Recommendation
System
&
Web App

**UP School** 

## **Know How**

- Projeye Genel Bakış
- Proje Metodu
- Uygulama
- Geliştirme
- Kazanımlarım

## **Know Self**

- Neler yaptık?
- Temel Kazanımlarım

## Projeye Genel Bakış

#### **Proje Amacı**

 E-ticaret sektöründe yer alan Trendyol firmasının sağladığı veriler ile kullanıcılara yönelik bir öneri sistemi geliştirilmesi



#### Veri Seti İncelemesi

	partition_date	orderparentid	user_id
0	20.08.2020 06:00	335057357	86386
1	24.08.2020 10:00	337401625	59469
2	26.08.2020 19:00	338457012	51248
3	22.08.2020 11:00	336681542	29380
4	20.08.2020 19:00	335736916	68368

user_id	productcontentid	brand_id
86386	39328996	919155
59469	31903343	121
51248	34726400	40
29380	32920640	7651
68368	39035716	3395

catego	ry_id	category_name	gender	price	color_id	business_unit
	418	Sandalet	Kadın	319.98	14.0	Branded Shoes B
	1827	Banyo Dolabi	Unisex	1195.56	3.0	Bahçe & Yapı Market & Hırdavat
	604	T-Shirt	Kadın	37.99	14.0	PL Woman
	604	T-Shirt	Kadın	24.99	6.0	Kadın A
	599	Kazak	Kadın	79.90	16.0	Kadın A

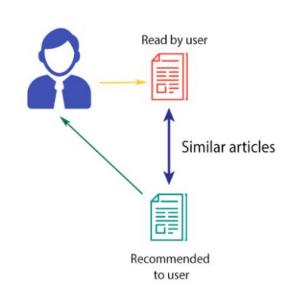
ImageLink
https://cdn.dsmcdn.com//ty1/product/media/imag
https://cdn.dsmcdn.com//assets/product/media/i
https://cdn.dsmcdn.com//assets/product/media/i
https://cdn.dsmcdn.com//assets/product/media/i
https://cdn.dsmcdn.com//ty1/product/media/imag

## Algoritma & Metod

#### COLLABORATIVE FILTERING

## Read by both users Similar users Read by her, recommended to him!

#### CONTENT-BASED FILTERING



Recommendation techniques

#### Proje Algoritmaları

- Popüler Ürünlerin Önerilmesi Cold Start Durumu
- İçeriğe Dayalı Öneri (Content Based Recommendation) Yöntemi
  - TF-IDF Metodu
    - Bir belge içinde geçen kelimelerin birbiriyle olan bağlantısını ölçen istatistiksel bir ölçüdür.

Tf(Term frequency) = Kelimenin ilgili dökümanda geçme sayısı/Kelime sayısı

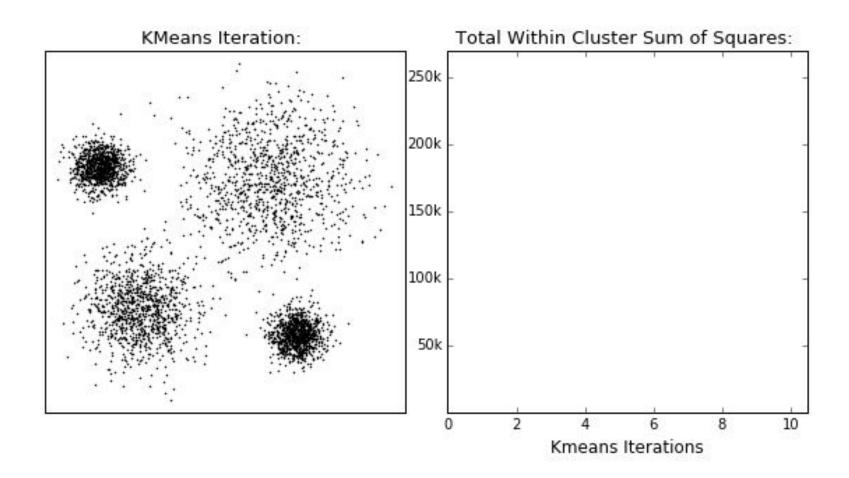
Idf(Inverse document frequency) = Toplam döküman sayısı/kelimenin geçtiği döküman sayısı

#### Proje Algoritmaları

- K-Means Kümeleme(Clustering) Metodu
  - Kümeleme, bir veri setinde benzerlik gösteren verilerin gruplara ayrılması işlemidir.
  - Gözetimsiz Öğrenme (Unsupervised Learning) algoritmasına ait bir metoddur.
  - K-Means' teki K sayısı küme sayısını belirler ve parametre olarak alması gerekir.



