ADA University
School of Business

1. Course Information

- DNSC 3603, Data Science, 6 credits
- Spring Semester 2022, Wednesday-Friday 15:30-16:45, B203
- Instructor Name and E-mail : Visiting Professor of Information Systems Hulisi 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-machine-learning
https://www.kaggle.com/learn/time-series
https://www.kaggle.com/learn/feature-engineering
https://www.kaggle.com/learn/data-cleaning

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

5. Assessment

HomeWorks and Cases (there will be around 11 homework): 30%
Project: 12% (It will be explained in class and sample reports will be distributed)
Midterm Exam :25%
Final Exam :30%
Attendance: 3%
For each day submission late, 5% will be deducted from the homework grade.

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. Additional Sources

Google Data Analytics Professional Certificate
https://www.coursera.org/professional-certificates/google-data-analytics#howItWorks

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

Energy Forecasting Lecture Notes (It will be distributed at the class)
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		Decision Trees +Random Forest	Lecture Notes
5		XGBoost+CatBoost+ LightGBM for binary output	Lecture Notes
6		XGBoost+CatBoost+ LightGBM for Multiclass Output and Regression+Parameter Optimization	Lecture Notes
7		Other Classification Techniques (Logistic Regression, kNN, SVM, Naive Bayes)	Lecture Notes
8		PCA+ Data Reduction Techniques	Lecture Notes
9		Review+ Midterm	
10		Clustering+ Segmentation	Lecture Notes
11		Neural Network	Lecture Notes
12		Text Mining	Lecture Notes
13		Forecasting	Lecture Notes
14		Forecasting	Lecture Notes
15		Recommendition System (If time permits) +Review	Lecture Notes