International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 5, Issue 3, May-June 2016 ISSN 2278-6856

Envisaging the R & D Impact by correlation of R & D Personnel and Patent Output

K.S. Oza[1], R.S. Kamath[2], R.K. Kamat[3]

[1]Department of Computer Science, Shivaji University, Kolhapur Maharashtra, India

[2]Department of Computer Science, CSIBER, Kolhapur Maharashtra, India

[3]Department of Electronics, Shivaji University, Kolhapur Maharashtra, India

Abstract

Translating technological know-how in terms of financially viability is one of the major attributes for furthering the competitiveness of modern economies in the knowledge era. Patents have been traditionally used to showcase the R & D competency of a country and serve as the significant indicator of the innovation competency of the R & D personnel of the respective country. They are now been increasingly used by the decision makers for the research policy formulation. However evaluating the country specific performance of the patents in light of the R&D personnel involved is no easy task. Moreover while analyzing on a global scale and benchmarking the performances necessitates adoption of standard methodologies such as correlation and clustering as presented in the present communication.

Keywords: R&D, patents, clustering, correlation, weka

1. INTRODUCTION

Innovation, inventiveness, technical novelty, R& D output and its qualitative appraisal of any research community can be significantly evidenced through the patents. Throughout the history of science, patents are used as a useful indicator of the fulfillment of R & D technical objectives, especially if the objectives was to develop a new invention or a significant departure from the prior art[1]. Though the patent statistics is being traditionally used to substantiate the R & D output of any country especially its technological output, its productivity and innovation competence, it is a matter of fact to ponder over the efficacy to convert the R &D output systematically in terms of filing the patents. In this context a pragmatic analysis of the productivity of the R & D workforce can be perhaps envisaged by correlating the scientific workforce of a country standardized in terms of R & D personnel i.e. the persons employed directly on R & D as well as those providing direct services such as R&D managers, administrators, and clerical staff [2] as against the number of patents filed.

Igor Prodan has reported a model which portrays dependency of the number of patent applications on R&D expenditure[3]. Research concludes the positive correlation between patent applications and R&D expenditure, patent applications with a delay which varies from country to country and the quantity of patent applications in developed countries depends more on R&D expenditure. Another paper by Meo and Usmani presented the impact of R&D expenses on research publications, patents and high technology exports among 47 European countries [4]. This study was based on the information regarding per capita GDP, R&D expenditure, scientific journals, technology exports and number of patents. Web of Science, World Bank, Thomson Reuters and SCImago/Scopus were source of information for this research. Yet another paper by Marlena reports comparative study of patent activity and R&D expenditure in European Union[5]. Author has applied Ward's hierarchical cluster analysis for this study.

A comparative study of Research and development spending and patents of India among SAARC and BRICS countries was reported by Janodia[6]. This study reveals that it is essential to increase in R&D expenditure by the Government of India which promotes research leading to innovation, increasing patenting and larger number of publications. The performance in terms of R&D expenditure and patents is strong among SAARC countries, whereas it is miserable among the countries of BRICS. Expenditure on R&D, scientific indexed journals and research publications are the most significant contributing factors towards a knowledge economy which in turn give a boost to patent applications, high technology exports and ultimately GDP[4].

In the backdrop of the research endeavors portrayed above, the present communication aims to correlate the patent activity vis-à-vis the R & D personnel on a global scale taking into consideration the individual country performance. While the computation of correlation has been undertaken herein to imply the R & D personnel in lockstep or otherwise with respect to the number of patents generated, a more significant analysis by means of clustering has also been presented using WEKA (Waikato Environment for Knowledge Analysis)[7] a popular open source free suite of machine learning software written in Java, developed at the University of Waikato, New Zealand available under the GNU General Public License. The said software suite contains tools for data preprocessing, classification, regression, clustering, association rules, and visualization [8]. The datasets used in the present communication have been taken from Knoema [9] a free to use web based public and open data platform launched for the purpose of statistical and infographics analysis.

2. Analysis through Correlation and Clustering

In the present analysis, two datasets have been used viz. patent applications (resident), which is the number of patent applications filed through Patent Cooperation Treaty procedure or with national patent office of the respective country. Second data set is related to Researchers in R&D i.e. the number of researchers engaged in research work. Both the dataset were downloaded and preprocessed for analysis. During data preprocessing selected data sets of the years 2012, 2011, 2010, 2009, 2008, 2005 and 2000 have been used as the datasets for both the parameters i.e. number of patents and R & D personnel are available for the years mentioned. Moreover the same has been done with a view that by taking the year 2000 as baseline the progressive analysis would be much more worthwhile. Table 1 depicts the snap shot of correlated data.

