

Q. 05 Which of the following methods do we use to best fit the data in logistic regression -

- a) Least square error b) Maximum likelihood
c) Jaccard distance d) Both (a) and (b)

लॉजिस्टिक रिग्रेशन में डेटा को सर्वोत्तम रूप से फिट करने के लिये हम निम्नलिखित किस विधि का प्रयोग करते हैं -

- a) कम से कम वर्ग त्रुटि b) अधिकतम संभावना
c) जैककार्ड दूरी d) (a) और (b) दोनों

[Section - B]

This section contains **Short Answer Type Questions**. Each question carries **5 Marks**.

इस खण्ड में लघुउत्तरीय प्रश्न हैं। प्रत्येक प्रश्न 5 अंकों का है।

Q. 1 What do you mean by Data Science ? What are the different kinds of data in Data Science.

डाटा साईंस से आप क्या समझते हैं ? डेटा साईंस में विभिन्न प्रकार के डेटा क्या है।

OR

Explain in detail about mean, median and mode measure of central tendency.

केन्द्रीय प्रवृत्ति के माध्य, माध्यिका और बहुलक माप के बारे में विस्तार से समझाएं।

Q. 2 What are the different objects of R programming language ?

R प्रोग्रामिंग लैंग्वेज के विभिन्न ऑब्जेक्ट क्या है ?

OR

How do we store tabular data in R ?

हम R में सारणीबद्ध डेटा कैसे स्टोर करते हैं ?

Q.3 What are the steps of data cleaning in Data Science ?

डेटा साईंस में डेटा क्लीनिंग के चरण क्या हैं ?

OR

What do you mean by Outliers ? Explain with an example.

आउटलायर्स से आप क्या समझते हैं ? एक उदाहरण के साथ समझाएं।

Q. 4 Explain machine learning in details ? What are its applications ?

मशीन लर्निंग के बारे में विस्तार से बताएं ? इसके अनुप्रयोग क्या हैं ?

OR

Explain feature selection method in machine learning ?

मशीन लर्निंग में फीचर सिलेक्शन मेथड समझाइये । ?

Q. 5 What is the difference between logistic regression and linear regression give an example ?

लॉजिस्टिक रिग्रेशन और लीनियर रिग्रेशन में क्या अंतर है, एक उदाहरण देकर समझाइये ?

OR

Explain Classification Algorithm ? What are the different types of classifications ?

वर्गीकरण एल्गोरिथम के बारे में बताएं ? विभिन्न प्रकार के वर्गीकरण क्या है ?

[Section - C]

This section contains **Essay Type Questions**. Each question carries **11 marks**.

इस खण्ड में दीर्घउत्तरीय प्रश्न हैं। प्रत्येक प्रश्न 11 अंकों का है।

Q. 6 Calculate the mean, variance and standard deviation for the following data -

2, 7, 3, 12, 9

निम्नलिखित आंकड़ों के लिये माध्य, प्रसरण और मानक विचलन की गणना कीजिये -

2, 7, 3, 12, 9

OR

Explain about Pie chart, Bar graph and Pareto chart. Why do we need these graphical representation ?

पाई चार्ट, बार ग्राफ और पैरेटो चार्ट के बारे में बताइये ? हमें इन चित्रमय प्रतिनिधित्व की आवश्यकता क्यों है ?

Q. 7 What are the functions of reading and writing data into R.

R में डेटा पढ़ने और लिखने के लिये किन फंक्शन का इस्तेमाल होता है।

OR

What are the different control structures used in R programming.

R में उपयोग की जाने वाली विभिन्न नियंत्रण संरचनाएं क्या है।

Q. 8 What is Feature Engineering ? What are the main steps of feature engineering. Explain in details.

फीचर इंजीनियरिंग क्या है ? फीचर इंजीनियरिंग के मुख्य चरण क्या है ? विस्तृत में बताएं।

OR

P.T.O.

How to remove missing values and repeated values from a data set ? Explain कैसे एक डाटा सेट से लापता मूल्यों और दोहराया मूल्यों को खोजते हैं ? विस्तार से समझाइये।

Q. 9 The sales of a company (in million dollars) for each year are shown in the table below -

X (year)	2005	2006	2007	2008	2009
Y (year)	12	19	29	37	45

- Find the least square regression line $y = a x + b$
 - Use the least square regression line as a model to estimate the sales of the company in 2012.
- प्रत्येक वर्ष के लिये एक कम्पनी की बिक्री (मिलियन डॉलर में) नीचे दी गई तालिका दिखाई गई है -

X (वर्ष)	2005	2006	2007	2008	2009
Y (वर्ष)	12	19	29	37	45

- सबसे छोटी वर्ग समाश्रयण रेखा $y = a x + b$ ज्ञात कीजिये।
- 2012 में कम्पनी की बिक्री का अनुमान लगाने के लिये मॉडल के रूप में कम वर्ग प्रतिगमन रेखा का उपयोग करें।

OR

Explain in details about K-Nearest Neighbors Algorithm (KNN) ?

K निकटतम पड़ोसी (KNN) एल्गोरिद्म के बारे में विस्तार से समझाइये ?

Q. 10 What is decision tree algorithm for classification ? What are the advantages of decision tree ?

वर्गीकरण के लिये डिसिज़न ट्री एल्गोरिद्म क्या है ? डिसिज़न ट्री के फायदे क्या हैं ?

OR

What is Regularization in logistic regression ? Explain with an example.

लॉजिस्टिक रिग्रेशन में नियमितीकरण क्या है ? उदाहरण देकर समझाइये।