



# Makine Öğrenmesi Destekli Anlamsal Ürün Arama Motoru

BIL 496  
İlk Sunum

Simge Sarıçayır

Proje Danışmanı: Fatih Erdoğan Sevilgen  
Kasım 2019



# Proje Şeması ve Tanımı

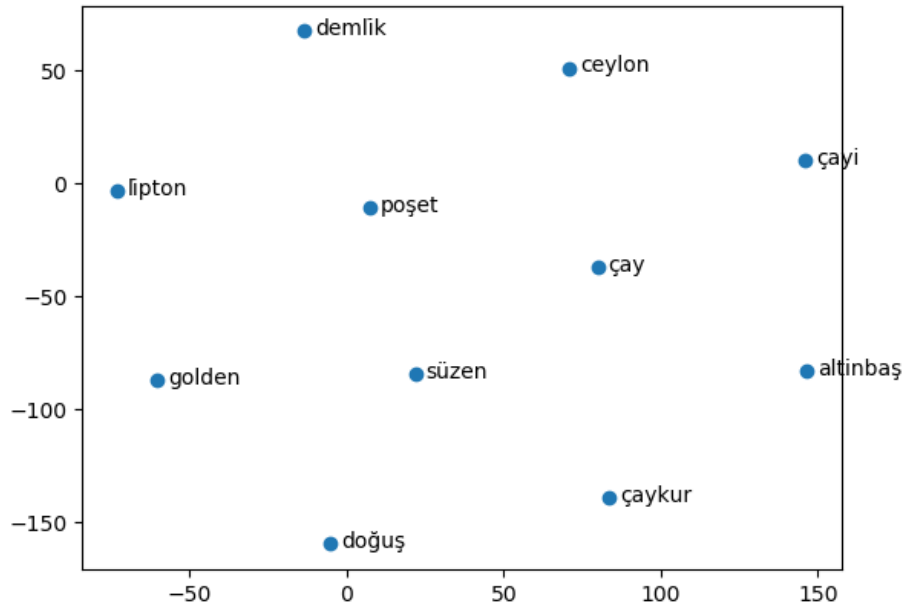


- Proje nedir?

E-ticaret siteleri için makine öğrenmesi destekli anlamsal ürün arama motoru.

- Günümüzde yaygınlaşan e-ticaret sitelerinde doğru ürün listeleme ile satışın artırılması.
- Satın alınmak istenen ürünün arama sonucunda doğru sırada bulunamaması.
- Yapılan listelemede istenen ürünün yer almaması.

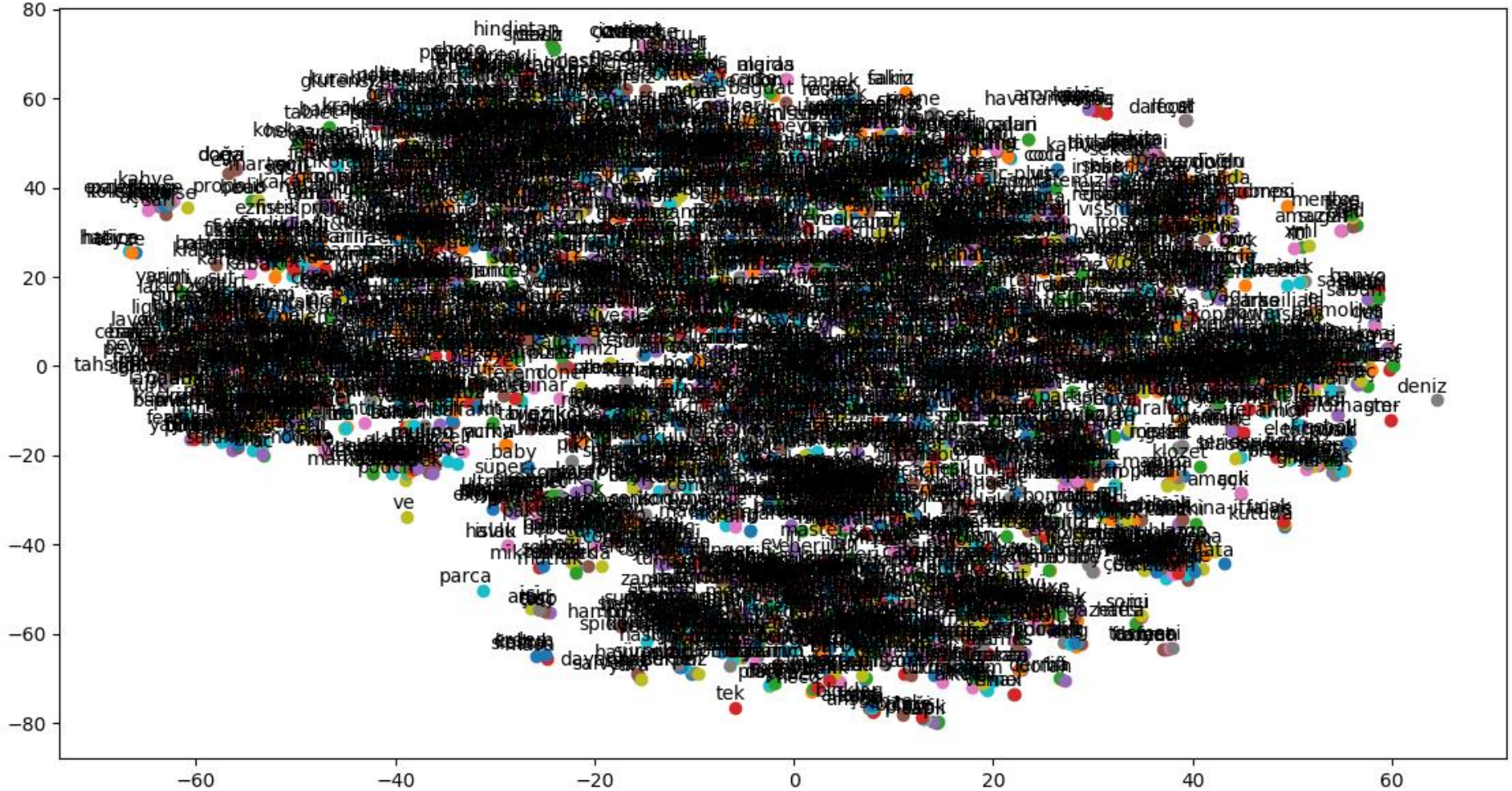
- Ürün isimleri üzerinde ön işleme yapılarak noktalama işaretlerinden arındırıldı.
- Ön işleme yapılmış olan veriler kullanılarak gensim kütüphanesi ile Word2Vec modeli eğitildi.



