

## **FIT3163 - Data Science Project 1 - Assignment 2 - PM Case Study**

**Workshop: Friday, 2pm - 4pm**

**Group: 4**

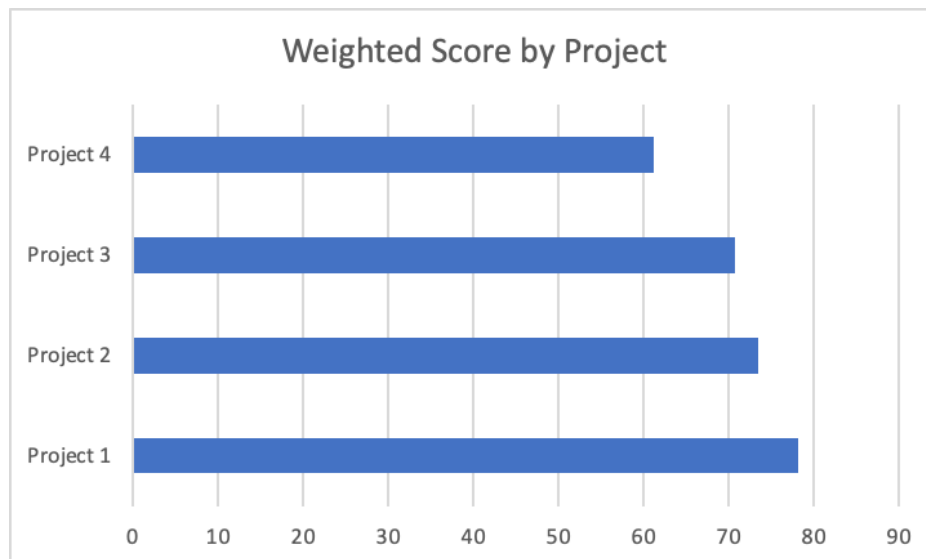
**Members: Elaine Liong (ID: 29942357), Jack (ID: 29037077), Vionnie Tan (ID: 30092809)**

---

### **Case 1 : Project Selection and Developing a Business Case**

**Task 1: Weighted Scoring Model for Data mining technique to detect cancer using predictive modelling**

Criteria	Weight	Project 1 - Cancer Prediction	Project 2 - Covid 19	Project 3 - Water Management	Project 4 - Zoom killer
Has low risk in meeting scope, time, and cost goals	15%	85	65	90	65
Can be implemented in one year or less	15%	80	30	75	40
Technical Learning	30%	75	90	50	40
Support Key Business Objectives	25%	70	75	85	80
Uses realistic level of technology	15%	90	90	65	90
Weighted Project Scores	100%	78.25	73.5	70.75	61.25



After carefully considering different criterias - risks, technicality, objectives and based on the scores from the weighted scoring model, Project 1 - Data mining technique to detect cancer using predictive modelling with a score of 78.25 seems like a viable choice for us.

### Justification

**Project 1** - High risk because we are dealing with health issues where there are many lives at stake. Scope is small because the main topic is only cancer. We also agreed that it can be implemented within a year due to the small scope needed for the project. We rated Technical Learning at 75 because there is a limited amount of new scopes that could be modelled. It uses a high level of technology because to develop a predictive model, we would need the latest packages to develop a better model.

**Project 2** - We rated Covid project having relatively low risk, due to the many models that are currently being developed across the world. But, we are unsure that the covid project could be implemented within a year due to the still ongoing course of the pandemic across the world. Technical Learning is high because upon the success of the project we'll be able to get rid of covid19 fully. It supports key business objectives as we'll essentially be trying to suppress the coronavirus pandemic. It uses a very high technology, as we would also need the newest form of technologies - especially considering the very new nature of covid19

**Project 3** - Water project has high risk due to the brittle nature of the project and we're highly uncertain about being able to complete the project within a year due to wide scopes that would need to be covered during the project. There's not much technical learning as we'll essentially just be trying to increase efficiency of water management systems. It wouldn't also need the latest technologies due to the unchanging nature of the project scope.

**Project 4** - We rated Zoom project as having low risk because developing a brand new application - especially one that has already existed in the market wouldn't pose a high risk. But, developing an application within a year would be a hard task, as most new start up applications might take extensive years of work to be optimized. There's not much to be

learned from this project as it's not developing an app that has not been on the market. But, it would need the extensive use of high end technologies to develop it.

