

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)
MCA DEGREE EXAMINATION DECEMBER 2024
(Third Semester)

Branch – **COMPUTER APPLICATIONS**
PYTHON FOR MACHINE LEARNING

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** marks

(10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	Find the open source tool used for Data Mining. a) Rapid miner b) Apple c) 7 point d) MangoDB	K1	CO1
	2	Extend a machine learning model has been run on clean data does not mean that the work of the ____ is done and dusted. a) data scientist b) data analytics c) deep learning d) data science	K2	CO1
2	3	What an important fact to remember about ____ is that it is a dynamically typed language. a) Python b) Data type c) Strings d) variables	K1	CO2
	4	Explain ____ enables us to carry out computational tasks with data in a very effective way. a) linear algebra b) non - linear c) algebra d) non- linear algebra	K2	CO2
3	5	Choose the field of ____ is thus interested in the systematic detection of regularities in a dataset, based on the use of algorithms. a) pattern recognition b) pattern evaluation c) pattern matching d) pattern letters	K1	CO3
	6	Demonstrate a common way to create new features is via ____ that make the variables suitable for exploitation by a particular algorithm. a) mathematical transformations b) transformations c) mathematical fuzzy d) fuzzy set	K2	CO3
4	7	____ is a data exploration technique. a) supervised b) un-supervised c) k-means d) clustering	K1	CO4
	8	____ similarity is given by the distance between data instances. a) cross-validation b) distance measure c) bias low d) KNN	K2	CO4
5	9	The result of hierarchical clustering are represented by a ____. a) Tree b) Mapping c) Decision –making d) dendrogram	K1	CO5
	10	____ minimises the generalisation error associated with the geometrical notion of a margin. a) Bagging b) Boosting c) Random forests d) support vector machine	K2	CO5

Cont...

SECTION - B (35 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Spell out the objective of the data science team.	K1	CO1
		(OR)		
	11.b.	Outline the data science workflow in details.	K2	
2	12.a.	Find out the python basic types.	K1	CO2
		(OR)		
	12.b.	Simplify the linear algebra in machine learning.	K4	
3	13.a.	Explain the recognizing of patterns in machine learning.	K2	CO3
		(OR)		
	13.b.	Classify the k-fold cross validation of machine learning.	K4	
4	14.a.	Identify the k-means in the action of clustering.	K3	CO4
		(OR)		
	14.b.	Explain the Naïve Bayes in Action of classification.	K2	
5	15.a.	Explain the hierarchical clustering.	K2	CO5
		(OR)		
	15.b.	Justify the SVM in action of classification.	K5	

SECTION - C (30 Marks)

Answer ANY THREE questions

ALL questions carry EQUAL Marks

(3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	List the characteristics of a data scientist.	K1	CO1
2	17	Classify the control flow functions and scripts.	K2	CO2
3	18	Explain the machine learning and data science in the process of feature selection.	K2	CO3
4	19	Analyze the clustering techniques with k-means algorithm.	K4	CO4
5	20	Build the SVM and kernel methods in details.	K6	CO5

Z-Z-Z

END