

Assignment Project Exam Help

COMP9318 Tutorial 2: Classification

<https://eduassistpro.github.io>

Add WeChat edu_assist_pr

Q1 I

Consider the following training dataset and the original decision tree induction algorithm (ID3).

Risk is the class label attribute. The *Height* values have been already discretized into disjoint ranges.

1. Calculate the information gain if *Gender* is chosen as the test attribute.
2. Calculate the information gain if *Height* is chosen as the test attribute.
3. Dr
4. Ge

F	(1.5, 1.6]	Low
M	(1.9, 2.0]	High
F	(1.8, 1.9]	Medium
F	(1.8, 1.9]	Medium
M	(1.6, 1.7]	Low
M	(1.8, 1.9]	Medium
F	(1.5, 1.6]	Low
M	(1.6, 1.7]	Low
M	(2.0, ∞]	High
M	(2.0, ∞]	High
F	(1.7, 1.8]	Medium
M	(1.9, 2.0]	Medium
F	(1.8, 1.9]	Medium
F	(1.7, 1.8]	Medium
F	(1.7, 1.8]	Medium

Q2 I

Consider applying the SPRINT algorithm on the following training dataset

<i>Age</i>	<i>CarType</i>	<i>Risk</i>
23	family	High
17	sports	High
43	sports	High
68	family	Low

Answer

1. Write down the attribute lists for the root node and its two child nodes. Do this actively.
2. Assume the first split criterion is $Age < 27.5$. Write down the attribute lists for the left child node (i.e., corresponding to the partition $Age < 27.5$).
3. Assume that the two attribute lists for the root node are stored in two relational tables name AL_Age and $AL_CarType$, respectively. We can in fact generate the attribute lists for the child nodes using standard SQL statements. Write down the SQL statements which will generate the attribute lists for the left child node for the split criterion $Age < 27.5$.
4. Write down the final decision tree constructed by the SPRINT algorithm.

Consider a (simplified) email classification example. Assume the training dataset contains 1000 emails in total, 100 of which are spams.

1. Calculate the class prior probability distribution. How would you classify a new incoming email?

2. A friend of you suggests that whether the email contains a \$ char is a good feature to detect spam emails. You look into the training dataset

a

\$

a

<https://eduassistpro.github.io>

SPAM	91	9
NOT SPAM	63	837

Describe the (naive) Bayes Classifier you can build on “evidence”. How would this classifier predict the class of an incoming email that contains a \$ character?

3. Another friend of you suggest looking into the feature of whether the email's length is longer than a fixed threshold (e.g., 500 bytes). You obtain the following results (this feature denoted as L (\bar{L})).

Assignment Project Exam Help

Class	L	\bar{L}
-------	-----	-----------

<https://eduassistpro.github.io>

incoming email that contains a \$ character and is shorter than the threshold?

Add WeChat edu_assist_pr

Based on the data in the following table,

1. estimate a Bernoulli Naive Bayes classifier (using the add-one smoothing)
2. apply the classifier to the test document
3. estimate a multinomial Naive Bayes classifier (using the add-one smoothing)
4. ap

You do
test do

	docID	words in document	ina?
training set	1	Taipei Taiwan	
	2	Macao Taiwan Shanghai	
	3	Japan Sapporo	No
	4	Sapporo Osaka Taiwan	No
test set	5	Taiwan Taiwan Taiwan Sapporo Bangkok	?

Assignment Project Exam Help

Consider a binary classification problem.

1. Fir

2

a

<https://eduassistpro.github.io>

to the positive class?

2. We then identify a feature x , and rearrange t based on their x values. The result is shown in the t

Add WeChat edu_assist_pr

x	y	count
1	-	6
1	+	2
2	-	5
2	+	2
3	-	7

Assignment Project Exam Help

<https://eduassistpro.github.io>

Table: Training Data

Add WeChat edu_assist_pro

For each of the group of training examples with the same x value, compute its probability p_i and $\text{logit}(p) := \log \frac{p}{1-p}$.

- What is your estimate of the probability that a novel test instance belongs to the positive class if its x value is 1?
- We can run a linear regression on the (x, logit) pairs from each group. Will this be the same as what Logistic Regression does?

Assignment Project Exam Help

Consider two-dimensional vectors $\mathbf{A} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ and $\mathbf{C} = \mathbf{A} + \mathbf{B}$.



<https://eduassistpro.github.io> hat if
me as

- ▶ Can you construct a matrix \mathbf{M} such that its i
in polar coordinates exhibit "linearity"? i.e.,

Add WeChat edu_assist_pr

Assignment Project Exam Help

Consider a set of d -dimensional points arranged in a *data matrix*

\mathbf{o}_1
 \mathbf{o}_2

$\mathbf{X}_{n \times d}$

s to

a m -d

$\pi(\mathbf{o}_i)$

<https://eduassistpro.github.io>

or

- Computer $r := \frac{\|\pi(\mathbf{o}_i)\|^2}{\|\mathbf{o}_i\|^2}$. Can you guess what will be

minimum values of r

Add WeChat edu_assist_pro