

# Decoding the Roads: An Infrastructure-Driven Machine Learning Approach to Predict Safety

Akanksha Agarwal<sup>1</sup>, Davy Janssens, Geert Wets, Tom Bellemans

Transportation Research Institute (IMOB), Hasselt University, [akanksha.agarwal@uhasselt.be](mailto:akanksha.agarwal@uhasselt.be)

Transportation Research Institute (IMOB), Hasselt University, [davy.janssens@uhasselt.be](mailto:davy.janssens@uhasselt.be)

Transportation Research Institute (IMOB), Hasselt University, [geert.wets@uhasselt.be](mailto:geert.wets@uhasselt.be)

Transportation Research Institute (IMOB), Hasselt University, [tom.bellemans@uhasselt.be](mailto:tom.bellemans@uhasselt.be)

## Abstract

In 2024, traffic-related injuries resulted in the loss of more than 20,000 lives across Europe (European Transport Safety Council, 2025). Belgium alone recorded approximately 470 fatalities (Statbel, 2025). For decades, road safety systems have relied heavily on historical crash data for safety interventions. Although such systems have demonstrated great potential, they require a significant number of crashes and conflicts to occur before any intervention can take place. Beyond being ethically tenuous, these methods often suffer from statistical insignificance in regions with low traffic volumes. Moreover, the traditional systems generally fail to account for the user's perception of safety. This is a crucial component of traffic safety research due to its correlation with road user behaviour.

The recent years have seen a fundamental shift from historically reactive systems towards more proactive and apologetic systems. The rapid advancements in machine learning and artificial intelligence have opened new horizons for proactively assessing infrastructural safety and road environments. We propose a methodological framework to assess the perceived safety of road segments. The primary goal of this study is to analyse the relationship between subjective perceptions and objective attributes of a road network. It presents an innovative approach to using the underlying information in the physical design of road networks to identify dangerous locations. Our primary objective is to explore the abilities of advanced machine learning models in estimating the safety of road segments across Flanders, Belgium. By leveraging data-driven learning, this study aims to design a tool that identifies latent risks before they manifest as crashes.

At the core of this approach lies a detailed dataset capturing both infrastructural and perceptual aspects of road networks. The dataset used for this study was curated by combining two primary data sources - Route2School and OpenStreetMap. The Route2School (R2S) dataset integrates the geocoded representation of road networks with safety labels, reflecting the subjective perceptions of road users. The dataset comprises home-to-school routes in approximately 80 cities across Belgium. These routes are stored as a series of road segments, each annotated with a safety label: SAFE, OCCASIONALLY\_UNSAFE, and UNSAFE. This dataset was enriched with infrastructural attributes sourced from OpenStreetMap. OpenStreetMap (OSM) is a public dataset built and maintained by a community of mappers. It contains spatial data about roads, buildings, railways, and other infrastructure features worldwide.

A systematic data preprocessing phase was completed, resulting in a training dataset comprising approximately 70,000 labelled points, each associated with nearly 400 potential infrastructural attributes. Although the dataset is significant in scale, the data exploration step revealed a substantial class imbalance. The set contains almost five times more safe locations than unsafe and occasionally unsafe locations. Additionally, as the OSM dataset is crowd-sourced, it is often inconsistent and incomplete, resulting in numerous missing values.

To address these gaps, we adopted a semi-supervised learning approach. The core of the methodology was a two-step learning framework that allowed the model to learn from both the observed values of OSM attributes and the inherent missingness patterns in the dataset. The first step involved training a denoising autoencoder in a self-supervised setting to capture intrinsic relationships in the dataset. The goal of this step

---

<sup>1</sup> Akanksha Agarwal. Tel: + 32 11 26 91 05  
E-mail Address: [akanksha.agarwal@uhasselt.be](mailto:akanksha.agarwal@uhasselt.be)

was to replace the raw, incomplete input variables with a more compact and information-rich feature set. In the second step, the methodology focused on leveraging the resulting feature set to train a supervised model to predict the predefined safety labels.

As a preliminary step in developing the framework, we implemented a data preparation pipeline. Firstly, all the OSM attributes containing metadata, such as names, survey dates, and other descriptive tags, were removed to ensure the dataset contained only analytical attributes. Since a considerable fraction of the attribute columns contains missing values, only the columns containing fewer than 5 non-null values were dropped to retain maximum information. This resulted in a cleaner, more structured dataset with 270 OSM attribute columns as features and a safety label representing the target variable. A crucial task at this stage involved extracting a missingness mask for both the numerical and categorical columns. The underlying idea of this process was to broaden the input dataset by adding information about the innate patterns in the data missingness. The encoded input, along with this missingness mask, was provided as input to the autoencoder.

