BIA 660 - Project: TEAM 5

TABLE OF CONTENTS

Project Proposal

1. Change Points

"Can you detect when things change suddenly (i.e. a business coming under new management)?"

2. Events

"Can you see when a city starts going nuts over cronuts?"

Algorithm

Refer Presentation Slides (under '/Presentation' folder)

Run Program

To run the program,
1. Download and Copy the yelp dataset to /scripts/data folder
2. open the '/scripts' folder in the terminal window and execute the RUN COMMAND (given below).

The program will take ~2 mins to execute completely, so please be patient.

Run Command

python Main.py --bucket quarter --averagestars 2.0 --confidencescore 74.9 --granularity general --kneighbours 5 --eventthreshold 300 --dispalyeventforbucket 2016-Q2

Run Command Configuration

Command	Value options	Default value	Description
--bucket or -b	day, month, year, quarter, semester	quarter	It is the date range, by which the data is partitioned and grouped
--averagestars or -a	0.0 to 5.0	2.0	A threshold value to filter changepoints that are less than equal to this value
--confidencesc	0.0 to 100.0	75.0	A threshold value to filter

ore or -c			changepoints that are less than equal to this value
--reviewcount or -rc	0 to 50	10	A threshold value to filter business that have less than this value
--traindataset or -t	File name	reviews-train	Training Dataset file name
--granularity or -g	general, detailed	general	Value specifies output detail level
--kneighbours or -k	Whole Numbers 1, 2, 3, 4 ...	5	Number of neighbours to use in KNN
--eventthreshold or -e	Greater than 0.0	300	A threshold value to filter events that have less than this value
--eventstartbucket	Bucket	2015-07-01	Value used to find events starting this bucket
--eventendbucket	Bucket	2015-12-31	Value used to find events ending this bucket
--dispalyeventforbucket	Bucket	None	Value used to display all events for this bucket. If the value is None then display events for all buckets.

Solution

Refer PDF under '/Results' folder