

ENEIDA.IO

Road to ZEN

Zero Emission Neighbourhoods

Eneida DeepGrid®

Detection of electric vehicle charging events in low voltage networks:
enhancing load visibility for network planning

CIREN 2024 Vienna

Adriana Leal
aleal@eneida.io

Outline

1. Introduction
2. Dataset
3. Methodology
4. Results
5. Conclusions

Introduction

Background

- Electric vehicles (EVs) introduce **non-linear loads** in low-voltage (LV) networks increasing the risk of equipment overloading.
- Distributed System Operators (DSOs) have **no visibility into EV mobility**: how many EVs are charging in the grid at a given time at a particular substation transformer's feeder?
- EVs sales are increasing globally in the major markets (China, Europe and United States). As such, **monitoring EV charging is critical** for the LV network management.

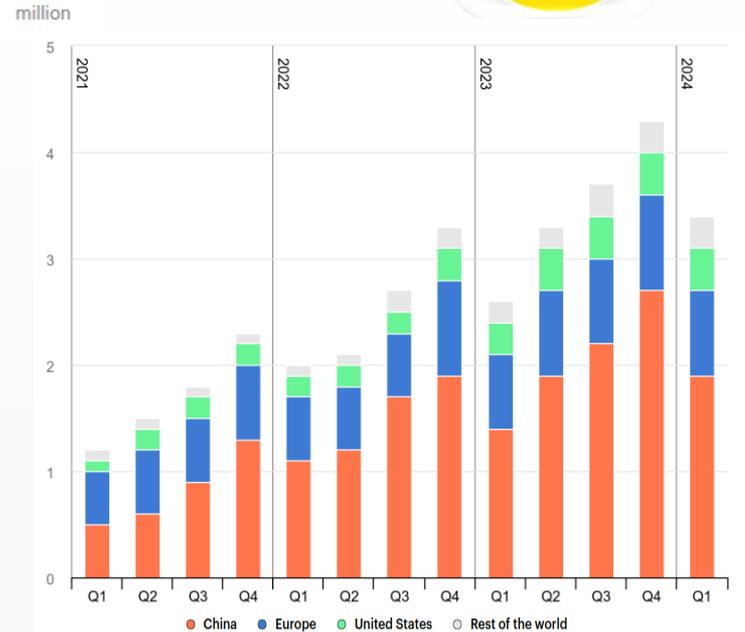


Figure 1: Quarterly electric car sales by region, 2021-2024.

Source: [Global EV Outlook 2024](#)

Introduction

Aims

- **Main goal:** Detect EV charging events in the LV network and ensure that the grid power quality standards are being met. Accurately produce an estimation of future EV charging load.
- **Detailed goal:** train supervised machine learning algorithms with samples labelled as "EV is Charging" or "EV is Not Charging", respectively indicating whether an EV is connected to the EV charging station or not.
- **Important to achieve high specificity:** never detect charging events that didn't occur (zero false positives) at the cost of missing charging events.
- **Application scenario:** one feeder feeding 10 to 12 houses in a neighbourhood possibly with one EV in each house. Domestic chargers would operate on single-phase mode.

Dataset

Acquisition setup

Data source: Eneida Deepgrid® EdgeSense device at the Instituto Superior Técnico (IST).

Charger data:

Feeder 1 collecting from charger 1 (tower 1) with two charging stations

Feeder 2 collecting from charger 2 (tower 2) with two charging stations.



(a) Transformer.



(b) Charger switchboard.



(c) Charging towers.

Dataset

Acquisition setup

- Data from *July 20, 2023 to March 11, 2024 (~27 complete weeks)*
- Only Feeder 1 and 2 were collecting charger information, Feeder 3 captures Neutral information
- Charger data shows single-phase and three-phase charging events.