Considering the central objective was to develop a model for dealing with the real-world incompleteness in the dataset, we employed a special type of autoencoder called a Denoising Autoencoder (DAE). The core principle of DAE is to extract meaningful representations from a corrupted input. Therefore, a DAE architecture was trained on the input dataset in a self-supervised manner. The observed values in the input dataset were corrupted randomly, and the DAE was trained to reconstruct the input from this corrupted version. A hybrid activation function comprising a linear transformation and a Rectified Linear Unit (ReLU) was employed in all embedded layers. The input data was corrupted by randomly dropping observed values to simulate missingness and train the DAE to learn hard samples. Given the mixed data types in the input, the autoencoder was optimised using a combination of Mean Squared Error (MSE) and Cross-Entropy loss. The autoencoder model was then trained to compress the input data into a 16-dimensional latent vector representation. This vector was then leveraged as the sole input for the subsequent classification algorithm, effectively decoupling the feature engineering and predictive modelling stages.

The supervised learning task was implemented using a Gradient Boosting (GB) Classifier. It is an ensemble learning technique that iteratively builds a strong predictive model by adding multiple weak learners, generally decision trees. It is an instance of the boosting paradigm wherein each weak learner improves the errors of the previous learner, thereby reducing the overall prediction error. The model selection was motivated by the established capabilities of GB classifiers in approximating non-linear relationships and mitigating the effects of class imbalance. The selection was further validated through comparisons with standard baseline models, such as Logistic Regression, where the Gradient Boosting classifier consistently demonstrated better results. The classification model was implemented using the Scikit-learn library. A stratified 5-fold cross-validation step was also implemented to fine-tune the hyperparameters based on a GridSearch algorithm and maximise the model's F1 score.

To evaluate the generalisation capabilities of our final model, its performance was estimated on unseen test data. The model achieved an overall accuracy of 77.1% when evaluated on the hold-out test set. This indicates a substantial correlation between subjective safety and infrastructural attributes. However, because of a considerable class imbalance in the dataset, accuracy can be a misleading metric. Hence, the model was further validated using more granular metrics, like precision, recall, and F1 score. The macro precision score of 0.73 demonstrated the model's efficiency in identifying samples belonging to the minority class. The macro recall score for the model was relatively lower (0.57), resulting in a macro F1 score of 0.61.

The core purpose of this research was to determine whether perceived safety can be modelled as a function of objective link attributes. The results confirm this hypothesis and the merits of integrating such disparate data sources. Our analysis revealed the significant potential of our model to accurately identify safe, less safe, and downright unsafe locations. The study also highlighted that, though not all unsafe locations in the network were detected, those that the model did identify were indeed unsafe in 73% of the cases. This can be valuable for road safety engineers and practitioners, as it helps reduce the resources spent on verifying false positives. The architectural reliance on OSM data is another key strength of the presented framework. The utilisation of globally standardised and widely available geospatial data assures easy implementation of our model in new regions, requiring only the integration of local safety labels to calibrate the model. Once calibrated, the model can also be used as a tool to assess safety in regions where crash data may be missing or underreported. Beyond static risk mapping, the proposed model can also be extended to safe routing systems. The presented model can be utilised to score entire networks, thereby supporting integration with routers to generate safety-aware routes. This application is particularly relevant for vulnerable road users, such as cyclists or children, allowing the proactive avoidance of high-risk links.

Despite exhibiting notable performance, the suggested framework is not without limitations. As evident from the attained F1 score, the model requires further optimisation to learn complex patterns and identify all unsafe locations. This implies that although the model is an effective screening tool for identifying high-risk nodes, it cannot yet completely replace expert audits and site inspections. Another constraint of the proposed methodology is the intrinsic lack of interpretability of autoencoders. While the latent layers are highly efficient in capturing complex non-linear relationships, they make it challenging to isolate the specific OSM attributes that contribute the most to an unsafe classification.

While there are certain constraints, the proposed methodology shows great potential and lays the groundwork for future research. Although the specific weights of the model are not directly transferable, the methodological framework itself can be readily transferred with minimal calibration. A possible further experiment would be to extend the model to generate safety scores instead of labels. This would facilitate the integration of the model with routing systems to generate safe routes. As the presented model is trained to estimate subjective perceptions of safety, a follow-up investigation could introduce objective measures of safety into the model for a more holistic estimation of safety. In conclusion, the study presents a promising data-driven approach for predicting public perceptions of safety based on the physical attributes of a road network.

**Keywords:** Subjective Safety; OpenStreetMap; Predictive Modelling; Semi-supervised classification

## Acknowledgements



Funded by the  
European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101119590. The opinions expressed in this document reflect only the authors' view and in no way reflect the European Commission's opinions. The European Commission is not responsible for any use that may be made of the information it contains.

## References

European Transport Safety Council. (2025, June 24). *19th Annual Road Safety Performance Index (PIN Report)*. ETSC. Retrieved from <https://etsc.eu/19th-annual-road-safety-performance-index-pin-report/>

Statbel. (2025, June 24). *470 road traffic deaths in Belgium in 2024*. Belgian Federal Government. Retrieved from <https://statbel.fgov.be/en/news/470-road-traffic-deaths-belgium-2024>

OpenStreetMap contributors. (2025). *OpenStreetMap* [Data set]. OpenStreetMap Foundation. Retrieved from <https://www.openstreetmap.org>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830. <https://jmlr.org/papers/v12/pedregosa11a.html>

Route2School. (n.d.). *Route2School*. <https://www.route2school.be/>