

Introduction to Data Mining and its Applications

(with Hands-on Session using WEKA)

(Half Day Tutorial)

Ahmed M. Zeki
(*PhD in Computer Science*)

Faculty of Information and Communication Technology
International Islamic University Malaysia (IIUM)

amzeki@iiu.edu.my

Tel: +603-6196 5657, Fax: +603-6196 5179, Mobile: +6012-6599 649

Introduction	Data Mining (DM) is the process of knowledge discovery or retrieval of hidden information from datasets. DM is becoming an essential instrument for researchers. It comprises of many techniques used for knowledge discovery. DM can process and analyze engineering data, medical data, business data, etc. In a recent study, DM was listed among the 12 IT skills that employers can't say no to ¹ .
Objectives	This tutorial gives an introduction to DM. It also explains briefly how some of the major DM algorithms work. It gives a hands-on experience on the use of WEKA. At the end of the tutorial, participants will be able to understand DM better, and will be able to select the appropriate approach to a particular problem.
Conference Track	The proposed tutorial falls under the track of "The Decision Processing in Management".
Who should attend?	Researchers, scientists, engineers, academicians and postgraduate students who seek to discover hidden knowledge in data.
Background of participants	The tutorial does not require the participants to have any background in DM.
Number of participants	Ideal number of participants is 15 because it contains hands-on session. However it should not exceed 25 participants.
Style	Participants must bring their own laptops. Instructions will be provided on how to install WEKA.

¹ <http://www.globalknowledge.com/training/generic.asp?pageid=1543&country=United+States>

Why WEKA?	WEKA is a powerful open source DM system. It is used by thousands of researchers all over the world.	
Hands-out	The attendees will receive a copy of the slides that cover the topics introduced during the workshop with exercises.	
Flexibility	Note that the tutorial can be extended to 1-day workshop, however, it is not recommended to reduce it less than half day event.	
Contents	9:00 - 10:30	Introduction to DM <ul style="list-style-type: none"> - What is DM? - Applications - How to mine data <ul style="list-style-type: none"> - Data Preparation - Defining a study - Data Mining - Analysis of Results - Assimilation of Knowledge DM Strategies and Techniques <ul style="list-style-type: none"> - Supervised learning - Unsupervised clustering
	11:00 – 1:00	Introduction to WEKA <ul style="list-style-type: none"> - Interface - Reading data - Building a model - Analysis Example 1 (Business Data) Example 2 (Medical Data)

Dr. Ahmed M. Zeki is an assistant professor at the department of computer science, International Islamic University Malaysia (IIUM). He teaches advanced computer science subjects such as Data Mining, Artificial Intelligence, Data Structures and Algorithms and others. He has been a trainer for several workshops conducted at IIUM.



He is familiar with several Data Mining packages including SPSS Clementine, SAS data miner and WEKA. His research is mainly focused on Pattern Recognition, Image Processing and Data Mining. He has published a number of papers at several local and international conferences, and received several awards.