**Project Proposal**

**Analyze Customer Behavior**

**Client:** Mr. Hong

**Supervisor:** Ms. Nhan Le

**Team members:** Hieu Tran | Nhiem Le | Nhu Le | Tien Thai | Vinh Truong

*Version 1.2*

*13 October 2018*

# Table of Contents

[Table of Contents 2](#_Toc527240354)

[History of Change 3](#_Toc527240355)

[Executive Summary 4](#_Toc527240356)

[I. Term of Reference 5](#_Toc527240357)

[II. Rationale 6](#_Toc527240358)

[1. Issues 6](#_Toc527240359)

[2. Opportunities 7](#_Toc527240360)

[III. Scope and Objectives 8](#_Toc527240361)

[IV. Project Approach 9](#_Toc527240362)

[V. Project Plan 12](#_Toc527240363)

[1. Sprint Timetable 13](#_Toc527240364)

[2. Role Details 14](#_Toc527240365)

[3. Training IT-specific skills plan 14](#_Toc527240366)

[4. Risk Management 15](#_Toc527240367)

[VI. Skills Analysis 16](#_Toc527240368)

[VII. Estimate All Costs Incurred 17](#_Toc527240369)

[VIII. Appendix 18](#_Toc527240370)

[References 19](#_Toc527240371)

[Marking Schedule Research & Development Project Proposal 20](#_Toc527240372)

[Marking Schedule 21](#_Toc527240373)

# History of Change

|  |  |  |
| --- | --- | --- |
| Version | Date | Change Log |
| 1.0 | 28/09/2018 | Making proposal |
| 1.1 | 11/10/2018 | Adjust:   * Project Plan - Sprint planning * Estimate costs |
| 1.2 | 13/10/2018 | Adjust:   * Scopes * Project Plan – Sprint timetable * Training IT-specific skills plan * Role details |

# Executive Summary

Our client is TMA Solution that is a company provides quality software outsourcing services to leading companies worldwide. It is one of the largest software outsourcing companies in Vietnam with 2,400 engineers.

The project that we are working on is **Analyze customer behaviors** (in 6-8 months)

Consumer behavior is the study of when, why, how, and where people do or do not buy a product. It blends elements from **Psychology**, **Sociology**, **Social Anthropology** and **Economics**. It attempts to understand the buyer decision making process, both individually and in groups. Customer behavior study is based on consumer buying behavior, with the customer playing the three distinct roles of user, payer and buyer.

The study allowed us to identify some of the following (Scope):

* Popularity of particular product categories purchased online
* Motivation behind online shopping
* Average level of spending on e-shopping
* Popularity of price comparison websites and auction platforms
* Customers' preferences regarding the form of payment and shipment collection
* Assessment of practicality and functionality of online retailers and solutions best suited to the client's needs.

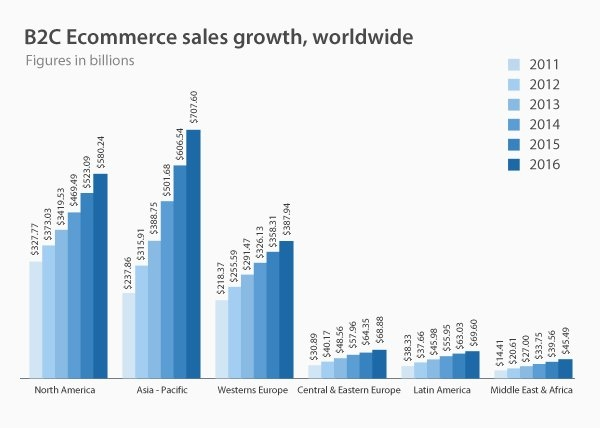
Our group is going to use Scrum as management method for all project activities.

* We will work 4 days per week and a daily meeting.
* Each a week we going to have a group sharing knowledge meeting.

We estimate that this project could take almost 5 million VND included the transportation and food and drink bills.

