

Роман Дропа

Кафедра систем автоматизованого проектування, Національний університет “Львівська політехніка”,
вул. С. Бандери, Львів, Україна, E-mail: roman.i.dropa@lpnu.ua, ORCID 0009-0005-0810-2081

ПРОГНОЗУВАННЯ ЗАЛИШКОВОГО ПРОБІГУ ЕЛЕКТРОМОБІЛЯ З УРАХУВАННЯМ ДЕГРАДАЦІЇ БАТАРЕЇ НА ОСНОВІ ДАНИХ МАШИННОГО НАВЧАННЯ

Отримано: Серпень 14, 2025 / Переглянуто: Серпень 25, 2025 / Прийнято: Вересень 15, 2025

© Дропа Р., 2025

<https://doi.org/10.23939/cds2025.02.055>

Анотація. Прогнозування залишкового пробігу електромобіля (EV) – критично важливе завдання для підвищення ефективності планування поїздок, зниження ризику повного розрядження батареї та покращення користувацького досвіду. Одним із основних факторів, що ускладнює формування точних прогнозів, є деградація літій-іонних батарей, яка поступово знижує їхню ємність і впливає на запас ходу. У статті досліджено вплив урахування деградації батареї через показник стану здоров'я (SoH) на точність прогнозування пробігу за допомогою алгоритмів машинного навчання. Виконано порівняльний аналіз моделей з і без урахування SoH на основі реальних телеметричних даних. Результати показують, що інтеграція показників деградації батареї у модель істотно підвищує точність прогнозування залишкового пробігу. Наведені висновки можуть бути корисними для розробників систем енергоменеджменту та програмного забезпечення для електромобілів.

Ключові слова: електромобіль, залишковий пробіг, машинне навчання, деградація батареї, гібридна модель, стан здоров'я батареї (SoH), прогнозування енергоспоживання, телеметрія.

Вступ

Зростання популярності електромобілів (EV) зумовлене прагненням зменшити викиди парникових газів та залежність від викопних палив. Однак одним із ключових викликів для широкого впровадження EV є точне прогнозування залишкового пробігу, що безпосередньо впливає на довіру споживачів та ефективність експлуатації транспортних засобів. Точність таких прогнозів ускладнюється деградацією літій-іонних батарей, яка з часом зменшує їхню ємність і, відповідно, пробіг автомобіля на одному заряді [1].

Деградація батареї – складний процес, що залежить від багатьох факторів, зокрема кількості циклів заряджання – розряджання, температури експлуатації та режимів використання [2]. Традиційні методи оцінювання стану здоров'я батареї (SoH) часто не враховують усі ці змінні, що може призводити до неточностей у прогнозуванні залишкового пробігу. Використання методів машинного навчання для аналізу великих обсягів даних телеметрії відкриває нові можливості для точнішого прогнозування, оскільки ці підходи можуть виявляти складні нелінійні залежності між параметрами експлуатації та деградацією батареї.

Останні дослідження демонструють ефективність застосування алгоритмів машинного навчання для оцінювання SoH та прогнозування залишкового пробігу. Наприклад, у роботі [3] запропоновано підхід, що поєднує доменні знання із нейронними мережами для прогнозування деградації батареї, що дає змогу підвищити точність та інтерпретованість результатів. Інше дослідження [4] використовує методи глибокого навчання для оцінювання залишкової ємності батареї на підставі даних про попередні цикли заряджання – розряджання.

Отже, актуальність цього дослідження полягає в оцінюванні ефективності застосування методів машинного навчання для прогнозування залишкового пробігу електромобіля з урахуванням деградації батареї. Мета роботи – аналіз відомих підходів та визначення оптимальних методів для підвищення точності прогнозів, що сприятиме покращенню користувацького досвіду та ефективності експлуатації електромобілів.

Огляд сучасних джерел інформації за тематикою публікації

Прогнозування залишкового пробігу електромобіля (EV) є важливим аспектом для підвищення довіри користувачів та оптимізації енергоспоживання. Традиційно для цієї мети використовували аналітичні моделі, ґрунтовані на фізичних характеристиках транспортного засобу та умовах експлуатації. Ці моделі враховують такі параметри, як маса автомобіля, аеродинамічний опір, профіль маршруту та стиль водіння, для оцінювання енергоспоживання та прогнозування пробігу.

