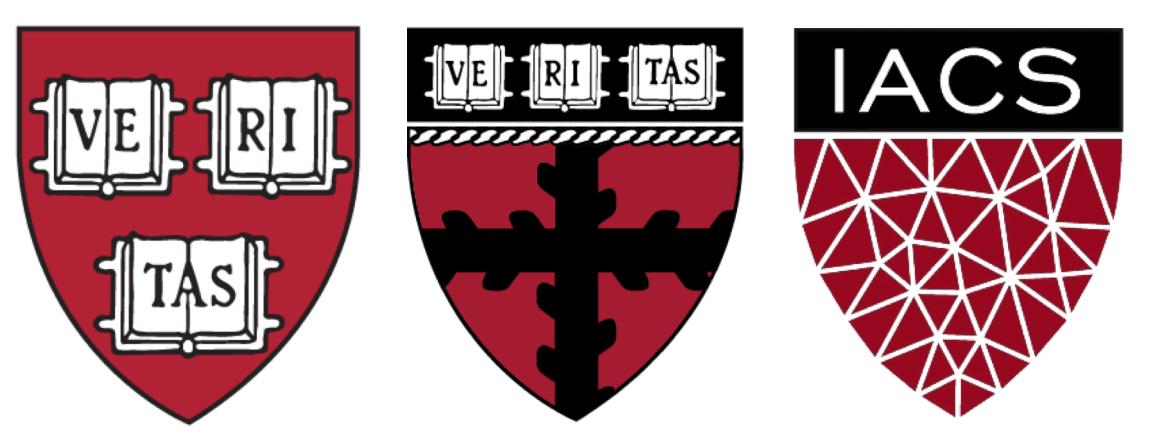


Estimating Off-Street Parking in Somerville

Joshua Feldman, Lipika Ramaswamy, Anthony Rentsch
AC297r, Institute for Applied Computational Science, Harvard University



In collaboration with
City of Somerville

INTRODUCTION

- The city of Somerville is conducting a comprehensive audit of parking. This includes on-street parking, parking garages, and residential off-street parking.
- We contributed to this audit by estimating the city's off-street, residential driveway parking capacity and providing an inventory of driveways.

GOAL: Predict probability of each parcel having a driveway

DATA

FEATURE DATA SOURCES

- Parcel and building polygons
- Administrative tabular data:
 - Property tax assessment
 - Parking permits by street address

There are 11,704 residential parcels in Somerville.

LABELLING DATA SOURCES

Aerial Imagery
(MassOrtho Imagery 2015)

