

Prediction method for classification of software engineering projects performance in Nigeria

Abstract

Software engineering projects in Nigeria have been classified generally as failure, challenged or successful with no proof that the projects really fall into these categories. The main focus has been on cost and time overrun and attention has not been given to check whether projects truly fall within the given categories. Discriminant analysis was employed to determine how the 30 selected projects in public sector in Nigeria can be correctly classified. This study developed a method of determining the actual category software engineering projects belong to giving the characteristics of projects as failure, challenged, and successful. The developed model was used to reclassify the 30 projects and it was discovered that 21 projects were correctly classified giving 70% of correctly classified projects while 9 were wrongly classified giving 30%. It is possible for projects to satisfy the established success criteria (requirements met on time and within budget) whereas the stigma of failure still exists in its result.

Keywords: Software engineering projects, project performance, project failure, challenged project, successful project, discriminant analysis

1. Introduction

Software projects success/failure rate has been on the increase since 1994 when Standish Group revealed the first ever Chaos report of software project woes. Many organizations had carried out studies to ascertain the level of success/failure rates in software engineering projects. Standish Group (2015) studied 50,000 software projects in 2014 and reported that 71% of the projects either challenged or failed to accomplish the project objectives and only 29% were successful. The report also revealed that large projects had higher failure rate of 94%, medium projects had 91% failure rate while all projects put together had 71% failure. Over 50% of the projects studied had cost and time overrun. According to Florentine (2017), Innatas by Planview carried out studies in 2013 on success/failure rates of software engineering projects and reported that 50% of the projects studied experienced failure within the last 12 months. McKinsey and Company in collaboration with the BT Centre of University of Oxford studied 5400 IT projects in 2012 and concluded that 45% of the projects studied experienced cost overrun, 7% experienced time overrun and 56% delivered less functionalities than expected (Bloch, Blumberg & Laartz, 2012). Software project performance has been a thing of concern to stakeholders due to the resources invested into such projects.

PMI (2013) tagged a project successful when it is completed on schedule and within budget, that is, the resources assigned from the inception were enough to carry out the project and the features and functionalities specified for the project were implemented. Montequin , Cousillas, Alvarez , and Villanueva (2016) said that challenged projects are those either

delivered late or over budget with fewer features and functions than initially specified and failed projects are cancelled before completion or implementation. Most software projects reported in Standish Group Chaos research as failure or challenged experienced time and cost overrun and content deficiency or incomplete features and functions (Standish Group, 2015). On time delivery of software engineering projects and within budget is becoming more important to stakeholders at this era of fast moving business environment. The duration and budget of software engineering projects are the main concerns of stakeholders (Davis, 2014). Project time and cost overrun is a globally well-known problem in every industry such as Road construction (Senouci, Ismail & Eldin, 2016), software (Suliman & Kadoda, 2017), and building (Bowen, Hall, Edwards, Pearl & Cattell, 2002). Cost and time overrun instigate projects failure, abandonment, and challenge and stakeholders are disappointed when a project could not achieve the desired objectives (Davis, 2014). Unfortunately many software engineering projects in Nigeria exceed scheduled duration and estimated budget (Eberendu, 2015). Eberendu (2015) also discovered that over 86% of software engineering projects in public sector in Nigeria had experienced poor project performance. Even the projects classified as failure, challenged or successful may not be as categorized. Project managers capitalize on already known factors to determine whether a project had failed, challenged, or succeeded. It is likely for a project to fulfill all established success criteria (requirements met on time and within budget) whereas the stigma of failure still exists in its result. Also perception can be reality, that is to say, project stakeholders might think a project has failed but chances are good that it has failed or it has not failed. The study tries to answer the question of how the performance of software engineering projects in Nigeria can be classified to determine whether a project has failed, challenged or succeeded. In order to achieve this, a model is developed to predict the category a project belongs when it reaches the deadline.

