



A comparison of classification methods for modelling urban mode choice

Tim Hillel

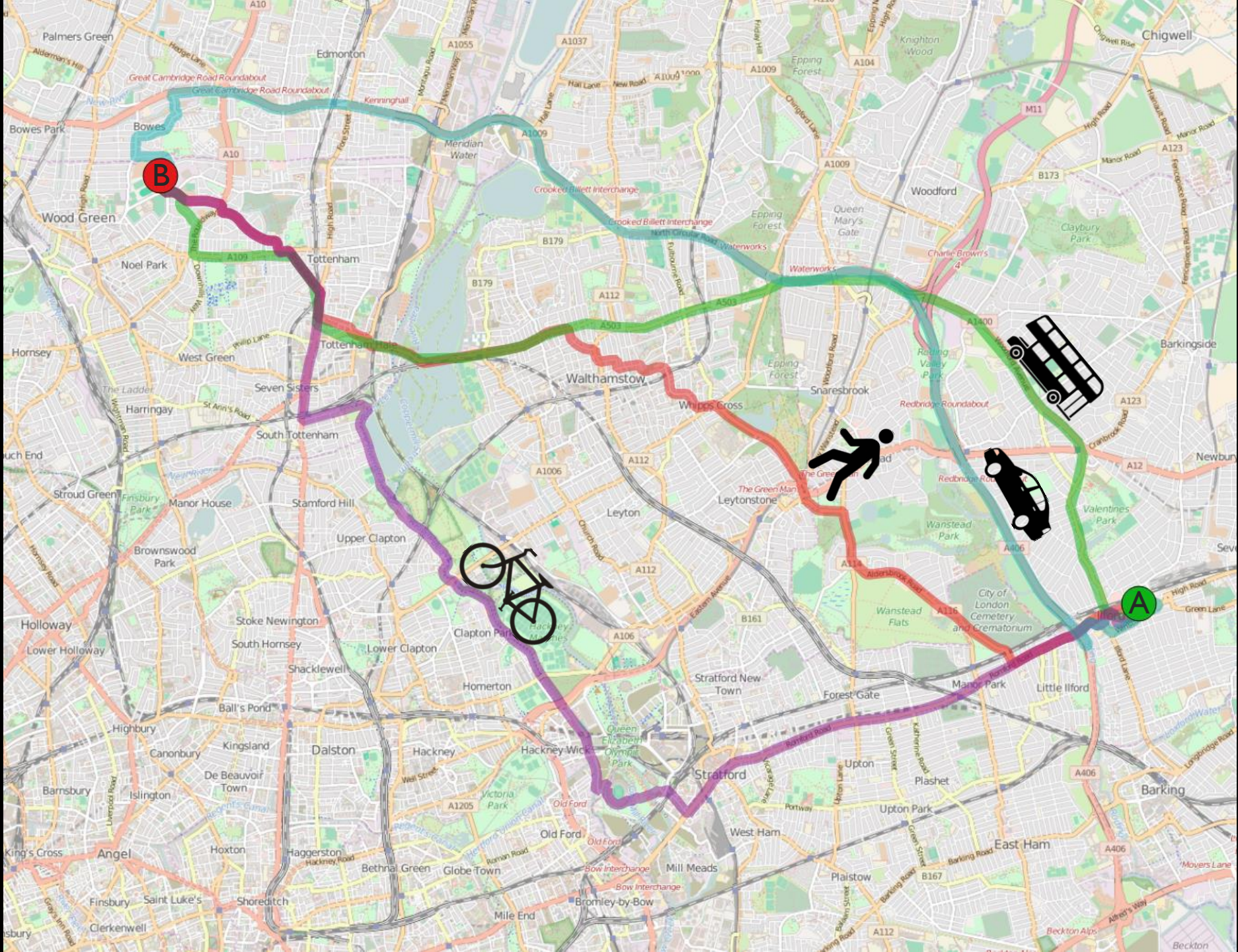


UNIVERSITY OF
CAMBRIDGE



EPSRC

Engineering and Physical Sciences
Research Council



$$\Pr(Y_i = K) = \frac{1}{1 + \sum_{k=1}^{K-1} e^{\beta'_k \cdot X_i}}$$



Population increase of
25% by 2036



Emissions cut **80%** by
2050

How can we move *more people,*
more efficiently?



