NOVA COLLEGE-WIDE COURSE CONTENT SUMMARY

ITD 140 – MACHINE LEARNING I (3 CR.)

Course Description

This is an introductory course in artificial intelligence and machine learning that covers basic theory, algorithms, and applications. It focuses on feature engineering and machine learning applications within the larger world of artificial intelligence. Lecture 3 hours per week.

General Course Purpose

To give the student competence in describing, choosing, training and testing basic machine learning algorithms through applications to common use cases.

Course Prerequisites/Co-requisites

Prerequisites: ITE 152 – Introduction to Digital Literacy and Computer Applications or equivalent

Course Objectives

A. Describe and identify basic artificial intelligence approaches
B. Describe and differentiate between supervised and unsupervised learning techniques
C. Describe and apply feature extraction and engineering techniques
D. Identify, explain and apply basic machine learning algorithms, both supervised and unsupervised
E. Identify, evaluate and apply performance metrics to improve models

Major Topics to be Included

A. Overview of artificial intelligence approaches
B. Feature extraction and engineering
C. Supervised learning with example implementations
D. Unsupervised learning with example implementations
E. Performance metrics and hyperparameter tuning

Student Learning Outcomes

A. Machine Learning
    1. Define and explain the purpose of machine learning
    2. Define and explain the purpose of artificial neural networks
B. Feature Engineering
    1. Define Feature Engineering, Imputation
    2. Define and apply mean substitution, back/forward-fill substitution
    3. Define raw features
    4. Define and explain how to create derived features, and define:
        a. Binarization, Rounding, Binning, Fixed-Width Binning
    5. Identify and use continous numerical data
    6. Define a qualitative variable as it applies to machine learning
    7. Define a quantitative variable as it applies to machine learning
        a. Define a continuous quantitative variable as it applies to machine learning
        b. Define a discrete quantitative variable
    8. Define, create and use categorical data
    9. Define, explain and apply one-hot encoding
    10. Define and perform simple feature scaling
    11. Define and apply min-max scaling (i.e., normalization)

12. Define and perform standardization

C. Performance Metrics and Tuning
   1. Define and evaluate basic performance metrics, including accuracy, the confusion matrix, precision, recall, F1 score, and AUROC
   2. Define and explain bias, variance and the bias-variance trade-off
   3. Define loss functions
      a. Define and calculate L1 Norm (least absolute error) and L2 Norm (least squares error)
   4. Defne and explain the significance of hyperparameters

D. Supervised Learning
   1. Define and explain the purpose of supervised learning
   2. Identify supervised learning algorithms and when they should be used
   3. Define, explain and perform linear and multivariate regression
   4. Define classification as a supervised learning prediction task
      a. Discuss when it is appropriate to use classification
      b. Define and explain the k nearest neighbor (knn) and decision tree algorithms
      c. Apply classification to a dataset using, e.g., knn and decision trees
   5. Neural Networks
      a. Define and explain the function of perceptrons
      b. Define and explain the structure and function of artificial neural networks
      c. Define and explain the purpose of deep learning

E. Unsupervised Learning
   1. Define and explain the purpose of unsupervised learning
   2. Identify unsupervised learning algorithms and when they should be used
   3. Clustering
      a. Define and explain the purpose of clustering
      b. Identify when clustering is appropriate
      c. Apply k-means clustering to a sample dataset
   4. Describe and apply basic dimensionality reduction

Required Time Allocation per Topic

To standardize the core topics of this course, the following student contact hours per topic are required. Each syllabus should adhere as closely as possible to these allocations. Topics are not necessarily to be taught in the order shown. There are normally 45 student contact-hours per semester for a three-credit course. Sections of the course offered in alternative formats (i.e. not standard 15-week) still meet for the same number of contact hours. The final exam is not included in the timetable. The quickly evolving nature of artificial intelligence—and machine learning in particular—means that some content noted in this document may be superseded or made obsolete. As such, it is important to include such changes in individual syllabi. Additionally, time is allocated for additional and optional topics to provide flexibility to instructors in tailoring the course to special needs or resources.

Topics	**Hours**	**Percentage**
Overview of artificial intelligence approaches	3	7%
Supervised learning	12	26%
Unsupervised learning	9	20%
Feature extraction and engineering	15	33%
Testing to include quizzes, tests and exams (excluding final exam)	3	7%
Other optional topics	3	7%
Total	**45**	**100%**