Retaining Quality of Service using Predictive Data Analytics

Group 24

Kaavya Rekanar
940521-7184
kare15@student.bth.se

Sushmitha Donthula
940303-8483
sudo15@student.bth.se

Group Members Participation

The work distribution among team members,

Group Member	Literature Search and Review	Research Design for Documentation	Report Writing and Preview
Kaavya Rekanar	45%	55%	50%
Sushmitha Donthula	55%	45%	50%

Table 1: Group Member's Participation

Abstract
The act of administering all activities and tasks needed to maintain a desired level of excellence for a product or project or initiative is called Quality Management. This involves creating and implementing the planning and assurance, quality control and improvement. The most important thing to understand about maintaining quality performance- or maintaining an organization is that you can never stop working at it. It takes constant care and hard work to maintain quality. This paper deals with the quality performance in an organization and a research proposal for carrying out this study to maintain quality using Predictive data analytics, a concept related to big data and one of it's many uses.

Author Keywords
Quality, Performance, maintaining, organization, constant work, Predictive Data Analytics
ACM Classification Keywords
Quality (Service, Management); Lifecycle (Deming, Test Driven Development); Data (Predictive Analytics, Big data)
General Terms
Management, Maintenance

I. Introduction

Analysis of data is a process of auditing, cleaning, transforming, and modelling data with the goal of discovering useful information, suggesting conclusions, and supporting decision- making[11]. One of the many ways where data analysis can be utilized is predicting the future outcome with the help of available information. The future of a company, i.e., the sustainability, stability and productivity can be calculated with the help of predictive analytics.

Besides various fields where predictive data analytics can be used, one of them is maintaining the Quality of Service provided for any product by predicting the complications and trying to overcome them by enhancing the outcome which helps eliminate mistakes, remove bugs(if any), and consequently reduces costs and time delays [2].

Background and Motivation

Predictive analysis can be used in almost any domain of work in the present market scenario. As discussed earlier, this paper deals mainly with the usage of Predictive analytics pertaining to the maintenance of Quality of Service in the end product given to the client.

New products evolve only when the old products (services) are analysed thoroughly and effective methods are developed to improvise it. Analysis of small amount of data can be done easily and it does not really require any special algorithms or mechanisms; but when data in huge amounts is involved, we need appropriate methods to modulate and manage the data. Hence effective algorithms are developed in order to customize and inculcate the commonsensical algorithmic behaviour into the machines so as to gain high productivity and aims to reach the goals of the customer.

This process of extracting requirements from the clients and analysing them prior to the occurrence of an event came into light because of the fast paced trending lifestyle in our contemporary world. Multinational companies and the evolved developed sectors these days, use such techniques in order to offer better services to their clients.

The dimensions involved in maintaining the Quality of Service are Reliability, Assurance and Responsiveness [5]. The perceived risk factors involved in this issue are

financial risk, performance risk, physical risk, social risk, psychological risk and time risk [5].

Objectives and Scope

Predictive analytics works efficiently only when used in a well- defined direction. It is used where data is involved in huge amounts and the seller or initiator or beneficiary needs to know the future of their own assets.
Customer Happiness Index (CHI) has been initiated by the company Hotspot and published in many journals. It states that, for a company even if a customer buys a product and consumes it, that doesn't mean they are satisfied and happy. They may consume the product due to several reasons; some of them varying from cheapest price to a good amount of marketing and advertising but the customer could be satisfied only when they get what they expect, which may or may not happen often. Hence, the customer happiness needs to be a key point in maintaining quality of any service provided.

II. Literature Review

A. Search Strategy

Initially, we searched for journals which were related to Predictive Data Analytics as we thought it would give a broad scope for Research and Development for Master Thesis. It also meant that both the authors of this report would have an opportunity to work in different fields; hence, we found it practically feasible at the time of search.
But as the search for proper data continued, we found that Predictive Data Analytics as a whole cannot be explained in the form of a report. It was a question of sanity and reliability on the topic to complete this work. The huge scope diverted the main cause of the report to something more and it became a hassle to focus on one particular topic for research.

