CENG574 - Statistical Data Analysis

Instructor: Hüseyin Aydın (email: huseyin@ceng.metu.edu.tr · Office: B211)

Lectures: Fri 13:40 - 16:30 · BMB3

Office Hours: By appointment

1. **Catalog Data:** Statistical Data Analysis · (3 - 0) · 3
   The objective of this course is to introduce the concepts and techniques of clustering and multivariate and exploratory data analysis. This course offers an opportunity to perform data analysis by using data visualization, projection, and embedding.

   Prerequisites: Knowledge of programming, probability and linear algebra.

2. **Main Reference Book:** E. Alpaydın, Introduction to Machine Learning, The MIT Press. [3rd Edition (2014) or 4th edition (2020)]

3. **Course Outline:**

   Schedule:

Week	Date	Topics	Assignments
1	Feb 20	Introduction	#1
2	Feb 27	Input representation; distance metrics and similarity measures	#2
3	Mar 6	Probability and linear algebra; linear projections of data; PCA	
4	Mar 13	MDS; clustering	#3
5	Mar 20	No Lecture	
6	Mar 27	Hierarchical clustering; k-means clustering	#4
7	Apr 3	Evaluation and validity of clusters	#5
8	Apr 10	Overview and discussion of covered topics	
9	Apr 17	Midterm Exam (in-class)	
10	Apr 24	Other clustering algorithms	
11	May 1	No Lecture	
12	May 8	Non-linear projections	#6
13	May 15	Presentations of dataset analysis prelim. work and discussion	
14	May 22	Data stream analysis	
15	May 29	No Lecture	
16	Jun 5	Presentations of the final analysis of datasets	

   Evaluation:

   (a) Attendance and class participation (5 %)

   (b) Assignments (25 %): #1-2: 3 % each · #3-5: 5 % each · #6: 4 %

   (c) Midterm exam (20 %)

   (d) Final dataset analysis (20 %)

   (e) Final exam (30 %)

4. **Notes & Remarks:**

- Assignments will be done on individual basis.
- Final dataset analysis will be performed in a team setting of 2 persons.
- `Python` is the main programming language for the applied part of the course, yet `R` is also accepted.
- You have a total of 4 days of late submission for the assignments.
- Communication is via ODTUClass (`https://odtuclass.metu.edu.tr`).
- All submissions must reflect your own effort. Using AI agents for code/text generation is prohibited.
- Academic Integrity Guide for Students: `http://oidb.metu.edu.tr/sites/oidb.metu.edu.tr/files/Academic%20Integrity%20Guide%20for%20Students.pdf`
- This syllabus is tentative, and changes can be made during the semester.