



1. Course Information

- DNSC 3603, Data Science, 6 credits
- Spring Semester 2023, Tuesday-Thursday 10:00-11:15, A110
- Instructor Name and E-mail: Professor of Information Systems Hulusi Ogut, hogut@ada.edu.az
- Office , Office Hours and Phone: D313, After every class or get appointment via email, 489
- Position in curriculum: Free Elective for 4th year Bachelor of Business Administration (BBA) students.
- Pre-requisite:

2. Course Description

This course covers modern statistical and machine learning methods for working with small and big data. Some of major topics in this class are classification, regression, forecasting, clustering and dimensionality reduction techniques. The Python programming language will also be used to teach the essential skills for data wrangling, application and deployment of techniques.

3. Course Learning objectives

The goals of this course are:

- To appreciate the enhanced data rich environment of today's global economy and get exposed to the related business intelligence service opportunities that exist.
- To provide a practical understanding of the key methods of classification, prediction, reduction and exploration that are at the heart of data science.
- To decide when to use which technique.
- To implement major techniques using software
- To become smart/skeptical consumers of statistical techniques.
- To gain the intellectual capital required to provide business analytics services.

Jupyter Framework for Google Colabs (<https://colab.research.google.com/>) will be used for exercises and implementation of course topics at Python.

4. Course Literature

Kaggle Courses <https://www.kaggle.com/learn/python>
<https://www.kaggle.com/learn/pandas>
<https://www.kaggle.com/learn/intro-to-machine-learning>
<https://www.kaggle.com/learn/intermediate-machinelearning>
<https://www.kaggle.com/learn/time-series>
<https://www.kaggle.com/learn/feature-engineering>
<https://www.kaggle.com/learn/data-cleaning>

Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 3rd Edition.

Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Python
<https://www.dataminingbook.com/book/python-edition>

Rob J Hyndman and G. Athanasopoulos, Forecasting Principles and Practice (3rd ed),
<https://otexts.com/fpp3/>

5. Assessment

HomeWorks and Cases: 25%
Midterm Exam :32%
Final Exam :40%
Attendance: 3%

If homework submission is late, 5% will be deducted from the homework grade for each day.
No makeup exam will be given for midterm exam. If medical documents are provided, I will increase the weight of final exam to 75%.

6. Student code of ethics

All students are required to uphold and embody the requirements and principles stated in the ADA Honor Code. You are responsible for reading the Honor Code in detail and obey it at all times during the course of your studies at ADA, as it is an institutional document which applies to all classes and other activities at ADA University.

7. Github Site of Class Material

<https://github.com/ogut77/DataScience>

8.Tentative Course Schedule

Week	Chapter	Subjects	Books
1		Course Introduction+ Python Tutorial	Lecture Notes
2		Data Processing and Data Wrangling	Lecture Notes
3		Linear Regression and Regularization Techniques	Lecture Notes
4		Bias and Variance Tradeoff, Decision Trees +Bagging (Random Forest)	Lecture Notes
5		Boosting Techniques (XGBoost+CatBoost+ LightGBM) for classification and classification metrics	Lecture Notes
6		Boosting Techniques for Regression and Parameter Optimization	Lecture Notes
7		Other Classification and Regression Techniques (Logistic Regression, kNN, Naive Bayes)	Lecture Notes
8		Review+ Midterm	Lecture Notes
9		Other Classification and Regression Techniques (SVM, Neural Network)	
10		AutoML Tools (AutoGluon and PyCaret)	
11		PCA+ Data Reduction Techniques	Lecture Notes
12		Clustering+ Segmentation	Lecture Notes
13		Forecasting using Univariate Data	Lecture Notes
14		Forecasting (ARIMA, SARIMA and other techniques for multivariate data)	Lecture Notes
15		Review	