GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

ECE 4803: Fundamentals of Machine Learning (3-0-3)
Syllabus – Spring 2022

Course Instructor: Prof. Ghassan AlRegib | alregib@gatech.edu | www.ghassanalregib.info

TAs: TBA

Course Days/Times: TBA

Office Hours: TBA

Textbook: No required textbook but the following books are excellent references for this class:

1. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, The MIT Press (November 18, 2016), ISBN-10: 0262035618, ISBN-13: 978-0262035613
   Available online: https://www.deeplearningbook.org/
2. Aggelos Konstantinos Katsaggelos, Jeremy Watt, and Reza Borhani, *Machine Learning Refined: Foundations, Algorithms, and Applications*, Cambridge University Press; 2 edition (March 12, 2020), ISBN-10: 1108480721, ISBN-13: 978-1108480727

Prerequisite: Undergraduate Semester level MATH 2550 Minimum Grade of C or Undergraduate Semester level MATH 2551 Minimum Grade of C) and Undergraduate Semester level CS 1301 Minimum Grade of C
The main pre-requisite is a course in multivariable calculus (MATH 2551 or equivalent) and linear algebra. Students should feel comfortable dealing with vectors and matrices. Students are expected to have the basic Python programming skills similar or equivalent to CS 1301.

Course Objective: An introduction to the fundamentals and applications of Machine Learning

Academic Honesty: All violations of the Georgia Tech Honor Code will be handled by referring the case directly to the Dean of Students for investigation and penalties.

Homework: Check the Homework link on the main course Canvas page for all assignments often. Exam problems tend to be simple version of homework problems. Students are encouraged to form **groups** to discuss homework problems, but they are required to formulate their own write-ups. Georgia Tech's Honor Code will be strictly enforced and students are required to observe the code all the time. Check Canvas Assignment often for due dates and submission instructions.

Exams: There will be two mid-term exams, amounting to **10 course points each out of 100**. Moreover, one final exam will be given at semester end and it accounts for **20 course points out of 100**. All exams will be **conducted remotely** through Canvas Assignment. Some digital proctoring tools may be used. Georgia Tech's Honor Code will be strictly enforced and students are required to observe the code all the time. Students are expected to be available at their

registered lecture times for exams. They are also required to be available for their assigned time slots to take the finals. The given exams and final exam duration will be longer than what are specified in the official schedules to allow for downloading, ordering, scanning and uploading exam papers. Check Canvas Exam link often for instructions.

Assignments: Assignments will have both analytical part and a hands-on part. The hands-on parts are coding assignments and they will look like mini projects. Students are expected to have some background in Python. Students are encouraged to utilize Google's Colab for their codes in this course. These hands-on assignments vary between adding to existing codes, writing codes from scratch, searching the literature for codes for a specific application, or summarizing papers in the literature.

Grading:

Homework Assignments	60%	Exams	40%

In all cases we will at least follow the basic traditional grading scale where: A=90-100; B=80-89; C=70-79; D=60-69; F=0-59. While in all cases a grade of 90-100 will be assigned an 'A', the boundaries/cutoffs for the other grades will be determined at the end of the semester based on the overall class performance. It is impossible to determine what the exact "cutoffs" will be for each grade but you can be assured that your assigned grade will never be lower than the "traditional" grading scale described above based on your final class average.

Programming Language: We will utilize Python throughout the course. We will also utilize a library of Jupyter notebooks in Colab.

Canvas: Course website on Canvas: when clicking the course, students will see an overall view of all the course components, including Syllabus, Lecture, Homework, Exam, Piazza, Supplements, etc. Go to https://canvas.gatech.edu/ and if you do not see the class page, make sure you are registered for the course.

Piazza: Students are expected to utilize PIAZZA platform to post questions and engage into online discussions. Make sure you enroll into the course site on Piazza. Check for sign-up announcements. If you have any problems or feedback for the developers, email team@piazza.com. Find our class page by clicking the Piazza link on the left navigation menu on Canvas. If Georgia Tech ends Piazza and moves to another platform, we will be utilizing the new platform.

