IT INDUSTRY PROJECT - PROJECT PLAN

BUDGET UTILISATION - Data Science Project

Team Techmates

1. Drashti Nayak (n10599568)
2. Roberto Carlos da Silva junior (n10374647)
3. Harshal Majithiya (n10550453)
4. Sarjak Tapodhan (n10553916)
5. Sing Yin Chan (n9317007)

Project Supervisor: Dr Venkat Venkatachalam

Industry Partner: Leap in!
Industry Supervisor: Jane Sheehy (CTO)

This project plan is submitted to Science and Engineering faculty in partial fulfilment for the Master of Information Technology degree.

The plan intends to deliver a machine learning predictive model for industry partner Leap In, which will help the organisation continuously track member activities and, based on pattern, provide targetable actionable recommendations.

Table of Contents

I. Introduction

1.1. Introduction to the Industry Partner

There are around 4.3 million Australians who have any permanent or significant disability. The National Disability Insurance Scheme (NDIS) is a national scheme that provides reasonable and necessary support to those with any disability to live an ordinary life—given a plan of support developed and tailored to their individual needs. And when required, it funds necessary supports that help participants achieve their goals (National Disability Insurance Agency, 2021).

Leap In is an NDIS not-for-profit plan manager that supports its members to utilise their fully NDIS budgets and; it does it by collecting NDIS participant's data and how they use their annual budget. Then mines their claims to look for patterns in the data, linking member budgets to other attributes such as location, life stage, nature of the disability, available funds, category of spend, and opportunities to improve the NDIS utilisation. Each month the organisation sends a report to members on their spending with the use of fundamental analysis. In this report, the spending's have showcased into three supports: core supports, capacity building supports, and capital supports. These support further differentiated into minor categories about budget allocated to members. The fundamental analysis describes details regarding the current budget, what spent, what's available, revised weekly budget and basic interpretation in cases of underspending and overspending.

1.2. Project Aim:

On average, each NDIS participant receives $67,000 in support budget annually. However, the way this financial support used may be an issue as it stipulated just 68% of this annual budget used (NDSP Plan Managers, 2020) because some NDIS participants don't know how to spend the total amount as some can't find supports that are available in their area, or just because they aren't happy with the supports provided (Leap In, 2021). Leap In found that members have problems with underspending rather than overspending. So an opportunity is seen to arise to support the need, improve and automate the budget handling for members using data analysis and machine learning concepts. The automated derived model must help identify early markers of underspending/overspending by continuously monitoring user activity and recommending related services that turn into actionable feedback for members.

1.3. Project Objectives:

Concisely, the project deliverables include:

1. Based on the data received, a visualising dashboard showcases insights based on past data created with the help of data analysis. This analysis will help make automated model features and help the client internally understand the data and use it for future decision-making.

2. A machine learning model can track user activity and return the relevant result indicating if the user might go off-track and affect the utilisation process.
3. If the time constraint is affordable, the team can add a recommendation system for members based on longitudinal information.

The objectives of this project are defined with a fact-based and data-oriented solution. Using SMART techniques, goals described into measurable, assignable and real-time deliverables as follows:

Specific	● Based on past spending data of members of Leap-In, we would like to accomplish a machine learning model which can predict whether a member is about to go off track in the time frame of 7 weeks. ● The project will be collaborative between QUT university students and Leap-In partners. ● The project is online-based; hence the team can access it on work from home basis. But, weekly meetings will take the palace face-to-face in case needed. ● This solution helps organisations advance and automate operational processes for accurate insights. Additionally, increased utilised budgets for members.
Measurable	To measure our progress, we have set milestones that must be achieved to accomplish the final deliverable. ● Data analysis and visualisation: After receiving clean data, when the team will carry out mathematical calculations. The insights received help understand affecting factors, variance, and ways to teach data to read and give significant models. ● Milestones will be the primary data analytic cycle processes followed step by step and include

	machine modelling processes.
Achievable/Assignable	● To achieve the desired outcome, we will require open-source tools like python, cloud service and Tableau. Not all team members are comfortable with python as a language and Tableau. Hence, some initial introduction and Bootcamp training provided where needed.
Relevant/Real	● The solution is kept relevant by focusing on data to be directed to member's needs and not the organisations employees need to handle budgeting. Also, keeping in mind that they don't need a communication medium but a prescriptive model which combines analysis and analytics.
Time-Bound	● The project will run from April to June 2021, with the allocation of a business analyst, a data analyst, a front-end developer, a back-end programmer and a machine learning engineer. ● In 7 weeks of duration, the first two weeks will handle past data and analyse spending patterns. For future predictions, these results will be used for the next five weeks with the help of a machine learning model.

(Table No. 1 SMART Technique)

1.4. Potential Impact:

The solution will have many positive impacts on the organisation's operating environment and also on its customers.

- AI is still in its very early stages, where people primarily work on weak AI projects. But with the use of data, industries are booming with possible profitable strategies and solutions for the company. This opportunity has reduced the bridge of understanding

between provider and customer and the potential experience of users. And the same benefit out solution will provide to the organisation and their members.

- Leap-In will gain a competitive advantage by automating the time consuming and complex data structures into streamlines predictive platforms, which can help them increase their member's budget utilisation strategy and, internally, much corporate decision making.
- The solution will showcase predictions and suggest targeted actionable suggestions based on forecasts for members to look for services that might be useful to them.

2. Project Scope

2.1. Functional Requirements:

1. **Identify and group members by under or overspent vs spending on track based on past data :**

   A correlation matrix table will be created to show the coefficients between variables that represent the users' spending so that a clustering algorithm will interpret the data and segment users by these variables. They aim to summarize users' spending data by correlation and segment users by under and overspend groups from their past spends.

2. **Identify potential factors that affect under or overspend :**

   Descriptive analytics is often used to interpret historical data and understand changes and shows its results by visualizations such as pie charts, bar charts, line graphs, tables (S. Song et al., 2013). In this project, descriptive analytics in Python will be used to interpret users' spending to draw a comparison between under or overspend that will outcome graphs that show potential factors that took a user to under or overspent his budget.

3. **Identify the under or overspend makers (Where the value/money was spent):**

   Exploratory data analytic is an analyse tactic for analysing and investigating data sets, summarising their key characteristics using data visualisation methods as needed (IBM Cloud Education, 2021). This approach will assist in terms of detecting anomalies, validating assumptions, and knowing the effect of the factors. In this project, an exploratory data analytic report will be published to identify markers of over or under spending and members spending on the track.

4. **Early (What it means in this case) identify members who are at risk in utilising their budgets:**

   Machine learning algorithms provide the ability to automatically learn from experience by training and test models (Burrell, J., 2016). In this project a machine learning model will be trained to identify members who are at risk in utilising their budget so that guidance can be provided to those to avoid under or overspend.

