

Medicaps University

Oral Examination December - 2021

Paper Code: CS3EA07
Paper Name: Machine Learning

From Each Unit: Minimum 120 questions and as many as possible, if it is 240 questions and more it will be well

Sl No	Unit	Question No	Question
1	I	1	What is Machine Learning ?
2	I	2	What are the types of machine learning algorithms?
3	I	3	Where is ML used?
4	I	4	Who invented machine learning?
5	I	5	How many algorithms are there in machine learning?
6	I	6	How does ML algorithm work?
7	I	7	Is machine learning AI?
8	I	8	How is machine learning used in everyday life?
9	I	9	What is the use of ML?
10	I	10	What apps use machine learning?
11	I	11	What is supervised learning explain?
12	I	12	What are the steps of machine learning?
13	I	13	What is the primary objective of supervised learning?
14	I	14	What are the steps of supervised learning?
15	I	15	How do you train ML algorithm?
16	I	16	What is the aim of machine learning?
17	I	17	What are the elements of supervised learning?
18	I	18	What problems are suitable for supervised machine learning?
19	I	19	What is regression in machine learning?
20	I	20	Why is it called regression?
21	I	21	Why do we use regression?
22	I	22	What is linear regression?
23	I	23	How do you explain linear regression?
24	I	24	How do you interpret regression results?
25	I	25	What do regression coefficients tell us?
26	I	26	What is a linear regression used for?
27	I	27	How is linear regression used in business?
28	I	28	What are the advantages of regression?
29	I	29	What is difference between correlation and regression?
30	I	30	What is a good R squared value?
31	I	31	What is the disadvantage of linear Regression?
32	I	32	What are the limitations to linear regression?
33	I	33	How is linear regression used in real life?
34	I	34	Do you want R-squared to be high or low?
35	I	35	What does R Square indicate?
36	I	36	How is R ² calculated?
37	I	37	What is R vs r-squared?
38	I	38	How do you interpret R and r-squared?
39	I	39	Is SST the same as SSR?

40	I	40	What is SSE?
41	I	41	What is SST equal to?
42	I	42	What is the SSR?
43	I	43	What is meant by logistic regression?
44	I	44	What is the main purpose of logistic regression?
45	I	45	What are the assumption of logistic regression?
46	I	46	What is the output of logistic regression?
47	I	47	What are the applications of Logistic Regression?
48	I	48	How do you analyze Logistic Regression?
49	I	49	What is the equation of Logistic Regression?
50	I	50	How is Logistic Regression used in business?
51	I	51	What are the limitations of logistic regression?
52	I	52	How many predictors can be used in logistic regression?
53	I	53	What is Multicollinearity in logistic regression?
54	I	54	What is Z value in logistic regression?
55	I	55	Why is it called logistic regression?
56	I	56	Which type of dataset is used for logistic regression?
57	I	57	How do you fit data in logistic regression?
58	I	58	How do you stop Overfitting in logistic regression?
59	I	59	What is the difference between linear regression and logistic regression?
60	I	60	What is Multiple Linear Regression?
61	I	61	What does a multiple linear regression tell you?
62	I	62	What is the difference between multiple and linear regression?
63	I	63	Why do we use MLR?
64	I	64	Why is multiple regression more accurate?
65	I	65	What are R Squared and RMSE?
66	I	66	How do businesses use multiple regression?
67	I	67	What is MAE and MSE?
68	I	68	What MAE tells us?
69	I	69	Is RMSE or R2 better?
70	I	70	What is a good MSE?
71	I	71	What MAPE means?
72	I	72	What does MAPE tell a forecaster?
73	I	73	What is an advantage of the MAPE?
74	I	74	What is KNN?
75	I	75	Is KNN predictive model?
76	I	76	What is K value in KNN?
77	I	77	Is KNN supervised learning?
78	I	78	Is it possible to penalize KNN?
79	I	79	What is KNN good for?
80	I	80	Why KNN is called lazy?
81	I	81	What is the disadvantage of KNN?
82	I	82	Is KNN greedy?
83	I	83	How do you calculate KNN from K?
84	I	84	How do I find the nearest Neighbour analysis?

85	I	85	What are the steps of KNN?
86	I	86	What is gradient descent in ML?
87	I	87	Why is gradient descent useful?
88	I	88	What is J in gradient descent?
89	I	89	Why does gradient descent not converge?
90	I	90	What is gradient descent formula?
91	I	91	How does gradient descent avoid local minima?
92	I	92	Is gradient descent used in linear regression?
93	I	93	What is the advantage of using interactive algorithm like gradient descent?
94	I	94	What is Alpha in gradient descent?
95	I	95	What is learning rate in gradient decent?
96	I	96	How does learning rate affect gradient descent?
97	I	97	What does flooring learning rate in gradient descent lead to?
98	I	98	What is Cost Function?
99	I	99	What is convex and non-convex in machine learning?
100	I	100	Which of the following is not an assumption for simple linear regression? Normally distributed variables Multicollinearity Linear relationship Constant variance Normally distributed residuals
101	I	101	Continuous predictors influence the _____ of the regression line, while categorical predictors influence the _____. slope, intercept intercept, slope R2, p-value p-value, R2
102	I	102	SSE can never be a. larger than SST b. smaller than SST c. equal to 1 d. equal to zero
103	I	103	The R^2 is the squared correlation of which two values? y and the predicted values of y y and each continuous x b and t b and se
104	I	104	Regression modeling is a statistical framework for developing a mathematical equation that describes how a. one explanatory and one or more response variables are related b. several explanatory and several response variables response are related c. one response and one or more explanatory variables are related d. All of these are correct.