З розвитком технологій з'явилися статистичні та машинно-навчальні підходи до прогнозування пробігу. Статистичні моделі аналізують історичні дані про енергоспоживання та пробіг, виявляючи закономірності та тренди. Моделі машинного навчання, такі як нейронні мережі та методи глибокого навчання, використовують великі обсяги даних для навчання та здатні враховувати складні нелінійні залежності між різними факторами, що впливають на пробіг [5].

Сучасні дослідження також зосереджуються на поєднанні фізичних моделей із алгоритмами машинного навчання для підвищення точності прогнозів. Такий гібридний підхід дає змогу враховувати як теоретичні аспекти енергоспоживання, так і реальні дані експлуатації [6].

Стан здоров'я батареї (State of Health, SoH) є критичним показником, що відображає поточну ємність батареї стосовно її номінального значення. Точна оцінка SoH необхідна для ефективного управління енергоспоживанням та прогнозування залишкового пробігу EV.

Існують різні підходи до оцінювання SoH:

- Методи прямого вимірювання: передбачають лабораторні тести, такі як повний цикл заряджання – розряджання, для визначення залишкової ємності. Хоча ці методи точні, вони не придатні для реального часу через тривалість та складність процедур [7].
- Моделі на основі фізичних процесів: використовують математичні моделі, що описують електрохімічні процеси в батареї. Вони дають змогу оцінювати SoH на основі вимірних параметрів, таких як напруга, струм та температура. Однак точність цих моделей залежить від якості параметризації та може знижуватися через індивідуальні відмінності між батареями.
- Дані, що ґрунтуються на методах машинного навчання: застосовують алгоритми, які навчаються на історичних даних для прогнозування SoH. Ці методи можуть враховувати складні залежності між різними параметрами та забезпечувати високу точність оцінки. Проте їх ефективність залежить від якості та обсягу навчальних даних [8].

Машинне навчання відіграє все важливішу роль у прогнозуванні енергоспоживання електромобілів. Алгоритми машинного навчання можуть аналізувати великі обсяги даних, ураховуючи інформацію про стиль водіння, дорожні умови, погодні фактори та стан батареї, для створення моделей, що точно передбачають енергоспоживання та залишковий пробіг [9].

Зокрема, глибокі нейронні мережі використовують для моделювання нелінійних залежностей між вхідними параметрами та енергоспоживанням. Ці моделі можуть адаптуватися до індивідуальних особливостей водія та умов експлуатації, що підвищує точність прогнозів.

Крім того, методи машинного навчання застосовують для оцінювання SoH батареї на основі експлуатаційних даних. Це дає змогу в реальному часі відстежувати стан батареї та враховувати його, прогножуючи пробіг [10].

Незважаючи на значний прогрес у сфері прогнозування пробігу EV та оцінки SoH, існують певні прогалини, що спонукають до подальших досліджень:

- Інтеграція різних джерел даних: більшість навних моделей не враховують одночасно всі фактори, що впливають на енергоспоживання та деградацію батареї, такі як погодні умови,

топографія маршруту та індивідуальний стиль водіння. Інтеграція цих даних може покращити точність прогнозів.

- Універсальність моделей: багато моделей створюють для конкретних типів транспортних засобів або умов експлуатації, що обмежує їх застосування. Розроблення універсальних моделей, здатних адаптуватися до різних сценаріїв, є актуальним завданням.

Цілі та проблеми дослідження

Зважаючи на зростання кількості електромобілів (EV) на дорогах та потребу в надійному прогнозуванні залишкового пробігу, актуальним є вивчення впливу деградації батареї на точність таких прогнозів. Більшість наявних моделей зосереджуються переважно на миттєвих експлуатаційних параметрах, нехтуючи динамічним зниженням ємності акумулятора з часом, що в умовах реального трафіку може призводити до значних похибок.

Метою цього дослідження є дослідження впливу інтеграції індикаторів деградації батареї, зокрема стану здоров'я (SoH), у моделі машинного навчання для підвищення точності прогнозування залишкового пробігу електромобіля.

Виклад основного матеріалу

У цьому розділі докладно висвітлено побудову моделі прогнозування залишкового пробігу електромобіля з урахуванням деградації батареї. Описано джерела даних, методи формування ознак, вибрані алгоритми машинного навчання, процедури навчання моделей та методи оцінювання їх ефективності.