Device	Acquisition/communication rate		Feeder information	Phase information	Purpose in the machine learning pipeline
	Normal Data	NPQ Data			
Transformer	1 min	5 min	Feeder 1	L1, L2, L3 and Neutral	Feature dataset
Charger	1 min	10 min	Feeder 1, 2 and 3	L1, L2, L3 and Neutral	Ground truth

Dataset

Analysed features

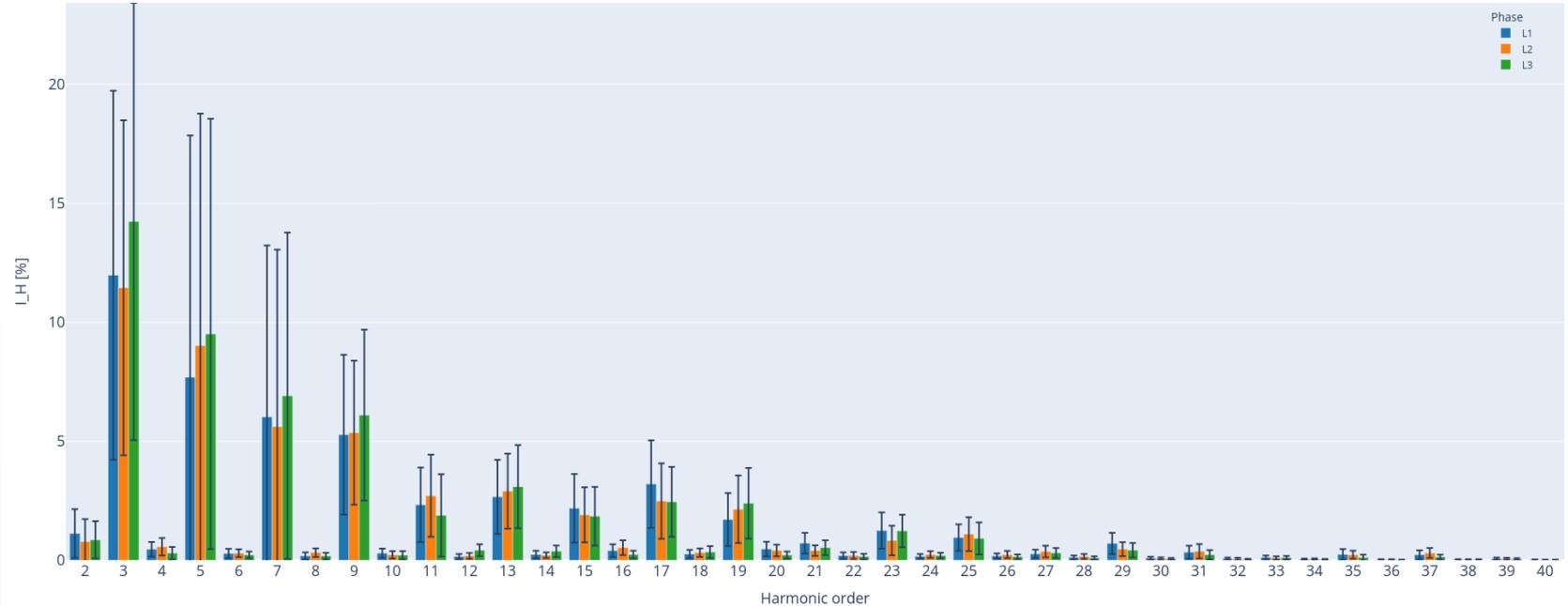
	Energy	Power	Current	Voltage
Normal Data	Active energy Active energy direction Reactive energy Reactive energy direction	Active power Apparent power Reactive power Power factor Power factor direction	Frequency Minimum Maximum RMS THD Phase angle	Frequency Minimum Maximum RMS THD Phase angle

	Current		Voltage	
	Percent of the fundamental frequency	Phase angle	Percent of the fundamental frequency	Phase angle
NPQ Data	I_H1 [%], ..., I_H40 [%]	I_H1Ang [°], ..., I_H40Ang [°]	V_H1 [%], ..., V_H40 [%]	V_H1Ang [°], ..., V_H40Ang [°]

