Department of Computer Science & Information Systems
Second Semester: 2016-2017
Course Handout: Part-II

Date: 11/01/2017

In addition to part-I (General handout for all courses appended to the timetable) this portion gives further specific details regarding the course:

Course No. : BITS F464
Course Title : Machine Learning
Instructor-in-Charge : Kamlesh Tiwari, (Kamlesh.tiwari@)

1. Objective and Scope of the Course

One of the most significant developments in current technological platform is the availability of huge volume of data. This is possible due to the devices that are enabled to automatically collect and store data. Availability of sufficient amount data brings the power to develop efficient algorithms for very complex tasks that is not possible otherwise. The process of making algorithm better based on data is typically called learning and is the subject matter of this course. Machine Learning as a discipline is devoted to design algorithms that allow itself to learn patterns and concepts from data without being explicitly programmed. This course will introduce some of the principles and foundations of Machine Learning algorithms along with their real -world applications. The course would be introductory in nature, and would not assume any prior exposure of its audience towards Machine Learning.

The course will cover the major approaches to learning namely, supervised, unsupervised, and reinforcement leaning. The course emphasizes various techniques, which have become feasible with increased computational power and our ability to produce and capture huge volumes of data. The topics covered in the course include regression, decision trees, support vector machines, artificial neural networks, Bayesian techniques, Hidden Markov models, genetic algorithms etc. Some advanced topics like active and deep learning will also be covered.

2. Course Material

Text Book:

- *Tom M. Mitchell, Machine Learning, The McGraw-Hill Companies, Inc. International Edition 1997.*

Reference Books:

1. Christopher M. Bhisop, Pattern Recognition and Machine Learning, Springer, 2006.
2. N. J. Nilson, Introduction to Machine Learning, Stanford, Online Link http://robotics.stanford.edu/people/nilsson/mlbook.html

3. D. Michie, D.J. Spiegelhalter, C.C. Taylor (eds), Machine Learning, Neural and Statistical Classification, Ellis Horwood publishers, Online Link http://www.amsta.leeds.ac.uk/~charles/statlog/

4. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer, 2009. Online Link http://statweb.stanford.edu/~tibs/ElemStatLearn/printings/ESLII_print10.pdf

5. Hal Daume III, A Course in Machine Learning, 2015. Online Link http://ciml.info/

6. Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012 Online Link https://mitpress.mit.edu/books/machine-learning-0

3. Course Plan

Lecture	Topic(s) to be discussed	Reference
1	Introduction to Machine Learning	TB[Ch-1]
	Probability theory, Decision theory, Information theory, Linear Algebra	Self Study + R1[Ch-2], TB[Apndx-C]
2-5	MAP Hypothesis, Minimum Description Length (MDL) principle, Expectation Maximization (EM) Algorithm, Bias-variance Decomposition, Lagrange Multipliers, Mixture of Gaussians, PCA and SVD	TB[Ch-6], class notes, R1[Apndx-E]
6-8	Liner Models for regression: Linear basis function models, Bayesian linear regression	R1[Ch-3]
9-12	Liner Models for classification: Discriminant Functions, Probabilistic Generative Classifiers, Probabilistic Discriminative Classifiers	R1[Ch-4]
13-14	Bayesian Learning Techniques: Bayes optimal classifier, Gibbs Algorithm, Naive Bayes Classifier	TB[Ch-6]
15-21	Non-linear Models: Model Selection & Decision Trees, Ensemble Classifiers, Neural Networks, Multilayer Perceptron, Network training, Error back-propagation, Instance-based Learning, K-NN, Case-based Reasoning	TB[Ch-3], TB[Ch-4], R1[Ch-5], TB[Ch-8]
22-24	Margin/Kernel Based Approaches: Support Vector Machines	Class Notes, R1[Ch-7]
25-28	Graphical Models: Bayesian Belief Networks, Hidden Markov Models	TB[Ch-6], class notes
29-30	Unsupervised Learning: Mixture Models, K-means Clustering, Self-organized Maps (SOM	TB[Ch-6], R1[Ch-9]
31-32	Genetic Algorithms: Hypothesis space search, Genetic programming, Models of evaluation & learning	TB[Ch-9]
33-34	Reinforcement Learning: Q Learning, Non-deterministic rewards & actions, Temporal difference learning, Generalization	TB[Ch-13]
35-38	Advanced Topics: Active Learning, Deep Learning, Metric Learning	Class Notes
39-40	Application Examples: Speech Recognition, Image Retrieval	Class Notes
41-42	Big Data Challenges: Machine Learning for Big Data	Class Notes

4. Evaluation Scheme

S.No.	Evaluation Component	Weight	Information
1.	Mid-sem Test (Open Book)	25%	Duration would be 120 Min
2.	Notes Scribing (Open Book)	5%	-
3.	Quiz-01 (Close Book)	5%	Duration would be 30 Min
4.	ML Competition: *24 hour ML competition would be organized where students would participate in teams. Top two teams would also get goodies.*	5%	One Day *Grading would be done based on the relative ranking and performance of the team*
5.	Term Project: *Would be done on individual basis. A list of titles would be advertised on NALANDA from where students can select one. This may require lot of coding, experimentation and mathematical analysis.*	30%	Would be evaluated based on the report and viva for each of the following components a) literature survey (Jan 29) 5% b) Implementation (Feb 26) 8% c) Innovation (Mar 19) 8% d) Results (Apr 08) 4% e) Final Report (April 15) 5%
6.	Comprehensive Exam (Close Book)	30%	Duration would be 180 Min

5. Honor Code

No form of plagiarism shall be tolerated. Student shall be awarded ZERO marks and case may be reported to the appropriate committee of the Institute for appropriate action.

6. Notices

All notices would be put on Nalanda

7. Make-up Policy

To be granted only in case of serious illness or emergency on case by case basis for Mid-sem Test and Comprehensive Exam only.

8. Chamber Consultation Hours: Mon/Wed 5-6 PM (6120-N @ NAB)

Instructor-in-Charge