Для дослідження використано дані із кількох відкритих джерел. Основним джерелом даних про деградацію батареї слугував набір EVBattery, який містить історію зарядно-розрядних циклів, напругу, струм, температуру, State of Charge (SOC), а також оцінку State of Health (SoH) батарей [1]. Для доповнення інформації про експлуатаційні умови та телеметрію транспортних засобів залучено дані Vehicle Energy Dataset (VED), які містять інформацію про швидкість руху, GPS-координати, зміну висоти маршруту, температуру навколишнього середовища та споживання енергії [2]. Використано також набір EVIoT-PredictiveMaint, що містить додаткову інформацію про режим експлуатації електромобілів [3]. З метою уніфікації аналізу для дослідження було вибрано поїздки завдовжки 250–300 км, що дало змогу оцінити точність прогнозування залишкового пробігу в умовах тривалих поїздок.

Попереднє опрацювання даних передбачало видалення записів з відсутніми значеннями, що перевищували 10 % обсягу записів, синхронізацію часових рядів з частотою вибірки 1 секунда, нормалізацію числових ознак за допомогою стандартизації:

$$x_{norm} = \frac{x - \mu}{\sigma}, \quad (1)$$

де μ – середнє значення ознаки у вибірці; σ – стандартне відхилення. Також здійснено агрегацію даних для узгодження різних часових інтервалів реєстрації через обчислення середніх, медіанних значень та стандартного відхилення для добових і годинних проміжків.

На основі аналізу літературних джерел і попередніх досліджень визначено перелік ознак, які формують вхідний простір моделей. Ознаки поділено на три групи:

- Операційні параметри: SOC, швидкість руху, струм заряджання / розряджання, напруга та температура батареї. Ці параметри безпосередньо впливають на моментальне енергоспоживання та динаміку розряджання акумулятора [4].

- Параметри деградації: ключовим показником є SoH, що характеризує залишкову ємність батареї у відсотках від номінальної. Додатково враховано кількість завершених циклів заряджання – розряджання, середню глибину розряду (DoD) та середньодобову температуру експлуатації [5].

- Контекстуальні параметри: профіль маршруту (висота, нахил), погодні умови (температура повітря, вологість) та середня швидкість на маршруті [6].

Такий вибір ознак дає змогу враховувати як фізичні характеристики батареї, так і експлуатаційні умови, які впливають на її деградацію та актуальну енергетичну ефективність транспортного засобу.

У дослідженні розглянуто моделі, побудовані на основі як традиційних деревних методів, так і нейронних мереж. Для аналізу вибрано такі алгоритми:

- XGBoost – алгоритм градієнтного бустингу на деревних структурах, що демонструє високу точність під час роботи з табличними даними [7].

- LSTM (Long Short-Term Memory) – рекурентна нейронна мережа, яка добре моделює часові залежності у даних, що важливо для аналізу історії деградації батареї та динаміки енергоспоживання [8].

- Гібридна модель XGBoost+LSTM – об'єднання можливостей обох підходів для одночасного врахування короткотермінових та довготермінових ефектів експлуатації [9].

Формально прогнозована модель залишкового пробігу $\hat{R}$ для базової моделі визначається такою функцією:

$$\hat{R}_{base} = f(X_{oper}, X_{context}), \quad (2)$$

де X_{oper} – вектор операційних параметрів, $X_{context}$ – вектор контекстуальних параметрів.

Розширена модель враховує додатковий вплив деградації:

$$\hat{R}_{SoH} = f(X_{oper}, X_{context}, X_{degrad}), \quad (3)$$

де X_{degrad} – параметри деградації, зокрема SoH.

Для навчання моделей використано поділ набору даних у співвідношенні 60 % для тренування, 20 % – для валідації та 20 % – для тестування остаточної точності. Для забезпечення стабільності оцінок застосовано методику перехресної перевірки (k-fold cross-validation) з кількістю блоків $k = 5$.

Якість моделі оцінювали за допомогою класичних метрик:

- MAE (Mean Absolute Error) – середня абсолютна похибка:

$$MAE = \frac{1}{n} \sum_{i=1}^n |R_i - \hat{R}_i|, \quad (4)$$

- RMSE (Root Mean Squared Error) – середньоквадратична похибка:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_i - \hat{R}_i)^2}, \quad (5)$$

- Коефіцієнт детермінації (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^n (R_i - \hat{R}_i)^2}{\sum_{i=1}^n (R_i - \bar{R})^2}, \quad (6)$$

Після навчання та тестування моделей отримано результати, наведені у таблиці.