B. Refining search strategy

The topic of Predictive Data Analytics has be narrowed down to one subject among all the derivatives.
Maintaining Quality of Service provided to many clients has become a challenge these days. Thus, computer experts are working on algorithms to predict such problems and overcome them before they demand huge attention.
This led us to narrow down our scope and refine our search. The scientific databases, Engineering Village, IEEE Explorer, Emerald Fulltext and IOS Press were searched with relevant keywords and the articles retrieved were considered for conducting the Systematic Literature Review.

The initial key words were: Quality of Service and Predictive data Analytics
A search in the database initially showed 71965 results.

S.No	Database	Results
1.	Inspec (Engineering Village)	128
2.	IEEE Explorer	25621
3.	Emerald Fulltext	399
4.	IOS Press	517
5.	Google Scholar	45300

Table 2: Initial Search Results

The search has been repeated. The keywords used were: (Predictive analytics*) AND((Big data) AND (Quality of Service*)). Refining and Re-refining of the search resulted in 13 relevant papers after applying exclusion criteria. The results are shown in Table 3.

S.No.	Database	Results
1.	Inspec(Engineering Village)	2
2.	IEEE Explorer	3
3.	Emerald Fulltext	1
4.	IOS Press	3
5.	Google Scholar	4

Table 3: Results after applying Exclusion Criteria

C. Data Extraction Strategy

We initially created a data extraction form, which consisted of the research questions that need to be addressed. This form was helpful to prepare the literature review taking into consideration all the relevant information. The keywords were selected and reviewing the abstracts of the available papers helped for a detailed search. Those papers, which were found most relevant, were considered for further study.
Papers, which were dated older than 2000 had to be excluded. This is due to the fact that algorithms keep changing day by day and updating accordingly is required for the report.

D. Inclusion/Exclusion Criteria

The inclusion and exclusion criteria, depending on the search results are:
Inclusion Criteria:

- Papers that focused on a particular service also were taken into consideration since it explained about certain algorithms being used in the industry in a very detailed manner.
- Papers with relevant vocabulary (to the research questions) were selected.

- As an exception, papers related to predictive data Analytics and Data Mining have also been considered, to get a better insight.

Exclusion Criteria:

- Unrelated papers and articles that were almost similar to each other were excluded.
- Articles where only abstract was available (some papers required purchasing to get the full text) were eliminated.
- Papers that were not published in English were excluded.
- Papers older than 2000 were excluded.

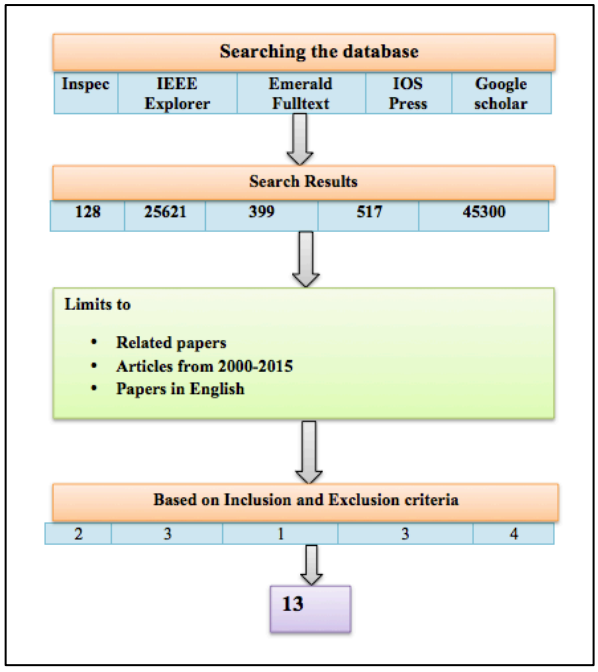

Figure 1: Inclusion and Exclusion criteria

E. Quality of Service in Real-life

a. In the field of Medicine

According to Kent Bottles, When Stanford researcher Michael Snyder sequenced his own 'whole' genome, he found he was at risk for developing type 2-diabetes. Genomic study is not a simple analytical study, since it involves statistical as well as biological aspects. By periodically having his blood tested for about 40,000 proteins, Snyder developed his own integrative personal genomics [11] profile consisting of 30 terabytes of data about how his body functions. The diabetes prediction was surprising because Snyder is slender, has no family history of diabetes, and has never had elevated blood glucose readings. This is why "predictive" analytics is said to goals and targets, which were once unreachable to the society.

