Proceedings of
the 32nd International Business Information Management Association Conference

(IBIMA)

15-16 November 2018
Seville Spain

ISBN: 978-0-9998551-1-9

Vision 2020: Sustainable Economic Development and Application of Innovation Management from Regional expansion to Global Growth

Editor

Khalid S. Soliman

International Business Information Management Association (IBIMA)

Copyright 2018

Table of Content

The Economic Transformation of the World Economy System..... 1
Inessa Sytnik, Artem Stopochkin and Bogdan Sytnik

Le Marketing Territorial En Pleine (R) Evolution ; Vers Un Marketing Identitaire…….. 15
Yasmine Alaoui and Rihab Abba

White Land Tax: The Case of Saudi Arabia..... 31
Nadisah Zakaria, Zaiton Ali and Mohd Zamri Awang

The Impact of Credit Risk Management on Financial Performance of Commercial Banks in Malaysia: Pre- and post-financial-crisis Period…….. 41
Chat F. Tham and F. Hashim

Development of Venture Financing to Ensure Economic Security of a Country…… 51
Ravil Gabdullaevich Akhmadeev, Mikhail Evgenievich Kosov, Olga Alekseevna Bykanova and Tatiana Borisovna Turishcheva

Serviceable Results of Five Automotive Companies: Comparisons Concerning Certain Originally Calculated Indicators..... 57
Alexandru Jivan, Ruxandra Ioana Curea-Pitorac and Diana Tînjală

Subjectivity Analysis of Restaurants Reviews in Myanmar Language..... 76
Yu Mon Aye and Sint Sint Aung

Evaluation of the Future Development of The Digital Economy in Russia……... 88
Alexey V. Bataev, Alexandr A. Gorovoy and Angela Mottaeva

Digital Transformation of the Financial Sector in Russia and The World……. 102
Alexey V. Bataev, Alexandr A. Gorovoy and Angela Mottaeva

Success Factors for Effective Use of KM Systems in SMEs…….. 115
Shiang-Yuan Kao, Peter Busch, Gustavo Guzman and Louis Sanzogni

A Conceptual Framework for Collaborative Cloud Manufacturing..... 128
Tatiana Galutskaia, Peter Busch and Małgorzata Sobińska

Can We Use Big Data Analytics to Leverage Tourism in Rural Tourism Destinations? ……. 145
Carlos R. Cunha, Elisabete Paulo Morais and Catarina Martins

Eliminating Multicollinearity and Considering Microeconomic Indicators in a Multifactorial Model for Predicting Bank Bankruptcy……... 157
Igor Lyukevich, Dmitry Rodionov and Alina Kindeeva

Improving Cyber Security Awareness among Generation X……. 170
Charles Van Der Vyver and Patricia Mcdonald

Conditions for Enhancing the Competitiveness of the Romanian Higher Education System: Case Study…… 182
Simona Rus, Matei Tamasila, Marian Mocan, Larisa Ivascu and Attila Turi

Effectiveness of Social Media for KM Tools in SMEs…….. 191
Keren-happuch Oritogun, Peter Busch and Winnie Picoto

Coping Strategies of Military Students in the Army of The Czech Republic…… 2099
Roman POSPÍŠIL, Radek MITÁČEK, Tomáš ZEMAN, Pavla MACHÁČKOVÁ and Bohumil PTÁČEK

Modernization of the Labor Market in Kazakhstan: Finding Ways of Implementing…….. 2107
Seisembay JUMAMBAYEV

Does Engage Youth in Agriculture Support National Food Sovereignty? : Empirical Evidence from Indonesian Rural Area…… 2115
Vanda NINGRUM

Determinants of Migration Following the EU Enlargement: A Panel Data Analysis…… 2123
Sanja FRANC, Anita ČEH ČASNI and Antea BARIŠIĆ

Private Label Perception of Food Related Lifestyle Segments In Croatia……. 2124
Sandra HORVAT and Nikolina ŽALAC

