### **MATLAB ile Metin Sınıflandırma: Ham Kod Geliştirme (Scripting) ve Classification Learner (Apps) Yöntemlerinin Karşılaştırmalı Analizi**

### **Hazırlayan: (Efe Can Kara / 2022481031)**

### **PROBLEMİN TANIMI :** Proje, sosyal medya yorumlarından (52.681 adet) oluşan bir veri setini 7 karmaşık ve anlamsal olarak benzer duygu sınıfına (Örn: 'Anxiety', 'Stress', 'Depression') ayırmayı hedefler.

Asıl teknik problem, MATLAB ortamında bu görevi gerçekleştirmek için kullanılan iki temel yaklaşım olan **Ham Kod (Script)** ve **Classification Learner (Apps)** yöntemlerinin; büyük ve yüksek boyutlu (14.414 özellik) verideki **eğitim hızı**, **başarı oranı** ve **geliştirme verimliliği** açısından adil bir şekilde karşılaştırılmamış olmasıdır.

### **ÇÖZÜM ÖNERİSİ:** İki yöntem arasında adil bir kıyaslama (apples-to-apples comparison) yapmak için aşağıdaki adımlar önerilmektedir:

1. **Merkezi Veri Hazırlığı:** Tüm veri seti ham kod ile tek seferde işlenerek 14.414 özellikli bir TF-IDF matrisi (X) ve etiket vektörü (Y) oluşturulacaktır. Bu, her iki yöntemin de **birebir aynı veriyi** kullanmasını garanti edecektir.
2. **Yöntem 1 (Ham Kod):** X ve Y verileri kullanılarak, fitctree (Karar Ağacı) modeli eğitilecektir. 5-Katlı Çapraz Doğrulama (crossval) uygulanacaktır. Modelin **eğitim süresi** (tic/toc ile) ve **başarı oranı (%)** kaydedilecektir.
3. **Yöntem 2 (Apps):** Classification Learner uygulamasına, ham kodda oluşturulan **aynı X ve Y** değişkenleri yüklenecektir. Ham kodla aynı ayar olan 5-Katlı Çapraz Doğrulama ve **"Medium Tree" (Karar Ağacı)** modeli seçilerek eğitim yapılacaktır. Modelin **eğitim süresi** ve **başarı oranı (%)** kaydedilecektir.

### **AVANTAJLARI**

* MATLAB'ın "Kod" ve "Uygulama" yaklaşımları arasında, büyük veri üzerinde **net bir hız ve doğruluk kıyaslaması (benchmark)** sağlar.
* MATLAB Apps'in, ham kodun varsayılan ayarlarına kıyasla ne kadar **optimize** olduğunu (veya olmadığını) sayısal olarak ortaya koyar.
* Geliştiricilere, benzer projeler için hangi yöntemin (hızlı geliştirme/Apps mi, esnek kodlama/Script mi) daha verimli olacağı konusunda kanıta dayalı bir rehber sunar.

### **SONUÇ RAPOR FORMATI :** Sonuç raporu, iki yöntemin bulgularını özetleyen **Nihai Karşılaştırma Tablosu** üzerine odaklanacaktır. Bu tablo şu sütunları içerecektir: Yöntem (Ham Kod / Apps), Algoritma (Karar Ağacı), Veri Seti (%100), Eğitim Süresi (Saniye), Başarı Oranı (Accuracy %)

Rapor, bu tablodaki **Hız** ve **Başarı** farklarının nedenlerini analiz ederek (Örn: Apps optimizasyonları, problemin zorluğu) son bulacaktır.