# Term of Reference

Nowadays, some online shopping websites or marketing companies are working on analyzing about customer behaviors to discover deep and actionable insights into their customer segments, customers’ value and how to best communicate with customers in order to edge over their competitors.

Figure 1 – World E-commerce Statistics

*(Source:* [*www.vpnmentor.com*](http://www.vpnmentor.com)*)*

To satisfy this requirement, TMA company is opening Analyze Customer Behavior Project for dynamic-micro segmentation in e-commerce/retail business. So that, our group co-operate with TMA to develop this service.

Therefore, we want a service that applied Data science and AI/ML to analyze customer behaviour, finding purchasing patterns and classify customers based on dynamic-micro segmentation for marketing purpose.

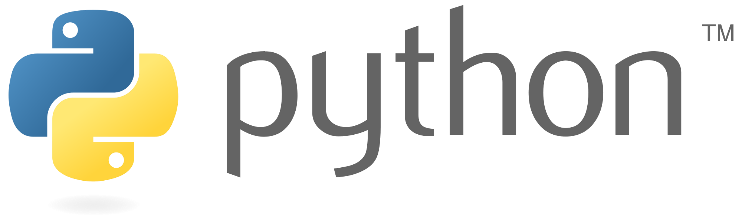
# Rationale

## Issues

In today’s highly competitive and connected environment, modern businesses can no longer survive with generic, static online content. Furthermore, marketing strategies using traditional tools are often expensive, hard to implement, and do not produce the desired return on investment. These systems often fail to take full advantage of the data collected to create customer segmentation to understand their customers and achieve effective personalized customer marketing.

Shopping online, e-Commerce is development and becoming trend over the world, but the question is “How can they predict the favorite and need of consumers?” Prediction is needed so we must have new system to solve these problems to guess all consumers behavior. Customers have access to information anywhere, anytime including where to shop, what to buy, how much to pay and so on. This makes it increasingly important to utilize predictive analytics and data to forecast how customers will behave when interacting with brands.

The goal of customer analytics is to create a single, accurate view of a customer to make decisions about how best to acquire and retain customers, identify high-value customers and proactively interact with them. The better the understanding of a customer's buying habits and lifestyle preferences, the more accurate predictive behaviors become and the better the customer journey becomes. Without large amounts of accurate data, any insight derived from analysis could be wildly inaccurate.

Without Data Science, programming languages will be complicated, wasting human power and time. Example if the coder coding Java it will cost 4 or 5 line to solve the issues but with Data Science, the programming language people often use is Python, it only code 1 line to solve problems.

In the modern world and age of 4.0 data size will increase day by day, with old devices and techniques they will face hard problems to solve or never be. However, it cannot deal with these problems if there’s not using Big Data to stores and manages. Therefore, in marketing world will using new ways to find out solutions and saving money for develop new devices or techniques.

## Opportunities

* Analyze customers’ behaviors to understand customers and achieve affective personalized customer marketing and communicate with customers base on their unique preferences and behaviors help business owners maximize customer engagement, loyalty and lifetime value.
* This service can be used to provide the best products based on customer demands classification and also help business realize market demands.
* We can let the tracking and analyzing of data for the machine and focus on creativity.
* Deploying new products and developing marketing strategy.

# Scope and Objectives

The main objective of the Analyze Customer Behavior System is to analyze customer behaviors for dynamic-micro segmentation in e-commerce retail business.

The goals are:

* Create system to analyze the customer behavior
* Create assumption database to integrate into the website use to analyze
* Create a dashboard that has visualize data that can customize by consumers
* Predict customer favorite based on data analyzed
* Giving extra sell offers for customers