Dataset

Analysed features

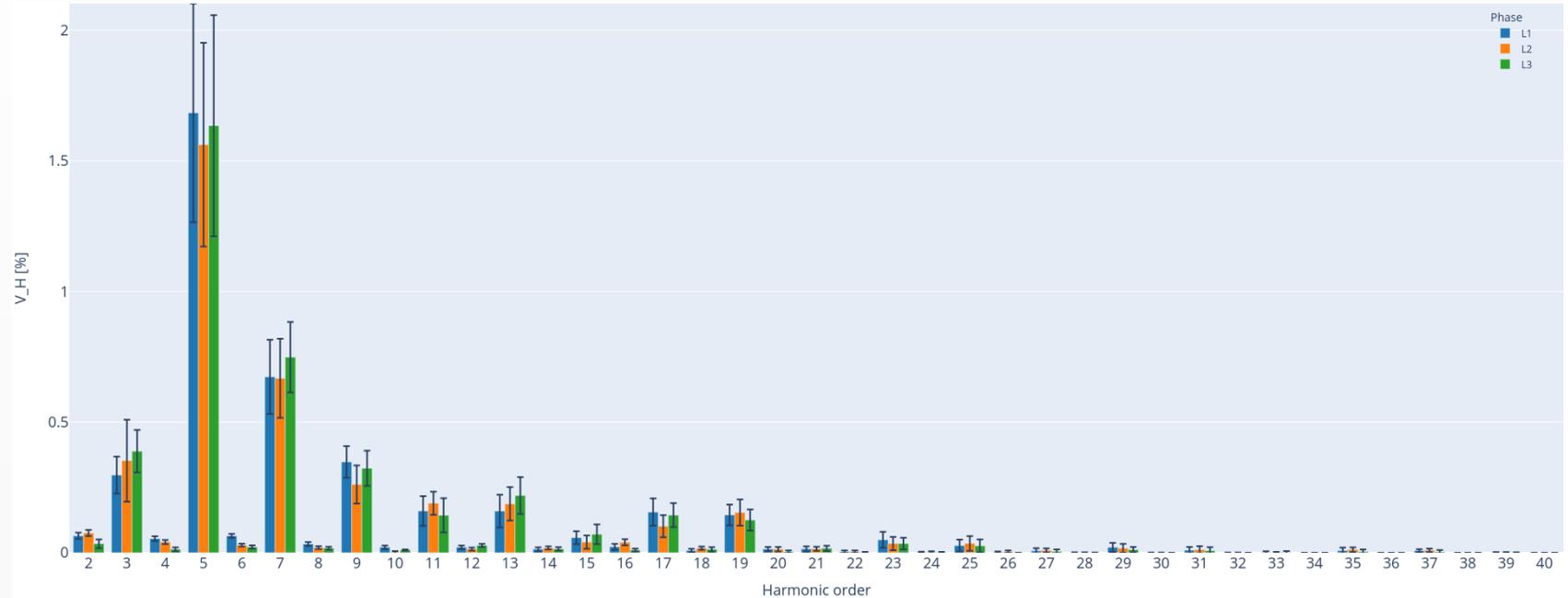
I_H [%] mean and standard deviation for each harmonic and transformer phase