When he consulted with Stanford endocrinologist Sun Kim, Snyder was told, "There is no way you have diabetes."

A three-hour fasting blood sugar test revealed an elevated initial blood glucose level of 127 (normal 70 to 99 mg per decilitre)[11], and later haemoglobin A1C tests were elevated at 6.4 percent and 6.7 percent (normal between 4 and 6 percent) establishing a diagnosis of diabetes[11]. Snyder has changed his diet and lost 15 pounds; his blood glucose levels have returned to normal. However, for life insurance purposes, he has diabetes and his rates have gone up(It has to be understood that a life insurance company also grants insurance for the future and not the present, hence it requires more amount of analysis even before it grants insurance for a person).

Snyder states, "Every medical professional I encountered said there was no way I could have diabetes. But soon the volume of available data is going to overwhelm the ability of physicians to be gatekeepers of information. This will absolutely change how we do medicine." Cardiologist Eric Topol [11] has called Snyder's integrative personal genomics profile study "a landmark for personalized medicine." The Snyder diabetes example exemplifies the transition we are experiencing from a traditional medical paradigm of "diagnose and treat"[12] to a digital personalized medicine paradigm of "predict and prevent" [12].

This is a live example in order to understand the standards of predictive analytics. It can change the medical practices vigorously and can improvise the present situations completely giving a new dimension to the discipline. The research programs using analytics can eventually find cure for many insanely harmful and contagious diseases. It can find treatments, which are affordable to the common man. Pharmaceutics believe in 'analyze and prevent' rather than 'consult and cure'.

b. Maintaining QoS in Retail Industry

In retail system it often happens that, no matter how much efficient a product is, but the quality of service is always an important factor. The Customer happiness Index(CHI) and Quality of Service(QoS) are often mistaken for the same.

The article written by Giuseppe Cicotti, Luigi Coppolino [1], Salvatore D'Antonio, and Luigi

Romano gives an architectural view in predicting QoS.
This prediction framework takes advantage of the qualitative and quantitative analysis performed by a probabilistic model-checking technique. These techniques make use of highly effective mathematical operations that can account to result into a QoS meter.
Thus, this type of analysis predicts and notifies about the QoS of the system/service in advance and also takes care of the pros and cons since it uses QoS related predictive indicators.
Every domain needs to check their QoS. In order to understand the service, quality, reaching to the customers levels, everything needs to be analyzed. So does the architecture stream. Hence this is the strategy being used in predicting and monitoring architectures. Usually the architectural domain needs lot more analysis because it is not like software and cannot be modified once built.

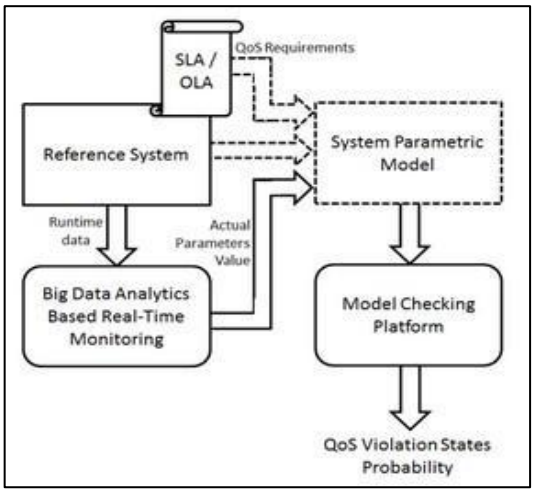

Figure 2: The QoS Monitoring and Prediction architecture [1].

