001. ___ useses a small labelled training set to build an initial model, which is then refined using the unlabelled data. **C**

A Supervised learning B Unsupervised learning
C Semi-supervised learning D Supervised Predictive learning

002. Which of the following sentence is FALSE regarding regression? **D**

A It relates inputs to outputs. B It is used for prediction.
C It may be used for interpretation. D It discovers causal relationships

003. Which of the following model output involves the target variable 1. Predictive model 2. Descriptive model **B**

A 1- True and 2-True B 1- True and 2- False
C 1- False and 2-True D 1- False and 2- False

004. Learning from unlabelled data is called___. **B**

A Supervised learning B Unsupervised learning
C Clustering D Regression

005. A _____ is an abstract representation of a problem. **D**

A Data B Features
C Item D Task

006. ___ describe the relevant objects in a domain. **B**

A Data B Features
C Item D Task

007. ___ is output of a machine learning algorithm applied to training data. **A**

A Model B Data
C Feature D Item

008. A function that calculates the dot product in feature space directly from the vectors in the original space is called a _____ **A**

A Kernel B Variance
C Dot product D Population mean

009. Positive covariance means that_____ **A**

A Both variables tend to increase or decrease together B Both variables tend to increase together
C Both variables tend to decrease together D Both variables will not change together

010. The objects of interest in machine learning are usually referred to as_____. **B**

A Facts B Instances
C Clusters D Concepts

011. Which of the following model output does not involves the target variable. 1. Predictive model 2. Descriptive model **C**

A 1- True and 2-True B 1- True and 2- False
C 1- False and 2-True D 1- False and 2- False

012. The process of forming general concept definitions from examples of concepts to be learned. **D**

A Deduction B Formation
C Conjunction D Induction

013. Learning from labelled data is called___. **C**

A Clustering B Regression
C Supervised learning D Unsupervised learning

014. Computers are best at learning **A**

A Facts B Concepts
C Procedures D Principles

015. _____ view learning as a process of reducing uncertainty using data **B**

A Geometric models, B Probabilistic models
C Logical models D Geometric and grading

016. Negative covariance means that **C**

A If one variable increases, the other B If one variable decrease, the other

also increase
tends to decrease

C If one variable increases, the other tends to decrease
D If one variable increases, the other will not change

017. If-then rules are used in ___ model **C**

A Geometric models,
B Probabilistic models
C Logical models
D Geometric and grading

018. The set of all possible instances is called the _____ **B**

A Problem space
B Instance space
C Feature space
D Vector space

019. _____ Data is used to build a data mining model **B**

A Validation data
B Training data
C Testing data
D Hidden data

020. _____ can often reveal useful hidden structure. **C**

A Association rules
B Regression
C Matrix decomposition
D Clusters

021. _____ models are constructed in Cartesian instance spaces, such as planes and distances. **A**

A Geometric models,
B Probabilistic models
C Logical models
D Geometric and grading

022. In machine learning settings, the task of supervised learning of predictive models is___ **B**

A Predictive clustering
B Classification and regression
C Subgroup discovery
D Association rule discovery

023. What is the purpose of performing cross-validation? 1. To assess the predictive performance of the models 2. To judge how the trained model performs outside the sample on test data **A**

A 1- True and 2-True
B 1- True and 2- False
C 1- False and 2-True
D 1- False and 2- False

024. Which of the following statement is True? 1. Models lend the machine learning field diversity 2. Tasks and features give machine learning field unity. **A**

A 1- True and 2-True
B 1- True and 2- False
C 1- False and 2-True
D 1- False and 2- False

025. The label space is used in _____to label the examples **A**

A Supervised learning
B Unsupervised learning
C Supervised learning and unsupervised learning
D Un supervised predictivelearning

026. A model is a mapping from the_____ space to the ____ space **A**

A Instance , output
B Label, instance
C Output, label
D Instance , label

027. A partial order is a binary relation that is_____. **D**

A Antisymmetric and transitive
B Reflexive and transitive
C Reflexive and antisymmetric
D Reflexive, antisymmetric and transitive

028. An equivalence relation is a binary relation that is ____ **B**

A Symmetric and transitive
B Reflexive, symmetric and transitive
C Reflexive and antisymmetric
D Reflexive, antisymmetric and transitive

029. Which of the following statements are TRUE? 1. Descriptive models identify interesting structure in the data. 2. Predictive models output involve the target variable **A**

A 1- True and 2- True
B 1- True and 2- False
C 1- False and 2- True
D 1- False and 2- False

030. In machine learning settings, the task of supervised learning of Descriptive models is___ **C**

A Predictive clustering
B Classification and regression
C Subgroup discovery
D Association rule discovery

031. Which of the following statements are True? 1. Logical models are typical examples of grouping models. 2. Geometric models are typical examples grading models. **A**

A 1- True and 2- True
B 1- True and 2- False
C 1- False and 2- True
D 1- False and 2- False

032. Suppose your model is overfitting. Which of the following is NOT a valid way to try and reduce the overfitting? **B**

A Increase the amount of training data.
B Improve the optimisation algorithm being used for error minimisation.
C Decrease the model complexity.
D Reduce the noise in the training data.

033. In machine learning settings, the task of Unsupervised learning of Predictive models is___ **A**

A Predictive clustering
B Classification and regression
C Subgroup discovery
D Association rule discovery

034. Which is not a supervised learning technique? **C**

A Decision tree induction
B Rule induction
C Clustering
D Bayesian learning

035. Which of the following are not unsupervised techniques? 1. Clustering 2. Dimensionality reduction 3. Support vector machines 4. Model ensembles **D**

A 1 and 2
B 1and 3
C 2 and 3
D 2 and 4

036. Which of the following statement is True about SVM? 1. Support vector machines are a powerful kind of linear classifier that find a decision boundary whose margin is as large as possible. 2. Support vector machines are a powerful kind of linear classifier that find a decision boundary whose margin is as small as possible. **B**

A 1- True and 2- True
B 1- True and 2- False
C 1- False and 2- True
D 1- False and 2- False

037. The most widely used metrics and tools to assess a classification model are: 1.Confusion matrix 2. Cost-sensitive accuracy 3. Area under the ROC curve **D**

A 1 only
B 2 only
C 1,2 only
D 1,2, and 3

038. The margin of a linear classifier is the distance between ____ **B**

A The two margines.
B The decision boundary and the closest instance.
C The decision boundary and the distant instance.
D The two instance.

