|  |  |
| --- | --- |
| **SESSION** | **JUL - AUG 2024** |
| **PROGRAM** | **MASTER OF BUSINESS ADMINISTRATION (MBA)** |
| **SEMESTER** | **IV** |
| **COURSE CODE & NAME** | **DADS402 UNSTRUCTURED DATA ANALYSIS** |
|  |  |
|  |  |

**Assignment Set – 1**

**1. (a) List down a few differences between structured and unstructured data.**

**(b) Explain the application of NLP and Taxonomies.**

**Ans 1.**

**Differences Between Structured and Unstructured Data**

Structured and unstructured data are two distinct forms of data, each with unique characteristics and applications.

Structured data is highly organized and adheres to a predefined schema. It is stored in tabular formats such as rows and columns within relational databases, making it easy to search, manage, and analyze using SQL and other query languages. For example, a database of customer information that includes fields like name, age, email, and purchase history is structured data. This type of data is typically numeric or text-based and is ideal for use cases requiring rigorous analysis, such as financial reporting, inventory tracking, and customer

Its Half solved only

Buy Complete assignment from us

**Price – 190/ assignment**

**MUJ Manipal University Complete SolvedAssignments session JULY-AUG 2024**

buy cheap assignment help online from us easily

we are here to help you with the best and cheap help

**Contact No – 8791514139 (WhatsApp)**

**OR**

**Mail us-** **bestassignment247@gmail.com**

**Our website -** [**www.assignmentsupport.in**](http://www.assignmentsupport.in)

**2. (a) What is a word cloud? Discuss some libraries that you need to import to create a word cloud in python?**

**(b) Demonstrate some common business applications of text classification.**

**Ans 2.**

**Word Cloud and Python Libraries for Its Creation**

**Word Cloud**

A word cloud is a popular visualization tool that highlights the most frequent words in a text corpus by displaying them in varying font sizes. Larger fonts represent higher word frequencies or importance, while smaller fonts denote less frequent terms. Word clouds are particularly useful for summarizing unstructured text data, such as customer reviews, speeches, or social media posts, as they offer a quick and engaging way to grasp the key themes or topics.

Word clouds are often used in marketing and social media analytics to identify trending keywords, in educational contexts to summarize lecture notes, and in business reports to present

**3. (a) How do you perform sentiment analysis using python?**

**(b) What is latent dirichlet allocation (LDA)?**

**Ans 3.**

**(a) Performing Sentiment Analysis Using Python**

Sentiment analysis, also known as opinion mining, is a process of determining the emotional tone behind a piece of text. It is commonly used to analyze social media posts, customer reviews, and other forms of unstructured text to gauge opinions or sentiments as positive, negative, or neutral. In Python, sentiment analysis can be performed using libraries such as **TextBlob**, **VADER (Valence Aware Dictionary and sEntiment Reasoner)**, and **NLTK (Natural Language Toolkit)**.

**Assignment Set – 2**

**4. (a) How NoSQL databases different from relational databases?**

**(b) What is the main feature of MongoDB that sets it apart from relational databases?**

**Ans 4.**

**(a) Differences Between NoSQL and Relational Databases**

Relational databases and NoSQL databases represent two fundamentally different approaches to data storage and management.

Relational databases, such as MySQL, PostgreSQL, and Oracle, are based on a tabular structure where data is stored in rows and columns. These databases use a predefined schema, which enforces strict rules about data types and relationships between tables. Relational databases rely on Structured Query Language (SQL) for data manipulation and querying. They are ideal

**5. (a) How can you visualize an audio signal?**

**(b) What is Acoustic Data Classification?**

**Ans 5.**

**(a) Visualizing an Audio Signal**

Audio signal visualization involves graphically representing sound waves to analyze their features, such as amplitude, frequency, and duration. This is essential for understanding the characteristics of audio signals and processing them in various applications, including speech recognition, music analysis, and sound engineering. Audio signals are typically represented as waveforms, spectrograms, or frequency-domain plots.

In Python, libraries like **Matplotlib**, **Librosa**, and **SciPy** are commonly used to visualize audio

**6. (a) How does histogram equalization work?**

**(b) What are the key components of a CNN for image classification?**

**Ans 6.**

 **(a) How Histogram Equalization Works**

Histogram equalization is a technique used in image processing to enhance the contrast of an image by redistributing its pixel intensity values. The goal is to achieve a more uniform histogram, where the intensity levels are spread out across the full range of possible values (e.g., 0 to 255 for 8-bit images). This enhances details in both dark and bright areas of an image, making it visually clearer.

The process begins by calculating the **histogram** of the image, which represents the frequency