A Conceptual Model of Social Media Marketing Strategy…… 2134
Andreea PACHIȚANU, Ștefan Claudiu CĂESCU, Margareta Stela Florescu, Andreea Mihaela BARBU and Mihai Cristian Orzan

Using the Methods of the Company's Capitalization Optimal Management……. 2144
Daniel S. DEMIDENKO, Valery V. KULIBANOVA and Vladislav G.MARUTA

Factors Influencing Customer Loyalty towards Private Labels …… 2150
Marie SLABÁ, Petra MARTÍŠKOVÁ and Roman ŠVEC

Service Recovery Influence on Consumer's Complaint Intention to Obtain Benefits. The Case of the Cruise Ship Industry……... 2160
Marcya Stefany GONZALES SANTIAGO and Ricardo Fontes CORREIA

Analysis of India Ecosystem for Startup with Using Data Mining: Settlement of Big Data……. 2176
Aziza ZHUPAROVA, Rimma SAGIYEVA, Dinara ZHAISANOVA

The Impact of New Technologies on Employment in Kazakhstan…… 2184
Seisembay JUMAMBAYEV, Ardak TURGINBAYEVA and Aisulu MOLDABEKOVA

Internal Factors Environment and Small Medium Enterprises (SMEs) Performance: Malaysian Evidence…. 2189
Mohamad Ridhuan MAT DANGI, Aida Hazlin ISMAIL, Shukriah SAAD and Razana Juhaida JOHARI

Intellectual Capital as a Basis of Competitiveness Development: Qualitative Research of Retail Companies…….. 2201
Igor A. ARENKOV, Iana Y. SALIKHOVA and Dinara V. YABUROVA

The Issue of CO2 Emissions in the Context of Sustainable Development - A Regional Approach with the Example of Poland…….. 2212
Piotr KUŁYK and Łukasz AUGUSTOWSKI

Factors that Impact Employee Performance: The Empirical Case of Two Telecommunication Companies in Palestine…….. 2222
Razi AQEL, Paula O. FERNANDES AND Alcina NUNES

Factors Affecting the Innovation Implementation in Small and Medium-Sized Enterprises in the Czech Republic…….. 2233
Jan MRÁZEK

Analysis of India Ecosystem for Startup with Using Data Mining: Settlement of Big Data

Aziza ZHUPAROVA, Rimma SAGIYEVA, Dinara ZHAISANOVA

Al-Farabi Kazakh National University, Almaty, Kazakhstan,
aziza_z@mail.ru, rimmasagiyeva@gmail.com, zhaisanova15@gmail.com

Abstract

In conditions of knowledge-based and digital economy research ecosystem of the startup have to make some decisions every day and in the framework of it analyze factors influencing them and estimate the potential of an ecosystem. An important feature of the ecosystem of startup development is understanding of what field is popular among the investors and estimate the level of involvement of different areas of the country in the startup ecosystem. Data mining techniques are a necessary approach for accomplishing practical and effective solutions to this problem. Indian startup has been an obvious target for big data. This work directly includes the existing literature about ecosystem concept and Big data application.

Keywords: Big data, data mining, startup ecosystem

Introduction

The challenge of modern life requires a flexible approach and quick reaction to changes in the world. Today we are building a training system based on professional standards. However, this situation becomes irrelevent for professionals in the field of knowledge-based and digital economy, because environment requirements change faster than standards. In this way, some researchers believe that we have become information dependents of the twenty-first century, living in an on-command, on-demand Big data world (Hota C. and others, 2015). Big data receives data from all fronts and it would be a helpful decision-making tool to use if it can be analyzed properly sure, but it also has huge data complexity which in turn make the computation of the data complex (Doruk et al, 2016). Besides, during the time of development knowledge management academies and government makers face the problem of decision making in any way. The lack of a synthesized overview of existing definitions and knowledge of the startup ecosystem framework indicates a research gap and poses a challenge in the conduct of further studies (Tripathi et al, 2018). According to this research gap, it is important to understand the potential of Indian startup ecosystem in the context of what type of startups are getting funded in the last few years, and what field is popular among the investors. Big data about the Indian startup scene, which could give the possibility to estimate the funding situation and help to make a forecast in the framework of business development. The main question, which clearly been researched in this paper: Data mining is an instrument for managing and predicting the strategy of development for governments and other involved participants of the startup sphere. Addition to that, some scholars consider, that big data has not only transformed the business models but also paved the way for organizational strategic decision-makers to act in a timely manner (Beath et al., 2012; McAfee and Brynjolfsson, 2012).