105	I	105	<p>Which of the following step / assumption in regression modeling impacts the trade-off between under-fitting and over-fitting the most.</p> <p>A. The polynomial degree</p> <p>B. Whether we learn the weights by matrix inversion or gradient descent</p> <p>C. The use of a constant-term</p>
106	I	106	<p>Which of the following is/ are true about “Maximum Likelihood estimate (MLE)”?</p> <p>MLE may not always exist</p> <p>MLE always exists</p> <p>If MLE exist, it (they) may not be unique</p> <p>If MLE exist, it (they) must be unique</p> <p>A. 1 and 4</p> <p>B. 2 and 3</p> <p>C. 1 and 3</p> <p>D. 2 and 4</p>
107	I	107	<p>Let’s say, a “Linear regression” model perfectly fits the training data (train error is zero). Now, Which of the following statement is true?</p> <p>A. You will always have test error zero</p> <p>B. You can not have test error zero</p> <p>C. None of the above</p>
108	I	108	What is the difference between lazy learning and eager learning?
109	I	109	What is least square regression in machine learning?
110	I	110	Why is the equation of regression line important?
111	I	111	What is gaussian function in machine learning?
112	I	112	What is Gaussian distribution in ML?
113	I	113	What is outlier?
114	I	114	How do you determine an outlier?
115	I	115	What is independent variable in machine learning?
116	I	116	What is dependent variable in machine learning?
117	I	117	What is a predictor variable?
118	I	118	What is a target variable?
119	I	119	What is train test split in ML?
120	I	120	Why is data splitting necessary?
121	II	1	What is parametric method?
122	II	2	What is overfitting?
123	II	3	What are different scenarios in which machine learning models overfitting can happen?
124	II	4	How can we avoid Overfitting?
125	II	5	When does regularization come into play in Machine Learning?
126	II	6	What is Bias, Variance

127	II	7	what do you mean by Bias-Variance Tradeoff?
128	II	8	How can we relate standard deviation and variance?
129	II	9	. A data set is given to you and it has missing values which spread along 1 standard deviation from the mean. How much of the data would remain untouched?
130	II	10	Is a high variance in data good or bad?
131	II	11	What are different scenarios in which machine learning models underfitting can happen?
132	II	12	If your dataset is suffering from high variance, how would you handle it?
133	II	13	A data set is given to you about utilities fraud detection. You have built a classifier model and achieved a performance score of 98.5%. Is this a good model? If yes, justify. If not, what can you do about it?
134	II	14	Explain the handling of missing or corrupted values in the given dataset.
135	II	15	Can you mention some advantages and disadvantages of decision trees?
136	II	16	What is a confusion matrix and why do you need it?
137	II	17	What is the difference between regularization and normalisation?
138	II	18	Explain the difference between Normalization and Standardization.
139	II	19	Differentiate between regression and classification.
140	II	20	What does the term Variance Inflation Factor mean?
141	II	21	Which machine learning algorithm is known as the lazy learner and why is it called so?
142	II	22	Differentiate between K-Means and KNN algorithms?
143	II	23	How does the SVM algorithm deal with self-learning?
144	II	24	What are Kernels in SVM? List popular kernels used in SVM along with a scenario of their applications.
145	II	25	What is Kernel Trick in an SVM Algorithm?
146	II	26	What are overfitting and underfitting
147	II	27	Why does the decision tree algorithm suffer often with overfitting problem?
148	II	28	What is OOB error and how does it occur?
149	II	29	How do you handle outliers in the data?
150	II	30	List popular cross validation techniques.
151	II	31	Define precision, recall and F1 Score?
152	II	32	What is Bayes' Theorem?
153	II	33	State at least 1 use case of bayes theorem with respect to the machine learning context?
154	II	34	What is Naive Bayes?
155	II	35	Why is Naive Bayes classifier considered as Naive?
156	II	36	Explain how a Naive Bayes Classifier works.
157	II	37	What do the terms prior probability and marginal likelihood in context of Naive Bayes theorem mean?
158	II	38	Explain the difference between Lasso and Ridge?
159	II	39	What's the difference between probability and likelihood?
160	II	40	Why would you Prune your tree?
161	II	41	Model accuracy or Model performance? Which one will you prefer and why?
162	II	42	Is linear SVM a parametric model?
163	II	43	What is the significance of Gamma and Regularization in SVM?
164	II	44	Define ROC curve work