2.2. Non-Functional Requirements:

1. **Cloud service:**
   Cloud computing provides computer research such as databases over the internet. Cloud load balancing is a process of distributing workloads and computing resources in a cloud computing environment. It split off allocating resources among multiple computers which helps enterprises achieve high-performance levels for potentially lower costs as one of the payment forms is by time response (Kansal, N. & Chana, I., 2012). This project will deal with cloud databases and also load balancing based in response time charging. What will provide an efficient data storage with high performance and low cost compared to traditional storage methods.

2. **Encryption data:**
   Mathematics might be used to develop algorithms for encryption data purposes. What helps to prevent data fraud or hacker's data understanding. All the data manipulated by the software to be developed will be encrypted to provide a high data security.

2.3. Scope

Scope	**Deliverables**	**Total hours by employees** **Total Hours spent on project = 140 * 5 members = 700 hours**	**Complexity**	**Effort Estimation**
In Scope	1. Data Cleaning and Data Pre-processing: When received data in RDBMS converted into preferred format for python by programmer and software developer.	50 Hours total 25 Hours -> Programmer 25 -> Software Developer	1.5	0.7142
	2. Data Manipulation and Analysis: Apply mathematics to extract patterns and relevant information using python. Done by Data analyst and ML engineer.	60 Hours total 20 -> ML engineer 20 -> Data Analyst 20 -> Business Analyst	2	0.8571

	3. Data Visualisation: Create visualisation dashboard using tableau by business and data analyst using Tableau.	50 Hours total 25 -> Data Analyst 25 -> Business Analyst	1.5	0.7142
	4. Machine Learning Model created in google collaboration using python as a language by ML engineer, programmer and software developer. Data and business analyst help set control and variables.	500 Hours total 150 -> ML engineer 150 -> Programmer 100 -> Software Developer 50 -> Data Analyst 50 -> Business Analyst	4	7.1428
	5. Documentation continuously updated and referred by data and business analysts using google docs.	40 Hours total 8 -> Data Analyst 8 -> Business Analyst 8 -> ML engineer 8 -> Programmer 8 -> Software Developer	1	0.5714
Out Scope	Establishing a front-end UX is transferred to out-scope as Clients already have engineers. If the model is accurate and accepted, the engineers can implement it into the system.	N/A	N/A	N/A
Scope Risks	Individual team members may add features which are not defined or not needed. Requirements may not be changed later or not defined properly. Scope cannot be flexible incases of change control process or change management. How the data will react to different mathematical calculations and machine learning how it predicts is uncertain. It might risk the whole scope.			

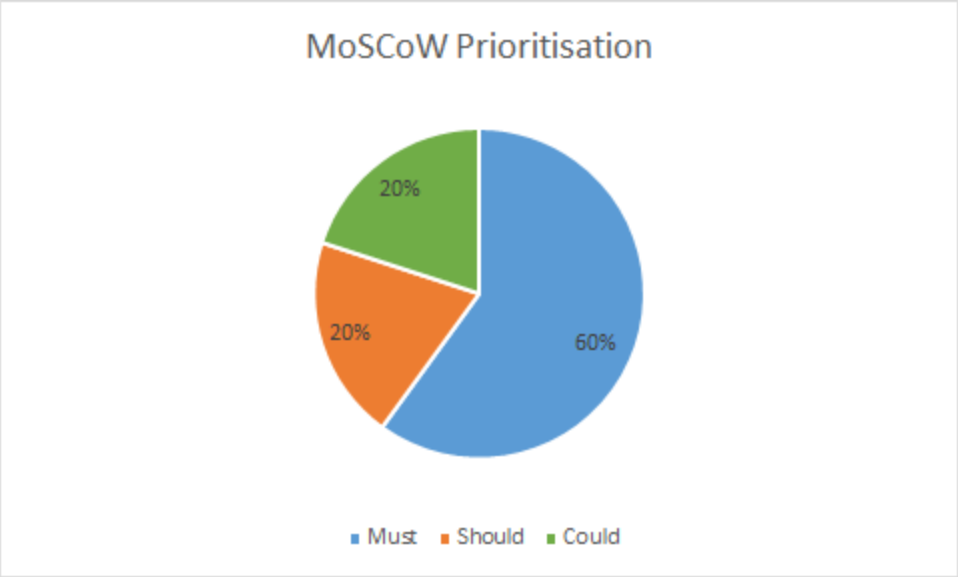
(Table No. 2 Scope Definition, Complexity and Effort estimation)

MoSCoW Prioritisation:

- 60% OR less Must = “Deliver it or we cancel the Project”
- ~ 20% Should = “Important but not vital”
- ~ 20% Could = “ Nice to Have”

Sr. No.	User Stories	Priority	Points	Reasoning
1.	As a Business Sponsor, I would like to have an automated system which can flag my members budget plan before they go off track.	Must	6	ML model is the main deliverable and must be deployed in the given time frame. ML models help predict based on data given for budget utilisation.
2.	As a Business Advisor, I would like a detailed and visual report on past member data. So we can make better strategies in future enhancement lists.	Must	4	A visualised dashboard shows insights from patterns extracted in data mining which help companies internally make decisions.
3.	As a member, I need my data to be secured and encrypted. My personal information must not be disclosed.	Must	2	Secure System is a must for any company who is data-centered. And this facility is important to have.
4.	As a Business Sponsor, I would like to give my members recommendation on various services which might be useful for them to utilise their budget.	Should	4	Recommendation System is a service beneficial to members when they are underspending and can explore services they can use which they didin't know about.
5.	As a member, I would like to rate the services I used so that it can be helpful to others.	Could	2	The Review System will help identify genuine service providers and safe for diabled people.
6.	As a member, I would like an flexible user experience to showcase	Could	2	An ease of access is very important for saving time and efforts of disabled people.

(Table No. 3 MoSCoW Prioritisation)



(Figure No. 1 MoSCoW Prioritisation pie chart)

3. Chosen Project Management Approach:

The project management methodology selected to develop **Budget utilisation** is DSDM. The key reasons for choosing DSDM are:

1. The necessity of the stakeholders to have a collaborative and cooperative attitude with full involvement in the project.
2. It has a solid focal point in motivating the stakeholders' communication and participation in the project development.
3. It provides a short-term learning time through the retrospectives of iteration as the solution is developed in iteration.
4. Its abilities to "dodge" unforeseen circumstances due to the complexity of the variables involved in the project.
5. It allows the project to be testing and integrating throughout the life cycle of the project.