# Project Approach

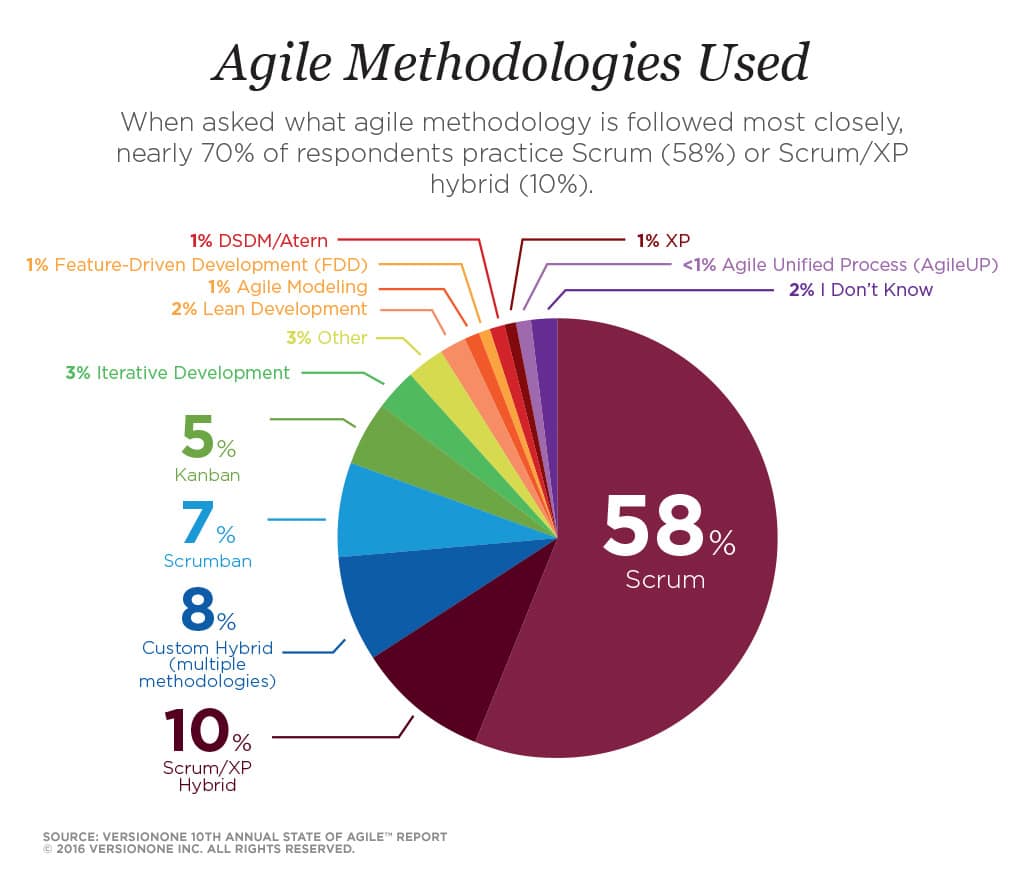
After meeting with client to get the requirements for project, our team decide to apply Scrum methodology to manage this project. First of all, Scrum is not only the most popular method in series of Agile methodologies in 2016 (58%) but also its framework fit to our project and client requirements.

Figure 3 – Agile Methodologies Used

As a subset of the Agile framework, Scrum model is iterative flexible. To apply Scrum for this project, we are going to redefine core values, specific artifacts, roles and events of Scrum.

1. Core values:

* Transparency: The most important values of Scrum method. All information evolve with development process must clear and transparency. For example, it could be vision of product, requirement of client, work in progress, limitation and problems. Understanding more about necessary information that can help to improve the performance. All tools and meeting must ensure the transparency of project for stakeholders.
* Inspection: Continuous inspection of Scrum activities ensure that team can detect issues and solution as soon as possible.
* Adaption: Base on the inspection and development, Scrum can response to the changes of requirements more quickly, so that it can deliver most values for product owner.

1. Artifacts: Scrum use simple tools to support the workflows more effective.

* Product backlog: The list of priority feature or requirements of project. Product Owner responsible for arrange the priority of Product Backlog Item (usually based on business value).
* Sprint backlog: the result of each Sprint Planning. Development Team analyze product Backlog base on the priority level and turn it to the task in TODO list.
* Burndown Chart: The chart provides the trend of project base on necessary time to complete task as Sprint Burndown Chart or Project Burndown Chart.

