Appendix to the manuscript Probabilistic Spatial Modelling of Travel Mode Choices with Synthetic Instances

No Author Given

No Institute Given

1 Decision Tree

Figure 1 presents a simplified decision tree model used in this work for travel mode choice prediction. The Decision Tree model was optimised based on grid search performed with `caret`. It has been used for estimating probabilities for individual trips present in synthetic trip data set.

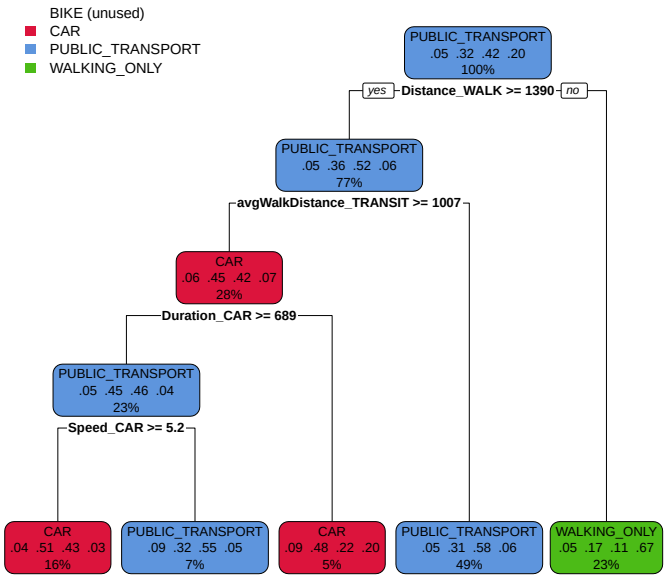


Fig. 1: The decision tree classification model, constructed using the `rpart` library with `cp = 0.01` selected in the process of hyperparameter tuning.

2 Probabilistic Spatial Modelling of Travel Mode

The figures presented in this section demonstrate the probabilities associated with the selection of various transport modes, as estimated by the decision tree and logistic regression models. The figures presented illustrate the aggregated probability values calculated using the minimum (`min`), maximum (`max`), median (`median`), and average (`avg`) functions for four distinct modes of transportation: public transport (`PT`), car (`CAR`), walking only (`WALK`) and bike (`BIKE`).

