Functional Document

Advanced Solar Power Forecasting System

1. **Introduction**

The Solar Power Forecasting System is an **AI-powered** solution designed to predict photovoltaic energy output with **high precision**. This document outlines the functional requirements for Sprint 1, which focuses on establishing the core forecasting engine, data pipeline architecture, and minimum viable integration capabilities. The system will serve as a critical decision-support tool for renewable energy integration, helping to balance supply and demand in smart grid environments while reducing reliance on fossil fuel-based peaker plants.

2. **Product Goal**

Primary Objectives:

- Develop an LSTM neural network capable of 24-hour-ahead forecasting with ≤15% mean absolute percentage error (MAPE).
- Create a robust data ingestion pipeline processing 10,000+ data points/minute from heterogeneous sources.
- Implement API endpoints for seamless integration with existing energy management systems.

Success Criteria:

- Achieve **95.87% accuracy** in clear sky conditions (R^2 Score of 0.9587).
- Maintain **>90% accuracy** during variable weather patterns (R^2 Score of 0.9587).
- Reduce forecast errors by **60%** compared to current ARIMA-based solutions (based on the lower RMSE of your custom LSTM model).
- Process real-time data with **<500ms latency** (assuming your model can meet this requirement in a practical setup).

3. **Demography (Users, Location)**

User Profiles:

1. Grid Operators

Responsibilities: Maintain grid stability, dispatch generation resources

Usage Patterns:

- Continuous monitoring during daylight hours
- Requires 15-minute forecast updates
- Needs alerting for rapid generation drops

2. Solar Farm Managers

Responsibilities: Optimize plant performance, schedule maintenance

Usage Patterns:

- Daily generation planning
- Performance benchmarking
- Fault detection through forecast deviations

3. Energy Traders

Responsibilities: Buy/sell energy in wholesale markets

Usage Patterns:

- Uses day-ahead forecasts for bidding
- Requires probabilistic forecasting (P10/P90 scenarios)
- Needs API integration with trading platforms

Geographic Specifications:

Phase 1 Deployment:

- 5 solar farms across Tamil Nadu (1GW total capacity)
- Focus on tropical climate patterns

Phase 2 Expansion:

- Desert climates (Rajasthan)
- Temperate zones (Himachal Pradesh)
- International markets (Middle East, Southeast Asia)

4. **Business Processes**

4.1 Data Management Workflow

Data Acquisition:

- SCADA system integration via OPC UA
- Weather API polling (15-min intervals)
- Satellite cloud imagery processing.

Data Validation:

- Range checking (0-100% of nameplate capacity)
- Temporal consistency verification.

Data Transformation:

- Temporal alignment (all sources → 15-min intervals)

- Feature engineering (clear sky index, persistence models)
- Normalization (MinMaxScaler per plant)

4.2 Forecasting Workflow

Model Execution:

- Scheduled runs (00:00 UTC daily)
- Event-triggered updates (severe weather alerts)

Result Generation:

- Point forecasts (best estimate)
- Probabilistic forecasts (confidence intervals)
- Scenario analysis (storm preparedness)

Result Dissemination:

- REST API (JSON payloads)
- MQTT for real-time updates
- PDF reports for regulatory compliance

5. **Features**

1: Core Forecasting Engine:-

Technical Specifications:

- Architecture: 3-layer stacked LSTM

- 128 neurons per layer
- Dropout rate: 0.2
- Dense output layer with linear activation

Training Configuration:

- Batch size: 64
- Epochs: 100 (early stopping)
- Loss function: Huber loss (δ=1.0)

Input Features:

- Historical power (48 timesteps)
- GHI (Global Horizontal Irradiance)
- Cloud opacity
- Panel temperature

User Stories:

- *"As a grid operator, I need 95% reliable day-ahead forecasts to optimize thermal unit commitment."*
- *"As an asset manager, I want to compare actual vs forecasted generation to identify underperforming arrays."*

2: Anomaly Detection Module

Technical Specifications:

Methodology:

- Isolation Forest for point anomalies
- LSTM autoencoder for contextual anomalies

Thresholds:

- 3σ for production outliers
- 20% deviation from expected clearsky

Alert Types:

- Equipment failure patterns
- Soiling detection
- Communication failures

User Stories:

- *"As a maintenance supervisor, I need immediate alerts when inverters show abnormal output patterns."*
- *"As a performance analyst, I want automated detection of panel degradation trends."*

6. **Authorization Matrix**

Define the roles and their corresponding access levels:

Role	Data Access	System Access	Action Permissions
System Admin	Full	SSH, Dashboard	Model retraining, User management.
Data Engineer	Raw datasets	ETL tool	Pipeline configuration, Data QA
Energy Trader	Forecast results only	Trading API	Query forecasts, Set alerts

Role	Data Access	System Access	Action Permissions
Field Technician	Equipment-level data	Mobile app	Anomaly acknowledgment, Report generation
Read-only User	Aggregated reports	Dashboard view	No write operations

7. **Assumptions**

Technical Assumptions:

Data Quality:

- SCADA systems provide ≥1 year of historical data at 15-min resolution
- Weather APIs deliver forecast updates at least hourly
- <5% missing values in operational data streams

Infrastructure:

- AWS EC2 p3.2xlarge instances available for training
- 10Gbps network connectivity between data centers
- 99.9% uptime for critical dependencies.

Business Assumptions:

Regulatory:

- No special permits required for data collection.
- Forecasts qualify as "non-firm" information per grid codes

Operational:

- Plant operators will provide ground truth validation
- Market rules won't change during development.