

# A Novel Approach for Successful Delivery of Big Data Projects in Software Industries

Chandrasekhar Banda\*, Phani Kumar S

Department of Computer Science Engineering, GITAM University, Telangana, India

## Research Article

**Received:** 31-Dec-2021, Manuscript No. grcs-22-50690; **Editor assigned:** 03- Jan-20212, Pre QC No. grcs-22-50960 (PQ); **Reviewed:** 05- Jan-2022, QC No. grcs-22-50960; **Accepted:** 10- Jan-2022, Manuscript No. grcs-22-50960 (A); **Published:** 17-Jan-2022, DOI: 10.4172/ 2229-371X.13.1.001

**\*For Correspondence:**

Chandrasekhar Banda, Department of Computer Science Engineering, GITAM University, Telangana, India

**E-mail:**

**bandachandrasekhar2020@gmail.com**

**Keywords:** Prediction; Big data

analytics; Smart; Project management

## ABSTRACT

The changing needs of the people are being supported by the changing technological advancements. The changing technologies have always been a golden duck to the software industry. Big data has forced all the investors to change their mindset and focus on the projects from various domains such as finance, health, and many more on analytics projects based on big data. Though there are various strategies being adopted by most of the software companies, those strategies have failed in successful delivery of big data projects. Hence, in this paper, we have proposed a comprehensive profitmaking framework to implement on Big data and Analytics projects. The proposed model had all the advantageous aspects in delivery the projects successfully within schedule and cent percent customer satisfaction and highest prediction accuracy.

## INTRODUCTION

The term big data has gained a prominent position in the current technological era. The world is worried about big data management (viz., storage), but even processing such huge data has become complex. The software industries have started developing various tools which are capable of holding big data. The area where big data arises is Health care, Finance Marketing, and Education, etc. The software market has huge potential for developing projects in these domains <sup>[1]</sup>.

Nevertheless, the major challenge is developing projects related to big data. Very little research work was done in this area. As per the Gartner survey, it is found that about 80% of the big data projects have failed. The reasons for

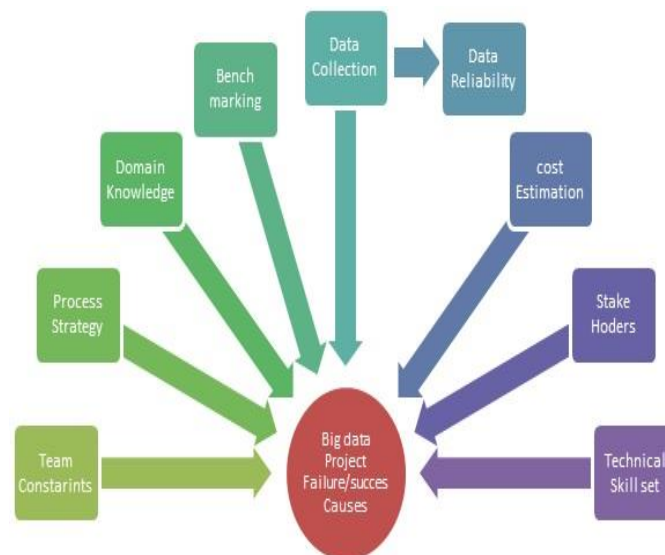
failure have been analysed by various researchers. However, the reasons for failure investigated by them are substantially not enough to reduce the failure rate [2].

Software industries usually followed traditional process models viz., the waterfall model adopted from other sectors like construction or engineering industries. All the traditional process models were observed to be successful when the customer requirements are predefined, static, and known apriori. However, the changing needs, dynamism in requirements, lack of clarity in requirements for the end-user, etc., have forced the software industry to adopt a different approach agile model where the customer is also made as a part of the project development and to a product developed iteratively.

The major focus is to what extent the traditional models suit the current trending projects based on big data. As per the Gartner 2018 report, it is observed that about 60% of the big data projects fail at the initial level only [3]. Though Gartner's report 2018 seems to be too submissive, stating 60% project failure, in reality, it is observed that the failure rate is about 80% [NVP 2019]. The failure of projects might happen for various reasons, such as:

1. Lack of customer needs properly.
2. Not asserting the future requirements of the customer.
3. Software cost estimation is done improperly.
4. Lack of experts in the relevant domain.
5. Data.

