Prom02 – Appendices

Student Name: Yue Fung (Lewis) Student Number: 199124958 Programme: MSc Data Science

1. Project Achievements, Challenges and Impacts

1.1 Achievements

After all, Ambros instructed me to prepare such research, and their main goal might be to develop an option which relied on professional services which could be less depend on machine learning result and build a data product. Hence, I have three critical points that I should communicate my success in this research to drive the new product – Machine Learning as SaaS instead of developing a professional service team.

Achievement	ievement Significance		Evidence		
The result provided a great	According to the practice of this	-F&B domain knowledge	Brown and Mitchell, 1993		
machine learning model for	research, it may become F&B	-Excel, Python and Data cleaning	Ayyadevara, 2018		
customer who needs to	industry knowledge for	-Data analysis	Kim and Gu, 2006		
understand the pattern of	restaurants to source a new	-ML algorithm selection	Refer to Chapter 4		
opening a new restaurant.	location.	-Problem solving skills			
Better prediction of restaurant	For restaurant operators who	-Feature engineering	Tanizaki et al., 2019		
sales and customer service based	need to plan forward for future	-ML algorithm selection and	<u>Lasek et al., 2016a</u>		
upon earlier customer behaviour	expansion or cost savings,	cross validation	Refer to Chapter 3		
data.	restaurant sales forecasts are a	-Data Modelling			
	critical consideration factor.	-Time management			
Ambros will offer restaurant	It enables Ambros to be more	-Algorithms & Optimization	Mathur, 2018,		
owner truly personalized services	competitive in the industry by	-Python programming	<u>Lian et al., 2017b</u> , <u>Lian et al., 2017a</u> ,		
recommendations	adding additional value to their	-Data structures	Refer to Chapter 2 and 3		
	existing SaaS product.	-Problem solving skills			

1.2 Challenges

Challenge	Explanation	Approach & Solutions	Evidence
Data Collection	Find data, consolidate those data from Ambros, Government and Property agency from web scraping spent a lot of time which over 50%-60% of the research. Once the data is collected, validation of those data quantity is sufficient also.	-Data cleaning with excel and Python tools -Dimensionality reduction, remove some variable which may not affect the resultInstance selection, it aims to reduce a dataset's height. Diverse and include random selection, genetic algorithm based selection, progressive sampling, using domain knowledge and cluster sampling.	L'heureux et al., 2017 Refer to Chapter 2.2
Selection of Machine Learning Models	Model selection is the process of evaluating or assessing candidate models in order to select the best one. Whereas, once a model has been selected, it can be evaluated in order to communicate how well it is expected to perform in general.	-A model that satisfies the needs and constraints of Ambros expectationSplit the data into training, validation, and test sets, then fit candidate models on the training set, evaluate and select them on the validation set, and report the performance of the final model on the test set.	Kim and Gu, 2006 Refer to Chapter 2.3
Python programming and debugging	When fine tuning the data csv during the training phase. The python code is essentially changing or evolving as the model trains. A lot of parameters or wights in the model are changing in every step during training. It shops changing once training is complete, at which pint what bugs I didn't catch during the training are part of the model.	I must be able to capture model and optimizer specific information during training. Monitor and react to changes in the capture data with assertion rules that get triggered when a condition is met, eg. over fitting.	Mathur, 2018 Refer to Chapter 3

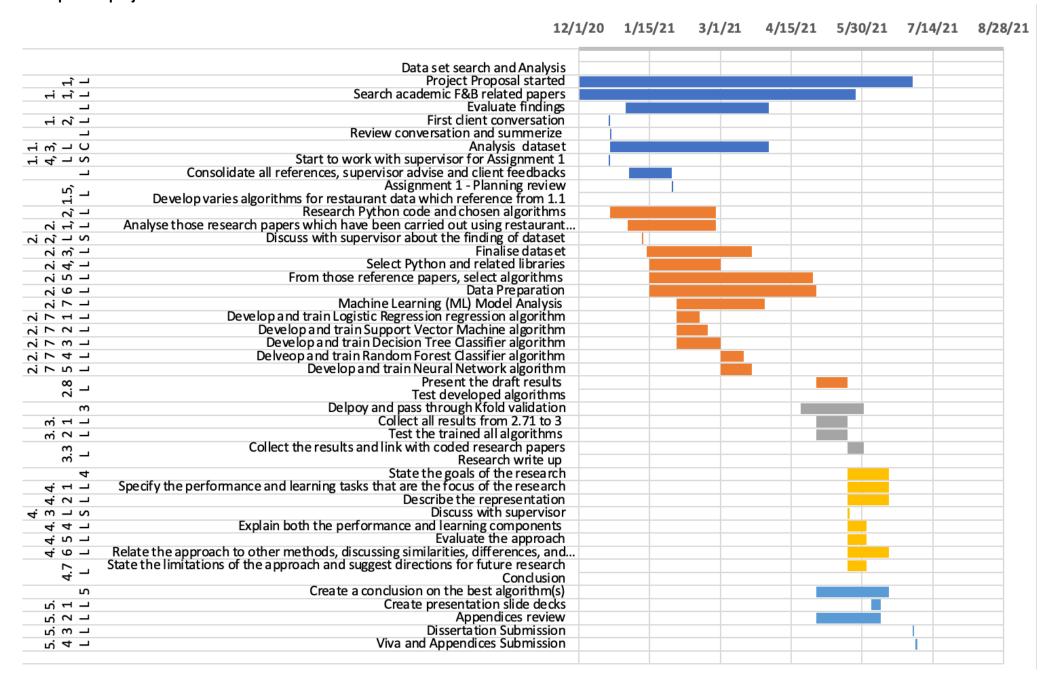
1.3 The Project/Research - Advancing Knowledge and Skills

