

MAKİNE ÖĞRENMESİ DERSİ PROJE/PERFORMANS ÖDEVİ

KESİNLİKLE DİKKAT EDİLMESİ GEREKEN HUSUSLAR:

- **MS-Word** formatında yazılı bir raporunuz ve proje için yazdığınız python kodlarınız tek bir **zip** dosyası olarak **SABİS** sistemine yüklenmelidir. Veri setinizi **yüklemeyeceksiniz**. Sadece rapor içinde veri setinin açıklandığı kısımda veri setinin linkini paylaşmanız yeterli olacaktır.
- Proje raporu son teslim tarihi: **01 Aralık 2025 23:50** dir. Sisteme proje rapor ve kod dosyasını yüklemeyen Proje Grubu Proje/performans ödevinden **0 (sıfır)** almış olarak notlandırılacaktır.
- Proje grubu **4 veya 5** kişilik olmalıdır. Grup üyeleri 1. Öğretim ve 2. Öğretim karışık olabilir.
- Proje grubunun bir sorumlusu olacak ve **raporu sadece o kişi sabis sistemine yükleyecektir**. Diğer üyeler kesinlikle **yüklememelidir**. Raporun sisteme yüklenmesi tüm grup üyelerini etkilediği için raporun yüklemesinin tamamlanması grubun sorumluluğundadır.
- Kodlarınız Jupyter Notebook veya Python script formatında olmalıdır. Ama Jupyter Notebook'da açıklamalı bir şekilde yapılmasını tavsiye ederim.
- Sunum yüz yüze ve ders saatinde yapılacaktır. Online kesinlikle kabul **edilmemektedir**.
- Sunumda grubun tüm üyeleri gelmeli ve sunumda yaptıkları kısımları anlatmalıdır. Grubun içinden sunuma gelmeyen öğrenci **GR (Girmedi)** almış olarak notlandırılacak, kendi görevli olduğu iş paketini anlatmayan kişiler ise **0 (sıfır)** almış olarak notlandırılacaktır.
- Rapor kapak sayfasında projenin adı, proje grup sorumlusu, proje grup üyelerinin adı, öğrenci numarası mutlaka yazılmalıdır.
- Proje sunumunları **için tarih ayrıca duyurulacaktır**.

Makine Öğrenmesi Proje/Performans Ödevi Kapsamı

Proje grubu ve proje grubundaki öğrencilerin herbirinin görevi ve proje kapsamında yaptığı işler belirtilip açıklanacaktır. Projenin amacı ve hedeflenen sonuç nedir belirtilecektir.

Sonrasında seçilen veriseti için aşağıdaki adımları gerçekleştireceklerdir.

1. Veri Seti Seçimi ve Tanıtımı

- Uygun bir veri seti bulun. Örneğin, Kaggle, UCI Machine Learning Repository veya diğer açık kaynaklardan faydalanabilirsiniz.
- Seçtiğiniz veri setinin kaynağını ve ne amaçla kullanıldığını açıklayın.
- Veri setinin yapısını tanıttın (veri boyutu, özellik sayısı, veri türleri, sınıf dağılımı vb.).
- Veri seti içinde anormal ve problemli veriler varsa bunları açıklayıp ne olduğunu ve neler yapıldığını belirtin.
- Veri setindeki hedef değişkeni açık bir şekilde belirtin.

2. Veri Ön İşleme (Preprocessing)

- **Eksik Veriler:** Veri setindeki eksik veya tutarsız verileri tespit edin ve bunları doldurma, düzeltme veya kaldırma vb. yöntemlerden birini uygulayın.
- **Normalizasyon ve Standardizasyon:** Gerekliğinde özelliklerin değer aralıklarını normalize edin veya standardize edin.
- **Feature Engineering:** Gerekirse yeni özellikler oluşturun.
- **Feature Selection/Importance:** Önemli özellikleri belirlemek için uygun teknikleri uygulayın ve bunu raporlayın.

3. Veri Analizi

- Veri setindeki sınıf dağılımını ve dengesizliği analiz edin.
- Eğer veri dengesiz ise, SMOTE gibi veri dengeleme tekniklerini kullanarak dengeli bir veri seti oluşturun.
- Veriyi anlamak için grafiksel analizler yapın (örneğin, histograms, scatter plots, correlation heatmaps).

4. Makine Öğrenmesi Modelleri

- En az 4 farklı makine öğrenmesi modeli seçin (örneğin, Decision Tree, Random Forest, SVM, KNN, Logistic Regression, XGBoost, vb.).
- Bu modelleri veri setinize uygulayın.

- Uygulama yaparken farklı hiperparametreler için de denemeler yaparak her model için en iyi hiperparametreler belirlenmeli.
- Model eğitimi sırasında eğitim/test oranını belirtin. Cross Validation veya Hold-Out hangi yöntem ile verileri böldüğünüzü belirtin ve adımları açıklayın.
- Modelde overfitting varsa overfitting'ı önlemek için gerekli yöntemleri uygulayın ve yaptığınız uygulamayı ve modelin durumunu açıklayın.
- Model performansını ilgili metriklerle göre değerlendirin. Sonuçları verin ve açıklayın.

5. Açıklanabilir Yapay Zeka

- Açıklanabilir yapay zeka yöntemleri araştırılacak ve uygun bir yöntem seçilip model için uygulanacaktır. Elde edilen sonuçlar proje dosyasında ve sunumda açıklanacaktır.
(Bunu nasıl yapmanız gerektiği ve nasıl uygulamanız gerektiği sizin öğrenmeniz gereken kısım. O nedenle bu kısım ile ilgili bana nasıl yapmanız gerektiğini sormayınız. Bu sizin ödev olarak araştırıp yapmanız gereken kısım.)

6. Sonuçların Karşılaştırılması

- Modellerin sonuçlarını karşılaştırmalı bir tablo ile sunun.
- Her bir modelin avantajlarını ve dezavantajlarını açıklayın.
- Performans sonuçlarını grafiklerle gösterin (örneğin, bar chart, line graph, vb.).

7. Sonuç ve Öneriler

- Çalışmanızın genel sonuçlarını özetleyin.
- Model seçiminde neden bazı modellerin daha iyi performans gösterdiğini tartışın.