

THE UNIVERSITY OF TEXAS AT TYLER
Soules College of Business
CSCI5350 – Machine Learning
Spring 2024

Instructor: Nary Subramanian, Ph.D.
COB 315.11
Email (*preferred way to contact*): nsubramanian@uttyler.edu
Phone: 430-558-1330

Office Hours: (Physical) Tues and Thurs 8:00 am to 9:30 am; by Zoom and email.

Text (Required): *Machine Learning for Business Analytics: Concepts, Techniques and Applications in RapidMiner*, 1st Edition, by Galit Shmueli, Peter C. Bruce, Amit V. Deokar, and Nitin R. Patel. ISBN: 978-1119828792.

Course Description: In today's rapidly evolving technological landscape, machine learning empowers computers to learn and adapt from data, enabling them to make intelligent decisions without explicit programming. This course is designed to equip you with the essential knowledge and skills to apply artificial intelligence, where machine learning plays a central role. Understanding machine learning is crucial for professionals across various domains, as it is important in data analysis, predictive modeling, natural language processing, and autonomous systems. In this course we will study various machine learning algorithms and applications which will give you the expertise needed to employ data-driven decision-making in your academic and professional careers. We will be programming using the *no-code* programming environment called **RapidMiner Studio** (download from www.rapidminer.com) which is free to use for educational purposes and makes our programming rather easy. All course material and videos will be posted on Canvas.

Grading: Grading will be based on homework assignments. Assignments are open book. Assignment submissions should be made electronically to Canvas. Late submissions will not be graded and there is **no** penalty for early submissions. Assignments will be available from 12pm on Sunday and will be due by 8am on Wednesday; only in the last week you will have a *second* assignment that will appear at 8am that Wednesday due 8am that Friday. Each assignment will be worth 20 points and there will be seven assignments in all – the lowest assignment grade will be dropped. Sum of the six best assignment grades (maximum possible score = 120) will be used to compute your final letter grade as per the grading policy given below.

Grading Policy:

Total Points	Letter Grade
≥105	A
≥90, < 105	B
≥80, < 90	C

Course Objectives:

1. Identify methods for data cleaning, replacing missing data, and normalization.
2. Develop models for supervised classification using linear regression, K-nearest neighbors, logistic regression, and neural networks.
3. Develop clustering models using hierarchical and K-means clustering.
4. Develop software to analyze data using decision trees.

Make-up Policy

There will be no make-ups for missed homework assignments; missed assignments will get a grade of zero.

Tentative Schedule:

<u>Week</u>	<u>Chapter</u>	<u>Topic</u>
1/16 – 1/21	1, 2	Introduction to Machine Learning, Overview of the Machine Learning Process
1/22 – 1/28	6	Multiple Linear Regression
1/29 – 2/4	7	K-Nearest Neighbors
2/5 – 2/11	9	Classification and Regression Trees
2/12 – 2/18	10	Logistic Regression
2/19 – 2/25	11	Neural Networks
2/26 – 3/2	16	Cluster Analysis

University Policies

University policies can be seen at <https://www.uttler.edu/academic-affairs/files/syllabuspolicy.pdf>.