## **Task 2 - Business Case for Data Mining Techniques to detect Cancer using Predictive Modelling**

### **1.0 Introduction/ Background**

According to the American Cancer Society (2018), cancer dates back to as early as 3000 BC, where the first cases of known cancer were documented on manuscripts with descriptions of tumors on the breast. Since then, cancer treatment has changed drastically. Now, the most common way to treat cancer is through chemotherapy and medications. In Australia, cancer is one of the leading causes of death - accumulating to over 50,000 deaths in 2019 (Cancer Council Australia, n.d.). The Australian Institute of Health and Welfare (2020) predicts that in 2020, new cases of cancer are expected to rise to up to 150,000 cases.

Predictive Model has been around for quite some time, with the famous models being risk grouping, probability tables, classification and regression tree, etc. Some cancers do not show symptoms at earlier stages, and having predictive models could determine the likelihood of the cancer and lower the risk of having to deal with other treatments. By developing this model, we aim to enhance the model to determine early stages of cancer so health institutions can provide this information to patients for immediate treatments which can reduce the rate of cancer and also save a lot of lives

### **2.0 Business Objective**

Our Business Objective is to utilize predictive modelling to determine the early stages of cancer. The predictive model will be tested on various other models, and we will aim to develop a model with the greatest accuracy. Upon successful completion of the predictive model, it can be given to health institutions for their use

### **3.0 Current Situation and Problem/Opportunity Statement**

We are currently finding contributing factors that lead to the development of cancer. There is an opportunity for us to develop this model as cancer is one of the leading causes of death in Australia. However, in the current situation, we are limited by the small dataset provided.

### **4.0 Critical Assumption and Constraints**

The current clients and consultant must actively support the project. The new system must run on existing hardware and software, and it should require minimal technical support. There is an open data source available to the public but there are limitations of datasets that covers only gastrointestinal cancer.

## **5.0 Analysis of Option and Recommendation**

- Training
  - Option 1: Train on given dataset
  - Option 2: Train using classification and forecasting model
  - Preferred option: Train on given dataset
- Deploying:
  - Option 1: Website
  - Option 2: Mobile Application
  - Preferred option: Website

## **6.0 Preliminary Project Requirements**

The preliminary project requirements of this project are as following:

- Valid Dataset to be used
- Good Knowledge on coding and capable to code based on project requirements
- Not changing teams from start to end of project
- Constant support from consultant

### 7.0 Schedule Estimate

- First six months dedicated to project planning
- Last six months dedicated to further development of the predictive model (Coding)
- Expected to finish within a year
- The approximate schedule is as follows:

	Week	Task
Semester 1	2 - 4	<ul style="list-style-type: none"><li>• Getting to know members of the team</li><li>• Preliminary Project Management, creating a business case and problem statement for our project</li></ul>
	5 - 6	<ul style="list-style-type: none"><li>• Project Scope Management</li></ul>
	7-9	<ul style="list-style-type: none"><li>• Project design</li></ul>
	9 - 12	<ul style="list-style-type: none"><li>• Presentation of current project (where we are in the timeline)</li><li>• Literature Review for Final Submission</li></ul>
Semester 2	1 - 6	<ul style="list-style-type: none"><li>• Coding</li></ul>
	7 - 11	<ul style="list-style-type: none"><li>• Product Testing</li></ul>
	12	<ul style="list-style-type: none"><li>• Final Presentation</li></ul>

### 8.0 Potential Risks

There are several risks when carrying this project. The leading risk is that we are unable to develop a successful predictive model from the data and it leads us to develop an unsuccessful predictive model. The model may not be up to the standard of the business objective and it may also be an inaccurate model which means it cannot pass it on to the health institutions. Besides, there might be a risk that one of the team members has decided to leave the team halfway through the project. Moreover, there might be some technical issue in choosing which software to develop the model.

## Case 2: Project Management Methodology

Project Management Methodologies have become an ongoing concern throughout public and private industry sectors due to their steadily high failure rates (Serrador & Pinto, 2015). Project Management, with its associated risks needs to be analyzed and understood as it could foster upon uncertainties and unfavourable situations in regards to project success (Buganová and Šimíčková, 2019). These, coupled with the ever-changing nature of the world served as a motivation to researchers and analysts alike to look for alternative methods for newer and efficient methods of project management methods (Serrador & Pinto, 2015). One of the newer approaches researchers concluded with is called Agile Methodology, which focuses on the flexibility of project scopes and adapting towards continuous changes throughout the project. It is often described as an iterative approach over traditional project management methodologies that takes into account the rigorous importance of initial planning and an unchanging scope (Serrador & Pinto, 2015). Although Agile Methodologies have started being incorporated into a much larger project scope, their effectiveness is still debatable over traditional project management methodologies (Serrador & Pinto, 2015). The purpose of this report is to examine the pros and cons of Agile methodologies and traditional methodologies and provide justification over which approach would suit our team's approach towards our project.

