REVIEW OF INTELLIGENT DATA PROCESSING JANUARY 2018

|  |  |
| --- | --- |
| Question |  |
| Data Mining is a part of Artificial Intelligence | **F** |
| Many statistical techniques are part of data mining | **V** |
| Knowledge extraction and data mining are one | **F** |
| The data to be processed in DM should always be numeric | **F** |
| To work with data mining is not necessary from a dataset | **V** |
| Continuous data may be too detailed to be used by some techniques DM | **V** |
| The principal component analysis works well with continuous variables or as much whole | **V** |
| The cálculode frequencies involves discretizing continuous variables | **V** |
| Pooling attempts to classify items into classes already established | **F** |
| Binary data have specific distance measures based on matching values | **V** |
| The method of k-means works poorly when groups are '' spherical '' | **F** |
| The DBSCAN need to know the number of groups to start | **F** |
| Hierarchical methods are easily scalable | **F** |
| The method of the fuzzy k-means forces every point belonging to at least one group | **V** |
| Unsupervised as the kindness of a grouping can not measure well | **F** |
| Entropy based measures used to measure the quality of a grouping without prior information | **F** |
| Entropy based measures allow comparison coherence between two clusters | **V** |
| Sort equivalent predict a categorical variable | **V** |
| No method of classification gives information about how dependent classes of variables involved | **F** |
| Neural Networks methods do not provide information on the influence of the variables in the classification | **V** |
| To choose the node to repartition must take into account some measure of diversity set representing | **V** |
| áboles can not be applied to continuous variables decison when internvienen ranking | **F** |
| You can not generate classification rules if not from a decision tree. | **F** |
| As are based on probabilities, Naive Bayes-continuous variables methods only apply | **F** |
| The KNN algorithm can only act with continuous data and Euclidean distance | **F** |
| When applying KNN algorithm must take into account the nature of the variables | **V** |
| ROC curves can be applied to any classification problem | **F** |
| Predicting a variable means finding a function to represent | **F** |
| The correlation coefficient measures the linear relationship between two variables | **V** |
| The correlation coefficient can not be extended to several variables | **F** |
| It can be used as a logistic regression model with continuous binary classification independent variables | **V** |
| Association rules to apply to numeric attributes must first discretize | **V** |
| Apriori algorithm calculates the frequent itemsets using pruning strategies | **V** |
| When the consequent of a rule is very common trust it is very low | **F** |
| Like many association rules are obtained is not easy to select the best easily | **F** |
| The OLAP OLAP Mining consists of quantized data | **F** |