039. We can construct a model from single features, such setting that could be described as___. **A**

A Univariate machine learning
B Linear models
C Grading models
D Grouping models

040. _____ is also called as false alarm rate. **B**

A True negative rate
B False positive rate
C False negative rate
D True positive rate

041. In machine learning settings, the task of Unsupervised learning of Descriptive models is___ **D**

A Predictive clustering
B Classification and regression
C Subgroup discovery
D Association rule discovery

042. Which of the following statements are machine learning models?I. Geometric models, II. Probabilistic modelsIII. Logical models. **D**

A Only 1
B Only 2
C 1 and 2
D 1,2 and 3

043. In which of the following scheme, it train k(k 1)/2 binary classifiers, one for each pair of different classes. 1. One-versus-rest 2. One-versus-one **B**

A 1 only
B 2 only
C Both 1 and 2
D Neither 1 nor 2

044. If we have k classes, performance of a classifier can be assessed using a ___contingency table **C**

A k-by-n B n-by-k
C k-by-k D l-by-k

045. The one-versus-rest and one-versus-one schemes are the most commonly used ways to turn _____ classifiers into multi-class classifiers. **A**

A Binary B Unary
C Simple D Complex

046. Which of the following is a good test dataset characteristic?I. It should be Large enough to yield meaningful results II. It is representative of the dataset as a whole **B**

A 1 only B 2 only
C Both 1 and 2 D Neither 1 nor 2

047. In classification, _____ and ___ are called as sensitivity and specificity , can be seen as per-class accuracies. **A**

A True positive rate , true negative rate B True positive rate , false negative rate
C False positive rate , true negative rate D False positive rate , false negative rate

048. Which of the following statements are False? 1. Grouping models divide the instance space into segments which are determined at training time. 2. Logical models are typical examples of grading models **B**

A 1- True and 2- True B 1- True and 2- False
C 1- False and 2- True D 1- False and 2- False

049. Which of the following is TRUE about role of features in machine learning:? 1. No model can exist without features. 2. Sometimes a single feature is enough to build a model 3. Features constructed from data 4. To analyze model performance, understand in what way the features need to be improved. **D**

A 1 and 2 B 2 and 3
C 1,2 and 4 D 1,2,3 and 4

050. ___ is the number of examples of each class that are classified as positive by the binary classifier **B**

A Instances B Coverage counts
C Classes D Features

051. A concept is complete if it covers _____example. 1. All positive 2. All negative **C**

A 1 only B 2 only
C 1 and 2 D Neither 1 nor 2

052. Which of the following algorithms are proposed to address multi-class classification problems 1. Neural networks 2. Decision trees 3. K-nearest neighbors 4. SVM **A**

A All 1,2,3 B Only 2,3
C Only 3,4 D Only 1

053. To decompose a k-class task into l binary classification tasks is by means of ____ **C**

A Output matrix. B Code matrix.
C Output code matrix. D Input code matrix

054. Which of the following values are entries of output code matrix? 1. +1 2. -1 3. 0 **D**

A 1 only B 2 only
C 1,2 only D All 1,2 and 3

055. Using code matrix we calculate distance d(s,c j) = i L(si c j i), and we assign the class which minimises this distance,this way of arriving at a multi-class decision from binary scores is called_____. **A**

A Loss-based decoding B Distance-based decoding
C Decoding D Encoding -decoding

056. Loss-based decoding assumes that each binary classifier scores on the ___ scale. **B**

A Unique B Same
C Multiple D Same or different

057. Which of the following sentence is FALSE regarding regression? **D**

A It relates inputs to outputs. B It is used for prediction.
C It may be used for interpretation. D It discovers causal relationships

058. An upward path through the hypothesis space corresponds to a _____ **B**
A Function estimator B Coverage curve
C Closed concept D Version space

059. The space of possible concepts_____ **D**
A Initial space B Sample space
C Input space D Hypothesis space

060. A concept is consistent if it covers ___. **A**
A None of the negative examples B All of the negative examples
C None of the positive examples D All of the positive examples

061. The version space is the set of all__ 1. Complete concepts. 2. Consistent concepts. 3. Incomplete concepts. 4. Unconsistent concepts. **A**
A 1 and 2 B 3 and 4
C 1 and 3 D 2 and 4

062. A concept that includes _____ is called a closed concept. **B**
A Complete concepts B All implicitly understood conditions
C Consistant concets D Negative examples

063. Suppose your model is overfitting. Which of the following is NOT a valid way to try and reduce the overfitting? **B**
A Increase the amount of training data. B Improve the optimisation algorithm being used for error minimisation.
C Decrease the model complexity. D Reduce the noise in the training data

064. Which of the following clustering model suitable for matrix representation? **C**
A Predictive B Descriptive
C Soft & descriptive D Predictive and descriptive

065. Which of the following clustering model row-normalized matrix representation? **A**
A Soft B Predictive
C Descriptive D Predictive and soft

066. Which of the following clustering model use partition matrix representation? **B**
A Predictive B Descriptive
C Soft & descriptive D Predictive and descriptive

067. In which of the following the label space was a discrete set of classes. 1. Classification 2. Scoring 3. Ranking4. Probability estimation **D**
A 1 , 2 B 2, 3
C 3, 4 D 1,2,3 and 4

068. A function estimator, also called a _____ **C**
A Classification B Clustering
C Regressor D Scoring

069. In regression to avoid overfitting, it is advisable to choose the degree of the polynomial _____ **C**
A 0 B 1
C As low as possible D As high as possible

070. Expand LGG. **B**
A Left General Generalization B Least General Generalization
C Like General Generalization D Lost General Generalization

071. Which of the following technique is hierarchical classification techniques? 1. One vs one 2. One vs rest3. K- nearest neighbors 4. Tree model **D**
A 1 and 2 B 3 and 4
C Only 3 D Only 4

072. In which of the following cases will K-means clustering fail to give good results?1. Data points with outliers2. Data points with different densities3. Data points with nonconvex shapes **C**
A 1 only B 1, 2 only

C 1,2 and 3 D Only 3

073. Which of the following technique is not transformed to binary classification techniques? **B**
1. One vs one 2. One vs rest3. K- nearest neighbors 4. Tree model
A 1 and 2 B 3 and 4
C Only 3 D Only 4

074. Which of the following clustering model not suitable for matrix representation? **A**
A Predictive B Descriptive
C Soft D Predictive and descriptive

075. K-means learning is _____ clustering model **A**
A Predictive B Descriptive
C Soft D Descriptive and soft

076. Rand Index is called as accuracy in _____ **B**
A Clustering B Classification
C Regression D Clustering and regression

