHAP Feature Importance Analysis

SHAP (SHapley Additive exPlanations) reveals which factors most strongly influence EV charging demand predictions in our linear regression model.

1. **Renewable Energy Usage (%)**

   Dominant predictor with strongest impact. Higher renewable usage correlates with reduced grid demand, highlighting sustainability's role in demand patterns.

2. **Electricity Price ($/kWh)**

   Second most influential feature. Price signals drive user behavior, with higher prices reducing charging demand through load shifting.

3. **Carbon Emissions (kgCO2/kWh)**

   Third key driver. Higher emissions indicate grid stress, correlating with increased charging demand during peak fossil fuel generation.

4. **Adjusted Charging Demand (kW)**

   Engineered feature capturing optimized load patterns. Reflects smart charging strategies and demand response programs.

SHAP Model Interpretation

Key Insights

The SHAP force plot reveals how individual features push predictions above or below the baseline expected value of 0.139 kW.

- Renewable energy usage reduction: -0.018 kW impact
- Electricity price increase: -0.030 kW impact
- Carbon emissions increase: contributing positive shift
- Final prediction: 0.091 kW (34% below baseline)

Actionable Insights

Model transparency enables grid operators to understand prediction drivers and implement targeted interventions for demand management.

Strategic Recommendations

Maximize Renewable Integration

Prioritize charging during high renewable availability windows. Nighttime and morning periods show strongest renewable production, reducing grid stress and carbon footprint.

Dynamic Pricing Incentives

Leverage electricity price signals to shift demand. Time-of-use rates can flatten demand curves by incentivizing off-peak charging behavior.

Smart Charging Optimization

Implement demand response programs using adjusted charging demand metrics. Optimize load distribution to minimize peak stress and maximize renewable utilization.

Carbon Footprint Reduction

Target low-emission charging windows. Invest in energy efficiency measures and cleaner technologies to support sustainability objectives while managing demand.

Neural Network Architecture

Multi-layer perceptron (MLP) regression model predicts EV charging demand using five key economic and environmental features.

Input Layer

Five normalized features: Net Energy Cost, Carbon Footprint Reduction, Electricity Price, Carbon Emissions, and Renewable Energy Usage.

Hidden Layer 1

64 neurons with ReLU activation. Captures complex non-linear relationships between economic and environmental factors.

Hidden Layer 2

32 neurons with ReLU activation. Further feature compression and pattern extraction for demand prediction.

Output Layer

Single neuron producing continuous kW demand prediction. Linear activation for regression output.

Neural Network Performance

The MLP model demonstrates exceptional predictive accuracy, converging rapidly during training with minimal loss.

0.9978

R^2 Score

Explains 99.78% of variance in charging demand, indicating near-perfect fit to underlying patterns.

0.0032

MAE (kW)

Mean absolute error of just 3.2 watts demonstrates precise predictions at hourly granularity.

0.0040

RMSE (kW)

Root mean squared error of 4.0 watts confirms minimal prediction deviation from actual values.

13

Training Epochs

Model converged in just 13 iterations, demonstrating efficient learning from strong feature relationships.

Training Convergence

Loss curve analysis reveals rapid optimization with early stopping after convergence detection.

Training Dynamics

- Initial loss: 0.00084 (iteration 1)
- Sharp decline in first 3 epochs
- Stabilization by epoch 10
- Early stopping at epoch 13
- Final loss: 0.00002

Interpretation

The steep initial drop followed by gradual refinement indicates the model quickly learned primary relationships between economic/environmental features and demand. Early convergence suggests the five selected features capture the essential dynamics of charging behavior without requiring complex non-linear transformations.

Prediction Accuracy Visualization

Actual vs. predicted values demonstrate the neural network's precise tracking of demand fluctuations across 100 test samples.

> The model's predictions (marked with X) closely follow actual demand values (marked with circles) across diverse operating conditions. Minimal deviation indicates robust generalization to unseen data, validating the model for deployment in real-world forecasting applications.

