(b) Explain different types of data used in Data Science. Discuss their significance in analysis and modeling. 7

डेटा विज्ञान में उपयोग किए जाने वाले विभिन्न प्रकार के डेटा की व्याख्या कीजिए । विश्लेषण और मॉडलिंग में इनके महत्व पर चर्चा कीजिए।

Unit II (इकाई II)

4. Explain the need for encoding categorical variables in data preprocessing. Discuss various techniques to perform Categorical Data Encoding using examples. **14**

डेटा प्रीप्रोसेसिंग में श्रेणीबद्ध चरों को एन्कोड करने की आवश्यकता समझाइए । उदाहरणों के साथ श्रेणीबद्ध डेटा एन्कोडिंग करने की विभिन्न तकनीकों पर चर्चा कीजिए ।

5. (a) Explain Exploratory Data Analysis and its objectives. How it helps in understanding patterns, trends, and anomalies in datasets ? Discuss. 7

No. of Printed Pages : 07 Roll No. .....

55691

M.C.A. (NEP 2020) EXAMINATION, 2026

(Fourth Semester)

DATA SCIENCE

Time : 3 *Hours*] [*Maximum Marks :* 70

Before answering the question-paper, candidates must ensure that they have been supplied with correct and complete question-paper. No complaint, in this regard will be entertained after the examination.

Note : Attempt *Five* questions in all, selecting *one* question from each Unit. Q. No. **1** is compulsory.

प्रत्येक इकाई में से **एक** प्रश्न चुनते हुए, कुल **पाँच** प्रश्नों के उत्तर दीजिए । प्रश्न क्र. **1** अनिवार्य है ।

Compulsory Question (अनिवार्य प्रश्न)

1. (a) Discuss the advantages of using Python in Data Science and its ecosystem for data analysis and scientific computing. **3.5**

डेटा विज्ञान में पायथन के उपयोग के लाभों और डेटा विश्लेषण तथा वैज्ञानिक कंप्यूटिंग के लिए इसके इकोसिस्टम पर चर्चा कीजिए ।

(b) Differentiate between normalization and standardization and discuss their importance in data preprocessing. **3.5**

सामान्यीकरण और मानकीकरण के बीच अंतर स्पष्ट कीजिए और डेटा प्रीप्रोसेसिंग में इनके महत्व पर चर्चा कीजिए ।

(c) What is the difference between null and alternative hypotheses ? Explain. **3.5**

शून्य और वैकल्पिक परिकल्पना में क्या अंतर है ? समझाइए ।

(d) Explain the features of Plotly for interactive data visualization. **3.5**

इंटरैक्टिव डेटा विजुअलाइजेशन के लिए प्लॉटली की विशेषताएं समझाइए ।

Unit I (इकाई I)

2. (a) Explain Data Science and its major components. Discuss the importance of data science in extracting meaningful insights from large volumes of data. **8**

डेटा विज्ञान और इसके प्रमुख घटकों की व्याख्या कीजिए । बड़ी मात्रा में डेटा से सार्थक अंतर्दृष्टि निकालने में डेटा विज्ञान के महत्व पर चर्चा कीजिए ।

(b) Discuss the applications of Data Science in various domains. **6**

विभिन्न डोमेनों में डेटा विज्ञान के अनुप्रयोगों पर चर्चा कीजिए ।

3. (a) Discuss different methods of data collection and explain various sources of data. **7**

डेटा संग्रहण की विभिन्न विधियों पर चर्चा कीजिए और डेटा के विभिन्न स्रोतों की व्याख्या कीजिए ।

एक्सप्लोरेटरी डेटा एनालिसिस और इसके उद्देश्यों की व्याख्या कीजिए । यह डेटासेट में पैटर्न, रुझान और विसंगतियों को समझने में कैसे मदद करता है ? चर्चा कीजिएं ।

(b) Discuss different types of correlation and methods used for correlation analysis. **7**

सहसंबंध के विभिन्न प्रकारों और सहसंबंध विश्लेषण के लिए उपयोग की जाने वाली विधियों पर चर्चा कीजिए ।

Unit III (इकाई III)

6. Describe logistic regression and decision tree algorithms used in supervised learning. Compare their working principles, advantages, and applications in classification problems. **14**

सुपरवाइज्ड लर्निंग में उपयोग किए जाने वाले लॉजिस्टिक रिग्रेशन और डिसीजन ट्री एल्गोरिदम का वर्णन कीजिए । वर्गीकरण समस्याओं में इनके कार्य सिद्धांतों, लाभों और अनुप्रयोगों की तुलना कीजिए ।

7. Explain unsupervised learning and discuss clustering techniques & dimensionality reduction methods along with their role in pattern discovery and data visualization. **14**

अनसुपरवाइज्ड लर्निंग की व्याख्या कीजिए और क्लस्टरिंग तकनीकों तथा डायमेंशनलिटी रिडक्शन विधियों पर चर्चा कीजिए, साथ ही पैटर्न खोज और डेटा विजुअलाइजेशन में इनकी भूमिका बताइए ।

Unit IV (इकाई IV)

8. (a) Explain the principles of effective data visualization. **6**

प्रभावी डेटा विजुअलाइजेशन के सिद्धांतों की व्याख्या कीजिए ।

(b) Describe the role of Seaborn in data visualization. Explain how Seaborn simplifies statistical plotting and enhances graphical representation. **8**

डेटा विजुअलाइजेशन में सीबॉर्न की भूमिका का वर्णन कीजिए । समझाइए कि सीबॉर्न सांख्यिकीय प्लॉटिंग को कैसे सरल बनाता है और ग्राफिकल प्रतिनिधित्व को कैसे बेहतर करता है ।

9. (a) Discuss the importance of data privacy in Data Science applications. Explain algorithmic bias, its causes, and its impact on fairness and decision-making systems. **8**

डेटा विज्ञान अनुप्रयोगों में डेटा गोपनीयता के महत्व पर चर्चा कीजिए । एल्गोरिदमिक पूर्वाग्रह, इसके कारणों और निष्पक्षता तथा निर्णय लेने वाली प्रणालियों पर इसके प्रभाव की व्याख्या कीजिए ।

(b) Explain responsible AI and its significance in modern data-driven systems. **6**

जिम्मेदार एआई और आधुनिक डेटा-संचालित प्रणालियों में इसके महत्व की व्याख्या कीजिए ।

❖❖❖❖❖❖❖