077. In regression, Finding the right balance between over- and underfitting is sometimes called the _ **A**
A bias variance dilemma B multi-class scores
C error-correcting output codes D real-valued target variables

078. _____clustering methods represent a cluster by their centroid or exemplar **A**
A Predictive B Descriptive
C Soft D Predictive and soft

079. The representation of clustering models depends on whether they 1. Predictive 2. Descriptive3. Soft **D**
A 1 only B 1, 2 only
C 3 only D 1,2 and 3

080. Which of the following are Predictive tasks? 1. Regression2. Association Rule Discovery3. Classification4. Clustering **C**
A 1 only B 2 only
C 1 and 3 D 2 and 4

081. Which of the following are descriptive tasks? 1. Regression (predictive) 2. Association Rule Discovery (descriptive) 3. Classification (predictive) 4. Clustering (descriptive) **D**
A 1 only B 2 only
C 1 and 3 D 2 and 4

082. Association Rule Discovery and Clustering are ___ and _____ **D**
A Predictive and Supervised B Predictive and unsupervised
C Descriptive and Supervised D Descriptive and unsupervised

083. a concept class is PAC-learnable if and only if its VC-dimension is ____ **C**
A Known B Zero
C Finite D infinite

084. In regression, A piecewise constant model with n segments has___ parameters **B**
A 2n B 2n 1
C 2n +1 D 2n n

085. In the Horn algorithm, in order to learn a theory consisting of m clauses and n Boolean variables, the algorithm requires___ equivalence queries and___ membership queries. **B**
A O(mn) And O(mn) B $O(mn)$ And $O(m^2n)$
C $O(m^2n)$ And $O(m^2n)$ D $O(m^2n)$ And $O(mn)$

086. There is a well-defined algorithm for computing LGGs of first-order literals called ___ **B**
A Unification B anti-unification
C propositional D First order predicate logic

087. Which of the following statements are true about biasvariance dilemma? 1. A low-complexity model suffers less from variability due to random variations in the training data. 2. A low-complexity model introduces a systematic bias that even large amounts of training data cant resolve. 3. A high-complexity model eliminates systematic bias but can suffer non-systematic errors due to variance. **D**

A 1, 2 are True B 2, 3 are True
C 1, 3 are True D 1,2 and 3 are True

088. If we want to calculate multi-class scores and probabilities from binary classifiers, we use__ 1. We can use the distances obtained by loss-based decoding and turn them into scores by means of some appropriate transformation. 2. Naive Bayes or tree models. 3. Derived scores from coverage counts **A**

A All 1,2 and 3 B Only 2
C Only 1, 2 D Only 2, 3

089. Which of the following is not a multiclass classification technique? 1. Transformation to binary classification 2. Extension from binary classification 3. Binary classification 4. Hierarchical classification **C**

A 1 and 2 B 3 and 4
C Only 3 D Only 4

090. Which of the following technique is not Extended from binary classification techniques? 1. Navie Bayes 2. One vs rest3. K- nearest neighbors 4. Tree model **B**

A 1 and 3 B 2 and 4
C Only 3 D Only 1

091. The soft clustering approach can be addressed in ____ 1. Expectation-Maximisation2. matrix decomposition **C**

A Only 1 B Only 2
C Both 1 and 2 D Nether 1 nor 2

092. A backtracking search algorithm can return an optimal solution, at the expense1. increased computation time 2. memory requirements **C**

A 1 only B 2 only
C 1 and 2 D Neither 1 nor 2

093. Which of the following is a disadvantage of decision trees? **C**

A Factor analysis B Decision trees are robust to outliers
C Decision trees are prone to be overfit answer D Use awhite boxmodel

094. Decision tree learning is generally best suited to problems with which of the following characteristics?1. Instances are represented by attribute-value pairs. 2. The targetfunction has discrete output values. 3. Disjunctive descriptions may be required. 4. The training data may contain errors. **D**

A 1 and 2 B 3 only
C 1,2 and 3 D 4 only

095. Regarding bias and variance, which of the follwing statements are true? (Here high and low are relative to the ideal model.)1. Models which overfit have a high bias.2. Models which overfit have a low bias.3. Models which underfit have a high variance.4. Models which underfit have a low variance. **C**

A 1 and 2 B 2 and 3
C 2 and 4 D 1 and 4

096. Which of the following statements are true about loss functions?1. Classification loss functions and regression loss function both are same . 2. Classification loss functions and regression loss function both are different . **B**

A 1 is true B 2 is true
C 1 and 2 both True D Neither 1 nor 2 is true

097. Which of the following sentence is True? 1. The classification rule setting is called learning from entailment. 2. Horn theories are superficially similar to classification rule models **D**

A 1- False and 2- False B 1- False and 2-True
C 1- True and 2- False D 1- True and 2-True

098. Which of the following is true about Divide-and conquer algorithm? 1. It is greedy algorithm 2. It is non recursive algorithm 3. It is a type of tried-and-tested technique **B**

A 1-False, 2- True and 3- True B 1-True, 2- False and 3- True

C 1-True, 2- True and 3- False D 1-True, 2- True and 3- True

099. Which of the following statements are true? 1. tree models can be turned into rankers. 2. tree models can be turned into probability estimators. **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

100. To avoid decision tree from overfittingwe remove the branches that make use of features having low importance.This method is called as_____ **B**

A Choosing B Pruning
C Selecting D Firing

101. Decision trees where the target variable can take continuous values (typically real numbers) are called_____. **D**

A Binary tree B classification trees
C clustering tree D regression tree

102. Tree models where the target variable can take a discrete set of values are called_____ **B**

A Binary tree B classification trees
C clustering tree D regression tree

103. Expand CART? **B**

A Clustering and Regression Tree B Classification And Regression Tree
C Clustering Analysis representation Tree D Classification Analysis representation Tree

104. The ordering of coverage curve segments is purely based on ____ 1. the class distributions in the leaves. 2. the tree structure. **B**

A 1 only B 2 only
C Both 1,2 D Neither 1 nor 2

105. A _____ is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. **A**

A Decision tree B Graphs
C Trees D Neural Networks

106. Which of the following statements are True? 1. Regression trees are susceptible to overfitting. 2. Decision trees cannot be utilized for regression **B**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

107. _____ it uses the Gini index as impurity measure **A**

A CART B Classifier
C Regression D Clustering

108. _____ sensitive to fluctuations in the class distribution **C**

A Entropy B Gini index
C Entropy and Gini index D Minority class

109. The ranking obtained from the empirical probabilities in the leaves of a decision tree yields a _____ on the training data. **A**

