IT University of Copenhagen		
Study Programme: KSD	**Course:** Data Mining, MSc.	
Class Code: KSDAMIN1KU	**Semester:** Autumn 2024	**Size in ECTS:** 7.50
Course Manager: Fabricio Batista Narcizo (narcizo[at]itu[dot]dk)		
Teaching Assistant: Hermano Dantas Farias (fariashermano[at]gmail[dot]com)		

Lecture Plan

Week	***Date***	***Syllabus***
01	30/08/2024	**Getting Started** DAMIN Course Organization, Introduction to Data Mining, Development Environment Setup, Visual Studio Code, Data Version Control, GitHub, Miniconda, Python Programming Language, Python Libraries (*Numpy*, *Scipy*, *Matplotlib*, *Plotly*, *Pandas*, *Seaborn*, *Scikit-Learn*, *PyTorch*, *TensorFlow*, *Keras*, *OpenCV*, *dLib*, *Google MediaPipe*, *Jupyter Notebook*)
Introduction		
02	06/09/2024	**Introduction to Python Programming Language** Hello World!, Variables and Types, Basic Operations, String, Control Flow, Conditions, Data Structure (*Lists*, *Sets*, *Maps*), Loops, Functions, Tuples and Dictionaries, Classes and Objects, Generators, Modules and Packages, Serialization and Files
03	13/09/2024	**Introduction to Linear Algebra** Basic Concepts and Notation, Scalars, Vector Operations, Matrix Operations, Special Matrices, Linear Equations, Determinants, Eigenvalues and Eigenvectors, Linear Transformations
Exploratory Data Analysis		
04	20/09/2024	**Data Preprocessing** Challenges, Data Quality, Data Cleaning, Data Integration, Data Reduction, Data Transformation, Feature Engineering, Data Sampling, Data Storage, Tools and Libraries, Best Practices
05	27/09/2024	**Data Exploration and Visualization** Exploratory Data Analysis (EDA), Descriptive Statistics, Data Visualization Principles, Basic Visualization Techniques, Advanced Visualization Techniques, Using Visualization for Data Exploration, Tools and Libraries

Lecture Plan		
Week	***Date***	***Syllabus***
Supervised Learning Algorithms		
06	04/10/2024	**Classification: Basic Concepts** Types of Classification, Decision Trees, k-Nearest Neighbors (k-NN), Logistic Regression, Naive Bayes, Model Evaluation Metrics (*Accuracy*, *Precision*, *Recall*, *F1-Score*), Confusion Matrix, ROC Curve and AUC, Data Preprocessing for Classification
07	11/10/2024	**Classification: Advanced Techniques** Support Vector Machines (SVM), Neural Networks, Deep Learning, Popular Tools and Libraries (*Scikit-learn*, *TensorFlow*, *Keras*, *PyTorch*, *PyTorch Lightning*)
Autumn Break (18/10/2024)		
08	25/10/2024	**Regression Analysis** Basic Concepts, Types of Regression, Simple Linear Regression, Multiple Linear Regression, Non-Linear Regression, Correlation Analysis, Evaluating Regression Models
		Assignment #01 Deadline: Thursday, 24 October 2024 at 23.59 via learnIT
Unsupervised Learning Algorithms		
09	01/11/2024	**Clustering: Basic Concepts** Types of Clustering, k-Means Clustering, Hierarchical Clustering, Evaluation of Clustering Results, Visual Evaluation (*Scatter Plots*, *Heatmaps*)
10	08/11/2024	**Clustering: Advanced Techniques** DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*), Gaussian Mixture Models (GMM), Self-Organizing Maps (SOM), Clustering in High-Dimensional Data
11	15/11/2024	**Dimensionality Reduction** Concepts and Benefits, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), t-Distributed Stochastic Neighbor Embedding (t-SNE)
Data Mining in Practice		
12	22/11/2024	**Anomaly Detection** Statistical Methods, Distance-Based Methods, Clustering-Based Methods, Machine Learning Methods, Evaluation Metrics for Anomaly Detection

Lecture Plan		
Week	***Date***	***Syllabus***
13	29/11/2024	**Association Rule Mining** Market Basket Analysis, Apriori Algorithm, FP-Growth Algorithm, Evaluation of Association Rules, Advanced Topics in Association Rule Mining
14	06/12/2024	**Final Exam Project and Course Evaluation** There will be no lecture. We will shortly repeat the requirements for the exam project, and help you to finish your projects before the deadline. In the end, we will have the course evaluation.
		Assignment #02 Deadline: Sunday, 20 December 2024 at 13.59 via learnIT