The impact of this research on me are that best practice flow is used to calculate the restaurant status and predict the sales forecast of a restaurant using machine learning regression models. These studies can also include a literature review and demonstrate the whole technology for achieving the best results for certain industries, eg. F&B. This journey begins only and every year the field of machine training grows fast. I should stay on top of all related news and learn how to produce the best machine learning algorithms in varies industries.

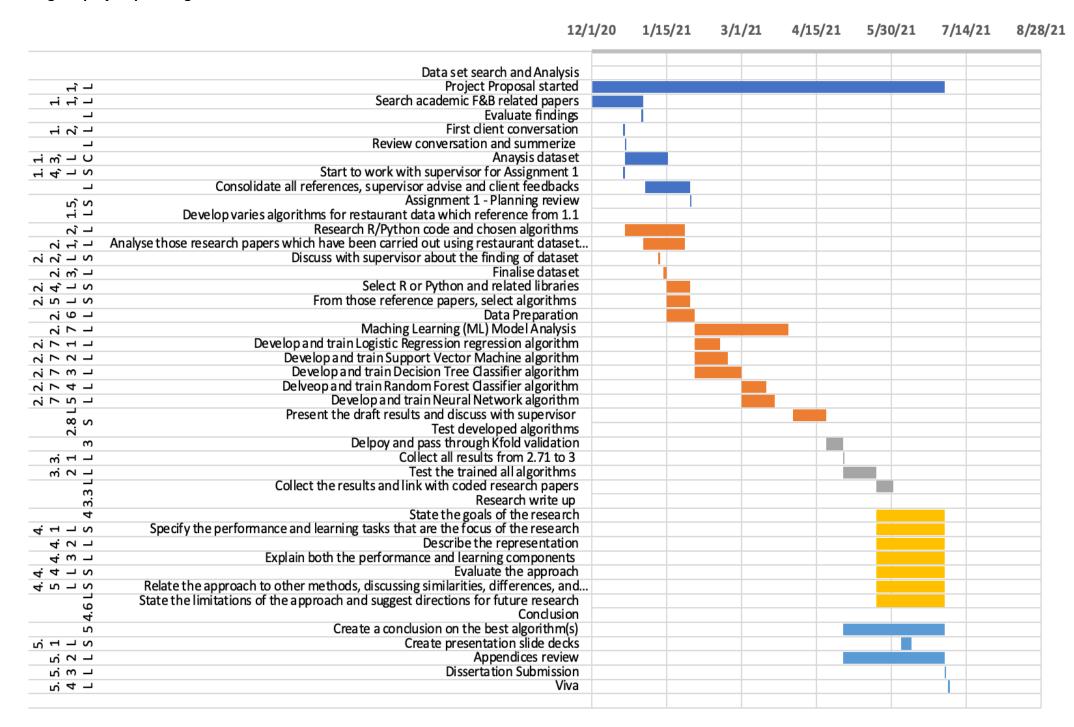
2. Project Management

2.1 Forecast and Actual delivery - Diagram & Explanation

The updated project schedule and deliverable – Final Gantt Chart version 2.0



March Marc		Project Start	12/13/2020	1				columns u	ised to cre	ate the cha	rt			
1.1.1. Project Proposed started 19/100 19/2071 50.0 10 0.0		TASK NAME	START	END	HOURS	DELIVERABLE	COLOR	Start		Red				Purple
1.1. Search academic PAB related papers 12/12/0 75/25/1 75/0	1.1		12/1/20	6/30/21	513.0	Prom02 research	Blue	12/1/20	212	0	0	0	0	0
1,2,1, Proceeds						Select usable papers								
1,2, Fast client conversation 1202/020 1202/020 3.0 0.	L	Evaluate findings	12/31/20	3/31/21	30.0		Blue	12/31/20	91	0	0	0	0	0
L. Review conversation and numerative	1.2, L					Get the dataset from						0		
15.1. C Averlyone of the control of	L	Review conversation and summerize	12/21/20	12/21/20	3.0	List out the discussion key	Blue	12/21/20	1	0	0	0	0	0
L. Coreadated at references, supervisor advisor and claim feedbacks 1,5,1. Assignment 1 - Planning review 1,1221 1 12221 2 2 0 0 0 0 0 0 0 0 0 0 0 0	1.3, L C	Analysis dataset	12/21/20	3/31/21	4.0	attributes of acquired	Blue	12/21/20	101	0	o	0	o	0
1.5, L. Assignment 1 - Remang review 1/28/21 1/28/21 2.00 Devolop varies algorithms for restaurant data which reference (wm to 1.7) 1.00 1.0	1.4, L S	Start to work with supervisor for Assignment 1	12/20/20	12/20/20	2.5	Prom02 research	Blue	12/20/20	1	0	0	0	0	0
Second process against 1- instant of data which reference 1/20/21 2.06 1/20/21 2.06 2.06 2.07	L	Consolidate all references, supervisor advise and client feedbacks	1/2/21	1/28/21	20.0	Prom02 research	Blue	1/2/21	27	0	0	0	0	0
From 1.1	1.5, L	Assignment 1 - Planning review	1/29/21	1/29/21	20.0		Blue	1/29/21	1	0	0	0	0	0
2.1. L Notable in Professional and enclosers algorithms in a least algorithms and enclosers algorithms are enclosed finally algorithms. And enclose the final enclosers algorithms are enclosed finally algorithms. And enclose the final enclosers algorithms are enclosed finally algorithms. And enclose the final enclosers algorithms are enclosed finally algorithms and enclosers algorithms. And enclose the final enclosers algorithms are enclosed finally algorithms. And enclose the final enclosers algorithms. And enclose the final enclosers algorithms are enclosed finally algorithms. And enclose the final enclosers algorithms. And enclose the final enclosers