A convex ROC curve B concave ROC curve
C ROC curve D convex curve

110. As already noted, the segment ordering cannot be deduced from the tree structure. Even if the place of the parent segment in the coverage curve is known, its children may come ___in the ordering. **C**

A much earlier B much later
C much earlier or later D in between

111. In general, if a feature tree has 4 leaves and we have 2 classes, then the number of possible labellings of leaves with classes is ___ **D**

A 2 B 4
C 8 D 16

112. The operation of merging all leaves in a sub-tree is called ____ the subtree. **A**

A Pruning B Plucking
C Padding D Mearging

113. Choose from the following that are Decision Tree nodes? 1. Decision Nodes 2. End Nodes 3. Chance Nodes **D**

A 1 only B 2 only
C 2,3 only D 1,2 and 3

114. The split with the ____ will be taken as the first split and the process will continue until___ . **A**

A highest information gain, the information gain is 0. B highest information gain, the information gain is 1.
C Lowest information gain, the information gain is 0. D Lowest information gain, the information gain is 1

115. In supervised rule learning, finding a combination of literals is called___ **B**

A Head of the rule B Body of the rule
C Rule list D Rule set

116. Employing the empirical distributions in the leaves of a feature tree is to build____ **C**

A probability estimators B rankers
C rankers and probability estimators D rankers or probability estimators

117. The algorithm, which is the basis for the majority of rule learning systems, is called __. **B**

A divide-and conquer algorithm. B covering algorithm
C separate-and-conquer algorithm D classification algorithm

118. A complete feature tree using the rules as features is called as_____ **C**

A complete tree B feature tree
C rule tree D binary tree

119. Which of the following statements are True? 1. Rule trees are expensive as their size is exponential in the number of rules. 2. the coverage counts have to be obtained in a separate step, after the rules have been learned in rule tree. **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

120. The class distributions in the leaves of an unlabelled feature tree can be used to turn the tree into1. a decision tree 2. a ranking tree,3. a probability estimation tree **D**

A 1 only B 1, 2 only
C 3 only D 1,2 and 3

121. A model consisting of an ordered sequence of rules which we placed is _____ **C**

A Head of the rule B Body of the rule
C Rule list D Rule set

122. What is Decision Tree? 1. Flow-Chart 2. Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label 3. Flow-Chart & Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label **C**

A 1 only B 1, 2 only
C 3 only D 2,3 only

123. Which of the following statements are True? 1. Pure means, in a selected sample of dataset all data belongs to same class. 2. Impure means, data is mixture of different classes. **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

124. Which of the following statements are True about Decision Trees? 1. Decision Tree is a display of an algorithm. 2. Decision Trees can be used for Classification Tasks **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

125. Which of the following statements are True about Gini impurity? 1. If our dataset **C**

isPurethen likelihood of incorrect classification is high.2. If our sample is mixture of different classes then likelihood of incorrect classification will be high.

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

126. Which of the following is a advantage of decision tree? Are simple to understandand interpret.Help determine worst, bestand expected values for different scenarios.Can becombined with other decision techniques **D**

A 1 only B 2 only
C 3 only D 1,2 and 3.

127. In which of the following, Decision tree is used? 1. Statistics 2. Data mining 3. Machine Learning **D**

A 1 only B 2 only
C 2,3 only D 1,2 and 3

128. To turn a feature tree into___ , we order its leaves on non-increasing empirical probabilities, which is provably optimal on the training set. **A**

A a ranker B a decision
C a probability estimation D a classifier

129. For a rule set consisting of r rules there are up to ___ different ways in which rules can overlap. **B**

A 2r B 2^r
C 2+r D 2^{r+1}

130. collections of rules is called _____ **D**

A Head of the rule B Body of the rule
C Rule list D Rule set

131. Decision Trees are anon-parametric_____ learningmethod used for __ **B**

A Supervised, classification task B Supervised, classification and regression tasks
C Unsupervised, classification task D Unsupervised, classification and regression tasks

132. Most rule learning algorithms now proceed as follows: they remove the examples covered by the rule just learned from consideration, and proceed with the remaining examples. This strategy is called _____ **A**

A separate-and-conque B Divide-and-conque
C covering algorithm/ Divide-and-conque D clustering algorithm

133. Which of the following statements are true? 1. Using precision as search heuristic is that it tends to focus a bit too much on finding pure rules. 2. A small beam size would allow us to find the more general rule. **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

134. To turn the tree into___, we choose the operating conditions and find the operating point that is optimal under those operating conditions **D**

A a ranker B a decision
C a probability estimation D a classifier

135. Which of the following statements are True? 1. In decision tree learning we are interested in the purity of both children. 2. In rule learning, we are only interested in the purity of one of the children. **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

136. To turn the feature tree into____, we predict the empirical probabilities in each leaf, applying Laplace or m-estimate smoothing to make these estimates more robust for small leaves **C**

A a ranker B a decision

C a probability estimation D a classifier

137. Which of the following statement True about Rule Trees? 1. Rule trees are inexpensive even though their size is exponential in the number of rules.2. The coverage counts have to be obtained in a separate step after the rules have been learned. 3. Rule trees can unlock the full potential of rule overlap, and give access to exact coverage counts for each area. **C**

A 1, 2 are true B 3 only true
C 2, 3 are true D 2 only false

138. While subgroup discovery,A good subgroup is___ **A**

A one whose class distribution is significantly different from the overall population. B one whose class distribution is significantly different from the sub population.
C one whose class distribution is significantly same from the overall population. D one whose class distribution is significantly same from the sub population.

