1. **Course Number and Course Title:**

   COE 637 – Advanced Machine Learning and Data Mining

2. **Credit Hours:**

   3 – 0 – 3

3. **Prerequisites and/or Co-Requisites:**

   Prerequisite: Approval of the CSE Head of Department
   Co-requisites: None
   Competencies: Undergraduate-level knowledge of statistics and programming

4. **Name and Contact Information of Instructor:**

   Name: Dr. Michel Pasquier

5. **Course Description (Catalog Description):**

   Presents the principles of machine learning and data mining. Covers key topics including data preparation and visualization, supervised and unsupervised learning, experimental validation and model interpretation. Examines various techniques from decision trees and rule induction to probabilistic methods and regression as well as association mining and clustering. Studies examples of data mining applications using state-of-the-art software such as R or Weka.

6. **Textbook and other Supplemental Material:**

   Textbook:
   - Witten I.H., Frank E., and Hall M.A., *Data Mining: Practical Machine Learning Tools and Techniques,* Morgan Kaufmann, 4th ed., 2016.

   Other supplemental material:
   - Alpaydin E., *Introduction to Machine Learning,* MIT Press, 3rd ed., 2014.
   - Hastie T., Tibshirani R., and Friedman J., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* Springer, 2nd ed., 2009.
   - Han J., Kamber M., and Pei J., *Data Mining: Concepts and Techniques,* Morgan Kaufmann, 3rd ed., 2012.
   - Machine learning and data mining software (free, open source); articles and data sets.

7. **Course Learning Outcomes:**

   Upon completion of the course, students will be able to:
   1. Describe the importance of and issues with machine learning and data mining for society.
   2. Apply appropriate machine learning models and data mining techniques for various tasks.
   3. Compare machine learning algorithms such as rules, trees, probabilistic, clustering.
   4. Evaluate the performance of machine learning techniques to solve real world problems.
   5. Use modern software to experiment with machine learning and data mining techniques.
   6. Critically review current research work in machine learning and data mining.
   7. Conduct independent research in the field of machine learning.

**8. Teaching and Learning Methodologies:**

Methods include lectures, problem and project based learning methods (assignments, exams, survey paper, research project, presentation), and class discussions.

**9. Course Topics and Schedule:**

| Topic/Activity | Weeks |
|---|---|
| Introduction: machine learning and data mining principles | Week 1 |
| Input data: concepts, instances, attributes, preparation | Week 2 |
| Output knowledge: tables, trees, rules, visualization | Week 3 |
| Basic methods: decision trees and probabilistic modeling | Week 4 |
| Basic methods: rule induction and association mining | Week 5 |
| Basic methods: linear models, regression, and clustering | Week 6 |
| System evaluation: testing, validation, comparison | Week 7 |
| System evaluation: performance and costs + Midterm exam | Week 8 |
| Data transformation: discretization, sampling, projections | Week 9 |
| Advanced methods: trees, numeric attributes, pruning | Week 10 |
| Advanced methods: rules, exceptions, optimization | Week 11 |
| Advanced methods: non-linear models, kernel methods | Week 12 |
| Advanced methods: numeric prediction, instance-based learning | Week 13 |
| Ensemble learning: combining models, bagging, boosting | Week 14 |
| Review and evaluation, class presentations | Week 15 |
| Final Exam | Week 16 |