165	II	45	What is the default method of splitting in decision trees?
166	II	46	How is p-value useful?
167	II	47	Name a few hyper-parameters of decision trees?
168	II	48	How to deal with multicollinearity?
169	II	49	What are the hyperparameters of an SVM?
170	II	50	If we have a high bias error what does it mean?
171	II	51	How to treat high bias error?
172	II	52	What is a false positive?
173	II	53	What is a false negative?
174	II	54	What's the difference between Type I and Type II error?
175	II	55	What do you understand by L1 and L2 regularization?
176	II	56	Which one is better, Naive Bayes Algorithm or Decision Trees?
177	II	57	What do you mean by the ROC curve?
178	II	58	What do you mean by AUC curve?
179	II	59	What are the advantages of SVM algorithms?
180	II	60	What is the difference between SVM Rank and SVR (Support Vector Regression)?
181	II	61	What is the difference between the normal soft margin SVM and SVM with a linear kernel?
182	II	62	How is linear classifier relevant to SVM?
183	II	63	What are the advantages of using a naive Bayes for classification?
184	II	64	In what real world applications is Naive Bayes classifier used?
185	II	65	Is naive Bayes supervised or unsupervised?
186	II	66	What do you understand by Precision and Recall?
187	II	67	What is Kernel SVM?
188	II	68	How can you avoid overfitting?
189	II	69	How can you identify a High Bias model?
190	II	70	How can you fix high bias model?
191	II	71	How to reduce High Variance?
192	II	72	What is Irreducible Error?
193	II	73	How to identify High Variance?
194	II	74	How to identify High Bias?
195	II	75	What are the problems results from different Bias - Variance combinations?
196	II	76	what are the examples of Low Bias models and High Bias models?
197	II	77	what are the examples of Low Variance models and High Variance models?
198	II	78	Decision tree algorithm belongs to the_____family. a) Supervised b) Unsupervised
199	II	79	Can we use a decision tree for both Regressions as well as Classification? a) True b) False

200	II	80	_____ is to create a training model that can be used to predict the class or value of the target variable by learning simple decision rules. a) Random Forest b) Decision Tree c) Both a and b d) None
201	II	81	Identify the type of a decision tree _____. a) Categorical b) Continuous c) Both a and b d) None
202	II	82	Are parent node and root node the both same in the decision tree? a) Yes b) No
203	II	83	Are child node and branch nodes the same in the Decision tree? a) Yes b) No
204	II	84	_____ is used for cutting or trimming the tree in Decision trees. a) Pruning b) Stemming
205	II	85	Does the decision of making strategic splits heavily affects a tree's accuracy? a) No b) Yes c) Maybe d) Don't know
206	II	86	_____ measure of the randomness in the information being processed in the Decision Tree. a) Entropy b) Information Gain
207	II	87	_____ is a statistical property that measures how well a given attribute separates the training examples according to their target classification. a) Entropy b) Information Gain
208	II	88	Decision tree is a _____. a) Non-linear ML technique. b) Non-Parametric technique. c) Supervised Learning technique. d) All of the above.
209	II	89	Choose the correct statement from below – a) A decision tree is a graphical representation of all the possible solutions to a decision based on certain conditions. b) Decision Trees usually mimic human thinking ability while making a decision, so it is easy to understand. c) A decision tree model consists of a set of rules for dividing a large heterogeneous population into smaller, more homogenous (mutually exclusive) classes. d) All of the above
210	II	90	Decision tree is also referred to as _____ algorithm. a) Recursive partitioning b) Non-Recursive partitioning. c) Variable partitioning d) None of the above.
211	II	91	Decision tree is used for _____. a) Only Regression b) Only Classification. c) Both (a) and (b) d) None of the above.

212	II	92	In the Decision tree, one rule is applied after another, resulting in a hierarchy of segments within segments. The hierarchy is called a _____, and each segment is called a _____. a) Node, Tree b) Tree, Node. c) Branch, Node d) None of the above.
213	II	93	Choose the correct sequence of typical decision tree structure – (I) Take the entire data set as input (II) Divide the input data into two part (III) Reapply the split to every part recursively (IV) Stop when meeting desired criteria (V) Cut the tree when we went too far while doing splits(pruning) a) (I), (II),(V),(IV),(III). b) (V),(I),(III),(II),(IV). c) (I),(III),(II),(V),(IV). d) (I),(II),(III),(IV),(V).
214	II	94	_____ denotes the entire population or sample and it further divides into two or more homogeneous sets. a) Leaf node b) Terminal node. c) Root node d) None of the above.
215	II	95	he Process of removing sub-nodes from a decision node is called _____. a) Splitting b) Breaking. c) Pruning d) None of the above.
216	II	96	In Decision tree pruning methods include a) No pruning b) Reduced error pruning. c) Bagging d) All of the above
217	II	97	In decision tree we only use discrete data ? a. True b.False
218	II	98	Which algorithm is most prone to overfitting? a. Link analysis b. Neural Analysis c. Decision Tree d. None of the above
219	II	99	In which pruning method we allow our tree to grow till last including all the observations and variables? a. post pruning b. pre pruning c. both a & b d. None of the above
220	II	100	Pre-pruning the decision tree may results in a. Overfitting b. Underfitting
221	II	101	Statement : Missing data can be handled by the DT. reason : classification is done by the yes or no condition. a. statment is True, Reason gives the correct explanation of statement b. statement is false, Reason is correct c. statement is false, reason is incorrect d. statement is true, reason is false explanation of statement
222	II	102	The minimum time complexity for training an SVM is $O(n^2)$. According to this fact, what sizes of datasets are not best suited for SVM's?