2. Related Literature

The definition of the outcome of a project depends on the expertise of the project team and the type of project they are handling. Mir and Pinnington (2014) saw on-time and within budget as been of great concern to project organizations and stakeholders are dissatisfied when projects are not delivered as expected. One of the main causes of failure in software project development is cost overrun and developers are concerned with project overall profitability (Montequin, et al, 2016). According to Senouci et al (2016), project cost is determined from the cost of individual project activities and the cost baseline is derived from the estimated cost and the actual cost of the project. Software project organizations need to have a way of estimating cost during software development (Suliman & Kadoda, 2017) and the cost estimation method is used to ensure that projects are developed within estimated budget (Rezaian, 2011)

Time management is a success factor in gaining competitive advantage in today's business. The time or duration of a project represents the scheduled dates of each project activities and their milestones (Ebbesen & Hope, 2013). Beleiu, Crisan and Nistor (2015) opined that most clients consider timely completion of a project as a major factor of project success. According to Bloch et al (2012), the time a project activity or task can take depends on some factors such as the competence and skills of the development team, and availability of resources (Hussein, 2013). Jørgensen (2014) said that software projects have been criticized due to failure to deliver projects within deadline and it takes a well-organized project team to deliver projects on time and within budget. If project team reduces project duration, there is

need to increase manpower (labor), tools and equipment, improved management (Savolainen, Ahonen & Richardson, 2012), and this increases the cost of completing the project (Jørgensen, 2014). Senouci et al (2016) concluded that when cost is tampered with in project development, duration and quality are compromised. A direct relationship exists between project cost and effectiveness (Bowen et al., 2002). On the other hand, minimizing cost jeopardizes deadlines too. Proper utilization of project time and cost determines how the product should be (Suliman & Kadoda, 2017). El-Emam and Koru (2008) revealed that projects that experience cost under run are usually behind schedule. According to Atkinson (1999), assessing the extent to which projects follow schedule and budget is an indication of whether the team will meet stakeholder expectations or not. According to Bowen et al (2002), project managers usually classify the performance of projects depending whether the projects was delivered on-time and within budget.

3. Methodology

A total of 30 software engineering projects in Nigerian public sector were selected and analyzed based on what was tagged “project status”: successful, failed, or challenged. Discriminant analysis was used to describe the characteristics that are specific to distinct project groups and classify the selected projects into pre-existing groups (failure, challenged, and success). Discriminant analysis is an approach recommended to maximally separate groups, determine the most practical way to separate groups and remove variables which do not relate to a particular group (Hair, Babin, Money & Samouel, 2013). According to Uddin, Meah, and Hossain (2013), discriminant analysis forms one or more weighted linear combinations of discriminator variables called discriminant functions and have the general form:

$Z = \alpha + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n$, where Z is the discriminant score, α is the constant or Y-intercept of the regression line, β is the discriminant function coefficient or the weight of the variable, x is the discriminator variable or respondents raw score for the variable, and n is the number of discriminator variables.

4. Results and Discussion

A total of 30 projects were used in this study and the organization had already tagged them as successful, challenged or failed. The projects comprised 7 failed, 14 challenged and 9 successful projects. The data was analyzed using discriminant analysis and the result is clearly illustrated below.

Table 1: Prior Probabilities for Groups

ProjStu	Prior	Cases Used in Analysis	
		Unweighted	Weighted
-1.00	.333	7	7.000
0.00	.333	14	14.000
1.00	.333	9	9.000
Total	1.000	30	30.000

The system assigned probabilities to each group of project outcomes: Failure is assigned -1, challenged is assigned 0 while success is assigned 1 as shown in Table 1. Each of these groups was given equal probabilities of 0.333 and weights were assigned to the cases used in the analysis.

Table 2: Functions at Group Centroids

Project Status	Function	
	1	2
-1.00	.684	.681
0.00	.157	-.533
1.00	-.777	.299

The Function at the Group Centroids gives the average discriminant score of each of the three groups which established the threshold for classifying the cases and it is presented in table 2. The centroids for projects that failed is 0.684, challenged projects = 0.157 and successful projects = -0.777. Since the three groups (Failure, Challenged, and Success) are not equal (there are 7 failed projects, 14 challenged projects and 9 successful projects from the dataset), the weights on the centroids are used to find the dividing point.

