

CAPSTONE PROJECT: FINAL REPORT

LOST PETS LOCALISATION AND SEGMENTATION

1. PURPOSE

2. INTRODUCTION

Missed and rejected pets have become a severe problem in many cities. Pets owners usually take a long time to find them or even they may never find their pets because, for example, these pets might be moved far from where they used to live. On the other hand, ONG's have high difficulty to rescue rejected pets, so these ONGs are not often alerted about them or when are these pets have already moved to another place.

The mobile phones have already become part of the day-to-day of people's lives, who might provide using their mobile phones the geographic location and a photo of these missed or rejected pets. In its turn, pets owners and ONGs could search and visualize the geographic location of missed or rejected pets including points of references such as cafes and restaurants. Furthermore, governmental agencies could visualize and compare areas of the city with the highest incidence of missed ou rejected and could segment areas to uncover patterns which could guide marketing campaigns to mitigate the problem.

3. DATA DESCRIPTION AND ACQUISITION

This prototype will make use of the following data sources:

Animal Services of The City of Toronto

Data will be scraped from <https://www.toronto.ca/data/mls/animals/strayanimals.html>

The Stray Animals Report provide by The Animal Services of The City Toronto displays stray animals (cats and dogs) received in the last 5 days. This report contains the following information:

- **Date:**
- **Breed**
- **Approximate Age**

- **Sex**
- **Colour**
- **Receiving Shelter**
- **Animal ID Number**
- **Crossing Intersection**

Localisation of Receiving Shelters

Data will be scraped from link

Toronto Venues nearby Crossing Intersections from FourSquare API (FourSquare website: www.foursquare.com)

I will be using the FourSquare API to explore neighborhoods in selected towns in Singapore. The Foursquare explore function will be used to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. The following information are retrieved on the first query:

Venue ID Venue Name Coordinates : Latitude and Longitude Category Name

Another venue query will be performed to retrieve venue ratings for each location. Note that rating information is a paid service from FourSquare and we are limited to only 50 queries per day. With this constraint, we limit the category analysis with only one type for this demo. I will try to retrieve as many ratings as possible for each retrieved venue ID.

4. METODOLOGY

5. SEGMENTATION AND CLUSTERING

6. DISCUSSION AND CONCLUSION