Assignments Submission: All homework assignments need to be submitted on Canvas. Read the instructions of each assignment carefully.

Attendance: Your attendance and participation are strongly encouraged. Check the Institute Absence Policy at: http://www.catalog.gatech.edu/rules/4/.

By January 22, 2022 (tentative), we must submit a report on Verification of Participation. We will use lecture attendance and/or other metrics (e.g., TurningPoint survey responses and Piazza activities) to decide if you are participating or not.

Communications: All communication is expected to be conducted on Piazza. One can utilize the private message option. Emails are discouraged.

Announcements: Official announcements will be posted on Canvas and/or Piazza or announced during lectures.

Academic Honesty: All violations of the Georgia Tech Honor Code will be handled by referring the case directly to the Dean of Students for investigation and penalties. The complete honor code can be found at this link: http://www.policylibrary.gatech.edu/student-affairs/academic-honor-code

Available Resources:

- The Center for Academic Success has programs to help students improve their study habits and time management: http://www.successprograms.gatech.edu/.
- The Dean of Students Office helps students who have personal or medical issues that impact their academic performance: http://www.deanofstudents.gatech.edu/

Office of Disability Services: If you are a student registered with the Office of Disability Services (ODS), please make sure the appropriate forms and paperwork are completed with the instructor within the first week of classes. The instructor will abide by all accommodations required by ODS. The schedule for exams is posted in the syllabus and any potential modifications or changes will be made with at least one week's notice. It is the responsibility of the student to properly arrange test accommodations for each exam with ODS in sufficient time to guarantee space for exam administration. ALL exam accommodations must be handled through ODS. If the student does not register accommodations with ODS for the taking of an exam, then they will have to take the exam at the normally scheduled times without any additional accommodation unless the instructor is given specific directive from ODS on the student's behalf due to a mitigating circumstance. (https://disabilityservices.gatech.edu/)

Topical Outline:

- Overview
    - History of Pattern recognition, Development of an ANN
    - Types of Learning i.e., Supervised, Semi-supervised, Weakly supervised, Un-supervised
    - General features of a supervised learning system i.e. features, training/validation set, labels, model complexity and overfitting etc.
    - Simple overview of Optimization
- Classification
    - Algorithms: Nearest Neighbors, Logistic Regression, Decision Trees, Random Forest, SVM, ANN
    - Classification Performance Evaluation
        - Cross-Validation using k-fold, Confusion Matrix, Precision, Recall, and F1 score, ROC,
    - A set of hand-on exercises on Colab
- Regression:
    - Linear Regression

        - Linear Regression, Performance Measures
        - Cost Function
    - Polynomial Regression
        - High Degree Polynomial Regression
    - Regularized Linear Models
        - Lasso vs. Ridge Regularization (L1/L2 regularization)
        - Dealing with high dimensional feature space – PCA
    - A set of hand-on exercises on Colab
- Clustering
    - Introduction
    - Proximity Measures
    - Similarity vs. Dissimilarity
    - Distance Measures
    - Common Clustering Methods
        - k-Means, GMM, Mean-shift, Spectral Clustering
    - Evaluating Clustering Performance
    - Image Segmentation as a clustering problem
    - A set of hand-on exercises on Colab
- Neural Networks
    - Introduction to Artificial Neural Network:
        - Artificial Neuron
        - PyTorch
    - Non-linearity, Activations, Losses
    - ConvNets
        - convolutional layer, pooling, FC, training
    - A set of hand-on exercises on Colab
- Autoencoders
    - Fully Connected autoencoders, Conv AE, VAE
    - A set of hand-on exercises on Colab
- Advanced Topics (To be covered as time permits)
    - Generative models and GANs
    - Sequence models
        - Plain RNN
        - GRUs and LSTMs
    - Boosting, Bagging, Stacking
    - Transfer learning
    - Data augmentation