

MACHINE LEARNING

1. D

2. D

3. C

4. B

5. D

6. C

7. B

8. A

9. A

10. A

11. A

12. B

13. The importance of clustering is that they can cluster different customer types into one group based on different factors, such as purchasing patterns. It is important in data analysis and data mining applications. It is the task of grouping a set of objects so that objects in the same group are more similar to each other than to those in other groups.

14. Clustering is an unsupervised machine learning methodology that aims to partition data into distinct groups, or clusters. clustering performance can easily be improved by applying ICA blind source separation during the graph Laplacian embedding step.

WORKSHEET 3 SQL

1. B

2. C

3. A

4. A

5. C

6. B

7. B

8. D

9. A

10. It is a mathematical formula for determining conditional probability. It is named after 18th-century British mathematician Thomas Bayes. Bayes theorem provides a way to revise existing predictions given new or additional evidence. The theorem is also called Bayes' Rule or Bayes' Law and is the foundation of the field of Bayesian statistics. The theorem has become a useful element in the implementation of machine learning.

11. Z-score is also called a standard score. It can be placed on a normal distribution curve and also ranges from -3 standard deviations up to +3 standard deviations. To use a z-score, one needs to know the mean μ and also the population standard deviation σ . Z-scores are a way to compare results to a normal population.

The basic z score formula for a sample is: $z = (x - \mu) / \sigma$

12. T-test is a statistical test that is used to compare the means of two groups. It is also used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another. Let's take an example: you might flip a coin 1,000 times and find the number of heads follows a normal distribution for all trials. There are three t-tests to compare means: a one-sample t-test, a two-sample t-test and a paired t-test.

13. Percentile can be defined as a comparison score between a particular score and the scores of the rest of a group. It is the value below which a given percentage falls under. It is used to determine the performance of a person over others. The percentile formula is used in finding where a student stands in the test compared to other candidates. In our day-to-day life, percentile formulas are usually helpful in finding the test scores or biometric measurements. Formula = Percentile = (Number of Values Below “x” / Total Number of Values) × 100

14. Analysis of variance, or ANOVA, is a statistical method that separates observed variance data into different components. A one-way ANOVA is used for three or more groups of data, to gain information about the relationship between the dependent and independent variables. Statistical tests like ANOVA help us justify if sample results are applicable to populations.

15. ANOVA is useful for testing three or more variables. The groups differences by comparing the means of each group and includes spreading out the variance into diverse sources. It is particularly useful when analyzing the multi-item scales common in market research. It is also used to reduce or increase some treatments.

WORKSHEET 3 SQL

1. To create table for Customers

```
CREATE TABLE Customer(
```

```
    FirstName ,
```

```
    Address ,
```

```
    City )
```

2. mysql> create table `order`

```
- > (
```

```
- > order number,
```

```
- > required date
```

```
- > )
```

3. /*Select all the columns

of all the records

in the Orders:*/

```
SELECT * FROM Orders;
```

5. SELECT COUNT(*)

FROM orders

```
WHERE ord_date='xxxx-xx-xx';
```

6. SELECT (employee_number, first_name, last_name)

```
FROM employees ORDER BY department;
```

7. . SELECT ("order_number","customer_name")

```
FROM *customer
```

9. SELECT COUNT(*)

```
FROM payment
```

```
WHERE(TO_CHAR(payment_date, 'total_pymment'))
```

10. SELECT pro_name, pro_price, company_description