algorithms algorithms. And enclose the final enclosers algorithms are enclosed final enclosers. And enclose the final enclosers algorithms algorithms. And enclose the final enclosers algorithms are enclosed final enclosers. And enclose the final enclosers algorithms algorithms. And enclose the final enclosers algorithms. And enclose the final enclose the final enclose the final enclosers algorithms. And enclose the final enclose the fina														
Analyse Proce research papers with have been carried out using restaurant obtained in different adoptimes 11/1021 1/1021	2, L	Research Python code and chosen algorithms	12/21/20	2/25/21	20.0		Red	12/21/20	0	67	0	0	0	0
2.3. L Select Python and related strarties	2.1, L		1/1/21	2/25/21	5.0	Assignment 3-	Red	1/1/21	0	56	0	0	0	0
2.4, L Salect Python and related Bibraries 1/15/21 2/26/21 5.0 Confirm number of 2.26 L Salect Python and related Bibraries 1/15/21 2/26/21 8.0 Confirm number of 2.26 L Salect Python and related Bibraries 1/15/21 4/26/21 8.0 Confirm number of 2.26 L Salect Lands M. Salect Market Salect Sa	2.2, L S	Discuss with supervisor about the finding of dataset	1/10/21	1/10/21	1.0		Red	1/10/21	0	1	0	0	0	0
2.5 L From those reference papers, select algorithms	2.3, L	Finalise dataset	1/13/21	3/20/21	5.0		Red	1/13/21	0	67	0	0	0	0
2.5 L From those reference papers, select signorithms 1/15/21 4/28/21 8.0 M. algorithms to be Red 1/15/21 0 104 0 0 0 0 2.6 L Date Preparation 1/15/21 4/30/21 15.0 Clean-deatest Red 1/15/21 0 106 0 0 0 0 2.7 L Mochine Learning (ML) Model Analysis 2/1/21 3/28/21 15.0 Select usable ML Red 2/1/21 0 15 0 0 0 0 2.7 L Develop and train Logistic Regression regression algorithm 2/1/21 2/29/21 8.0 Select usable ML Red 2/1/21 0 15 0 0 0 0 2.7 L Develop and train Support Vector Machine signrithm 2/1/21 2/20/21 8.0 Select usable ML Red 2/1/21 0 20 0 0 0 0 2.7 L Develop and train Support Vector Machine signrithm 2/1/21 2/28/21 8.0 Select usable ML Red 2/1/21 0 20 0 0 0 0 2.7 L Develop and train Rendom Forest Classifier algorithm 3/1/21 3/26/21 8.0 Select usable ML Red 2/1/21 0 20 0 0 0 0 2.7 L Develop and train Neural Network algorithm 3/1/21 3/26/21 8.0 Select usable ML Red 3/1/21 0 20 0 0 0 0 2.7 L Develop and train Neural Network algorithm 3/1/21 3/26/21 10.0 Select usable ML Red 3/1/21 0 20 0 0 0 0 2.8 L Present the draft results 5/1/21 5/20/21 30.0 Find the best models Select usable ML Red 3/1/21 0 20 0 0 0 0 2.8 L Present the draft results 5/1/21 5/20/21 30.0 Find the best models Select usable ML Red 3/1/21 0 20 0 0 0 0 3.1 L Colect lar results from 2.7 It is 3 5/1/21 5/20/21 30.0 Find the best models Select usable ML Red 3/1/21 0 0 0 0 0 0 3.1 L Collect the results and lark with coded research papers 5/21/21 5/20/21 15.0 Confirm the best Select usable ML Red 3/1/21 0 0 0 0 0 0 0 0 3.1 L Collect the results from 2.7 It is 3 5/1/21 5/20/21 5/20/21 20.0 Analysis the results Green 5/1/21 0 0 0 0 0 0 0 0 0 3.1 L Collect the results and lark with	2.4, L	Select Python and related libraries	1/15/21	2/28/21	8.0	Confirm R or Python	Red	1/15/21	0	45	0	0	0	0