Dataset

Analysed features

V_H [%] mean and standard deviation for each harmonic and transformer phase



Methodology

Data preparation

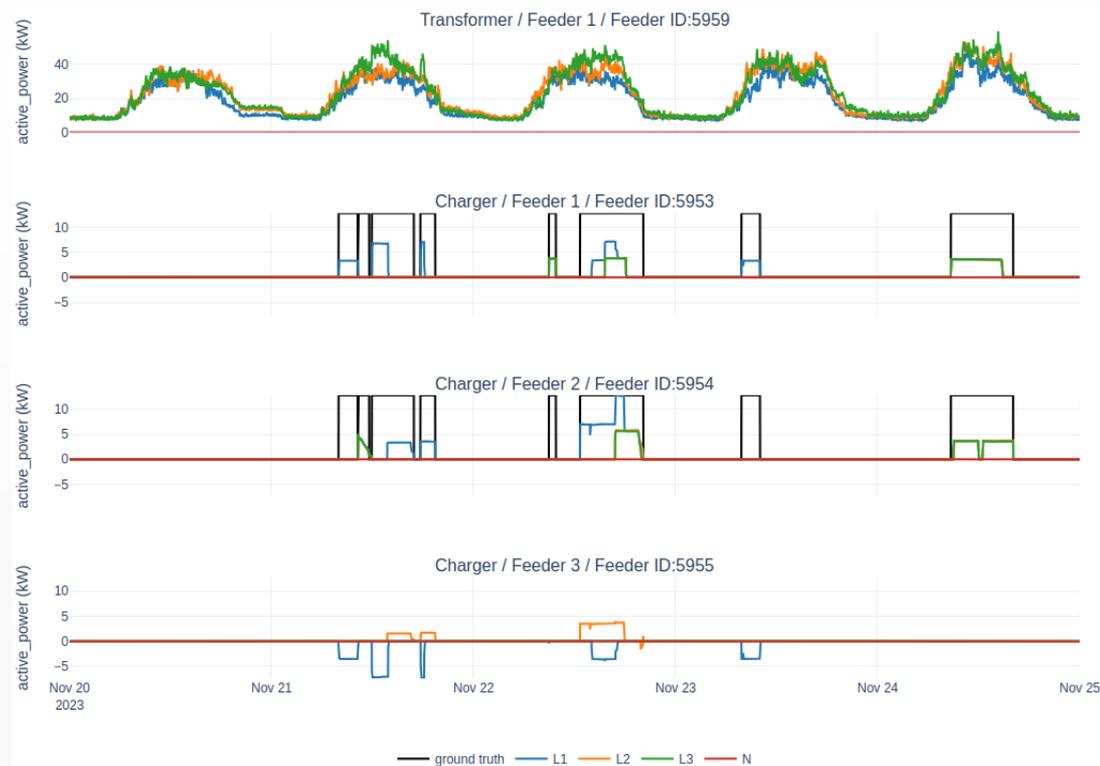
- Angle encoding through circular representation (current and phase angles were encoded using trigonometric functions sine and cosine)
- Phase matching between transformer and charger (L1, L2 and L3 in transformer correspond to L3, L2 and L1 in charger)
- Remove constant features
- During **weekends**, in the IST building, no other loads will be present in the grid, therefore, these data is not useful to discriminate EV loads from other loads and was **not used to train EV detection models**.

Methodology

Data preparation

- **Ground truth** data corresponds to all charging events (three-phase and single-phase) occurring in **active power** time series.

$\text{threshold} = \text{mean}_{\text{active_power}} + \text{std}_{\text{active_power}}$
 if $L1_t > \text{threshold}$:
 ground truth_t = 1
 else:
 ground truth_t = 0
 (the same is applied to each phase)

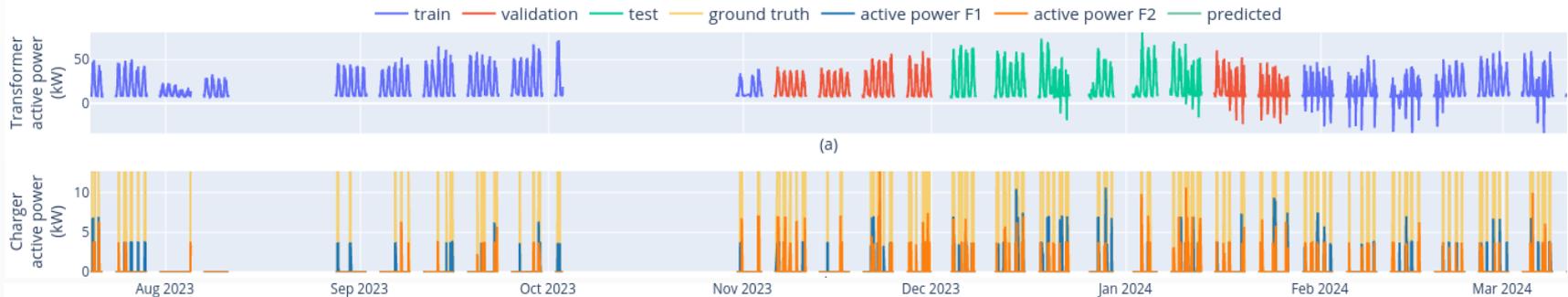


Methodology

Data partition

- **Ideally:** chronological data splitting to account for time causality and therefore simulate real-life and avoid time leakage
- **What was done:** due to the photovoltaic (PV) activity in the last weeks of data, training data also included weeks with PV patterns to increase model generalisation.