This accuracy enables grid operators to anticipate demand with confidence intervals narrow enough for operational decision-making, including capacity planning, renewable integration scheduling, and dynamic pricing implementation.

Decision Tree: Day Period Classification

A decision tree classifier achieves perfect accuracy (100%) predicting time-of-day periods using only month and hour features.

Perfect Classification

Accuracy: 1.00. The model correctly classifies all test samples into Night, Morning, Afternoon, and Evening periods.

Simple Decision Rules

Hour-based thresholds provide interpretable splits: Night (0-3, 21-23), Morning (4-11), Afternoon (12-16), Evening (17-20).

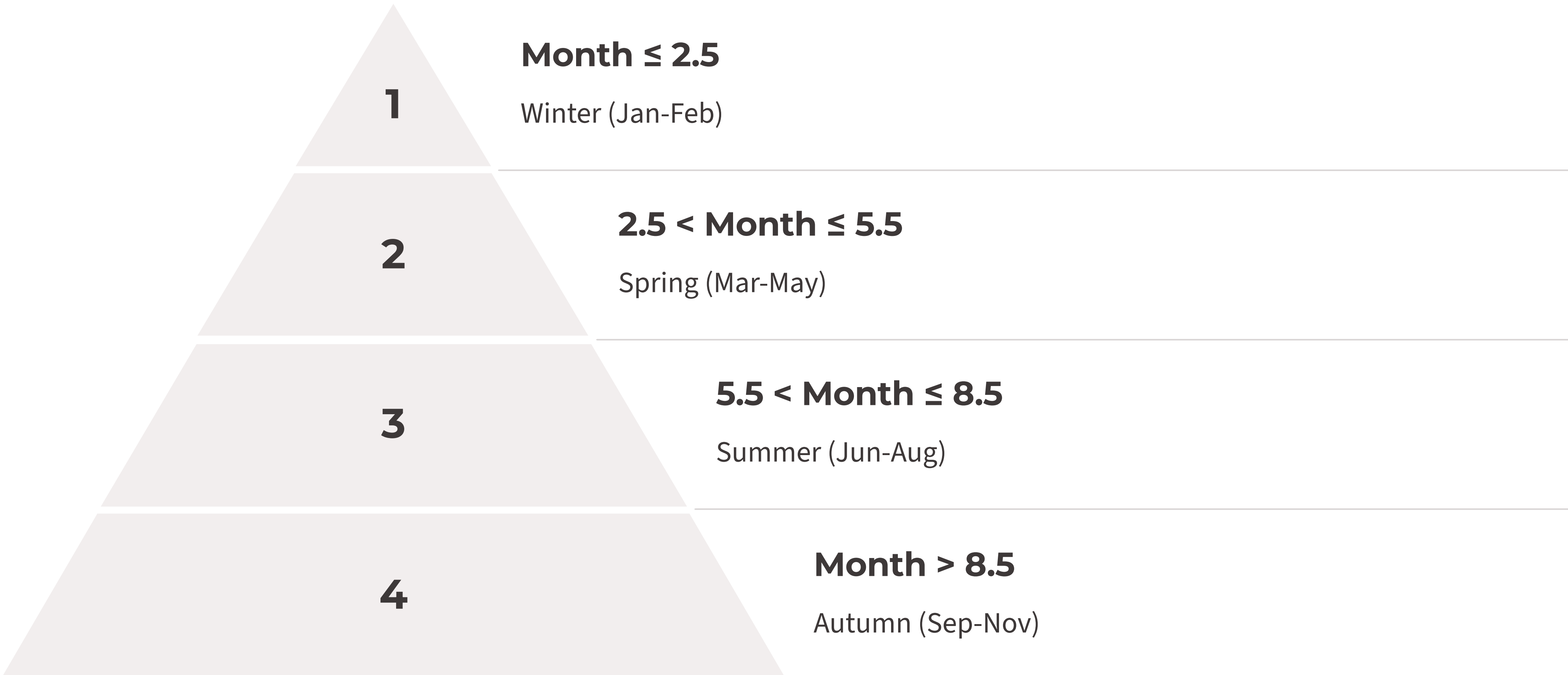
10 Leaf Nodes

Model complexity constrained to maintain interpretability while capturing all essential temporal patterns.