**Figure 1.** Factors influencing the big data projects.



According to the literature available (Figure 1), various agencies conducted many surveys; there should be a much more sophisticated, feasible, and robust model that should help the software industries handle big data projects to successfully deliver the products with maximum confidence of the product and the customer well [4].

In this paper, we are attempting to reconcile the aspects which are a point of focus for the software industries while accepting the projects and predicting the software project cost estimation and failing in cost estimation, along with various other factors that are impacting upon the failure of big data projects.







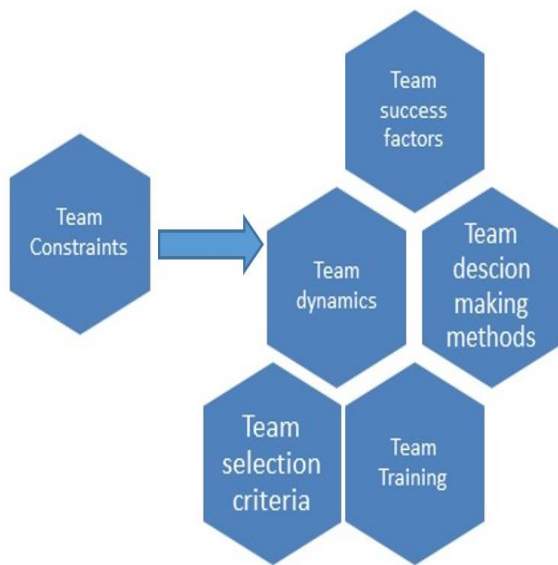








Figure 5. Team constraints.



**Stakeholders:** Stakeholders play a pivotal role in finalizing the requirements from time to time and assessing the project's performance and usability. The project development inputs have to be gathered from various sources such as the end users, customers, and various business organizations working with similar projects. This process shall help enhance the project's performance and provide scope for providing additional features into the projects that shall, in turn, help improve the client's satisfaction levels (Figure 5).

Figure 6. Sources of data for the analytics projects.



In today's analytical world, data is treated to be a vital component in the analytics based companies. So, most of the businesses are turning out to be data-driven. Getting data from various sources is also challenging task. Especially, identifying the right person for gathering the data is the most challenging issue.

“The voice of the customer guides world-class leaders’ every action and decision”

-Yu Sang Chang, George Labovitz, and Victor Rosansky.

The Figure 6 above gives an idea of gathering data from various sources like, client's voice is the first most important aspect, then gathering the information from the business organizations and involving the most important target group i.e., end user who is the ultimate beneficiary. Identifying and selecting an appropriate data collection methods (*viz.*, interviews, surveys, focus groups, observations) helps in improving the projects deliveries more efficiently and successfully.

