

# RAPOR: SocialScanAI

**Kapsam:** Çok modlu ürün arama (görsel + metin), Pricelens ile mağazalar arası kıyas, B2C/B2B bildirim katmanı.

**Durum:** 10 günde geliştirilmiş demo; kontrollü veri seti üzerinde değerlendirme.

## İçindekiler

- Özet
- Problem Tanımı ve Hedef
- Demo Kapsamı ve Veri
- Yöntem
  - Çok Modlu Temsil
  - Sıralama Birleştirme
  - Pricelens Entegrasyonu
  - Türkçe Yorum Duygu Analizi
- Sistem Bileşenleri ve Akış
  - Pipeline
  - Arama Katmanı
  - Bildirim Katmanı
- Değerlendirme
- Kısıtlar
- Gelecek Çalışmalar
- Sonuç

---

## 1. Özet

**SocialScanAI**, çok modlu (görsel + metin) arama, mağazalar arası teklif kıyaslama ve bildirim mekanizmalarını tek akışta birleştiren bir prototiptir. Sistem; (i) görsel/metin tabanlı sorgularla ürün eşleştirme, (ii) **Pricelens** ile üç mağaza arasında en uygun alternatifi seçme, (iii) **B2C** (stok/fiyat) ve **B2B** (satıcı içgörülere) bildirimleri üretme işlevlerini sağlar. Çalışma, 10 gün içinde, kontrollü bir demo verisi üzerinde hayata geçirilmiştir.

## 2. Problem Tanımı ve Hedef

- Kullanıcı:** Ürün adını bilmeden betimleyici metin veya görselle arama yapmak; mağazalar arası fiyat/algı dengesine göre tek ve gerekçeli öneri almak; stok/fiyat değişimlerinde gürültüsüz bildirim almak.
- Satıcı:** Talep sinyallerinden (toplu takip davranışları) eyleme dönük öneriler (stok tamamlama, kampanya) elde etmek.

**Hedef:** Kayıp satış riskini azaltan, karar süresini kısaltan, açıklanabilir ve maliyet-duyarlı bir arama–kıyas–bildirim hattı sunmak.

### 3. Demo Kapsamı ve Veri

- **Mağazalar:** ecommerce1, ecommerce2, ecommerce3 (aynı ürün kümesi; farklı fiyat/stok/rating/yorum).
- **Kategoriler:** ayakkabi, sapka (tek beden), sweat, tshirt (S, M, L, XL).
  - **Ayakkabı beden aralığı:** 37–42 (dahil).
- **Ölçek:** ~65 ürün; görseller manuel toplanmıştır.
- **Dosya düzeni:**
  - dukkans/<shop>/backend/data/product/<kategori>.json
  - dukkans/<shop>/backend/data/comments/<kategori>\_comments.json
- **Ürün alanları (çekirdek):** id, name, brand, model, category, price, images[], colors[], description, tags[], stock[]
  - **stock örneği:** [{ "size": "39", "isAvailable": true }, ...]
- **Türetilmiş alanlar:** clip\_vector, text\_vector\_clip, text\_vector\_st, combined\_vector, pricelens\_score, rating\_count, rating\_stats{...}, pricelens\_metadata{...}.

### 4. Yöntem

#### 4.1. Çok Modlu Temsil

Ürünler hem görselden hem metinden temsil edilir ve normalize edilir. Görsel (CLIP image) ve metin (CLIP text) temsilleri, boyut uyumu kontrol edilerek **ağırlıklı birleştirme** ile tek bir **combined** temsile dönüştürülür. Bu strateji, görsel benzerlik ile metinsel bağlamı dengeli kullanarak hatalı eşleşmeleri azaltmayı hedefler. (ST temsili, uzay uyumsuzluğu nedeniyle combine’a dahil edilmez; aramada ayrı bir sinyal olarak kullanılır.)

#### 4.2. Sıralama Birleştirme

Üç bağımsız yöntemden (ST, CLIP-text, Combined) elde edilen sıralamalar **Reciprocal Rank Fusion** ile tek bir nihai listeye dönüştürülür. Yöntemlerin ortak biçimde üst sıralara koyduğu adaylar nihai sonuçta öne çıkar; böylece tek bir yöntemin önyargısı dengelenir.

#### 4.3. Pricelens Entegrasyonu

Nihai arama sonucunda belirlenen hedef ürün için üç mağazadaki varyantlar toplanır. Ürün JSON’larında tutulan özet metrikler (örn. ortalama derecelendirme, yorum yoğunluğu sinyali, değer skoru, güven göstergeleri) **tek bir karar skoru** altında birleştirilir ve kullanıcıya “**en uygun teklif**” olarak sunulur. Skor bileşenleri şeffaftır; formül detayları bu rapor kapsamı dışındadır.