2.71 L Machine Learning (M.) Model Analysis 2/1/21 3/28/21 20.0 Select usable ML Red 2/1/21 0 55 0 0 0 0 0 0 2.71 L Develop and train Logistic Regression regression algorithm 2/1/21 2/15/21 5.0 Select usable ML algorithms Red 2/1/21 0 15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.5 L	From those reference papers, select algorithms	1/15/21	4/28/21	8.0	ML algorithms to be	Red	1/15/21	0	104	0	0	0	0
271 Develop and train Support Vector Machine algorithm 271/21 271/21 271/21 5.0 Select usable ML elgorithms Red 271/21 0 15 0 0 0 0 0 0 0 0 0	2.6 L	Data Preparation	1/15/21	4/30/21	15.0		Red	1/15/21	0	106	0	0	0	0
2/12 2/12 3.0 algorithms Red 2/1/12 0 20 0 0 0 0 0 0 0	2.7 L	Machine Learning (ML) Model Analysis	2/1/21	3/28/21	20.0		Red	2/1/21	0	56	0	0	0	0
273 Develop and train Patients apport Vector Machine apportune 27/21 22/22 8.0 Select usable ML sejorithms Red 27/21 0 28 0 0 0 0 0 0 0 0 0	2.71 L	Develop and train Logistic Regression regression algorithm	2/1/21	2/15/21	5.0		Red	2/1/21	0	15	0	0	0	0
2.74 Delveop and train Rendom Forest Classifier algorithm 3/1/21 3/15/21 8.0 Select usable ML algorithms Red 3/1/21 0 15 0 0 0 0 0 0 0 0 0	2.72 L	Develop and train Support Vector Machine algorithm	2/1/21	2/20/21	8.0		Red	2/1/21	0	20	0	0	0	0
2.75 L Develop and train Neural Network algorithm 3/1/21 3/20/21 10.0 Select usable ML algorithms Red 3/1/21 0 20 0 0 0 0 0 0 0	2.73 L	Develop and train Decision Tree Classifier algorithm	2/1/21	2/28/21	8.0		Red	2/1/21	0	28	0	0	0	0
2.8 Present the draft results S/1/21 S/20/21 S	2.74 L	Delveop and train Random Forest Classifier algorithm	3/1/21	3/15/21	8.0		Red	3/1/21	0	15	0	0	0	0
Present the draft results	2.75 L	Develop and train Neural Network algorithm	3/1/21	3/20/21	10.0		Red	3/1/21	0	20	0	0	0	0
Delpoy and pass through Kfold validation	2.8 L	Present the draft results	5/1/21	5/20/21	30.0	lowest error from	Red	5/1/21	0	20	0	0	0	0
Second S		Test developed algorithms				0 - f t - b t								
3.2 L Test the trained all algorithms	3		4/21/21	5/30/21	15.0	model	Green	4/21/21	0	0	40	0	0	0
3.3 L Collect the results and link with coded research papers 5/21/21 5/30/21 20.0 Analysis the results Green 5/21/21 0 0 0 0 0 0 0 0 0						,								
State the goals of the research 5/21/21 6/15/21 20.0 Dissertation Brown 5/21/21 0 0 0 26 0 0												_		
4.1 L Specify the performance and learning tasks that are the focus of the research 4.2 L Describe the representation	4		5/21/21	6/15/21	20.0	Dissertation	Brown	5/21/21	0	0	0	26	0	0
