Computer Science Department
cs.salemstate.edu

CSC 455 Machine Learning **4 cr.**

Instructor: TBA **Office:** location **Phone:** (978) 542-extension
email: TBA@salemstata.edu **Office Hours:** days and times

Section	Time	Room	Final Exam
nn	days and times	location	Date and time

Catalog description:
This course provides an introduction to machine learning algorithms. Machine learning is focused on finding patterns in data to solve complex problems. Instead of explicitly programming computers to perform a task, machine learning lets us program the computer to learn from examples and improve over time without human intervention. The course covers a broad cross-section of models and algorithms for machine learning. Topics include data preprocessing techniques, supervised learning algorithms, unsupervised learning algorithms, deep learning algorithms, and reinforcement learning algorithms. Four lecture hours per week.

Prerequisite(s): CSC 260 and MAT 147.

Course Goals:
The purpose of this course is to develop student's knowledge and understanding of various machine learning algorithms. Specific goals are to:

CG01: present the basic concepts of machine learning
CG02: understand the importance of data pre-processing techniques
CG03: understand various supervised learning, unsupervised learning, and reinforcement learning algorithms
CG04: present the commonly used libraries for performing machine learning tasks
CG05: discuss the application of machine learning algorithms to real-world problems

Course Objectives:
Upon successful completion of the course, a student will have:

CO01: developed an appreciation for data-based learning models
CO02: gained experience with a wide variety of machine-learning algorithms
CO03: developed an understanding of machine learning model evaluation
CO04: gained hands-on experience in applying machine learning algorithms to real-world problems
CO05: demonstrated through projects and written assignments the ability to apply methods and techniques of machine learning

Student Outcome (SO) vs. Course Objectives matrix

SO	CO01	CO02	CO03	CO04	CO05
SO-1	✓	✓		✓	✓
SO-2				✓	✓
SO-3	✓				
SO-4					
SO-5	✓				
SO-6	✓	✓	✓		✓

Notes:

SO-1: Analyze a complex computing problem and to apply principles of computing and other relevant disciplines to identify solutions.

SO-2: Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.

SO-3: Communicate effectively in a variety of professional contexts.

SO-4: Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles.

SO-5: Function effectively as a member or leader of a team engaged in activities appropriate to the program's discipline.

SO-6: Apply computer science theory and software development fundamentals to produce computing-based solutions.

Course topics:

- **Machine Learning Basics** **IS4(0, 1, 0)**
  - Supervised Learning
  - Unsupervised Learning
  - Semi-supervised Learning
- **Data Pre-Processing** **IS10(0, 0, 2)**
  - Feature Extraction
  - Sampling
  - Data Transformation
  - Outlier Removal
  - Normalization
  - Feature Selection and Data Reduction
  - Feature Scaling
  - Discretization
  - Categorical Coding
- **Supervised Learning**
  - Classification and Regression **IS4(0, 16, 0)**
    - Linear Regression
    - Linear Regression with Regularization
    - Polynomial Regression
    - Logistic Regression
    - Decision Trees
    - Nearest Neighbor Classifier
    - Bayesian Models
    - Support Vector Machines
    - Discriminant Analysis
  - **Ensemble Learning and Random Forests** **IS4(0,2,0)**
    - Voting Ensembles
    - Bagging
    - Boosting
- **Evaluation and Comparison of Learners** **IS4(0, 2, 0)**
  - Overfitting and Underfitting
  - Errors to Costs
    - Loss
    - Cost
    - Score
  - Validation
    - Cross-validation
    - Leave-One-Out Cross Validation
  - Bias-Variance Tradeoff
    - Variance of the Data
    - Variance of the Model
    - Bias of the Model
- **Evaluation of Classifiers and Regressors** **IS4(0, 2, 0)**
  - Baseline Classifiers and Regressors
  - Accuracy
  - Confusion Matrix

- Receiver Operating Characteristic (ROC)
- Measures for Regression
- Residual Plots

- **Unsupervised Learning** **IS10(0, 0, 8)**
  - Clustering
  - Overview of Basic Clustering Methods
    - Partitional Clustering
    - Hierarchical Clustering
    - Self-Organizing Maps
  - K-Means Clustering
  - Expectation-Maximization (EM) Algorithm
  - Gaussian Mixture Clustering
- **Neural Networks and Deep-Learning** **IS10(0, 0, 16)**
  - Introduction to Artificial Neural Networks
    - Activation Functions
    - Feature Learning
    - Hyperparameters
  - Convolutional Neural Networks
    - Convolutional and Pooling Layers
    - Implementing CNNs
    - Computer Vision Using CNNs
  - Recurrent Neural Networks
    - Recurrent Neurons and Layers
    - Natural Language Processing with RNNs
  - Autoencoders
    - Stacked Autoencoders
    - Convolutional Autoencoders
    - Variational Autoencoders
  - Generative Adversarial Networks (GANs) and Diffusion Models
- Reinforcement Learning **IS10(0, 0, 6)**
  - Elements of Reinforcement Learning
  - Basics of Dynamic Programming
    - Finding Optimal Policies
    - Value Iteration
    - Policy Iteration
  - Temporal Difference Learning
    - Q-Learning
    - Generalization

Assignments and Examination:

The emphasis of this course in on the understanding of the basic principles of machine learning. Various supervised and unsupervised machine learning algorithms used for regression and classification are presented and discussed with emphasis on working with real-world data. Several case studies of classification and regression are discussed to fully understand the concepts.