Thus the QoS needs to be analyzed very efficiently using the defined models and the errors need to be modelled accordingly either on the paper or in the software. Every step needs compatibility to the past and future-amending steps, which cannot be done to the real material; hence, it is tested in a virtual environment.

c. Advocacy

A good quality program responds effectively to the needs it was designed to meet. Ethically, a firm is bound to provide the absolutely best quality of service or advocacy a firm can. If the staff members and volunteers know that they and the organization, i.e., the law firm are doing the best job possible, it builds the morale and makes them get a feeling of accomplishment towards themselves and the organization.

d. Community Development

A quality program is totally consistent with the mission and philosophy of the organization or initiative. Quality makes a group more effective at meeting the ends it is concerned with. Striving for quality helps to develop organizational and individual competence at a community level, thus continually improving the organization. Quality adds strength and credibility to the organization or an initiative.

e. Human Service Organizations
A quality program is a model of ethical behaviour. Quality is more economical in the long run and a quality program continually increases its performance level and improves its service delivery, which gives the organization's credibility and ultimately benefits the target audience.

F. Theories about Quality of Service

a. The Deming Cycle

The Deming Cycle is a process of creating and selling a product with an acceptable quality. The steps involved in the cycle are:

i. **Plan** - conduct a consumer research on their requirements and use that for planning the development of the end product.

ii. **Do** - produce the product.

iii. **Check** - check the product to make sure it has been produced in accordance with the plan made.

iv. **Act** - market the product .

v. **Analyze** - analyze the sales of the product and the customer's experience in terms of quality, cost and other data.

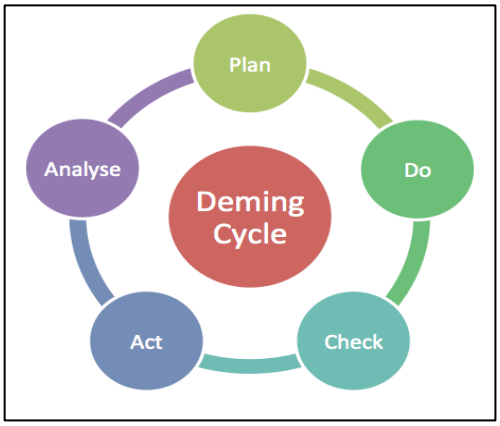

Figure 3: The Deming Lifecycle

b. Relevance of Quality Management with the Organization:

The principles of QoS have been precisely designed for the business sector, so while some aspects are relevant to the organizations concerned with advocacy, community development, health and human services; others may be contrary to the goals of those organizations.

Some elements of Quality Management that could work towards quality in any environment:

- ✓ Careful planning, monitoring, and evaluation over the product's development.
- ✓ Teamwork and empowerment of the organization
- ✓ Constant education and training for all the staff of the organization.
- ✓ Identifying and changing the loopholes wherever development is necessary.
- ✓ Attention to the needs of the target population.
- ✓ Encouraging and rewarding new ideas whenever possible.
- ✓ Developing an organization-wide culture for quality
- ✓ Maintain the same quality for a long term.

Some elements which may not work toward Quality Management are:

- ✓ Emphasis on products and production, which may take the focus off human needs and the consequences that are concerned with it.
- ✓ Assumption of hierarchical structure where leadership comes at the top most and achievement of quality on the product is not a real priority when compared to the advancement.
- ✓ The definition of everyone will actually differ from manager to a worker, where everyone merely can mean every participant taking part in development of the end product.

c. Maintain Quality Performance

While maintenance of quality is inbuilt into the Deming Cycle to some extent, it requires some commitment and action.

i. Dynamism:

An organization always needs to be dynamic; seeking new targets and an attempt to improvise at every level. Everyone in the organization have to be encouraged to take advantage of the opportunities available to learn and develop accordingly.

This includes grasping new ideas from every possible area and carefully evaluating them before implementation. Always aware of the goal, the organization should be able to take the risk of change if necessary. It should not be consistent with the principles and philosophies if they need a change from their originality for a better outcome.

ii. Long Range Strategic Planning:

Quality can only be maintained only when the organization plans long term and plans accordingly using strategic methods to secure it's future.

The role and future of the company should be questioned frequently in order to remain on par with the ever-growing market. Re-examining the current strategies of the organization for meeting the goals set up earlier can help them assess the progress till date.

iii. SWOT Analysis:

This is a structured planning method to evaluate the *Strengths, Weaknesses, Opportunities* and *Threats* involved in the project venture.

