BUSINESS INTELLIGENCE and DATA MINING

Instructor: Prof. Sameek Ghosh

Introduction

Information Technology has created new opportunities and newer challenges to handle vast data created every day. Managers are facing challenges in collecting and analysing the data. Low-cost massive data storage technologies and Internet connectivity have made available large amounts of data. The challenge is how to make use of the available data to make better decisions. Business Intelligence (BI) helps here. Business Intelligence (BI) refers to technologies, applications and practices for the collection, integration, storage, access, analysis, and presentation of business information to help users make better decisions.

This course seeks to develop skills in using data for supporting business decisions. The course also gives an introduction to the topics of statistics and data mining. The course will be application driven rather than the theory underlying the tool. However, necessary statistical foundation will be built in the class.

We will be using primarily R for statistical analysis, MS-Excel 2013 as an add-on statistical software and XLMiner for Data mining. Students are requested to procure the same.

Prerequisites for the course

No pre-requisite for this course

Expectation from students

Students are expected to know basic descriptive statistics. The students are expected to attend all the sessions and actively participate in the discussions. Timely completion and submission of all assignments is mandatory.

Course Contents and session plan

Session	Topics
1	**Introduction –** Course, Faculty, Students, Business Intelligence & Data Mining
2	**Statistical Discussion** – To develop a common ground
3	**Refreshing Correlation, Co-efficient of Variation, t-test (1 sample & 2 sample)** - using R statistical software
4	**ANOVA**
5	**Simple Linear Regression**
6	**Simple Linear Regression**
7	**Simple Linear Regression**
8	**Business Intelligence Farming –** Finding and forming specific BI problems
9	**Business Intelligence Farming –** Finding and forming specific BI problems

10	**Introduction to Datamining –** Data Analysis and Data Mining
11	**Partition Data** – Standard Data Partition
12	**Partition Data** – Partitioning with oversampling
13	**Classify –** Discriminant Analysis
14	**Classify –** Classification Tree
15	**Prediction –** k-nearest method prediction
16	**Prediction –** Regression Trees
17	**Association Rules**
18	**Mid Term Test**
19	Data Analysis – Transform
20	Data Analysis – Transform
21	
22	Data Analysis – Cluster Analysis
23	Data Analysis – Text Mining
24	**Classify –** Logistic Regression
25	**Classify –** Naïve Bayes
26	**Prediction** – Multiple Linear Regression
27	**Term Project Discussion**
28	**Term Project Discussion**
29	**Term Project Discussion**
30	**Term Project Discussion**

Student consultation hours

Students are welcome to consult me after seeking an appointment, or via electronic media like email and whatsapp.

Evaluation

The following components. The respective weights of each component are given alongside. Any change in this will be notified to you in advance.

Assignments (3 nos) - 30%

Mid-Term - 20%

Project - 30%

End-Term – 20%

Indicative list of References

R Programming – Step by Step guide.

Datamining – Data pattern evaluation

http://www.sthda.com/english/wiki/what-is-r-and-why-learning-r-programming

Rud, Olivia. Par. 2009. Business intelligence success factors – Tools for aligning your business in the global economy. USA: John Wiley and Sons Inc.

Soman, K. P., Diwakar, Shyam., & Ajay, V. 2006. Insight into data mining theory and practice. India: Prentice Hall of India Pvt Ltd.