

ITCertMagic

ITCertMagic

HOME

ALL VENDORS

★ GUARANTEE

? FAQ

TESTIMONIALS

CART (0)



Try **PDF Demo** before you buy

28 Top Certifications

Apr

- ▶ HP CSE ▶ Avaya Specialist
- ▶ ACE InDesign ▶ LPIC Level1
- ▶ Apple Certified Pro ▶ VCP6-CMA
- ▶ JNCDA ▶ Aruba Certification ▶ CCA XP
- ▶ ICND1 ▶ RCSP ▶ GAQM LCP
- ▶ JNCDS-SEC ▶ Fireware Essentials
- ▶ Oracle Spatial 11g

28 Top Vendors

Apr

- ▶ ISM ▶ HRCI
- ▶ Palo Alto Networks ▶ NSCA
- ▶ SUN ▶ ISQI ▶ Huawei
- ▶ American College ▶ IIA ▶ ARM
- ▶ Pegasystems ▶ OMG ▶ Simens ▶ GRE
- ▶ HAAD ▶ PCI ▶ BBPSD ▶ SCO
- ▶ SugarCRM ▶ Logical Operations ▶ IIBA
- ▶ Altiris ▶ Alfresco ▶ AMA ▶ Informatca

What Client's Say

“ There are some less than 8 new questions, so this 70-695 dump is still mostly valid. Wrote the exams today and passed. ”

Timothy ★★★★★

<http://www.itcertmagic.com/>

Pass-Guaranteed Certification Exam Questions | Exam Dumps - ITCertMagic

NO.1 Suppose you have made a model for the rating system, which rates between 1 to 5 stars. And you calculated that RMSE value is 1.0 then which of the following is correct

- A.** It means that your predictions are on average two star off of what people really think
- B.** It means that your predictions are on average four star off of what people really think
- C.** It means that your predictions are on average one star off of what people really think
- D.** It means that your predictions are on average three star off of what people really think

Answer: C

NO.2 In which of the scenario you can use the linear regression model?

- A.** Predicting tumor size reduction based on input as number of radiation treatment
- B.** Predicting Home Price based on the location and house area
- C.** Predicting demand of the goods and services based on the weather
- D.** Predicting sales of the text book based on the number of students in state