Порівняння точності моделей

Модель	MAE, км	RMSE, км	R^2
XGBoost (без SoH)	8,2	11,5	0,82
XGBoost (з SoH)	6,7	9,3	0,89
LSTM (без SoH)	7,5	10,8	0,85
LSTM (з SoH)	5,9	8,7	0,91

На рис. 1 наведено порівняння розподілу похибок прогнозування, що демонструє помітне зменшення помилок у разі урахування SoH у моделі.

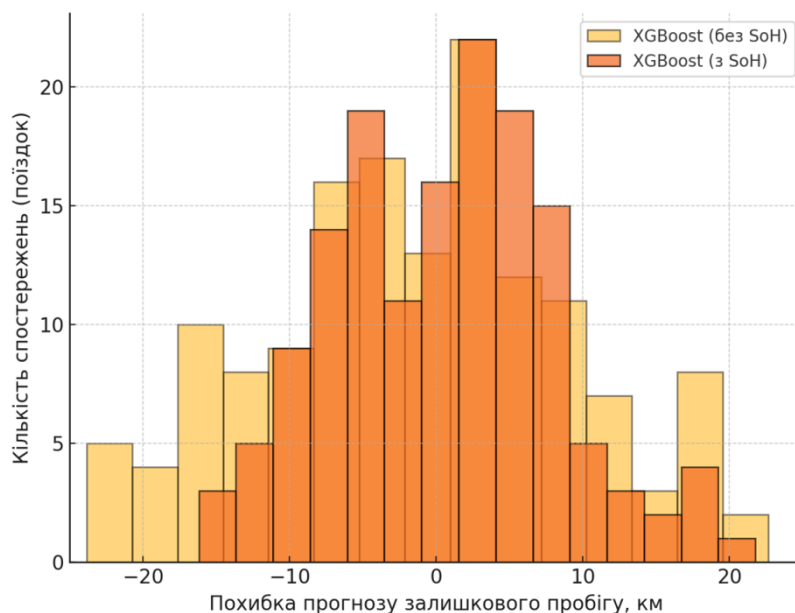


Рис. 1. Розподіл похибок прогнозу для моделей без SoH і з SoH

Графік демонструє, що модель з урахуванням SoH має помітно вужчий розподіл похибок, тобто абсолютна похибка більшості прогнозів менша. Вужча гістограма свідчить про меншу дисперсію помилок і вищу стабільність моделі.

Особливо важливо, що модель із SoH практично позбавлена великих відхилень (хвости розподілу зменшені), що свідчить про зниження ризику критичних помилок у прогнозі пробігу, які можуть призвести до розрядження батареї на маршруті.

На другому графіку (рис. 2) подано співвідношення фактичного та прогнозованого залишкового пробігу для моделі з урахуванням SoH.

- Кожна точка відповідає одному прогнозу для конкретної поїздки.
- Червона пунктирна діагональ відповідає ідеальному прогнозу (прогноз точно дорівнює фактичному значенню).

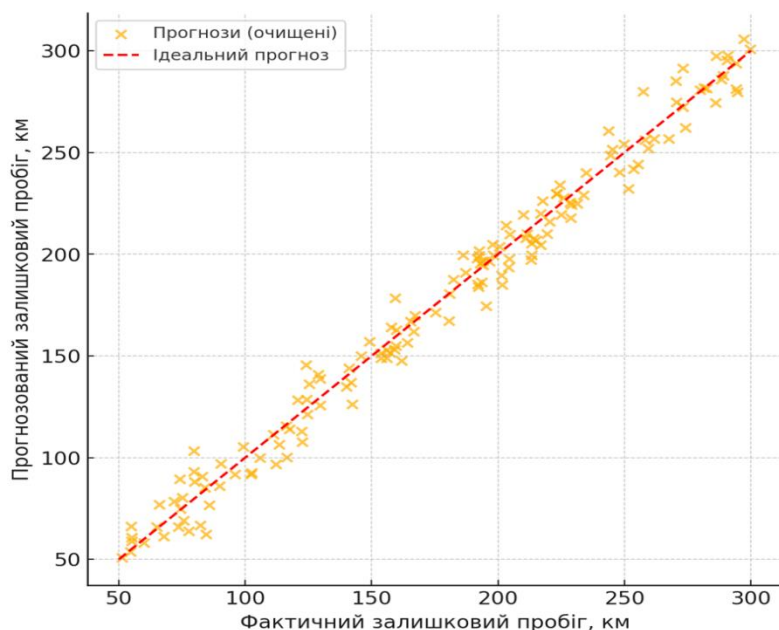


Рис. 2. Діаграма розсіювання фактичного та прогнозованого пробігу (модель з SoH).