While Strengths and weaknesses are considered as internal factors; opportunities and threats are external factors.

Strengths and weaknesses define the relationship with the competitors of the organization. It is not necessarily the strength of teamwork in the organization, but it describes the productivity of work produced as a team.

Opportunities are to be seized as they appear in a world full of competition. It is all the more similar in the field of business, where opportunities have to be well utilized in order to sustain in the market. Threats are the challenges that go along with any opportunity. Threats may sometimes come unaccompanied by opportunities; they have to be identified and terminated as soon as possible.

G. Quality Assessment Criteria

The assessment of the paper's quality was done based on two criterions.
QC 1: How far is the paper giving accurate information about Quality of Service using Predictive Data Analytics?
QC 2: Is the paper discussing any new approach to predict data or is it giving an overall view of the topic as a summary?

The selected papers have been graded on four levels- Excellent, Good, Average, Low- based on their feasibility for the report.

Papers	QC 1	QC 2
Ref [1]	Excellent	Good
Ref [2]	Good	Good
Ref [3]	Good	Low
Ref [4]	Average	Good
Ref [5]	Low	Good
Ref [6]	Average	Low
Ref [7]	Low	Average
Ref [8]	Low	Average
Ref [9]	Average	Good
Ref [10]	Low	Good
Ref [11]	Good	Low
Ref [12]	Low	Average
Ref [13]	Low	Good

Table 4: Quality Assessment of the Literature

The reasons for choosing these criteria are:

- We wanted papers that were relevant or at the least close to our concept of research. Due to wide scope of Predictive Analytics in daily life, it is easy to deviate.
- Most of the papers gave an overall view on Data Analytics and its uses, but there was not much information relating to maintaining Quality of Service in Predictive Data Analytics.

H. Validation of the Protocol

We peer reviewed each other's opinion on every search result obtained. The other companion was asked to retest the search results using the same search string to be sure about the results and vice versa.

I. Data Extraction Process

The diagram describes the Data Extraction process in the form of a flow chart.

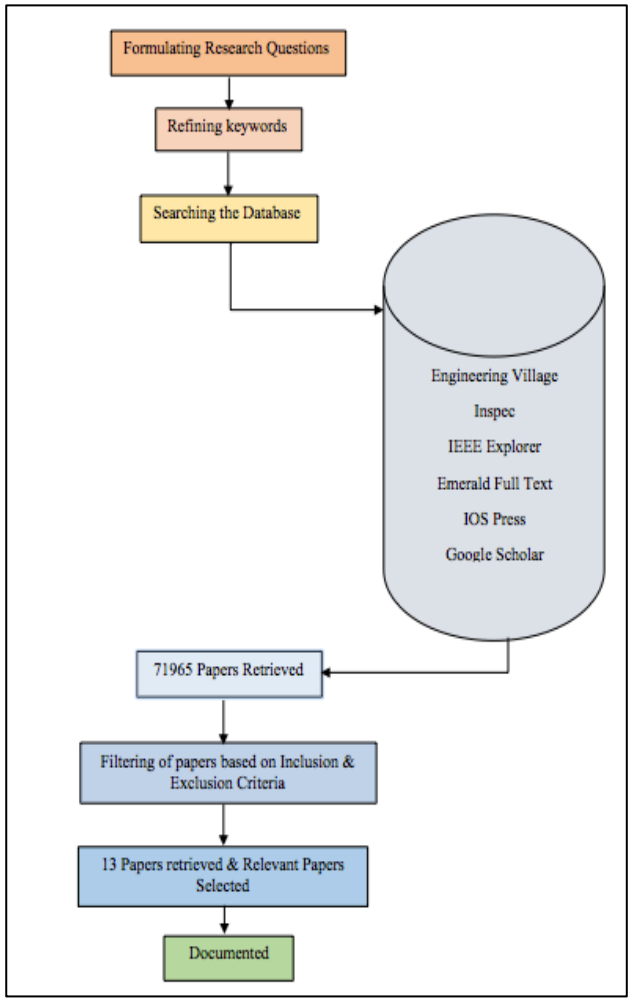

Figure 4 : Data Extraction Process

III. Research Questions

RQ 1: How QoS affects productivity of a product? How is QoS related to the organisation's growth?