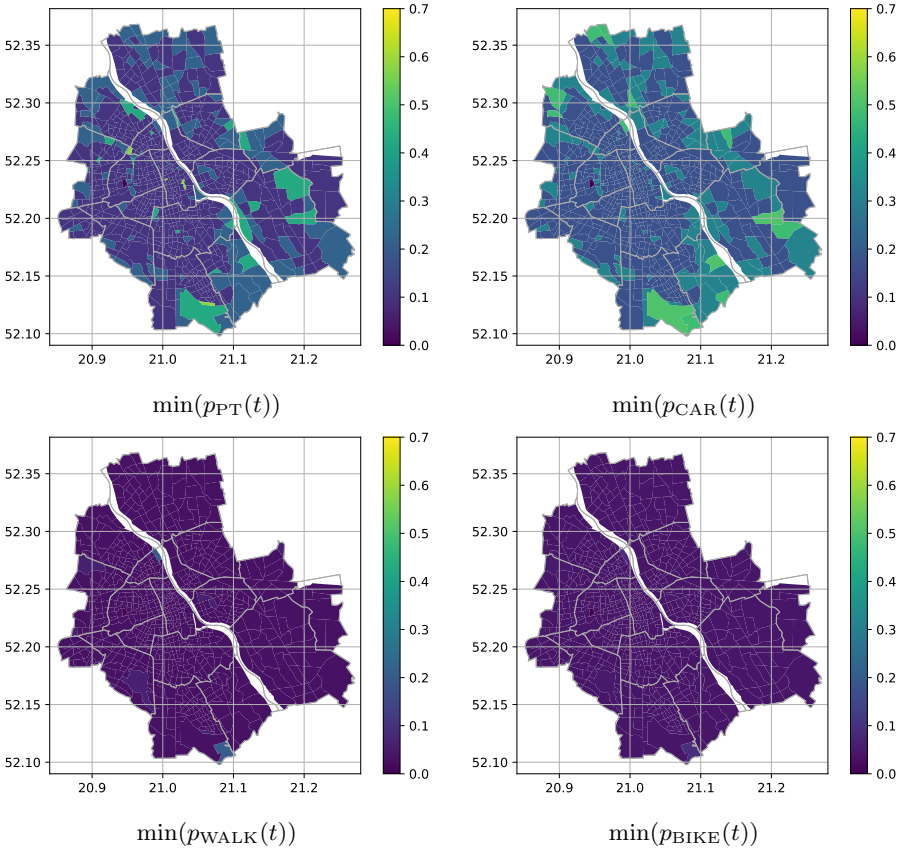


Fig. 2: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the decision tree model and aggregated by the `min` function.

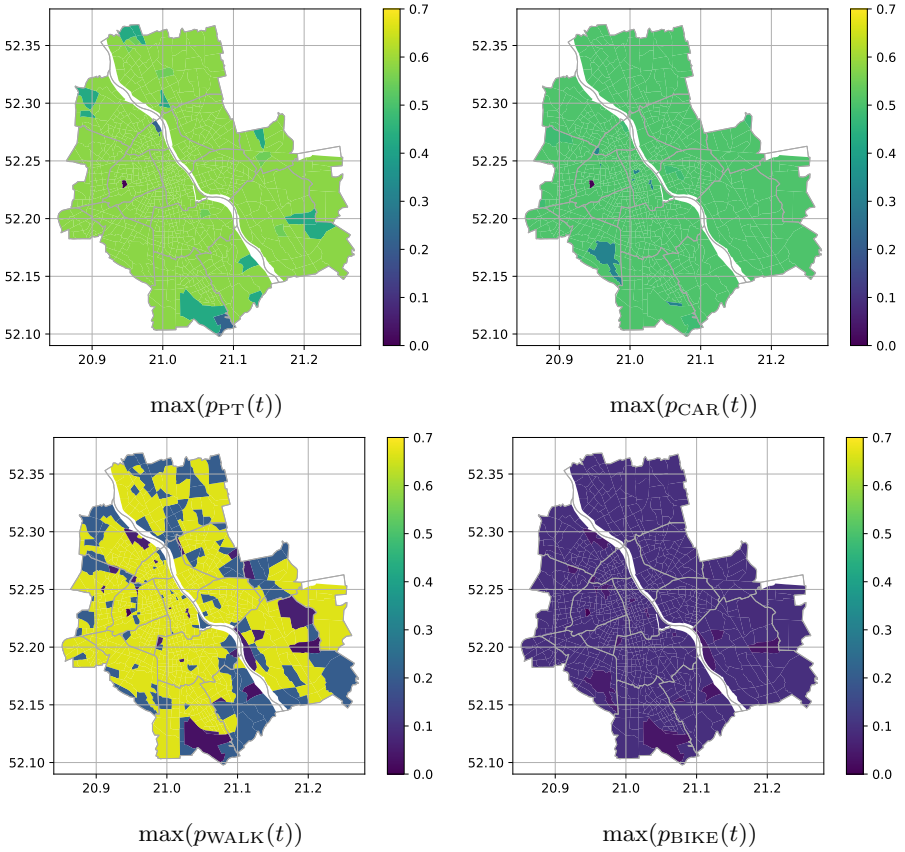


Fig. 3: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the decision tree model and aggregated by the `max` function.

