

# CMPE 407 - Machine Learning

## Homework 01

Deadline: 29.03.2020 23.59

### RULES

- Codes without documentation (comments) will not be evaluated.
- This work is an individual study. Plagiarism is strictly prohibited, involved students will get zero.

### Submission

- Submit your solution as [id\_name\_surname\_cmpe407\_hw01].ipynb. Do not use Turkish characters in file naming.
- Solution without comments will not be evaluated.
- Late submissions will not be accepted. Submission system will be closed after deadline. Submissions via e-mail will not be accepted.

### QUESTION

1. (10 point) Read the Life Expectation dataset uploaded in Learn<sup>1</sup>. Provide analysis about statistics of the dataset.
2. (30 point) Visualize the features by using different plots. Provide discussion for each plot.  
Plots without discussion will **not** be evaluated.
3. (20 point) Make proper pre-processing operations to prepare the data to feed models.
4. (30 point) Train regression models on Life Expectation dataset. Make different experiments with different hyper-parameters or models.
5. (10 point) Test your models and provide discussion about your results.

**Note: Ensure that your programs are fully documented, using comments.**

---

<sup>1</sup>For details about dataset: <https://www.kaggle.com/kumaraajarshi/life-expectancy-who>