A science-based approach to classifying light vehicles in Europe:

Methodology and case study

Authors

Lorenzo **Laveneziana**[1]
Andres L. **Marin**[2]
Dermot **O'Brien**[3]
Matteo **Prussi**[1]
Georgios **Fontaras**[3]

[1] Politecnico di Torino
[2] Universitat Politecnica de Valencia
[3] Joint Research Centre

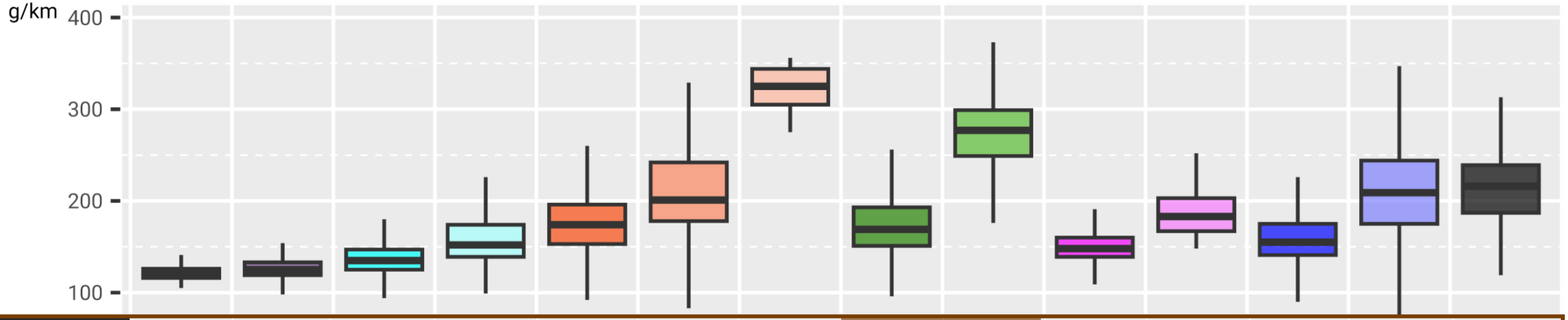
30th September 2024

Politecnico di Torino

Contex providing scientific bases for vehicle classifications

Variation of declared WLTP emissions in different segments

ICE - Gasoline



Extended Eurocar	A City	B Small	C Medium	D Large	E Executive	F Luxury	I Super Luxury	G Sport	H Supercar	J MPV (small)	K MPV (large)	L SUV (small)	M SUV (large)	N Van
Eurocar Segment	**A**	**B**	**C**	**D**	**E**	**F**		**S**		**MPV**		**SUV**		**-**
NCAP	City		Small	Large	Executive			Coupè	-	*MPV (small)*	*MPV (large)*	*SUV (small)*	*SUV (large)*	Van
ACEA	*Small*		*Lower Medium*	*Upper Medium*	*Luxury*			-		*MPV*		*SUV*		-

- Choice of classification scheme has **profound impact on segment emissions**
- **Classification criteria** are often **opaque**, vehicles are classified based on the similarity **with other vehicles on the market**

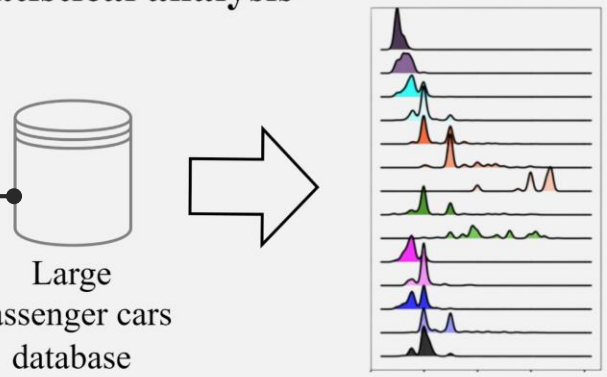
- Can we provide a **transparent, standardised and reproducible methodology** for vehicles classification?
- Can we **explain** current classification through **technical specifications**?