The confusion matrix shows zero misclassifications across all four day period categories, confirming the model's deterministic nature when temporal boundaries are well-defined.

Decision Tree: Seasonal Classification

Season prediction using month, hour, charging demand, and electricity price achieves perfect accuracy (100%), with month being the dominant feature.



The tree structure reveals that month alone sufficiently determines season with perfect accuracy, though additional features (demand, price) were available. This demonstrates the deterministic relationship between calendar month and meteorological season.

Understanding Gini Impurity

Technical Definition

Gini impurity measures node purity in classification trees, ranging from 0 (pure) to 0.5 (maximum impurity for binary classification).

- Gini = 0: All samples belong to one class
- Higher Gini: Mixed class distribution
- Decision trees minimize Gini at each split

Cricket Ball Analogy

Imagine sorting cricket balls by color into boxes. A box with only red balls has Gini = 0 (perfect). A box with equal red and white balls has high Gini (impure).

The decision tree asks questions like "Is the ball red?" to create progressively purer boxes until each contains only one color (class).

Grid Availability Prediction

Decision tree classifier predicts grid availability status using electricity price, hour, month, and grid stability index, achieving 95% accuracy.

1

Primary Feature

Electricity price ($/kWh) provides the first split. Price ≤ 0.191 strongly indicates grid availability.

2

Secondary Feature

Grid stability index refines predictions. Higher stability (>0.516) confirms availability when price is low.

3

Temporal Factors

Hour and month provide final classification refinement for borderline cases with mixed indicators.

Grid Availability Decision Rules

The decision tree reveals critical thresholds that determine grid availability status through interpretable if-then rules.

Low Price + High Stability = Available

When electricity price ≤ \$0.191/kWh and grid stability index > 0.516, the grid is reliably available. This represents optimal operating conditions.

High Price = Likely Unavailable

Prices exceeding \$0.191/kWh indicate grid stress, increasing unavailability risk regardless of other factors.

Temporal Edge Cases

Evening hours (>17:30) and certain months (>6.5) elevate unavailability risk when combined with borderline stability metrics.

These rules enable proactive grid management by identifying high-risk conditions before outages occur.

Model Comparison Summary

Model	Task	Accuracy/R^2	Key Features
Neural Network	Demand Prediction	$R^2 = 0.9978$	5 economic/environmental
Linear + SHAP	Demand Interpretation	R^2 = Variable	31 features (full dataset)
Decision Tree	Day Period	100% accuracy	Month, Hour
Decision Tree	Season	100% accuracy	Month (primary)
Decision Tree	Grid Availability	95% accuracy	Price, Stability, Time

Each model serves distinct operational purposes: neural networks for precise forecasting, SHAP for transparency, and decision trees for rapid classification and rule extraction.

Conclusions & Future Directions

Key Findings

Renewable energy usage and electricity pricing dominate demand patterns. Neural networks achieve near-perfect prediction accuracy. Temporal and grid stability factors enable reliable availability forecasting.

Operational Impact

Models enable proactive grid management, renewable integration optimization, and dynamic pricing strategies. Decision tree rules provide actionable thresholds for real-time operations.

Next Steps

Deploy models in pilot programs for demand response validation. Integrate weather forecasting for extended prediction horizons. Expand dataset to additional geographic regions and vehicle types.

This analysis demonstrates the power of combining interpretable and high-performance machine learning approaches to solve complex energy infrastructure challenges.