

**Ödev Konusu: Optimizasyon algoritmalarını karşılaştırma**

Ödev, 2 kısımdan (A ve B) oluşmaktadır.

**A: Metin sınıflandırma**

C programlama dilinde yazılacak.

2 sınıflı her sınıfta en az 100 örnek içeren bir metin veri kümesi oluşturun.

Metinleri okuyup bir sözlük oluşturun.

Metinleri 1-hot vektörler haline getirin

Eğitim ve test kümesine böl (%80 / %20)

Model:

A sınıfı için çıkış 1, B sınıfı için çıkış -1

$\text{çıkış} = \tanh(w \cdot x)$  boyutlar:  $(1 \cdot v)(v \cdot 1) = (1 \cdot 1)$  v: sözlük boyutu

x: 1-hot vektör

w: öğrenilecek parametreler

5 farklı ilk w değeri için GD, SGD ve Adam'ı karşılaştır (yatay eksen: süre, güncelleme sayısı, dikey eksen: eğitim ve test başarıları kriterlerine göre)

Süre vs. loss ve Epoch vs. loss grafikleri çiz.

Bonus: Token olarak kelimeler yerine harf 3-gramlarını kullanmanın etkisini araştırmak

**B: Optimizasyon sürecinin 2 boyutta gösterimi**

5 farklı ilk w değeri için  $w_{1:t}$  değerlerini kaydet

$w_{1:t}$  leri T-sne ile 2 boyuta indir, görselleştir. 1 grafikte 5 yörünge olacak.

Grafikleri yorumla.

**Not: B kısmı ve grafik çizme haricindeki tüm kodlar C de yazılacaktır.**

**Ödevin Son Teslim Tarihi: 1 Aralık 2023 saat 09:00**

**Ödevin Teslim Şekli: [online.yildiz.edu.tr](https://online.yildiz.edu.tr)**

Bu ödevde en fazla 2 kişilik gruplar halinde çalışacaktır.

**Ödevde teslim edilecekler:**

Ödev raporu (PDF formatında 1 adet belge) içeriği:

1-A ve B kısmında yaptıklarınız, elde ettiğiniz görseller, yorumlar (3P)

2-Açıklamalı program kodlarınız (2P)

3-Projenizi anlattığınız en az 5, en çok 8 dakika uzunluğunda, sizin görüntünüzü de içeren ve youtube'a yüklenmiş videonun adresi (5P)