

SCHOOL OF COMPUTER TECHNOLOGY

EXAMINATION FOR THE BSC (HONS) INFORMATION SYSTEMS; YEAR 3

ACADEMIC SESSION 2011; SEMESTER 8

CTB3308 BUSINESS INTELLIGENCE SYSTEMS

FINAL - JULY 2011 EXAM CYCLE

TIME: 2 HOURS + 10 MINUTES READING TIME

INSTRUCTIONS TO CANDIDATES

This question booklet contains two sections.

Section A: Answer the **compulsory** questions.

Section B: Answer **two** questions out of **three** questions

All answers must be written in the answer booklets provided using blue or black INK

IMPORTANT NOTES TO CANDIDATES

Materials Allowed

Standard Items : Pen, Pencil, Eraser or Correction Fluid, Ruler

Special Items : Non Programmable Calculators

It is your responsibility to ensure that you do **NOT** have in your possession any unauthorised notes or any other means that would improperly help you in your work. If you have any unauthorised materials with you, hand it to the invigilator **BEFORE** reading any further.

DO NOT REMOVE THIS QUESTION PAPER FROM THE EXAMINATION HALL

[This paper contains FOUR questions printed on FOUR pages]

Section A Compulsory section
--

Question 1 (50 marks)**HP and Text Mining**

Hewlett-Packard Company (HP), founded in 1939 by William R. Hewlett and David Packard, is headquartered in Palo Alto, California. The company provides products, technologies, solutions and services to individuals, small- medium-sized businesses and large enterprises on a global basis. HP also offers management software solutions that enable enterprise customers to manage their IT infrastructure, operations, applications, IT services and business processes as well as carrier grade platforms for various applications. Some of HP's popular product categories are commercial and consumer personal computers, workstations, handheld computing devices, inkjet printers, laser printers, digital entertainment systems, calculators and related accessories, software and services, digital photography and entertainment. Retailers form the distribution channel of the company. The company also sells through distribution partners, original equipment manufacturers and systems integrators.

Text Mining

Customers of HP communicate with the company through millions of e-mails. Structured data analysis is effective in finding out parameters such as from whom, when, where and how the messages originated. A wealth of information would be available if mining techniques could find out why these e-mails were sent. One of the common interaction points between the customer and the company is the call center. HP is impressed by the business insights that could be culled from communications such as word documents, e-mail and other sources. The combination of the structured and unstructured data can create a tremendous potential for companies to find valuable business insights.

System

The standard tools that HP previously used could not report useful information from customer-related communications. Now, HP used SAS Institute's Text Miner to uncover analytical insights from customer-related data in call center applications and then standardized those insights. HP implemented Text Miner to combine structure data set that is stored in a Microsoft SQL Server Database with an Analysis Services Online Analytic Processing engine. The system today, encompasses 300,000 text documents and is roughly 50 gigabytes in size, covering an 18month period and 3 call centers. HP implemented Executive Viewer a web-based tool developed by Temtec that enables HP to augment the OLA cubes with predictive modelling, loyalty scores, and customer differentiations created by SAS Enterprise Miner.

Process

Various concepts, such as products used, call frequency and common customer issues are used to aggregate text data; the result is consolidated into probabilistic text clusters. This consolidated cluster is then combined with the structured data from third party providers. HP can now combine and analyze structured data such as revenue with customer desires attitudes and needs.

Text analysis is challenging due to dimensionality and data dispersal. Different customer databases contain different structured information that could be integrated without much difficulty. The challenge lies in combining the structured data with unstructured data from text. SAS Text miner uses a technique called singular value decomposition. Text mining software includes a prebuilt dictionary of words and synonyms lists it is an overwhelming task for the organization to customize the text information generated in its business environments. Text data are available at various sources that are outside the realms of

traditional data warehousing. Some of the largest challenges that SAS Text Miner faces are customer activity at the HP web site as well as finding insights into the businesses of HP's customers.

In addition to the major application of text mining, SAS Text Miner could be used proactively on customer websites to generate insights into the customer needs that HP could satisfy. The tool could also be used to analyze multiple suppliers with various numbers and description in text.

Results

SAS Text Miner was successfully able to develop standard data definitions and product classification models with more than 80 percent accuracy. The system is now being used to support HP in contributing to the top line through improved cross-selling, targeted marketing, customer retention and better anticipation of customer needs. The information generated from structured / unstructured data now supports multiple business users in various departments.

Sources based on M. Hammond, "BI Case Study: What's in a Word? for Hewlett Packard , It's Customer Insight," Business Intelligence Journal, Vol 9, No.3, Summer 2004, pp.48-51

- a) What is the practical application of text mining and explain the importance's of it? (8 marks)
- b) How do you use text mining techniques for in other businesses? (14 marks)
- c) What were HP's challenges in text mining? How did they overcome it? (8 marks)
- d) In what other areas , in your opinion can HP use text mining ? (10 marks)
- e) Give TWO(2) popular application areas of text mining? (10 marks)

Section B

Answer any TWO questions from this section.

Question 2 (25 marks)

Data warehouse is a pool of data produced to support decision making; it is also a repository of current and historical data of potential interest to managers throughout the organization.

- a) Describe the data warehousing process. (7 marks)
- b) Describe the THREE (3) steps of the ETL (Extraction, Transformation , Loading) process. (6marks)
- c) Compare Inmon and Kimball approaches in terms of : (12 marks)
 - a. Complexity of the method
 - b. Overall approach
 - c. Architecture structure
 - d.Data orientation
 - e.End- user accessibility
 - f.Audience

Question 3 (25 marks)

- a) What are the major differences between a scorecard and a dashboard ? (10 marks)
- b) Most organizations collect a wide range of operational metrics. As the name implies, these metrics deal with the operational activities and performance of a company. Whether an operational metric is strategic or not depends on the company and its use of the measure. Explain with examples THREE (3) variety of operational areas covered by this metrics. (15 marks)

Question 4 (25 marks)

Social networking is built on the idea that there is structure to how people know each other and interact. The basic premise is that social networking gives people the power to share, making the world more open and connected.

- a) Facebook's early success was de to the close affiliation of its members' networks. How does Facebook expand into new markets without losing what originally made the site popular and alienating existing users? (8 marks)

Alkem Distribution center is a well known books distributor. The CEO of the company would like to implement business intelligence into their company to accommodate their growing business and retaining their customers. They have hired you as their business intelligence consultant.

- b) How would you advise this company the major types of BI implementation influencing factors and their relation to Alkem Distribution center business model? (15 marks)
- c) Based on the above scenario, list TWO(2) major issues implementing BI? (2 marks)

~ END OF PAPER ~