Розташування точок близько до діагоналі свідчить про високу точність моделі. Як видно з графіка (рис. 2), більшість прогнозів містяться у межах невеликого відхилення від ідеальної лінії, що додатково підтверджує ефективність введення показника SoH до моделі.

Результати та обговорення

У ході дослідження було побудовано дві групи моделей прогнозування залишкового пробігу електромобіля: базові моделі, що враховують лише операційні та контекстуальні фактори експлуатації, та розширені моделі, які додатково інтегрують показники деградації батареї, зокрема State of Health (SoH).

Результати оцінювання точності моделей наведено в таблиці. Як видно з отриманих даних, введення SoH до складу вхідних параметрів істотно підвищило точність прогнозування залишкового пробігу. Зокрема, для моделі XGBoost середня абсолютна похибка (MAE) зменшилась з 8,2 км до 6,7 км, а для LSTM – з 7,5 км до 5,9 км. Аналогічно покращились показники визначення RMSE та коефіцієнта детермінації R^2 .

Такі результати підтверджують гіпотезу, що деградація батареї є істотним фактором, який необхідно враховувати для точного прогнозування запасу ходу. Особливо важливо, що моделі з урахуванням SoH зменшили кількість великих негативних відхилень у прогнозах, що критично для забезпечення надійності електромобіля під час планування тривалих поїздок.

Додатковий аналіз розподілу похибок наведено на гістограмі (рис. 1). Модель з урахуванням SoH демонструє компактніший розподіл похибок, зменшуючи частку випадків із великими відхиленнями – понад ± 20 км. Це свідчить про підвищену стабільність прогнозів за різних умов експлуатації, оскільки найбільші ризики для користувача пов'язані саме із великими відхиленнями у бік недооцінювання пробігу.

Зокрема, кількість випадків з похибкою понад -20 км зменшилась із 14 випадків для базової моделі до лише 3 випадків для моделі з урахуванням SoH (див. таблицю розподілу похибок). Така стабілізація прогнозу є важливою перевагою для практичного застосування в системах енергоменеджменту EV.

Діаграма розсіювання на рис. 2 демонструє точність моделі із SoH. Більшість точок розташовані поблизу лінії ідеального прогнозу (пунктирна червона лінія), що вказує на високу відповідність між прогнозованим та фактичним залишковим пробігом. Лише поодинокі випадки демонструють істотні відхилення, що корелює із розподілом похибок.

Отримані результати підтверджують, що урахування деградації батареї істотно поліпшує якість моделей прогнозування пробігу, особливо в умовах тривалого терміну експлуатації, коли накопичений ефект втрати ємності батареї стає визначальним.

Особливо важливо відзначити такі аспекти:

- Практична значущість: моделі з SoH забезпечують зменшення критичних помилок, які потенційно можуть призводити до розряджання батареї на маршруті.
- Гнучкість моделей: запропонований підхід легко масштабується на різні типи транспортних засобів за умови наявності якісної телеметрії та моніторингу SoH.
- Обмеження: актуальною проблемою залишається наявність повноцінних відкритих наборів даних, що містять одночасно і телеметрію, і дані про SoH, а також виклики щодо реального онлайн-оцінювання SoH під час руху.

Висновки

У статті досліджено завдання прогнозування залишкового пробігу електромобіля з урахуванням деградації батареї за допомогою інтеграції індикатора стану здоров'я (SoH) до моделей машинного навчання. Порівняльний аналіз базових та розширених моделей дав підставу для таких узагальнень:

- Введення показника SoH до складу вхідних ознак забезпечує істотне зниження середньої абсолютної похибки (MAE) та кореневої середньоквадратичної похибки (RMSE), а також підвищення коефіцієнта детермінації (R^2), що свідчить про підвищення загальної точності моделі.

- Особливо значущим є зменшення кількості великих відхилень у прогнозах, що знижує ризик неправильного планування маршруту через недооцінку залишкового пробігу.

- Інтеграція параметрів деградації дає змогу формувати стабільніші моделі, які демонструють надійні результати в умовах різних профілів експлуатації, ураховуючи зміни температури, стилю водіння та тривалості використання батареї.

- Запропоновану методику можна використовувати як компонент сучасних систем енергоменеджменту електромобілів, що забезпечить точнішу інформацію для водіїв, диспетчерських служб та систем планування маршрутів.

У подальших дослідженнях доцільно розглянути:

- Побудову персоналізованих моделей для окремих користувачів з урахуванням стилю водіння.