4.3 L S Discuss with supervisor 5/21/21 5/21/21 1.0 Dissertation Brown 5/21/21 0 0 1 0 0 4.4 L Explain both the performance and learning components 5/21/21 6/1/21 20.0 Dissertation Brown 5/21/21 0 0 0 12 0 0 4.5 L Explain both the performance and learning components 5/21/21 6/1/21 20.0 Dissertation Brown 5/21/21 0 0 0 12 0 0 4.5 L Evaluate the approach to other methods, discussing similarities, differences, and advances over previous research 5/21/21 6/15/21 20.0 Dissertation Brown 5/21/21 0 0 26 0 0 4.7 L State the limitations of the approach and suggest directions for future research 5/21/21 6/15/21 20.0 Dissertation Brown 5/21/21 0 0 26 0 0 4.7 L State the limitations of the approach and suggest directions for future research 5/21/21 6/1/21 <td>4.1 L</td> <td>Specify the performance and learning tasks that are the focus of</td> <td></td>	4.1 L	Specify the performance and learning tasks that are the focus of												
4.4 L Explain both the performance and learning components 5/21/21 6/1/21 20.0 Dissertation Brown 5/21/21 0 0 12 0 0 4.5 L Evaluate the approach 5/21/21 6/1/21 20.0 Dissertation Brown 5/21/21 0 0 0 12 0 0 4.6 L Relate the approach to other methods, discussing similarities, differences, and advances over previous research 5/21/21 6/15/21 20.0 Dissertation Brown 5/21/21 0 0 26 0 0 4.7 L State the limitations of the approach and suggest directions for future research 5/21/21 6/1/21 20.0 Dissertation Brown 5/21/21 0 0 0 12 0 0 Conclusion Conclusion 5/21/21 6/1/21 20.0 Dissertation Dissertation 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0														
4.5 L Evaluate the approach to other methods, discussing similarities, differences, and advances over previous research 4.7 L State the limitations of the approach and suggest directions for future research 5/21/21 6/15/21 20.0 Dissertation Brown 5/21/21 0 0 0 0 26 0 0 0 0 0 0 0 0 0 0 0 0 0 0										_				
## differences, and advances over previous research		Evaluate the approach												
Solution	4.6 L		5/21/21	6/15/21	20.0	Dissertation	Brown	5/21/21	0	0	0	26	0	0
5 Create a conclusion on the best algorithm(s) 5/1/21 6/15/21 20.0 Dissertation Orange 5/1/21 0 0 0 46 0 5.1 L Create presentation slide decks 6/5/21 6/10/21 10.0 Viva Orange 6/5/21 0 <td< td=""><td>4.7 L</td><td>future research</td><td>5/21/21</td><td>6/1/21</td><td>20.0</td><td>Dissertation</td><td>Brown</td><td>5/21/21</td><td>0</td><td>0</td><td>0</td><td>12</td><td>0</td><td>0</td></td<>	4.7 L	future research	5/21/21	6/1/21	20.0	Dissertation	Brown	5/21/21	0	0	0	12	0	0
5.1 L Create presentation slide decks 6/5/21 6/10/21 10.0 Viva Orange 6/5/21 0 0 0 0 6 0 5.2 L Appendices review 5/1/21 6/10/21 10.0 Dissertation Orange 5/1/21 0 0 0 0 41 0 5.3 L Dissertation Submission 7/1/21 7/1/21 5.0 Dissertation Orange 7/1/21 0 0 0 0 0 1 0 5.4 L Viva and Appendices Submission 7/3/21 7/3/21 5.0 Viva and Appendices Orange 7/3/21 0 0 0 0 0 1 0	5		5/1/21	6/15/21	20.0	Dissertation	Orange	5/1/21	0	0	0	0	46	0
5.3 L Dissertation Submission 7/1/21 7/1/21 5.0 Dissertation Orange 7/1/21 0 0 0 1 0 5.4 L Viva and Appendices Submission 7/3/21 7/3/21 5.0 Viva and Appendices Orange 7/3/21 0 0 0 0 1 0		Create presentation slide decks					Orange							
5.4 L Viva and Appendices Submission 7/3/21 7/3/21 5.0 Viva and Appendices Orange 7/3/21 0 0 0 1 0														
	5.4 L	11			5.0	viva and Appendices	Orange	7/3/21	0	0	0	0	1	U