Data mining technique includes methods of classification, modeling and forecasting, based on the use of decision trees, artificial neural networks, genetic algorithms, evolutionary programming, associative memory, fuzzy logic and statistical methods (descriptive analysis, correlation and regression analysis, factor analysis, variance analysis, component analysis, discriminant analysis, time series analysis, survival analysis). However, it takes some knowledge of the data being and analyzing.

Literature Review

There are different methodologies developed and evaluated by the researchers all over the world in the field of startup. One of them: Researcher like Srivardhini (Srivardhini, 2018) is analyzed entrepreneurial ecosystem in India. He distinguished main domains, which includes access to markets, finance, and quality human capital progressive policy framework and a range of institutional support as key factors of a strong ecosystem. In the framework of research in the area of 'entrepreneurial ecosystems' (Feld, 2012; Isenberg, 2011; Mason and Brown, 2014; Neck et al., 2004) takes an embedded view of new business ventures and their evolution. In this way, ecosystem approach supports firms to develop in the comprehensive development in the social, cultural and institutional context that increase their possibility for success.

Most development efforts have been focused on the industrialization and commercialization of data technologies and infrastructure, while few efforts have addressed the immense repercussions of the social dynamics and organizational, political, and managerial decisions inherent in the development of big data (Housley et al, 2014). The ecosystem concept is also used in the business field, such as a business ecosystem forming a network of companies that collaborate to produce systems that hold value for customers (James, 1997).

Bahrami and Evans (1995) were one of the first in the academic entrepreneurship literature to invoke the term "ecosystem" in their study of Silicon Valley. At the same time, Spilling (1996: 91) emphasized the "entrepreneurial system," describing it as the actors, roles, and environmental factors that interact to determine the entrepreneurial performance of a region. Because when entrepreneurs identify, create, and respond to market opportunities, they exhibit intentionality in formulating implicit and explicit action plans (i.e., projective linkages of actions to objectives; Encinar& Muñoz, 2006; Zapkau, Schwens, Steinmetz, &Kabst, 2015). In addition, a growing demand for big data is a noticeable note, because, nowadays, big data analytics has become a trendy practice in business intelligence encompassing combination of massive data sets and advanced analytics techniques, and it plays a key role in influencing aspects of business activities and customer choice (Russom, 2011).

At this point, Sheng (Sheng et al, 2017) made an analysis of big data's significance in business and management improvement. He demonstrates in his work that besides the data-driven approach the technological innovation promotes changes in managing an organization, operation, marketing and other business activities. In the process of using big data into management mechanism, it is estimated additional values, and it was created and realized in business development. Furthermore, he distinguishes that firms need to formulate and implement a data-driven strategy and their head of management should organize the strategy taking into account the investments in IT innovation and data analytics skills development. So as high level of techniques to analyze and interpret big data, executives would be able to make decisions in a strategic, timely, and flexible manner, which gives competitive advantages overall.

Methodology

The objective of the proposed work is to analyze the startup environment data using data mining techniques. In the proposed work, startup data has been collected from the following sources: The dataset for research aim [https://kaggle.com]. Input dataset consists of 3-year data for the period from January 2015 to August 2017. It includes columns with the date funded, the city the startup is based out of, the names of the funders, and the amount invested (in USD).