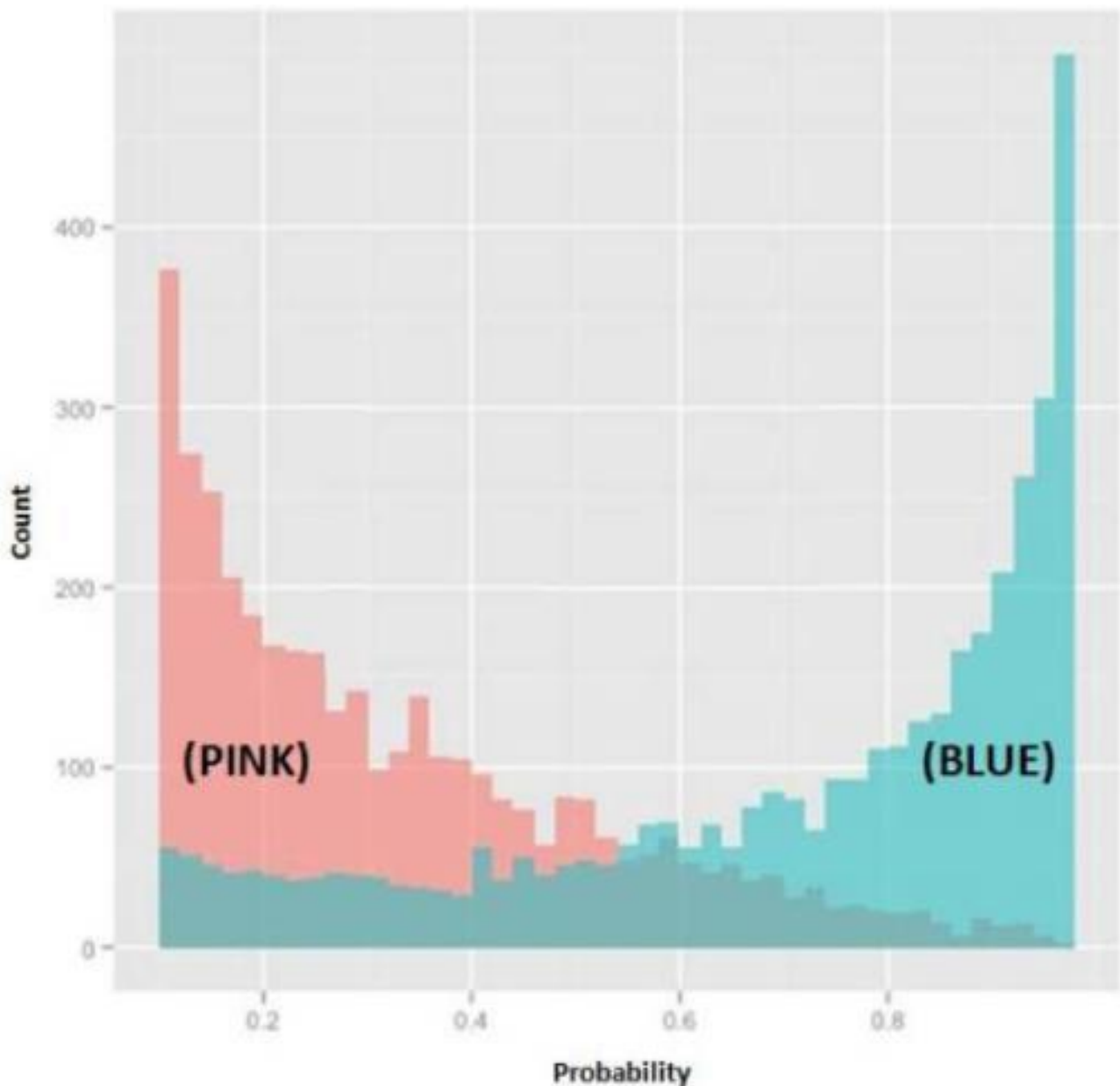
Answer: A,B,C,D

Explanation : You can use the linear regression model for predicting the continuous output variable based on the input variables. In all the cases mentioned in the question option, you can see that output can be predicted based on the input variable.

Option-A: Input: Location, House Area and Output: House Price

Option-B : Input: Weather condition, Output: Demand for the goods and services
Option-C : Input: Number of Radiation Session Output: Tumor Size Reduction
Option-D : Input: Number of students and Output: Sale quantity of text book

NO.3 Refer to Exhibit



In the exhibit, the x-axis represents the derived probability of a borrower defaulting on a loan. Also in the exhibit, the pink represents borrowers that are known to have not defaulted on their loan, and the blue represents borrowers that are known to have defaulted on their loan. Which analytical method could produce the probabilities needed to build this exhibit?

- A. Logistic Regression
- B. Association Rules
- C. Linear Regression
- D. Discriminant Analysis

Answer: A

NO.4 Question-3: In machine learning, feature hashing, also known as the hashing trick (by analogy to the kernel trick), is a fast and space-efficient way of vectorizing features (such as the words in a language), i.e., turning arbitrary features into indices in a vector or matrix. It works by applying a hash function to the features and using their hash values modulo the number of features as indices directly, rather than looking the indices up in an associative array. So what is the primary reason of

the hashing trick for building classifiers?

- A. It reduces the non-significant features e.g. punctuations
- B. Noisy features are removed
- C. It requires the lesser memory to store the coefficients for the model
- D. It creates the smaller models

Answer: C

Explanation

This hashed feature approach has the distinct advantage of requiring less memory and one less pass through the training data, but it can make it much harder to reverse engineer vectors to determine which original feature mapped to a vector location. This is because multiple features may hash to the same location. With large vectors or with multiple locations per feature, this isn't a problem for accuracy but it can make it hard to understand what a classifier is doing.

Models always have a coefficient per feature, which are stored in memory during model building. The hashing trick collapses a high number of features to a small number which reduces the number of coefficients and thus memory requirements. Noisy features are not removed; they are combined with other features and so still have an impact.

The validity of this approach depends a lot on the nature of the features and problem domain; knowledge of the domain is important to understand whether it is applicable or will likely produce poor results. While hashing features may produce a smaller model, it will be one built from odd combinations of real-world features, and so will be harder to interpret.

An additional benefit of feature hashing is that the unknown and unbounded vocabularies typical of word-like variables aren't a problem.

NO.5 You are creating a regression model with the input income, education and current debt of a customer, what could be the possible output from this model.

- A. Customer fit as a good
- B. Customer fit as acceptable or average category
- C. 1 and 3 are correct
- D. expressed as a percent, that the customer will default on a loan
- E. 2 and 3 are correct

Answer: D

Explanation

Regression is the process of using several inputs to produce one or more outputs. For example The input might be the income, education and current debt of a customer The output might be the probability, expressed as a percent that the customer will default on a loan. Contrast this to classification where the output is not a number, but a class.