1. Roles: As Scrum team, it usually includes three main roles and responsibilities to ensure optimize workflow for each progress, as our R&D Project:

* Product Owner: (Mr. Tran Quang Thang) who responsible for the success of project: identify the project requirements and assess the final deliverable of Development team.
* Scrum Master: Ensure the team can perform as well as with Scrum method.
* Development Team: Transform product Backlog to the function of system.

1. Events: Scrum define rules for four key events to create an environment, operational and collaborative framework for project members.

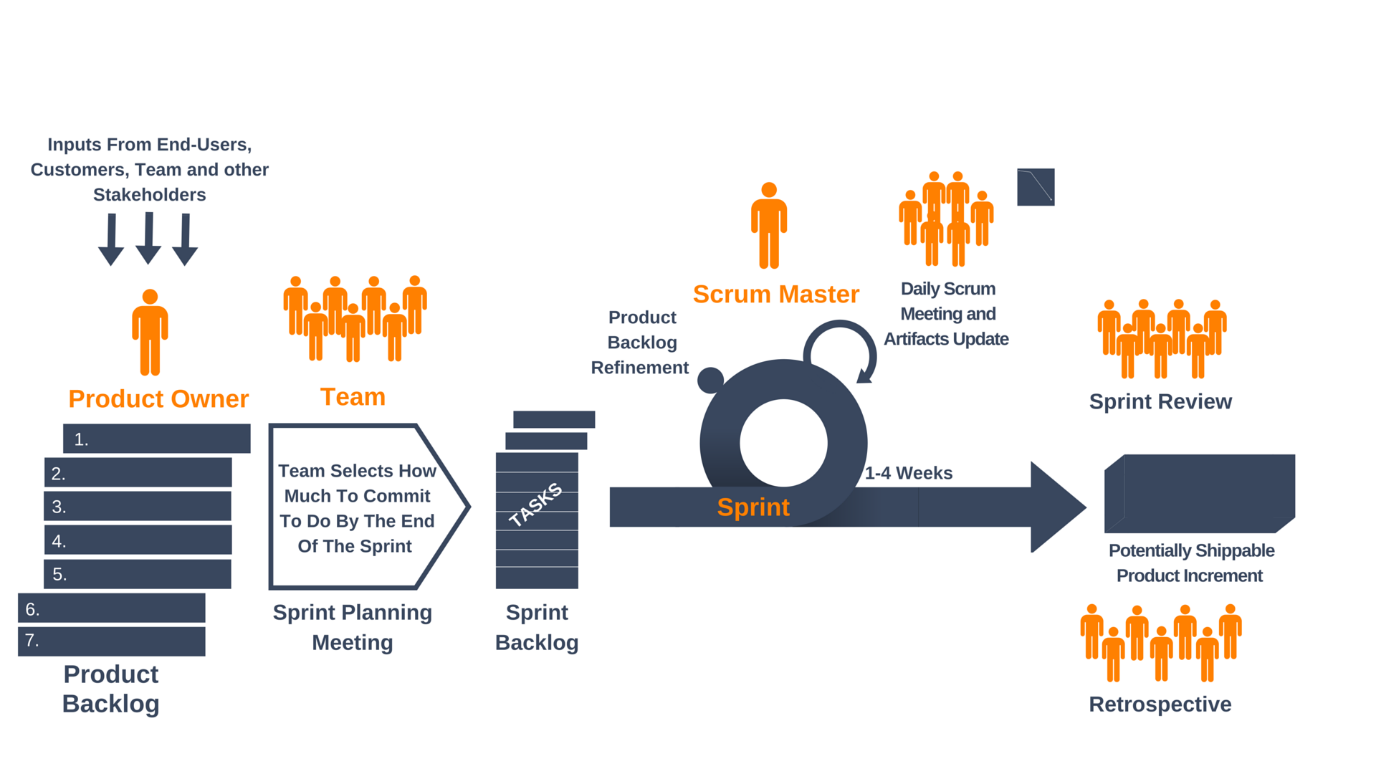
* Sprint Planning: The Development team meet Product Owner to set up the work flow for one Sprint. Work planning include select requirements to develop, analyze and realize the task should be done with estimated time and human power.
* Daily Scrum: Scrum Master celebrate daily meeting for development team about 15 minutes to share about the work in progress and problems during a Sprint.
* Sprint Review: At the end of each Sprint, Development Team with Product Owner will inspect the done works to suggest the modifications or changes that are necessary for project.
* Sprint Retrospective: With the support of Scrum master, the development Team will inspect again whole Sprint to find out the methods to improve the process or themselves.