#### 4.4. Türkçe Yorum Duygu Analizi

Klasik “pozitif/negatif” ayrımının ötesinde; yorumlar için **yön (polarity)**, **şiddet (intensity)** ve **yoğunluk (density)** sinyalleri üretilir.

- **Normalize:** Küçük harf, diakritik sadeleştirme, noktalama/boşluk temizliği.
- **Cache:** Normalize metnin özetiyle önbellek; tekrar eden içerikler yeniden modele gönderilmez.
- **Toplu işleme:** 5’li gruplar; limitlerde geri-bekleme (backoff).
- **Boş/kısa yorumlar:** Nötr kabul edilir, alanlar boş/varsayılan değerlerle işaretlenir.

Bu sinyaller Pricelens tarafında gürültüyü azaltır; “gerçekten sevilen” ürünlerin seçimine katkı sağlar.

#### 5. Sistem Bileşenleri ve Akış

##### 5.1. Pipeline

- **Değişiklik tespiti:** Ürün ve yorum dosyaları hash’lenerek izlenir.
- **Ürün hattı:** Açıklama/etiket üretimi → görsel/metin embedding → birleşik temsil.
- **Yorum hattı:** Normalize/cache → duygu sinyalleri → özet metrikler ve derecelendirme güncellemeleri.
- **Çıktı:** Zenginleştirilmiş alanlar ürün/yorum JSON’larına geri yazılır.

##### 5.2. Bildirim Katmanı

- **B2C:** /api/track ile takip (tip: **stok/fiyat**). Stok girişi veya fiyat düşüşünde **WhatsApp** mesajı gönderilir; kayıt pasifleştirilir.
- **B2B:** Aynı shop+ürün+takip\_tipi için kullanıcı takibi eşik değeri aşınca satıcıya eylem önerisi iletilir (demo’da eşik 2).

#### 6. Değerlendirme

- **Arama kalitesi:** Görsel ve metin sinyallerinin birlikte kullanılması, tek yönleme göre daha kararlı ve düşük önyargılı sonuç üretir.
- **Algı sinyali:** Yorumlarda şiddet/yoğunluk sinyalleri, yalnızca “pozitif/negatif” etiketine göre daha ince taneli ürün algısı sunar.
- **Maliyet & hız:** Önbellekleme + toplu istek stratejileri, LLM çağrı maliyetini düşürürken pratik yanıt süreleri sağlar.
- **Uygulanabilirlik:** Gerçek e-ticaret API’leri olmaksızın kontrollü veriyle uçtan uca akışın kurulması, yaklaşımın **yapılabilirliğini** gösterir.

## 7. Kısıtlar

- Demo verisi sınırlıdır; gerçek dünyadaki çeşitlilik ve gürültü tamamen temsil edilmez.
- Türkçe yorum analizi yönteminin üretim-seviyesi doğruluğu garanti değildir; etiketli veriyle yerel model eğitimi önerilir.
- Ağırlıklar ve birleştirme parametreleri sabittir; kategori/koşullara adaptif değildir.
- Gerçek zamanlı tetikleme (watchdog/event) yoktur; pipeline manuel çalıştırılır.

## 8. Gelecek Çalışmalar

1. **Kategoriye duyarlı birleşim:** Görsel/Metin ağırlıklarını kategoriye göre dinamik belirleme.
2. **Adaptif birleştirme:** Veri kalitesi ve liste uzunluğuna göre parametrelerin otomatik ayarlanması.
3. **Bağlamsal önbellek:** Aynı cümlenin farklı ürün bağlamlarını ayırt edecek hash stratejisi.
4. **Watchdog & olay tabanlı tetikleme:** Değişikliklerde otomatik ve izlenebilir akış.
5. **Yerel TR modeli:** Etiketli veri ile kalıcı duygu modeli (maliyet ↓, tutarlılık ↑).
6. **Görsel arama UX'i:** Görselden otomatik kategori/öğe tespiti ve hızlı filtre butonları.
7. **Pricelens şeffaflığı:** Skor bileşenlerinin kullanıcıya açıklanabilir alt kırılımlarla sunulması.

## 9. Sonuç

SocialScanAI, çok modlu temsil, sıralama birleştirme ve bildirim üretimini bir araya getirerek kullanıcıya tek bakışta “**en uygun**” alternatifini sunmayı amaçlayan modüler bir çözümdür. Kısa geliştirme süresine rağmen arama–kıyas–bildirim hattı uçtan uca çalışır hâle getirilmiştir. Üretim yol haritası; kategoriye duyarlı parametreler, adaptif birleştirme ve yerel Türkçe modellerle derinleşecektir.

-  
-