# PTE-588 (3) Smart Oilfield Data Mining

## Course Objectives:

Data mining is the process of discovering information in large data sets. Using a combination of machine learning, statistical analysis, modeling techniques and database technology, data mining finds patterns and relationships in data and infers rules that allow predication of future results. Typical applications include reservoir characterization, simulation/workover candidate selection, failure detection and failure prediction.

At the end of the course, students should have an understanding of the data mining process as applied to petroleum engineering problems, of how to apply data mining algorithms to petroleum engineering data sets, and of key data mining algorithms and data structures.

## Instructor:

Ke-Thia Yao, kyao@isi.edu

| USC ISI Office: | USC UPC Office: |
|---|---|
| 4676 Admiralty Way, Suite 1227<br>Marina del Rey, CA 90292<br>(310) 448-8297 | 302 HED<br>(213) 740-0016 |


## Office Hours:

Thursdays 10:00AM-11AM at 302 HED
And by appointment

## Topics:

| Date | Topic | Lecturer |
|---|---|---|
| 29-Aug-13 | Data Mining Concepts and techniques | Ke-Thia Yao |
| 5-Sep-13 | Oilfield Data Systems | Jim Crompton |
| 12-Sep-13 | Data Mining Process | Ke-Thia Yao |
| 19-Sep-13 | Classification: Decision Trees, and Model Evaluation | Ke-Thia Yao |
| 26-Sep-13 | Classification: Linear Models, Neural Networks, SVM | Ke-Thia Yao |
| 3-Oct-13 | Classification: Rules, Bayesian Classifiers | Ke-Thia Yao |
| 10-Oct-13 | Clustering Analysis: Concepts and Algorithms | Ke-Thia Yao |
| 17-Oct-13 | Presentations: Formulating Petroleum Problems as Data Mining Problems | Student Presentation |
| 24-Oct-13 | Midterm | |
| 31-Oct-13 | Zonal Allocation and Increased Production Opportunities Using Data Mining | Andrei Popa |
| 7-Nov-13 | Control Performance Monitoring of Excessive Oscillations of an Offshore Production Facility | Joe Qin |

| | | |
|---|---|---|
| 14-Nov-13 | Failure Prediction in Rod Pumps | Yintao Liu |
| 21-Nov-13 | Presentations: Formulating Petroleum Problems as Data Mining Problems | Student Presentation |
| 28-Nov-13 | Thanksgiving recess | |
| 5-Dec-13 | Project Presentations | Student Presentation |

## Textbook:

Pang-Ning Tan, Michael Steinbach, Vipin Kumar
Introduction to Data Mining
Pearson Education/Addison Wesley
ISBN 0-321-32136-7

Optional:

Ian H. Witten & Eibe Frank
Data Mining: Practical Machine Learning Tools and Techniques
Morgan Kaufmann
ISBN 0-12-088407-0

## Software:

Weka: Data Mining Software in Java

http://www.cs.waikato.ac.nz/ml/weka/

## Grading:

Homework: 25%
Presentations: 25%
Final Project: 25%
Midterm: 25%

## Important dates:

| | |
|---|---|
| Aug 29 | First day of class |
| Sept 13 | Last day to add class, or drop without “W” |
| Oct 24 | Midterm |
| Nov 15 | Last day to drop with “W” |
| Nov 27-30 | Thanksgiving recess |
| Dec 5 | Final Project Presentations |