Methodology: overview and data sources

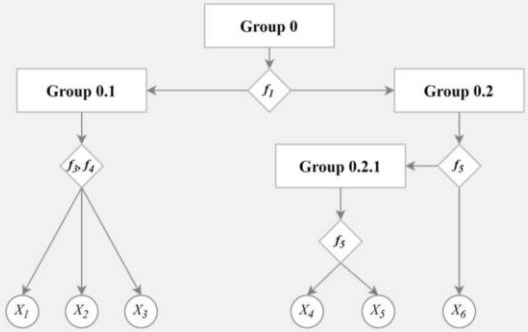
300’000+ vehicles

Features Selection

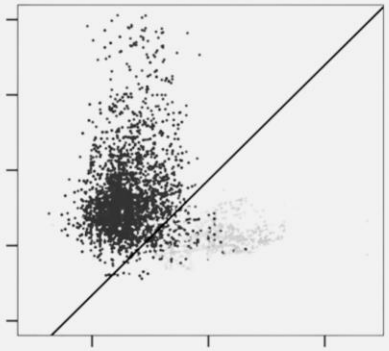
1. Statistical analysis



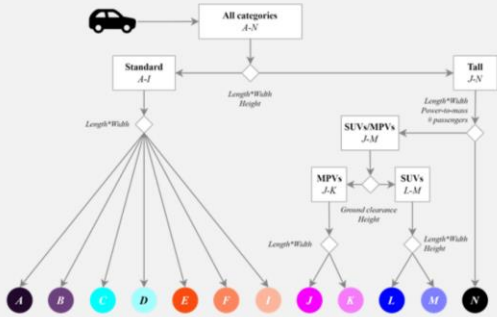
2. Dimensionality reduction