3.1. Justification:

Our project Budget Utilisation is a machine learning project with lots of experiments and different ways to feature models for better accuracy. This requires control over the project by deploying quality solutions in a given timeframe. The project involves client involvement and customer service orientation, which DSDM focuses on as customer collaboration. The project comes with different requirements of automated ML, which keeps on updating and analysing, including a data analytic dashboard to understand data patterns and also a recommendation system based on longitudinal data present for which DSDM is adaptable to change and can handle multiple user requirements.

3.2. How the approach is adopted:

Domain	Cause and Effect	Risk Types	Cost of Change
Complicated	Sense, Analyse and Respond	Known-Unknowns	High

(Table No. 4 Cynefin Framework)

Based on the Cynefin framework, the project falls in the Complicated domain. This is because even though the data is known, and approach and outcomes are known. We never know how data reacts to various mathematical calculations and how models will behave. Hence it has many known and unknown. It requires pre-knowledge in data analytics and machine learning using python. Trello is also used for task management with differentiating between what's done, work in progress and tasks in the backlog.

4. Team's Solution

4.1. Potential Technology solution:

Our solution is an automated machine learning model which will continuously update itself automatically and analyse member data including various variables. The result we receive will predict not a member is underspending/ overspending but is about to go off track. This will help in better utilisation of the member's budget.

- But before creating an ML model, the data scientists prefer a structured lifecycle to enhance performance and longevity of machine learning models.
- Therefor, We first analyse the data, understand how data will be useful in predicting and later strategize ML model based on insights
- The resources are online and can be worked from home considering post-pandemic situations.
- We must consider security implications because the scope is data-oriented and which must be protected. It consists of the member's personal information, diability information, services used and budget spent information.
- We must also differentiate the results of the ML model into the factors that affected the member to go off-track. This helps the report system create better utilisation plans for members.
- This can be understood by the Netflix recommendation system use-case which explores user activity, location and personal information to extract relevant patterns that help recommend future personalised experience on the system.
- If time constraint allows, a review system for members is must which can help provide genuine service providers for them and understand their needs. This can be implicated in the recommendation system.

4.2. Justify solution and technology choices:

We will follow the data analytics cycle and machine learning cycle to fulfill the requirements stated in scope. The cycles are shown in the figures below:

There are two phases of these project:

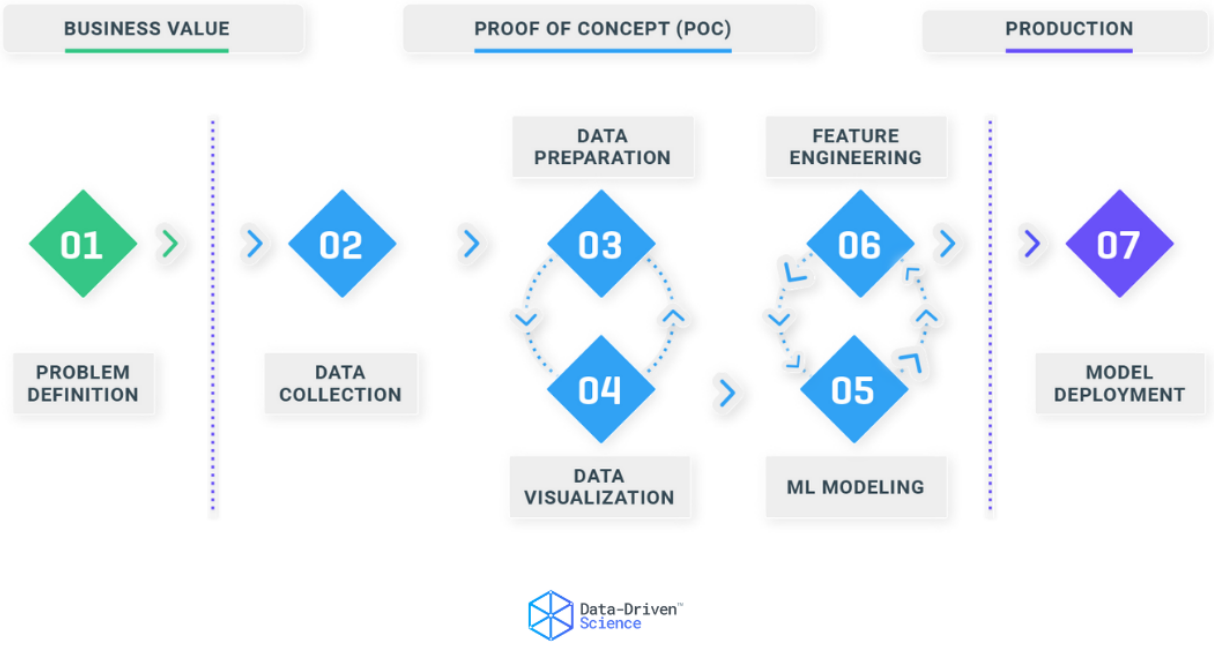
1. **Data Analysis:**

(Figure No. 2 Data Analysis Cycle)

- Here, first when we receive data in RDBMS (relational database) accessed in MySQL. We need to convert the data which can be readed in a python file to further pre-process it.
- Later, data is converted and taken as input in Google Collab.We use python language to clean it, impute the missing values and preprocess it. Here, we need to know that missing values are empty because the member is not spending and understand how we will create meaningful features from it. We might need to normalise data because the data is huge and wide with multiple variables.
- Then, we will perform statistical analysis on the refined data to understand the data and correlation between variables. Understand factors affected in changes and relate the variance of it.

- Later a dashboard is provided to Leap In organisation which includes results observed from their past data. This will help companies internally strategise even better for future endeavours. For dashboard, we use Tableau public source as per client's request.