Thus, the dividing rule will be
$$\frac{n_1(\text{lowerCentroid}) + n_2(\text{UpperCentroid})}{n_1 + n_2}$$

The group centroids between successful and challenged projects

$$= \frac{-0.777 * 9 + 0.157 * 14}{9 + 14} = \frac{-6.993 + 2.198}{23} = -0.2848$$

The group centroids between challenged and failed projects

$$= \frac{0.157 * 14 + 0.684 * 7}{14 + 7} = \frac{2.198 + 4.788}{21} = 0.3327$$



Fig 1: Group Centroids to formulate the decision rule

Figure 1 denotes that if the score of a new project is negative and below -0.2848, such project will be classified as successful; but if the score lies between -0.2848 and 0.3327, then the project will be classified as challenged, any other positive score outside this range will classify the project as a failure.

The decision rule is formulated as follows:

- A project is predicted and classified as successful if the discriminant score Z falls between -0.777 and -0.2848
i.e. $-0.777 < Z < -0.2848$
- A project is predicted and classified as challenged if the discriminant score Z falls between -0.2848 and 0.3327
i.e. $-0.2848 < Z < 0.3327$

- c) A project is predicted and classified as failure if the discriminant score Z falls between 0.3327 and 0.684
i.e. $0.3327 < Z < 0.684$

To verify the predictive capacity of the discriminant function, the values of the data collected is substituted in the discriminant function and the decision rule is used to classify the overall performance.

The 30 projects used in this study were also classified as ‘-1’ for those projects that experienced failure, ‘0’ for those that are challenged and ‘1’ for those that are successful, but the system prediction showed different result of the project overall performance. The system used the decision rules to reclassify the projects to different groups. According to the reclassification, 21 projects were correctly classified given 70% of correctly classified projects while 9 were wrongly classified given 30% as displayed in table 3.

Table 3: Summary of Reclassification Result

	Project Status	Predicted Group Membership		Total	%
Original count		Correct	Incorrect		
	Successful	6	3	9	30%
	Challenged	9	5	14	46.7%
	Failure	6	1	7	23.3%
Total		21	9	30	
(%) Percentage		70	30	100	

The predicted group membership gave the predicted frequencies showing the number that was classified correctly and those that were incorrectly classified and the last column of table 3 gives the percentage of observations in a given group. In other words, out of the 30 projects used in the analysis, 6 projects have been correctly classified as successful projects, 9 projects also were correctly classified as challenged projects and 6 projects were correctly classified as failed projects. This gives high result that using the decision rule developed; the project manager can come out with correct classifications of successful projects, challenged projects, and failed projects. This shows that the estimated model is good for predicting project overall performance.

Comparing prior classification and reclassification

The details of the comparison between prior classification and reclassification are given in table 4. The table 4 also indicated the probabilities of a project either being in Group -1 for those that experienced failure, Group 0 for challenged projects, and Group 1 for successful projects. It also shows that a project might be tagged as failure, that is, being in Group -1 but there are some little corrections to be made for it to be in Group 0 or Group 1 and that gives its probability of being in the other groups. For instance, additional resources or improved stakeholder management could have helped the project to move from failure or challenged group to success group. A project that was classified as success might have met some of the expected functionalities, and the unmet functionalities might linger to the next phase

196 attracting extra cost and time and that shows the probability of its membership in Group -1
 197 and Group 0. Most software engineering projects are tagged as successful if it is delivered
 198 within budget and on time while some of the functionalities are yet to be implemented.