3. Determination of boundaries



4. Stepwise classification algorithm

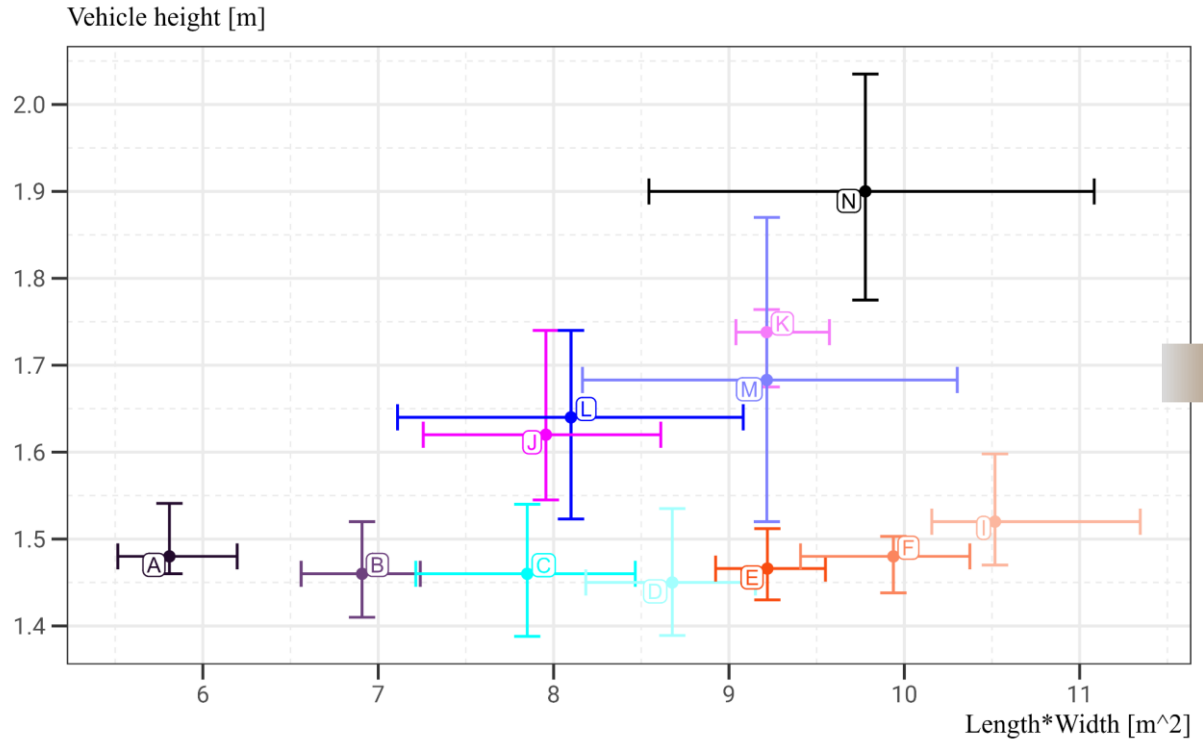


Definition of classification criteria

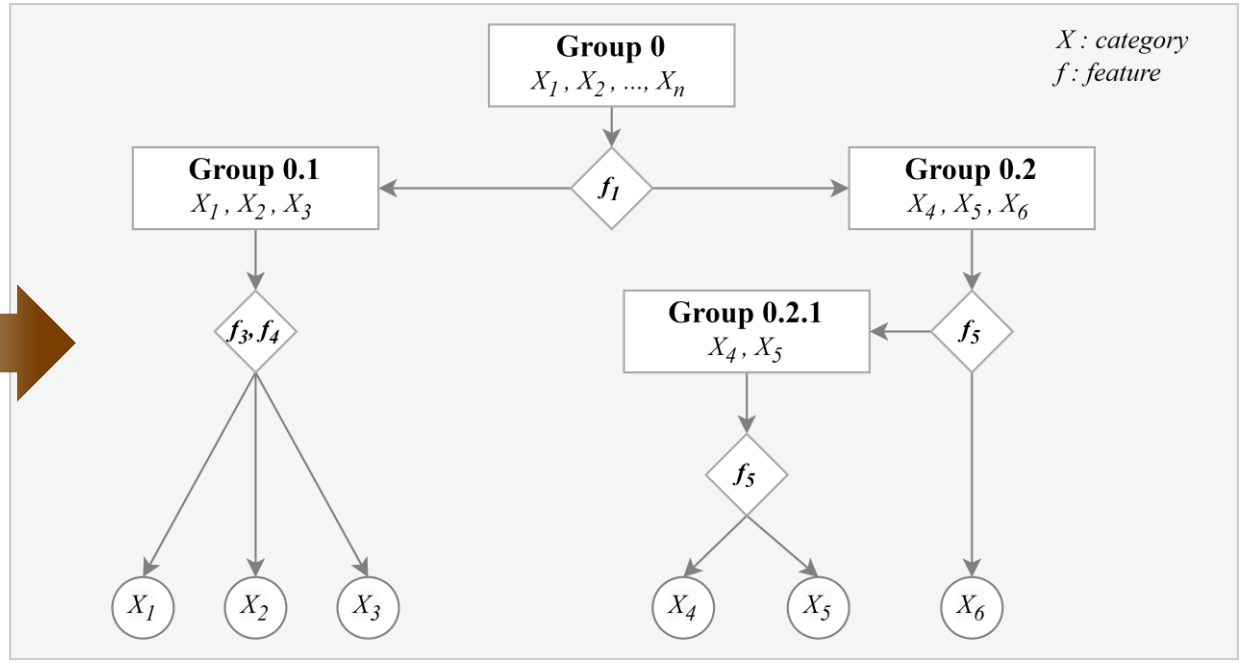
Development and verification classification algorithm