139. In Rule learning for subgroup discovery, a subgroups on the diagonal always have average recall___, regardless of the class distribution. **C**

A 0.1 B 0.2
C 0.5 D 1

140. Which of the following statements are True about seed example1. The advantage is that choice of seed example cuts back the search space. 2. The disadvantage is that the choice of seed example may be sub-optimal **A**

A 1-True and 2-True B 1-True and 2- False
C 1- False and 2-True D 1- False and 2-False

001. Task of inferring a model from labelled training data is called_____ **A**

A Supervised learning B Unsupervised learning
C Reinforcement learning D Prediction learning

002. Self organizing maps are an example of_____ **B**

A Supervised learning B Unsupervised learning
C Reinforcement learning D Prediction learning

003. Some telecommunication company wants to segment their customers into distinct groups in order to send appropriate subscription offers this is an example of _____ **B**

A Supervised learning B Unsupervised learning
C Data extraction D Data reduction

004. The pre processing of data in order to remove or reduce noise is known as _____ **C**

A Data Reduction B Relevance analysis
C Data cleaning D Data extraction

005. Prediction can be viewed as forecasting a _____value **C**

A Non-continuous. B Constant.
C Continuous. D Variable

006. The problem of finding hidden structure in unlabeled data is called_____ **B**

A Supervised learning B Unsupervised learning
C Reinforcement learning D Prediction learning

007. Classification is a two step process as_____ **C**

A Join and Search B Search and prune
C Learning and classification D Join and Prune

008. _____data mining analyzes the data in order to construct one or a set of models and attempts to predict the behaviour of new data sets **B**

A Classification B Predictive
C Learning D Regression

009. Which of the following is a predictive model? **D**

A Clustering B Association rules
C Summarization D Regression

010. Each leaf node holds a _____ **B**

A Attributes B Class label
C Data Instance D Rows

011. CART means_____ **A**

A Classification and Regression Trees B Classifier and Regression Trees
C Classification and Regression Top D Class and Regression Trees

012. Classification rules are extracted from_____ **D**

A Analysis B Regression
C Classifier D Decision trees

013. Which of the following criteria is not used to decide which attribute to split next in a decision tree **D**

A Entropy B Gini Index
C Information Gain D Scatter

014. Internal nodes of a decision tree correspond to _____ **C**

A Classes B Data Instances
C Attributes D Label

015. Branch of a decision tree represents an _____ **D**

A Classes B Attributes
C Data Instance D Outcome of the test

016. _____analysis can be used to identify whether any two given attributes are statistically related **D**

A Cluster B Evolution
C Outlier D Correlation

017. A predictive model makes use of _____ **B**

A Current data. B Historical data.

C Both current and historical data. D Assumptions

018. A decision tree is a _____like tree structure **A**

A Flowchart B Map
C Graph D Node

019. Attribute selection measures are also known as_____ because they determine how the tuples at given node are to be split **A**

A Splitting rules B Training tuples
C Splitting subset D Training rules

020. ID3 uses as its attribute selection measure **B**

A Induction B Splitting rules
C Information gain D Entropy

021. Gain(A) is calculated as **C**

A $Info_A(D)$ -Info(A) B Info- $Info_A(D)$
C Info(D)-$Info_A(D)$ D Info(A)-Info(D)

022. GainRatio (A) is calculated as **D**

A Gain(A)/Gain(A) B SplitInfo(A)/SplitInfo(B)
C SplitInfo(A)/Gain(A) D Gain(A)/SplitInfo(A)

023. Attribute selection method specifies a _____procedure for selecting the attribute that best discriminates the given tuples according to class **D**

A Measured B Optimistic
C Analytical D Heuristic

024. The attribute selection measure such as the gini index enforce the resulting tree to be _____ **B**

A Ternary B Binary
C Partitioned D Pure

025. A decision tree is pruned in order to _____ **B**

A Improve classification accuracy on training set B Improve generalization performance
C Reduce dimensionality of the data D Make the tree balanced

026. Rule based classification algorithms generate _____ to perform the classification **C**

A Logic B Code
C If Then rules D Tuples

027. CART is a _____approach **A**

A Top Down B Bottom Up
C Top Up D Bottom Down

028. RID stands for **B**

A Resource Identifier B Record Identifier
C Record Idle D Repetition Idle

029. AVC stands for **C**

A Attribute Very ClassLable B Attribute Value ClasslessLable
C Attribute Value ClassLable D According Value ClassLable

030. BOAT stands for **D**

A Bootstrapped Optional Algorithm for tree Construction B Bootstrapped Optimistic Algorithm for table Construction
C Bootstrapped Optimistic Analysis for tree Construction D Bootstrapped Optimistic Algorithm for tree Construction

031. Bayesian classifiers are _____ Classifier **A**

A Statistical B Augmented
C Pruned D Analytical

032. A _____of class labeled tuples is used to estimate cost complexity **D**

A Attributes B Analysis
C Noise D Pruning set

033. Repetition occurs when an attribute is _____ tested along a given branch of a tree **A**

A Repeatedly B Analytically
C Noisely D Pruned

034. When a decision tree is built, many of the branches will reflect anomalies in the training data due to _____ or _____ **A**
A Noise, outliers B Noise, analysis
C Outliers, analysis D Noise, attributes

035. Tree pruning address the problem of _____ the data **B**
A Underfitting B Overfitting
C Minimum fitting D Maximum fitting

036. Tree pruning approaches are **C**
A Noise, outliers B Documentation and Augmentation
C Prepruning and Postpruning D Analysis and Analytical

037. The _____of a classifier on a given test set is the percentage of test set tuples that are correctly classified by the classifier **C**
A Speed B probability
C Accuracy D Substitution

038. The _____ is a useful tool for analyzing how well your classifier can recognize tuples of different classes **D**
A Confusion class B Confusion row
C Confusion table D Confusion matrix

039. Senstivity is calculated as _____ **A**
A t_pos/pos B t_pos/neg
C t_neg/pos D t_pos/minus

040. Specificity is calculated as_____ **B**
A t_pos/pos B t_neg/neg
C t_neg/pos D t_neg/minus

041. Multiple linear regression is an extension of _____ **A**
A Straight line regression B Straight curve regression
C Star line regression D Stop line regression

042. Polynomial regression is the one where it has just _____predictor variable **A**
A One B Two
C Three D Four

043. Bayes classifier is also known as **A**
A Maximum aposteriori classifier B Maximum apriori classifier
C Most likely classifier D Maximum margin classifier

044. Which of the following statement is true about Bayes classifier? **B**
A It always provides zero error when class distributions are known
B It always provides the lowest possible error when class distributions are known
C It may not always provide the lowest possible error when class distributions are known
D It always provides the lowest possible error when class distributions are estimated

045. There are two binary attributes A and B. We know that the attributes are independent and probability P(A=1) = 0.7, and P(B=0) = 0.3. What is the probability that both A and B have values 1? **D**
A 0.11 B 0.33
C 0.40 D 0.49

046. A binary variable is _____if the outcomes of the states are not equally important **B**
A Symmetric B Asymmetric
C Unsymmetric D Predictive

047. PAM stands for **D**
A Participate Around Mediods B Partitioning Around Mean
C Partitioning Area Mediods D Partitioning Around Mediods