Big data analysis was made by authors using data mining techniques which are illustrated in figure 1. It begins with the tracking, which is meant a recognition of the pattern of the dataset. According to the

dataset about funding startup in India, the description of every indicator was learned. After that, it took place classification, which is used for collection of various attributes together into discernable categories, which you can then use to draw further conclusions or serve some function. Following that at the stage of association, you'll look for specific events or attributes that are highly correlated with another event or attribute. In turn of outlier detection, you have the possibility to exclude outliers in your data, because simply recognizing the overarching pattern doesn't enough for a clear understanding. Subsequently, clustering is necessary for data points. Cluster analysis divides data into well-formed groups (Majumdar et al, 2017). Finally, regression and forecasting used as a form of planning and modeling. Particularly, multiple linear regression (OLS) is the method used to model the linear relationship between a dependent variable and one or more independent variables. The dependent variable is volume funding startup and independent variables are investment type, industry, and location. Purpose of this work is to find suitable data models that achieve high accuracy and show for an interested person the picture of an ecosystem for startup. As it is known, least square method - the most well-known regression analysis method, which supports creating a global variable or process model that you want to study or predict.
Linear multiple regression model:

$$y = a + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m + \varepsilon$$

OLS makes it possible to obtain estimates of the parameters for which the sum of the squares of the deviations of the actual values of the effective characteristic from the calculated ones is minimal:

$$\sum_i \left(y_i - \hat{y}_{x_i} \right)^2 \rightarrow \min$$

At the result of OLS method, it is possible to make interpretation of variables and prediction, which is one of the most valuable data mining techniques, since it's used to project the types of data you'll see in the future.

Panel Data

Figure 2 contains log normal distribution data of funding amount of startup according to investment type for panel data. The data reveals subordination to the normal distribution law and the possibility of application of statistical methods of analysis. It can be clearly seen, that the majority of India startup was funded by private equity and seed type of investment, which may be due to certain conditions of the ecosystem for startup of the country. Moreover, it is noticeable that the share of crowdfunding in the funding amount is minor, that means the necessity to improve infrastructure for crowdfunding or improve another suitable type of investment for the country.

At the next stage of analysis, it is seen that the consumer internet industry has got the great demand among investors which could be connected with the globalization of the world economy and digital technologies. In this way, technology and e-commerce industry take the second and the third place among the investors by funding rate (figure 3).

On the other hand, it is important to take into account the location of startup financing. It is evidently, every region of the country needs support from the government with the aim of development and well-being. Krajcik (Krajcik&Formanek, 2015) described a regional startup ecosystem as follows: «A regional startup ecosystem is an effective method to endorse regional innovations and the development of the business environment along with securing the growth of the domestic product and employment in the given country». So in this way, we researched the visualization of funding rate according to location

(figure 4). As we see, four main cities of India at the top place of ranking because of their relatively well-developed infrastructure. However, it is noticeable that such cities like Vadodara, Pune/US, Chandigarth use the least demand in startup funding.

New players in the market in the face of innovative firms take a key role in the conditions of a knowledge - based economy, because they are one of the sources for creating new jobs, however, they are restricted by the lack of funding. Thus, in addition to our research, we have created the visualization of funding amount according to investors. It is noticeable, that the majority of startups is funded by undisclosed investors and Indian angel network (figure 5). The smaller part of the funding amount takes different venture companies and other private companies.
We chose the regression model in order to analyze the dependence between the variables of the dataset and provide coefficients that influence the result of the funding of startup projects. Moreover, in the process of building a regression model to predict funding rate startups with dummies variables.

Table 1 reveals the regression model results. The control variables Investment type and Investor name reach significant levels, p-value <0.05 and adjusted R-squared > 50%. Oppositely, control variables Industry Vertical, City, and Sub-industry show a weak dependence on funding amount with a low level of statistical significance (p-value >0.05) and R-squared.

The implication is that data mining technique can be useful as a tool for analysts of all the spheres. Particularly, such tool provides policy makers with the opportunity to better create and manage the state program of supporting ecosystem for startup. According to the figure 6, which illustrates the scheme of decision making for policy makers in the framework of supporting ecosystem for startup. As a result, the use of this kind of Big Data analysis has allowed managers to make decisions according to the strategy development in the country and choose appropriate mechanism of funding in appliance with the specific of the economy.

