Final Project
Computational Data Science
Professor Nitin Upadhyay

Is movie a Box Office Success? Analyzing Regional Search Queries to Predict Commercial Success

Prepared by
Maharshi Vyas 201401414
Keval Shah 201401418

Introduction

Previous Work

IMDB rating Prediction from Google Trends

Movie Success Prediction from IMDB, Wikipeadia

Movie Income Prediction using Google Trends

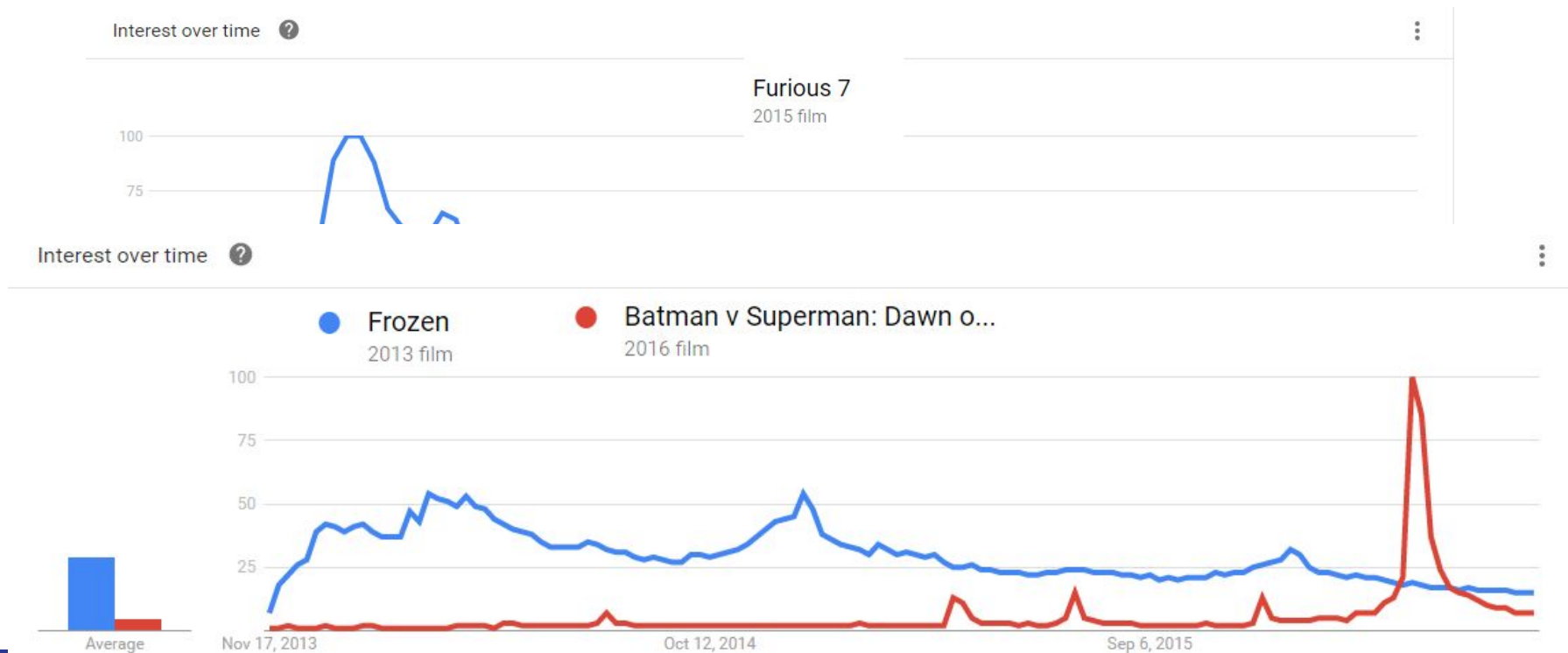Predicting Regional Interests.

Feature Selection

Google Trends:

Concept of SPI.