048. Pick out a hierarchical clustering algorithm **A**

A DBSCAN B BIRCH
C PAM D CURE

049. In BIRCH the size of the CF tree can be changed by modifying the _____ **B**
A Branching factor B Threshold
C Nonleaf node D Database

050. Minkowski distance is a generalization of both _____and_____distance **A**
A Euclidean and Manhattan B Euclidean and Binary
C Categorical and Manhattan D Euclidean and Ordinal

051. A binary variable is_____if both of its states are equally valuable and carry the same weight **A**
A Symmetric B Asymmetric
C Unsymmetric D Predictive

052. In the holdout method the given data are randomly partitioned into _____and _____ independent sets **C**
A Random, single B Training, Tutor
C Training, Test D Single, double

053. The process of grouping a set of physical or abstract objects into classes of similar objects is called as_____ **D**
A Regulation B Ruling
C Classification D Clustering

054. In machine learning clustering is an example of_____ **B**
A Supervised learning B Unsupervised learning
C Reinforcement learning D Prediction learning

055. In random sampling to obtain an unbiased sampling of the data we need to know the length of the _____ in advance **D**
A Elements B Sample
C Strom D Stream

056. Which technique will be used to select an unbiased random sample of s elements without replacement **A**
A Reservoir sampling B Random sampling
C Outlier sampling D Sliding sampling

057. Classification is a _____step process **B**
A One B Two
C Three D Four

058. The change in the nature of the data takes the form of changes in the target classification model over time and is referred to as _____ **C**
A Concept adapting B Concept oriented
C Concept drift D Concept step

059. _____, _____ and _____ are needed for effective processing of stream data **B**
A Sequence, techniques, algorithms B Data structures, techniques, algorithms
C Data structures, techniques, probability D Sequence, algorithms, techniques

060. _____allows to provide summaries of the data, which typically can be used to return approximate answers to questions **C**
A Sequence B Data structures
C Synopses D Methodology

061. A tree structure called _____ is used to represent the process of hierarchical clustering **C**
A Dendogram B Dendredogram
C Dendrogram D Dencogram

062. In sequential exception technique the task of finding an exception set can be_____ **D**
A NP hard graph B NP complete

C NP D NPhard

063. A database consists of sequence of ordered elements or events, recorded with or without a concrete notion of time **A**

A Sequence B Predefined
C Concrete D Relational

064. The data stream model of computation requires algorithms to make a pass over the data with boundary memory and limited processing time **C**

A Single B Two
C Three D Four

065. Clustering of data streams follows _____strategy **D**

A Backtracking B Dynamic Programming
C Greedy Method D Divide and conquer

066. By applying which algorithm in clustering data streams the micro clusters are formed? **A**

A Hierarchical bottom up clustering B Density based clustering
C Model based clustering D Partition based clustering

067. Which of the following are the algorithms for clustering data stream? **B**

A BIRCH,ROCK B STREAM, cluStream
C Hoeffding Tree D WaveCluster

068. CVFDT uses a _____ approach **D**

A Stop and wait B Automatic repeat request
C Negative Acknowledgement D Sliding Window

069. CVFDT achieves better _____ than VFDT with time changing data streams **B**

A Decision B Accuracy
C Splitting D Confidentiality

070. VFTD stands for **D**

A Very fast Data Tree B Very free Decision Tree
C Voting fast Decision Tree D Very fast Decision Tree

071. CVFTD stands for **B**

A Concept addressing Very fast Decision Tree B Concept adapting Very fast Decision Tree
C Concept adapting Voting fast Decision Tree D Concept adapting Very free Decision Tree

072. In Hoeffding Tree Algorithm if d is the number of attributes, v is the maximum number of values for any attribute, c is the number of classes and l is the maximum depth of the tree then the total memory required is _____ **A**

A O(ldvc) B O(dlvc)
C O(cdlv) D O(vcdl)

073. CluStream is divided into a statistical data collection component and an on-line analytical component based on a _____ window **B**

A Threshold B Pyramidal time
C Online component D User specified

074. Customers who buy a Canon digital camera are likely to buy an Hpcolor printer within a month is an example of which mining technique **C**

A Spatial aggregation pattern mining B Non Sequential pattern mining
C Sequential Pattern mining D Spatial data mining

075. The Sequential Pattern mining problem was first introduced by _____and_____ **D**

A Ester and Kriegel B Jaiwei and Hanber
C Agarwal and Sharma D Agarwal and Srikant

076. The number of instances of items in a sequence is called the _____of the sequence **A**

A Length B Measurement
C Coding D Perimeter

077. The two phases of on-line microcluster processing are **C**

A Statistical data collection and updating of macro clusters
B Analytical data collection and updating of micro clusters
C Statistical data collection and updating of micro clusters
D Time horizon and updating of micro clusters

078. Which clustering allows the user to explore the stream clusters over different time horizons **A**
A Maxclusters
B Miniclusters
C Macroclusters
D Microclusters

079. STREAM derives quality k-medians clusters with limited _____and_____ **A**
A Space and Time
B Flexibility and computation
C Frame and model
D Memory and Time

080. By using which clusters the online component computes and stores summary statistics about the data stream **D**
A Maxclusters
B Miniclusters
C Macroclusters
D Microclusters

081. Which one is the titled frame model **B**
A STREAM
B Progressive logarithmic model
C Model based clustering
D Hoeffding Tree

082. Which algorithm performs bidirectional search **C**
A A SPADE
B CloSpan
C BIDE
D BIRCH

083. In constraint based mining of sequential patterns the constraints relating to the maximal length can be treated as _____constraints **D**
A Antmono
B Antimonic
C Anmonotonic
D Antimonotonic

084. The statistical methods are the one which rely on _____domain transformations **A**
A Time -to - frequency
B Time -to - size
C Time -to - measurement
D Time -to -size

085. _____refers to the extraction of knowledge, spatial relationships, or other interesting patterns not explicitly stored in spatial databases **B**
A Text mining
B Spatial data mining
C Data mining
D Image mining

086. Which of the following is the basic search methodology of SPADE and GSP? **A**
A Breadth first search and Apriori pruning
B Depth first search and Apriori pruning
C Linear search and Sequential pruning
D Linear search and Apriori pruning

087. Which one is the closed sequential pattern mining method **B**
A SPADE
B CloSpan
C DBSCAN
D BRICH

088. Which Sequential Pattern mining algorithm adopts a candidate generate and test approach using horizontal data format **B**
A SPADE
B GSP
C DBSCAN
D BRICH