RQ 2: Is QoS directly related to Predictive Data Analytics when it comes to the Total Quality Management in an organisation?

RQ 3: How far is Predictive Data Analytics helping in maintaining the Quality of Service when compared to traditional methods? Is it successful in increasing the productivity and growth of the product in a regular market?

IV. Research Method

The research questions can be solved using many research methods. The methods which are available are Pilot Study, Case Study, Post Mortem, Experiment Method, and Survey

- Pilot Study is a small scale study conducted in order to check the feasibility of a project, to determine its'

costs, adverse events and statistical variability as an attempt to predict challenges and improve them prior to full scale performance of a product.
This method can be used for our research to a certain extent but it is not completely viable as we cannot practically set up an experiment due to the limited scope.

- Case Study is a process into the development of a particular product, which is also not practically possible as no company would be willing to give their financial secret studies to some student for their research work.
- Experiment is not selected because we did not propound any new method in this study.
- Post mortem cannot do justice to this study as we are dealing only with one are of research and it may affect the results of research questions.
- Survey is selected because the research question proposed clearly demands diverse opinions of people.

Sample Selection

Sampling is a technique where you choose a group of people you would like to conduct a survey on. As we have selected a topic that is relatively close to customer satisfaction, we thought it would be a good idea to consult managers in the maintenance area of a few selected companies.
We selected people from different fields in order to know problems in their respective areas. We also selected some students who are good at using technology, as they would know the loopholes in maintenance and can give a better overview about the quality of service.

V. Data Collection Method

The several methods of data collection available are: Registration, Questionnaires, Interviews, Direct observations, Reporting and many more.
We have chosen the survey method for our research, which basically includes answering questionnaires online developed using Google Docs and taking interviews of people relatively working closer to the research area.
This paper basically describes the maintenance of Quality of Service using

- ✓ Questionnaire- A questionnaire has been prepared by the authors and handed over to a group of people who answered the questions according to the best of their knowledge.
- ✓ Interviews- Interviews of certain assistant managers of three different companies have been taken to get a clear view of the task at a managerial level. Some customers have also been at interviewed at a local customer service center to know what they expect from the system and what they receive.

The reason for choosing survey as a data collection method is the limited amount of time and a need to as much data as possible for the research. Another major reason is that this particular research area of predicting the requirements of a customer even before they know about it and maintaining the quality in the service provided is a realm which requires a lot of knowledge about the ever changing market and the psychological thinking of the customer as well. Survey method helps us interact with people and find the requirements first-handed directly from the source unlike any other data collection mechanism.
Our survey was made using the following procedure:

i. Select the sampling group and prepare a list of all the respondents.
ii. Develop the questionnaire
iii. Recheck the questions.
iv. Broadcasted to the group using web services- email and social networks
v. The data is deduced and all the collected responses are interpreted.

The questions used in the survey are:

1. Why is quality important for a grass roots organization?
2. Does the program (or initiative or product) actually address the identified needs?
3. Does the program or initiative reach, or help participants reach the desired outcomes? Were these outcomes the right ones to aim for, or do they need to be changed?
4. Is the program or initiative consistent with the vision, mission, philosophies, and principles of the organization?
5. Is the program or initiative inclusive and respectful of the targeted community?

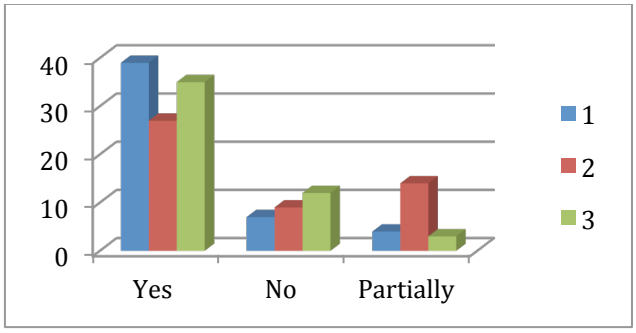

Figure 5: Graph showing the results of survey conducted

This graph is a depiction of the answers received in the survey; where 1, 2 and 3 in the graph represent answers to questions 2,4 and 5 respectively.