Security



Accessibility



Aging infrastructure



Disruptive innovation

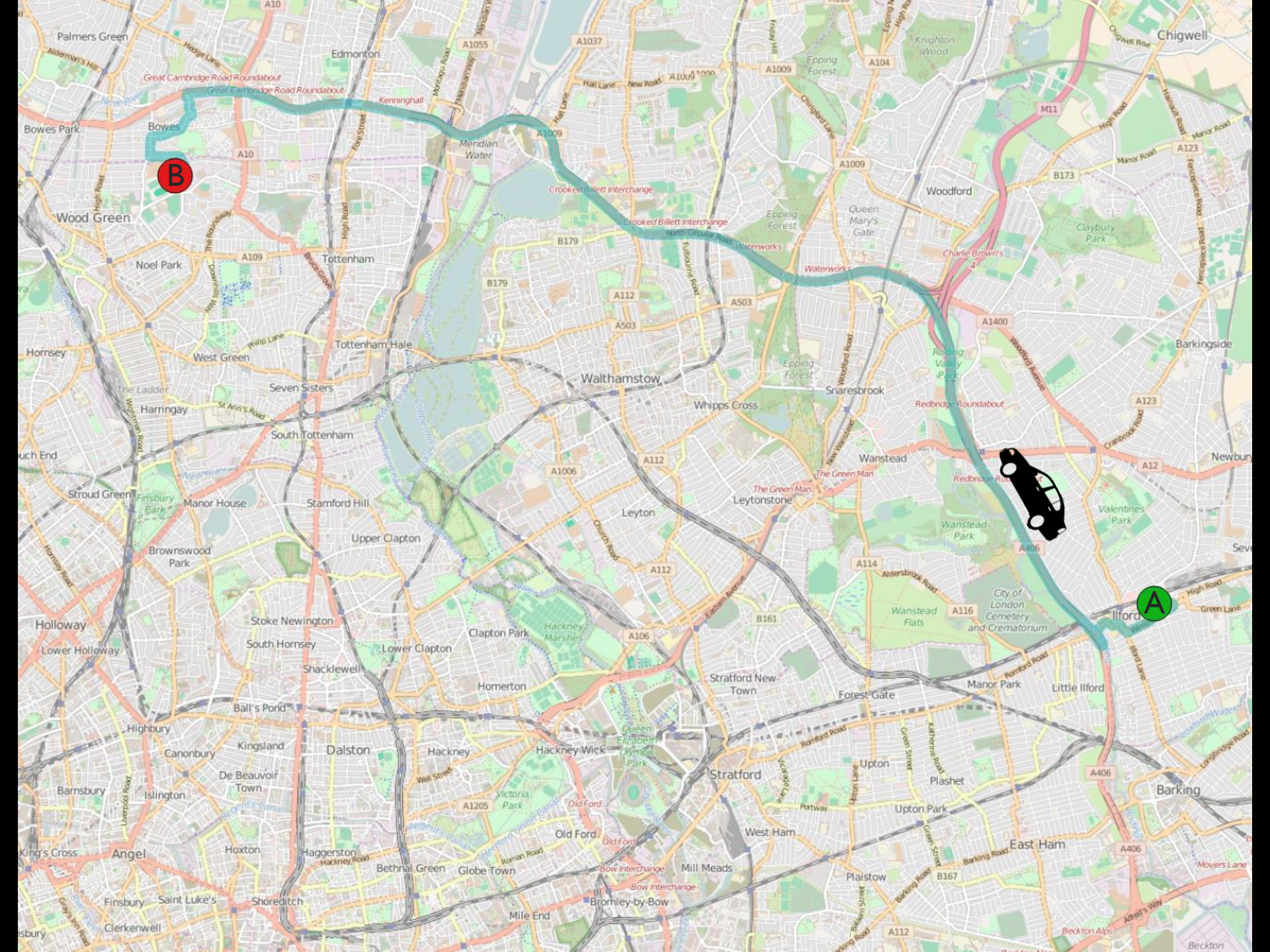


Automation

B



A





Historical
trip data

Trip details



Journey planner
service

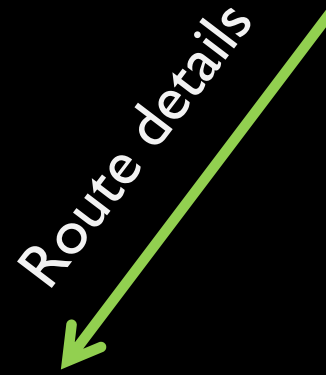


Historical
trip data

