Course Name	Course Code			
Data Mining & Warehousing	CPIT 440			
Prerequisite	Credit Units			
CPIT 340	Theory	Practice	Training	Credit
	3	1	0	3

اسم المقرر	الرمز			
مستودعات البيانات والتنقيب	تم ٤٤٠			
المتطلبات السابقة	الوحدات الدراسية المعتمدة			
تم ٣٤٠	نظري	عملي	تدريب	معتمد
	٣	١	٠	٣

Course Description:

This course introduces the principles of data mining including database systems, artificial intelligence, data retrieval and statistics. It also covers the necessary tools for knowledge exploration and data storage.

توصيف المقرر :

تهدف هذه المادة إلى تعريف الطالب بمفهوم ومبادئ تنقيب البيانات والتي تشمل العديد من المجالات مثل قواعد البيانات والذكاء الصناعي واسترجاع البيانات والإحصاء. وتهدف أيضا إلى إلمامه بأدوات استكشاف المعرفة وتخزين البيانات

Course Contents:

1. Principles and structures of data warehousing, methods of analysis process (OLTP) and direct (OLAP).
2. Extraction and representation, cleaning and delivery of data.
3. Digging and cutting, data management and data warehouse architecture.
4. Patterns of data warehouses (star - SnowFlake - tower).
5. Ways to represent data warehouses (ROLAP - MOLAP - HOLAP).
6. Performance Optimization and compilation of data warehouses.
7. The importance of clean data in data mining techniques and basic statistics.
8. Techniques for exploring data (tree resolution - neural networks - a market basket analysis).
9. Application areas such as cross-selling – customer retention - detect fraud - pricing - the allocation of sources).
10. Use of software ready to explore the data.

محتويات المقرر:

١. مبادئ تخزين البيانات وهيكلة البيانات، طرق التحليل العملية (OLTP) والمباشرة (OLAP).
٢. استخلاص وتمثيل وتنظيف وإيصال البيانات.
٣. حفر وتقطيع وتكعيب البيانات و إدارة وبنيان مستودعات البيانات.
٤. أنماط مستودعات البيانات (النجمة – ندفة الثلج – البرج) .
٥. طرق تمثيل مستودعات البيانات (ROLAP – MOLAP – HOLAP).
٦. تحسين الأداء وتجميع وتقسيم مستودعات البيانات.
٧. أهمية البيانات النظيفة في التنقيب عن البيانات وتقنيات الإحصاء الأساسية.
٨. تقنيات استكشاف البيانات (شجرة القرار – الشبكات العصبية – تحليل سلة السوق).
٩. مجالات التطبيق مثل البيع المتقاطع – الاحتفاظ بالزبون – كشف الاحتيال – التسعير - تخصيص المصادر).
١٠. استخدام البرمجيات الجاهزة لاستكشاف البيانات.

Course Outcomes:

1. To identify the foundations and techniques for integrating and cleaning of data.
2. To know the role of OLAP techniques and explore data to troubleshoot business information problems.
3. To design and establish data warehouses to troubleshoot the problems of different kinds of information.
4. To learn about the benefits and techniques of accessing data stores through the Web.
5. To know assessment techniques to support exploration and prospecting data (KDD).
6. To apply data mining tools on actual situations.

مخرجات المقرر:

١. تحديد الأسس والتقنيات لدمج وتنظيف البيانات والتركيز على شئون الجودة.
٢. معرفة دور تقنيات OLAP واستكشاف البيانات لحل مشاكل المعلومات التجارية.
٣. تمثيل وتصميم وإنشاء مستودعات البيانات لحل مشاكل المعلومات المختلفة.
٤. معرفة فوائد وتقنيات الوصول إلى مستودعات البيانات عن طريق الويب.
٥. تقييم التقنيات الداعمة لاستكشاف البيانات والتنقيب عنها (KDD)
٦. تطبيق أدوات التنقيب عن البيانات على الحالات الواقعية.

Textbook: **الكتاب المقرر :**

- W. H. Inmon, "Building the Data Warehouse", 3rd Ed., Wiley, 2002, 0471081302.

- A. Berson, S. J. Smith, "Data Warehousing, Data Mining, and OLAP", McGraw-Hill, 1997, ISBN 0070062722.

Extra Reference: **المرجع المساند :**

- R. Kimball, M. Ross, "The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling", 2nd Ed., Wiley, 2002, ISBN 0471200247.

- J. P. Bigus, "Data Mining With Neural Networks: Solving Business Problems from Application Development to Decision Support", McGraw-Hill, 1996, ISBN 0070057796.