223	II	103	The effectiveness of an SVM depends upon: A) Selection of Kernel B) Kernel Parameters C) Soft Margin Parameter C D) All of the above
224	II	104	Support vectors are the data points that lie closest to the decision surface. A) TRUE B) FALSE
225	II	105	The SVM's are less effective when: A) The data is linearly separable B) The data is clean and ready to use C) The data is noisy and contains overlapping points
226	II	106	The cost parameter in the SVM means: A) The number of cross-validations to be made B) The kernel to be used C) The tradeoff between misclassification and simplicity of the model D) None of the above
227	II	107	If I am using all features of my dataset and I achieve 100% accuracy on my training set, but ~70% on validation set, what should I look out for? A) Underfitting B) Nothing, the model is perfect C) Overfitting
228	II	108	Which of the following are real world applications of the SVM? A) Text and Hypertext Categorization B) Image Classification C) Clustering of News Articles D) All of the above
229	II	109	What is/are true about kernel in SVM? 1. Kernel function map low dimensional data to high dimensional space 2. It's a similarity function A) 1 B) 2 C) 1 and 2 D) None of these
230	II	110	What are kernel methods used for?
231	II	111	Is kernel methods can be used for supervised and unsupervised problems?
232	II	112	What is the use of kernel methods in SVM?
233	II	113	Which kernel is used in SVM?
234	II	114	What is gaussian kernel in SVM?
235	II	115	Why SVM is widely used in machine learning?
236	II	116	What is C parameter in SVM?
237	II	117	What is the main assumption used in naive Bayes?
238	II	118	Can naive Bayes be used for regression?
239	II	119	Why do naive Bayesian classifiers perform so well?
240	II	120	Is naive Bayes clustering algorithm?
241	III	1	What is unsupervised learning example?
242	III	2	What is difference between supervised and unsupervised learning?
243	III	3	How is unsupervised learning used?
244	III	4	Which learning is better supervised or unsupervised?
245	III	5	What is K means algorithm?
246	III	6	Is K-means supervised or unsupervised?
247	III	7	What are the challenges in unsupervised learning?

248	III	8	What is a clustering algorithm?
249	III	9	Is Kmeans deterministic?
250	III	10	Why is K-means clustering used?
251	III	11	What is clustering used for?
252	III	12	What is a cluster evaluation?
253	III	13	What is distance based clustering in machine learning?
254	III	14	What are the distance based clustering algorithms?
255	III	15	What is K means algorithm in machine learning?
256	III	16	Why choose K-means clustering?
257	III	17	What is distance measure in clustering?
258	III	18	What is distance matrix in clustering?
259	III	19	What is Manhattan distance formula?
260	III	20	Why is it called Manhattan distance?
261	III	21	Is Manhattan distance consistent?
262	III	22	How do you calculate Euclidean distance?
263	III	23	How does Euclidean distance work?
264	III	24	What is Euclidean distance in K NN?
265	III	25	Where is KNN algorithm used?
266	III	26	What are the advantages of KNN algorithm?
267	III	27	Can KNN be used for regression?
268	III	28	Can KNN be used for continuous target?
269	III	29	How do you calculate KNN from K?
270	III	30	What is meant by dimensionality reduction?
271	III	31	What is PCA used for?
272	III	32	Why do we use dimensionality reduction?
273	III	33	Which technique handles higher dimensionality data very well?
274	III	34	Can dimensionality reduction be reversed?
275	III	35	What is Multi Dimensionality means?
276	III	36	What are the limitations of PCA?
277	III	37	What is PC1 and PC2 in PCA?
278	III	38	What is PCA transformation?
279	III	39	Does PCA reduce accuracy?
280	III	40	Is PCA unsupervised?
281	III	41	What is PC1 vs PC2?
282	III	42	What is loading in PCA?
283	III	43	Does PCA reduce variance?
284	III	44	Does PCA improve prediction?
285	III	45	Does PCA really improve classification outcome?
286	III	46	What is the curse of dimensionality in machine learning?
287	III	47	What is eigenvalues in PCA?
288	III	48	What is eigenvector in PCA?
289	III	49	Does PCA create new features?
290	III	50	What is feature extraction in ML?
291	III	51	Why is feature extraction used?
292	III	52	Can we use PCA for feature selection?
293	III	53	What's the difference between feature selection and feature extraction?

