

## PROJE RAPORU İÇERİK VE RAPOR HAZIRLIĞI İÇİN YARDIM KILAVUZU

Proje raporu dönem boyunca proje kapsamında yaptığınız çalışmaların sonuçların sunulduğu rapordur. Aşağıda projede yapacağınız çalışmalar ve bu çalışmaların sonucunda üreteceğiniz çıktıların gösterimi ile ilgili yol gösterici bilgiler verilmiştir. Proje raporu aşağıdaki bölümlerden oluşmalıdır: (Yapılan çalışmaya ve problem tipine bağlı olarak bazı kısımlar atlanabilir veya yeni eklemeler yapılabilir)

### 1. Özet

### 2. Giriş – problem tanımı

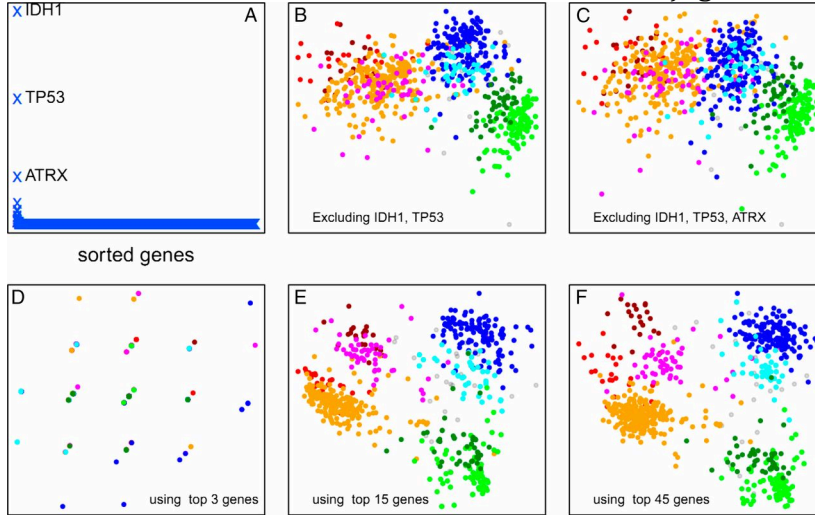
- Motivasyon
- Sınıflandırma / regresyon
- Amaç/Hedef
- Başarım metrikleri (%90 doğru sınıflandırma, vb.)

### 3. (Varsa) Literatür araştırması

### 4. Veri Seti, Veri özellikleri, Öznitelikler

- Veri kaynağı
  - o URL Link, in-house veriler, vb.
- Veri kümesi
  - o Kısaca öznitelik ve veri seti tanımları
- Ön işleme aşamaları (Bunu ayrı bölüm olarak da oluşturabilirsiniz)  
Burada eğer hazır veri seti kullanmıyorsanız (veya kullandığınız hazır veri seti ham veri seti ise) yaptığınız veri işleme işlemlerini açıklayın.
  - o (varsa) görüntü işleme aşamaları
  - o Ses işleme
  - o Veri işleme, vb
- Öznitelik açıklamaları
  - o Hatalı veriler, eksik veriler
  - o Structured / unstructured
  - o Sayısal değerler
    - Ordinal veriler
      - Sıralama, kategorizasyon
    - Nominal veriler
      - Veriler sıralı mı ? (Ağırlık, sıcaklık, vb)
      - Veriler karşılaştırılabilir değil (Şehir plaka no)
  - o Sayısal olmayan veriler
    - İki sınıflı veriler (erkek/kadın – 0-1 veya -1 – 1)
    - Çok sınıflı veriler
      - Sıralı mı ? (Harf notu gibi: AA, BA, BB, ...)
      - Sırasız mı ? (Ağaç çeşidi gibi: Çam, ardıç, söğüt,...)
    - Düz metin verileri
      - Kelimeler, tweetler, metinler, vb.
- Öznitelik seçimi veya düzenlemesi varsa bu işlemlerin detayları (PCA – Principal Component Analysis) veya ICA (Independent Component Analysis)

- Öznitelik çıkarımında başka probleme özgü işlemler
- Sınıf veri dağılımları bilgileri
  - o Data balanced / imbalanced ? Sınıflardaki veri sayılarını eğitim verileri için dengele
- Data normalizasyonu, öznitelik normalizasyonu (lineer, normal, log, özel bir dönüşüm, vb.)
- Verinin görsel olarak 2-D gösterimi (sınıf değerleriyle beraber) multidimensional data visualization Aşağıda örnek bir grafik var



- Öznitelikler arasındaki ilişkiler
  - o Öznitelikler arasında korelasyonlar
  - o Özniteliklerin sonuç ile korelasyonu
  - o sınıf dağılımları istatistiksel parametreleri (ne tip bir dağılım, Normal, durağan, Poisson, vb)
  - o Normal dağılım durumunda Kovaryans matrisi
- Verinin tamamını sınıf sayısı kadar kümeleye ayır (clustering) (normalizasyondan sonra k-means yapılabilir). Bu doğal kümeleme işleminin (natural clustering) gerçek sınıflar ile karşılaştırılması (Burada da verinin görsel olarak 2-D gösterimi (sınıf değerleriyle beraber) multidimensional data visualization kullanılarak sonuçlar gösterilebilir.

## 5. Kullanılan modeller

- Eğitim / çapraz doğrulama/test ayırımı (60,20,20)
- ikili ayırma, multiclass, vb.
- Kullanılan modellerle ilgili kısa tanımlar, metodolojiler, literatür bilgisi
- Neden bu modelleri seçtik ?

## 6. Test sonuçları (Results) ve sonuçların yorumları, Tartışma (Discussion)

- Sınıflandırma için Karmaşıklık matrisi, TP, FN, FP, FN
- Regresyon için MSE, RMSE, E (Error), PE (Percent Error), AE (Absolute Error), APE (Absolute Percent Error), vb performans metrikleri
- Modeller arası karşılaştırma tabloları ve grafikleri

- Değişik modeller arasında İstatistiksel doğruluk ve önem testi (statistical significance test, t-test)
- Aynı modeli 10-20 kere çalıştırınca elde ettiğiniz sonuçlar ve bunların birbirleri ile istatistiksel karşılaştırmaları (statistical significance)
- Hangi model neden daha iyi ? açıklama ?

## **7. Sonuçlar (conclusions)**

- Bu çalışmada kısaca ne yaptığımızın özeti
- Çalışma sonucunda ne öğrendik ? Kazancımız nedir ? Katkımız ne oldu ?
- Çalışmada yap(a)madıklarımız neler var ? Bunların sebepleri neler ?
- Gelecek çalışmalar neler olabilir ?

## **8. Referanslar**

Not: Rapor uzunluğu IEEE konferans bildiri formatında (Word ve Latex) 4-8 sayfa arasında olacaktır. Buna ek olarak Ekler (Appendices) adı altında bir bölüm daha oluşturup bunu bildiriden ayrı olarak verebilirsiniz. Appendix içeriğinde özellikle Önişleme, Öznitelik seçimi, korelasyonlar, vb. gibi bilgilerin yanı sıra program kodu da konabilir.