According to Serrador and Pinto (2015), Agile approach can improve the likelihood of project success considering the fact that it emphasizes on individuals and interactions over processes, customer collaboration over contracts or formal negotiations, as well as responsiveness over rigid planning. This approach also consists of frequent working software delivery which implies that there are "short feedback loops [that] lead to quick adaptation to changing requirements" (Cypress Data Defense, 2020). This also means that a project using the agile approach eliminates the risk of complete project failure for the reason that "[t]here is always a small part that can be salvaged and used in the future even if a particular approach doesn't go as planned" (Kissflow, 2021). However, short cycles in agile projects might not leave enough time for the design process and as a consequence, the product has to be redeveloped due to negative feedback from the customer (Olic, 2017). In addition, Lynn (n.d) mentioned that agile approach might have poor resource planning as the projects have no clear and succinct idea of what their end products should look like. Another disadvantage to this approach is that "factors such as the overly aggressive push to reduce costs are at work to distort estimates" and "[t]his can result in adverse outcomes such as incorrect projects being selected, over- runs costing more than the original rejected estimate

or projects that might have been more beneficial being passed over” (Serrador, 2015, p. 119).

Opposed to Agile approach, Predictive approach to project management is easier in terms of operation and monitoring the entire project as there are specific and dedicated timelines for each stage of the project (Cypress Data Defense, 2020). Moreover, team members will find it easier to understand and follow as each of the phases is initiated after another phase is complete (Cypress Data Defense, 2020). In contrast to agile approach, the predictive approach lacks customer’s opinion in its processes until the end product is ready. This will be a waste of time if the product does not satisfy the customer’s requirements and it is time-consuming as well as costly to go back and fix the product (Web Applications & Data Integration Company, n.d.). Additionally, Gemino et al. (2021) mentioned that less communication which leads to surprises surfacing later in the project can make sponsors and teams dissatisfied.

There are multiple pros and cons for both agile approach and predictive approach, however, agile approach is much more appropriate in this situation. According to Ungureanu and Ungureanu (2014), Agile approach is specific and applicable especially for collaborative teams. We are a collaborative team so that we can exercise agile methods to its full extent. Moreover, agile methods are not strongly constrained by budget for materials, which is also applicable for us, as a project given by university, no budgets are given to us so that in this case agile method works better than predictive or waterfall methods. Besides that, Ungureanu and Ungureanu (2014) also mentioned that agile methodologies propose teamwork and establish on-going collaboration and adaptability throughout the project life cycle and development iterated, as this project is separated into two different units, a project management method which is stronger in establishing on-going collaboration and adaptability is much more suitable in this situation which is agile methodologies. According to Gemino et al. (2021), agile methods are also useful in attaining the same level of stakeholder success while providing the same budget, time, scope, quality output than predictive approach. Moreover, according to Buganová and Šimíčková (2019), agile methods have higher adaptability to change and the contracting authority which is us students can adapt to the project to the current needs of the customer or the feedback which is the feedback given from tutor. Besides, Buganová and Šimíčková (2019) also stated that agile method is much more suitable than predictive method for smaller team that communicate with each other personally and have a close collaboration. The criteria is perfect for a team like us which is formed from 3 members.



## Case 3: Project Scope Management

**Task 1 - Develop a project scope statement for your project. Be as specific as possible in describing product characteristics and requirements, as well as all of the project's deliverables.**

### Scope Statement (Version 01)

**Project Title: Data Mining Technique To Detect Cancer Using Predictive Modelling**

**Date: 12 April 2021**

**Prepared by: Group 4**

#### **Project Justification:**

The Cancer Prediction Model would be a user friendly web-based application distributed to health institutions to help determine and detect cancer at an earlier stage. Detecting cancer at an earlier stage allows patients to receive earlier medications.

#### **Product Characteristics and Requirements:**

##### **IN-SCOPE:**

- Building a cancer predicting system to determine cancer earlier.
- A user friendly website for health institutions to access.
- Tutorial for users to use the system.

