## Tópicos Selectos en Minerias de Datos Avanzada

### Objetivo

The course is designed especially for computer science postgraduate students who need text mining and web mining techniques directly or indirectly in their thesis work. Basic concepts and the state-of-art techniques and representative algorithms will be introduced.

Description:
With the rapid advance of computer and internet technologies, a plethora of data accumulates. Data will not turn into knowledge no matter how long it is kept. Mining nuggets from data will add values to what we are currently doing in many areas. Data mining is a process that finds the valuables among the mountains of data.

We will review and examine the present techniques and the theories behind them and explore new and improved techniques for real world data mining applications. The arrangement of the course will encourage active class participation, creative thinking, and hands-on project development among the participants. Several course projects on some specific aspect of this emerging field will be required for each student to explore some in-depth issue(s) and gain data mining experience.

### Contenido:

1. Information retrieval
2. Text representation and search
3. Representative techniques of text processing
4. Evaluation techniques
5. Web Mining
6. Social Network Analysis
7. Web Crawling
8. Opinion Mining and Sentiment Analysis
9. Web Usage Mining

### Required text:

1.- Ricardo Baeza-Yates, Berthier Ribeiro-Neto (2010). Modern Information Retrieval. The Concepts and Technology behind Search. Second edition. Addison-Wesley.
2.- Bing Liu (2011). Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data. Springer. ISBN: 978-3-642-19459-7.
3.-Pang-Ning Tan, Michael Steinbach, and Vipin Kumar (2005) Introduction to Data Mining, Addison Wesley http://www-users.cs.umn.edu/~kumar/dmbook