Methodology: features selection

Height and Length x Width between standard and tall vehicles



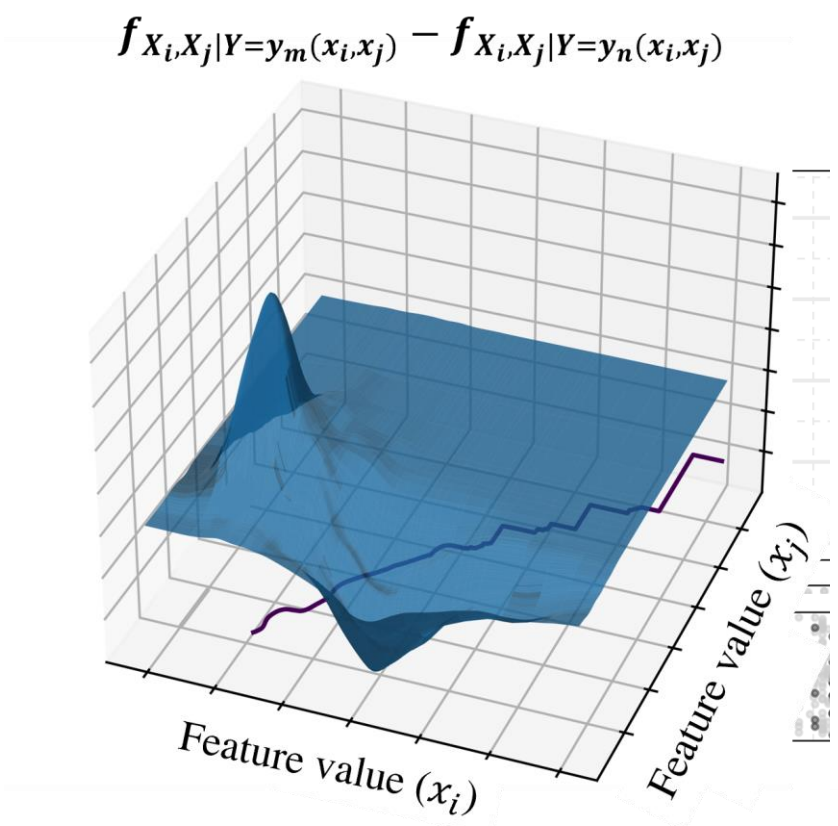
Stepwise classification approach

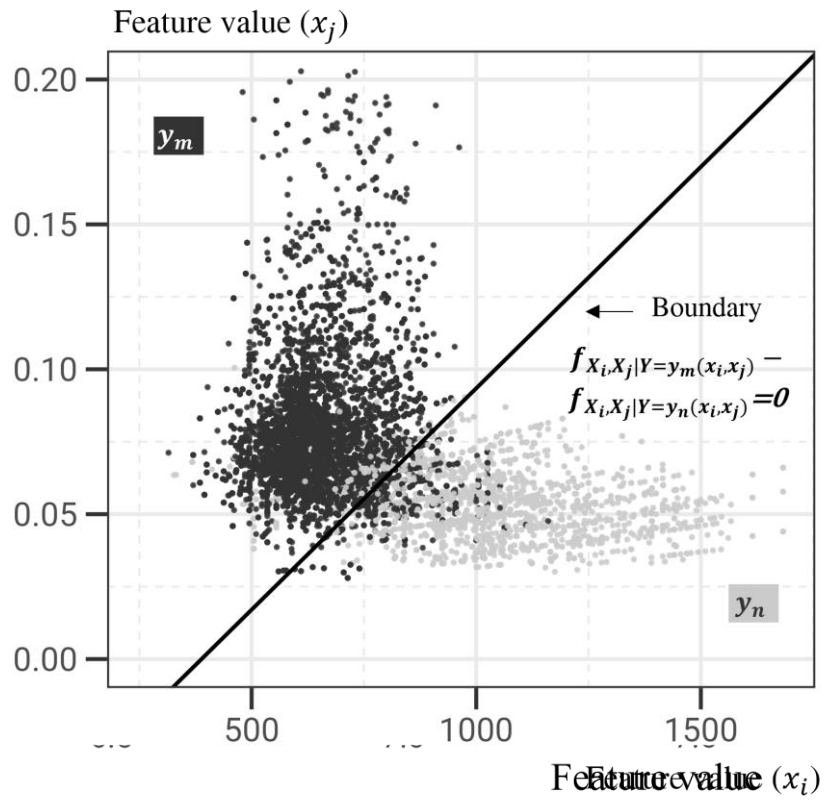


Methodology: determination of boundaries

«Classification criteria** are often **opaque**, vehicles are classified based on the **similarity with other vehicles on the market»