	Project Start	12/13/2020	1				columns	used to cr	eate the	chart			
Task ID with who is involved	TASK NAME	START	END	HOURS	DELIVERABLE	COLOR	Start		Red				
	Data set search and Analysis												
1, L	Project Proposal started	12/1/20	6/30/21	512.0	Prom02 research	Blue	12/1/20	212	0	0	0	0	0
1.1, L	Search academic F&B related papers	12/1/20	12/31/20	50.0	Select usable papers for reference	Blue	12/1/20	31	0	0	0	0	0
L	Evaluate findings	12/31/20	12/31/20	30.0	Bookmark findings	Blue	12/31/20	1	0	0	0	0	0
1.2, L	First client conversation	12/20/20	12/20/20	0.5	Get the dataset from Client	Blue	12/20/20	1	0	0	0	0	0
L	Review conversation and summerize	12/21/20	12/21/20	3.0	List out the discussion key points	Blue	12/21/20	1	0	0	0	0	0
1.3, L C	Anaysis dataset	12/21/20	1/15/21	4.0	Understand all attributes of acquired dataset	Blue	12/21/20	26	0	0	0	0	0
1.4, L S	Start to work with supervisor for Assignment 1	12/20/20	12/20/20	2.5	Prom02 research	Blue	12/20/20	1	0	0	0	0	0
L	Consolidate all references, supervisor advise and client feedbacks	1/2/21	1/28/21	20.0	Prom02 research	Blue	1/2/21	27	0	0	0	0	0
1.5, L	Assignment 1 - Planning review	1/29/21	1/29/21	20.0	Submit Assignment 1 - Planning Review	Blue	1/29/21	1	0	0	0	0	0
	Develop varies algorithms for restaurant data which reference from 1.1				A i + 2								_
2, L	Research R/Python code and chosen algorithms	12/21/20	1/25/21	20.0	Assignment 3- Appendices	Red	12/21/20	0	36	0	0	0	0
2.1, L	Analyse those research papers which have been carried out using restaurant dataset in different algorithms	1/1/21	1/25/21	5.0	Assignment 3- Appendices	Red	1/1/21	0	25	0	0	0	0
2.2, L S	Discuss with supervisor about the finding of dataset	1/10/21	1/10/21	1.0	Getting dataset ready	Red	1/10/21	0	1	0	0	0	0
2.3, L	Finalise dataset	1/13/21	1/14/21	5.0	dataset ready	Red	1/13/21	0	2	0	0	0	0
2.4, L	Select R or Python and related libraries	1/15/21	1/28/21	8.0	Confirm R or Python	Red	1/15/21	0	14	0	0	0	0
2.5 L	From those reference papers, select algorithms	1/15/21	1/28/21	8.0	Confirm number of ML algorithms to be used	Red	1/15/21	0	14	0	0	0	0
2.6 L	Data Preparation	1/15/21	1/31/21	15.0	Cleaned dataset ready	Red	1/15/21	0	17	0	0	0	0
2.7 L	Maching Learning (ML) Model Analysis	2/1/21	3/28/21	20.0	Select usable ML algorithms	Red	2/1/21	0	56	0	0	0	0
2.71 L	Develop and train Logistic Regression regression algorithm	2/1/21	2/15/21	5.0	Select usable ML algorithms	Red	2/1/21	0	15	0	0	0	0
2.72 L	Develop and train Support Vector Machine algorithm	2/1/21	2/20/21	8.0	Select usable ML algorithms	Red	2/1/21	0	20	0	0	0	0
2.73 L	Develop and train Decision Tree Classifier algorithm	2/1/21	2/28/21	8.0	Select usable ML algorithms	Red	2/1/21	0	28	0	0	0	0
2.74 L	Delveop and train Random Forest Classifier algorithm	3/1/21	3/15/21	8.0	Select usable ML algorithms	Red	3/1/21	0	15	0	0	0	0
2.75 L	Develop and train Neural Network algorithm	3/1/21	3/20/21	10.0	Select usable ML algorithms	Red	3/1/21	0	20	0	0	0	0
2.8 L	Present the draft results	4/1/21	4/20/21	30.0	Find the best model, lowest error from algorithms	Red	4/1/21	0	20	0	0	0	0
	Test developed algorithms				Confirm the best								-
3 3.1 L	Delpoy and pass through Kfold validation Collect all results from 2.71 to 3	4/21/21 5/1/21	4/30/21 5/1/21	15.0	model Analysis the results	Green	4/21/21 5/1/21	0	0	10	0	0	0
3.2 L	Test the trained all algorithms	5/1/21	5/20/21	5.0	Analysis the results	Green	5/1/21	0	0	20	0	0	0
3.3 L	Collect the results and link with coded research papers	5/21/21	5/30/21	20.0	Analysis the results	Green	5/21/21	0	0	10	0	0	0
	Research write up				,								
4	State the goals of the research	5/21/21	6/30/21	20.0	Assignment 3	Brown	5/21/21	0	0	0	41	0	0
4.1 L	Specify the performance and learning tasks that are the focus of the research	5/21/21	6/30/21	20.0	Assignment 3	Brown	5/21/21	0	0	0	41	0	0
4.2 L	Describe the representation	5/21/21	6/30/21	20.0	Assignment 3	Brown	5/21/21	0	0	0	41	0	0
4.3 L	Explain both the performance and learning components	5/21/21	6/30/21	20.0	Assignment 3	Brown	5/21/21	0	0	0	41	0	0
4.4 L 4.5 L	Evaluate the approach Relate the approach to other methods, discussing similarities, differences, and advances over previous research	5/21/21 5/21/21	6/30/21 6/30/21	20.0	Assignment 3 Assignment 3	Brown	5/21/21	0	0	0	41	0	0
4.6 L	State the limitations of the approach and suggest directions for future research	5/21/21	6/30/21	20.0	Assignment 3	Brown	5/21/21	0	0	0	41	0	0
-	Create a conclusion on the heat algorithm(s)	E14104	6/20/04	20.0	Assign	0	E/4/04	_			_	64	-
5 5.1 L	Create a conclusion on the best algorithm(s) Create presentation slide decks	5/1/21 6/5/21	6/30/21 6/10/21	20.0 10.0	Assignment 3 Assignment 3	Orange Orange	5/1/21 6/5/21	0	0	0	0	61 6	0
5.1 L	Appendices review	5/1/21	6/30/21	10.0	Assignment 3	Orange	5/1/21	0	0	0	0	61	0
5.3 L	Dissertation Submission	7/1/21	7/1/21	5.0	Assignment 3	Orange	7/1/21	0	0	0	0	1	0
5.4 L	Viva	7/3/21	7/3/21	5.0		Orange	7/3/21	0	0	0	0	1	0