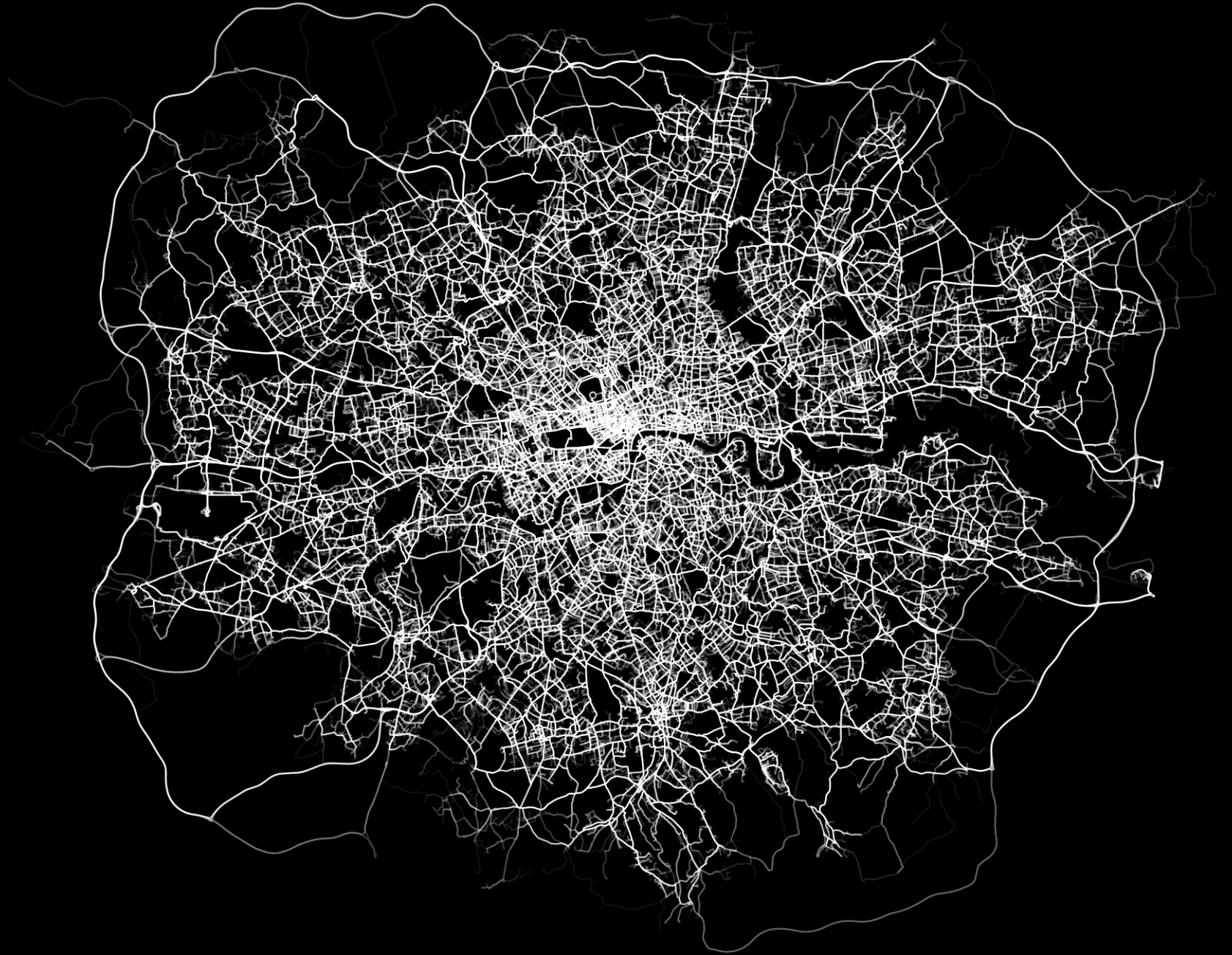
Trip details



Journey planner
service



Cost Model



Input data



Feature vector

- Start point
- End point
- Start time
- Journey purpose
- Car ownership
- Fare type

- For each mode:
 - Duration
 - Cost
 - Typical traffic
- Journey purpose
- Car ownership

