

Forecasting Patent Filing Volumes using Google Patents Data

Vishnu Murthy

Thomas Jefferson High School for Science and Technology
Computer Systems Lab

vishnurmurthy@gmail.com

September 7, 2018

Introduction

- United States Patent and Trademark Office (USPTO)
- Patent Examiner
 - Scientifically literate
 - Specialized
 - Search for existing claims
- Patent Application Process
 - Multiple rounds of Patent Examination
 - Back-and-forth with Patent Attorneys
 - Renewable Patents
 - Long Process - can take years to grant a patent
- Cooperative Patent Classification
 - Section, Class, Subclass, Group, Subgroups
 - Human Necessities, Agriculture, Accessories of agricultural machines, Hand tools, Attaching the blades or the like to handles

Motivation

- Large component of USPTO revenue is patent application filing fees.
[1]

- Creating a model that can
 - Forecast the number of patents filed
 - Forecast the number of patents filed per category
 - Be sensitive to fluctuations
 - Accurate

is valuable

- Current models use an econometric for prediction.
 - Forecasting Innovations: Methods for Predicting Numbers of Patent Filings by Hingley and Nicholas[2].
 - Uses economic indicators such as Gross Domestic Product, Research & Development expenditures of various companies
 - Inaccurate
 - Econometric models have difficulty in capturing short-term trends.

- Google Patents.
 - Used for searching existing patents
 - Used by patent examiners, patent attorneysis valuable
- Seasonal Autoregressive Model
 - Predicting the Present with Google Trends, Hal Varian et al. [3]
 - Predicting car sales using Google Trends data

Proposal

- Create a model using Google Patents Data
- Baseline: Seasonal Autoregressive Model

$$\bar{x} = ax_{t-1} + bx_{t-1} + \dots + kx_{t-11} + lx_{t-12} + \text{google_patents_data}$$

where \bar{x} is forecasted patent filing volume for month t , x_t is real patent filing volume in month t , $g_{n,t}$ is the normalized query share for Google Trend number n in month t , a, b, \dots, k, l are trainable parameters.

- Based off of work by Google's Chief Economist Hal Varian [3]
- Data
 - Previous Year's Patent Filing Volume Data: USPTO's Patent Examination Research Dataset (Public PAIR)
 - Google Patents Public Datasets on Google Cloud

Phase 1

- Forecast Number of Patents Filed
- Google_Patents_Data: of searches Google Patents per month


Phase 2


- Forecast Number of Patents Filed for specific patent classifications
- Google_Patents_Data: of searches certain search queries have on Google Patents

Treatments	Response 1
10/1	Get data on # of searches Google Patents has per month
11/1	Filter out those searches by patent examiners via IP address
12/1	Complete Phase 2 - Train Model
1/1	Get data on # Google Patents queries per month
2/1	Complete Phase 3 - Train Model
3/1	Implement USPTO patent search interface data with model
4/1	Continue making improvements on model

References

 O. of the Chief Financial Officer, Aug 2018.

 P. Hingley and M. Nicolas, *Forecasting innovations: methods for predicting numbers of patent filings*.
Springer, 2006.

 H. Varian and H. Choi, “Predicting the present with google trends,”
Apr 2009.

Questions?