Insert new rows above this one

According to above Gantt chart v2 vs v1, there are some major different from original planning to the actual work: -

- 1. The real time spent searching for a reference academic article was far longer than anticipated.
- 2. Data reduction and data transformation during data cleansing consume more than 60% of the total planning time.
- 3. Deployment of the test and training sets from Python also necessitates a significant amount of adjustment between dataset and Python code modifications. It also took significantly longer than intended.

2.2 Project Management Methods, Techniques and Tools

Method / Technique / Tool	Use & Rationale	Appropriateness	Evidence
Agile project method, MS excel	Machine learning project could	1.Roles – research owner,	Larson and Chang, 2016
or Trello is the project	use an agile approach, it is useful	process expert, data owner etc	
management tool that helps to	to take a step back and define	2.Eventsv - daily or weekly	
accomplish the research	the key concepts that should	reviews	
	drive a lean agile machine	3. Process to create and prioritize	
	learning project.	items – items/tasks are created,	
		prioritized and viewed on task	
		board in Trello.	
Data driven scrum (DDS), tool is	DDS helps to ensure that me and	1.Functional iterations 2. Flexible	Wei and Rana, 2019
scrum Kanban, DDS guide	the Ambros can achieve some	task estimation. 3. Iteration	
	key achievements	independent meeting	
		4.Collective analysis	

3. Evidence

Section 1 Evidence	Section 2 Evidence	Section 3 Evidence
Brown and Mitchell, 1993	L'heureux et al., 2017	Larson and Chang, 2016
Ayyadevara, 2018 Kim and Gu, 2006		
Tanizaki et al., 2019 Lasek et al., 2016a	Kim and Gu, 2006	Wei and Rana, 2019
Lasek et al., 2010a		
Mathur, 2018, Lian et al., 2017b, Lian et al., 2017a,	Mathur, 2018	

Presentation Slides:-

https://github.com/nunufung/PROM02/blob/main/Prom02%20ppt%20v1.46.pdf

Python code in github

https://github.com/nunufung/PROM02/blob/main/Restaurant Sales and Status Predictions v5.ipynb

Dataset CSV in github

https://github.com/nunufung/PROM02/blob/main/dataset%20v7 57k%20another%20pattern%203.csv

Books and reference in github

https://github.com/nunufung/PROM02

4. References

(Van Der Malsburg, 1986, Brown and Mitchell, 1993, Wolpert, 1996, Pinkus, 1999, Platt, 2001, Bilenko and Mooney, 2003, Kim and Gu, 2006, Cutler et al., 2007, Goldfarb and Tucker, 2011, Schrouff et al., 2013, Larson and Chang, 2016, Lasek et al., 2016a, Lasek et al., 2016b, Nam et al., 2016, Steele et al., 2016, Kirk, 2017, L'heureux et al., 2017, Lian et al., 2017a, Ayyadevara, 2018, Holmberg and Halldén, 2018, Khan and Roy, 2018, Mathur, 2018, Te et al., 2018, Adolfsson and Andersson, 2019, Tanizaki et al., 2019, Wei and Rana, 2019, Bruce et al., 2020, Dulhare et al., 2020, Gressling, 2020, Shrove and Jovanov, 2020, Varatharajan et al., 2020)