Results

We focused on two relative questions for this study: research perspectives of the Indian startup ecosystem and could data mining be the instrument of strategic management for startup ecosystem development. This work extends previous research on the two specific ways. Firstly, at the result of analyzing Indian startup dataset using data mining techniques we found out, what industry is funded by investors, in which area, and type of investment, which illustrated that innovative possibility of entrepreneurship was developed in the different levels among the areas of India. Secondly, we could estimate the ecosystem of Indian startup according to the dataset, which was analyzed by using a data mining technique in the settlement of Big data. We could be convinced, that Big data can provide benefits such as more focused advertising and marketing of products and services – leading to higher prosperity and more adaptable and sustainable products and services (Ryan and etc., 2018). As such, the information analyzing and big data helps investors and governments in creating the strategy of development and making a decision in the aim of improvement startup ecosystem. Furthermore, in this case, it takes place estimate efficiency of investment type, which influences to funding amount and successful realization of a startup.

We think this is a reasonable assumption for startups since information about the ecosystem of startups could be collected and tend to the right way of a good establishment. On the other hand, investors could find an answer to their questions and information assumption relevant even in such settings.

This study is limited to firms, which are interested in other information about funding, such as stage of financing and other information about factors affected successful funding of the firm. This study essentially parallels the work about the interaction between government and business in the framework of the knowledge-based economy conditions, so additional work is needed in a couple of areas. Besides,

technically it was difficult to process data because of incorrect data format. It took enough time for changing the format of data in order to analyze it. The nature of this dataset does not allow us to explore the efficiency of knowledge-intensive production, and estimate results of high tech innovation in industries. In fact, it is possible to accomplish some cutting-edge data mining with relatively modest database systems, and simple tools that almost any company will have.

However, data mining is the best collection of techniques which is popular among scholars for making the most out of the data. As long as we apply the correct logic, and ask the right questions, we can get conclusions that have the potential to revolutionize everything.

