Subject: Machine Learning
Type: Elective (CE/CS)
Credit: 3

Course code: COMP 484
FM: 100 (50 internal + 50 final)

Course Description:

The primary objective of the course is to introduce the students in the undergraduate level to the primary approaches to machine learning and the study of computer algorithms which improve automatically through experience. Students would be introduced to basic concepts from statistics, artificial intelligence, information theory and other disciplines with a balanced coverage of theory and practice. As part of the course, students would not only get the exposure to theoretical and practical know-how about building machine learning systems on real-world problems but also would be developing their own prototypes or proof-of-concepts in the form of project assignments.

Contents:

Lecture 1:

Introduction

Why is Machine Learning?
What is well-defined learning problem?
An example: learning to play checkers, what questions should we ask about Machine Learning?

Concept Learning
Learning from examples, General to specific ordering over hypotheses, Version spaces and candidate elimination algorithm, picking new examples, the need for inductive bias [**5 hrs**]

Lecture 2:
Decision Tree Learning
Decision tree representation, ID3 learning algorithm, Entropy, Information Gain, Overfitting

Artificial neural networks
Threshold units, Gradient descent, Multilayer networks, Backpropagation, Hidden layer representations, Example: Face Recognition, Advanced Topics [**7 hrs**]

Lecture 3:
Evaluating hypotheses
Sample error, True error, Confidence intervals for observed hypothesis error, Estimators, Binomial distribution, Normal Distribution, Central Limit Theorem, Paired t tests, Comparing Learning Methods

Bayesian Learning

Bayes Theorem, MAP, ML hypotheses, MAP learners, Minimum description length principle, Bayes optimal classifier, Naïve Bayes learner, Example: Learning over text data, Bayesian belief networks, Expectation Maximization algorithm [**10 hrs**]

Lecture 4:
Computational Learning Theory
Computational learning theory, Setting 1: learner poses queries to teacher, Setting 2: teacher chooses examples, Setting 3: randomly generated instances, labeled by teacher, probably Approximately Correct (PAC) learning, Vapnik-Chervonenkis Dimension, Mistake Bounds

Instance Based Learning
k-Nearest Neighbor, Locally weighted regression, Radial basis functions, Case-based reasoning, Lazy and eager learning [**9 hrs**]

Lecture 5:
Genetic Algorithms
Evolutionary computation, Prototypical GA, An example: GABIL, Genetic Programming, Individual learning and population evolution
Learning Sets of Rules
Sequential covering algorithms, FOIL, Induction as inverse of deduction, Inductive Logic Programming [**8 hrs**]

Lecture 6:
Reinforcement Learning
Learning task, Q Learning, Nondeterministic Rewards and Actions, Temporal Difference Learning, Generalizing from Examples, Relationship to Dynamic Programming [**8 hrs**]

Text book:

Machine Learning. Tom M. Mitchell. McGraw-Hill International Editions. Computer Science Series, 1997.

Reference books:

Machine Learning – A Probabilistic Perspective, Kevin P. Murphy. The MIT Press, Cambridge, Massachusetts. London, England. 2012.

Introduction to Machine Learning. Ethem Alpaydin. The MIT Press. Cambridge, Massachusetts. London, England. 2004

Deep Learning in Python. Francois Chollet, Manning, 2018.

Data Mining. Practical Machine Learning Tools and Techniques. Third Edition. Ian H. Witten, Eibe Frank, Mark A. Hall. Elsevier Inc., 2011.

Books for Lab works:

Building Machine Learning Systems Using Python, Second Edition, Luis Pedro Coelho, Willi Richart, Packt Publishing, 2015.

Mastering Machine Learning with Python in Six Steps – A Practical Implementation Guide to Predictive Data Analytics Using Python, Manohar Swamynathan, Apress, 2017.

Artificial Intelligence with Python – Build real-world Artificial Intelligence Applications with Python to intelligently interact with the world around you. Prateek Joshi, Packt Publishing, 2017.

Hands-On Machine Learning with Scikit-Learn and TensorFlow. Aurelien Geron, O'Reilly Media, Inc., 2017.

Data Science from Scratch. Joel Grus, O'Reilly Media, Inc., 2015.