2. Data Analytics:



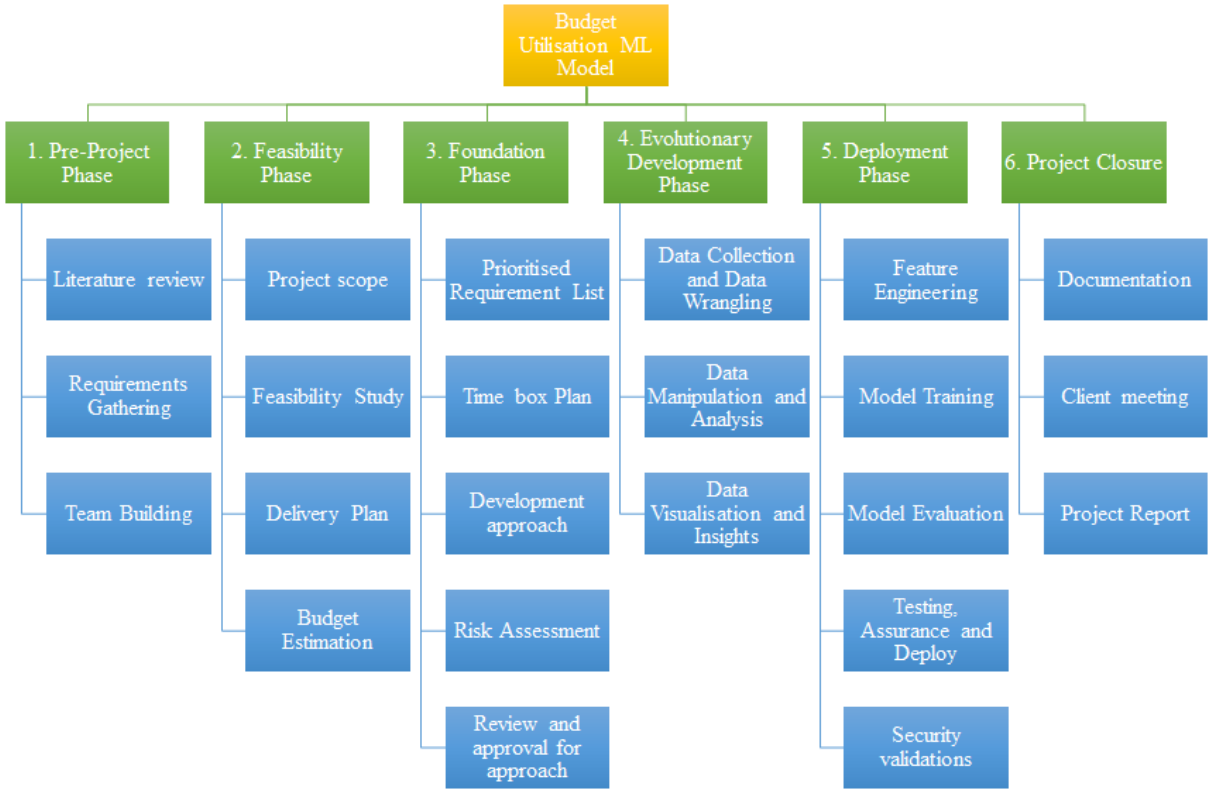
(Figure No. 3 Data Science cycle = Analysis + ML)

- In data analytics phase, we will work on the actual model which is the final deliverable.
- Based on the studied data in phase one; we can extract data and their variables can help in the feature engineering step. In feature engineering, we will extract important features using mathematical equations and generate feature algorithms.
- Based on these algorithms, the model is trained using python language. The whole project works on python using Google Collab because it provides facility of using GPU to increase runtime performance with these huge set of data
- We choose a logistic regression model because at the end the machine learning model predicts the probability of members going off track based on multi-variants. Logistic regression will have to be customised based on needs.
- The accuracy technique will be selected in a data analysis phase that suits and increases the ML positive predictability. Our goal is to achieve at least 90% accuracy at the end of the project in predicting the probability.

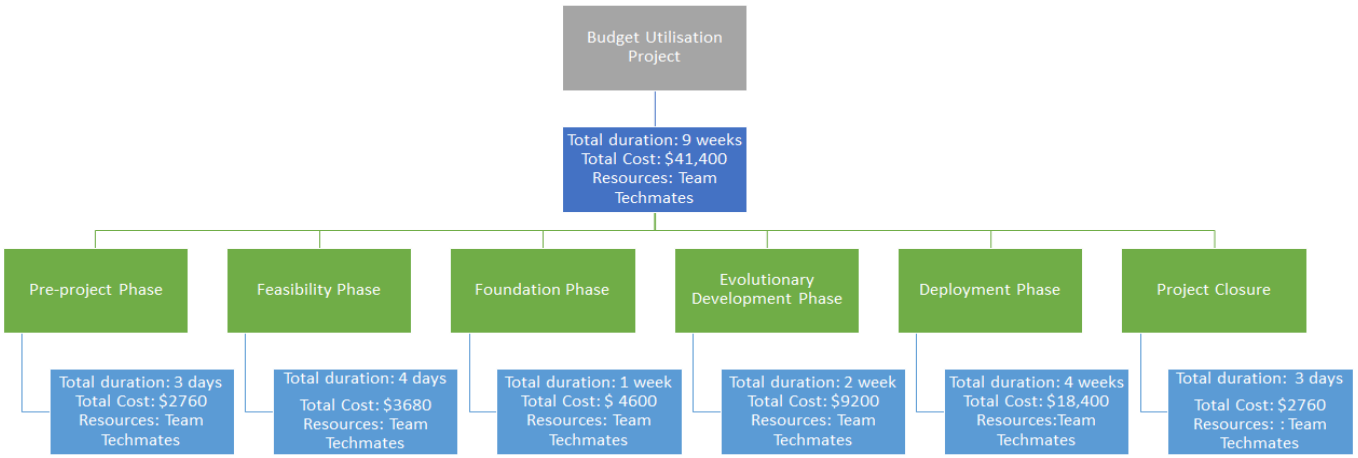
V. Schedule of Intermediate and Final Deliverables

5.1. Work Breakdown Structure

The work breakdown structure is as per follows:



(Figure No. 4 Work Breakdown Structure)

