

plag report

by Dded Dedded

Submission date: 20-Jun-2021 11:22AM (UTC-0500)

Submission ID: 1609472185

File name: Engineering_Data_Mining_Discussion_Questions_2.edited.docx (18.38K)

Word count: 664

Character count: 3819

Student Name:

Institutional Affiliation:

Professor Name:

Course, Department:

Date

Engineering Data Mining Discussion Questions

Data mining is the process of retrieving meaningful patterns, anomalies, and correlations from data sets within large databases (Suresh & Harshni, 2017). Text mining is a type of data mining that focuses on finding meaningful patterns from text data (Suresh & Harshni, 2017). Sentimental analysis is the process of identifying, classifying, and labeling data (Suresh & Harshni, 2017). Text mining specifically deals with unstructured text files. The data patterns and relationships are essential for decision-makers when making informed decisions towards their organizations (Suresh & Harshni, 2017).

Text mining helps in managing large unstructured and structured text files (Suresh & Harshni, 2017). It manages mainly text data that is generated in large quantities from daily operations and activities. The major difference between data mining and text mining is that data mining is based on text files, while data mining deals with numerical data files (Suresh & Harshni, 2017). Text mining is used for high text generating industries like research centers, lawcourts, finance reports, marketing reports, and technology hubs. Text mining is applicable in extracting relationships and information meaning by using keywords designed in sequences and patterns in those sectors. Text mining also tracks topics from documents, interests, and user profiles (Suresh & Harshni, 2017). In offices, text mining help in clustering, categorizing and matching text patterns based on predefined topics, themes, and patterns to help in answering different queries (Suresh & Harshni, 2017). Using text mining procedures and strategies, it is easy to connect related documents and text when making searches and drawing patterns.

Natural language is a branch of artificial intelligence that helps in text mining by connecting two topics (Deng & Liu, 2018). First, during text mining, the natural language process helps the text mining tools to understand the natural human language. Second, using models and numerical, natural language processing helps in processing the text files to

retrieve meaningful patterns. To fulfill its purpose, natural language uses computer programs and software to conduct data manipulation and processing. Using natural language, text mining helps give structure to information and data collected through classes, categorization, sequences, association, and clusters (Deng & Liu, 2018). However, text mining can be limited by factors like an error in text or speech, lack of research and development, ambiguity in some words, and contextual word phrases.

Exercise three

Human activity over the internet has led to the generation of a lot of data by internet users. Big platforms like eBay are experiencing the real data management challenge from the large suppliers like the social media and CRM platforms (Wang et al., 2021). Despite facing management challenges, data is essential for decision making and strategizing for the future by such companies making it basic need for the organization (Wang et al., 2021). eBay platform uses analytical tools to analyze and draw patterns that make it easy to manage and maintain customers through personalization and customer analysis (Wang et al., 2021). The organization has installed web analytics programs that are able to generate and handle thousands of terabytes of data. The analytics tool has the ability to handle and dig out patterns that are necessary for understanding customers.

Internet exercise 7

From the website, the following applications and software can help in data mining;

Classification software which uses neural network models and decision trees to combine and compare data patterns.

Clustering and segmentation software which is made of features like CMSR Data Miner IBM SPSS modeler, CLUTO, ELKI, and BayesiaLab.

Web Usage Mining Software which is made up of clickTracks, Download analyzer, and A1WebStats.

Reference

- Suresh, R., & Harshni, S. R. (2017, March). Data mining and text mining—a survey. In *2017 International Conference on Computation of Power, Energy Information and Communication (ICCPEIC)* (pp. 412-420). IEEE.
- Deng, L., & Liu, Y. (Eds.). (2018). *Deep learning in natural language processing*. Springer.
- Wang, T., Brovman, Y. M., & Madhvanath, S. (2021). Personalized Embedding-based e-Commerce Recommendations at eBay. *arXiv preprint arXiv:2102.06156*.

plag report

ORIGINALITY REPORT

0%

SIMILARITY INDEX

0%

INTERNET SOURCES

0%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

Exclude quotes On

Exclude bibliography On

Exclude matches Off

plag report

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4