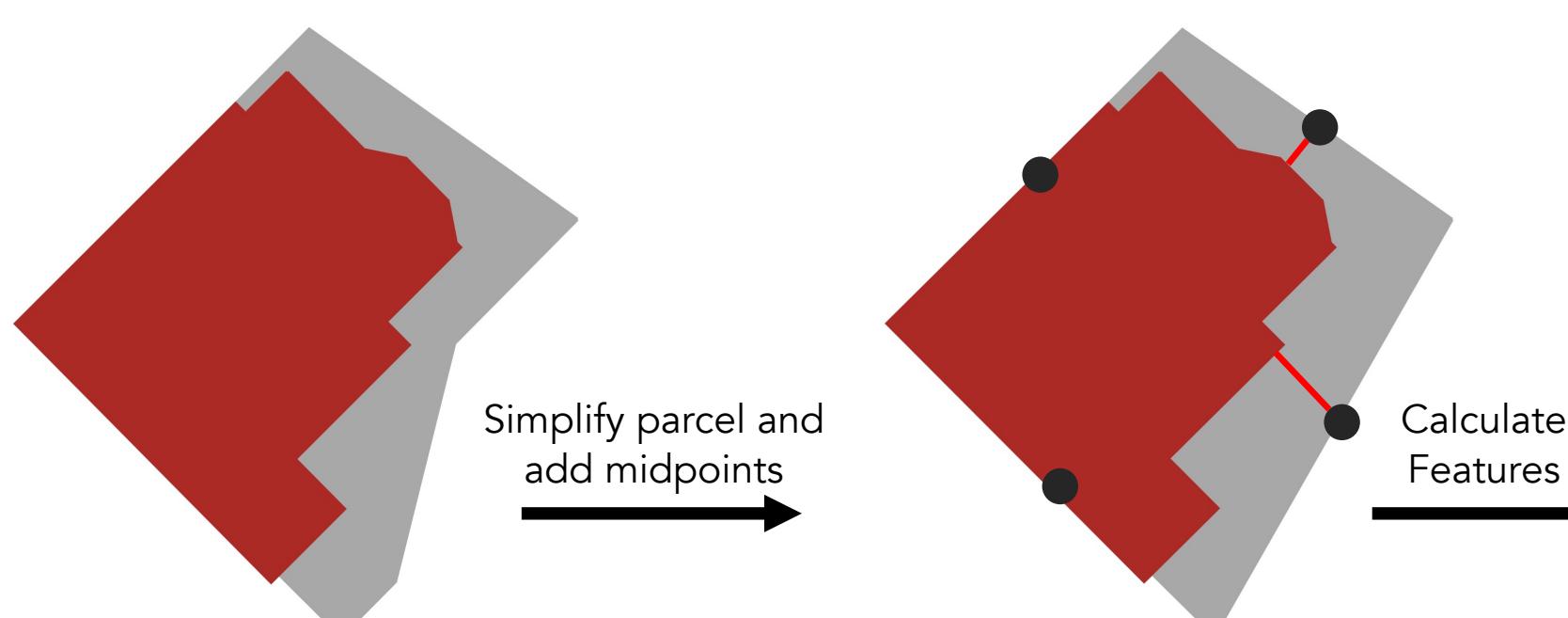


Labelled as having a driveway

We labelled 6,480 residential parcels to indicate whether or not they have driveways.

FEATURE ENGINEERING

Geometric Features



- Distance from building to side 1, side 2, front, and back of parcel
- Number of buildings
- How much closer building is to one side of parcel and to front of parcel
- Number of edges in parcel and building polygons
- Distance to nearest neighbors

Total number of features: 175

Other Features

- Ratio of building area to lot area
- Number of parking permits registered per address

Note: All data is publicly available via Google or Somerstat: Somerville's Open Data Repository.

METHODS

MODELLING

Machine learning with administrative data

- Framed the problem as a classification task. Classifiers implemented include Logistic Regression (with and without higher order terms), **Random Forest**, AdaBoost, XGBoost and Feed Forward Neural Network.
- Class Imbalance: Only 16% of hand labelled parcels did not contain a driveway. For model training, the minority class was upsampled to synthetically obtain balanced classes.
- Best performing model:** Random Forest with no bootstrapping, a max tree depth of 6, and 256 estimators (found through cross validation grid search).