Table1. Correlation of R&D personnel and Patent's filed

Countries	Correlation
Finland	-0.89
Iceland	0.32
Denmark	-0.65
Israel	-0.38
Singapore	0.89
Korea, Rep.	0.11

This correlated data was further clustered for classification of countries to know whether number of R&D personnel's have direct effect on number of patents filed. Data was clustered using an open source data mining tool Weka. Simple K-means clustering algorithm was used as data is of numeric nature. Data is clustered in to three clusters as we wanted to find whether the correlation is positive, negative or neutral. Following figure shows the output of clustering on correlated data.

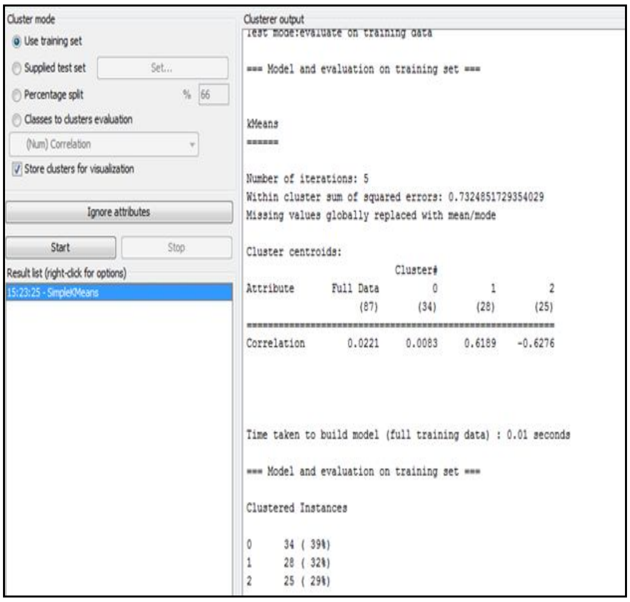
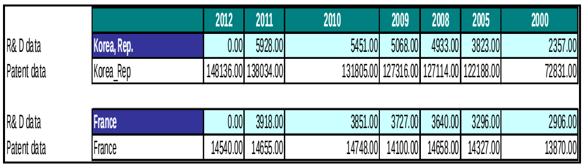


Fig. 1 Clustering output on Weka.

As observed in Fig. 1 three clusters are created to classify data in to three categories. There are 34 countries belonging to cluster-0, 28 to cluster-1 and 25 to cluster-2. It has within cluster sum of squared errors as 0.73 which is due to replacing blank data with zero.

Intra cluster analysis of cluster -0 shows that all the countries belonging to this cluster has shown weak correlation between R&D personnel's and number of Patents filed. This means as that there is very less linear relationship between R&D personnel's and Patent's filed. Here for example Korea and France belong to this cluster with coefficient 0.11 and 0.03.



		2012	2011	2010	2009	2008	2005	2000
R& D data	Korea, Rep.	0.00	5928.00	5451.00	5068.00	4933.00	3823.00	2357.00
Patent data	Korea_Rep	148136.00	138034.00	131805.00	127316.00	127114.00	122188.00	72831.00
R& D data	France	0.00	3918.00	3851.00	3727.00	3640.00	3296.00	2906.00
Patent data	France	14540.00	14655.00	14748.00	14100.00	14658.00	14327.00	13870.00

Fig. 2 Sample data of countries belonging to cluster-0

Here we can observe in Fig. 2 there is linear relationship between R&D personnel's and Patent's filed as the number of R&D personnel's are increasing the number of Patents filed are also increasing. But in the year 2012 there is no data available for R&D personnel's so these countries though have linear relationship have been put under this cluster.