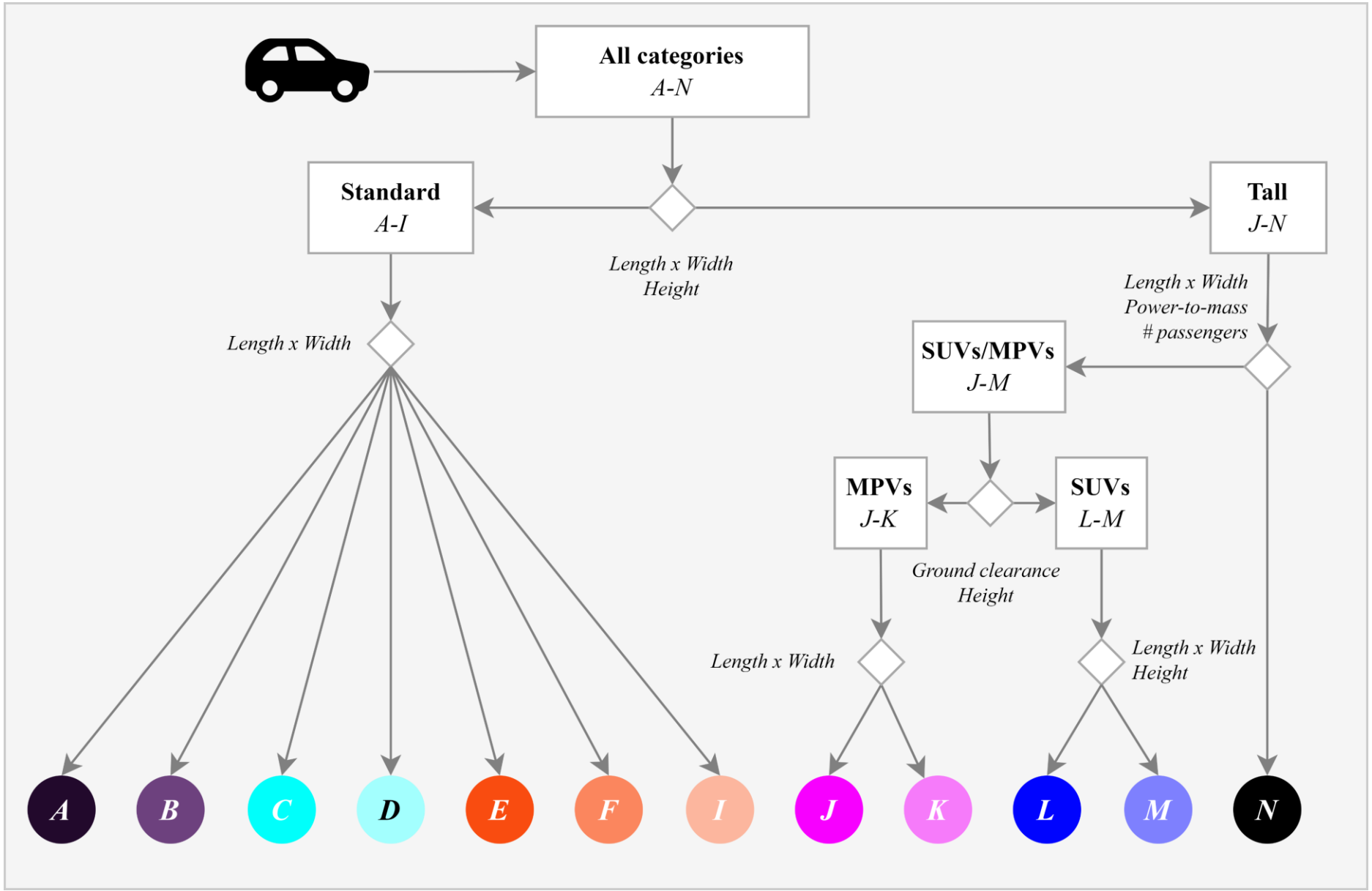
- The principle of **similarity** is **well-established** in statistic literature
- It can be formulated in mathematical terms using **Probability Density Functions (PDFs)**
- PDFs are defined for each feature and segment using the **distribution of the feature** in that segment
- The higher the value of the PDF, the higher the probability that the vehicle belongs to that segment
- The **boundary** between two groups can be drawn when PDFs have the same value
- The reasoning can be extended to n-dimensions





