

You Do #1

1. We Do sessionda modellediğimiz ev fiyatı tahminlemesiyle alakalı iş biriminden yeni bir talep var !!!

- Kendileri Data Science ekibiyle kahve içerken yaptıkları görüşmede fiyat tahminlemelerinde yaptığımız hataların, toplam hataya katkılarının hata miktarı arttıkça arttığını anlamışlar.
- Ancak geri gelen bildirimde belirli bir threshold sonrasında (θ) yapılan hataların aslında çok da fark yaratmadığını, örneğin gerçek fiyat 5 USD iken 8 USD tahmin yapmak ile 10 USD tahmin yapmanın eşit cezalandırılması gerektiğini iletmişler. En azından bunun mümkün olup olmadığını sormuşlar 😊
- Data Science ekibi olarak hızlıca söyle bir sprint dönelim istiyoruz
 - Müşterimizin tanımına uygun bir yeni bir convex (almost convex/not strongly convex or strongly convex) loss function tanımlayalım.
 - İşin matematiğine çok girmeden β_0, β_1 için fonksiyonun convex olduğuna kendimizi inandıralım.
 - Not: Ben görselden pek anlamıyorum dolayısıyla convexity ispatlamak için daha formal bir yoldan gitmek istiyorum diyenler more than welcome.
 - Loss function için tanımlı olan yerlerde β_0, β_1 için ∇ (gradient) hesaplarımızı yapıp gradient descent update fonksiyonlarımızı yazalım.
 - Verilen data üzerinden loss fonksiyonumuzu minimize edecek β_0, β_1 değerlerini bulalım.
 - Bir önceki aşamada geliştirdiğimiz modeli $L2$ regularized versiyon olarak genel bir forma getirelim.
 - Son olarak We Do modeli ve You Do da geliştirdiğimiz iki model plotly de nasıl görünüyor inceleyelim.

1 haftamız. Kolaylıklar diliyoruz...

Aşağıda veri yükleme için kullandığımız python kodunu bulabilirsiniz.

```
cal_housing = fetch_california_housing()
X = pd.DataFrame(cal_housing.data, columns=cal_housing.feature_names)
y = cal_housing.target

st.dataframe(X)

df = pd.DataFrame(
    dict(MedInc=X['MedInc'], Price=cal_housing.target))
```

Not: Kolaylık olması için HouseAgeGroup konusunu dışarıya çıkarttık. Lütfen sadece **MedInc** ve **Price** üzerinden ilerleyelim