Partition	Number of samples	% of samples
Train	24759	57
Validation	9495	22
Test	8952	21

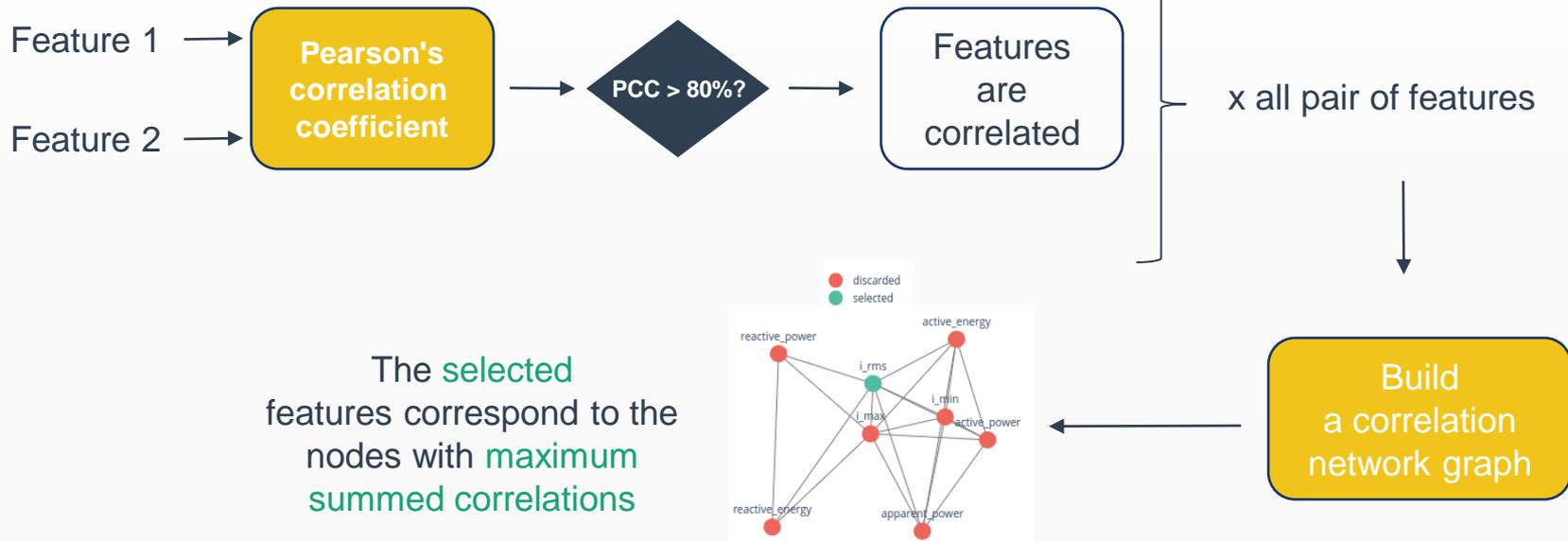


Time series plot of the feature active power in transformer and charger from (a) July 20th, 2023 to March 11th, 2024. There are weeks in August and October where data were not eligible to be analysed. In the transformer active power, we show the data partition: train (blue), validation (red) and test (green). In the charger active power, we show the feature values for feeder 1 (dark blue) and 2 (orange) and the final ground truth (yellow). PV activity can be seen in (b) when active power amplitude drops below zero.