Figure 4 – Scrum framework

*(Source:* [*www.yanado.com*](http://www.yanado.com)*)*

Each Sprint will repeat again and again until all the items of Product Backlog are completed or when the Product Owner decide to stop the project. Because Scrum apply the strategy” focus on complete priority task more than first things must do”, so the most valuable tasks always completed first. Therefore, Scrum usually bring highest benefit to the investor of project. As the process is always improving, the Scrum team often has very high productivity. These are two great benefits that Scrum brings to the organization.

As those advantages of Scrum mentioned above, if our team apply this method into the development AI system to analyze customer behavior, our team are able to manage this project well, improve each personal skill and guarantee the benefits of Product Owner during the project.

# Project Plan

**8 Oct**

**Nov**

**Dec**

**Jan**

**Feb**

**Mar**

**May**

**26 Sep**

Training

**Sprint 1**

Data

analysis

**Sprint 2**

Request for Proposal

**Sprint 0**

Implement

A

**Sprint 3**

Implement

B

**Sprint 4**

Testing

**Sprint 7**

Review,

Fix bugs

**Sprint 8**

Implement

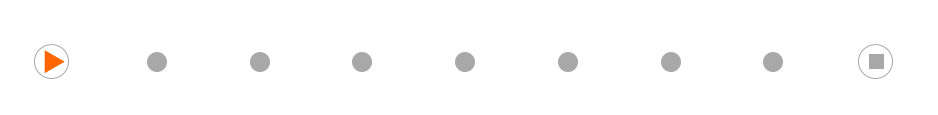
c

**Sprint 5**

Implement

B

**Sprint 6**



**Apr**

INITIATING

EXECUTING

CLOSING

**Step 1**

Gathering data from various sources

**Step 2**

Cleaning data

to have homogeneity

**Step 3**

Model Building

*Selecting the right*

*ML algorithm*

**Step 4**

Gaining insights from the model’s results

**Step 5**

Data Visualization *Transforming results into visuals graphs*

***Machine Learning Process***

## Sprint Timetable

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sprint Name | Details | Start Date | End Date | Resource Names |
| **Sprint 0** | Making proposal | 26/09/2018 | 28/09/2018 | All members |
| **Sprint 1** | Acquire Python | 08/10/2018 | 30/11/2018 | All members |
| Machine Learning | Hieu, Nhiem |
| Data Analysis | Vinh, Tien |
| Data Visualization | Nhu |
| **Sprint 2** | Acquire Python | All members |
| Machine Learning | Hieu, Nhiem |
| Data Analysis | Vinh, Tien |
| Data Visualization | Nhu |
| **Sprint 3** | Data collection and description | 03/12/2018 | 04/12/2018 | All members |
| Data cleaning | 06/12/2018 | 07/12/2018 | All members |
| Data integration | 10/12/2018 | 14/12/2018 | Vinh, Tien |
| Data transformation | 17/12/2018 | 21/12/2018 | Vinh, Tien |
| Data reduction | 24/12/2018 | 25/12/2018 | Vinh, Tien |
| Data discretization | 27/12/2018 | 28/12/2018 | Vinh, Tien |
| **Sprint 4** | ML algorithm | 31/12/2018 | 11/01/2019 | Hieu, Nhiem |
| Model building | 14/01/2019 | 15/02/2019 | Tien |
| Testing | 18/02/2019 | 22/02/2019 | All members |
| **Sprint 5** | Conducting Experiment | 25/02/2019 | 29/03/2019 | All members |
|  |  |  |  |
| **Sprint 6** | Dashboard | 01/04/2019 | 26/04/2019 | Nhu |
| **Sprint 7** |  |  |  |  |
| **Sprint 8** |  |  |  |  |