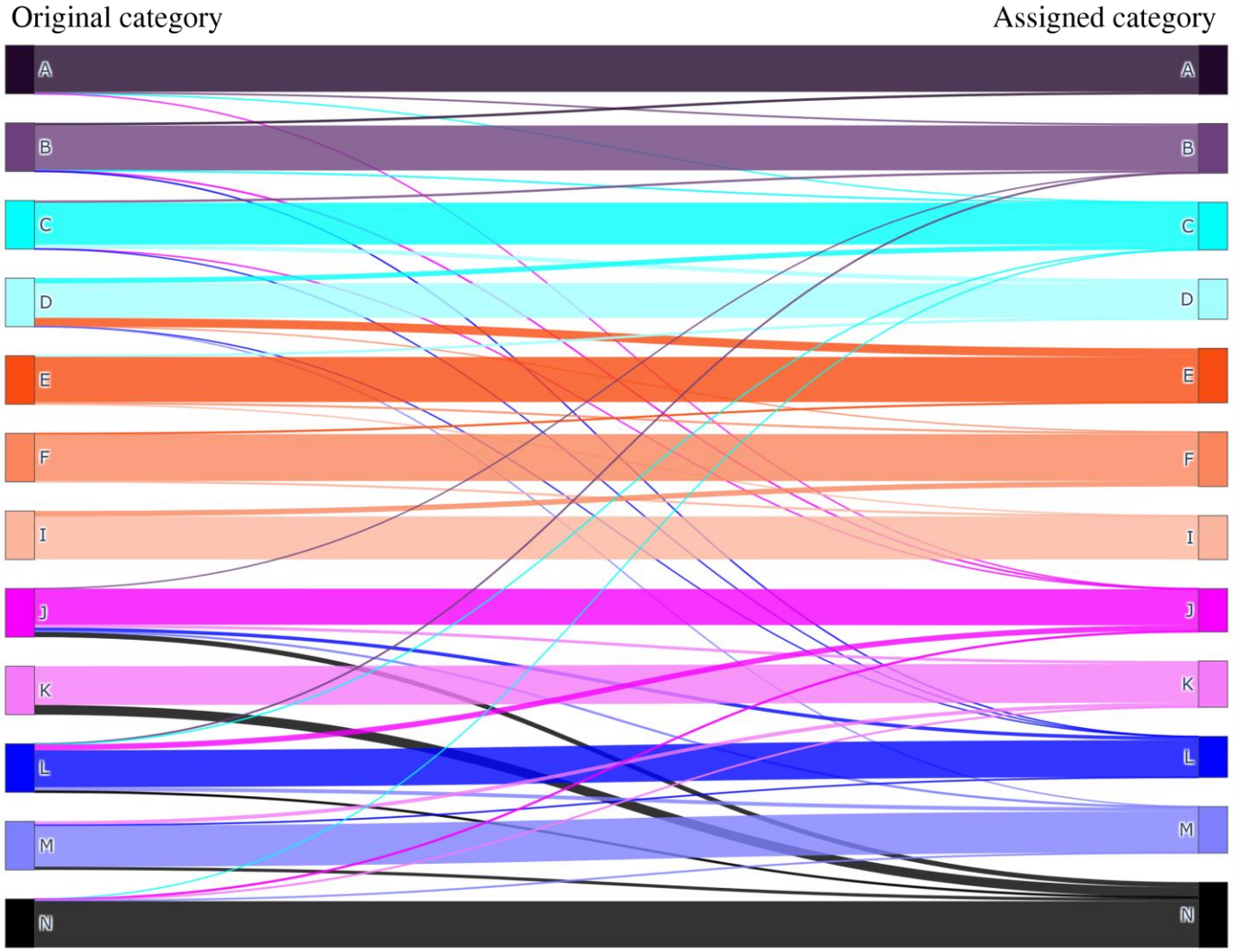
Results

stepwise classification algorithm



Results

classification accuracy



Comparison with Machine Learning algorithms

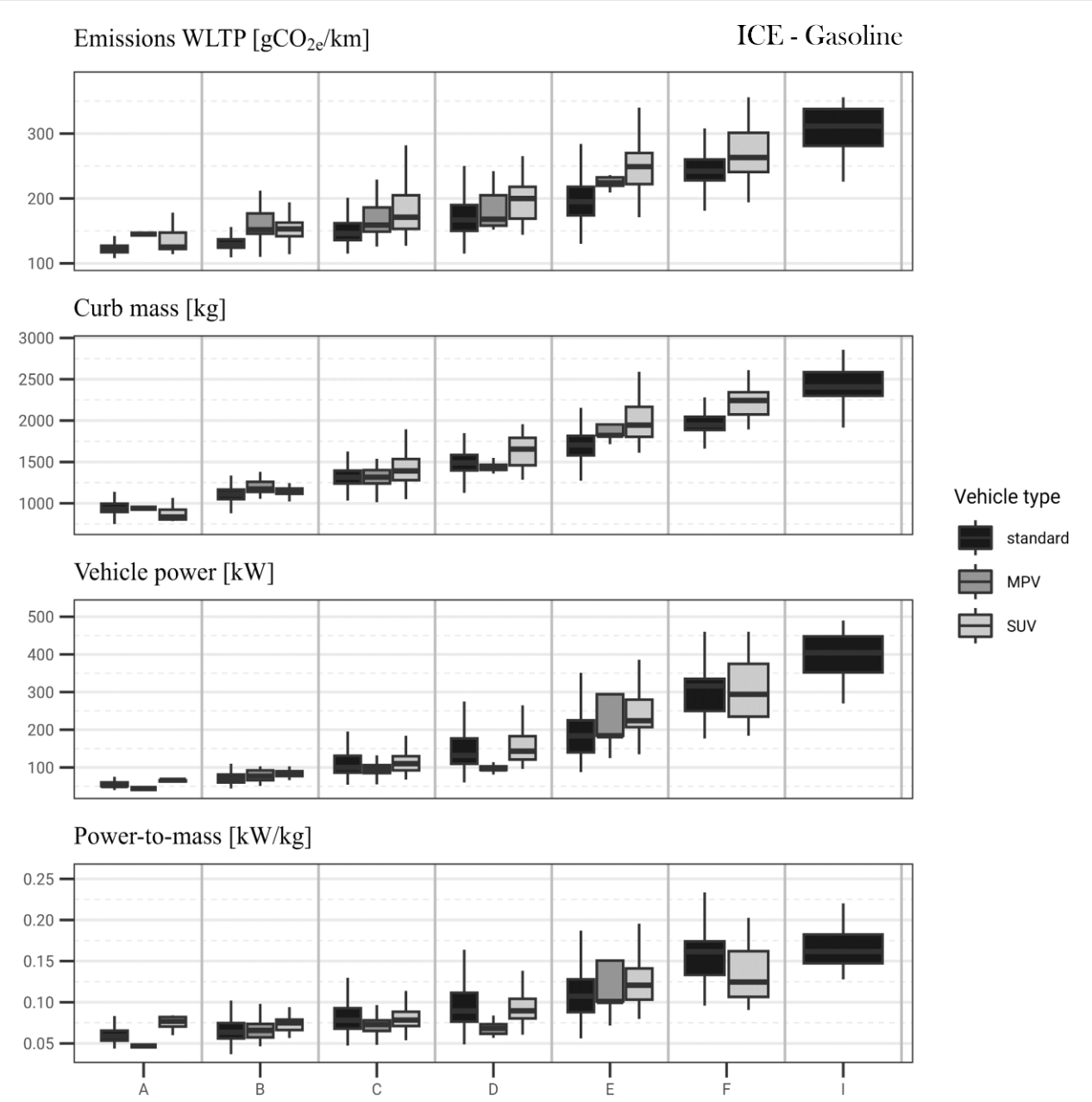
Vehicle Segment	Accuracy			
	This study	SVM	xGBoost	Dummy
A	89.5%	94.6%	100%	0%
B	87.6%	89.4%	100%	5.6%
C	81.8%	90.2%	99%	16.5%
D	70.6%	80.8%	100%	15.3%
E	92.9%	99.0%	100%	11.3%
F	96.6%	95.5%	100%	0%
I	88.7%	-	-	-
J	74.4%	72.7%	100%	8.5%
K	80.0%	87.5%	100%	0%
L	76.4%	94.1%	100%	23%
M	85.0%	81.3%	99%	9.3%
N	93.2%	91.7%	99%	5.9%
Overall	82.2%	88.9%	99.6%	12.1%

Main conclusions:

- The **designed methodology** for vehicles classification is:
  - Transparent, reproducible and replicable
  - Mathematically robust
  - **Deterministic,** with respect to other empirical approaches
  - **Compatible** with current classification approaches

- The **designed classification criteria:**
  - Was **verified** against a large set of data
  - Was able to **explain** the bulk of the extended Eurocar segment based on technical specifications of the vehicles.
  - Showed that the **distinction of MPVs and SUVs** is the most challenging part

- **MPVs and SUVs** classes still present large variability. A further classification based on the vehicle size would **improve the classification scheme**



Politecnico
di Torino