Figures and Tables

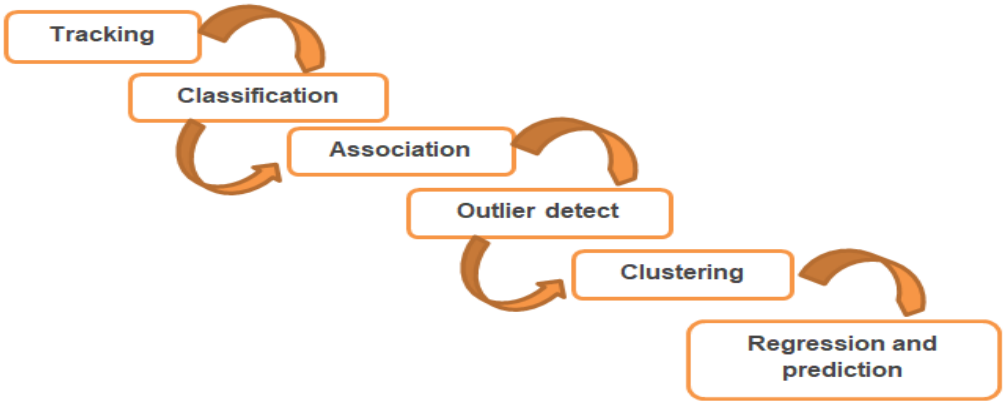

Figure 1: Data mining techniques

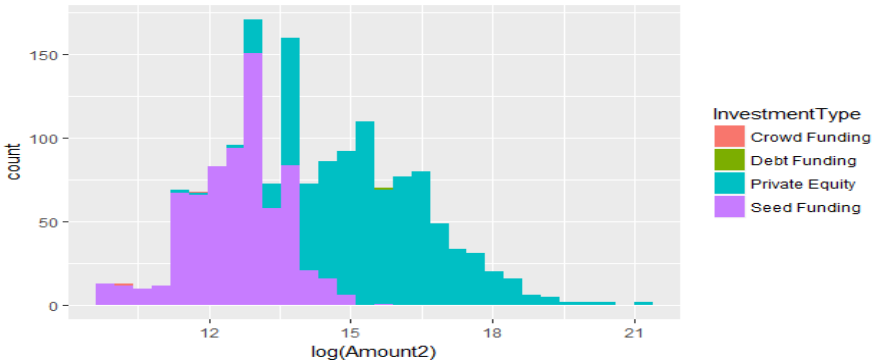

Figure 2: Data funding amount

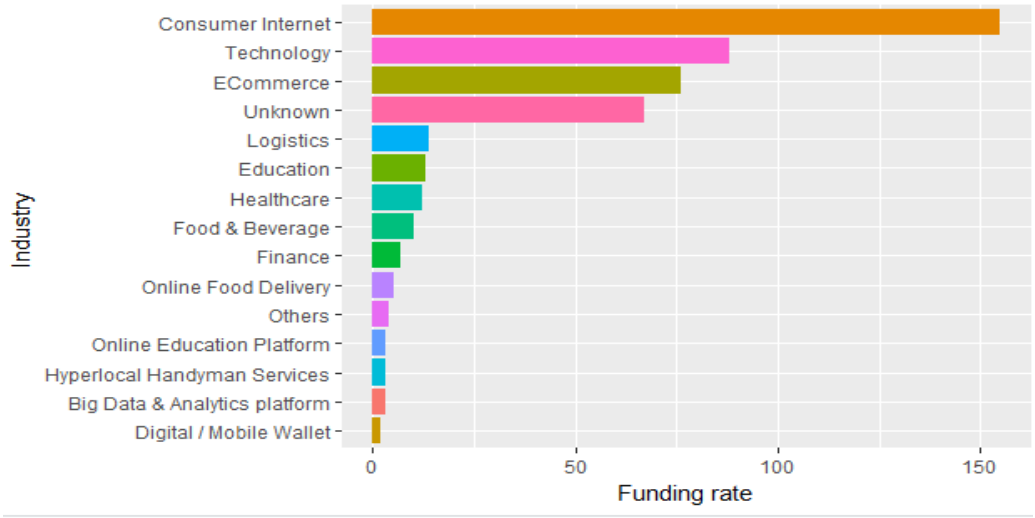

Figure 3: Visualization of funding rate according to industry

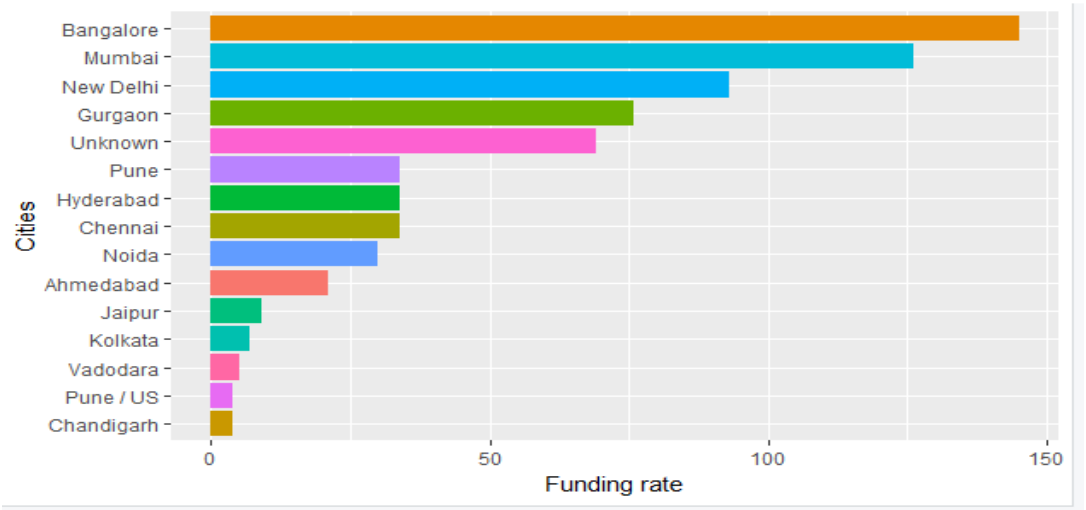

Figure 4: Funding rate according to location

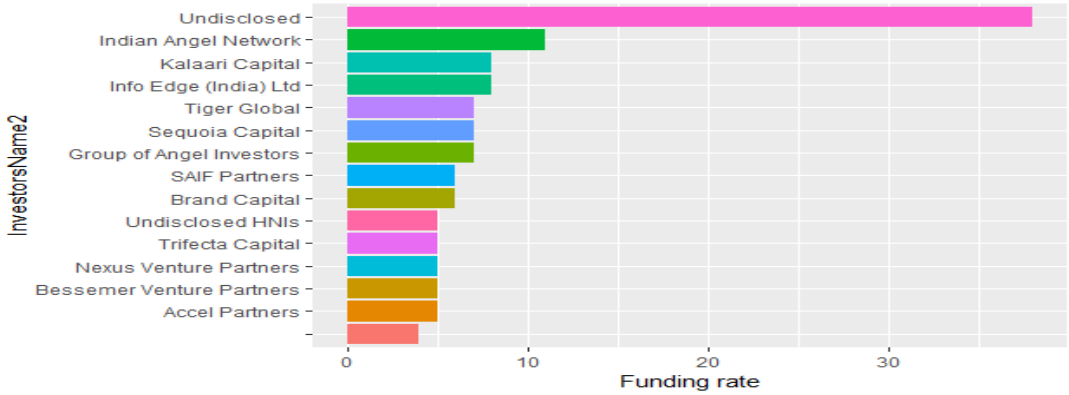

Figure 5: Funding amount according to investors

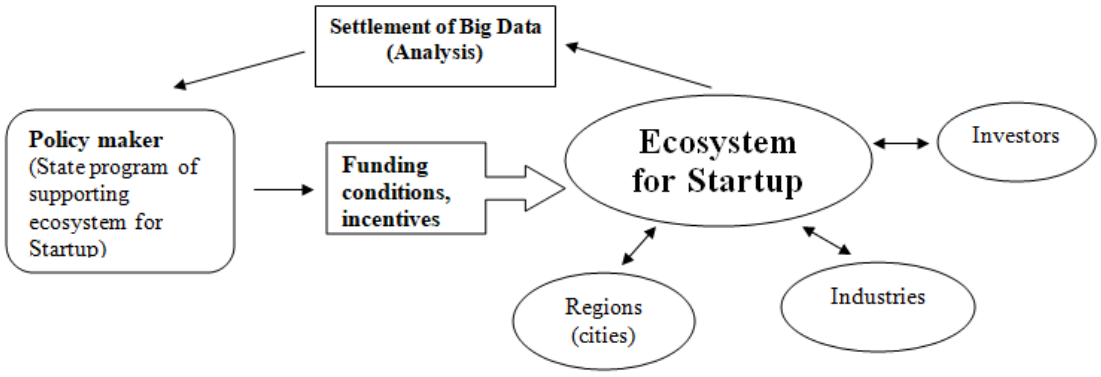

Figure 6: Scheme of decision making for policymakers

Table 1: The results of the regression model of India startup dataset

	Variable name	Type of variable	Adjusted R^2	p-value	F-statistic
1	Amount of Funding	number			
2	Investment type	dummie	58%	2.2e-16	698.8
3	Industry Vertical	dummie	1,4%	0.2889	1.043
4	City	dummie	1,9%	0.005895	1.578
5	Sub Industry	dummie	- 2,8%	0.7635	0.9493
6	Investor Name	dummie	70%	2.2e-16	3.883

References

Bahrami, H. and Evans, S. (1995), 'Flexible re-cycling and high-technology entrepreneurship,' *California Management Review*, 37(3), 62–89.