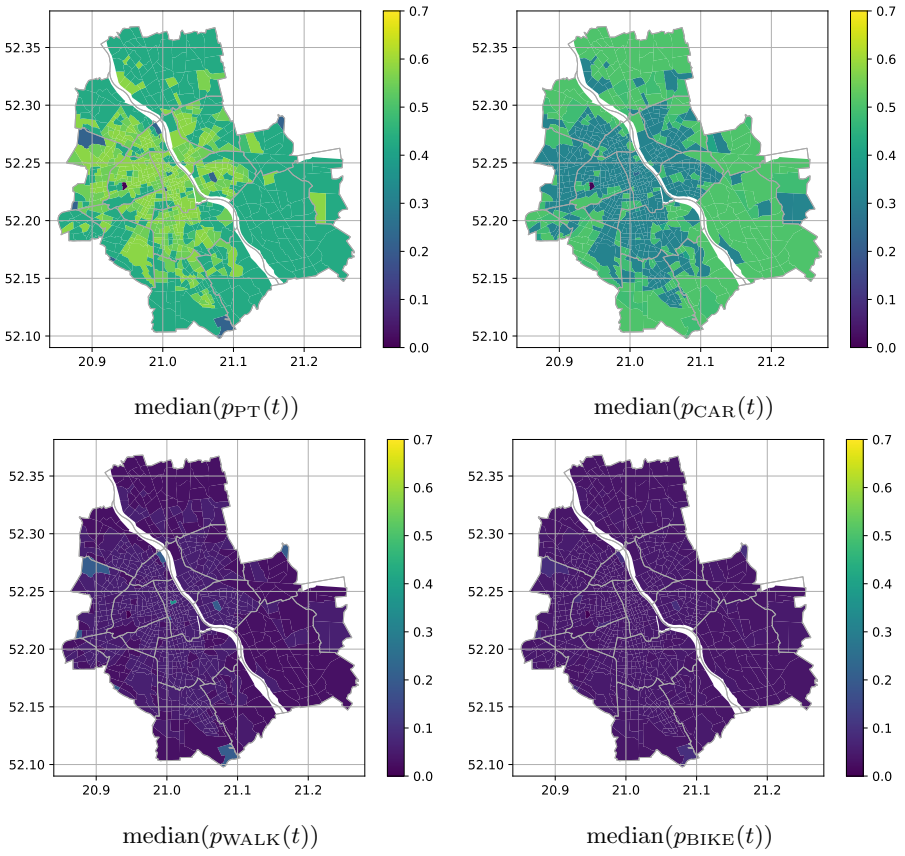


Fig. 4: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the decision tree model and aggregated by the `median` function.

