

Kolektif Öğrenme 2024/1

Ödev No:2

Ödev Konusu: Temsillerin / Kararların Birleştirilmesi

Farklı temsil yöntemlerini farklı ML algoritmalarıyla kullanmak

5 temsil * 3 alg (SVM, RF, MLP) = 15 sonuç.

Aynı temsil yöntemine ait sonuçları (5 adet ensemble) birleştirmek

Aynı algoritmaya ait sonuçları (3 adet ensemble) birleştirmek

15 sonucu birleştirmek (1 ensemble)

Yukarıdaki 9 sonucu karşılaştırmak

Temsil yöntemleri için örnekler:

<https://huggingface.co/sentence-transformers/all-MiniLM-L12-v2>

<https://huggingface.co/jinaai/jina-embeddings-v3>

<https://huggingface.co/intfloat/multilingual-e5-large-instruct>

<https://huggingface.co/BAAI/bge-m3>

<https://huggingface.co/thenlper/gte-large>

<https://huggingface.co/nomic-ai/nomic-embed-text-v1>

<https://huggingface.co/ytu-ce-cosmos>

<https://huggingface.co/dbmdz/bert-base-turkish-uncased>

Seçtiğiniz **en az 2 Türkçe metin veri kümesi** (her biri en az 5 bin örnek içeren) üzerinde karşılaştırma yapılacaktır. Performans ölçütü olarak accuracy / F1 kullanılabilir. Her veri kümesi için test örneklerini tüm işlemlerden önce ayırıp tüm yöntemlerde aynı test kümelerini kullanınız.

Sonuç birleştirmede demokrasi usulü kullanımı yeterlidir.

Grafikler (tablolar) sizin bulgularınız. Bunlara dair yorum yapınız. Beklentilerinizle uyuşanlar uyuşmayanlar vb.

Sınıfta sunum yapmak %70 (yapmayanlar en fazla %30 alabilirler)

Önemli Not: Deney sonuçlarınızı, bulgularınızı iyi ifade eden tablolar ve grafiklerle sununuz.

Ödevin Son Teslim Tarihi: 3 Aralık 2024 saat 09:00

Ödevin Teslim Şekli: online.yildiz.edu.tr

Bu ödevde **en fazla 2 kişilik gruplar** halinde çalışabilirsiniz.

3 Aralık'taki derste raporunuz üzerinden ödevinizi sunmanız beklenmektedir.

Ödevde Teslim Edilecekler: 1-Açıklama içeren kodlarınız (değişkenlerin ne için kullanıldıkları, algoritmanın adımları) 2-Ödev raporu