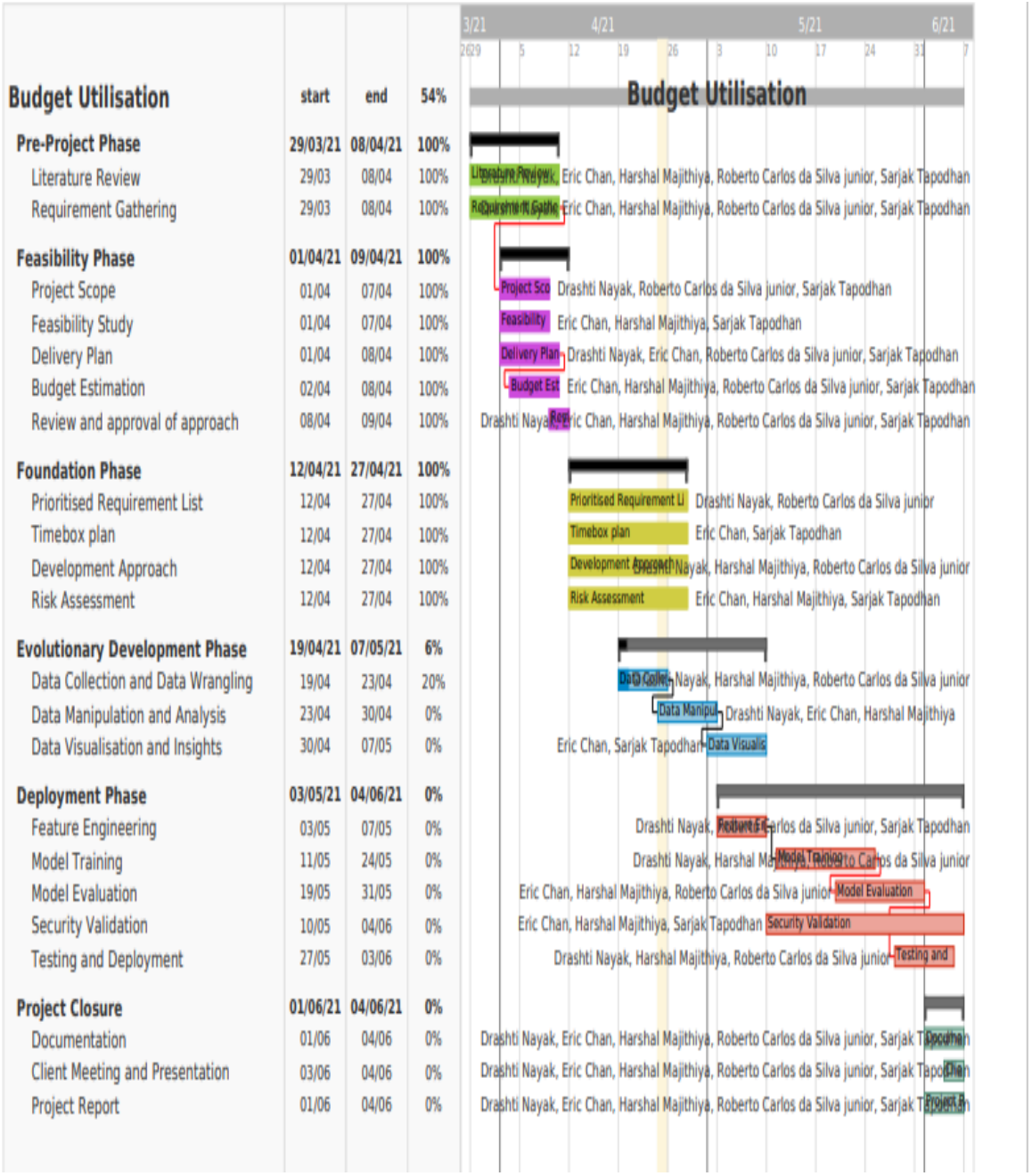


(Figure No. 5 Resource, time and cost allocation to phases)

The cost is calculated based on human resources and services like internet, electricity and tableau.

5.2. Gantt chart

Gantt chart depicts the timeline of our project with tasks, dates and team as resources.



(Figure No. 6 Gantt chart with team as resource)

VI. Manage Communications and Risks

6.1. RACI matrix & Communication Plan

RACI Matrix

IFN711 Team TechMates

	Deliverable/ Task	Description	Communication method	Frequency	Business Sponsor (Jane Sheedy)	Business Advisor (Anton Lord)	Business Analyst (Sarjak Tapodhan)	Data Analyst (Eric Chan)	Solution Developer (Drashti and Harshal)	Quality Assurance/ Tester (Roberto)	Project Manager (Drashti)	Project Supervisor (Venket Venkatachalam)
			ROLES		Client Team		Business Team		Technical Team		Management	
Occur in every phase	Daily Updates	Daily updates will happen in every phase and include what is done, what is going to be done in the next activity, whether there is any problem	WhatsApp group	Daily	X	X	I	I	I	I	A	X
PRE-PROJECT PHASE												
	Market Research	Similar to literature review	Online	Once - at the commence of the project	X	X	R	R	R	R	R/A	X
	Rules and Regulations	Understand terms and Conditions regarding the client and data policies	Zoom meeting		C	C	R	I	I	I	A	C
	Requirement Gathering and Classification	Consult with Business sponsor and advisor regarding the requirements of their workflow and mangement	Zoom meeting		C	C	R	I	I	I	A	C
FEASIBILITY PHASE												
	Feasibility Check from both Technical and Business sides	Establish relevant solution with both team	Face to face	Weekly	I	I	C	C	C	C	R/A	I
	Project Outline	Project brief to all team members	Face to Face	Fortnightly	I	I	C	C	C	C	R/A	I
	Project Cost Report	Cost estimation report of project	Face to Face Email	Weekly	I	I	C	C	C	C	R/A	I
	Delivery Plan	Includes sources and strategies in delievry plan	Face to Face Zoom meeting	Monthly	I	I	C	C	C	C	R/A	I
FOUNDATIONS PHASE												
	Solution Architecture Definition	showcase the high level design framework of the solution	Face to Face	Once	X	X	C	C	R	R	A	I
	Development Approach Definition	showcase high level definition of the tools, techniques, practices, standards and technical approaches	Face to Face	Once	X	X	C	C	R	R	A	I
	Priortised Requirement List	MoSCoW Prioritisation list	Face to Face Online	Monthly	X	X	R	C	C	C	R/A	I
	Management Approach Definition	showcase the approaches for management using trello	Online	Once	X	X	R	C	I	I	A	I
	Delivery Plan	Showcase the high level delivery plan of Pr	Face to Face	Monthly	I	X	C	C	C	C	R/A	I
	Timebox Plan	Timebox showcase tasks "done" till now	Face to FaceTrelloVideo	Weekly	X	X	R	I	C	I	A/R	X
	Timebox Review Record	access the quality and risks related	Face to FaceTrelloVideo	Weekly	X	X	R	I	C	I	A/R	X
EVOLUTIONARY DEVELOPMENT PHASE												
	Data collection and Data	Data will be provided in RDBMS which	Google CollabZoom M	Once	X	X	I	C	R	R	A	X
	Analysis	Exploratory Data Analysis	Google CollabZoom M	e (later if require	X	X	C	R	R	I	A	X
	Data visualisation and	A visual dashboard to study past busi	Tableau online	Once	X	X	R	R	C	C	A	X
	Workshop Facilitated	Provide the direct and indirect benefits to	Face to FaceVideo Confe	Weekly	X	X	I	I	I	I	R/A	X
DEPLOYMENT PHASE												
The product created in this phase is the result of the following iterative procress: **Assemble - Review - Deploy**	Feature Engineering	Feature selection process for models based on filtered data	Google CollabZoom M	Daily	X	X	C	C	R	C	A	X
	Model Training	Model algorithm, parameter tuning an	Google CollabZoom M	Daily	X	X	X	C	R	R	A	X
	Model Evaluation	Validate developed model	Google CollabZoom M	Daily	X	X	X	I	R	R	A	X
	Technical Testing & Assurance	Test the user interface and functionality of the product.	Face-to-face	Last two weeks	X	X	I	I	R	R	A	I
	Model Deployment	Deploy final model	Conference Call Online	Once	I	I	I	I	C	R	R/A	I
POST PROJECT PHASE												
	Maintanance	Post deployment upgrades	Face to Face (onsite)Vid	Every 6 Months	C	C	C	C	R	R	A	I
	Benefit Assessment	Realize the benefits in relation to the busi	Face to FaceTemplate	If required	X	X	R	R	I	I	A	I
	Support	Resolve the Bugs and Fixes.	Face to Face (onsite)Vid	If requested	C	C	I	C	R	R	A	I