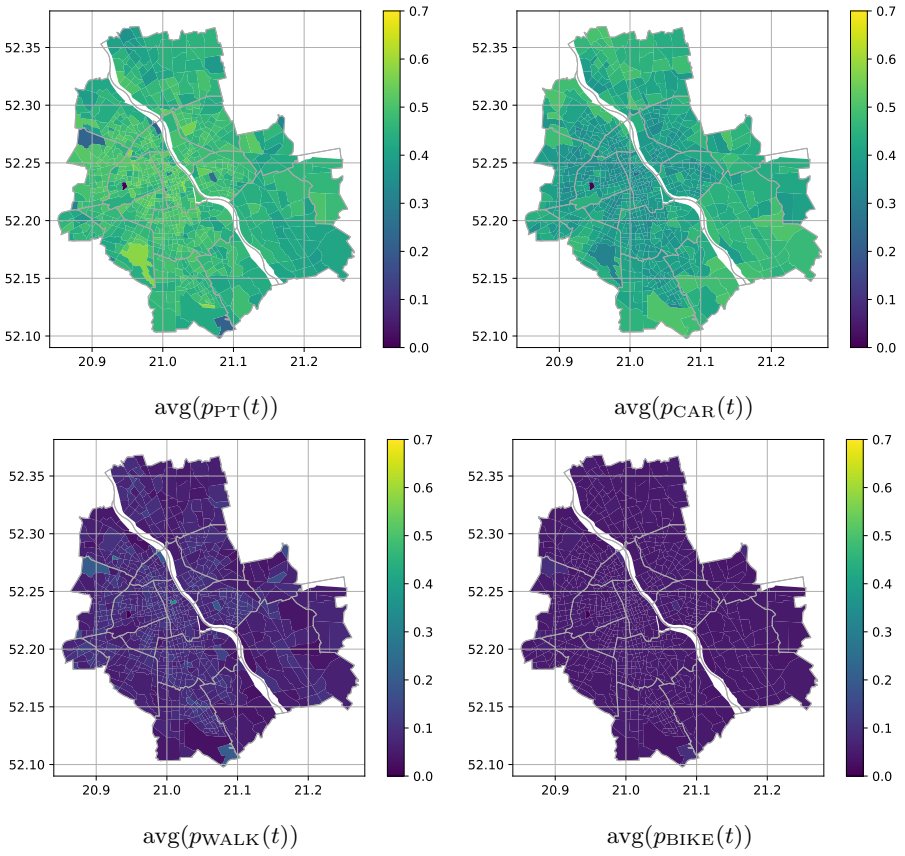
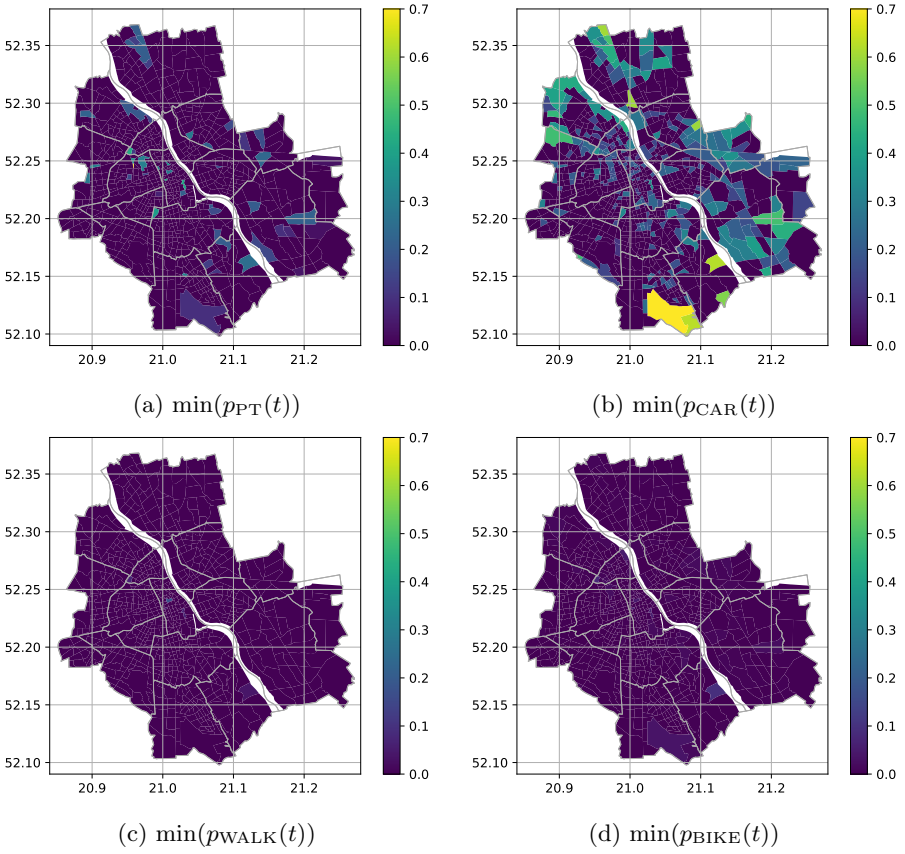