### DISCUSSION

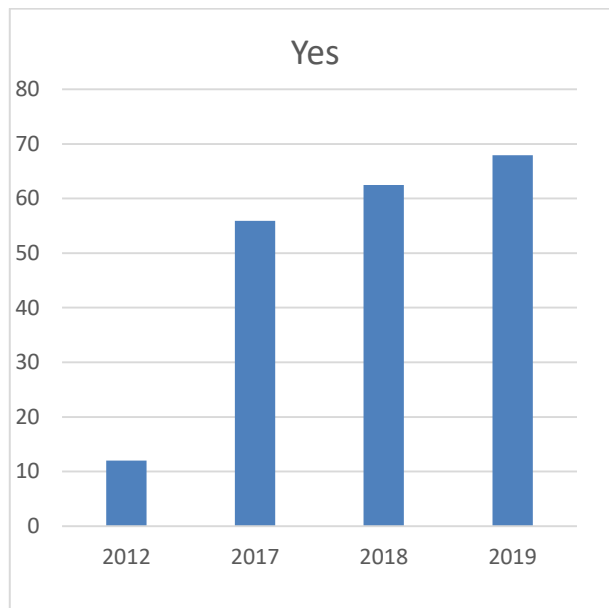
#### Data management

**Specialized persons:** Successful delivery of the software projects highly depend on the expertise and skillset of the employees. Certain projects shall need to deploy highly efficient and specialized persons. Similarly, the projects related to big data analytics shall also need to have deployed specialized persons like Chief Data Officer, Analytics officer. As per our proposed model, the team should consist of a perfect balance of the right people. It is observed that apart from the skilled persons, the big data projects are vulnerable to receive huge amounts of data at a very higher rate with varieties of information, etc. Using the data as is received shall not serve the purpose of the big data projects and may lead to wrong predictions as there might be lot of missing values or other issues which are discussed in the next subsection. Hence, to handle all these issues related to data, it is essential to deploy an expert in data management. As per a survey conducted by NVP, 2017, it is observed that those companies which have employed Chief Data Officer (CDO), Chief Analytics Officer (CAO), Chief Information Officer (CIF), and Head of Big Data (HBD) shall reduce the risk of improper data leading to mispredictions. Deploying experts into the project shall accelerate the pace of the project completion and also increases the accuracy in prediction <sup>[23]</sup>. The recruitment of CDOs has been increased drastically since 2012. Table 4 clearly demonstrates that the deployment of specialized person shall increase the probability of the project getting successfully delivered. The graph in Figure 7 shown represents the increasing in the specialized persons in the company for big data analytics persons.

**Table 4.** Statistics of appointing specialized person in company.

Appointment of a expert	Chief data officer
2012	12
2017	55.9
2018	62.5
2019	67.9

Figure 7. Graph showing increase in appointing CDOs in companies.



**Data treatment:** The term Big Data was used initially coined in about 1990s and has gained wider prominence since 2012 ranging from terabytes to zettabytes transforming to heavier data. Now, the data is increasing drastically with a higher velocity and veracity having various varieties of data [24]. Data has to be obtained from various sources such as from Focus groups, interviews, observations on the domain from the domain experts, etc., as shown in Figure 8. It is essential to identify the right stakeholders for getting the data.

Figure 8: Data collection from various sources.



After acquiring the data, it is also essential to check for reliability of the data. As the data is acquired from survey/questionnaire, there is a chance of encountering reliability issues shown in Figure 9. Reliability issues are found due to the issues like few questions might be irrelevant or ambiguous and few questions may not be understood by the responder. Few questions may be compound. Also, it is very important to analyze the answers/responses in the survey form because of the reason that the answers might be biased or misinterpreting. Hence, data treatment has to be performed to avoid mispredictions.

**Figure 9.** Causes for data reliability issues.



**Required parameters:** Prediction is defined to be a statement made by someone thinks shall happen in future based on certain assumptions or statistics or following certain patterns. But, it is essential assess the necessity of parameters which shall influence the prediction analysis. Prediction accuracy is completely dependent on the parameters used and the patterns identified. The reliability on any analytics projects lies in the accuracy of prediction. The higher the accuracy is, the higher will be the customer satisfaction which shall in-turn help in expanding the business. Retaining or withstanding the confidence of the customer is the most difficult task in the Hence, in this paper we suggest the best way of identifying the parameters.

In our proposed model, the parameters may be classified into 3 levels such as direct linked parameters, indirectly linked parameters and hidden parameters. The details of parameter categorization are discussed in the next section [25-28].

**Project management:** Project management is the most crucial aspect in handling the big data & analytics projects. Any company shall have an assessment model for the PMV (Project Management Values) for the measurement of the performance of project management. The stronger the project management, the higher the cost to benefits of the project. The effectiveness of the project management is determined by the various aspects which can be classified as:

- a. Financial aspects: Financial aspects are those measures which show the variation in Return on Investment (ROI), Productivity, Cost savings, Earnings and cash flow per share, Economic value addition, growth of sales in terms of both increase in volumes of product and the revenue inflow.
- b. Customer aspects: Customer is treated to be the highly prioritized person in the entire project development process. He gains utmost importance among all the people involved in a project. Customer satisfaction, his profitability, customer retention, acquisition of new customers, market share of customer's and his use.
- c. Project or Process aspects: The success of project does not just depend on the financial aspects and customer aspects, but are even dependent on Project and process measures. Project performance, projects risk, etc.
- d. Employee aspects: Successful delivery of any project lies also with employees involved in respective projects. Employee Buy-in is the most essential aspect that shall definitely lead to successfully completing the project leading to timely delivery and 100% customer satisfaction. The growth of any company lies with few more aspects like employee productivity, empowerment, employee's turnover and motivation.

### CONCLUSION

The increase in investments of industries on analytics based projects in software industries signifies that most of the sectors are gearing up predicting the risk well in advance to reduce the loss and also look for increase in profits. As per the statistics revealed by various organizations like Gartner, NVP etc., it is observed that the failure rate of big data based analytics projects is more than 85% which is impacting on the economy of the companies investing on big data projects. There is need for an efficient model which can lead to successful delivery of the big data projects with higher prediction accuracy.