199

200 **Table 4: Reclassification of Project Status**

Projects	Prior Classification of Project status	Reclassification of Project based on Analysis 1	Discriminant Scores from Function 1 for Analysis 1	Discriminant Scores from Function 2 for Analysis 1	Probabilities of Membership in Group -1 for Analysis 1	Probabilities of Membership in Group 0 for Analysis 1	Probabilities of Membership in Group 1 for Analysis 1
FUTO	0	1	-2.41618	-0.93377	0.01407	0.21331	0.77262
ECR	0	1	-1.72301	-0.55748	0.04009	0.26715	0.69275
DCCA	1	1	-1.06273	-0.38482	0.09107	0.34743	0.5615
LAG-PAY	1	1	-1.73201	-0.62615	0.0381	0.27722	0.68468
LAG-PMIS	1	1	-1.60131	0.54381	0.08236	0.13526	0.78238
UNIBEN	1	1	-1.45798	-0.22446	0.06579	0.25449	0.67971
UNIMAID	1	1	-1.82626	-1.17311	0.02423	0.35989	0.61588
UNIBUJA	0	0	-0.45174	-1.73077	0.05161	0.73049	0.21789
NIS-IP	0	-1	-0.00731	1.26475	0.50038	0.14786	0.35175
CLEAR	0	1	-0.19212	0.79398	0.3733	0.21528	0.41142
NHIS	0	0	0.62931	-1.18308	0.1717	0.70707	0.12123
IFEM	0	0	0.59688	-0.5124	0.29158	0.54132	0.1671
FTMS	1	-1	0.26509	1.26601	0.57886	0.14794	0.2732
DBSET	-1	-1	0.86488	0.47693	0.57115	0.27717	0.15169
SIAMS	1	-1	0.69823	0.8733	0.61775	0.20235	0.1799
NUCDB	0	-1	0.01732	1.36653	0.52354	0.13497	0.34149
AK-ELIB	-1	-1	0.64897	0.66973	0.56542	0.24334	0.19124
NPC-BIO	0	0	1.17059	-0.2849	0.44087	0.45893	0.1002
E-AGRIC	1	1	-0.73145	1.22978	0.28543	0.12892	0.58566
US-SCH	-1	-1	-0.13016	1.42637	0.48781	0.12641	0.38578
US-ELIB	-1	-1	0.86902	1.18537	0.71103	0.14575	0.14322
NIMC	-1	0	1.17544	-1.18767	0.22607	0.70193	0.072
JAMB-BF	1	-1	0.45976	1.18703	0.61734	0.15671	0.22595
KADPOLY	0	0	0.56295	-0.8274	0.22477	0.62258	0.15265
DISICS	-1	-1	-0.1154	2.0779	0.58134	0.0678	0.35086
BENSU	-1	-1	1.47736	0.11666	0.59933	0.32603	0.07465
IPSSS	0	0	0.11712	-0.8635	0.16718	0.61202	0.2208
IMT	0	0	1.09639	-1.2608	0.20387	0.72118	0.07495
FRSC-DL	0	0	0.92281	-0.71039	0.29657	0.58956	0.11387
FUOKE	0	0	1.87552	-2.01744	0.1422	0.83544	0.02236

201

202

203 5. Conclusion

204 This work has suggested a prediction model to determine project overall performance using
 205 discriminant analysis for software engineering projects in Nigeria. The discriminant function

thus developed was subjected to predict how many of these projects were successful, challenged, or failure. The prediction, based on this discriminant function, was compared with the actual information from the data collected. The original value was the same in some cases as that of score used in the data collected but some showed misclassification. The literatures revealed many cases of projects that fall short of stakeholders' expectations in terms of schedule, budget, performance, and customer satisfaction, yet the developed team tagged the project as successful. The study revealed that a project might be termed successful by its development team but other stakeholders might prefer the old system. This gives management the impression that the system could not perform as expected. This scenario is the case of Nigerian Permanent Voters Card (PVC) system that gave the project poor reputation and makes other stakeholders to conclude that the project is a failure. Further studies can be carried out to develop new prediction model using Naïve Bayesian Classification, artificial neural network model, or Decision Tree method.