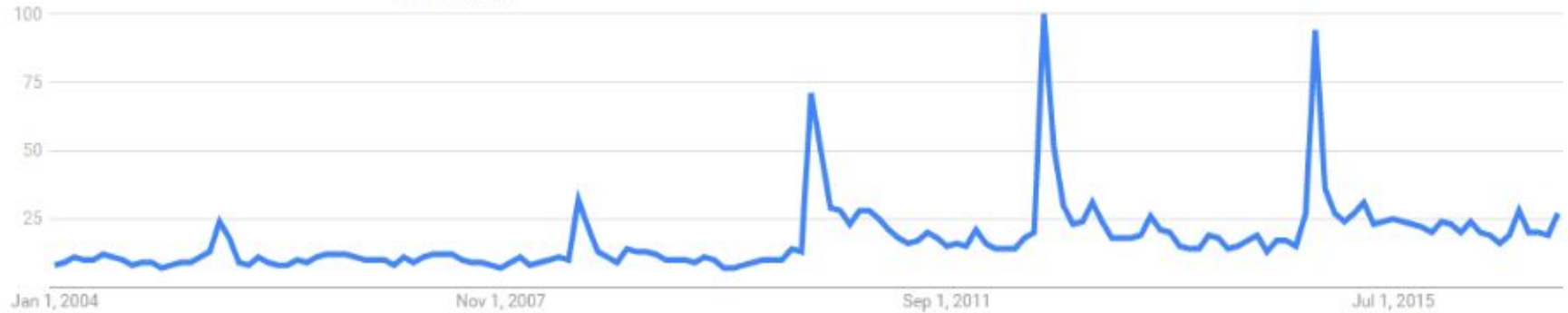Limited Use of SPI.

Introduction of RPI.

Feature Selection

Feature Selection

Interest over time

Christopher Nolan
Film director

100		
75		
50		
25		

Jan 1, 2004   Nov 1, 2007   Sep 1, 2011   Jul 1, 2015

Selected Features

Movie Name Query(M)

Director Name Query(D)

Actor #1 , Actor #2 Name Query(A1, A2)

Genre #1 ,#2, #3 Name Query

Number of theatres

Number of weeks lasted in theatres

Average over a time.For each country independent
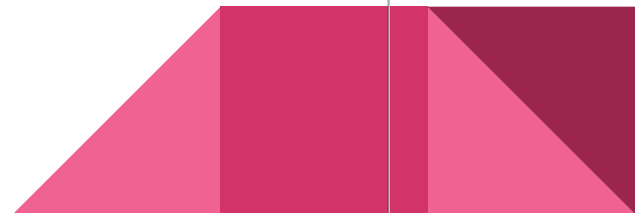
What to Predict?

Popularity Index

Predicts average popularity or buzz of movie for people in a country.

False assumptions.