294	III	54	Which is better feature extraction or feature selection?
295	III	55	What is feature selection ?
296	III	56	Is feature selection necessary for deep learning?
297	III	57	What is matrix factorization in machine learning?
298	III	58	Where is matrix factorization used?
299	III	59	Is matrix factorization supervised or unsupervised?
300	III	60	What is recommender system in AI/ML?
301	III	61	Which ML algorithm is used for recommendation system?
302	III	62	How do you create a recommendation system in machine learning?
303	III	63	What is the use of recommender system?
304	III	64	Is recommender system supervised or unsupervised?
305	III	65	Does Netflix use collaborative filtering?
306	III	66	What are the benefits of recommender systems?
307	III	67	What makes a good recommendation engine?
308	III	68	How do recommender systems relate to classification?
309	III	69	Why recommender systems are being used in e commerce?
310	III	70	What type of machine learning is recommender system?
311	III	71	What is collaborative filtering in ML?
312	III	72	What is the difference between content-based filtering and collaborative filtering?
313	III	73	What is model based CF?
314	III	74	What is content based recommender system?
315	III	75	Which is an example of content based recommendation system?
316	III	76	Which is better content-based or collaborative filtering?
317	III	77	Is content filtering supervised or unsupervised?
318	III	78	What is content recommendation?
319	III	79	What is cluster sampling?
320	III	80	Which of the following is required by K-means clustering? a) defined distance metric b) number of clusters c) initial guess as to cluster centroids d) all of the mentioned
321	III	81	Point out the wrong statement. a) k-means clustering is a method of vector quantization b) k-means clustering aims to partition n observations into k clusters c) k-nearest neighbor is same as k-means d) none of the mentioned
322	III	82	Which of the following combination is incorrect? a) Continuous – euclidean distance b) Continuous – correlation similarity c) Binary – manhattan distance d) None of the mentioned
323	III	83	K-means is not deterministic and it also consists of number of iterations. a) True b) False

324	III	84	<p>Movie Recommendation systems are an example of:</p> <p>Classification Clustering Reinforcement Learning Regression</p>
325	III	85	<p>Sentiment Analysis is an example of:</p> <p>Regression Classification Clustering Reinforcement Learning</p>
326	III	86	What is the minimum no. of variables/ features required to perform clustering?
327	III	87	<p>For two runs of K-Mean clustering is it expected to get same clustering results?</p> <p>A. Yes B. No</p>
328	III	88	<p>Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means</p> <p>A. Yes B. No C. Can't say D. None of these</p>
329	III	89	<p>Which of the following can act as possible termination conditions in K-Means?</p> <p>For a fixed number of iterations. Assignment of observations to clusters does not change between iterations. Except for cases with a bad local minimum. Centroids do not change between successive iterations. Terminate when RSS falls below a threshold.</p>
330	III	90	<p>Which of the following algorithm is most sensitive to outliers?</p> <p>A. K-means clustering algorithm B. K-medians clustering algorithm C. K-modes clustering algorithm D. K-medoids clustering algorithm</p>

331	III	91	<p>How can Clustering (Unsupervised Learning) be used to improve the accuracy of Linear Regression model (Supervised Learning):</p> <p>Creating different models for different cluster groups. Creating an input feature for cluster ids as an ordinal variable. Creating an input feature for cluster centroids as a continuous variable. Creating an input feature for cluster size as a continuous variable.</p>
332	III	92	<p>Which of the following are the applications of clustering?</p> <p>Identifying patterns of crime in different regions of a city and managing police enforcement based on frequency and type of crime Looking at social media behaviour to find out the types of online communities that exist All of the above Correct answer. Identifying consumer segments and their properties to position products appropriately</p>
333	III	93	<p>Following the steps to run a PCA's algorithm, why is so important standardize your data?</p> <p>a. Make the training time more fast b. Standardize data allows other people understand better your work c. Find the features which can best predicts Y d. Use the best practices of data wrangling</p>
334	III	94	<p>When we need to use PCA?</p> <p>a. Everytime before uses a Machine Learning algorithm b. When my data is small and with a few features c. When I have a overfit case d. You want to find latent features and reduce dimensionality</p>
335	III	95	<p>PCA reduces the dimension by finding a few_____.</p> <p>A. Hexagonal linear combination B. Orthogonal linear combinations C. Octagonal linear combination D. Pentagonal Linear Combination</p>
336	III	96	<p>PCA is a _____.</p> <p>A. Non linear method B. Linear method C. Continuous method D. Repeated method</p>
337	III	97	<p>Which of the following is not kernel method?</p> <p>A. linear B. polynomial C. gaussian D. Continuously</p>
338	III	98	<p>_____ is non-zero vector that stays parallel after matrix multiplication.</p> <p>A. Eigen value B. Eigen vector C. Linear value D. None of these</p>

339	III	99	<p>_____ basically known as characteristic roots. It basically measures the variance in all variables which is accounted for by that factor</p> <p>A. Eigen value B. Eigen vector C. Linear value D. None of these</p>
340	III	100	<p>Euclidean distance measure is</p> <p>A stage of the KDD process in which new data is added to the existing selection.</p> <p>The process of finding a solution for a problem simply by enumerating all possible solutions according to some pre-defined order and then testing them</p> <p>The distance between two points as calculated using the Pythagoras theorem</p> <p>none of above</p>
341	III	101	<p>_____ metric is examined to determine a reasonably optimal value of k.</p> <p>Mean Square Error Within Sum of Squares (WSS) Speed None of these</p>
342	III	102	<p>_____ recommend items based on similarity measures between users and/or items.</p> <p>Content Based Systems Hybrid System Collaborative Filtering Systems None of these</p>
343	III	103	<p>There are _____ major Classification of Collaborative Filtering Mechanisms</p> <p>1 2 3 none of above</p>
344	III	104	<p>Movie Recommendation to people is an example of</p> <p>User Based Recommendation Item Based Recommendation Knowledge Based Recommendation content based recommendation</p>
345	III	105	<p>_____ recommenders rely on an explicitly defined set of recommendation rules</p> <p>Constraint Based Case Based Content Based User Based</p>

