

Roll No

CD-404 (GS)**B.Tech. IV Semester**

Examination, June 2023

Grading System (GS)**Introduction to Data Science***Time : Three Hours**Maximum Marks : 70***Note:** i) Attempt any five questions.

किन्हीं पाँच प्रश्नों को हल कीजिए।

ii) All questions carry equal marks.

सभी प्रश्नों के समान अंक हैं।

iii) In case of any doubt or dispute the English version question should be treated as final.

किसी भी प्रकार के संदेह अथवा विवाद की स्थिति में अंग्रेजी भाषा के प्रश्न को अंतिम माना जायेगा।

1. a) Explain the following terms with suitable examples: 7

i) Reading and Displaying of dataset

ii) Rearranging and Ranking of data.

उपयुक्त उदाहरणों के साथ निम्नलिखित पदों की व्याख्या कीजिए।

i) डाटा सेट को पढ़ना और प्रदर्शित करना

ii) डाटा को पुनर्व्यवस्थित करना और रैंकिंग करना

b) Explain Exploratory Data Analysis and discuss its role in data science. 7

अन्वेषणात्मक डाटा विश्लेषण की व्याख्या करें और डाटा विज्ञान में इसकी भूमिका पर चर्चा करें।

2. a) Define Data Security. Why is Data Security Important? List the different issues and challenges in Data Security. 7

डाटा सुरक्षा को परिभाषित करें। डाटा सुरक्षा क्यों महत्वपूर्ण है? डाटा सुरक्षा में विभिन्न मुद्दों और चुनौतियों की सूची बनाएं।

b) What is the purpose of data discretization explain its process. 7

डाटा डिसक्रेटाइजेशन का उद्देश्य क्या है इसकी प्रक्रिया की व्याख्या कीजिए।

3. a) Why is Data Cleaning is so Important? Explain Data Cleaning Process in detail. 7

डाटा क्लीनिंग इतनी महत्वपूर्ण क्यों है? डाटा क्लीनिंग प्रोसेस को विस्तार से समझाइए।

b) Explain dimensionality reduction and numerosity reduction in detail. 7

डायमेंशनलिटी रिडक्शन और न्यूमेरोसिटी रिडक्शन को विस्तार से समझाइए।

4. a) Write about Method in Descriptive Statistics and Explain the role of inferential statistics for Analysis for various datasets. 7

वर्णनात्मक सांख्यिकी में विधि के बारे में लिखें और विभिन्न डाटासेट के विश्लेषण के लिए अनुमानित सांख्यिकी की भूमिका की व्याख्या करें।

b) Elaborate how heatmap is applied in Data Science project. 7

डाटा साइंस प्रोजेक्ट में हीटमैप को कैसे लागू किया जाता है, इसका विस्तार से वर्णन करें।

5. a) Write short note on : 7
 i) The significance of kurtosis
 ii) Different measures of kurtosis
 संक्षिप्त टिप्पणी लिखिए।
 i) कर्टोसिस का महत्व
 ii) कर्टोसिस के विभिन्न पैमाने
- b) Illustrate and describe any two regression models used in Data analysis. 7
 डाटा विश्लेषण में उपयोग किए जाने वाली किन्हीं दो प्रतिगमन मॉडलों का चित्रण और वर्णन करें।
6. a) Explain polynomial regression model. Also compare it with linear regression models. 7
 बहुपद प्रतिगमन मॉडल की व्याख्या करें। इसकी तुलना रैखिक प्रतिगमन मॉडल से करें।
- b) How is model evaluation done by using Data Visualization? Explain with an example. <https://www.rgpvonline.com> 7
 डाटा विजुअलाइज़ेशन का उपयोग करके मॉडल मुल्यांकन कैसे किया जाता है? उदाहरण सहित स्पष्ट कीजिए।
7. a) Explain how can over fitting and under fitting issues are handled in Regression modeling. 7
 प्रतिगमन मॉडलिंग में ओवरफिटिंग और अंडरफिटिंग मुद्दों को कैसे संभाला जा सकता है, इसकी व्याख्या करें।
- b) Explain the use of Grid Search in testing Multiple parameters. 7
 मल्टीपल पैरामीटर्स की टेस्टिंग में ग़्रिड सर्च के उपयोग की व्याख्या करें।

8. Write a short note on : (any two) 7 + 7 = 14
 a) Data Reduction
 b) Polynomial Regression and Pipelines
 c) ANOVA
 एक संक्षिप्त टिप्पणी लिखिए। (किन्हीं दो पर)
 अ) डाटा कटौती
 ब) बहुपद प्रतिगमन और पाइपलाइन
 स) एनोवा

<https://www.rgpvonline.com>
 Whatsapp @ 9300930012
 Send your old paper & get 10/-
 अपने पुराने पेपर्स भेजे और 10 रुपये पायें,
 Paytm or Google Pay से