The course grade will be determined using the following approximate weights: final exam: 25%, midterm exam: 25%, written homework 25%, and projects: 25%.

Course Objective / Assessment Mechanism matrix

	Homework	Projects	Midterm Examination	Final Examination
CO01	✓	✓		
CO02	✓			
CO03	✓	✓		
CO04		✓		
CO05	✓	✓	✓	✓
CO06	✓	✓	✓	✓
CO07		✓		

Bibliography:

1. Christopher M. Bishop. **Pattern Recognition and Machine Learning**. Springer International Publishing 2006.
2. Dirk P. Kroese and Zdravko I. Botev. **Data Science and Machine Learning: Mathematical and Statistical Methods**. CRC Press 2020.
3. Ethem Alpaydin. **Introduction to Machine Learning**. Third Edition, MIT Press 2014.
4. Jon Krohn, Grant Beyleveld, and Aglae Bassens. **Deep Learning Illustrated: A Visual, Interactive Guide to Artificial Intelligence**. Addison-Wesley Publications 2020.
5. Laura Graesser and Wah Loon Keng. **Foundations of Deep Reinforcement Learning: Theory and Practice in Python**. Addison Wesley Publications 2020.
6. Mark E. Fenner. **Machine Learning with Python for Everyone**. Addison Wesley Publications, 2019.
7. Michael Bowles. **Machine Learning with Spark and Python: Essential Techniques for predictive analytics**. Second Edition, Wiley Publications 2019.
8. Miroslav Kubat. **An Introduction to Machine Learning**. Second Edition, Springer International Publishing, 2017.
9. Mohseen Mohammed, Muhammad Badrudin Khan, and Eihab Bashier Mohammad Bashier. **Machine Learning: Algorithms and Applications**. CRC Press 2017.
10. M. Gopal. **Applied Machine Learning**. McGraw-Hill Education, 2019.
11. Peter Wlodaraczak. **Machine Learning and its Applications**. CRC Press/Taylor & Francis Group 2020.
12. Simon Haykin. **Neural Networks and Learning Machines**. Third Edition, Pearson Publications 2008.
13. Simon Rogers and Mark Girolani. **A First Course in Machine Learning**. Second Edition, CRC Press 2017.
14. Stephen Marsland. **Machine Learning an Algorithmic Perspective**. CRC Press 2009.
15. Wei-Meng Lee. **Python Machine Learning**. Wiley Publications 2019.

Study Groups:

While I strongly encourage study groups, for non-group assignments I require that each student hand in his/her answers in his/her own words **- if two answers come out exactly the same or highly similar, neither will receive credit and/or further actions will be taken** (such as reporting to the department and/or university). Given the nature of most of the projects, homework questions and writing assignments, it will be almost impossible for two people to come up with highly similar answers UNLESS they copy.

Academic Integrity:

Academic Integrity Policy and Regulations can be found in the University Catalog and on the University's website (http://catalog.salemstate.edu/content.php?catoid=13&navoid=1295#Academic_Integrity). The formal regulations are extensive and detailed - familiarize yourself with them if you have not previously done so. A concise summary of and direct quote from the regulations: "Materials (written or otherwise) submitted to fulfill academic requirements must represent a student's own efforts". *Submission of other's work as one's own without proper attribution is in direct violation of the University's Policy* and will be dealt with according to the University's formal Procedures.

"Salem State University is committed to providing equal access to the educational experience for all students in compliance with Section 504 of The Rehabilitation Act and The Americans with Disabilities Act and to providing all reasonable academic accommodations, aids and adjustments. Any student who has a documented disability requiring an accommodation, aid or adjustment should speak with the instructor immediately. Students with Disabilities who have not previously done so should provide documentation to and schedule an appointment with the Office for Students with Disabilities and obtain appropriate services."

In the event of a university declared critical emergency, Salem State University reserves the right to alter this course plan. Students should refer to http://www.salemstate.edu for further information and updates. The course attendance policy stays in effect until there is a university declared critical emergency. In the event of an emergency, you will be informed through Canvas/Email. Students should review the plans and gather all required materials when an emergency is declared.

Note: This syllabus represents the intended structure of the course for the semester. If changes are necessary, students will be notified in writing and via all regular class communication mechanisms.