Group 20 Meeting Minutes 01

Subject:	MAST90107 Data Science Project P2
Group name/ identifier	Group20
Meeting Location, Date & Time	Online (Zoom) 14:00 – 15:00, 02/08/2022
Group members present	Justin Beaconsfield (761885) Hongyang Li (1074379) Zhaopeng Li (983522) Meng Yang (1193990) Jingya Liu (1286109)
Apologies	None

1 Actions from previous meetings

Task	**Who is responsible**	**Deadline**
NA	NA	NA

2 Items discussed and decisions made

Item	**Decision (if applicable)**
Review the outcomes and procedures from the previous semester	
Improve the performance of Principle Component Analysis (PCA)	Try Robust PCA (RPCA)
How to model with RPCA	

3 Agreed actions from this meeting

Task	**Who is responsible**	**Deadline**
Take meeting minutes	Jingya Liu	03/08/2022
Modelling with RPCA	Zhaopeng Li	15/08/2022
Consider other modelling techniques	Everyone	01/10/2022

4 Next meeting

Location	Online (Zoom)
Date & Time	14:00 – 15:00, 16/08/2022
Goals	Examine the efficiency of RPCA

Group 20 Meeting Minutes 02

Subject:	MAST90107 Data Science Project P2
Group name/ identifier	Group20
Meeting Location, Date & Time	Online (Zoom) 14:00 – 15:00, 16/08/2022
Group members present	Justin Beaconsfield (761885) Hongyang Li (1074379) Zhaopeng Li (983522) Meng Yang (1193990) Jingya Liu (1286109)
Apologies	None

5 Actions from previous meetings

Task	**Who is responsible**	**Deadline**
Meeting records	Jingya Liu	03/08/2022
Implement RPCA modification	Zhaopeng Li	15/08/2022

6 Items discussed and decisions made

Item	**Decision (if applicable)**
Check the achievement of RPCA	
Compare RPCA and PCA	The results are similar
Other methods of enhancement	Try supervised learning
What kinds of supervised learning methods are useful	

7 Agreed actions from this meeting

Task	**Who is responsible**	**Deadline**
Take meeting minutes	Hongyang Li	17/08/2022
Explore the possibility of using supervised learning methods	Meng Yang	29/08/2022
Consider other modelling techniques	Everyone	01/10/2022

8 Next meeting

Location	Online (Zoom)
Date & Time	14:00 – 15:00, 30/08/2022
Goals	Examine the efficiency of supervised learning

Group 20 Meeting Minutes 03

Subject:	MAST90107 Data Science Project P2
Group name/ identifier	Group20
Meeting Location, Date & Time	Online (Zoom) 14:00 – 15:00, 30/08/2022
Group members present	Justin Beaconsfield (761885) Hongyang Li (1074379) Zhaopeng Li (983522) Meng Yang (1193990) Jingya Liu (1286109)
Apologies	None

9 Actions from previous meetings

Task	**Who is responsible**	**Deadline**
Meeting records	Hongyang Li	17/08/2022
Supervised learning model construction	Meng Yang	29/08/2022

10 Items discussed and decisions made

Item	**Decision (if applicable)**
Logistic regression (LR), support vector machine (SVM), random forest and decision tree are tested	
Evaluate the execution ability of supervised learning models	High performances are given
Is there any overfitting problem	
New methods for building models	Using K-means Clustering and K-fold Cross Validation with 3 groups

11 Agreed actions from this meeting

Task	**Who is responsible**	**Deadline**
Take meeting minutes	Jingya Liu	31/08/2022
Study the causes of overfitting	Everyone	12/09/2022
Carrying on modelling associated with K-means Clustering and K-fold Cross Validation with 3 groups	Zhaopeng Li	12/09/2022
Consider other modelling techniques	Everyone	01/10/2022

12 Next meeting

Location	Online (Zoom)
Date & Time	14:00 – 15:00, 13/09/2022
Goals	Examine the efficiency of K-means Clustering and K-fold Cross Validation with 3 groups