Methodology

Feature selection

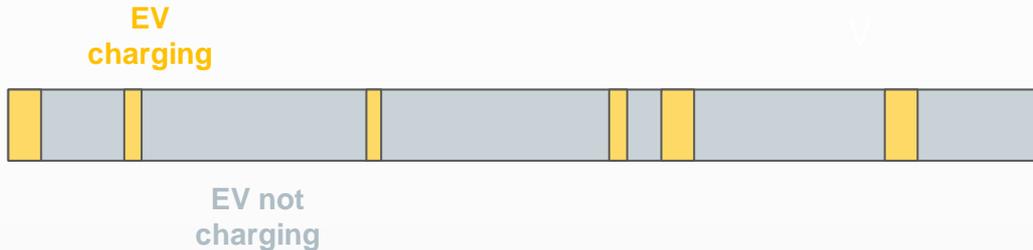
The feature dataset was assessed regarding feature redundancy.



Methodology

Dealing with class imbalance

Percentage of samples belonging to class **EV Charging: 15.31 %**
Percentage of samples belonging to class **EV NOT Charging: 84.69 %** ⇒



the model tends to be biased towards the majority class, as it has more samples to learn from



tendency to have higher sensitivity and lower specificity



in this problem we want high specificity (avoid false positives)

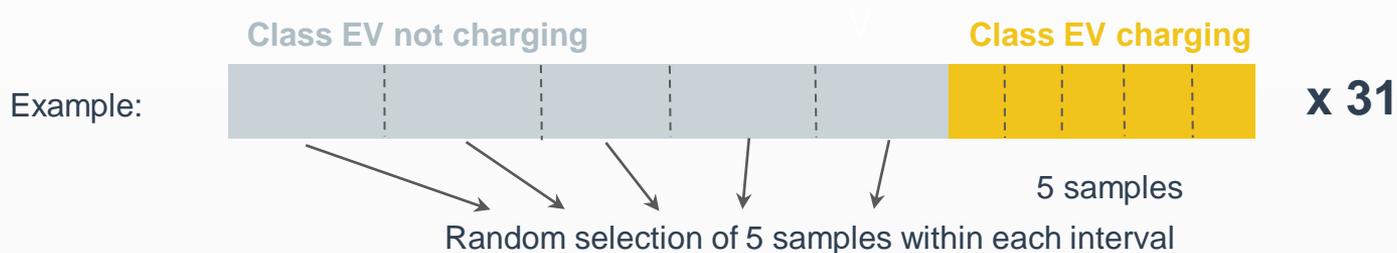
- **OPTION 1:** assign different weights according to the class prevalence (LogReg weights, Linear SVM weights, Random Forest weights)
- **OPTION 2:** use sequential random undersampling (LogReg, Linear SVM, Random Forest, kNN)

Methodology

Dealing with class imbalance

1. In train dataset: random undersampling

Perform sequential random undersampling (LogReg, Linear SVM, Random Forest, kNN)



Methodology

Dealing with class imbalance

2. In validation and test datasets: evaluate performance using geometric mean

CONFUSION MATRIX

		Predicted value	
		EV charging	EV not charging
True value	EV charging	True positives (TP)	False negatives (FN)
	EV not charging	False positives (FP)	True negatives (TN)

$$\text{sensitivity (SE)} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{correctly predicted EV charging}}{\text{all true EV charging}}$$

$$\text{specificity (SP)} = \frac{\text{TN}}{\text{TN} + \text{FP}} = \frac{\text{correctly predicted EV not charging}}{\text{all true EV not charging}}$$

$$\text{geometric mean (GM)} = \sqrt{\text{SE} \times \text{SP}}$$

3. In all partitions: weekend data were not used as there is already EV not charging events representativity in data

Results

Classification results for the three datasets

Partition	Dataset	Best model	SE (%)	SP (%)	GM (%)
Validation	Current, voltage and power	Random Forest	84.2	74.7	79.3
	NPQ	Linear SVM weights	86.5	76.4	81.3
	Current, voltage, power and NPQ	Linear SVM	86.1	78.3	82.1
Test	Current, voltage, power and NPQ	Linear SVM	77.7	83.0	80.2

In the test dataset the model was able to correctly identify 77.7% of the EV charging events.