## Role Details

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Role  Name | Team Leader | Business Analyst | Data Scientist | Data Engineer | Developer | Tester |
| Hieu Tran | ✓ |  | ✓ |  | ✓ | ✓ |
| Nhiem Le |  |  | ✓ |  | ✓ | ✓ |
| Nhu Le |  | ✓ |  |  | ✓ | ✓ |
| Tien Thai |  |  |  | ✓ | ✓ |  |
| Vinh Truong |  |  |  | ✓ | ✓ |  |

## Training IT-specific skills plan

|  |  |  |
| --- | --- | --- |
| Name | Training Skills | Time Requirement |
| Hieu Tran | * Python * Machine Learning | 08/10/2018 – 30/11/2018 |
| Nhu Le | * JavaScript * Business Analysis * Data Visualization | 08/10/2018 – 30/11/2018 |
| Nhiem Le | * Python * Machine Learning | 08/10/2018 – 30/11/2018 |
| Tien Thai | * Python * Data Analysis | 08/10/2018 – 30/11/2018 |
| Vinh Truong | * Python * Data Analysis | 08/10/2018 – 30/11/2018 |

## Risk Management

1. **High** – Riskhas great impacts to project performance that cost a huge amount of time.
2. **Medium** – Risk has slight impacts to project performance that cost a moderate amount of time.
3. Low – Risk has little impacts to project performance that cost a little amount of time.

|  |  |  |
| --- | --- | --- |
| Risk | Impact | Solution |
| Member quit the project | **High** | Seek advice from supervisors, advisors and keep working until the deadline |
| Client change Project requirement | **Medium** | All members must have a meeting and decide whether to accept the requirement depend on the difficulty, importance, time, feasibility level. |
| Problem with learning new coding language (TensorFlow) | **Medium** | Ask supervisor or team member will support each other to get the solution. |
| Internal issues | **High** | Team leader will provide a solution or ask supervisors for further discussion. |
| Technical issues | **High** | Leader need to contact to supervisor for solution |
| Time requirement | Medium | All members need to work at least 350 hour and show the contribution in Gantt chart as evidence |

# Skills Analysis

|  |  |
| --- | --- |
| Skills | Knowledge |
| Professional Skills | * Teamwork * Critical thinking * Communication * Decision making * Research and analysis * Problem solving * Coding and fixing * Carefulness and responsiveness |
| Technical Skills | Programming language:   * Python * Machine Learning and AI   Database:   * SQL * NoSQL * Data Visualization * Unstructured data   Management Tools:   * Trello * Google Drive   Communication Tools:   * Skype * Facebook   Coding Tools:   * Pentaho * Atom |

# Estimate All Costs Incurred

***Working labor time formula:***

**Working time** = **8 hours/day** *(8h30 - 12h, 13h30 - 18h)* \* **4 days/week** *(Monday, Tuesday, Thursday, Friday)*

|  |  |  |  |
| --- | --- | --- | --- |
| Member | Initiating Phase | Executing Phase | Closing Phase |
| Nhu Le | 32 hours/week | 23 hours/week | 25 hours/week |
| Vinh Truong | 32 hours/week | 28 hours/week | 20 hours/week |
| Nhiem Le | 32 hours/week | 28 hours/week | 20 hours/week |
| Tien Thai | 32 hours/week | 28 hours/week | 20 hours/week |
| Hieu Tran | 32 hours/week | 28 hours/week | 20 hours/week |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Name | Unit | Quantity | Price per unit | Total |
| Hardware | PC/Laptop | Device | 5 | 0 | 0 |
| Peripherals | Device | 10 (5 mouses + 5 keyboards) | 0 | 0 |
| Software | Office 2016 | Package | 5 | 0 | 0 |
| G Suite (Doc, Drive, Gmail,…) | Package | 5 | 0 | 0 |
| Wrike | Online tool | 5 |  |  |
| Trello | Application | 5 | 0 | 0 |
| Skype | Application | 5 | 0 | 0 |
| Others | Internet fee | Monthly | 5 | $2 | $60 |
| Electricity fee | Monthly | 5 | 0 | 0 |
| Working place | Room | 1 | 0 | 0 |
| Fuel fee | Litter | 100 | $1 | $100 |
| Lunch | Meal |  |  | $200 |
| All cost total | | | | | $360 |

