

## **Title: Machine learning for Churn Prediction in Telematics for Cartrack (2019–2025):**

### **Introduction & scope**

We will state the problem, B2B/B2C scope in telematics, aims, and contributions, and justify organisational relevance using Cartrack scale and context using Karooooo, N.D.

### **Method (PRISMA – explain)**

A PRISMA-aligned process will document search sources, strings, screening, and reasons for exclusion, ensuring transparency and reproducibility (Akl et al., 2024). Searches will prioritise Google Scholar and recent, domain-relevant work (2019+) with clear methods/metrics; preprints may be included where they add recency but will be critically appraised (Google, n.d.; Imani et al., 2025).

### **Algorithms & models' basics**

The review will compare logistic baselines, gradient-boosted trees (XGBoost/LightGBM) and DL for sequences/text; evidence generally favours tree ensembles on tabular data, with DL competitive when behaviour sequences are informative (Hasan, 2025; Imani et al., 2025).

### **Features & data design**

Feature groups reflect Cartrack's systems: subscriptions/contracts, client-level usage/telemetry, device health, tickets latency, billing/arrears, and repairs. Literature supports behavioural/transaction variables and early-warning signals for churn targeting (Bolaji et al., 2023). Design principles include strict time cut-offs to prevent leakage, monthly snapshots, label windows (e.g., churn within 90 days), and CRISP-DM-aligned preparation/feature stores for reproducibility (IBM, 2021).

### **Evaluation & validation**

Define the evaluation plan (PR-AUC, precision@K, cost curves) and probability calibration/assessment procedures to be applied across models (Văduva, Oprea, Niculae, Bâra, & Andreescu, 2024).

### **Explainability & governance**

State how SHAP-based explanations and documentation will be produced and connected to the CRISP-DM lifecycle for deployment readiness (IBM, 2021).

### **Gaps & agenda for telematics**

Outline how we will identify evidence gaps and inconsistencies and translate them into a research agenda specific to telematics and Cartrack (Hasan Mamun, 2025; Imani, Beikmohamadi, & Arabnia, 2025).

### **Implications for Cartrack**

We will describe how findings will be converted into a practical implementation roadmap (model stack, metrics, processes) aligned to CRISP-DM phases and operations (IBM, 2021).

### **Conclusion**

Summarise expected outputs (synthesis + roadmap) and how they support early detection and profitable interventions within Cartrack's B2B/B2C mix.

### **References**

Karooooo (N.D.) Welcome to Karooooo. Available from: <https://karooooo.com/> [Accessed 15 August 2025]

Akl, E. et al. (2024) Extension of the PRISMA 2020 statement for living systematic reviews (PRISMA-LSR): checklist and explanation. Available from: <https://www.bmj.com/content/387/bmj-2024-079183> [Accessed 14 August 2025]

Google (N.D.) Google Scholar. Available from: <https://scholar.google.com/> [Accessed 14 August 2025]

Imani, M., Beikmohamadi, A., & Arabnia, H. (2025) Customer Churn Prediction: A Review of Recent Advances, Trends, and Challenges in Conventional Machine Learning and Deep Learning. Available from: <https://www.preprints.org/manuscript/202503.1969/v2> [Accessed 15 August 2025]

Hasan Mamun, H. (2025) ADVANCEMENTS IN MACHINE LEARNING FOR CUSTOMER RETENTION: A SYSTEMATIC LITERATURE REVIEW OF PREDICTIVE MODELS AND CHURN ANALYSIS. Journal of Sustainable Development and Policy, 1(01), pp. 250–284. Available from: <https://jsdp-journal.org/index.php/jsdp/article/view/11> [Accessed 15 August 2025]

Bolaji, A., Chukwuma-Eke, E., Balogun, E., & Ogunsola, K. (2023) Improving Customer Retention Through Machine Learning: A Predictive Approach to Churn Prevention and Engagement Strategies. Available from: [https://www.researchgate.net/publication/390113580\\_Improving\\_Customer\\_Retention\\_Through\\_Machine\\_Learning\\_A\\_Predictive\\_Approach\\_to\\_Churn\\_Prevention\\_and\\_Engagement\\_Strategies](https://www.researchgate.net/publication/390113580_Improving_Customer_Retention_Through_Machine_Learning_A_Predictive_Approach_to_Churn_Prevention_and_Engagement_Strategies) [Accessed 15 August 2025]

IBM (2021) CRISP-DM Help Overview. Available from: <https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview> [Accessed 15 August 2025]

Văduva, A.-G., Oprea, S.-V., Niculae, A.-M., Bâra, A., & Andreescu, A.-I. (2024) Improving Churn Detection in the Banking Sector: A Machine Learning Approach with Probability Calibration Techniques. Available from: <https://www.mdpi.com/2079-9292/13/22/4527> [Accessed 15 August 2025]

Feedback:

Hello, Matthew, thank you for submitting the LR draft outline for comments.

I think you overall understood what was required at this stage and completed a satisfactory draft. While the format is correct, however, the outline is a bit synthetic to be really able to evaluate the effort, even though it is quite promising.

Here are some suggestions to improve and “populate” it for the graded submission in Week 7.

1. Work on your research questions (RQs) narrowing them down as much as you can. Stated or not explicitly, a literature review is always about research questions and the way they have been treated /explored by others before you. The clearer and more focused RQs are, the better the LR.
2. Remember that a LR is not just a collection of articles/sources. You need to construct a coherent narrative, and show you are able to offer critical analysis of them
3. The bulk of your references should be academic, supported by the appropriate industry/technical references. This depends a lot on the specific topic.
4. The purpose of the conclusion is to wrap up your ideas/discussion and reinforce the key concepts expressed in the body of the LR.

So, no original contributions are expected here, but you should highlight how the LR supports the research gaps you have identified when stating your research questions.

I look forward to reading the final version.