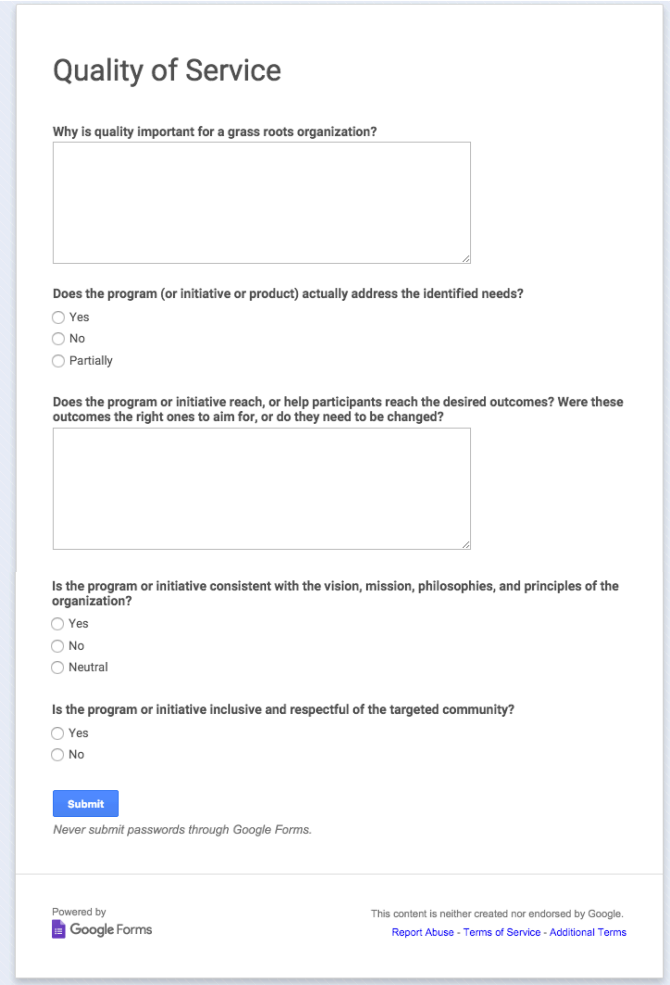

Figure 5: A picture of the questionnaire

Risk Management and Limitations

Deceptive Data:
Since information gathering is done using the survey method, sometimes they may be misleading due to different reasons, which can range from personal pressures to having fun and trying to mess with someone as an innocent prank. In order to overcome this problem, we will have to check the feasibility of the data and rule out the irrelevant data.

Inelegant questions:
As we are not really experienced in asking questions and extracting the needed information, there are chances of that we may not frame proper questions. There is also a chance that the person answering those questions may actually not understand what we want. But, reading the guidelines in implementing the survey and having the questions peer-reviewed before publishing as an extra pair of eyes can always help mitigate this risk.

Improper Responses:
There is a high chance of no response from the managers of companies as they may be busy in their own professions and cannot spare time for student projects. They may not take such things seriously and will not be interested in giving a feedback.

VI. Time and Activity Plan

Time	Plan
20th May, 2015	Submit research proposal
10 days	Search literature to answer the prepared research questions
15 days	Validate literature with questions
20 days	Conduct survey and interviews
20 days	Gather and scrutinise the received data
25 days	Document the report and submit

Table 5: Time plan of research work for the report

VII. Threats and Validity

Omission bias: The researcher may tend to prefer inaction to action for risky issues. This may affect the productivity of the outcome.

Keeping in mind that the topic of research is based solely on a person's review over a certain product; we cannot guarantee that the reviewer is not partial while giving the feedback. Hence, the only way to mitigate this risk is to choose only the feasible data while considering the feedbacks.

This is considered as a threat because, in the process of trying to eradicate deceptive data, the researcher may neglect something very important from the customer's side, thinking of it as a threat and preferred inactivity towards contemplation itself, thus affecting the productivity and accuracy of the outcome.