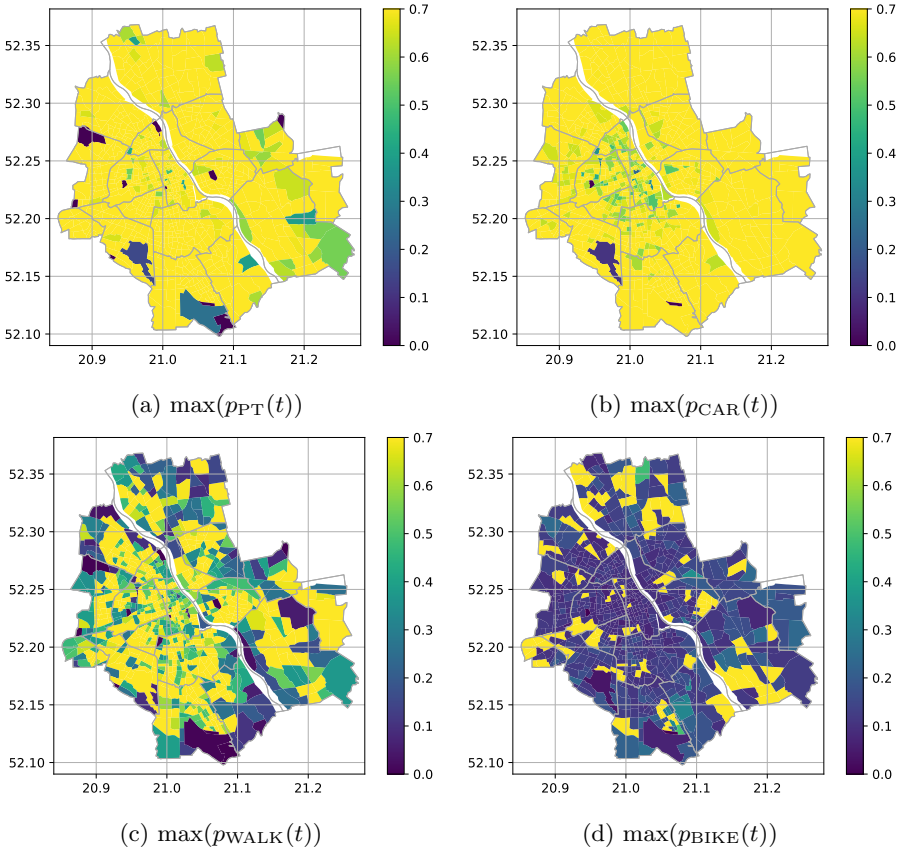
Fig. 5: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the decision tree model and aggregated by the `avg` function.



(a) $\min(p_{\mathrm{PT}}(t))$ (b) $\min(p_{\mathrm{CAR}}(t))$

(c) $\min(p_{\mathrm{WALK}}(t))$ (d) $\min(p_{\mathrm{BIKE}}(t))$

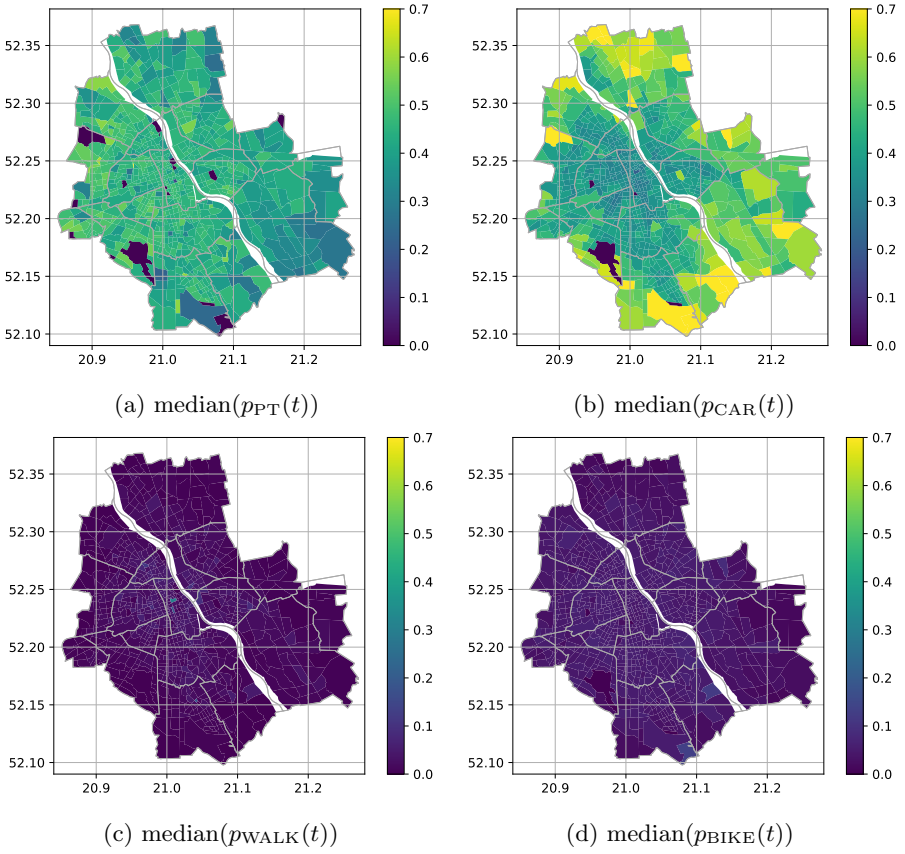
Fig. 6: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the logistic regression model and aggregated by the `min` function.



