Title: Machine learning: tools, techniques, applications

Course No: CS771

Units: 3-0-0-(4) (9)

Proposer: Harish Karnick

Others who may teach the course:

Amitabha Mukerjee, Arnab Bhattacharya. Vinay Namboodiri

Pre-requisites:

CS210/ESO211/ESO207A, MSO201A/equivalent
Desirable: CS365/equivalent or instructor's consent;

About the course:

This course introduces the main tools and techniques of machine learning that are widely used in applications. The emphasis is on methods. Just enough theory will be discussed so that the basis for the method or technique is understood. Tools like ML libraries Weka, R, Octave, Scilab will have to be used in the assignments and project

A subsequent course covers the theoretical foundations in more depth.

Topics:

- Machine learning - what, how, where.
- Supervised, unsupervised and semi-supervised learning.
- Training, validation, testing, generalization, overfitting.
- Features and feature engineering.
- Decision trees, random forests.
- Linear classifiers.
- Kernel based methods and SVMs.
- Nearest neighbour methods.
- Hidden Markov models.
- Neural and deep networks.
- Ensemble methods - boosting, bagging, voting schemes.
- Distance metrics and clustering.
- Methods for semi-supervised learning.

References:

1. G James, D Witten, T Hastie, R Tibshirani, Introduction to Statistical Learning, Springer, 2014 (corrected printing available online).
2. Yaser S Abu-Mostafa, Malik Magdon-Ismail, Hsuan-Tin Lin, Learning from data, AMLBook, 2012. (see Mostafa video on based on book)
3. David Barber, Bayesian reasoning and machine learning, Cambridge University Press, 2012.
4. Kevin Murphy, Machine learning: a probabilistic perspective, MIT Press, 2012.
5. Christopher Bishop, pattern recognition and machine learning, Springer, 2007.
6. S Kulkarni, G Harman, An Elementary Introduction to Statistical Learning Theory, Wiley, 2011.
7. Zhi-Hua Zhou, Ensemble methods: foundations and algorithms, Chapman and Hall, 2012.