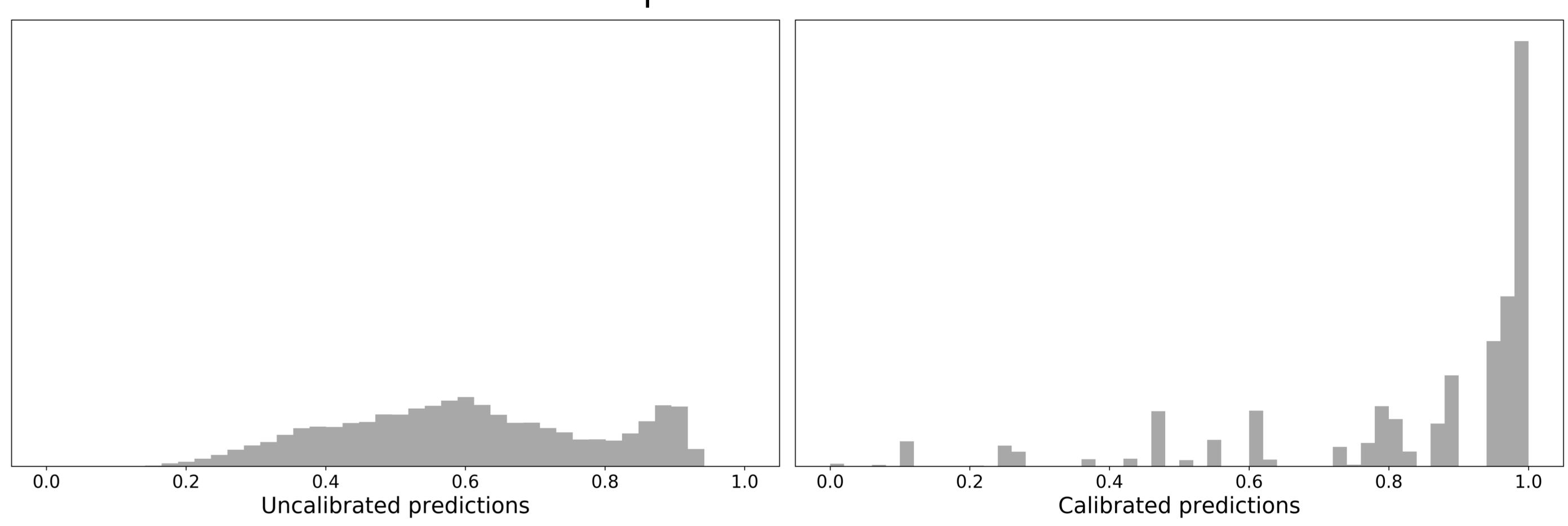
Deep learning with image data

- Framed the problem for convolutional neural networks to predict driveways from satellite and/or street view images.
- Attempted to train models with randomly initialized weights and using transfer learning with weights learned from aerial imagery.
- Highly unsuccessful due to limited labels, poor signal from low quality satellite images, and ambiguity in street view images.

POST-MODELLING

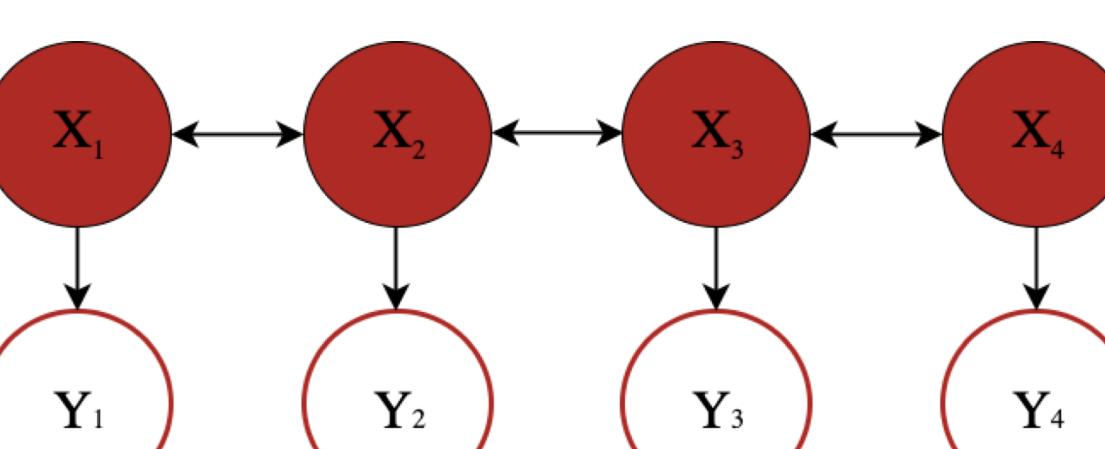
Probability calibration

- A well-calibrated probability is one that reflects the true likelihood of an event.
- Since many machine learning models produce scores that are not well calibrated, a recalibration step was performed to turn raw model scores into well-calibrated probabilities.