089. GSP stands for **C**
A Generalized Sequence Patterns
B Generalized Sequential Path
C Generalized Sequential Patterns
D General Sequential Path

090. Which one is an Apriori based vertical data format sequential pattern mining algorithm **A**
A SPADE
B GSP
C DBSCAN
D BRICH

091. Which mining optimization method can be adopted in spatial association analysis? **D**
A Segmentation
B Factor analysis
C Regression
D Progressive refinement

092. _____ identifies clusters or densely populated regions according to some **A**

distance measurement

A Spatial data clustering B Spatial data correlation
C Spatial data Segmentation D Spatial data association

093. The trend analysis detects changes with _____, such as the changes of temporal patterns in time series data **B**

A Speed B Time
C Value D Space

094. Spatial database consists of which type of data? **C**

A Spatial B Nominal
C Vector D Ordinal

095. The efficient implementation of spatial data cubes and spatial OLAP, generalization based descriptive spatial mining such as _____and_____can be performed efficiently **A**

A Discrimination and Characterization B B Association and Correlation
C Discrimination and Selection D Discrimination and Correlation

096. is_a(X,school) close_to(X,sports_center) close_to(X,park) [0.5%,80%] is an example of which rule? **C**

A Co-location association B Non Spatial association
C Spatial association D location rule

097. The following is the example of which measure A measure in a spatial data warehouse could be the monthly_revenue of a region so that a roll up may compute the total revenue by year, by country **C**

A Nonspatial B Spatial
C Numerical D Spatial to spatial

098. A _____ measure contains a collection of pointers to spatial objects **D**

A Nonspatial B Spatial to spatial
C Numerical D Spatial

099. OLAP stands for **D**

A Online Analytical Process B Online Analysis Processing
C Online Analyzer procedure D Online Analytical Processing

100. Document ranking methods use the _____ to rank all documents **D**

A Attribute B Table
C Schema D Query

101. To avoid indexing useless words, a text retrieval system often associates _____ with a set of documents **A**

A Stop list B Directory list
C Attribute list D Word list

102. IDF stands for **B**

A Inverse document facilitate B Inverse document frequency
C Inverse domain frequency D Input domain frequency

103. The inverted index technique maintains how many hash indexed tables? **B**

A One B Two
C Three D Four

104. Recall is formally defined as **A**

A |{Relevant}{Retrieved}| / |{Relevant}| B |{Relevant}{Retrieved}| / |{Retrieved}|
C |{Retrieved}{Retrieved}| / |{Retrieved}| D |{Relevant}{Relevant}| / |{Retrieved}|

105. In order to measure the quality of a ranked list of documents, it is common to compute which one of the following at all the ranks where a new relevant document is returned **C**

A Average of Measure B Average of Methods
C Average of precision D Average of document

106. _____ is the one which consist of large collection of documents from various sources such as news articles, research papers, books. **D**

A Vector mining B Nominal mining
C Spatial mining D Text mining

107. IR stands for **A**

A Information Retrieval B Information Range
C Inter Race D Identifying Range

108. Precision is formally defined as **B**

A |{Relevant}{Retrieved}| / |{Retrieval}| B |{Relevant}{Retrieved}| / |{Retrieved}|
C |{Retrieved}{Retrieved}| / |{Retrieved}| D |{Relevant}{Relevant}| / |{Retrieved}|

109. In keyword based association analysis a set of frequently occurring consecutive or closely located keywords may form a _____ **D**

A Process B Parse
C Data D Phrase

110. Which dimensionality reduction technique is used by the spectral clustering method ? **A**

A Nonlinear embedding B Principal component analysis
C Linear dimensionality analysis D Canonical correlation analysis

111. In the following list find the most popular technique that can be used for effective document classification **B**

A Latent semantic indexing. B Bayesian classification
C Support vector machines(SVM) D Artificial neural network.

112. _____is the integration of information gathered by traditional data mining methodologies and techniques with information gathered over the World Wide Web **C**

A Text Mining B Data Mining
C Web Mining D Spatial Mining

113. Latent semantic indexing is based on which matrix decomposition method? **A**

A Singular value decomposition B Multiple value decomposition
C Binary value decomposition D Octal value decomposition

114. _____ is an essential process where intelligent methods are applied to extract data patterns. **C**

A Data warehousing B Data mining
C Text mining D Data selection

115. Which one of the following is the disadvantage of signature file in text indexing technique? **C**

A Multiple-to- triple mapping B Multiple-to- Multiple mapping
C Multiple-to-one mapping D Triple-to- one mapping

116. Pseudo feedback is also known as **D**

A Absolute feedback B Program feedback
C Code feedback D Blind feedback

117. Which one is not the dimensionality reduction technique? **B**

A Latent semantic indexing B Document indexing
C Probabilistic latent semantic analysis D Locality preserving indexing

118. In indexed based web search engine the index web pages, build and store huge _____ based indices that help locate sets of web pages containing certain keywords **D**

A Sets B Indices
C Pages D Keywords

119. In simple keyword based search engine many documents that are highly relevant to a topic may not contain keywords defining them this is referred to _____problem **A**

A Polysemy B Polyconic
C Pollywog D Polybreadth

120. DOM stands for **C**

A Document on model B Document over model
C Document object model D Document other model

121. In which type of structure does the DOM structure of a web page represents? **B**

A Graph B Tree
C Node D List

122. _____ is a link from a hypertext document to another location, activated by clicking on a highlighted word or image **B**

A Web link B Hyperlink
C Cascade style sheets D Java script

123. _____ is a service that allows Internet users to search for content via the World Wide Web **C**
A Browser B Web content
C Search Engine D Hyperlink

124. Which of the following mining is not a part of the web mining **D**
A Webusagemining, B Webstructure mining,
C Web content mining D Web data mining

125. Which of the following is not a web data content? **D**
A Text, image B Audio, Video
C Metadata, hyperlinks D Network, application

126. WWW stands for **A**
A World wide web B World wide word
C World wide window D Window wide world

127. What does the web consists? **C**
A Sets, notations B Pages, notations
C Pages, hyperlink D Pages, methods

128. _____is one or a set of web pages that provides collections of links to authorities **D**
A Gateway B Router
C Switch D Hub

129. HITS stands for **A**
A Hyperlink Induced Topic Search B Hyperlink Inductive Topic Search
C Hyperlink Induced Top Search D Hyperlink Induced Topic Survey

130. The HITS collects pages from an index based search engine those pages are known as_____ **A**
A Root set B Sub root set
C Inner root set D Outer root set

