Introduction to Data Science – Spring 1403

Instructors	Behnam Bahrak (bahrak@gmail.com) – Yadollah Yaghoobzadeh (y.yaghoobzadeh@ut.ac.ir)
Lectures	Sunday, Tuesday: 15:00-16:30, Class 5

Syllabus

- Introduction – Data science lifecycle
- Types of Data, Web scraping and Data Collection, Sampling, Sampling Bias, Observational and Experimental Studies, Confounding Variable
- Data Visualization, Visualizing numerical data, Visualizing categorical data, Storytelling with data
- A review of probability theory, conditional probability, Bayes theorem, random variable, expectation and variance, Normal, binomial, Poisson, and geometric distributions.
- Central limit theorem, confidence interval, hypothesis testing, power of a test
- Linear Regression and Logistic regression
- Databases and SQL
- Cognitive Development Theories, Social Learning, Information Processing Theories, Learning Models

- Human-Computer Interaction
- BigData, Spark, Kafka
- MLOps, Kubernetes, Docker, Ci/Cd
- NumPy for Array Operations
- Data Manipulation with Pandas
- Data Visualization with Matplotlib
- Data Exploration and Preprocessing
- Feature Engineering
- Machine learning and predictive analytics
- An end-to-end data science project
- Regression and Classification
- Model Training
- Decision Trees and Interpretability
- Ensemble methods and random forest
- Dimensionality reduction
- Unsupervised learning
- Deep learning
- Computer vision, CNN
- NLP and LLM
- Generative AI
- Privacy,Ethics, Trustworthy

Announcements   Official announcements are made through eLearn **only**.

Calendar   The course calendar is accessible on the course page on eLearn. Please note that the calendar is subject to change and may change throughout the semester.

Grading   Your final grade out of 20 is calculated as follows:

- 7*1.5 for assignments
- 4.5 for project
- 5 for final exam
- 1 for contest

Exams   One final exam is scheduled at the times announced in the course calendar. Please note that according to the university law, your presence at the final exam is **mandatory** for the course to be recorded in your transcript.

Assignments

There will be 7 assignments and a project throughout the course. You can do them **individually** or in a group of **maximum 3 students**. If done in a group, each of the members should be able to answer the questions asked about that assignment, or their score will decrease. You can change your group for each assignment, but all of the members for the assignment should be specified when it is about to be delivered.

Late Policy
- For every 1 day delay, there will be a 10% penalty on the Assignment final score.
- You will have a total of 8 days grace period.
- The final project will not be accepted after its initial deadline i.e. there will be no late and grace policy!

Cheating Policy

Your work is considered cheating if:
- You use the ready-made code available on the internet as a solution.
- You use someone else's code.
- You mistakenly upload someone else's code instead of your own!
- Copy the answers generated by Chat GPT without understanding

Penalties for cheating, especially to prevent the spread of cheating, are considered in this course as follows:
- Cheating in an assignment results in a 100% negative score for that assignment for each student involved, whether they are the cheater or the cheated upon;
- Cheating may be detected at any time during the term. If you have once performed an unacceptable act and have not been caught immediately, it does not mean that this check will never occur.
- Cheating in exams is subject to educational regulations and may result in the subject being referred to the educational committee and ultimately the disciplinary committee.

| **Resources** | • [Github Page](#) |
| | • [Previous Semester Github Page](#) |

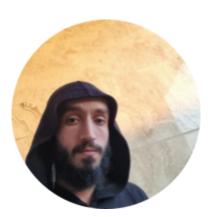| **Environment** | [Jupyter Notebook/JupyterLab](#) |

Staff:

Mohammad Amanlou

([Mohammad Amanlou | LinkedIn](#))

Chief Teaching Assistant

Mohammad Reza Nemati

([Mohammad Reza Nemati | LinkedIn](#))

Teaching Assistant

Shahriar Attar

([Shahriar Attar | Github](#))

Teaching Assistant

Mohammad Reza Mohammad Hashemi

Teaching Assistant

Fatemeh Mohammadi

(fatemeh mohammadi | LinkedIn)

Teaching Assistant

Ali Karimi

(Ali K. | LinkedIn)

Teaching Assistant

Mina Shirazi

([Mina Shirazi | LinkedIn](#))

Teaching Assistant

Mobina Mehrazar

([Mobina Mehrazar | LinkedIn](#))

Teaching Assistant

Javad Kavian

([Javad Kavian | Github](#))

Teaching Assistant

Mahdi Ebrahim Soltani

(Mahdi Ebrahim Soltani | Github)

Teaching Assistant

Hamid Salemi

(Hamid Salemi | LinkedIn)

Teaching Assistant

Mohammad Ranjbar

(Mohammad Javad Ranjbar | LinkedIn)

Teaching Assistant

Reihaneh Yourdkhani

(Reihaneh Yourdkhani | LinkedIn)

Teaching Assistant

Matin Bazrafshan

(Matin Bazrafshan | LinkedIn)

Supervisor Teaching Assistant

Mohammad Reza Alavi

(MohammadReza Alavi | LinkedIn)

Teaching Assistant

Farbod Azim Mohseni

(Farbod | Github)

Teaching Assistant

Javad Razi

(Javad Razi | LinkedIn)

Teaching Assistant

Hesam Ramezanian

(Hesam Ramezanian | LinkedIn)

Teaching Assistant

Ali Hamzepour

(Ali Hamzehpour | LinkedIn)

Teaching Assistant

Mohammad Amin Yousefi

(Amin Yousefi | LinkedIn)

Supervisor Teaching Assistant

AmirMahdi Farzaneh

(Amirmahdi Farzane | LinkedIn)

Teaching Assistant