ESTIMATION

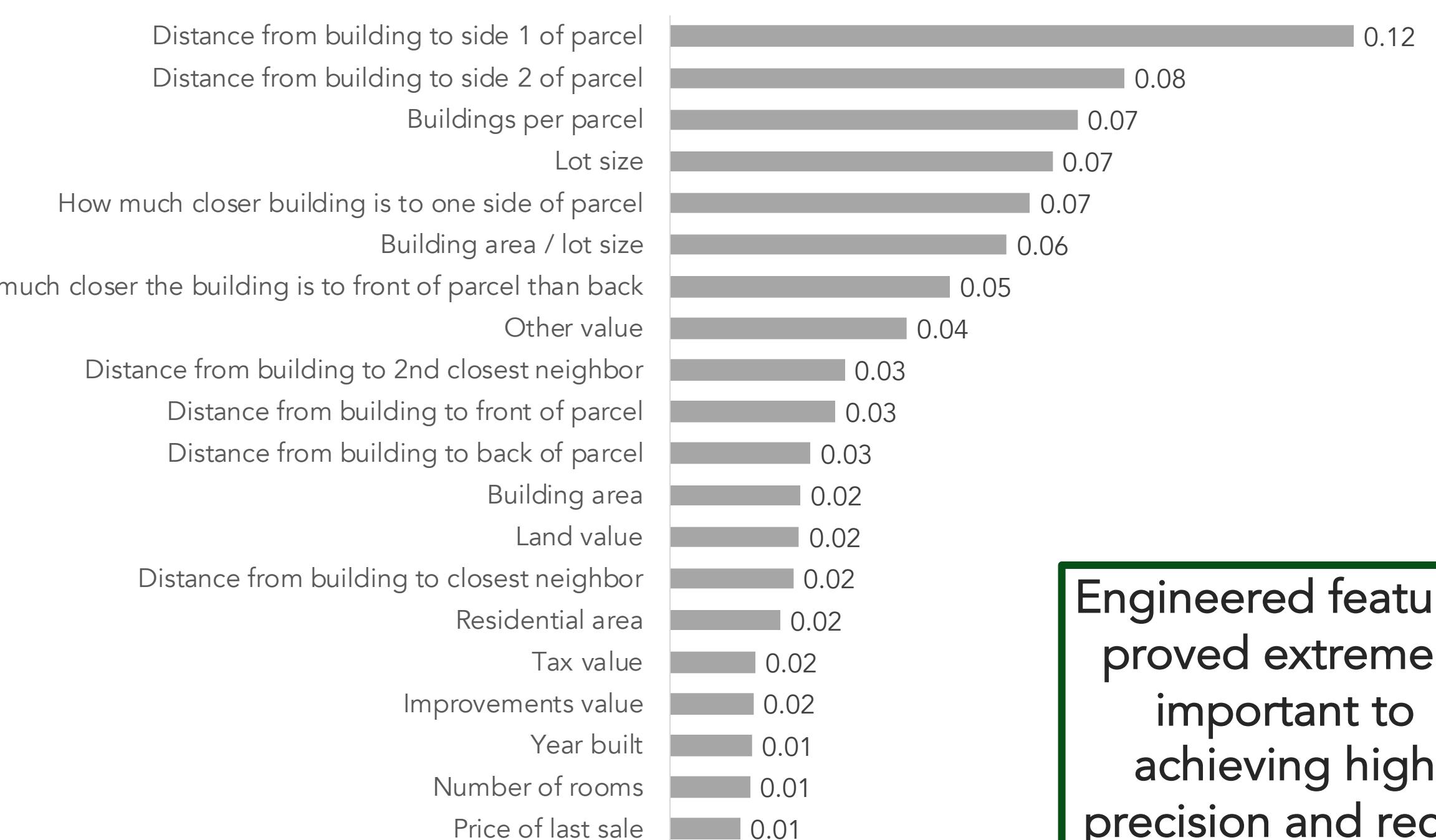
- Estimating uncertainty is complicated by spatial correlation
- We assume that conditional on observed covariates X , the presence of a driveway Y is independent from parcel to parcel.