Group 20 Meeting Minutes 04

Subject:	MAST90107 Data Science Project P2
Group name/ identifier	Group20
Meeting Location, Date & Time	Online (Zoom) 14:00 – 15:00, 13/09/2022
Group members present	Justin Beaconsfield (761885) Hongyang Li (1074379) Zhaopeng Li (983522) Meng Yang (1193990) Jingya Liu (1286109)
Apologies	None

13 Actions from previous meetings

Task	**Who is responsible**	**Deadline**
Meeting record	Jingya Liu	31/08/2022
K-means Clustering and K-fold Cross Validation with 3 groups methods achievement	Zhaopeng Li	12/09/2022

14 Items discussed and decisions made

Item	**Decision (if applicable)**
Present the results of K-means Clustering and K-fold Cross Validation with 3 groups models	
Compare these new models with other previous models	A big improvement
Labour division	Details listed in part (15)

15 Agreed actions from this meeting

Task	**Who is responsible**	**Deadline**
Take meeting minutes	Hongyang Li	14/09/2022
Compare ratios to current results	Justin Beaconsfield	26/09/2022
Comparative metrics for clusters (distance/position)	Jingya Liu	26/09/2022
Visualizations for clusters	Hongyang Li	26/09/2022
Consider other modelling techniques	Everyone	01/10/2022

16 Next meeting

Location	Online (Zoom)
Date & Time	14:00 – 15:00, 27/09/2022
Goals	Check current progresses

Group 20 Meeting Minutes 05

Subject:	MAST90107 Data Science Project P2
Group name/ identifier	Group20
Meeting Location, Date & Time	Online (Zoom) 14:00 – 15:00, 27/09/2022
Group members present	Justin Beaconsfield (761885) Hongyang Li (1074379) Zhaopeng Li (983522) Meng Yang (1193990) Jingya Liu (1286109)
Apologies	None

17 Actions from previous meetings

Task	**Who is responsible**	**Deadline**
Meeting record	Hongyang Li	14/09/2022
Present ratios comparison	Justin Beaconsfield	26/09/2022
Measure the probability of Euclidean distance	Jingya Liu	26/09/2022
Focus on visualization	Hongyang Li	26/09/2022

18 Items discussed and decisions made

Item	**Decision (if applicable)**
Apply Euclidean distance on models	Try it with logistic regression of 3 clusters model
Put forward the idea of AA/EPA (the ratio of Arachidonic acid over Eicosapentaenoic acid)	
How to do presentation and report	

19 Agreed actions from this meeting

Task	**Who is responsible**	**Deadline**
Take meeting minutes	Jingya Liu	28/09/2022
Access AA/EPA	Meng Yang	10/10/2022
Modeling with logistic regression of 3 clusters model in term of Euclidean distance	Zhaopeng Li	10/10/2022
Give a brief description and demonstration about what we can do for the presentation and report	Hongyang Li Jingya Liu	10/10/2022

20 Next meeting

Location	Online (Zoom)
Date & Time	14:00 – 15:00, 11/10/2022
Goals	Differentiate the result of all models

Group 20 Meeting Minutes 06

Subject:	MAST90107 Data Science Project P2
Group name/ identifier	Group20
Meeting Location, Date & Time	Online (Zoom) 14:00 – 15:00, 11/10/2022
Group members present	Justin Beaconsfield (761885) Hongyang Li (1074379) Zhaopeng Li (983522) Meng Yang (1193990) Jingya Liu (1286109)
Apologies	None

21 Actions from previous meetings

Task	**Who is responsible**	**Deadline**
Meeting record	Jingya Liu	28/09/2022
Finished AA/EPA classification	Meng Yang	10/10/2022
Fit 3-clusters logistic regression model with Euclidean distance	Zhaopeng Li	10/10/2022
Think about the design of presentation and report	Hongyang Li Jingya liu	10/10/2022

22 Items discussed and decisions made

Item	**Decision (if applicable)**
Decide the final model	
Divide the part of presentation and report	
Arrange the time for presentation recording	

23 Agreed actions from this meeting

Task	**Who is responsible**	**Deadline**
Taking meeting minutes	Hongyang Li	12/10/2022
If there is any update feel free to share	Everyone	18/10/2022

24 Next meeting

Location	Online (Zoom)
Date & Time	14:00 – 15:00, 18/10/2022
Goals	This is the last meeting for presentation recording