Literature Review: Hybrid Physics + Machine Learning Approaches for Electric Vehicle Energy Consumption Prediction

P2I Project – Digital Twin for EV Energy Modeling

1 Introduction: Limitations of Classical Approaches

Predicting the energy consumption of an electric vehicle (EV) has traditionally relied on two main families of models. **Physics-based** models use mechanical laws (aerodynamic drag, rolling resistance, road grade, inertia, thermal effects). They are interpretable and robust but oversimplify real-world driving conditions. In practice, their prediction error commonly ranges from 12–15%.

In contrast, **data-driven** Machine Learning (ML) models can achieve very high accuracy. However, they require large datasets, lack interpretability, and often generalize poorly when deployed outside the distribution of the training data (new vehicles, unseen weather conditions, different driving styles).

These limitations have motivated the emergence of a third approach: **hybrid models that combine physics with Machine Learning**. This hybridization is increasingly central in modern Digital Twin systems and in energy modeling of EVs.

2 Hybrid Physics + Machine Learning Approach

The core idea behind hybrid models is the following:

$$E_{\text{true}} = E_{\text{physics}} + f_{\text{ML}}(X)$$

where:

- E_{physics} is the energy predicted by the physics-based model,
- $f_{\text{ML}}(X)$ is the data-driven correction learned from the residuals,

- X denotes the input features (average speed, ambient temperature, payload, road grade, etc.).

This decomposition provides several key advantages:

- **Improved accuracy**: the MAPE can be reduced from 12–15% to as low as 5–7%.
- **Interpretability**: the physical component remains the backbone of the prediction.
- **Reduced data requirements**: the ML model learns only the residuals, which are often simpler and closer to zero.

3 Hybrid Approaches in the Literature

3.1 Hybrid Linear Correction

A basic form of hybridization uses a linear regression model to learn the residuals. The residual prediction takes the form:

$$\text{residual} = \beta_0 + \beta_1 v + \beta_2 T + \beta_3 \text{payload} + \beta_4 \text{grade} + \epsilon$$

This method is simple, interpretable, and fast to train. Its main limitation is its inability to capture nonlinear behaviors present in real EV energy consumption patterns.

3.2 Hybrid EV Model with XGBoost

State-of-the-art hybrid approaches increasingly rely on **gradient boosting** algorithms such as XGBoost, because they:

- capture nonlinear relationships effectively,
- require minimal feature preprocessing,
- generalize well even with moderately sized datasets,
- are robust to noise and outliers.

The standard pipeline consists of:

1. computing the physics-based prediction E_{physics},
2. computing the residuals,
3. training the XGBoost regressor on these residuals,
4. producing the final prediction via the sum of physics + learned correction.

This strategy has become a reference method for building Digital Twins of EV energy consumption.

4 Role of Machine Learning in a Hybrid Model

In a hybrid framework, ML does **not** aim to replace physics. Instead, it serves to **correct what physics alone cannot model**, such as:

- real-world driving behavior (irregular acceleration, stop-and-go traffic),
- thermal effects influencing motor and battery efficiency,
- subtle road-vehicle interactions (wind, tire pressure, surface friction),
- nonlinear efficiency variations across operating conditions.

Thus, the physics model provides structure and coherence, while the ML layer provides fine-grained adjustments.

5 Towards Physics-Informed Neural Networks (PINN)

Physics-Informed Neural Networks (PINN) represent a natural extension of hybrid modeling. Instead of learning corrections on top of the physics, PINNs integrate physical laws directly into the neural network training objective:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda \mathcal{L}_{\text{physics}}$$

where:

- $\mathcal{L}_{\text{data}}$ measures the mismatch with the observed data,
- $\mathcal{L}_{\text{physics}}$ penalizes violations of physical equations (e.g., mechanical balance, energy conservation).

PINNs provide:

- improved generalization with limited data,
- physically consistent predictions,
- suitability for safety-critical domains such as automotive systems.

In this project, PINNs are considered as a future extension beyond the initial hybrid XGBoost phase.

6 Conclusion

Hybrid Physics + Machine Learning approaches currently represent the state-of-the-art solution for EV energy consumption prediction. They offer an optimal trade-off between accuracy, interpretability, and robustness. Physics-based models provide structure and explanatory power, while Machine Learning captures real-world nonlinearities and unmodeled effects. PINNs further extend this synergy by embedding physical constraints directly into the neural network training process.