##### **OUT OF SCOPE:**

- Side-effect of cancer predicted
- Respective medicine for cancer predicted
- Accuracy of the model > 98%

##### **LIMITATIONS/CONSTRAINTS:**

- Limited / Lack of relevant datasets for testing
- Coding Capabilities

#### **Summary of Project Deliverables**

##### **Project management-related deliverables:**

1. Business Case
2. Weighted Scoring Model
3. Scope statement
4. Requirements Traceability Matrix
5. Risk Register
6. Final Project Presentation
7. Final Project Report

##### **Product-related deliverables:**

1. Predictive model software code
2. Research Report
3. Website design document
4. Website software code

<b>Project Success Criteria:</b>

- |                                                                                                                                                                          |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>1. Met expectations with project sponsor</li><li>2. High accuracy predictive model</li><li>3. High customer satisfaction</li></ul> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

**Task 2 - Develop a requirements traceability matrix, which should include at least two functional requirements and two non-functional requirements.**

REQUIREMENTS TRACEABILITY MATRIX					
Project Name:	Data Mining Technique To Detect Cancer Using Predictive Modelling				
Project Manager Name:	Afsaneh Koohestani				
Project Description:	Building a predictive model to determine early stages of cancer				
ID	Requirements (Functional or Non-Functional)	Assumption(s) and/or Customer Need(s)	Category	Source	Status
001	Image Processing - Find relevant dataset that could be used for testing & training	Source of images are come from a legit source	Functional	Kaggle	In Progress
002	Programming Skills - Understanding topics regarding AI, Machine Learning, Deep Learning, Transfer Learning	Extensive knowledge of these programming skills increases the chance of efficiency of our predictive model	Functional	Online resources such as Stack Overflow, Leetcode, and Monash Units	Planning
003	User Interface - Allow login specific of the health institutions	Our stakeholders have to be able to easily access and understand the interface for them to use the predictive model	Non-functional	Project supervisor	Planning
004	Stakeholder expectations met	Accuracy of the Model has to be > 98%	Non-functional	Stakeholders	Planning
005	The software system should be integrated with health institutions's AP	If we are building on an existing model, then our predictive model must be integrated with the current system.	Functional	Stakeholders	Planning

006					
007					
008					
<b>Documentation: (Include any justification and assumptions made)</b>					

# References

- American Cancer Society. (2018, January 4). *Early History of Cancer. What is Cancer?*. Retrieved March 29, 2021, from <https://www.cancer.org/cancer/cancer-basics/history-of-cancer/what-is-cancer.html>
- Australian Institute of Health and Welfare. (2020, Nov 13). *Cancer data in Australia. Rankings*. Retrieved March 29, 2021, from <https://www.aihw.gov.au/reports/cancer/cancer-data-in-australia/contents/rankings>
- Buganová, K., & Šimíčková, J. (2019). Risk management in traditional and agile project management. *Transportation Research Procedia*, 40, 986–993. <https://www.sciencedirect.com/science/article/pii/S2352146519303060>
- Cancer Council Australia. (n.d.). *Cancer Statistics in Australia. Facts and Figures*. Retrieved March 29, 2021, from <https://www.cancer.org.au/cancer-information/what-is-cancer/facts-and-figures>
- Cypress Data Defense. (2020, March 02). *Predictive vs. Adaptive SDLC: What is the Difference?*. Retrieved March 29, 2021, from <https://www.cypressdatadefense.com/blog/predictive-vs-adaptive-sdlc/>
- Gemino, A., Reich, B. H., & Serrador, P. M. (2021). Agile, Traditional, and Hybrid Approaches to Project Success: Is Hybrid a Poor Second Choice?. *Project Management Journal*, 52(2), 161-175. <https://journals.sagepub.com/doi/pdf/10.1177/8756972820973082>
- Kissflow. (2021, March 21). *The 9 Key Benefits of Using the Agile Methodology*. Retrieved April 2, 2021, from <https://kissflow.com/project/agile/benefits-of-agile/>
- Lynn, R. (n.d.). *Disadvantages of Agile*. <https://www.planview.com/resources/articles/disadvantages-agile/>
- Olic, Aleksandar. (2017, June 21). *Advantages and Disadvantages of Agile Project Management [Checklist]*. Retrieved April 3, 2021, from <https://activecollab.com/blog/project-management/agile-project-management-advantages-disadvantages>
- Serrador, P. (2015). *Project Planning and Project Success: The 25% Solution*. CRC Press.
- Serrador, P., & Pinto J. K. (2015). Does Agile Work? - A quantitative analysis of agile project success. *International Journal of Project Management*, 33(5), 1040-1051. <https://www.sciencedirect.com/science/article/pii/S0263786315000071>
- Ungureanu, A., & Ungureanu, A. (2014). Methodologies used in Project Management. *Annals of Spiru Haret University Economic Series*, 14(2), 47-53.

[https://www.researchgate.net/publication/319229966\\_Methodologies\\_used\\_in\\_Project\\_Management](https://www.researchgate.net/publication/319229966_Methodologies_used_in_Project_Management)

Web Applications & Data Integration Company. (n.d.). *11 Advantages & 9 Disadvantages of Traditional Project Management*. Retrieved April 3, 2021, from <https://wadic.net/traditional-project-management-advantages-disadvantages/>