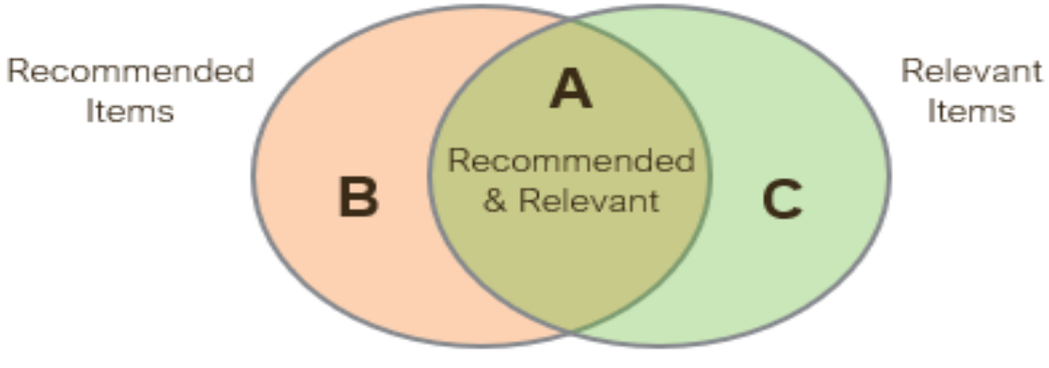
Q1

- The table below reports a ranked list of 10 TV show recommendations produced by a recommender system for a user, along with feedback indicating whether the user found each recommendation relevant or not. Based on the user's relevance feedback, calculate the Precision-at-*k* scores for *k*=1, *k=5*, and *k=10* recommendations

Rank	Show	Relevant
1	Game of Thrones	Yes
2	Arrow	No
3	House of Cards	Yes
4	Daredevil	No
5	The Returned	Yes
6	The Night Of	Yes
7	Mr. Robot	Yes
8	Supernatural	No
9	Lost	No
10	Alpha House	No

Q1



$$\text{Precision} = \frac{A}{\text{Recommended}}$$

1	Game of Thrones	Yes

At k=1
Recommended = 1
Recommended & relevant = 1
Precision-at-1 = 1/1 = 1.0

1	Game of Thrones	Yes
2	Arrow	No
3	House of Cards	Yes
4	Daredevil	No
5	The Returned	Yes

At k=5
Recommended = 5
Recommended & relevant = 3
Precision-at-5 = 3/5 = 0.6

1	Game of Thrones	Yes
2	Arrow	No
3	House of Cards	Yes
4	Daredevil	No
5	The Returned	Yes
6	The Night Of	Yes
7	Mr. Robot	Yes
8	Supernatural	No
9	Lost	No
10	Alpha House	No

At k=10
Recommended = 10
Recommended & relevant = 5
Precision-at-10 = 5/10 = 0.5

Q1 Marking Scheme

- This question is trivially easy, so the student must get each *k* absolutely correct. 1 marks for each precision-at-*k* correctly calculated. No partial marks for any other value of k

Q2

Calculating the predictions made by the model simply involves inserting the descriptive features from each query instance into the prediction model. The only extra thing that must be considered in this case is the categorical descriptive feature SOCIO ECONOMIC BAND.

We can note from the regression equation that this one feature has been expanded into two: SOCIO ECONOMIC BAND B and SOCIO ECONOMIC BAND C.

These are binary features, indicating that the original feature was set to the level ***b*** or ***c***. It is assumed that when both of these features are set to 0, then the original feature was set to ***a*** (the choice of which level to leave out is arbitrary)

Q2

1: $Logistic(-3.82398 + -0.0299 \times 56 + -0.09089 \times 1 + -0.19558 \times 0 + 0.74572 \times 1.6 + 0.02999 \times 109.32)$
$= Logistic(-1.12) = \frac{1}{1+e^{1.12}}$
$= 0.25 \Rightarrow no$

2: $Logistic(-3.82398 + -0.0299 \times 21 + -0.09089 \times 0 + -0.19558 \times 1 + 0.74572 \times 4.92 + 0.02999 \times 11.28)$
$= Logistic(-0.64) = \frac{1}{1+e^{0.64}}$
$= 0.35 \Rightarrow no$

3: $Logistic(-3.82398 + -0.0299 \times 48 + -0.09089 \times 1 + -0.19558 \times 0 + 0.74572 \times 1.21 + 0.02999 \times 161.19)$
$= Logistic(0.39) = \frac{1}{1+e^{-0.39}}$
$= 0.60 \Rightarrow no$

