Citizen science for traffic monitoring: Investigating the potentials for complementing traffic counters with crowdsourced data
Descriptions of the best regression models

Miha Janež[1], Špela Verovšek[2], Tadeja Zupančič[2], Miha Moškon[1,*]

[1]Faculty of Computer and Information Science, University of Ljubljana
[2]Faculty of Architecture, University of Ljubljana
[*] corresponding author; e-mail: miha.moskon@fri.uni-lj.si

This document is supplementing the paper *Citizen science for traffic monitoring: Investigating the potentials for complementing traffic counters with crowdsourced data* with a more detailed description of the best performing models together with their parameters. In each of the observed road segments either kernel ridge regression (krr) or gradient boosting regression (gbr) yields the best results. The details of the implementations of these two models together with the description of their parameters are available in scikit-learn documentation [1]. The description of kernel ridge regression implementation is available here. The description of gradient boosting regression is available here.

Dunajska (from centre)

Summary

- Model: krr
- Features: basic, 0656-1, 0655-2

Model parameters

- alpha = 0.001
- coef0 = 1
- degree = 3
- gamma = 0.001
- kernel = chi2

Dunajska (to centre)

Summary

- Model: krr
- Features: basic, 0656-2, 0655-1

Model parameters

- alpha = 0.001
- coef0 = 1
- degree = 3
- gamma = 0.001
- kernel = chi2

Ižanska (from centre)

Summary

- Model: krr
- Features: basic, 0820-1, 1506-1

Model parameters

- alpha = 0.0001
- coef0 = 1
- degree = 3
- gamma = 0.001
- kernel = chi2

Ižanska (to centre)

Summary

- Model: krr
- Features: basic, 0820-2, 1506-2

Model parameters

- alpha = 1e-05
- coef0 = 1
- degree = 3
- gamma = 0.001
- kernel = chi2

Slovenska (from centre)

Summary

- Model: krr
- Features: basic, 0619-1

Model parameters

- alpha = 1.0
- coef0 = 1
- degree = 3
- gamma = 0.1
- kernel = chi2

Slovenska (to centre)

Summary

- Model: krr
- Features: basic, 0619-2

Model parameters

- alpha = 0.0001
- coef0 = 1
- degree = 3
- gamma = 0.01
- kernel = chi2

Škofije (towards Koper)

Summary

- Model: gbr
- Features: basic, 1092-1

Model parameters

- alpha = 0.9
- ccp_alpha = 0.0
- criterion = friedman_mse
- learning_rate = 0.1
- loss = squared_error
- max_depth = 3
- min_impurity_decrease = 0.0
- min_samples_leaf = 1
- min_samples_split = 2
- min_weight_fraction_leaf = 0.0
- n_estimators = 100

- subsample = 1.0
- tol = 0.0001
- validation_fraction = 0.1
- verbose = 0
- warm_start = False

Škofije (towards Trieste)

Summary

- Model: krr
- Features: basic, 1092-2

Model parameters

- alpha = 0.0001
- coef0 = 1
- degree = 3
- gamma = 0.01
- kernel = chi2

References

[1] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.