- ADOLFSSON, R. & ANDERSSON, E. 2019. Improving sales forecast accuracy for restaurants.
- AYYADEVARA, V. K. 2018. Pro machine learning algorithms. Apress: Berkeley, CA, USA.
- BILENKO, M. & MOONEY, R. J. Adaptive duplicate detection using learnable string similarity measures. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 2003. 39-48.
- BROWN, K. A. & MITCHELL, T. R. 1993. Organizational obstacles: Links with financial performance, customer satisfaction, and job satisfaction in a service environment. *Human relations*, 46, 725-757.
- BRUCE, P., BRUCE, A. & GEDECK, P. 2020. Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python, O'Reilly Media.
- CUTLER, D. R., EDWARDS JR, T. C., BEARD, K. H., CUTLER, A., HESS, K. T., GIBSON, J. & LAWLER, J. J. 2007. Random forests for classification in ecology. *Ecology*, 88, 2783-2792.
- DULHARE, U. N., AHMAD, K. & AHMAD, K. A. B. 2020. Machine Learning and Big Data: Concepts, Algorithms, Tools and Applications, John Wiley & Sons.
- GOLDFARB, A. & TUCKER, C. E. 2011. Privacy regulation and online advertising. *Management science*, 57, 57-71.
- GRESSLING, T. 2020. 36 Second generation: classification. *Data Science in Chemistry*. De Gruyter.
- HOLMBERG, M. & HALLDÉN, P. 2018. Machine Learning for Restaurant Sales Forecast.
- KHAN, P. M. & ROY, K. 2018. Current approaches for choosing feature selection and learning algorithms in quantitative structure—activity relationships (QSAR). *Expert opinion on drug discovery,* 13, 1075-1089.
- KIM, H. & GU, Z. 2006. Predicting restaurant bankruptcy: A logit model in comparison with a discriminant model. *Journal of Hospitality & Tourism Research*, 30, 474-493.
- KIRK, M. 2017. Thoughtful machine learning with Python: a test-driven approach, "O'Reilly Media, Inc.".
- L'HEUREUX, A., GROLINGER, K., ELYAMANY, H. F. & CAPRETZ, M. A. 2017. Machine learning with big data: Challenges and approaches. *leee Access*, 5, 7776-7797.
- LARSON, D. & CHANG, V. 2016. A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36, 700-710.
- LASEK, A., CERCONE, N. & SAUNDERS, J. 2016a. Restaurant sales and customer demand forecasting: Literature survey and categorization of methods. *Smart City* 360°, 479-491.
- LASEK, A., CERCONE, N. & SAUNDERS, J. 2016b. Smart Restaurants: Survey on customer demand and sales forecasting. Smart Cities and Homes, 361-386.
- LIAN, J., ZHANG, F., XIE, X. & SUN, G. CCCFNet: a content-boosted collaborative filtering neural network for cross domain recommender systems. Proceedings of the 26th international conference on World Wide Web companion, 2017a. 817-818.

- LIAN, J., ZHANG, F., XIE, X. & SUN, G. Restaurant survival analysis with heterogeneous information. Proceedings of the 26th International Conference on World Wide Web Companion, 2017b. 993-1002.
- MATHUR, P. 2018. Machine Learning Applications Using Python: Cases Studies from Healthcare, Retail, and Finance, Apress.
- NAM, H., BAEK, M. & HAN, B. 2016. Modeling and propagating cnns in a tree structure for visual tracking. arXiv preprint arXiv:1608.07242.
- PINKUS, A. 1999. Approximation theory of the MLP model. Acta Numerica 1999: Volume 8, 8, 143-195.
- PLATT, J. C. 2001. Methods and apparatus for building a support vector machine classifier. Google Patents.
- SCHROUFF, J., ROSA, M. J., RONDINA, J. M., MARQUAND, A. F., CHU, C., ASHBURNER, J., PHILLIPS, C., RICHIARDI, J. & MOURÃO-MIRANDA, J. 2013. PRONTO: pattern recognition for neuroimaging toolbox. *Neuroinformatics*, 11, 319-337.
- SHROVE, M. T. & JOVANOV, E. 2020. Early Outcome Prediction of Software Projects using Software Defects and Machine Learning. *Journal of Software Engineering Practice*, **4**, 1-12.
- STEELE, B., CHANDLER, J. & REDDY, S. 2016. *Algorithms for data science*, Springer.
- TANIZAKI, T., HOSHINO, T., SHIMMURA, T. & TAKENAKA, T. 2019. Demand forecasting in restaurants using machine learning and statistical analysis. *Procedia CIRP*, 79, 679-683.
- TE, Y.-F., MULLER, D., WYDER, S. & PRAMONO, D. 2018. Predicting the Growth of Restaurants Using Web Data. *Economic and Social Development: Book of Proceedings*, 237-256.
- VAN DER MALSBURG, C. 1986. Frank rosenblatt: Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. Brain theory. Springer.
- VARATHARAJAN, M. N., GURUPRASAD, J. & MATHUMITHA, K. 2020. Restaurant Recommendation System Using Machine Learning. *International Educational Applied Research Journal*, 4, 1-4.
- WEI, W. & RANA, M. E. 2019. Software Project Schedule Management Using Machine Learning & Data Mining. *International Journal of Scientific & Technology Research*, 8, 1385-1389.
- WOLPERT, D. H. 1996. The lack of a priori distinctions between learning algorithms. *Neural computation*, 8, 1341-1390.