WINNING THE RAT RACE

WHERE ARE BOSTON'S RODENTS?

Faculty Advisor: Prof. Daniel Freund
City of Boston Team: Jeff Lambart, Hope Ruse, John Ulrich, Gina Belmonte, Kelly Mackey


Sara Pasquino


Sasha Lioutikova

PROBLEM STATEMENT

The City currently mitigates rodent activity in two ways:

REACTIVELY (via 311 reports): INCIDENT → 311 REPORT → INSPECTION → INTERVENTION

PROACTIVELY (via consistent activity): CONSISTENT ACTIVITY → INTERVENTION

Current operations are **BIASED** to locations that report more & might **MISS RODENT ACTIVITY**

→ **OBJECTIVE:** Equitably predict and explain rodent activity across Boston

DATA

SPATIAL DATA

ENVIRONMENT
Parks, sewers, restaurants, trash pickup, sidewalks

BUILDINGS
Condition, age, material

CENSUS DATA
Population density

RODENT DATA

311 REPORTS

VIOLATIONS

PROACTIVE BAITING

METHODOLOGY

SAMPLING
GOAL: Collect Ground Truth Data to Train and Validate

1) Find most informative locations

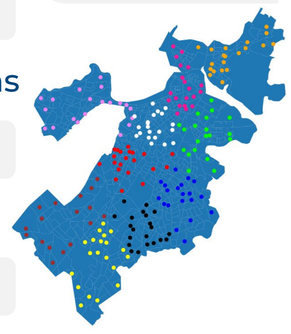
Select 1 Point per Block
↓ 600 candidate locations

Max(Spatial Uniformity)
↓ 400 intermediate locations

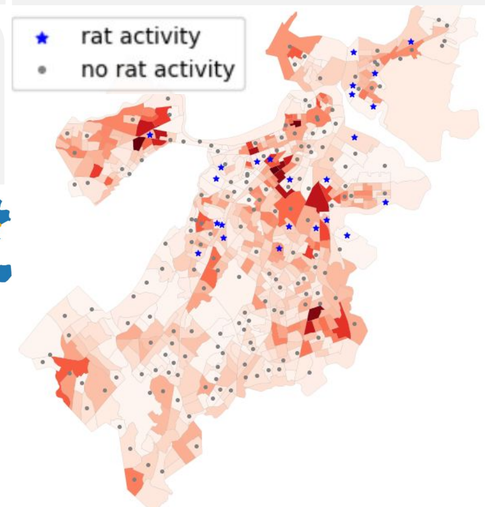
Max(Feature Uniformity)
↓ 200 final locations

Cluster

"We're definitely hitting areas in the city where people don't call in." - John Ulrich



2) Collect data



DEBIASING
GOAL: Quantify and Offset Bias in 311 Complaints

1) Estimate Reporting Bias per Tract

311 sidewalk reports → True sidewalk problems

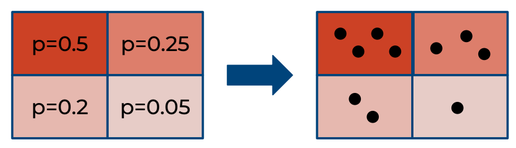
p_t

Proportion of sidewalk problems that are reported in tract *t*

Why sidewalks?
- Have ground truth and 311 data
- Similar "inconvenience" as rodents

2) Pseudosample Negatives

Sample **synthetic negatives** ("pseudosample") from each tract, **proportional to tract's reporting bias**



Negatives have **the same sampling bias** as positives

↓

Prevents model from learning bias

DESCRIPTIVE RODENT ACTIVITY MODEL

GOAL: Describe Drivers of Rodent Activity

FEATURES
- Spatial data
- Census data
- Building data

LABELS
SUMMER 2024
- 311 Complaints
- Violations
- Proactive Inspections
- Sampled Absences