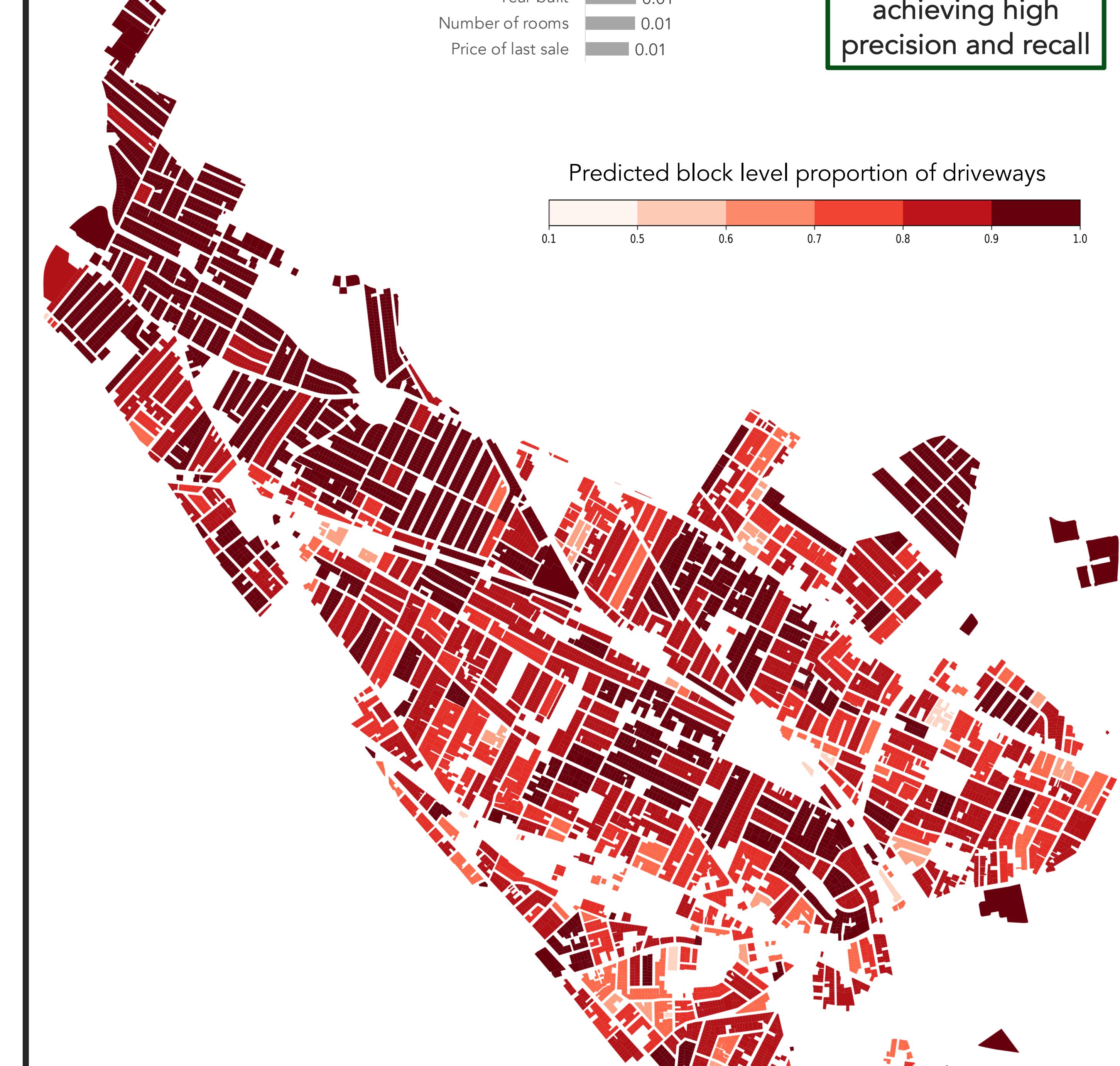
RESULTS

Test set:	Precision	Recall	F1 Score	Support
Driveway	0.78	0.82	0.80	1179
No Driveway	0.81	0.77	0.79	1179

Feature Importance



Engineered features proved extremely important to achieving high precision and recall



10,191 ± 63
parcels with driveways