131. VIPS stands for **A**
A Vision based page segmentation B Value based page segmentation
C Vision based perform segmentation D Vision based page series

132. In the semantic structure of a web page which extracted by the vision based page segmentation each node in the tree corresponds to _____ **B**
A Attribute B Block
C Coherence D Degree

133. What is the reason for avoiding the attributes property in the HTML DOM? **B**
A Found unnecessary B Attributes dont have attributes
C Attributes have attributes D Considered irrelevant

134. How is everything treated in HTML DOM? **A**
A Node B Attributes
C Elements D Arrays

135. How are the objects organized in the HTML DOM? **C**
A Class-wise B Queue
C Hierarchy D Stack

136. The main purpose for structure mining is to extract previously unknown relationships between **A**
A Web pages B Web hyperlinks
C Web data D Web content

137. The websites that improve themselves by learning from user access patterns are known as **B**
A Non adaptive sites B Adaptive sites
C Frequent sites D Non frequent sites

138. A set of pageviews requested by a single user from a Web server **C**
A Index page B Common log

C Session D Page frame

139. The automation of Web site adaptation involves creating and deleting **A**

A Index pages B Cookies

C Pageviews D Clickstreams

140. _____is the one which mines weblog records to discover user access patterns of web pages **C**

A Weblog record B Weblog entry

C Web usage mining D Web entry

141. A web server usually registers a weblog entry for every access of a _____ **D**

A Web doc B Web entry

C Web content D Web page

142. The text surrounding hyperlink definitions in web pages is known as _____ **B**

A Quota text B Anchor text

C Simple text D Variable text

143. First step for Semantic Web is **B**

A HTML B XML

C B2C e-commerce D B2B e-commerce

144. Business Intelligence solutions minimize the transformation time of data to_____ **C**

A Fact B Knowledge

C Information D Statistic

145. _____ is a record of a user &#39s activity on the internet, including every web site and every page of every web site that the users visits, how long the user was on a page or site, in what order the pages were visited, any newsgroups that the user participates in and even the email-addresses of mail that the users send and receive. **D**

A Text mining B Clustering

C Data mining D Click stream mining

146. _____ and _____ are capable of tracking a user &#39s click stream **A**

A ISPs and Individual websites B IP and Ethernet

C Websites and MAC address D Website and login details

147. Click stream mining is also referred to as **B**

A Web Mining B Web Usage Mining

C Text Mining D Data Mining

148. Balanced Scorecard measures with benchmark for performance in **B**

A Financial areas B Nonfinancial areas

C Development areas D Structural areas

149. Consensus gaining while strategy management in balanced scorecard is made at stage of **B**

A Feedback and learning B Translating the vision

C Business planning D Learning

150. Balanced scorecard can act as a powerful **D**

A Development framework B Operations framework

C Service framework D Organizing framework

151. Personal scorecard consists of information of **B**

A 2 levels B 3 levels

C 4 levels D 5 levels

152. The sales campaigns are carefully analysed to improve **B**

A Discounts B Profits

C Reductions D Cost

153. Which analysis will be done to improve the effectiveness of sales campaigns **C**

A Quantitative, Interdimensional B Categorical , Nominal

C Multidimensional ,Association D Categorical , Association

154. With the help of mining associations from sales records one can identify the particular _____of the customers purchasing of goods **D**

A System B Sequence

C Loo D Pattern

155. _____ is a useful technique in identifying various fraud detection processes **A**

A Data Dredging B Data Backup
C Data Modelling D Data processing

156. The retail industry must update and provide timely information regarding customer requirements, product sales and fashions along with _____ , _____and _____ of products **D**

A Discount , reduction, store B Quality, cost and discount
C Discount, cost and profit D Quality, cost and profit

157. What are the techniques which are used to promote products and entice new customers **A**

A Advertisements, discounts, coupons B Discount , reduction, store
C Advertisements, discounts ,store D Discount , reduction, coupons

158. Reduction in prices made on larger purchases is classified as **B**

A Bargained reduction B Discount
C Allowance D Price segment

159. Type of reduction made from list price is classified as **C**

A bargained reduction B discount
C allowance D price segment

160. CRM stands for **C**

A Customer relation management B Customer real management
C Customer relationship management D Control relationship management

161. This is an arrangement in which a company outsources some or all of its customer relationship management functions to an application service provider **C**

A Spend management B Supplier relationship management
C Hosted CRM D Customer Information Control System

162. In which stage does an organization requires finance to expand its business operations by entering into joint ventures, mergers, and acquisitions **A**

A Development stage B Inception stage
C Maturity stage D Business stage

163. The process of managing the funds of an organization is known as **B**

A Business Management B Financial Management
C Business Intelligence D Intelligence Management

164. The data usage of preprocessing techniques to detect and correct errors and fill missing or inconsistent data values is an example of which technique in fraud detection **A**

A Statistical technique B Artificial technique
C Machine technique D Analysis technique

165. The data analyst can consolidate the telecommunication data into huge data warehouses and can regularly analyze the multidimensional data using different _____ and _____tools **B**

A Dundas and Inetsoft B OLAP and Visualization
C Qlik and Rapid miner D Dundas and Rapid miner

166. In voters list fraud is detected b using _____in combination with symbolic and analog data mining **B**

A Automation B Neural networks
C Support vector machine D Vector Machine

167. CCTV stands for **C**

A Closed clue television B Closed coupling television
C Closed circuit television D Closed circuit telegraph

168. Corporate finance is which approach **B**

A Macro B Micro
C Minimum D Maximum

169. Which of the following areas are affected by BI **B**

A Revenue B CRM

C Sales D Mining

170. Companies that targets market very narrowly is called **D**

A massmarketing B segmentedmarketing

C nichemarketing D micromarketing

171. Toyota Corporation which produces several different brands of cars is an example of **B**

A micromarketing B mass marketing

C segmented marketing D niche marketing

172. _____deals with the important decisions related to finance that an organization makes in its day to day operations **B**

A Business finance B Corporate finance

C Organization finance D Admin finance

173. Business finance is which approach **A**

A Macro B Micro

C Minimum D Maximum

174. In which stage does the organization needs finance to sustain itself in the business market through effective advertisement and constant improvement in its products **C**

A Business Management B Financial Management

C Business Intelligence D Intelligence Management

175. What are the two perspectives in which the finance is defined **B**

A Corporate and decision B Corporate and business

C Corporate and Management D Corporate and Factor