Doctoral Dissertation
Doctoral Program in Control and Computer Engineering (38.th cycle)

Data science methods for complex mobility systems

Andrea Avignone

* * * * *

Supervisors
Prof. Silvia Chiusano, Supervisor
Prof. Tania Cerquitelli Co-supervisor

Doctoral Examination Committee:
Prof. Elisa Quintarelli, Referee, Università di Verona
Prof. Andrea Acquaviva, Referee, Università di Bologna
Prof. Ruggero Gaetano Pensa, Università degli Studi di Torino
Prof. Oscar Romero, Universitat Politècnica de Catalunya
Prof. Paolo Garza, Politecnico di Torino

Politecnico di Torino
2026

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . .

Andrea Avignone

Turin, 2026

Summary

Mobility systems are rapidly evolving under the combined influence of technological innovation, digitalization, and the growing availability of data. These transformations have made transport networks increasingly interconnected and adaptive, but they have also introduced new layers of complexity. Information is collected from heterogeneous sources such as vehicles, sensors, ticketing systems, and monitoring platforms, producing massive and irregular data streams with strong spatial and temporal dependencies. Although such data offer new opportunities for understanding mobility processes, they are difficult to access, manage, and integrate. Their effective use requires analytical and computational frameworks capable of processing large-scale, fragmented information and extracting consistent patterns to support actionable knowledge.

This thesis investigates how data science can address these challenges through the development of data-driven models and computational methods that learn from empirical evidence and transform complex data into interpretable representations of mobility phenomena. The research explores several domains: *pedestrian–vehicle interaction* in autonomous driving contexts under mixed traffic conditions, *demand forecasting* for public and shared transport, and *performance analysis* in competitive racing. In addition, it includes a methodological contribution on *data modeling* for document-oriented databases and *distributed learning strategies* for embedded devices, aimed at enabling efficient computation in data-intensive mobility applications. Despite their diversity in scope, these studies share common analytical principles and face similar challenges to integrate heterogeneous data sources and capture temporal dynamics.

By integrating machine learning with behavioural analysis, the thesis develops methods that extract structure from complex data and quantify how mobility systems evolve and interact under real operating conditions. These approaches make it possible to assess interactions, analyze system dynamics, and evaluate performance under real operating conditions. The resulting models offer a structured understanding of how mobility systems function and adapt, providing a basis for data-informed decision-making and management. The work advances the role of data science for translating empirical evidence into reliable knowledge that supports safety, efficiency, and the sustainable development of modern mobility.