# Appendix

While all due care and diligence will be expected to be taken by the students, (acting in software development, research or other IT professional capacities), and the Auckland University of Technology, and student efforts will be supervised by experienced AUT lecturers, it must be recognized that these projects are undertaken in the course of student instruction. There is therefore no guarantee that students will succeed in their efforts.

This inherently means that the client assumes a degree of risk. This is part of an arrangement, which is intended to be of mutual benefit. On completion of the project it is hoped that the client will receive a professionally documented and soundly constructed working software application, some part thereof, or other appropriate set of IT artefacts, while the students are exposed to live external environments and problems, in a realistic project and customer context.

In consequence of the above, the students, acting in their assigned professional capacities and the Auckland University of Technology, disclaim responsibility and offer no warranty in respect of the “technology solution” or services delivered, (e.g. a “software application” and its associated documentation), both in relation to their use and results from their use.

# References

*Customer Behavior Analytics.* (2018). Retrieved from NGDATA: https://www.ngdata.com/dictionary/customer-behavior-analytics/

Gökhan Sılahtaroğlu, H. D. (2015). *Analysis and Prediction of E-Customers’ Behavior by Mining Clickstream Data.* IEEE.

*Learn About Scrum*. (n.d.). Retrieved from Scrum Alliance: https://www.scrumalliance.org/learn-about-scrum

Matthias Volk, A. E. (2017). *New E-Commerce User Interest Patterns.* IEEE.

Nelson, P. (n.d.). *An Open Source Approach to Log Analytics with Big Data*. Retrieved from Search Technologies: https://www.searchtechnologies.com/blog/big-data-open-source-log-analytics

Rouse, M. (2017). What is Customer Analytics? *SearchBusinessAnalytics*.

*State of Scrum 2017 - 2018.* (2018). Retrieved from Scrum Alliance: https://www.scrumalliance.org/

*TMA Solutions*. (n.d.). Retrieved from TMA Solutions: https://www.tmasolutions.com/

*What Customer Behavior Analytics Can Do For You.* (2018, February 26). Retrieved from Datameer: https://www.datameer.com/blog/customer-behavior-analytics-can/

# Marking Schedule Research & Development Project Proposal

|  |  |  |
| --- | --- | --- |
| Title of project |  | |
| Student’s name |  | |
| Student ID number |  | |
| Team Members |  | |
| Date submitted |  | |
| Name and signature of Assessor |  |  |
| Name and signature of Moderator |  |  |

**Conditions/Suggestions (if any) from review panel:**

**Other Comments:**

# Marking Schedule

|  |  |
| --- | --- |
| **CRITERIA** | **FEEDBACK** |
| Proposal includes all required elements |  |
| Terms of reference and initial project scope clearly identified |  |
| Methodology / Approach clearly delineated and justified |  |
| Required costs and resources for project identified |  |
| Proposed actions (with deliverables) identified and justified |  |
| Skills and Know-how involved identified |  |
| Suitably professionally presented and referenced. |  |
| Quality of proposal presentation |  |

**Grade for Project Proposal**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **A+** | **A** | **A-** | **B+** | **B** | **B-** | **C+** | **C** | **C-** | **D** | **D** |
|  |  |  |  |  |  |  |  |  | Over 40% | Under 40% |

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Approved** | **Yes** 🞏 | **No** 🞏 | **Date:** …………………… |