References

- Atkinson, R. (1999) Project management: cost, time and quality, two best guesses and a phenomenon, it's time to accept other success criteria *International Journal of Project Management* 17(6), 337 - 342
- Beleiu, I., Crisan, E. & Nistor, R. (2015). Main factors influencing project success, *Interdisciplinary Management Research*, 11, 59-72.
- Bloch, M., Blumberg, S. & Laartz J. (2012). Delivering large-scale IT projects on time, on budget, and on value *Article - McKinsey Quarterly* <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights> accessed 25 June, 2018
- Bowen, P.A., Hall, K.A., Edwards, P.J., Pearl, R.G. & Cattell, K.S. (2002) Perception of time, cost, and quality management on building projects. *The Australian journal of construction economics and building* 2(2), 48 - 56
- Davis, K. (2014). Different stakeholder groups and their perceptions of project success, *International Journal of Project Management* 32, 189–201
- Ebbesen, J. B. & Hope, A. (2013) Re-imagining the Iron Triangle: Embedding sustainability into project constraints. *PM World Journal*, 2 (3).
- Eberendu A. C. (2015) *Evaluation of Software Project Failure and Abandonment in Tertiary Institutions in Nigeria*, *Journal of Information and Knowledge Management* 5(4), 29 – 40
- El Emam, K., & Koru, A. G. (2008). A replicated survey of IT software project failures. *IEEE software*, (5), 84-90.
- Florentine S. (2017) IT project success rates finally improving *CIO Newsletter*, February 27
- Hair, J., Babin, B., Money, A., & Samouel, P. (2013). *Essentials of business research methods*. Hoboken,

- 248 New Jersey: John Wiley & Sons, Inc.
- 249
- 250 Hussein, B. A. (September 2013) Factors Influencing Project Success Criteria. *The 7th IEEE*
- 251 *International Conference on Intelligent Data Acquisition and Advanced Computing Systems:*
- 252 *Technology and Applications, Berlin, Germany*
- 253 Jørgensen, M. (2014). Failure factors of small software projects at a global outsourcing
- 254 marketplace. *Journal of systems and software*, 92, 157-169.
- 255 Koelmans, R.G. (2004) Project success and performance evaluation. *International Platinum*
- 256 *Conference 'Platinum Adding Value'*, The South African Institute of Mining and
- 257 Metallurgy,
- 258 Mir, F.A.& Pinnington, A.H. (2014). Exploring the value of project management: Linking project
- 259 management performance and project success, *International Journal of Project Management*
- 260 32, 202–217
- 261 Montequin , V.R. Cousillas, S.M. Alvarez , V. & Villanueva, J. (2016) Success Factors and Failure
- 262 Causes in Projects: analysis of cluster patterns using self-organizing maps *Procedia Computer*
- 263 *Science 100* (2016), 440 – 448
- 264 Project Management Institute (PMI) (2013). A guide to the project management body of knowledge
- 265 (PMBOK Guide) PMI Standard Committee: Project Management Institute
- 266 **Projects**
- 267 Rezaian A. (2011) Time-Cost-Quality-Risk of Construction and Development Projects or Investment
- 268 *Middle-East Journal of Scientific Research 10*(2): 218-223,
- 269 Savolainen, P., Ahonen, J. J., & Richardson, I. (2012). Software development project success and
- 270 failure
- 271 from the supplier's perspective: A systematic literature review. *International Journal of*
- 272 *Project Management*, 30(4), 458-469.
- 273 Senouci, A., Ismail, A. A., & Eldin, N. (2016, June). Time and Cost Overrun in Public Construction
- 274 Projects in Qatar'. In *Creative Construction Conference, Budapest Hungary* (pp. 25-28)
- 275 Standish Group (2015). Chaos Report, Standish Group
- 276 Suliman, S. M. A., & Kadoda, G. (2017, November). Factors that influence software project cost and
- 277 schedule estimation. In *Computer Science and Information Technology (SCCSIT), 2017*
- 278 *Sudan Conference on* (pp. 1-9). IEEE.
- 279 **The Role of Requirements in the Success or Failure of Software**
- 280 Uddin, N., Meah, M. S., & Hossain, R. (2013). Discriminant analysis as an aid to human resource
- 281 selection and human resource turnover minimization decisions. *International Journal of*
- 282 *Business and Management*, 8(17), 153 - 169.