- Застосування глибших моделей гібридного навчання, що поєднують фізичні моделі деградації із нейромережевими структурами.

- Організацію експериментів із розширеними реальними наборами даних, які враховують широкий спектр кліматичних зон та умов експлуатації.

Отримані результати є важливим кроком у розвитку систем високоточних прогнозів енергоспоживання та залишкового пробігу для електромобілів, що сприяє підвищенню надійності та комфорту їх використання у повсякденній експлуатації.

Перелік використаних джерел

[1] Waldmann, T., Wilka, M., Kasper, M., Fleischhammer, M., and Wohlfahrt-Mehrens, M. (2014). Temperature dependent ageing mechanisms in Lithium-ion batteries – A Post-Mortem study. *Journal of Power Sources*, 262, 129–135. <https://doi.org/10.1016/j.jpowsour.2014.03.112>

[2] S. Saxena (2019). Accelerated degradation model for C-rate loading of lithium-ion batteries /et al. *International Journal of Electrical Power & Energy Systems*, Vol. 107, 438–445. <https://doi.org/10.1016/j.ijepes.2018.12.016>.

[3] Murgai, S., Bhagwat, H., Dandekar, R. A., Dandekar, R., and Panat, S. (2024). A Scientific Machine Learning Approach for Predicting and Forecasting Battery Degradation in Electric Vehicles. arXiv preprint *arXiv:2410.14347*. <https://doi.org/10.48550/arXiv.2410.14347>

[4] Lanubile, A., Bosoni, P., Pozzato, G., Allam, A., Acquarone, M., and Onori, S. (2024). Domain knowledge-guided machine learning framework for state of health estimation in Lithium-ion batteries (Version 1). *arXiv*, <https://doi.org/10.48550/ARXIV.2409.14575>

[5] Oh G., Leblanc D. J., Peng H. (2020). Vehicle Energy Dataset (VED), A Large-Scale Dataset for Vehicle Energy Consumption Research. *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11. <https://doi.org/10.1109/tits.2020.3035596>

[6] Zhang, S., Fatih, D., Abdulqadir, F., Schwarz, T., and Ma, X. (2022). Extended Vehicle Energy Dataset (eVED): an enhanced large-scale dataset for deep learning on vehicle trip energy consumption. arXiv preprint *arXiv:2203.08630*. Отримано з <https://arxiv.org/abs/2203.08630>

[7] An open tool for creating battery-electric vehicle time series from empirical data, emobpy / C. Gaete-Morales et al. *Scientific Data*, 2021, Vol. 8, No. 1. <https://doi.org/10.1038/s41597-021-00932-9>

[8] EV Battery Health Insights: Data From 10,000 Cars / Geotab. Geotab. URL: <https://www.geotab.com/blog/ev-battery-health/>

[9] Using Internal Data for Electric Vehicles: A Deep Dive. EV Fleet Solutions for Charging Management Systems Ampcontrol. URL: <https://www.ampcontrol.io/post/a-deep-dive-into-internal-vehicle-data-and-how-it-is-used-for-electric-vehicles>

[10] EVIoT-PredictiveMaint Dataset (2023). Kaggle. URL: <https://www.kaggle.com/datasets/datasetengineer/eviot-predictivemaint-dataset>

**MACHINE LEARNING-BASED PREDICTION OF ELECTRIC VEHICLE REMAINING RANGE
WITH CONSIDERATION OF BATTERY DEGRADATION**

Received: August 14, 2025 / Revised: August 25, 2025 / Accepted: September 15, 2025

© Dropa R.*, 2025

Abstract. Accurate prediction of the remaining driving range in electric vehicles (EVs) is critical for efficient trip planning, reducing the risk of battery depletion, and improving user experience. One of the significant challenges in achieving high prediction accuracy is battery degradation, which gradually reduces battery capacity and impacts the vehicle's range. This study uses machine learning algorithms to investigate the impact of incorporating battery degradation – expressed through the State of Health (SoH) indicator – into range prediction models. A comparative analysis is conducted between models with and without SoH input based on real-world EV telemetry data. The results demonstrate that integrating degradation metrics into the prediction model significantly enhances the accuracy of remaining range estimation. The findings may benefit developers of energy management systems and EV software applications.

Keywords: electric vehicle, range prediction, machine learning, battery degradation, state of health (SoH), energy consumption, telemetry data.

* Corresponding author



© The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution Licence 4.0 (<https://creativecommons.org/licenses/by/4.0/>)