(a) $\max(p_{\mathrm{PT}}(t))$ (b) $\max(p_{\mathrm{CAR}}(t))$

(c) $\max(p_{\mathrm{WALK}}(t))$ (d) $\max(p_{\mathrm{BIKE}}(t))$

Fig. 7: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the logistic regression model and aggregated by the `max` function.



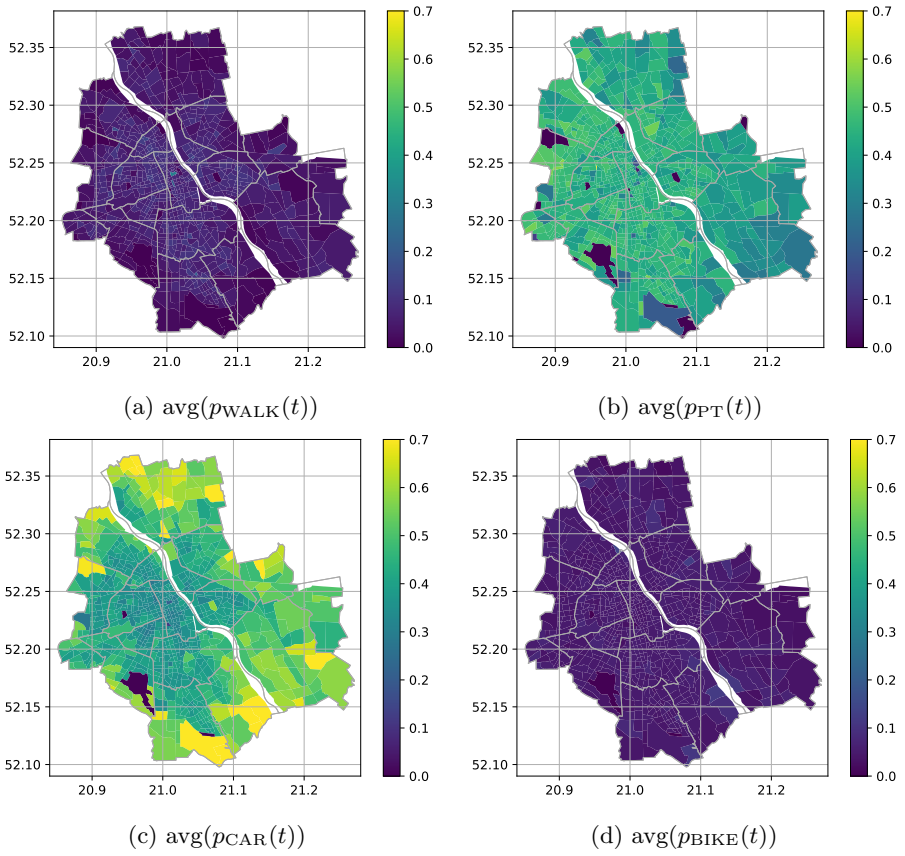
(a) median($p_{\mathrm{PT}}(t)$)

(b) median($p_{\mathrm{CAR}}(t)$)

(c) median($p_{\mathrm{WALK}}(t)$)

(d) median($p_{\mathrm{BIKE}}(t)$)

Fig. 8: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the logistic regression model and aggregated by the `median` function.



(a) avg($p_{\mathrm{WALK}}(t)$) (b) avg($p_{\mathrm{PT}}(t)$)

(c) avg($p_{\mathrm{CAR}}(t)$) (d) avg($p_{\mathrm{BIKE}}(t)$)

Fig. 9: The aggregated probability of choosing a travel mode per city zone. Probability values derived from the logistic regression model and aggregated by the `avg` function.