Cluster-1 has all the countries having correlation nearer to +1. This indicates there is strong relationship between number of R&D personnel's and number of Patents filed. Intra cluster analysis shows that as the number of R&D personnels increases number of patents filed also increase this shows linear relationship between R&D and Patents. For example Singapore and Portugal countries belong to this cluster with correlation of 0.89 and 0.98

		2012	2011	2010	2009	2008	2005	2000
R& D data	Singapore	6438.00	6494.00	6307.00	6150.00	5742.00	5292.00	4245.00
Patent data	Singapore	1081.00	1056.00	895.00	750.00	793.00	569.00	516.00
R& D data	Portugal	4781.00	4724.00	4368.00	4166.00	3823.00	2010.00	1624.00
Patent data	Portugal	621.00	571.00	499.00	571.00	381.00	158.00	81.00

Fig.3 Sample data of countries belonging to cluster-1

In the fig. 3 it can be observed that R&D personnel's in year 2000 are 4245 and patents filed are 516 . R&D personnel's has increased from 4245 to 5292 and Patent filed also has increased from 516 to 569 in the year 2005. This trend is observed till 2012.

Cluster -2 is the smallest cluster with only 25 countries belonging to it with correlation coefficient closer to -1. Here it shows negative correlation between R&D personnel's and number of Patents filed. Intra cluster analysis shows that as the number of R&D personnel's increases number of patents filed decreases it shows non-linear relationship between R&D personnel's and Patents filed. Here for example Finland and Germany belong to this cluster with correlation coefficient - 0.89 and -0.87

		2012	2011	2010	2009	2008	2005	2000
R& D data	Finland	7482.00	7423.00	7717.00	7644.00	7686.00	7545.00	6732.00
Patent data	Finland	1698.00	1650.00	1731.00	1806.00	1799.00	1830.00	2579.00
R& D data	Germany	4139.00	4085.00	3950.00	3814.00	3628.00	3246.00	3088.00
Patent data	Germany	46620.00	46986.00	47047.00	47859.00	49240.00	48367.00	51736.00

Fig. 4 Sample data of countries belonging to cluster-2

In fig 4, it can be observed that in year 2000 R&D personnel's were 6732 and patents were 2579 and in year 2005 the number of R&D personnel's increases from 6732 to 7545 but the number of patents filed decreases from 2579 to 1830. This trend can be observed for most of the years.

3. Conclusion

The analysis carried out can help countries belonging to cluster-0 and cluster-2 to re-plan their R&D funds and motivate the R&D researchers to file more and more patents. Although countries fit the correlation, the relationship between patents and R&D personnel does not appear to be straightforward in all of them. Moreover, the evolution of patent activity does not necessarily follow that of R&D personnel. Asthe countries presented in this publication vary in terms of economic size and innovative structure, the analysis should accordingly take these aspects into consideration. This analysis takes into account that R&D personnel includes not only researchers and scientists, but also people providing direct services such as R&D managers, administrators and clerical staff. On the other hand, it should be noted that not all inventions necessarily lead to the submission of patent applications. There are other ways to protect intellectual property, such as copyrights, trademarks and industrial designs. Sometimes inventors prefer keeping their invention secret and do not apply for a patent. Other researchers prefer to place their invention in the public domain in order to freely share it with other users while preventing future patenting by third parties

References

[1]. Bozeman, B., & Melkers, J. (1993). Evaluating R & D impacts. Boston: Kluwer Academic

[2]. Gaillard, J. (2010). Measuring Research and Development in Developing Countries: Main Characteristics and Implications for the Frascati Manual. Science Technology & Society, 15(1), 77-111. doi:10.1177/097172180901500104

[3]. Prodan Igor, Influence of Research and Development Expenditures on Number of Patent Applications: Selected Case Studies in OECD Countries and Central Europe, Applied Econometrics and International Development. AEID., 2005, Vol. 5-4.

[4]. Meo SA and Usmani Adnan Mahmood, Impact of R&D expenditures on research publications, patents and high-tech exports among European countries, European Review for Medical and Pharmacological Sciences, 2014, 18, 1-9.

[5]. Marlena P., Polarization of the European Union: Patent Activity and R&D Expenditure, Journal of Intellectual Property Rights, Vol 18, July 3013, pp330-337

[6]. Manthan D. Janodia, Research and development spending and patents: where does India stand among SAARC and BRICS, Current Science, 2015, 108.

[7]. Cs.waikato.ac.nz,. (2015). Weka 3 - Data Mining with Open Source Machine Learning Software in Java. Retrieved 20 August 2015, from http://www.cs.waikato.ac.nz/ml/weka/

[8]. Jiawei Han, Micheline Kamber and Jian Pei, Data Minining Concepts and Techniques, Third Edition, 2012 Elsevier Inc.

[9]. Retrieved 19 August 2015, from ï□http://knoema.com/atlas/topics/Research-and Development/Patent-applications/Patent-applications-residents