Çay kelimesine karşılık modeldeki yakın kelimelerin gösterimi



# Kelimelerin Vektör Uzayındaki Yerleri







- Eğitilen Word2Vec modeli kullanılarak arama yapılan kelimelere karşılık modelde en yakın kelimeler bulundu ve ürün listesinden bu kelimeleri içeren ürünler listelendi.
- Modelin iki kelime arasındaki benzerlik oranını vermesi ile ürün isimleri üzerinde bu benzerlik oranı kullanılarak sıralamada puanlama yapıldı.

Keyword: diş

('lux', 0.9232676029205322)  
( 'ipana', 0.9019483327865601)  
( 'signal', 0.9002252817153931)  
( 'pronamel', 0.893963098526001)  
( 'pro', 0.8931301236152649)  
( 'sensodyne', 0.8908178806304932)  
( 'total', 0.8875018358230591)  
( 'diş', 0.8871555924415588)  
( 'pro-expert', 0.886247992515564)  
( 'glamorous', 0.8778396248817444)  
( 'clean', 0.8733474016189575)  
( 'true', 0.8727699518203735)  
( 'beyazlik', 0.868601381778717)  
( 'hassasiyete', 0.8660621643066406)  
( 'white', 0.8402878642082214)  
( 'expert', 0.8391355872154236)  
( 'med', 0.8387041687965393)  
( 'boyutlu', 0.835991382598877)

İPANA 3D WHITE LUXE GLAMOROUS SHINE DİŞ MACUNU 75M  
PARODONTAX ULTRA CLEAN DİŞ MACUNU 75 ML  
COLGATE TOTAL PRO DİŞ ETİ SAĞLIĞI DİŞ MACUNU 75ML  
COLGATE OPTIC WHITE EXPERT WHITE DİŞ MACUNU 75ML  
İPANA PRO-EXPERT CLİNİC LİNE DİŞ MACUNU DİŞ K. 75M  
İPANA PRO-EXPERT DİŞ ETİ KORUYUCU DİŞ MACUNU 75 ML  
İPANA 3 BOYUTLU BEYAZLIK LUXE SİGARA İÇENLER İÇİN  
SIGNAL DİŞ MACUNU EXPERT PROTECTION BEYAZLIK 75ML  
SENSODYNE TRUE WHITE EKSTRA FRESH DİŞ MACUNU 75 ML  
SENSODYNE TRUE WHITE NANELİ DİŞ MACUNU 75 ML  
SENSODYNE TRUE WHITE SOFT DİŞ FIRÇASI  
COLGATE TOTAL PRO NEFES SAĞLIĞI DİŞ MACUNU 75ML  
COLGATE TOTAL PRO AKTİF ETKİ DİŞ MACUNU 75ML  
SENSODYNE TOTAL CARE-TAM KORUMA SOFT DİŞ FIRÇASI



- İki farklı e-ticaret sitesinden elde edilmiş olan ürün açıklama verileri ve arama motorunun gerçekleştirildiği ürün isimleri birleştirilerek başka bir Word2Vec modeli eğitildi.
- Oluşturulan diğer model ile aynı kelimeler üzerinde karşılaştırma yapıldığında ilk oluşturulan model daha iyi sonuç verdi.



# Yapılması Planlananlar

- Wikipedia'dan elde edilen türkçe kelimelerden oluşan corpus kullanılarak eğitilmiş model ürün isimleri kullanarak tekrar eğitilecek.
- İki farklı e-ticaret sitesinden elde edilmiş ürün açıklama verileri ve arama motorunun gerçekleştirildiği ürün isimleri kullanılarak eğitilmiş model üzerine tekrar eğitilecek.
- Elde edilen word2vec modeli kullanılarak arama sonuçlarının anlamsal yakınlıktaki ürünler getirmesi sağlanacak. Bunun için ürün isimlerindeki kelimeler bir bütün olarak ele alıp aranan kelime ile ortalama yakınlık sağlayacak algoritma geliştirilecek.





1. T. Mikolov, K. Chen, G. Corrado, J. Dean, "Efficient Estimation of Word Representations in Vector Space", arxiv.org., September 2013.
2. A McCallum, K Nigam, J Rennine, K Seymore - IJCAI, 1999 – Citese