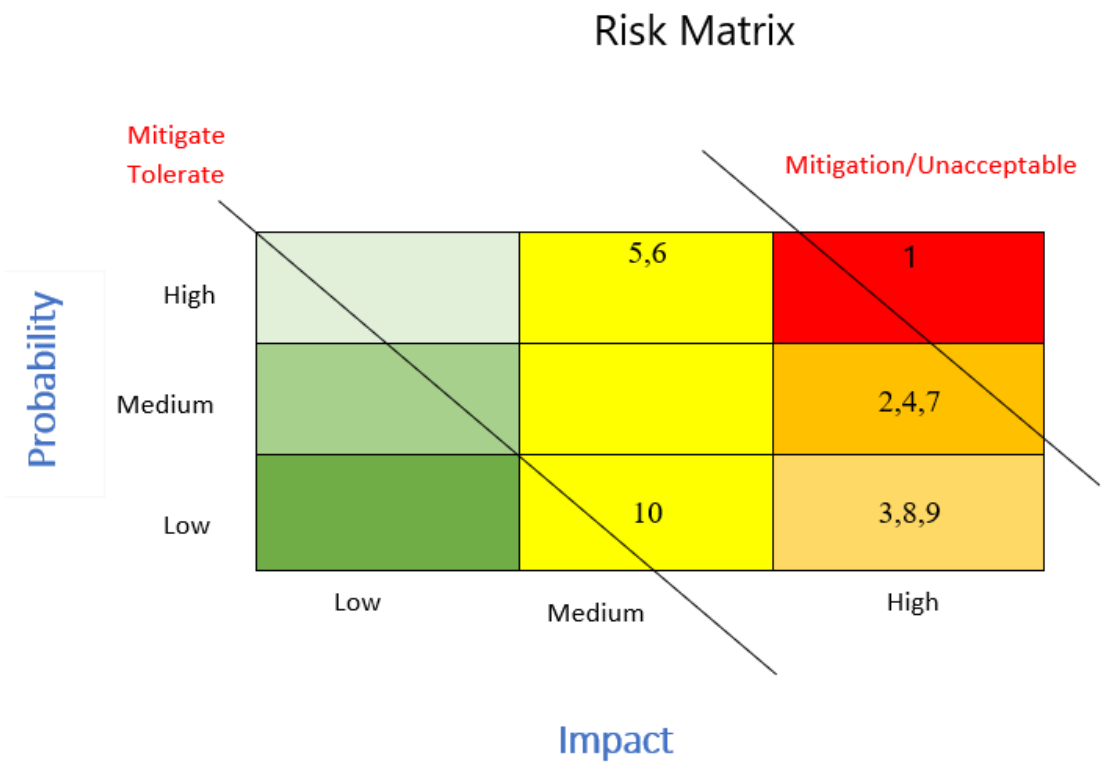
R	Responsible	Assigned to complete the task or deliverable.
A	Accountable	Has final authority and accountability of completed tasks and decison making.
C	Consulted	An adviser, client or subject matter expert who is consulted before a decision or action.
I	Informed	Must be informed after a decision or action.
X	Not Applicable	

(Figure No. 7 RACI matric including Communication Plan)

6.2. Risk Analysis Matrix

Risk Identi-fier	Author	Category	Risk Description	Probability	Impact	Proximity	Response Action
1	Project Manager	Organisational	Team behaviour/ Team member's absence	High	High	Implementation Stage	Outsourcing of required work during their absence.
2	Business Analyst	Technological	Hardware/Software Malfunction or Failure.	Med	High	Within project	Ensure to cover the cost of replacement and repair of hardware.
3	Operations Director	Technological	Server Failure	Low	High	Deployment stage	Transfer to a trustworthy cloud server (Amazon)
4	Project Manager	Organisational	Task not completed in allocated time. Or change in scope	Med	High	Implementation stage	Track the process of the project through Gantt Chart. And refine tasks that need to be done.
5	Quality Assurance Tester	Technological	Slow systems due to more concurrent users.	High	Med	Deployment Stage	Estimate the highest number of concurrent users based on historical data and update the server system.
6	Solution Developer	Technology	Unstructured /Missing data due to manual input	High	Med	Implementation Stage	Gather all the data and structure it accordingly in the database.
7	Project Manager	Organisational	Scope Creep	Med	High	Initiation Stage	Consult with the stakeholders on feasibility of the project.
8	Business Sponsor	Financial	Exceeding the estimated fund for completing the project plan	Low	High	Planning Stage	Structured timeline and budget allocation will be planned upfront to complete the project within specified criteria.
9	Operations Director	Organisational /Technological	System failure/database error.	Low	High	Deployment Stage	Test the unit code with various behavioural testing and reduce the risk.
10	Solution Developer	Technological	The ML model is malfunctioning.	Low	Med	Design Phase	Testing the design with a bunch of users and signing the user acceptance Test.

(Table no. 5 Risk Register)



(Figure No. 8 Risk Matrix)

Mitigation Strategy: According to the Risk matrix, staff/absence or irresponsible staff behaviour is highest at risk because without man work the project can't move further. In case, training can be regarding soft skills and boost their attitude towards projects and clients. Outsourcing can be kept as a last option.

VII. Cost Estimation:

7.1Cost estimation:

Sr. No.	Resource	Hrs./week	No. of Weeks	Price/Week	Quantity	Sub-total	Total
Planning and Documentation							
1.	Stationary	N/A	N/A	As required	5	$100	$150
2.	Printing	N/A	N/A	As required	N/A	$50	
Development tools required							
3.	Laptop	N/A	N/A	$1200	5	$6000	$8800
4.	Cloud Storage	N/A	N/A	$1000	1	$1000	
5.	Server	N/A	N/A	$1800	1	$1800	
Services							
6.	Internet	20	9	$15	5	$135	$500
7.	Tableau	20	2	$115 (One time)	5	$230	
8.	Electricity	20	9	$15	5	$135	
Human Resource							
9.	Drashti (ML Engineer)	20	9	$55/hour	N/A	$ 9900	$41,400
10.	Eric (Data Analyst)	20	9	$40/hour	N/A	$7200	
11.	Harshal (Web Developer)	20	9	$45/hour	N/A	$8100	
12.	Roberto (Programmer)	20	9	$50/hour	N/A	$9000	
13.	Sarjak (Business Analyst)	20	9	$40/hour	N/A	$7200	
Total							$50,850

(Table No. 6 Cost Estimation)

So, the total cost estimated for the project including resources and the salary of the team is approximated to be $50,000.

