

Advanced Database Demo

Nearest Neighbor Searches System in Spatial Databases Using R-Trees

Zhuoran Wu

Data

Yelp Business Data

174,000 Business

59106 Category

Around the World (Mostly in U.S.
and Canada)

Data Info

ID	business_id	name	neighborhood	addresses	city	state	postal_code	latitude	longitude	stars	categories
9	EsMcGiZa QuG1OOvL 9iUFug	"Any Given Sundae"	nan	"2612 Brandt School Rd"	Wexford	PA	15090	40.61 51022 445	-80.09 13487 465	5	Coffee & Tea;Ice Cream & Frozen Yogurt ;Food

Data Info

ID	business_id	name	neighborhood	address	city	state	postal_code	latitude	longitude	stars	categories
9	EsMcGiZa QuG10Ov L9iUFug	"Any Given Sunda e"	nan	"2612 Brandt School Rd"	We xfor d	PA	15090	40.61 51022 445	-80.09 13487 465	5	Coffee & Tea;lc e Cream & Froze n Yogurt ;Food

System

Search Based on Info

Back-end: Flask (RESTful)

Front-end: HTML + JQuery

Database: MySQL

The Methods

From MySQL

Select 5 nearest businesses according to Great-Circle Distance

From Rtree

Build a Rtree based on (business_id, (latitude, longitude)) As each tuple.

Then Select from MySQL according to business_id.

From my own Rtree

Write a Rtree tools and use the similar way as before.

Still Working on...

Experiment Result

From MySQL directly:

- Select one random business:
 - Real Time: **0.692s**
 - System Time: **0.015s**
- Select 10 random businesses:
 - Time: **7.031s**
 - System Time: **0.19s**

From Rtree package:

- Select one random business:
 - Real Time: **0.242s**
 - System Time: **0.002s**
- Select 10 random businesses:
 - Real Time: **2.130s**
 - System Time: **0.020s**



Demo

Future Work

Today Work

Build Rtree

Write as API

Search based on star and cuisine

Score result based on weight

Reference

[1] Jens Even Berg Blomsøy, *Evaluating Algorithms for Nearest Neighbor Searches in Spatial Databases Using R-Trees*, Norwegian University of Science and Technology.

[2] Rtree Package <http://toblerity.org/rtree/>

[3] Yelp Data <https://www.kaggle.com/yelp-dataset/yelp-dataset/data>





Thank You

Zhuoran Wu <zw118@georgetown.edu>