Beath, C., Becerra - Fernandez, I., Ross, J. and Short, J., (2012), 'Finding value in the information explosion,' *MIT Sloan Manag. Rev*, 53 (4), 18 – 20.

Chittaranjan Hota, Shambhu Upadhyaya and Jamal Nazzal Al – Karaki, (2015), 'Advances in secure knowledge management in the big data era,' *Inf Syst Front*, (17), 983–986.

Sen, D., Ozturkb, M. and Vayvay, O. (2016), 'An Overview of Big Data for Growth in SMEs,' 12th International Strategic Management Conference, ISMC 2016, 159 – 167.

Feld, B. 2012 Startup Communities: building an entrepreneurial ecosystem in your city. Wiley: NJ.

Isenberg, D. (2011). The entrepreneurship ecosystem strategy as a new paradigm for economic policy: principles for cultivating entrepreneurship, Babson Entrepreneurship Ecosystem Project, Babson College, Babson Park: MA.

James F. Moore.1997, The death of competition: leadership and strategy in the age of business ecosystems. Harper Paperbacks.
Tripathi, N., Seppanen, P. Boominathan,G., Oivo, M. and Liukkunen, K. (2018), 'Insights into Startup Ecosystems through Exploration of Multi-vocal Literature,' *Information and Software Technology*, [article in the press],[Retrieved August 20, 2018], https://doi.org/10.1016/j.infsof.2018.08.005.
McAfee, A. and Brynjolfsson, E., (2012), 'Big data: the management revolution,' *Harv. Bus. Rev*. 90 (10), 60 – 68.
Srivardhini K. Jha, (2018), Entrepreneurial Ecosystem in India, *IIMB Management Review*, 30 (2), 179-188.
Mason, C., & Brown, R. (2014). Entrepreneurial ecosystems and growth-oriented entrepreneurship. Final Report to OECD, Paris, 1-38.;
Neck, H.M., Meyer, G.D., Cohen, B., and Corbett, A.C. (2004), 'An entrepreneurial system view of new venture creation,' *Journal of Small Business Management*, 42, 190-208.
Housley, W., Procter, R., Edwards, A., Burnap P., Williams, A., Sloan, L., Rana, O., Morgan, J., Voss A. and Greenhill, A. (2014). 'Big and broad social data and the sociological imagination: A collaborative response,' *Big Data & Society,* July–December 2014: 1–15.
Spilling, O. R. (1996), 'The entrepreneurial system: On entrepreneurship in the context of a mega-event,' *Journal of Business Research*, 36(1), 91–103.
Encinar, M. I., and Muñoz, F. F. (2006), 'On novelty and economics: Schumpeter's paradox,' *Journal of Evolutionary Economics*, 16(3), 255–277.
Zapkau, F. B., Schwens, C., Steinmetz, H., and Kabst, R. (2015), 'Disentangling the effect of prior entrepreneurial exposure on entrepreneurial intention,' *Journal of Business Research*, 68 (3), 639–653.
Russom, P., (2011), Big data Analytics. TDWI Best Practices Report, Fourth Quarter.
Sheng, J., Amankwah-Amoah, J. and Wang, X. (2017), 'A multidisciplinary perspective of big data in management research,' *International Journal of Production Economics*, 97-112;
Majumdar, J., Naraseeyappa, S. and Ankalaki, S. (2017), 'Analysis of agriculture data using data mining techniques: application of big data,' *Journal of Big Data*, 4(20),1-15;
LaBrie, R.C., Steinke, G.H., Lib, X. and Cazier, J.A. 'Big data analytics sentiment: US-China reaction to data collection by business and government,' *Technological Forecasting & Social Change*, [article in the press], [Retrieved August 20, 2018], https://doi.org/10.1016/j.techfore.2017.06.029.
Krajcik, V. and Formanek, I. (2015), 'Regional startup ecosystem,' *European Business & Management*, 1(2), 14–18.