

Semisupervised Quiz

Due No due date	Points 3	Questions 3
Available after May 15 at 11:32		Time Limit None

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	less than 1 minute	0 out of 3

Submitted Jun 16 at 18:23

Unanswered

Question 10 / 1 pts

We are given the dataset below, where every row shows the probability of an instance belonging to the three classes y1, y2, and y3. Using the least confident query strategy, which instance should you query first?

	y1	y2	y3
Instance 1	0.01	0.9	0.09
Instance 2	0.1	0.5	0.4
Instance 3	0.6	0.2	0.2
Instance 4	0.45	0.51	0.04

☐ instance 4

☐ instance 1

Correct Answer

☐ instance 2

☐ instance 3

In the least confidence strategy, you select an instance for which the maximum probability of belonging to a class is the smallest. So your classifier will be least confident about classifying that instance.

You select an instance for which $1 - P_{\theta}(\hat{y}|x)$ is maximum, where \hat{y} is the class for which the instance has the maximum probability.

Instance 1: $\hat{y} = y_2$ and $1 - P_{\theta}(\hat{y}|x) = 0.1$

Instance 2: $\hat{y} = y_2$ and $1 - P_{\theta}(\hat{y}|x) = 0.5$

Instance 3: $\hat{y} = y_1$ and $1 - P_{\theta}(\hat{y}|x) = 0.4$

Instance 4: $\hat{y} = y_2$ and $1 - P_{\theta}(\hat{y}|x) = 0.49$

Unanswered

Question 2

0 / 1 pts

We are given the dataset below, where every row shows the probability of an instance belonging to the three classes y_1 , y_2 , and y_3 . Using the margin sampling query strategy, which instance should our active learner query first?

	y_1	y_2	y_3
Instance 1	0.01	0.9	0.09
Instance 2	0.1	0.5	0.4
Instance 3	0.6	0.2	0.2
Instance 4	0.45	0.51	0.04

Correct Answer

☐ instance 4

☐ instance 3

☐ instance 1

☐ instance 2

In the least margin sampling strategy, you select an instance for which the difference between the first and second most probable class is the smallest. In other words the instance is on the boundary between the two most probable classes.

You select an instance for which $P_{\theta}(\hat{y}_1|x) - P_{\theta}(\hat{y}_2|x)$ is minimum, where \hat{y}_1, \hat{y}_2 are the first and second most probable classes for the instance.

Instance 1: y2 and y3 are the most probable and
 $P_{\theta}(\hat{y}_1|x) - P_{\theta}(\hat{y}_2|x) = 0.9 - 0.09 = 0.81$

Instance 2: y2 and y3 are the most probable and
 $P_{\theta}(\hat{y}_1|x) - P_{\theta}(\hat{y}_2|x) = 0.5 - 0.4 = 0.1$

Instance 3: y1 and y2 are the most probable and
 $P_{\theta}(\hat{y}_1|x) - P_{\theta}(\hat{y}_2|x) = 0.6 - 0.2 = 0.4$

Instance 4: y2 and y1 are the most probable and
 $P_{\theta}(\hat{y}_1|x) - P_{\theta}(\hat{y}_2|x) = 0.51 - 0.45 = 0.06$

Unanswered

Question 3

0 / 1 pts

Which of the following statement(s) are true about semi-supervised learning and active learning?

☐

In supervised learning, we have a passive learner while semi-supervised and active learning are designed for active learners.

orrect Answer

☐

In self training, we move instances from unlabelled to the Labelled set, for which our classifier is most confident.

orrect Answer

☐

Self training is a strategy used in Semi-supervised learning.

☐

In active learning, we query instances for which our classifier is most confident.

Both semi-supervised and supervised learning strategies are designed for passive learners. An active learner will communicate with the oracle and has a say in the selection of instances for labelling.

In active learning, we are mainly interested in the labels for instances that our classifier is least confident about, acquiring labels for the most confident instances will be a waste of resources . On the other hand, in semi-supervised learning, we expand our labelled set by adding the unlabelled instances for which our classifier is most confident, in this way we won't introduce too much error into the labelled set.