4: $Logistic(-3.82398 + -0.0299 \times 37 + -0.09089 \times 0 + -0.19558 \times 1 + 0.74572 \times 0.72 + 0.02999 \times 170.65)$
$= Logistic(0.53) = \frac{1}{1+e^{-0.53}}$
$= 0.63 \Rightarrow yes$

5: $Logistic(-3.82398 + -0.0299 \times 32 + -0.09089 \times 0 + -0.19558 \times 0 + 0.74572 \times 1.08 + 0.02999 \times 165.39)$
$= Logistic(0.98) = \frac{1}{1+e^{-0.98}}$
$= 0.73 \Rightarrow yes$

Q2

Note in the last instance, that we multiplied the weights of both SOCIO ECONOMIC BAND B and SOCIO ECONOMIC BAND C by zero, because the original feature was set to a, encoded as setting both b and c to 0

Q2. Marking Scheme

- ½ mark for the correct calculation for each of the query instances. The student may have rounded up, in which case the calculation after two decimal places might be slightly off. As long as the calculation is reasonably correct, give marks

- ½ mark for the correct prediction. This has to be strictly correct for each query instance. The classification threshold is 0.62 and the student should be able to say “yes” or “no” very easily.

Q3

- The table below shows the number of correct and incorrect predictions made by an image classifier during a 10-fold cross validation experiment, where the goal was to classify 500 images into one of three categories: {cats, dogs, people}.

Fold	Class: Cats		Class: Dogs		Class: People	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
1	82	68	82	68	164	36
2	81	69	102	48	176	24
3	99	51	97	53	160	40
4	81	69	102	48	148	52
5	94	56	99	51	148	52
6	97	53	91	59	162	38
7	81	69	94	56	148	52
8	76	74	79	71	181	19
9	76	74	97	53	160	40
10	96	54	79	71	179	21

Q3(a)

a) What is the overall accuracy of the classifier based on the cross-validation results?

Fold	Class: Cats		Class: Dogs		Class: People		
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	**Accuracy**
1	82	68	82	68	164	36	65.6%
2	81	69	102	48	176	24	71.8%
3	99	51	97	53	160	40	71.2%
4	81	69	102	48	148	52	66.2%
5	94	56	99	51	148	52	68.2%
6	97	53	91	59	162	38	70.0%
7	81	69	94	56	148	52	64.6%
8	76	74	79	71	181	19	67.2%
9	76	74	97	53	160	40	66.6%
10	96	54	79	71	179	21	70.8%

Fold 1: (82+82+164)/(82+68+82+68+164+36) = 65.6% accuracy for fold 1 …

Q3(a)

a) What is the overall accuracy of the classifier based on the cross-validation results?

Fold	Class: Cats		Class: Dogs		Class: People		
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	**Accuracy**
1	82	68	82	68	164	36	65.6%
2	81	69	102	48	176	24	71.8%
3	99	51	97	53	160	40	71.2%
4	81	69	102	48	148	52	66.2%
5	94	56	99	51	148	52	68.2%
6	97	53	91	59	162	38	70.0%
7	81	69	94	56	148	52	64.6%
8	76	74	79	71	181	19	67.2%
9	76	74	97	53	160	40	66.6%
10	96	54	79	71	179	21	70.8%
Mean							**68.2%**

Fold 1: (82+82+164)/(82+68+82+68+164+36) = 65.6% accuracy for fold …

Overall: (65.6% + 71.8% + 71.2% + 66.2% + 68.2% + 70.0% + 64.6% + 67.2% + 66.6% + 70.8%)/10 = 68.2%

Marking Scheme

(a)

- 2 marks if the student has all calculations correct
- 1 mark if the student has at least 50% of calculations correct

Q3(b)

b) What conclusion might be draw about the different classes in the data, based on the results above?

