

BLM19462E
Term Project Instructions

For the final project, you must work in a group of up to three people. Each person in your group will choose one of machine learning algorithms that we covered in the course, experiment with applying that algorithm to a particular dataset or problem, and analyzing the results. Multi-person groups should work on the same dataset or problem and compare the results of their respective methods.

For your final project, your group will turn in a presentation:

- explain dataset /problem (number of features, number of data, number of missing value, distribution of data, range, min, max, mean, correlation between columns)
- briefly explain your algorithms
- describes your hypotheses and the experiments you did to test these hypotheses, (parameter selection)
- give the r, etc. esults of your experiments
- discuss the results with respect to your hypotheses,
- give a brief description of what you learned in doing the project

Important Notes:

1. Make sure to label the axes on any plots you give in the presentation. Unlabeled axes will decrease your grade!
2. Say what language you used. If you used external machine-learning libraries or other code, please cite what you used.

Good Luck
Berna Kiraz.

BLM19462E
Final Projesi Bilgileri

Final projenizde, en fazla 3 kişiden oluşan grupla çalışmanız gerekmektedir. Grubunuzdaki her bir kişi derste işlenen bir algoritmayı implement edecek, problem üzerinde deneyler yapacak ve sonuçları analiz edecektir. Gruplar aynı veri kümesi veya problem üzerinde çalışmalı ve sonuçlarını karşılaştırmalıdır.

Final projeniz için, grubunuz bir sunum hazırlayacaktır. Bu sunumda

- veri kümenizi açıklayınız (number of features, number of data, number of missing value, distribution of data, range, min, max, mean, correlation between columns)
- algoritmanızı kısaca açıklayınız
- hipotezinizi ve hipotezinizi test etmek için deneylerinizi tanımlayınız (parametre seçimi vs)
- Deneylerinizin sonucu veriniz
- Hipotezinize göre sonuçlarınızı tartışınız
- Projeyi yaparak ne öğrendiğinizi kısaca açıklayınız.

Önemli Notlar

1. Sunumunuzdaki grafiklerde eksenlere etiket eklemeyi unutmayınız. Aksi halde notunuz düşecektir.
2. Hangi dili kullandığınızı belirtiniz. Kullandıysanız hangi kütüphaneyi / başka kodları bunların referansını vermeniz gerekmektedir.

Sample Project Ideas:

Classification :

Titanic : <https://www.kaggle.com/c/titanic>

Forest Cover Type Prediction: <https://www.kaggle.com/c/forest-cover-type-prediction>

Don't Overfit 2 : <https://www.kaggle.com/c/dont-overfit-ii>

CareerCon 2019: <https://www.kaggle.com/c/career-con-2019>

Regression :

House Prices: <https://www.kaggle.com/c/house-prices-advanced-regression-techniques>

Bike Sharing Demand: <https://www.kaggle.com/c/bike-sharing-demand/data>

Predict Future Sales: <https://www.kaggle.com/c/competitive-data-science-predict-future-sales>

TMDB Box Office Prediction: <https://www.kaggle.com/c/tmdb-box-office-prediction>

Clustering:

Credit Card Dataset for Clustering: <https://www.kaggle.com/arjunbhasin2013/ccdata>

Medical Cost Personal Datasets: <https://www.kaggle.com/mirichoi0218/insurance>