346	III	106	<p>Why are recommendation engines becoming popular?</p> <p>Users have lesser time, more options and face an information overload</p> <p>It is mandatory to have recommendation engine as per telecom rules</p> <p>It is better to recommend than ask user to search on mobile phones</p> <p>Users don't know what they want</p>
347	III	107	<p>What kind of information does a Recommendation Engine need for effective recommendations?</p> <p>Users' explicit interactions such as information about their past activity, ratings, reviews</p> <p>Users' implicit interactions such as device they use for access, clicks on a link, location, and dates</p> <p>Other information about profile, such as gender, age, or income levels</p> <p>All of the above</p>
348	III	108	<p>What are the challenges in Content Based Filtering?</p> <p>Need to capture significant amount of users' information, which may lead to regulatory and pricing issues</p> <p>Need to have information of all users across different demographics</p> <p>Need to have lower number of categories for content based filtering to be effective</p> <p>Need to have user's social media and digital footprint</p>
349	III	109	<p>Which of the following are not used for filtering in a Recommendation Engine?</p> <p>Sine Similarity</p> <p>Cosine Similarity</p> <p>Jaccard Similarity</p> <p>Euclidean Distance</p>
350	III	110	<p>For an ecommerce website, which of the following is explicit data?</p> <p>Order history</p> <p>Page views</p> <p>Cart events</p> <p>Product feedback</p>
351	III	111	What are the advantages of content based filtering?
352	III	112	What are the disadvantages of content based filtering?
353	III	113	<p>Which of the following refers to the problem of finding abstracted patterns (or structures) in the unlabeled data?</p> <p>Supervised learning</p> <p>Unsupervised learning</p> <p>Hybrid learning</p> <p>Reinforcement learning</p>
354	III	114	Why clustering is important in real life?
355	III	115	What is the difference between clustering and classification?
356	III	116	When might an organization use clustering?
357	III	117	What are the similarity measures of clustering?
358	III	118	What is a similarity matrix used for?
359	III	119	Which is most commonly used measure of similarity in clustering exercise?
360	III	120	What is dissimilarity matrix?
361	IV	1	Explain Neural Network?
362	IV	2	Explain Biological Neural Network and Artificial Neural network?
363	IV	3	What do you mean by combination function in Neural Network?

364	IV	4	How can we count the layers in a neural network?
365	IV	5	Explain cases and variables?
366	IV	6	Define Activation function?
367	IV	7	Explain Loss function?
368	IV	8	Explain different kinds of Kohonen networks?
369	IV	9	Is ANN similar to standard computers? Explain?
370	IV	10	How is ANN useful in making a machine intelligent?
371	IV	11	How can we help artificial neurons in learning?
372	IV	12	What is the use of Artificial Neural Network?
373	IV	13	What is bias?
374	IV	14	Define Error function?
375	IV	15	Explain Supervised Learning
376	IV	16	Which optimization algorithm is best in Neural Network?
377	IV	17	How many types of the artificial neural network used in machine learning?
378	IV	18	Explain the feed-forward neural network?
379	IV	19	What is the convolutional neural network?
380	IV	20	Layers of CNN
381	IV	21	What is the basic concept of Recurrent Neural Network?
382	IV	22	For what RNN is used and achieve the best results?
383	IV	23	What AI techniques use ANN?
384	IV	24	What is ANN classifier?
385	IV	25	What is ANN in machine learning?
386	IV	26	How do you use ANN?
387	IV	27	Types of Artificial Neural Networks Currently Being Used in Machine Learning
388	IV	28	What do you mean by activation function?
389	IV	29	Which are all activation function?
390	IV	30	What is the purpose of activation function?
391	IV	31	What is CNN activation function
392	IV	32	What is epoch in neural network?
393	IV	33	Why activation functions are important in neural network?
394	IV	34	Why is nonlinearity important in neural networks?
395	IV	35	What is difference between Softmax and sigmoid activation functions?
396	IV	36	What is activation value?
397	IV	37	Why do we use nonlinear activation function?
398	IV	38	What is difference between epoch and iteration?
399	IV	39	What is Overfitting in neural network?
400	IV	40	Is more epoch better?
401	IV	41	Why is ReLU used in CNN?
402	IV	42	What is TensorFlow and PyTorch?
403	IV	43	What is linear and non linear activation functions?
404	IV	44	What gives nonlinearity to neural network?
405	IV	45	Is sigmoid a activation function?
406	IV	46	What is leaky ReLU activation and why is it used?
407	IV	47	What is linear activation?