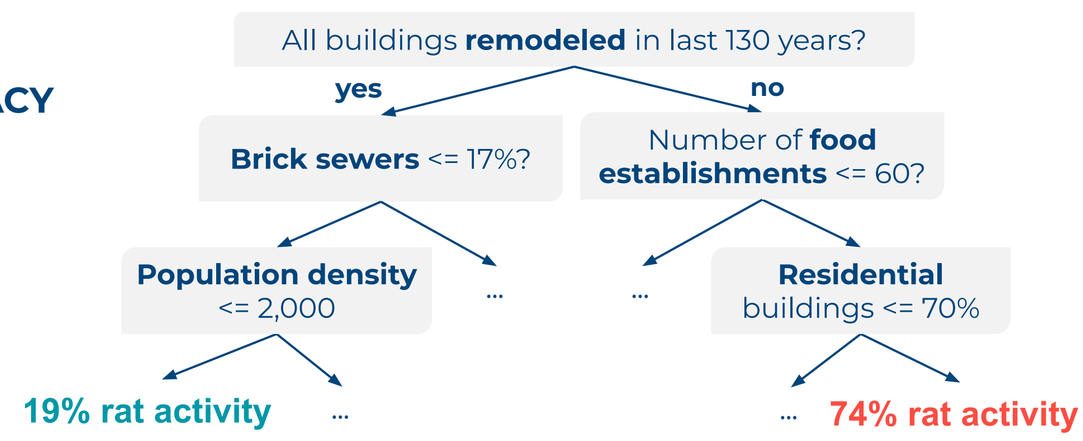
CLASSIFICATION TREE
WITH INTERPRETABILITY

RESULTS

77%
ACCURACY

82%
AUC

CLASSIFICATION TREE



PREDICTIVE RODENT ACTIVITY MODEL

GOAL: Predict Actionable Rodent Hotpots

FEATURES
- Spatial data
- Census data
- Building data
- Reporting bias metric

TRAIN LABELS
SUMMER 23 - SPRING 24
- Violations
- Proactive Inspections
- Pseudosampled Absences

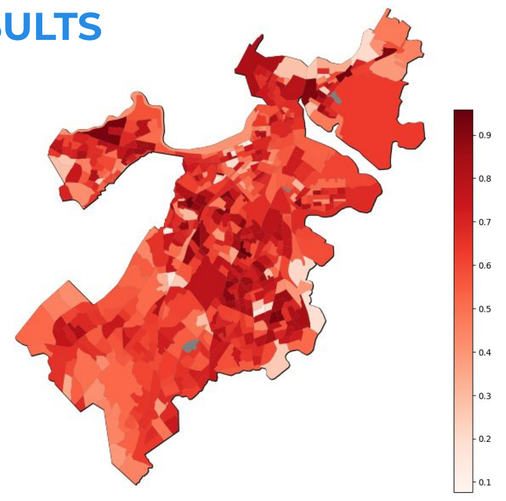
TEST LABELS
SUMMER 24
- Violations
- Proactive Inspections
- Sampled Absences

Informed by **Reporting Bias** from Debiasing methodology

RANDOM FOREST
WITH TIME SERIES CV

EVALUATION & RESULTS

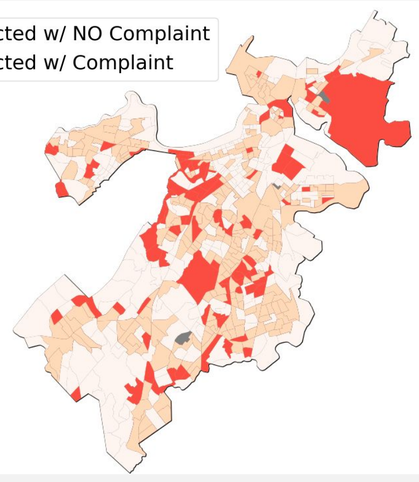
PRECISION
=
% of predicted locations have actionable activity



IMPACT

Support equitable rodent mitigation

Identified **96** non-reporting areas with high predicted risk of rodent activity



Increase inspection success rate

72%
MODEL PRECISION

↓

Increase from **baseline of 30%** from 311 complaints & baiting

Facilitate dialogue between departments

Identified & Described

Environmental causes of rodent activity