Travel Expense:
As we reside in Brisbane, travel expenses are set on employees, however travel expenses may be needed in case of hiring an external source. It will be included in the contingency budget.

7.2 Contingency Plan

We have estimated technical resources and use of cloud service for high-quality run-time and handle maximum traffic. In case, the resources fail, backup of server or help of cloud service and GPU can be taken to balance load.

Risk	Probability	Cost	Contingency
Hardware/Software/Server Failure	50%	$3500	It can be replaced, repaired or transferred to cloud services.
Outsourcing	60%	$7500	Hire an employee for a short term in urgency.
Training	60%	$2000	Train existing employees with new skills.
Power failure	10%	$500	A generator in case of emergency
Data storage	80%	Only pay for what you use $900/monthly	Cloud services used.
Data breach/ Cyber attack	70%	Millions	A cyber-attack or data breach can cost millions. Hence, secure authentication is must.

(Table No. 7 Contingency budget Plan)

VIII. Team Controls and Team Contract

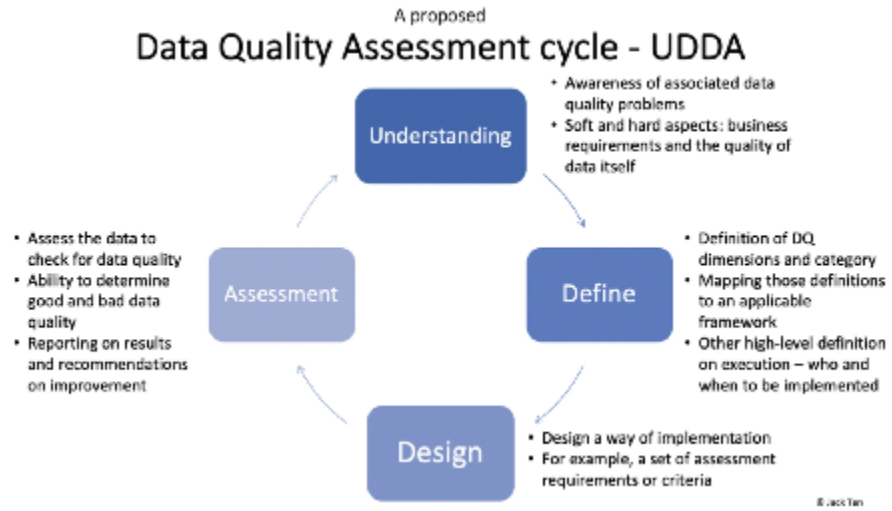
8.1. Project Controls:

Controls implemented to monitor and execute the project ensures that the project is on track. The controls considered in the table below have been considered around four project variables scope, time, cost and features.

Variable	Techniques to manage the variables

Scope	Open communication, proper documentation, ensure design options, monitoring and analyzing change
Time	Planning work ahead , set priorities per member, summarise each day with team and minimising interruptions
Cost	Cost estimation. Record actual costs, achieve baseline budget
Features	Features will be managed through DSDM timeboxing plan

(Table No.8 Project Control trade-offs)



(Figure No. 9 Project Control Cycle)

How we will ensure the quality of artefact: Choosing the right quantity data for the right algorithm will increase performance of the final model. A data size evaluation and heuristic rule carried out in the evolutionary development phase will help determine quantity of dataset.Quality of training data sets can be done with help of quality-assurance methods to unget unbiased results. A validation technique will help understand overall performance of ML models.

Tracking and reporting the progress: The progress of the report will be tracked with help of Gantt chart and Trello Dashboard. Weekly meetings with the tutor on Friday at 3.30 pm will give us insights about our work done. We have also established weekly meetings with the client, for analysing requirements.

8.2. Team Contract

For the team, contract refer to Appendix B

IX. References

- Anolytics. (2020). *How To Improve Machine Learning Model Performance: Five Ways.* https://www.anolytics.ai/blog/how-to-improve-machine-learning-model-performance/
- Baader, F., & Sattler, U, (2001). An overview of tableau algorithms for description logics *Studia Logica, 69*(1), 5-40
- Bisen, S, V. (2019). *How To Ensure Data Quality For Machine Learning And AI Projects.* Medium. https://medium.com/vsinghbisen/how-to-ensure-data-quality-for-machine-learning-and-ai-projects-c8af1fe18c57
- Burrell, J. (2016). *How the machine 'thinks': Understanding opacity in machine learning algorithms*
- Big Data & Society. https://doi.org/10.1177/2053951715622512
- Centric Consulting. (2021, March 26). Machine Learning: A Quick Introduction and Five Core Steps. Retrieved from https://centricconsulting.com/blog/machine-learning-a-quick-introduction-and-five-core-steps//
- Chabot, C., Stolte, C., & Hanrahan, P., (2003). *Tableau software* Tableau Software, 6
- Dhavel.M. (2018). *How to perform Quality Assurance for Machine Learning models.* Medium.https://medium.datadriveninvestor.com/how-to-perform-quality-assurance-for-ml-models-cef77bbbcfb
- Frameworks for Approaching the Machine Learning Process. (n.d.). Retrieved April 25, 2021, from https://www.kdnuggets.com/2018/05/general-approaches-machine-learning-process.html
- Hall, H. (2020, February 12). How to Identify Scope Risks. Retrieved from https://projectriskcoach.com/how-to-identify-scope-risks/#:%7E:text=Scope%20risks%20are%20uncertain%20events,related%20to%20the%20project%20scope.
- IBM Cloud Education. (2021). *Exploratory Data Analysis.* Retrieved from https://www.ibm.com/cloud/learn/exploratory-data-analysis
- Leap in. (2021). *Top 5 reasons why people don't spend their NDIS budgets.* https://www.leapin.com.au/5-reasons-ndis-budgets-unspent/
- **K**ansal, N. J., & Chana, I., (2012). *Cloud load balancing techniques: A step towards green computing*
- IJCSI International Journal of Computer Science Issues, 9(1), 238-246.
- Kurek, E., Johnson, J., & Mulder, H. (2017). Measuring the value of enterprise architecture on IT projects with CHAOS research Syst. *Cybern Inform, 15*(7),13–18
- Messenger, S. (2014). *DSDM Agile Project Framework Handbook.* The Agile Business Consortium. https://www.agilebusiness.org/page/ProjectFramework_00_welcome

