

مولانا آزاد نیشنل اردو یونیورسٹی

Semester Examination April 2026

Programme: M.Tech

پروگرام: ایم ٹیک

Sretseme: 2nd

سیسٹر:

Title & Paper Code: Big Data Analytics (MTCS246PET)

مضمون مع کوڈ:

Time: 3 Hours وقت:

Maximum Marks 70 جملہ نشانہات:

ہدایات:

یہ پرچہ تین حصوں پر مشتمل ہے: حصہ اول، حصہ دوم، حصہ سوم۔ ہر جواب کے لئے لفظوں کی تعداد شمار ہے۔ تمام حصوں سے سوالوں کا جواب دینا لازمی ہے۔

1. حصہ اول میں 10 لازمی سوالات ہیں جو کہ مختصر جواب والے سوالات ہیں۔ ہر سوال کا جواب دینا لازمی ہے ہر سوال کے لئے 1 نمبر مختص ہے۔
(10 x 1 = 10 Mskra)

2. حصہ دوم میں 8 سوالات ہیں، اس میں طالب علم کو کوئی پانچ سوالوں کے جواب دینے ہیں ہر سوال کا جواب تقریباً دو سو (200) لفظوں پر مشتمل ہے ہر سوال کے لئے 6 نمبرات مختص ہیں
(5 x 6 = 30 Mskra)

3. حصہ سوم میں پانچ سوالات ہیں۔ اس میں سے طالب علم کو کوئی تین سوالوں کے جواب دینے ہیں۔ ہر سوال کا جواب تقریباً پانچ سو (500) لفظوں پر مشتمل ہے۔ ہر سوال کے لئے 10 نمبرات مختص ہیں۔

(Marks 30 = 10 x 3)

حصہ اول

سوال 1

(1) RDD کیا ہیں RDD کے Operations کی وضاحت کریں۔

(2) اس تمام Big Data کا Analyze کرنے کے لیے، Hive ایک tool ہے جسے تیار کیا گیا ہے۔

A. True

B. False

(3) Master کے طور پر کام کرتا ہے، اور ہر cluster میں صرف ایک NameNode ہوتا ہے

A. Data Node

B. Block Size

C. Data block

D. NameNode

(4) HDFS اور Google File System میں فرق کریں۔

(5) Structured Data اور Unstructured Data میں فرق کریں۔

- (6) Sqoop اور Floom میں فرق کریں
- (7) Oozie کی تعریف کریں۔
- (8) Hadoop کے Components کیا ہیں۔
- (9) Big Data میں V5 کیا ہیں؟
- (10) CAP Theorem کا Concept تصور Big Data System میں کیوں Relevant ہے

حصہ دوم

- سوال (2) Big Data کیا ہے؟ ہمیں Big Data کا Analyse کرنے کی ضرورت کیوں ہے؟ Big Data کو ہینڈل کرنے میں Big Data اور Challenge کی Characteristics کی فہرست بنائیں
- سوال (3) New SQL کیا ہے؟ اس کی خصوصیات کیا ہیں۔ مناسب مثال کے ساتھ SQL، NOSQL اور New SQL میں فرق کریں۔
- سوال (4) Data Analytics کیا ہے؟ یہ Data Analysis سے کیوں مختلف ہے؟ م مثال کے ساتھ Data Analytics کی مختلف اقسام کی وضاحت کریں
- سوال (5) Big Data کی تین Dimension کے بارے میں وضاحت کریں۔
- سوال (6) Data کی Nature اور اس کے Applications کے بارے میں تفصیل سے وضاحت کریں۔
- سوال (7) Data Analytics Life cycle کے مراحل کا خاکہ بنائیں، Discovery سے لے کر Model Operation تک۔ Smart city Application کے لیے Case Study فراہم کریں۔
- سوال (8) Big Data میں Storage Considerations کے بارے میں تفصیل سے وضاحت کریں۔
- سوال (9) Hadoop کیا ہے؟ Hadoop Framework کے کیا Requirements ہیں۔ وضاحت کریں کہ Hadoop، Hardware کی ناکامیوں کو کیسے ہینڈل کرتا ہے۔ خاص طور پر، Data Replication اور Rack Awareness Policy پر تبادلہ خیال کریں۔

حصہ سوم

- سوال (10) CAP Theorem کیا ہے یہ ACID Properties سے مختلف کیوں ہے CAP Theorem کا Concept تصور Big Data System میں کیوں Relevant ہے وضاحت کریں
- سوال (11) "Word Count" کی مثال کا استعمال کرتے ہوئے MapReduce جاب (Map, Shuffle, Sort, and Reduce) کے مراحل کی وضاحت کریں۔
- سوال (12) MapReduce کیا ہے؟ MapReduce کا استعمال کرتے ہوئے Hadoop کے ساتھ Data Processing کی وضاحت کریں۔
- سوال (13) YARN کیا ہے؟ YARN کے Components کیا ہیں؟ YARN کے Architecture کی وضاحت کریں
- سوال (14) NOSQL ڈیٹا بیس کیا ہے؟ یہ RDBMS سے مختلف کیوں ہے؟ مناسب مثال کے ساتھ Document Oriented اور Key Value Store کی وضاحت کریں۔