Results

Feature selection results for the best model: linear SVM with $C = 2^{-5}$

	Energy	Power	Current	Voltage
Normal Data	Active energy Active energy direction Reactive energy Reactive energy direction	Active power Apparent power Reactive power Power factor Power factor direction	Frequency Minimum Maximum RMS THD Phase angle (sin, cos)	Frequency Minimum Maximum RMS THD Phase angle (sin, cos)

NPQ Data	I_H [%]	1, 4, 12, 13, 17, 18, 19, 21, 23, 25, 27, 28, 29, 31, 33, 35, 39	
	V_H [%]	1, 2, 3, 7, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40	
		sin	cos
	I_HAng [°]	2, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 15, 16, 17, 18, 19, 21, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 39, 40
	V_HAng [°]	4, 5, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40	2, 3, 4, 7, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40

■ Constant

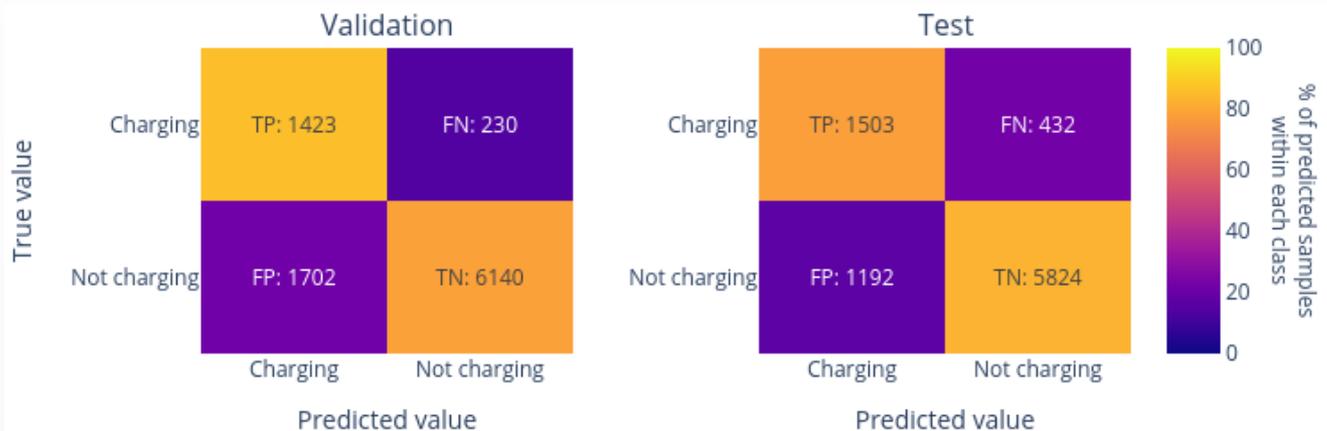
■ Discarded

■ Selected

Results

Classification results for the best model: linear SVM with $C = 2^{-5}$

Confusion matrices for validation and test datasets



Results

Classification results for the best model: linear SVM with $C = 2^{-5}$



(a) Data partition shown for active power. (b) Ground truth (EV charging vs EV not charging) shown for active power. PV activity can be seen when active power amplitude drops below zero. (c) Prediction results obtained for the best model (Linear SVM) on current, voltage, power and NPQ data.

Conclusions/Future work

Conclusions

- Best prediction performance in test dataset (**78% sensitivity and 83% specificity**) was achieved by training **a linear SVM model on all available data**.
- PV injection is negatively affecting model performance.

Future endeavours

- **Train models with transformer data collected at a neighbourhood area** (scenario where a feeder provides for 10 to 12 houses, possibly with one EV per house).
- Consider training new models that can more accurately **deal with the presence of PV activity in the grid**.
- Explore **new grid measures that are specific to EV activity** (load decoupling).

Acknowledgements



Financiado pela
União Europeia
NextGenerationEU

 **aliança para a
transição energética**

ENEIDA.IO

Road to ZEN

Zero Emission Neighbourhoods

Adriana Leal | aleal@eneida.io