**TOPKAPI UNİVERSİTESİ**

**MAKİNE ÖĞRENMESİ-Final Ödevi**

**\*FİNAL ÖDEVİ TOPKAPI ALMSCLOUD A YUKLENMELIDIR.**

**\*ODEVIN ÇIKTISI ALINARAK ELDEN TESLIM EDILMELI VE ODEV YOKLAMA TUTANAĞI İMZALANCAKTIR. TESLIM EDILMEYEN VE YOKLAMA IMZALANMAYAN ÖDEVLER GECERSIZ SAYILACAKTIR.**

**\*ÖDEV KODLARI VE PERFORMANS ÖLCÜTLERİ İÇİN HESAPLANACAK PERFORMANS METRİKLERİNİN (ROC EĞRİLERİ CONFUSION MATRIX VE SENSITIVITY, SPECIFICITY, RECALL VS) ÇIKTILARI GITHUB LINKI HALINDE VERILMELIDIR**

1. **Size atanan veri setine tüm algoritmalar için optimizasyon uygulayınız.**
2. **Veri setinizi rastgele olarak %70 eğitim %30 test olacak şekilde ayırınız. Eğitim veri seti için Naive bayes sınıflandırıcısını uygulayınız. Elde ettiğiniz sonucları raporlayınız.**
3. **Veri setinizi rastgele olarak %70 eğitim %30 test olacak şekilde ayırınız. Eğitim veri seti için K-en yakın komşuluk sınıflandırıcısını uygulayınız. En iyi k değerini belirleyerek Elde ettiğiniz sonuçları raporlayınız.**
4. **Veri setinizi rastgele olarak %70 eğitim %30 test olacak şekilde ayırınız. Eğitim veri seti için Multi-Layer Perceptron (MLP) ve Support Vector Machines (SVM) sınıflandırıcılarını uygulayınız. Eğitim ve test adımlarında elde ettiğiniz sonucları raporlayınız.**

**\*\*Maddelerde Sonuçları raporlama olarak adlandırılan kavram sonuçların konfuzyon matrisinin verilmesi, sensitivity/specificity/accuracy/f1-score veya ROC curve gibi parametreleri ile açıklanması anlamına gelmektedir.**

**Dataset**

Pima Indians Diabetes Dataset

The Pima Indians Diabetes Dataset involves predicting the onset of diabetes within 5 years in Pima Indians given medical details.

It is a binary (2-class) classification problem. The number of observations for each class is not balanced. There are 768 observations with 8 input variables and 1 output variable. Missing values are believed to be encoded with zero values. The variable names are as follows:

Number of times pregnant.

Plasma glucose concentration a 2 hours in an oral glucose tolerance test.

Diastolic blood pressure (mm Hg).

Triceps skinfold thickness (mm).

2-Hour serum insulin (mu U/ml).

Body mass index (weight in kg/(height in m)^2).

Diabetes pedigree function.

Age (years).

Class variable (0 or 1).