|  |  |
| --- | --- |
| **SESSION** | **July 2023** |
| **PROGRAM** | **MASTER OF BUSINESS ADMINISTRATION (MBA)** |
| **SEMESTER** | **IV** |
| **course CODE & NAME** | **DADS402 – Unstructured data analysis** |
| **CREDITS** | **04** |
| **nUMBER OF ASSIGNMENTS & Marks** | **02****30 MARKS EACH** |

Assignment Set – 1

**1 (a) Define unstructured data and give three real-world examples of where it can be found.**

 **(b) Differentiate between structured and unstructured data. Highlight three key differences.**

Ans -

1 (a) Unstructured data refers to information that does not have a predefined data model or format, making it challenging to organize, process, and analyze using traditional data management tools. It lacks a specific structure or schema, which means it doesn't fit neatly into rows and columns like structured data. Here are three real-world examples of unstructured data:

1. **Textual Data**: Large volumes of unstructured text can be found in documents, emails, social media posts, news articles, and customer reviews. This data lacks a standardized format and

Its Half solved only

Buy Complete from our online store

<https://smuassignment.in/online-store/>

MUJ Fully solved assignment available for**session July 2023.**

Lowest price guarantee with quality.

Charges**INR 200 only per assignment.**For more information you can get via mail or Whats app also

Mail id is aapkieducation@gmail.com

Our website www.smuassignment.in

After mail, we will reply you instant or maximum

1 hour.

Otherwise you can also contact on our

whatsapp no 8791490301.

**2 (a) Describe the process of feature extraction in textual data. Why is this step crucial in data analysis?**

 **(b) Discuss the role of pictorial data in unstructured data analysis. Give two scenarios where pictorial data analysis can be beneficial.**

**Ans 2 (a)**

Feature extraction is a crucial step in the analysis of textual data, as it involves transforming raw text into numerical or categorical features that can be used for various machine learning or statistical analysis tasks. Here's a step-by-step description of the process:

1. **Tokenization**: The first step is to break down the text into smaller units, typically words or tokens

**3 Describe different sentiment analysis models. How can one perform sentiment analysis using R?**

**Ans 3.**

Sentiment analysis is a natural language processing (NLP) task that involves determining the sentiment or emotional tone expressed in a piece of text, such as a review, tweet, or comment. There are various sentiment analysis models, and I'll describe three common ones: Rule-based, Machine Learning, and Deep Learning. Additionally, I'll provide a brief overview of how to perform sentiment analysis using R.

1. **Rule-Based Sentiment Analysis:**

**Assignment Set – 2**

**4 Explain the differences between SQL and NoSQL databases. Why might a business choose a NoSQL database for certain applications?**

**Ans 4.**

SQL (Structured Query Language) and NoSQL (Not Only SQL) databases are two broad categories of database management systems, each with its own set of characteristics and use cases. Here are the key differences between SQL and NoSQL databases, along with reasons why a business might choose a

**5 Briefly describe three types of NoSQL databases and provide a case scenario for each.**

**Ans 5.**

**NoSQL Databases**

NoSQL (Not Only SQL) databases have emerged as an alternative to traditional relational databases, especially in scenarios where scale, speed, and schema flexibility are critical. They are particularly suitable for dealing with vast amounts of data and for applications that require agile development and quick iterations. Here, we will describe three types of NoSQL databases and provide a

**6 (a) Discuss the significance of audio data in today's digital age. How is it generally processed for analytical purposes?**

**6 (b) Introduce the concept of video classification. What are the main challenges and techniques used in video data classification? 5+5**

**Ans 6 (a)**

**Significance of Audio Data in Today's Digital Age and Its Analytical Processing**

Audio data has carved a significant niche in today's digital age. With the proliferation of smart devices and platforms like podcasts, virtual assistants, and music streaming services, audio has become a staple in daily digital consumption. It holds importance in numerous sectors: from