Machine Learning in Action, Peter Harrington, Manning Publications, 2012.

Machine Learning in Python. Essential Techniques for Predictive Analysis, Michael Bowles, Wiley & Sons, 2015.

Machine Learning for Hackers, Drew Conway & John Myles White, O'Reilly, 2012.

Programming Collective Intelligence, Building Smart Web 2.0 Applications, Toby Segaran, O'Reilly, 2007.

Lab and practical assignments:

Note: The whole idea of the lab sessions is to engage the students with practical machine learning libraries and tools thus prepare them for the mini-project. Hence, the students are encouraged to go through the relevant chapters of the different books listed in this syllabus for the lab works besides the one prescribed for the lab sessions below.

Tasks	**Timelines**
Lab 1: Getting started with Python Machine Learning Introduction to NumPy, SciPy, and Matplotlib Installing with Python Chewing data efficiently with NumPy and intelligently with SciPy Learning NumPy, SciPy First application of machine learning - Reading in the data - Preprocessing and cleaning the data	**Week I** **Chapter 1 -** Building Machine Learning Systems Using Python

- Choosing the right model and learning algorithm	
Lab 2: Classifying with Real-world examples The Iris dataset Building more complex classifiers A more complex dataset and a more complex classifier Classifying with scikit-learn Binary and multiclass classification	**Week II** **Chapter 2** - Building Machine Learning Systems Using Python
Lab 3: Clustering – Finding related posts Measuring the relatedness of posts Preprocessing – similarity measured as a similar number of common words Clustering Solving our initial challenge Tweaking the parameters	**Week III** **Chapter 3** - Building Machine Learning Systems Using Python
Lab 4: Topic Modeling Latent Dirichlet allocation Comparing documents by topics Choosing the number of topics	**Week IV** **Chapter 4** - Building Machine Learning Systems Using Python
Lab 5: Classification – Detecting Poor Answers Sketching our roadmap Learning to classify classy answers Fetching the data Creating our first classifier Deciding how to improve Using logistic regression	**Week V** **Chapter 5** - Building Machine Learning Systems Using Python
Lab 6: Artificial Neural Networks Building a Perceptron based classifier Constructing a single layer neural network Constructing a multilayer neural network Building a vector quantizer Analyzing sequential data using recurrent neural networks	**Week VI** **Chapter 14** – Artificial Intelligence with Python **Chapter 6** – Mastering Machine Learning with Python in Six Steps

Visualizing characters in an Optical Character Recognition Building an Optical Character Recognition Engine Recurrent Neural Network Long Short-Term Memory (LSTM) Transfer Learning	
Lab 7: Reinforcement Learning Creating an environment Building a learning agent Example code for q-learning	**Week VII** **Chapter 15** – Artificial Intelligence with Python **Chapter 6** – Mastering Machine Learning with Python in Six Steps
Lab 8: Deep Learning with Convolution Neural Networks Building a perceptron-based linear regressor Building an image classifier using a single layer neural network Building an image classifier using a Convolution Neural Network CNN on CIFAR10 Dataset CNN on MNIST Dataset Visualization of Layers	**Week VIII** **Chapter 16** – Artificial Intelligence with Python **Chapter 6** – Mastering Machine Learning with Python in Six Steps

Mini-Project

This is an opportunity for you to explore interesting machine learning problems. To get some idea about the project, you may consult the page, http://www.cs.cmu.edu/~tom/10701_sp11/proj.shtml . Form a group of at most 3 people. Each group member should contribute equally to the project. The execution outline for the mini-project are as follows:

- Proposal submission
  - Project title
  - Data set
  - Project idea (approximately two paragraphs)

    - Software and tools
    - Teammates and work division
- Midterm milestone
    - Midterm report should be 4-5 pages
        - Proposed method and experiments done so far
        - Design of upcoming experiments
        - Plan of upcoming activities(align each project member with their respective activity)
- Final project demo and submission

Information regarding the deadline of the mini-project milestones will be provided separately.

Grading policy:

Final Exam: 50

Internals:

Lab and practical assignments: 12

First internal: 10

Second internal: 10

Quiz: 8

Mini Project: 10