VIII. Expected Outcomes

After this research, we can analyze the existing methods of maintaining quality for any product and give methods to revive them if necessary based on customer review of the service of respective companies.
We can further estimate the use of Predictive Data Analytics in something as important as upholding the future of the company's existence based on its presence.

IX. Conclusion

After having attained the desired level of excellence in the development of a product, another major step in carrying the legacy forward is maintaining the same level of service. The different approaches towards maintaining the quality of service using predictive data analytics have been studied

and a survey has been conducted to evaluate the possible best methods of improvement in this area.

This is as important and tough as is establishing the standards achieved in order to sustain in any business.

X. References

[1]. Vogt, E. R. Mattfeldt, G. Satzger, L. Luders, M. Piper, O. Gehb, and W. L. Jones, "Analytical support for predicting cost in complex service delivery environments," *IBM Journal of Research and Development*, vol. 58, no. 4, pp. 7:1–7:10, Jul. 2014.

[2]. E. Ronchieri, M. Canaparo, and D. Salomoni, "A Software Quality Model by Using Discriminant Analysis Predictive Technique," *J. Integr. Des. Process Sci.*, vol. 18, no. 4, pp. 25–29, Oct. 2014.

[3]. G. Cicotti, L. Coppolino, S. D'Antonio, and L. Romano, "Big Data Analytics for QoS Prediction Through Probabilistic Model Checking," *arXiv:1405.0327 [cs]*, May 2014.

[4]. R. Krishnamurthy and K. C. Desouza, "Big Data Analytics: The Case of the Social Security Administration," *Inf. Polity*, vol. 19, no. 3,4, pp. 165–178, Jul. 2014.

[5]. C. Gan, V. Limsombunchao, M. Clemes, and A. Weng, "Consumer choice prediction : artificial neural networks versus logistic models," Jul. 2005.

[6]. "Development of Pavement Performance Models by Combining Experimental and Field Data," *Journal of Infrastructure Systems*, vol. 10, no. 1, pp. 9–22, 2004.

[7]. D. W. Walker, L. Huang, O. F. Rana, and Y. Huang, "Dynamic Service Selection in Workflows Using Performance Data," *Scientific Programming*, vol. 15, no. 4, pp. 235–247, 2007.

[8]. Chuan Shi, Rajesh Jugulum, Harold Ian Joyce, Jagmet Singh, Bob Granese, Raji Ramachandran, Donald Gray, Christopher H Heien, and John R. Talburt, "Improving financial services data quality – a financial company practice," *Lean Six Sigma Journal*, vol. 6, no. 2, pp. 98–110, May 2015.

[9]. N. Ye, S. Yan, D. Huang, M. Baydogan, B. M. Aranda, A. Roontiva, and P. Hurley, "Models of Dynamic Relations Among Service Activities, System State and Service Quality on Computer and Network Systems," Jan. 2010.

[10]. J. D. Gibson and K. Kasravi, "Predicting the Future of IT Services with TRIZ," *J. Integr. Des. Process Sci.*, vol. 16, no. 2, pp. 5–14, Apr. 2012.

[11]. E. Siegel, *Prediction Effect : How Predictive Analytics Revolutionizes the Business World*. Somerset, NJ, USA: John Wiley & Sons, 2013.

[12]. P. C. Pendharkar and J. A. Rodger, "Probabilistic and Analytical Estimation of Software Development Team Size," *Int. J. Hybrid Intell. Syst.*, vol. 7, no. 2, pp. 137–153, Apr. 2010.

[13]. T. Heyman, D. Preuveneers, and W. Joosen, "Scalar: Systematic Scalability Analysis with the Universal Scalability Law," in *2014 International Conference on Future Internet of Things and Cloud (FiCloud)*, 2014, pp. 497–504.

[14]. Deming, W. Edwards. "Out of Crisis, Centre for Advanced Engineering Study."*Massachusetts Institute of Technology, Cambridge, MA* (1986).

[15]. Deming, William Edwards. *Quality, productivity, and competitive position*. Massachusetts Institute of Technology Center for Advanced En, 1982.

[16]. Goetsch, David L. *Introduction to total quality: quality management for production, processing, and services*. Prentice hall, 1997.