NO.6 You are working in a classification model for a book, written by HadoopExam Learning Resources and decided to use building a text classification model for determining whether this book is for Hadoop or Cloud computing. You have to select the proper features (feature selection) hence, to cut down on the size of the feature space, you will use the mutual information of each word with the label of hadoop or cloud to select the 1000 best features to use as input to a Naive Bayes model. When you compare the performance of a model built with the 250 best features to a model built with the 1000 best features, you notice that the model with only 250 features performs slightly better on our test data.

What would help you choose better features for your model?

- A.** Include the number of times each of the words appears in the book in your model
- B.** Decrease the size of our training data
- C.** Include least mutual information with other selected features as a feature selection criterion
- D.** Evaluate a model that only includes the top 100 words

Answer: C

Explanation

Correlation measures the linear relationship (Pearson's correlation) or monotonic relationship (Spearman's correlation) between two variables, X and Y.

Mutual information is more general and measures the reduction of uncertainty in Y after observing X. It is the KL distance between the joint density and the product of the individual densities. So MI can measure non-monotonic relationships and other more complicated relationships. Mutual information is a quantification of the dependency between random variables. It is sometimes contrasted with linear correlation since mutual information captures nonlinear dependence.

Features with high mutual information with the predicted value are good. However a feature may have high mutual information because it is highly correlated with another feature that has already been selected.

Choosing another feature with somewhat less mutual information with the predicted value, but low mutual information with other selected features, may be more beneficial. Hence it may help to also prefer features that are less redundant with other selected features.

NO.7 You are asked to create a model to predict the total number of monthly subscribers for a specific magazine.

You are provided with 1 year's worth of subscription and payment data, user demographic data, and 10 years worth of content of the magazine (articles and pictures). Which algorithm is the most appropriate for building a predictive model for subscribers?

- A.** Linear regression
- B.** TF-IDF
- C.** Logistic regression
- D.** Decision trees

Answer: A

Explanation : A data model explicitly describes a relationship between predictor and response variables.

Linear regression fits a data model that is linear in the model coefficients. The most common type of linear regression is a least-squares fit, which can fit both lines and polynomials, among other linear models.

Before you model the relationship between pairs of quantities, it is a good idea to perform correlation analysis to establish if a linear relationship exists between these quantities. Be aware that variables can have nonlinear relationships, which correlation analysis cannot detect. For more information, see Linear Correlation.

If you need to fit data with a nonlinear model, transform the variables to make the relationship linear.

Alternatively try to fit a nonlinear function directly using either the Statistics and Machine Learning Toolbox `nlinfit` function, the Optimization Toolbox `lsqcurvefit` function, or by applying functions in the Curve Fitting Toolbox.

79

NO.8 RMSE is a good measure of accuracy, but only to compare forecasting errors of different models for a _____, as it is scale-dependent.

- A. Particular Variable
- B. Among all the variables
- C. Between Variables
- D. All of the above are correct

Answer: A

Explanation : The RMSE serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power. RMSE is a good measure of accuracy, but only to compare forecasting errors of different models for a particular variable and not between variables, as it is scale-dependent.

NO.9 Your customer provided you with 2,000 unlabeled records three groups. What is the correct analytical method to use?

- A. Semi Linear Regression
- B. Linear regression
- C. K-means clustering
- D. Naive Bayesian classification
- E. Logistic regression

Answer: C

Explanation

k-means clustering is a method of vector quantization[^] originally from signal processing, that is popular for cluster analysis in data mining, k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster This results in a partitioning of the data space into Voronoi cells. The problem is computationally difficult (NP-hard); however there are efficient heuristic algorithms that are commonly employed and converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both algorithms. Additionally they both use cluster centers to model the data; however k-means clustering tends to find clusters of comparable spatial extent, while the expectation-maximization mechanism allows clusters to have different shapes. The algorithm has nothing to do with and should not be confused with k-nearest neighbor another popular machine learning technique.

NO.10 Suppose there are three events then which formula must always be equal to $P(E1 | E2, E3)$?

- A. $P(E1, E2 | E3)P(E2 | E3)P(E3)$
- B. $P(E1, E2 | E3)P(E3)$
- C. $P(E1, E2; E3)/P(E2, E3)$
- D. $P(E1, E2, E3)P(E2)P(E3)$
- E. $P(E1, E2, E3)P(E1)/P(E2; E3)$

Answer: C