408	IV	48	Which activation function is more prone to vanishing gradient problem?
409	IV	49	What does dropout layer do?
410	IV	50	What is Adam Optimiser?
411	IV	51	What is the first layer in a neural network?
412	IV	52	What is batch size?
413	IV	53	What is a good epoch number?
414	IV	54	What is Max pooling layer in CNN?
415	IV	55	Which activation function is best?
416	IV	56	Why use ReLU vs sigmoid?
417	IV	57	What is a kernel in CNN?
418	IV	58	What is Optimizer in neural network?
419	IV	59	Why activation functions are used in neural networks?
420	IV	60	What is theano and TensorFlow?
421	IV	61	What is GPU in TensorFlow?
422	IV	62	Can activation function be linear?
423	IV	63	What is the difference between PyTorch and TensorFlow?
424	IV	64	What is back-propagation and how is it used in a neural network?
425	IV	65	What is backpropagation and how does it work?
426	IV	66	What is the purpose of backpropagation?
427	IV	67	What backpropagation is usually used for in neural networks?
428	IV	68	What is backward pass in neural network?
429	IV	69	How do you explain backpropagation?
430	IV	70	What are the types of back propagation?
431	IV	71	What is the objective of backpropagation algorithm?
432	IV	72	Why is backpropagation efficient?
433	IV	73	Image result Disadvantages of Back Propagation Algorithm:
434	IV	74	What are the main steps of back propagation algorithm?
435	IV	75	What is the difference between backpropagation and gradient descent?
436	IV	76	What is chain rule in backpropagation?
437	IV	77	What is recurrent backpropagation?
438	IV	78	Does Perceptron require supervised learning?
439	IV	79	Is backpropagation an optimization algorithm?

440	IV	80	Is backpropagation deep learning?
441	IV	81	What is bias in backpropagation?
442	IV	82	What is SGD in deep learning?
443	IV	83	Does backpropagation use gradient descent?
444	IV	84	What is perceptron ML?
445	IV	85	What are weights in neural network?
446	IV	86	What is stochastic backpropagation?
447	IV	87	What is Adam optimization algorithm?
448	IV	88	Why optimizers are used in neural network?
449	IV	89	What is recall in neural network?
450	IV	90	Why is perceptron used?
451	IV	91	What is bias in CNN?
452	IV	92	What are the variants of back propagation?
453	IV	93	How many types of the artificial neural network used in machine learning?
454	IV	94	Explain the feed-forward neural network?
455	IV	95	What is the convolutional neural network?
456	IV	96	What is convergence in neural network?
457	IV	97	What is converge in ML?
458	IV	98	What is convergence in SGD?
459	IV	99	What is backward pass in neural network?
460	IV	100	What is leaky ReLU activation and why is it used?
461	V	1	In confusion matrix FP stands for...
462	V	2	In confusion matrix TP stands for...
463	V	3	In confusion matrix FN stands for...
464	V	4	In confusion matrix TN stands for...
465	V	5	What is precision?
466	V	6	What is recall?
467	V	7	recall=?
468	V	8	precision=?
469	V	9	accuracy=?
470	V	10	Precision is a useful metric in cases where False Positive is a higher concern than False Negatives.
471	V	11	Recall is a useful metric in cases where False Negative trumps False Positive.
472	V	12	ROC curve is plotted between?
473	V	13	PR curve is plotted between?
474	V	14	What Cross Validation technique?
475	V	15	What is validation set?
476	V	16	A lower correlation among ensemble model members will increase
477	V	17	In ensemble learning different learners can come from same algorithm with different hyper parameters
478	V	18	In ensemble learning Different learners can come from different algorithms

479	V	19	In ensemble learning Different learners can come from different training spaces
480	V	20	True or False: Ensemble learning can only be applied to supervised learning methods.
481	V	21	True or False: Ensembles will yield bad results when there is significant diversity among the models.
482	V	22	True or False: Ensemble of classifiers may or may not be more accurate than any of its individual model.
483	V	23	If you use an ensemble of different base models, is it necessary to tune the hyper parameters of all base models to improve the ensemble performance?
484	V	24	It can only be used in classification problem. is true about averaging ensemble?
485	V	25	It can only be used in regression problem. is true about averaging ensemble?
486	V	26	It can be used in both classification as well as regression.is true about averaging ensemble?
487	V	27	We want to give higher weights to better performing models.is true about Weighthed averaging ensemble?
488	V	28	Inferior models can overrule the best model if collective weighted votes for inferior models is higher than best model
489	V	29	A machine learning model is trained on predictions of multiple machine learning models.Is correct statement about stacking?
490	V	30	What are advantages of stacking?
491	V	31	Lower time of execution is the advantage of stacking?
492	V	32	Give an example of an ensemble method
493	V	33	Can Bagging be parallel?
494	V	34	Why Bagging can be parallel?
495	V	35	Bagging helps in reducing overfitting or underfitting?
496	V	36	The aim of bagging is to reduce bias not variance?
497	V	37	True or False: In boosting, individual base learners can be parallel.
498	V	38	True or False: Dropout is computationally expensive technique w.r.t. bagging
499	V	39	Which parameter can be tuned for finding good ensemble model in bagging based algorithms?
500	V	40	In machine learning, an algorithm (or learning algorithm) is said to be unstable if a small change in training data cause the large change in the learned classifiers.
501	V	41	True or False: Bagging of unstable classifiers is a good idea.
502	V	42	Decision Tree is not an example of an ensemble method?
503	V	43	Random Forest is not an example of an ensemble method?
504	V	44	Gradient Boosting is an example of an ensemble method?
505	V	45	Generally, an ensemble method works better, if the individual base models have _____?
506	V	46	Suppose, you are working on a binary classification problem. And there are 3 models each with 70% accuracy.If you want to ensemble these models using majority voting method. What will be the maximum accuracy you can get?