The proposed comprehensive model shall help in driving to successful project development with higher customer retention ratio and reliable tool. The proposed model is based on certain industrial practices and had opted for novel aspects.

The overall project implementation framework has been suggested for the companies investing on big data analytics projects. However, the project management strategy is left for the future scope. The suggested model is being experimented along with the novice PM methodology for future work.

### REFERENCES

1. Bagriyanik S, et al., Big data in software engineering: A systematic literature review. J Glob Inf Technol Manag. 2016; 6:107-116.
2. B. W. Boehm. Software Engineering Economics, Prentice Hall, Englewood Cliffs, NJ, USA. 1981.
3. Putnam L.H. A general empirical solution to the macro software sizing and estimating problem. IEEE TRANS. SOFTW. ENG. IEEE T SOFTWARE ENG.1978; 4:345-361. [Crossref][Google Scholar]
4. Reihaneh H, et al. Uncertainty in big data analytics: survey, opportunities, and challenges. Journal of Big Data. 2019; 6:44.

5. Feras A, et al. Predicting failures in agile software development through data analytics. *Softw Qual J.* 2018; 26:49-66.
6. Vikram S, et al. Project Analytics to Improve Project and Portfolio Decision Making, Project Management National Conference, India. 2017.
7. Mikalef P, et al. Big data analytics capabilities: A systematic literature review and research agenda. *INF SYST E-BUS MANAG.*2018; 16:547-578.
8. Thurber M. A Holistic Framework for managing data analytics projects. Whitepaper. 2017.
9. Premjith BPK. How to make an Agile Teamwork for Big Data Analytics, Whitepaper.2017.
10. Yonatan H. My best tips for Agile Data science research, Whitepaper. 2018.
11. Ori Cohen Data–Science? Agile? Cycles? My method for managing data-science projects in the Hi-tech industry, Whitepaper.2018.
12. Lewis M. Project Management Methodologies for Big Data Analytics, Whitepaper.2019.
13. Jiwat. What roles does Big Data have in shaping the future of Project Management? (Part-A), IPMA. 2017.
14. Jiwat Project Eco-system: Designing future developments through Big Data analytics (Part-B), IPMA. 2017.
15. Michael Bloch, et al., Delivering large-scale IT projects on time, on budget, and on value. *McKinsey on Business Technology.* 2012; 27:1-7.
16. Max Henrion. Why most of the Big Data analytics projects fail, *ORMS Today.* 2019.
17. Thomas H, et al. Big Data Executive Survey 2017, New Vantage Partners LLC. 2017.
18. Hassan N, et al. Predicting Software Projects Cost Estimation, Based on Mining Historical Data. *ISRN Software Engineering.* 2012; 1-8.
19. V R Basili, et al., Software Errors and complexity: an empirical investigation. *Commun. ACM.* 1984;27:42-52.
20. Halstead M H *Elements of Software Science.* Elsevier North-Holland. 1977.
21. Chidember SR, et al. Metrics suite for object-oriented design. *IEEE TRANS. SOFTW. ENG. IEEE T SOFTWARE ENG.* 1994; 20:476-493.
22. Stefan Studer, et al. Towards CRISP-ML (Q): A Machine Learning Process Model with Quality Assurance Methodology. *CoRR abs/2003.05155 (2020), arXiv: 2003.05155.*
23. McCabe T.J. A Complexity measure. *IEEE TRANS. SOFTW. ENG. IEEE T SOFTWARE ENG.* 1976; 2:308-320.
24. Fei YY, et al. Improvements about Halstead model in software science. *J Comput Appl.* 2004; 130-132.
25. Kosarenko Y. The majority of business analytics and AI projects are still failing, in data driven investors. 2020.
26. <https://www.datadriveninvestor.com/2020/04/30/the-majority-of-business-analytics-and-ai-projects-are-still-failing/#>.
27. Zicari RV, et al., Setting Up a Big Data Project: Challenges, Opportunities, Technologies, and Optimization. In: Emrouznejad A. (eds.) *Big Data Optimization: Recent Developments and Challenges. Studies in Big Data.* 2016:18.
28. Center for Business Practices Measures of Project Management Performance and Value, A Benchmark of Current Business Practices. 2005.