Fold	Class: Cats		Class: Dogs		Class: People		
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect	**Accuracy**
1	82	68	82	68	164	36	65.6%
2	81	69	102	48	176	24	71.8%
3	99	51	97	53	160	40	71.2%
4	81	69	102	48	148	52	66.2%
5	94	56	99	51	148	52	68.2%
6	97	53	91	59	162	38	70.0%
7	81	69	94	56	148	52	64.6%
8	76	74	79	71	181	19	67.2%
9	76	74	97	53	160	40	66.6%
10	96	54	79	71	179	21	70.8%
Mean	86.3	63.7	92.2	57.8	162.6	37.4	**68.2%**
Class Acc.	**57.5%**		**61.5%**		**81.3%**		

→ High accuracy for class "People", low accuracy for "Cats" and "Dogs". Suggests system is poor at distinguishing between these classes.

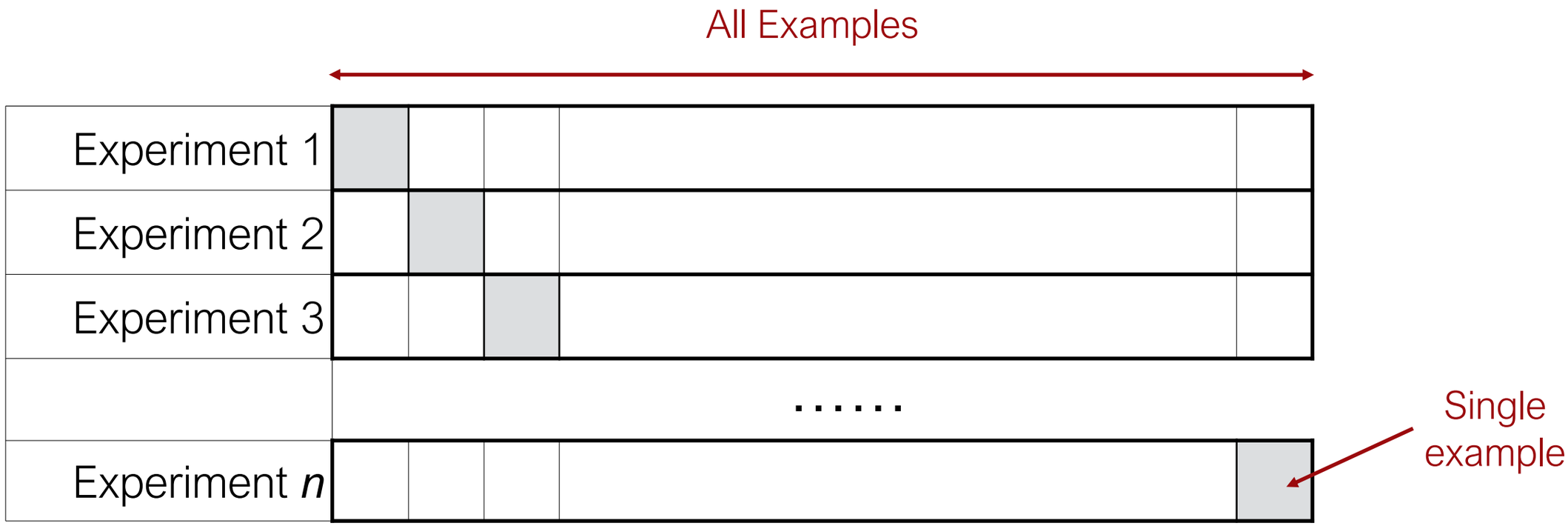
Marking Scheme

- 2 out of 3 classes have low accuracy, so the classifier is bad at distinguishing between them

  - 1 mark for correct conclusion
  - 1 mark for calculation of correct class-wise accuracy

Q3(c)

c) If a “leave-one-out cross validation” strategy had been used, how many folds would need to be created?



If each fold has 500 examples, then overall dataset has *n=5000* examples.

So leave-one-out would require running 5000 experiments where 1 example is left out each time.

Marking Scheme

- 1 mark for the correct number of folds.
- If the student has simply written the correct number, that is okay. Award marks.
- Basically, the student must demonstrate an understanding of what *leave-one-out* would imply, so please use your judgement if the student has tried to explain correctly, but got the calculation wrong. ½ mark may be given in this case.

Q4

- The fundamental assumption of using gradient descent in regression-based learning is that the error surface is *continuously differentiable* and *convex*. This means that when the slope of the gradient is zero, we have reached the minima of the error surface.
- When the partial derivative is non-zero, we want to move in the opposite direction of the gradient, hence we multiply the entire term by minus. This leads to the original term losing its negative sign.

Marking Scheme

- 1 mark for explaining the error surface properly
- 1 mark for explaining that the partial derivative is essentially in the opposite direction of the minima