Independent Models for each Country.
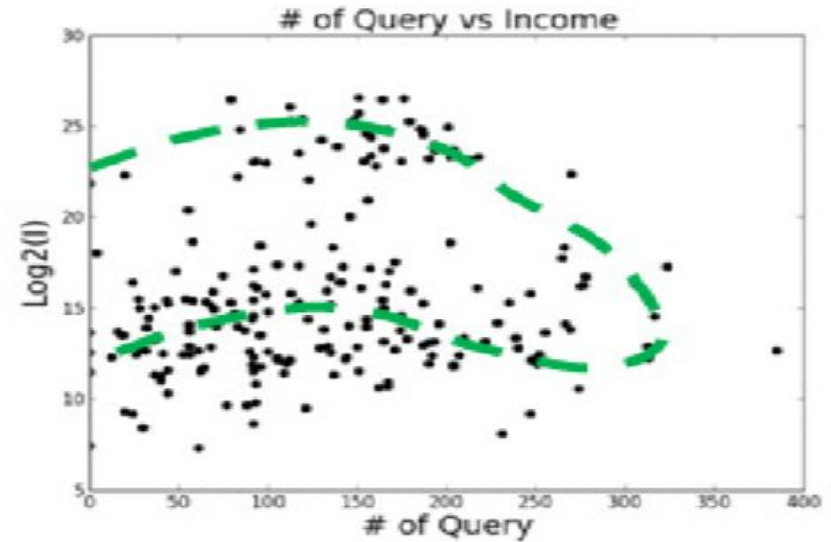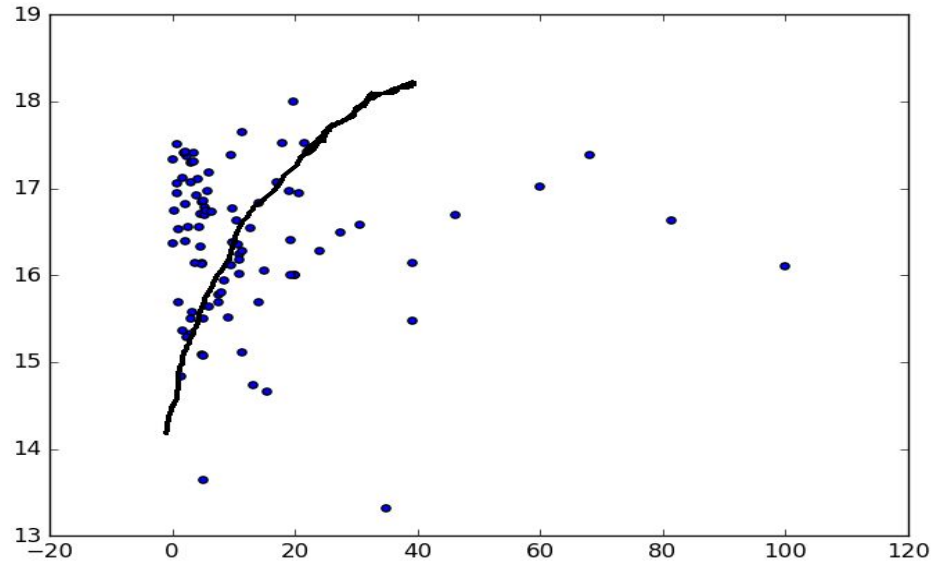
Overall Process.
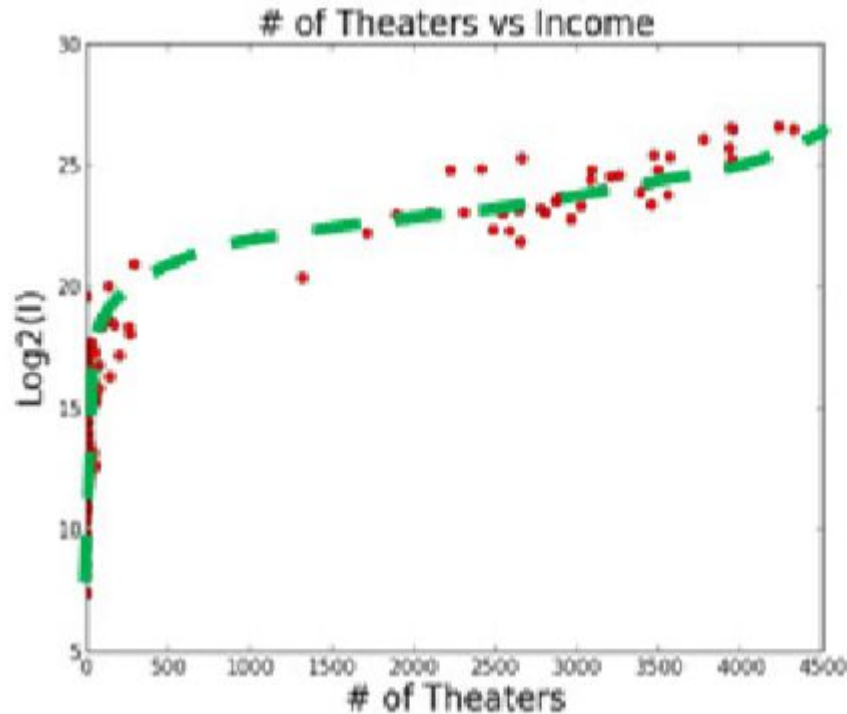
Prediction Methods and Results

		Linear Regression		Multilayer Perceptron		Radial Basis Functions	
		MSE	AME	MSE	AME	MSE	AME
US	Training	0.656006755	0.706006755	0.675944392	0.724504018	0.598553145	0.610811212
	Test	0.806763555	0.796743456	0.804633318	0.971971723	0.768878446	0.932416534
UK	Training	0.842786326	1.321880363	0.85863122	1.384144925	0.754562831	1.183944255
	Test	0.986593115	1.337418571	0.84043669	1.036522114	0.796157591	0.931764797
Germany	Training	0.86457465	1.262280643	0.847115928	1.354640248	0.775404355	1.223634987
	Test	0.957412408	1.338742732	0.783195942	0.991949837	0.738259418	0.886560773
Australia	Training	0.694468862	0.977721815	0.697060679	0.996316943	0.615959965	0.888159089
	Test	0.844290305	1.169708347	0.974795286	1.798722497	0.951269969	1.704012737
France	Training	0.694468862	0.977721815	1.047648168	1.852270297	0.901713737	1.466307775
	Test	0.844290305	1.169708347	0.877028686	1.562037831	0.85917394	1.477809933

Conclusions

SPI vs RPI

Conclusions,Discussions and Future Work

# of Theaters vs Income

References

[1] Deniz Demir, Olga Kapralova, Hongze Lai, "Predicting IMDB Movie Ratings Using Google Trends," Dept.Elect.Eng,Stanford Univ., California, December, 2012

[2]PREDICTING MOVIE SUCCESS FROM SEARCH QUERY USING SUPPORT VECTOR REGRESSION METHOD
Chanseung Lee1 and Mina Jung2

[3] Predicting Movie Success Based on IMDB Data Nithin VR, Pranav M, Sarath Babu PB, International Journal of Data Mining Techniques and Applications
ISSN: 2278-2419

[4] Choi, H., Varian, H., (2012). Predicting the present with google trends. Economic Record, 88 (s1), 2–9

Thank You