 K sayısı belirlendikten sonra algoritma rastgele K tane merkez noktası seçer. Bu rastgelelik problemlere yol açabilir. Çözüm K-Means++ yöntemidir.

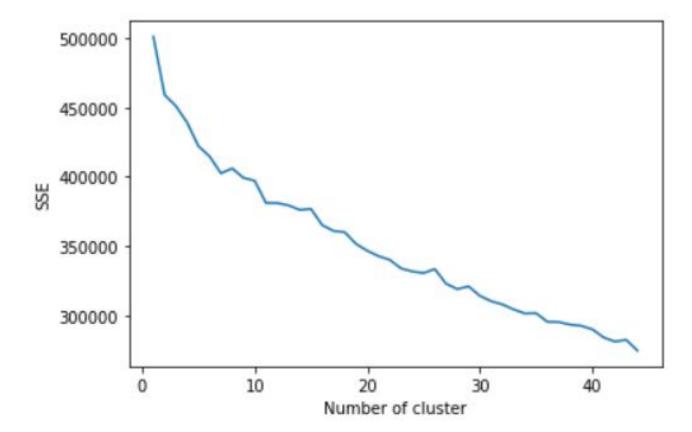


## Sonuçlar

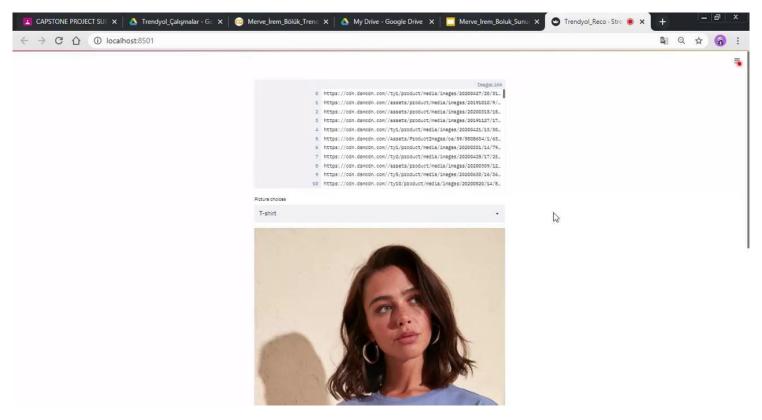
```
Top terms per cluster:
Cluster 0:
  medikalmaske
  gömlek
  pijamatakımı
  terlik
  jeans
  sweatshirt
  güneşgözlüğü
  poloyakatshirt
  tayt
  çorap
```

```
Cluster 1:
tshirt
sövale
gümüşlük
gözfarı
gözkalemi
gözlükaksesuarı
gözlükaksesuarı
gözlükkabı
gözmakyajtemizleyicileri
göğüspedi
göğüspompası
```

```
Cluster 2:
elbise
şövale
gümüşlük
gözfarı
gözkalemi
gözlükaksesuarı
gözlükkabı
gözmakyajtemizleyicileri
göğüspedi
göğüspompası
```



## **Uygulama**



## Geliştirme

- Standardizasyon
- Optimum metrik
- Farklı algoritmalar
  - Cosine similarity
  - Collabrative filtering

### **KAZANIMLARIM**

- Proje yönetimi
- Gerçek hayat problemi üzerine çalışma imkanı
- Ne-Neden-Nasıl pratiği
- Recommendation & WebApp ilk uygulama

## Biraz da mizah :)



## **Know Self**

#### Neler yapıyoruz?

- Seyirler
- MTC
- UPGrade Week



#### Temel Kazanımlarım

- Kendini keşif
- Feedforward
- İçten takdir
- Dayanışma
- İlham

# Teşekkürler