- National Disability Insurance Agency, (2021). *What is the NDIS?* https://www.ndis.gov.au/understanding/what-ndis
- NDSP Plan Managers, (2020). *Are You Using Your Plan Funds?* https://ndsp.com.au/are-you-using-your-plan-funds
- Public Services and Procurement Canada. (2019). *Scope Management Techniques*.https://www.tpsgc-pwgsc.gc.ca/biens-property/sngp-npms/bi-rp/conn-know/portee-scope/techniques-techniques-eng.html
- Rahm, E., & Do, H. H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull., 23*(4), 3-13.
- S. Song et al., (2013). Prescriptive Analytics System for Improving Research Power. *2013 IEEE 16th International Conference on Computational Science and Engineering, Sydney, NSW, Australia,* (2013, pp. 1144-1145), doi:10.1109/CSE.2013.169.
- Science, D. (2020, July 20). 7 Stages of Machine Learning — A Framework - Data-Driven Science. Retrieved from https://medium.com/@datadrivenscience/7-stages-of-machine-learning-a-framework-33d39065e2c9
- Snowden, J, D., & Boone , E, M. (2007). *A Leader's Framework for Decision Making*. Harvard Business Review https://hbr.org/2007/11/a-leaders-framework-for-decision-making
- Tan.J. (2020). *How to improve data quality for machine learning?* Towards Data Science https://towardsdatascience.com/how-to-improve-data-preparation-for-machine-learning-dde107b60091
- Templatelab. (2021). *40 Detailed Contingency Plan Examples (& Free Templates)*. https://templatelab.com/contingency-plans/
- TensorFlow. (2021). *Welcome To Colaboratory.* https://colab.research.google.com/notebooks/intro.ipynb

Appendix A: Glossary and Abbreviations

Machine Learning (ML): Machine learning is a subcategory of AI which is a model that learns through experience with large sets of data using algorithms.
Deep Learning: Deep learning is a subcategory of ML which is inspired by how our brains work. In short it connects the dots to get relevant information.
Data Analysis: Data analysis is a study of past or historical data to get insights and relevant information regarding past environments and how data has emerged.
Data Analytics: Data analytics is a study of present data to get predictions of the future based on fact-based data.
Data Mining: Data mining is a process of extracting patterns from data by manipulating and applying statistics on it.
Data Visualisation: Data visualisation is a process of representing the patterns and information from data mining. So, that it can be understood by business teams.

Logistic Regression: Logistic regression for multi-class can determine for probability for a scenario based on multi-variables.
Exploratory analysis: this analysis used to summarise the results of statistical techniques applied on datasets.
Prescriptive analysis: Prescriptive analysis is finding the best solution for a given problem with available data. It is a combination of descriptive and predictive analytics.
Accuracy: Accuracy defines how precise a machine learning model gives correct outcomes.
GPU: It is a specific processor used to handle large-size data sets and for machine learning.
Tableau: Tableau is a data visualisation tool that helps to create visual dashboards.

Appendix B: Team Contract

Team Contract

IFN711 IT Industry Project - Budget Utilisation

Project Supervisor: Dr Venkat Venkatachalam

Industry Partner: Leap in!
Industry Supervisor: Jane Sheehy (CTO)

Team Techmates

1. Drashti Nayak (n10599568)
2. Harshal Majithiya (n10550453)
3. Roberto Carlos da Silva junior (n10374647)
4. Sarjak Tapodhan (n10553916)
5. Sing Yin Chan (n9317007)

A. Commitment:

As a team, we will:

1. Be practical with planning and reporting important variables like scope, time, costs and features.
2. Become result-driven and take charge whenever a potential problem arises and work to prevent it beforehand or reduce the impact in case faced.
3. Keep all team member informed and notify the client of any slightest of changes.
4. Keep information and data received from client confidential.
5. Focus on what is best and see through a full completion.

B. Participation Rules:

As a team, we will:

1. Solve issues that arise with unbiased solutions and to the minimum.
2. Keep meetings open-minded, diverse, honest and practical.
3. Be flexible to allow each team member the opportunity to present their opinions and participation.
4. Consider New approaches and be open to listening to other team members opinions.
5. Never play blame-game and instead discuss and explore the possible solution when things go wrong.

C. Communication Methods:

As a team, we will:

1. Be clear and precise to understand and make others be understood. For this, the active listening skill must incorporate.

2. Accept different communication methods to communicate through text, phone calls, video or face-to-face with agreed time, location and medium of contact.
3. Keep the communicated ideas on track.

D. Problem Solving:

As a team, we will:

1. Ensure that each team member participates in each task.
2. Strengthen and encourage each other's new ideas rather than criticise. Problem-solving is a creative process.
3. Make use of tools and date whenever the problem arises and need facilitation to solve.

E. Decision Making:

As a team, we will:

1. Make decisions based on facts, data when needed.
2. Carry out research when information needed and when need to define different interpretations of data.
3. Make sure the project variables scope, time, cost and features considered before decision making.
4. Listen to each member's view before taking a decision.
5. Never communicate concerns with a third party without the knowledge of others in inappropriate ways.

F. Handle Conflicts:

As a team, we will:

1. Consider conflicts that arise as an occasion of growth.
2. Involve each party, and theirs desires and interest considered before reaching a solution.
3. Select preferred location and time to meet and solve conflict arose.
4. Acknowledge other member's input and state our perception without being ill-mannered or judgemental.
5. Overcome conflicts and must be on the same ground of agreement.

G. Meeting Guidelines:

1. Meetings will be held every Thursday at noon.
2. Client meeting will be held every Thursday at 10 am on Zoom.
3. The supervisor meeting will be held every Friday at 3.30 pm on Zoom.
4. Agendas will be issued every Monday at 11 am on WhatsApp group chat. In this meeting, an evaluation of the work done also carried out.
5. Meetings are called by Drashti Nayak (Project Manager).

6. The appointed team member must submit the meeting minutes with the client within two days of the meeting to other team members and supervisors.

H. Meeting Procedures:

1. Each team member must be present in all meetings and must come prepared.
2. In case of urgency, the team member must inform the project manager before one hour of the meeting time.
3. In case a team member is unable to finish a task, he/she must inform 2-3 days before the due date.
4. The team must keep off-topic ideas and concerns aside in different document for future evaluation. A record must include issues faced and was unable to solve.
5. If team members are unable to be present, they must select a member who will be authoritative to decide on behalf.

★ I agree to be participated in including the rules, processes, standards and timelines in this document.
★ I understand the terms and conditions and be obligated to abide by them.
★ I understand if I do not abide by the terms and conditions, I will be obliged to suffer the consequences given by the unit authorities.

Signatures

1. Drashti Kanubhai Nayak
   Date: 17 April 2021
   Sign:
   
2. Harshal Majithiya
   Date: 17 April 2021
   Sign:
   
3. Roberto Carlos da Silva junior
   Date: 17 April 2021
   Sign:

4. Sarjak Tapodhan
   Date: 17 April 2021
   Sign:
   
5. Sing Yin Chan
   Date: 17 April 2021
   Sign: