

A Predictive Analytics Approach to Project Management: Reducing Project Failures in Web & Software Development Projects

Tazeen Fatima

College of Electrical & Mechanical Engineering, National University of Sciences & Technology, Rawalpindi, Pakistan

Abstract - Use of Project Management in Web & Software Development Projects is very significant. It has been observed that even with the application of effective Project Management, projects usually do not complete their lifecycle and fail. To minimize these failures, Key Performance Indicators have been introduced in studies to Counter Project Failures. However, do we always identify the right KPIs? Despite of efforts, why are Projects still failing? Is there a way to identify and avoid these failures in the very beginning of the Project Lifecycle? We aimed to answer these questions by analyzing concept of Predictive Analytics which is a specialized technology and is very easy to use in this era of Computation. Project Organizations can use Data Gathering, Compute Power, and Modern Tools to render efficient Predictions. The research aims to identify such a Predictive Analytics Approach. The core objective of the study was to reduce failures and introduce effective implementation of Project Management principles. Existing Predictive Analytics Methodologies, Tools and Solution Providers were also analyzed. Relevant Data was gathered from projects and was analyzed via Predictive Techniques to make predictions well advance in time to render effective Project Management in Web & Software Development Industry.

Keywords: Project Management; Predictive Analytics; Predictive Analytics Methodology; Project Failures

1 Introduction

Project Failures is an issue that even big firms in the Web and Software Industry cannot 100% control. Project failure can be easily attributed to a number of factors. These can be lack of Proactive Risk Management; lack of effective Performance Measurement; inefficient Alignment of Organizational Strategy with the Project Strategy; absence of a well-defined Project Communication Methodology; ineffective Project Scope Definition; lack of Effective Leadership and more. On the contrary, businesses incessantly encounter increased investment demand despite of restricted capacity measures. Therefore, the need to render successful projects becomes an absolute necessity to avoid financial losses. In this environment, executing Projects Successfully and Reducing Project Failures are key Business Requirements

and essence of Successful Project Management. Therefore, this research was focused on assisting Project Managers to render successful projects in Web & Software Development Industry while aiming to reduce project failures. Numerous Projects fail at National and Corporate Levels. The idea behind this research is to minimize Project Failures. The study is initially aimed at Project Management for Web & Software Development Industry. However, with fruitful results, it will not only prove to improvise Project Management in Web & Software Industry but will also extend to other industries.

The research aims to: Reduce Project Failures in Web & Software Development Projects; Render Effective Project Management in Web & Software Development Projects; Ensure Efficient Delivery of Outcomes in Web & Software Development Projects and Introduce the implementation of Predictive Analytics Methodology & Technology to Reduce Project Failures in Web & Software Development Projects.

1.1 Previous work

A lot of studies and methodologies exist to render the idea of Reducing Project Failures and Rendering Successful Projects. In the fast growing world of today, where everything is being predicted in Advance through the Use of Intelligent Systems, the same concept can be used in Project Management for Web & Software Development to achieve substantial and profitable results. Predictive analytics is transforming into a remarkable advantage for acquiring a wide list of business goals including: increased customer satisfaction; financial profitability; effective business operations; higher market shares and more. Predictive analytics comprise of tools, methods and technologies for organizations to identify trends and patterns in a given data set, testing of huge quantities of variables, development of data models and data mining for innovative findings. Though, Predictive Analytics has been widely used in this aspect, the role of Predictive Analytics in Web & Software Development and Project Management is a new concept and is still in initial stages of development. Therefore, the Level of Research Already Carried out on the Proposed Topic is mostly conceptual. Deloitte have proposed a study on this concept back in 2010-2011. Though, there still remains a huge gap between the study and the pragmatic results of its implementation at an International Level. On the contrary,

detailed levels of Research have been carried out on Predictive Analytics in General as well as on the Reasons of Failure for Web & Software Development and Project Management. This has assisted in formalizing the fundamentals concepts to develop the study at hand.

1.2 Introduction to predictive analytics

Predictive project analytics work extremely effectively at the start of a project by helping you build the competencies you need to enhance success. However, the capability also adds value throughout the project lifecycle. By conducting this analysis during the course of a project, you can bring a floundering initiative back on track. Similarly, by assessing your project towards completion, you can avoid launching an initiative prematurely.¹ No matter when you leverage predictive project analytics; you position your organization to: Gain an objective overview of your project risks and exposure areas so you can close identified gaps; Enhance project outcomes by prioritizing actions most likely to help you achieve success; Realize early value from your projects by identifying specific measures you can take to keep projects on track; Minimize potential financial, productivity and reputation-based losses by improving project oversight; Drive greater organizational efficiency by eliminating unnecessary project characteristics and strengthening internal attributes that contribute to project success; Reduce risk by assessing your project relative to those with similar attributes and Benchmark your project against a database that includes engineering, capital works and technology projects and is effective regardless of project type.¹

There are seven reasons why we need Predictive Analytics today:

1. *Compete*: Secure the Most Powerful and Unique Competitive Stronghold
2. *Grow*: Increase Sales and Retain Customers Competitively
3. *Enforce*: Maintain Business Integrity by Managing Fraud
4. *Improve*: Advance Your Core Business Capacity Competitively
5. *Satisfy*: Meet Today's Escalating Consumer Expectations
6. *Learn*: Employ Today's Most Advanced Analytics
7. *Act*: Render Business Intelligence and Analytics Truly Actionable.²

1.3 Predict to reduce project failures in web and software development projects

There are number of approaches to address the research problem e.g. predicting why do projects fail? Predicting what drives project performance? Predicting risk via Risk Analytics and so on.³ Another approach that can be adopted is Big Data Predictive Analytics. Big Project Data (risks, causes of failures, wrong performance measures) can be used to leverage the approach to Predict to Reduce Project Failures in Web & Software Development Projects.

1.3.1 Use of IBM SPSS Modeler as predictive analytics tool

IBM SPSS Modeler is a powerful, versatile data mining workbench that helps you gain unprecedented insight from your data. Its breadth and depth of techniques allows you to build predictive models easily, efficiently and rapidly whether you are an expert data miner or a business analyst.⁴ Predictive analytics includes running numerous iterations of the most relevant data to attain optimal results. Big data can be complex and requires specific skills and scripts.⁵ The IBM SPSS Modeler solution is designed to support unstructured or semi-structured predictive analytics for big data. SPSS Modeler features a graphical interface that will put the power of data mining in the hands of project users. As a result, insights can be gained from our project data quickly and efficiently without complex programming packages or scarce expertise.⁵ With SPSS Modeler, you can use a broad range of data, including data stored in operational databases or flat files, as well as unstructured data such as call center notes, emails and survey responses. Its simple graphical interface quickly takes you through the entire data mining process from data access and preparation, through modeling and model evaluation, to model deployment. It has flexible deployment options from a simple desktop tool to advanced client server architecture. This feature allows organizations of all sizes to integrate predictive analytics into their everyday business environments and processes.⁴ Many FORTUNE 500 corporations, academic institutions and national and local government agencies worldwide rely on SPSS Modeler to unlock the value of their enterprise data, improve business processes and make more informed decisions⁴.

Furthermore, a predictive analytics approach will be proposed based on the various aspects that have been highlighted to Reduce Project Failures in Web and Software Development projects.

2 Methodology

2.1 Data identification

The first step in this study was to design and formalize a Project Profile Organization Matrix. The idea was to acquire thorough knowledge about the Projects in the study at hand. The matrix was formed based on the following Project Profile Fields: Project Name; Abstract; Technology; Business Value; No of Resources; Skillset of Resources; Budget; Duration; Project Results and Additional Comments

All of the above were textual fields. Each row represented a different Project while the columns represented the project profile fields above. The second step was to design and formalize the Project Primary Variables Matrix. The dataset comprised of information on the following primary variables which play a vital role in any projects success or failure: *Project Name*; *Project Result (Success/Fail)*; *Well Defined Scope*; *Effective Requirements Management*; *Effective Project*

Delivery Management; Effective Resource Management; Effective Risk Management; Effective Stakeholder Management; Effective Contract Management; Project Complexity (Conceptual Complexity; Stakeholder Complexity; Technological Complexity; Integration Complexity; System Development Complexity; Financial Complexity; Schedule Complexity; Organizational Complexity; Environmental Complexity); Project Schedule Management; Effective Change Management; Project Communication Management; Project Cost Management; Project Design Methods (Joint Application Design; Formal Design Methods; Structured Coding Methods); Project Quality Management (Formal Test Plans; Regression Testing; Performance Testing; User Acceptance Testing; Effective Project Controls; Effective Project & Process Measurement). All of the fields above had a numeric value of 1 to 10 except for the following fields: Project Name (Textual field) Project Result (Success/Fail Binary Value Field). Each row represented a different Project while the columns represented the primary variables pertaining to a single project. Each numeric variable was assigned a score option from 1 to 10 with 1 being the least value while 10 being the highest value. Input was gathered from users against each project on this textual and scoring scale to form a complete matrix. Both of these Matrices were designed using Microsoft Excel Technology.

2.2 Data gathering

Next, we identified the Software & Web Development Organizations which served as data sources. Collecting data was an essential part of this research. The idea was to identify the major sources and firms to give an access to at least 36 projects from the Web/Software Development and Information Systems sector which were willing to undergo this research. In order to accomplish this, organizations were contacted which were going to give an access to these projects residing within its domain. The idea was to convince the organization that the study will help them manage their projects better. Design of Effective Predictive Approach and Techniques was derived from the results of the study proposed as well as analyzing existing Predictive Models within the industry.

This data was gathered via Mixed Qualitative methods including Case Study Analysis; Focus Group Studies and Questionnaire technique.

A personal visit was conducted to Enekon and WeRPlay organizations to carry our Case Study Analysis and Focus Group Studies. Their Project Managers and SMEs were consulted in this aspect. Project Profile Organization Matrix and Project Primary Variables Matrix were discussed in length to analyze their practicality and authenticity in the design process of this research. Moreover, once the matrices were fully designed and approved, authentic data was gathered from these organizations to initiate the filling of the matrices with relevant information in Microsoft Excel.

In order to acquire results from the masses, a Questionnaire was developed based on Project Profile Organization Matrix and Project Primary Variables Matrix. The Questionnaire comprised of a Total of 25 Questions with 2 Multiple Choice; 2 Essay and 21 Scoring Questions based on a value from 1 to 10 with 1 being the least value while 10 being the highest value. Essay Questions dealt with acquiring Project Name; Business Value; Number of Resources; Skillset of Resources; Approximate Project Budget and Total Project Duration. Multiple Choice questions were Binary in Nature to acquire information such as if the person giving data is a Project Manager or Not and if the Project was a Success or a Failure.

The data set collected comprised of information on the primary variables which play a vital role in any Projects Success or Failure will be used as Predictor Variables in the Research. As the Research grew, List of Predictor Variables determining Project Success or Failure were narrowed down to improve results and analysis. Consequently, Key Predictor Variables were highlighted towards the end of the Research as part of a Predictive Approach to Predict a Projects Success or Failure in advance based on Data Analysis. This helped reduce Project Failures.

2.3 Identification of predictive approach

The objective of the study was to find answer to the question: How to Predict and Define a Predictive Approach? Predictive Analytics using *BigData Analytics* Tools & Technologies was employed. The idea here was to use existing technologies and tools to adopt a Predictive Analytics Approach to Reduce Project Failures in Web & Software Development projects. In this study, we used Predictive Modeling as part of IBM SPSS Modeler to Reduce Project Failures in Web & Software Development Projects. IBM SPSS Modeler utilizes Data mining techniques to predict the future. Project(s) Data was gathered from various data sources via Mixed Qualitative Methods. The sample size was a total of 36 Projects from Web & Software Development industry. Project(s) data was a set of primary variables which play a vital role in any Projects Success or Failure and was in the form of Project Primary Variables Matrix. This Matrix was at first Cleaned and Analyzed. Afterwards, the Matrix was then inserted as an input to IBM SPSS Modeler to identify relationship between different variables. The tool was used to run Predictive Modeling processes which applied a combination of the following Predictive Algorithms on the dataset *Logistic Regression*⁶; *Discriminant Analysis*⁷ and *CHAID*⁶. *K-Means Clustering*⁶ was also used as an additional Predictive modeling process to render prediction via Clustering for improved results. Target Variable defined was Project Result with Binary values Success or Fail while others were regarded as Predictor Variables in the data set to determine their percentage of contribution in Success Or Failure of the project. Moreover, different processes in IBM SPSS Modeler were incurred to identify Variable Types and Perform Data Auditing to have an insight into Variables Field,

Measurement, Values, Check, Role, Format, Category (Categorical/Continuous), Minimum Value, Maximum Value, Mean, Standard Deviation, Skewness, Median, Mode, Uniqueness and Validity. The idea was to ensure the quality of the dataset.

3 Results

3.1 Research findings

The main idea behind the study was to analyze which factors play the most significant role in Predicting Project Failures at an early stage of the project lifecycle. During the course of the project, there are many aspects that are overlapped and it is very critical for a Project Manager to concentrate on a few key items that can make a huge difference. The results gathered from the 31 successful and 5 failed projects in the study at hand can help determine some of these factors. A total of 16 Project Managers and 10 Project Management Office Members and Business Analysts took part in the study. Project Profile Organization Matrix helped gather and maintain Project basic level information for various projects.

3.1.1 Key projects

Following are some of the projects from which the data was gathered: Commission Management System; Enterprise Data warehouse and Business Intelligence Project; Use of Information Technology in schools as teaching aid and HDMS- Decision Modeling System Development.

3.1.2 Scope of web and software development technologies

Different Forms of Web and Software Development Technologies that encompass the projects within the study at hand include: iOS/Android Mobile Game Development; Projection Mapping; Java based Web Application; ETL & BO Reporting; Teradata Data warehouse Platform; Micro Strategy; BI Tool for Front End Reporting; Informatica for ETL; ASP and more.

3.1.3 Identifying Predictor Importance

Based on the data analysis and prediction using Logistic Regression, Discriminant Analysis and CHAID algorithms in Auto-classifier Predictive Model; primary variables *Regression Testing* and *Effective Contract Management* have the Most Predictor Importance with a rating of 0.04. Thus, they can be considered to be important Predictor Variables as shown in Table 1.

Tab.1 Important Predictor Variables

Field	Variable Information				
	Min	Max	Mean	Median	Mode
Contract management	0	10	5.917	7	8
Regression Testing	0	10	6.75	8	10

Tab.2 Predictor Importance Ratings of Some Other Primary Variables Analyzed in the Study

Primary Variable	Ratings
	Predictor Importance
Performance Testing	0.03
Schedule Complexity	0.03
Risk Management	0.03
User Acceptance Testing	0.03
Formal Design Methods	0.03
Resource Management	0.03
Joint Application Design	0.02

Prediction using K-Means Clustering rendered the following results:

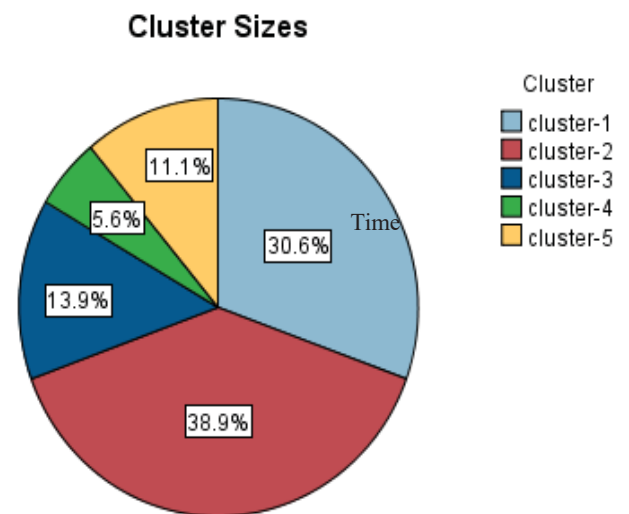


Fig.1 Cluster Sizes

Size of Smallest Cluster	2 (5.6%)
Size of Largest Cluster	14 (38.9%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	7.00

Fig.2 Smallest & Largest Cluster Comparison

The following key observations were made while analyzing data for all the five clusters as shown in Table 3. Predictor Importance value was common to all the five clusters. E.g. Cluster 1 with Mean 8.09 has predictor importance of 1.00 for User Acceptance Testing, Cluster 2 with Mean 8.50 also has predictor importance of 1.00 for User Acceptance Testing and so on. It also highlights the highest Predictor Importance values (in descending order with value > 0.50 being considered). Therefore, they are all significant Predictor Variables.

Tab.3 Cluster Wise Predictor Importance

Primary Variable	Importance					
	Predictor Importance	Cluster 1 Mean	Cluster 2 Mean	Cluster 3 Mean	Cluster 4 Mean	Cluster 5 Mean
User Acceptance Testing	1.00	8.09	8.50	9.00	2.50	0.25
Project Delivery Management	0.66	7.45	8.50	8.60	5.50	3.25
Formal Test Plan	0.65	5.45	8.64	8.00	2.50	1.75
Project & Process Measurement	0.62	6.64	8.00	6.60	3.00	3.00
Contract Management	0.59	4.45	8.57	5.80	7.00	0.25

The simulation modeling in IBM SPSS identified key Predictor Variables which can help in determining Reasons of Project Failures at an early stage of Project Lifecycle. Use of the combination of Logistic Regression, Discriminant Analysis and CHAID algorithms in Auto-classifier Predictive Model of IBM SPSS concludes that Primary Variables Regression Testing and Effective Contract Management have the Most Predictor Importance with a rating of 0.04. The Auto-

Classifier Predictive Model calculated the same Predictor Importance value of 0.03 for all other Primary Variables including Effective Project & Process Measurement; Effective Project Controls; Performance Testing; Formal Test Plans; Structured Coding Methods; Project Communication Management; Project; Cost Management; Project Schedule Management; Environmental Complexity; Schedule Complexity; Financial Complexity; Technological Complexity; Stakeholder Complexity; Conceptual Complexity; Effective Stakeholder Management; Effective Risk Management; Effective Resource Management; Effective Project Delivery Management; Effective Requirements Management; Well Defined Scope; System Development Complexity; User Acceptance Testing; Organizational Complexity; Integration Complexity; Formal Design Methods; and Effective Change Management. Some of these are shown in Table 2. Lastly, Joint Application Design had a predictor importance of 0.02. Two possibilities could be deduced from these results: 1. All Project Primary Variables ranging from Effective Project & Process Measurement to Effective Change Management should be given equal weightage by Project Managers to avoid Project Failures and Joint Application Design can be given the least importance. 2. The Predictor Importance values were really close together and did not present a significant variation to make pragmatic deductions from the given dataset. Therefore, more computations were needed. The latter was chosen to be an authentic inference. Therefore, it was decided to add K-Means Clustering to the Predictive Model to refine the results. This computation turned out to be useful. The algorithm distributed the data inputs into five distinct Clusters. Cluster Quality was Fair with Silhouette Measurements of Cohesion and Separation to be approximately 0.3 on a scale range of (-1.0 to 1.0). The five clusters had variable sizes. Predictor Importance values calculated for Primary Variables by K-Means Clustering as part of the Predictor Model was same for all the five clusters while the Mean Values were different for each cluster.

Two observations using K-Means Clustering reinforce the results obtained by combination of Logistic Regression, Discriminant Analysis and CHAID algorithms in Auto-classifier Predictive Model of IBM SPSS as shown below:

Tab.4 Cluster Wise Predictor Importance

Primary Variable	Importance	
	Predictor Importance via Auto classifier	Predictor Importance via K-Means Clustering
Regression Testing	0.04	0.53
Contract Management	0.04	0.59

Therefore, as mentioned before Regression Testing and Effective Contract Management can be used as means of predicting and mitigating project failures at an early stage of Web and Software Development Project Lifecycle. Other observations signify that *User Acceptance Testing (UAT)* is the most important Primary Variable as per K-Means Clustering with Predictor Importance of 1.00.

4 Conclusions

According to Predictive Model designed and proposed, if a project does not have specialized level of User Acceptance Testing, Project Delivery Management, Formal Test Plans, Project and Process Measurement, Effective Contract Management, Performance Testing, Cost Management and Regression Testing, the project is bound to fail. If Project Managers want to assess the course of a Web or Software Development Project Lifecycle to take actions in advance to avoid Project Failures, the above *Predictor Variables* need to be assessed and improved. Moreover, Successful Identification of Project Predictor Variables and Reasons of Project Failures for Web and Software Development projects in study with the help of IBM SPSS Modeler and Predictive Models and Techniques have also reinforced the notion that Predictive Analytics Methodology and Technology can apply effectively in Web & Software Development Projects to Reduce Project Failures. Predictive Analytics is a new area in Project Management and there is still a long way to go. Existing Predictive Analytics in Project Management is very limited and does not apply exclusively on Web and Software Development projects. The combination of these findings has helped to introduce a Predictive Analytics Approach in Project Management to Reduce Project Failures in Web & Software Development Projects. Similarly, other Tools and Technologies can also be analyzed to find out more ways of devising a Predictive Approach for Web Development and Software Projects Industry. Though, project management applies to all walks of business, the research aimed to focus only on project management in relation to Web and Software Development industry. Numerous projects fail in Web and Software Development and other industries. Deloitte in their paper establishes that 63% of IT projects have either failed (21%) or have been challenged (42%) according to Chaos Report, Standish Group International, 2010³. Moreover, according to the same report, projects between \$1M and \$3M have a chance of success estimated at 34%, and projects over \$10M only a 7% chance of success. Making Change Work, Survey of 1,500 change management executives, IBM, October 2008 suggests that 44% of all projects failed to meet time, budget or quality goals, while 15% were either stopped or failed to meet all objectives³. In the midst of these prevailing facts, it becomes an extreme need of the hour to propose an approach to reduce project failures at a very early stage of a web and software development project lifecycle to ensure that the maximum number of projects meet their

objectives and client expectations. The study at hand is targeted right towards achieving this approach. The concept of Predictive Analytics is thus introduced to fill in the gaps and to offer something extra in Project Management for Web and Software Development projects to reduce project failures. The concept is relatively new and requires further research but foundation was laid to insinuate an added dimension in traditional project management. As part of Predictive Analytics, the research introduced IBM SPSS Modeler as the key predictive tool to predict reasons of project failures at an early stage of the project lifecycle. In addition, the importance of Predictive Algorithms such as Logistic Regression, Discriminant Analysis, CHAID and K-Means Clustering was also highlighted.

5 References

- [1] Scott Whelehan, Predictive Project Analytics Deloitte, Capturing a Moving Target: Change Management, Consultants News. Retrieved from: <http://www.changemanagementresources.com/2011/10/04/the-beginningof-change-management/>, A Comparison of Project Complexity between Defence and Other Sectors, The Helmsman Institute, Sidney, Australia. www.helmsman-institute.com, pp. 2-8, 2011
- [2] Eric Siegel, IBM, Seven Reasons You Need Predictive Analytics today, Prediction Impact White Paper, pp. 6
- [3] Jack Su, Enabling Risk Analytics for Project Success, Deloitte, IIA Luncheon, pp.1, March 2013.
- [4] Solve your toughest challenges with data mining, www01.ibm.com/software/analytics/spss/products/modeler/downloads.html, IBM Software, Business Analytics, IBM SPSS Modeler:
- [5] IBM SPSS Modeler and SPSS Analytic Server: A powerful combination, IBM Software, Business Analytics, Predictive Analytics for Big Data.
- [6] <http://www.saedsayad.com/modeling.htm>
- [7] Dr. Syed Athar Masood, National University of Sciences & Technology (NUST), Analysis and Development of a Model for the Causes of Cost and Time Overruns in Engineering Projects, , pp.1
- [8] C. A. Silva and Paulo Ferrão, MIT-Portugal Program - Instituto Superior Técnico, Lisbon, Portugal, A Systems Modeling Approach to Project Management: The Green Islands Project example.
- [9] Ž. Antolić, R&D Center, Ericsson Nikola Tesla, An Example of Using Key Performance Indicators for Software Development Process Efficiency Evaluation, d.d pp.1
- [10] An Oracle White Paper, Why Projects Fail: Avoiding the Classic Pitfalls, pp.2, 2011.
- [11] Mike Gualtieri and Rowan Curran, The Forrester Wave™: Big Data Predictive Analytics Solutions, pp. 9 – 15, Q2 2015, Updated: April 1, 2015,
- [12]https://en.wikipedia.org/wiki/Discriminant_function_analysis
- [13]<https://msdn.microsoft.com/en-us/library/aa292167%28v=vs.71%29.aspx>

- [14] <http://smallbusiness.chron.com/eight-steps-effective-contract-management-24948.html>
- [15] <https://www.techopedia.com/definition/3887/user-acceptance-testing-uat>
- [16] <http://it.toolbox.com/blogs/delivery-doctor/what-is-delivery-management-10341>
- [17] <http://softwaretestingfundamentals.com/test-plan/>
- [18] Dr. Christof Ebert, Vector Consulting Services , Software Project and Process Measurement, pp.1
- [19] <https://msdn.microsoft.com/en-us/library/bb924357.aspx>
- [20] <http://www.projectinsight.net/>
- [21] Charles Elkan, Predictive Analytics and Data Mining, pp. 15 – 31, May 8, 2013