DIYguru
innovations in future mobility

**Submitted by:**
Aravindh G *(DIYguru username: aravindhkrishnan09)*
Linkedin | GitHub | E-mail

# Capstone Project: Intelligent EV Battery Health Prediction System

- **Strategic Objective** – Predictive Maintenance for EV Batteries | SoH Estimation & RUL Forecasting
- **Core Problem Areas** – Complex .mat File Structures | High-Frequency Sensor Noise | Non-Linear Degradation Dynamics

**Project Links:**

- Live Dashboard: https://ev-battery-health-prediction-app-v2.streamlit.app/
- RAG AI Assistant: https://ev-battery-rag-1023905756100.us-central1.run.app/
- Code & Notebooks: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Capstone_Project/README.md
- Project Presentation: https://ev-battery-health-predic-z4pcfbk.gamma.site/

**End-to-End Predictive Maintenance Architecture**

- **Dataset Coverage** – NASA PCoE | 2.1M+ Measurements | 636 Charge–Discharge Cycles
- **Data Engineering Layer** – Nested .mat Parsing (SciPy) | Hierarchical Flattening | Analytical Pandas DataFrames
- **Cycle-Level Feature Aggregation** – Mean Voltage | Mean Current | Mean Temperature | Capacity Fade Indicators
- **Physics-Based Feature Design** – Voltage Drop (ΔV) | Temperature Rise (ΔT) | Internal Degradation Signals
- **Scalable Data Pipeline** – Automated AWS S3 Ingestion | CSV & Parquet Storage | Dockerized ETL Container

**Predictive Modeling & Learning Strategies**

- **Tree-Based Baselines** – Random Forest | XGBoost | Non-Linear Regression Benchmarks
- **Deep Learning Models** – LSTM | GRU | Sequential Time-Series Degradation Modeling
- **Temporal Learning Design** – Sliding Windows | Sequence-to-One SoH Prediction | Multi-Cycle Context
- **Model Optimization Levers** – Time Steps 30 | Tanh Activation | 128/64 Units | 100 Training Epochs
- **Complementary Learning** – K-Means Aging Stage Clustering | Q-Learning Charging Policy Optimization

**Model Performance, Improvements & Comparative Analysis**

- **Primary Accuracy Target** – SoH Prediction Within ±5% Absolute Error on Unseen Batteries
- **Baseline Strength** – Random Forest | $R^2$ ~0.94 | MAE 1.70% | Strong Generalization Anchor
- **Gradient Boosting Alternative** – XGBoost | $R^2$ 0.9276 | MAE 1.76% | Robust Non-Linear Capture
- **Deep Learning Advancement** – GRU as Top Performer | Lowest MAE 1.50% (B0018)
- **Cross-Battery Consistency** – GRU | $R^2$ 0.79 (B0007) | $R^2$ 0.92 (B0018) | Stable Across Degradation Profiles
- **Architectural Improvements** – Time Steps 30 | Tanh Activation | 128/64 Units | 100 Epochs
- **LSTM Comparison** – Variable $R^2$ 0.29–0.81 | Sensitivity to Degradation Pattern Diversity
- **Quantified Improvement** – GRU MAE 2.90% vs LSTM MAE 5.73% on B0007 | ~50% Error Reduction

**Impact, Reliability & Practical Significance**

- **Goal Exceeded** – Substantially Better Than ±5% SoH Error Threshold
- **Operational Impact** – Early Degradation Detection | Reduced Unplanned Battery Failures
- **Model Reliability** – Consistent Unseen-Battery Performance | Lower Overfitting Risk
- **Decision Trust** – Physics-Informed Features | SHAP-Based Interpretability
- **RUL Evaluation Insight** – Test Batteries Maintaining Health Beyond Failure Threshold

**GenAI-Augmented Intelligence Layer**

- **RAG Architecture** – Dockerized Containers on Google Cloud Run
- **Core Stack** – Vertex AI Search | Gemini 2.5 Flash | Containerized Inference Services

- **Knowledge Enablement** – Grounded Technical Answers | Citation-Backed Reasoning | Model Context Awareness

**Deployment, Monitoring & Explainability**

- **Production Deployment** – Streamlit Dashboard | Docker Containers | Cloud-Native Hosting
- **User Interaction Layer** – Real-Time SoH & RUL | CSV Uploads | AWS S3 Integration
- **Model Explainability** – SHAP Values | Feature Attribution | Transparent Predictions

**Knowledge Systems & Engineering Automation**

- **ML Reference Handbook**: https://docs.google.com/document/d/124kgXhx3bXfq7GbmAUBpu5sGk1QKvqUau468rDnHRVk/edit?tab=t.bwyqhzzhi4q9
- **Custom AI Agents**: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/tree/main/AI
- **Automation Stack** – Jupyter notebook → Quarto → Gamma | Docker-Based Reporting Pipelines

**Final Takeaway**

- **Technically Mature** – Physics-Guided Features + Advanced Time-Series Models
- **Measurable Impact** – ~50% Error Reduction vs LSTM on Hard-to-Predict Batteries
- **Production-Grade** – Dockerized | Container-Orchestrated | Interpretable | GenAI-Enhanced

---

**<u>Other Main Projects:</u>**

**General Data Analysis**

Summary published to site via Gamma: https://engineering-materials-da-wus3wrn.gamma.site/

Readme file: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Projects/Main_Project_Data_Analysis/README.md

Jupyter notebook: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Projects/Main_Project_Data_Analysis/Engineering_Materials__Project_AravindhG_JupyterNotebook.ipynb

**EV Charging Patters**

Summary published to site via Gamma: https://ev-charging-demand-forec-5hppct2.gamma.site/

Readme file: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Projects/Main_Project_ML/EV%20Charging%20Patterns/README.md

Jupyter notebook: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Projects/Main_Project_ML/EV%20Charging%20Patterns/EV_Charging_Patterns.ipynb

**Vehicle Energy Dataset (VED) Analysis & Machine Learning**

Summary published to site via Gamma: https://vehicle-energy-dataset-a-6knezqt.gamma.site/

Readme file: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Projects/Main_Project_ML/VED%20Analysis/README.md

Jupyter notebook: https://github.com/aravindhkrishnan09/Data-Science-and-Engineering-Analytics/blob/main/Projects/Main_Project_ML/VED%20Analysis/VED_ML.ipynb