75%
Training data

25%
Testing data

Algorithm

Predictive accuracy

SVM

75.6%

RF

80.0%

GB

82.1%

NN

76.6%

BNN

77.5%

GP

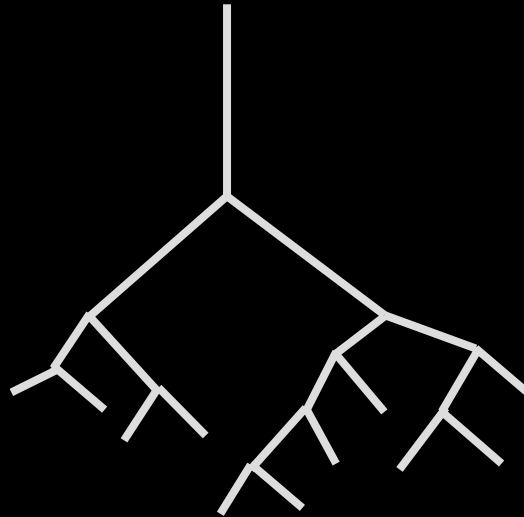
75.0%

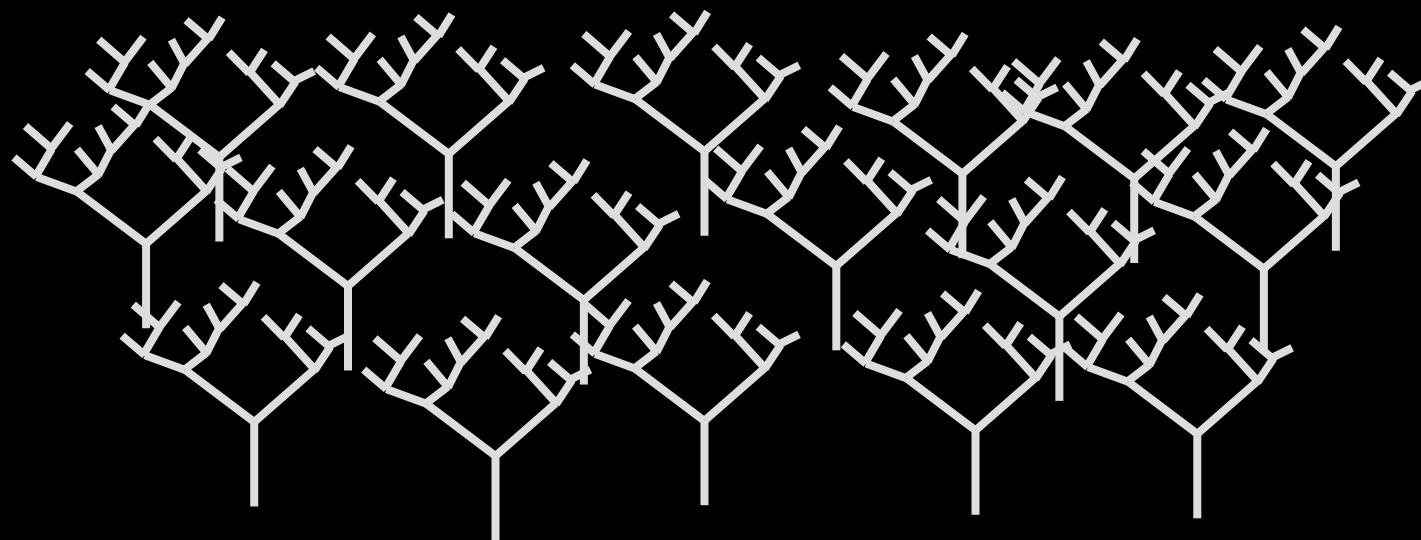
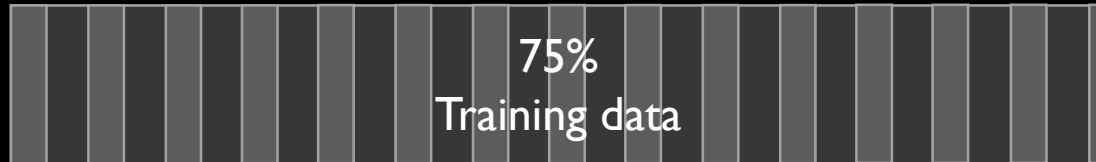
MNL

75.0%

Algorithm	Predictive accuracy
SVM	75.6%
RF	80.0%
GB	82.1%
NN	76.6%
BNN	77.5%
GP	75.0%
MNL	75.0%

75%
Training data





Input data



82% accuracy

- Start point
- End point
- Start time
- Journey purpose
- Car ownership
- Fare type

Feature Importances