507	V	47	How can we assign the weights to output of different models in an ensemble?
508	V	48	differentiate between test set and validation set.
509	V	49	True negative=correctly rejected
510	V	50	False negative=correctly rejected
511	V	51	False positive=correctly identified
512	V	52	Which is a common error measure?
513	V	53	The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are
514	V	54	“sigmoid” activation function can’t be used at output layer to classify an image ?
515	V	55	“ReLU” activation function can’t be used at output layer to classify an image ?
516	V	56	In the neural network, every parameter can have theirlearning rate.
517	V	57	Sentiment analysis using Deep Learning is a many-to one prediction task
518	V	58	What steps can we take to prevent overfitting in a Neural Network?
519	V	59	What is learning in deep learning?
520	V	60	What is representation in deep learning?
521	V	61	What is hypothesis space in deep learning?
522	V	62	What is deep in deep learning?
523	V	63	What is depth in deep learning?
524	V	64	What is shallow learning in deep learning?
525	V	65	What are neural networks in deep learning?
526	V	66	Deep learning is a mathematical framework for learning representations from data
527	V	67	Deep learning is a biological framework for learning representations from brain Data
528	V	68	Deep learning is an analogue framework for learning representations from data
529	V	69	Deep learning is a digital framework for learning representations from data
530	V	70	What are layers in deep learning?
531	V	71	Learning means finding a set of values for the weights of all layers in a network, such that the network will correctly map example inputs to their associated targets
532	V	72	Is loss function same as objective function?
533	V	73	What is loss function in deep learning?
534	V	74	The fundamental trick in deep learning is to use this score as a feedback signal to adjust the value of the weights a little, in a direction that will higher the loss score
535	V	75	What is training loop in deep learning?
536	V	76	What is decision boundary in classification problems?
537	V	77	What is Random Forest?
538	V	78	How deep learning learns from data?
539	V	79	The two key ideas of deep learning for computer vision?
540	V	80	Three technical forces are driving advances in machine learning
541	V	81	Which is the second-best algorithm for any shallow machine-learning task
542	V	82	Deep learning learns from input data only

543	V	83	What is overfit model?
544	V	84	CNN stands for ?
545	V	85	RNN stands for ?
546	V	86	Practical applications of Semi-Supervised Learning?
547	V	87	What is semi-supervised learning?
548	V	88	What is Continuity Assumption?
549	V	89	What is Cluster Assumption ?
550	V	90	What is Manifold Assumption?
551	V	91	the Google search algorithm uses a variant of Semi-Supervised learning to rank the relevance of a webpage for a given query.
552	V	92	What is internet Content Classification?
553	V	93	Reinforcement learning is-
554	V	94	Reinforcement learning is a unsupervised learning?
555	V	95	Reinforcement learning is a supervised learning?
556	V	96	Which is an application of reinforcement learning?
557	V	97	Upper confidence bound is a
558	V	98	you have a task which is to show relative ads to target users. Which algorithm you should use for this task?
559	V	99	Type1 is known as false positive and Type2 is known as false negative.
560	V	100	Type1 is known as false negative and Type2 is known as false positive.
561	V	101	Type1 error occurs when we reject a null hypothesis when it is actually true.
562	V	102	NLP stands for
563	V	103	Accuracy metric is not a good idea for imbalanced class problems.
564	V	104	Accuracy metric is a good idea for imbalanced class problems.
565	V	105	Precision and recall metrics are good for imbalanced class problems.
566	V	106	Precision and recall metrics aren't good for imbalanced class problems.
567	V	107	Increase in K will result in higher time required to cross validate the result.
568	V	108values of K will result in higher confidence on the cross-validation result as compared to lower value of K.
569	V	109	If $K=N$, then it is called.....cross validation, where N is the number of observations.
570	V	110	What is inductive machine learning?
571	V	111	What is Model Selection in Machine Learning?
572	V	112	What is ensemble learning?
573	V	113	Why ensemble learning is used?
574	V	114	When to use ensemble learning?
575	V	115	What are the two paradigms of ensemble methods?
576	V	116	What is an Incremental Learning algorithm in ensemble?
577	V	117	What is bias-variance decomposition of classification error in ensemble method?
578	V	118	What is sequence learning?
579	V	119	Give name of component of relational evaluation techniques?
580	V	120	Give a popular application of machine learning that you see on day to day basis?