**TEKNOLOJİ KAPTANLARI YARIŞMASI PROJE KAPSAM**

EVE MAĞAZACILIK (EVESHOP)

ANALİTİK CRM PROJESİ

Proje Adı: **Analitik CRM**

Proje Başlangıç Tarihi: 01.04.2021

Proje Bitiş Tarihi: 31.12.2022

Projede Görev Alan Personel Sayısı: **5**

Proje Süresi Adam/ Ay: **20**

|  |
| --- |
| **PROJE ÖZETİ**Analitik CRM Projesi, online ve offline kanallardan toplanan müşteri verilerinin KVKK'ya uygun bir şekilde yönetimini ve tekilleştirilmesini sağlayarak, müşteri ilişkileri yönetimini daha stratejik bir seviyeye taşımayı hedeflemektedir. Proje kapsamında, sms ve ETK izinleri gibi müşteri verileri tek bir platformda birleştirilerek veri tutarlılığı sağlanmıştır. Aynı müşteriye ait farklı kaynaklardan gelen veriler tek bir profil altında toplanarak, daha bütünsel bir müşteri bakış açısı elde edilmiştir. Bu süreç, verilerin doğruluğunu artırırken, veri tabanındaki karmaşayı da önlemiştir.Proje, aynı zamanda RFM (Sıklık, Miktar, Son Satın Alma) analizi ile müşteri segmentasyonu yapmayı sağlamaktadır. Müşteriler, satın alma sıklıkları, toplam harcama miktarları ve son satın alma tarihlerine göre sınıflandırılarak, işletmelere müşteri davranışlarını ve ihtiyaçlarını daha derinlemesine anlama fırsatı sunulmuştur. Bu segmentasyon sayesinde, her bir müşteri grubuna özel kişiselleştirilmiş kampanyalar tasarlanarak doğru hedef kitleye, doğru zamanda ve uygun teklifler ile ulaşılması sağlanmıştır. Kişiselleştirilmiş kampanyalar, müşteri memnuniyetini artırırken aynı zamanda satışların ve gelirin de yükseltilmesine katkı sağlamaktadır. Projenin önemli bir diğer boyutu ise KVKK gibi yasal düzenlemelere uyum sağlanarak müşteri izin yönetiminin gerçekleştirilmesidir. Bu adım, işletmelerin yasal yükümlülüklerini yerine getirirken, müşteri güvenini artırma ve marka değerini koruma gibi avantajlar sunmaktadır. Böylece veri güvenliği ve gizliliği konusunda projenin, işletmelere rekabet avantajı kazandıracağı öngörülmektedir.Sonuç olarak, "Analitik CRM" projesi, müşteri verilerinin yasal düzenlemelere uygun olarak toplanması, tekilleştirilmesi, segmente edilmesi ve kişiselleştirilmiş kampanyaların kurgulanması süreçlerini kapsamaktadır. Bu kapsamlı yaklaşım, işletmelerin müşteri ilişkilerini daha bilinçli ve stratejik bir şekilde yönetmelerine olanak tanırken, müşteri bağlılığını ve memnuniyetini de artırmaktadır.**PROJE HAKKINDA**Analitik CRM projesi hem online hem de offline kanallar aracılığıyla müşteri üyeliği alınan verilerin, KVKK (Kişisel Verilerin Korunması Kanunu) izinlerine uygun şekilde analiz edilmesini ve müşteri davranışlarına dayalı bir öneri sistemi geliştirilmesini hedeflemektedir. Müşteri numaralarıyla veri tabanına kaydedilen müşterilerin satın aldıkları ürünler üzerinde yapılacak trend analizleri, müşterilere kişiselleştirilmiş ürün önerileri sunmak için kullanılacaktır.E-ticaret platformlarının hızla büyümesi, daha etkili öneri sistemlerine olan ihtiyacı artırmıştır. Geleneksel işbirlikçi filtreleme ve içerik tabanlı yaklaşımlar, özellikle marka sadakati ve satın alma sıklığının önemli olduğu perakende sektöründe, kullanıcılar ve ürünler arasındaki karmaşık ilişkileri yakalamakta yetersiz kalmaktadır.Proje içerisinde kullanılan metotlar; Doğal Dil İşleme (NLP) alanındaki son gelişmeler, özellikle all-MiniLM-L6-v2 gömme (embedding) yöntemleri, öneri sistemlerini daha nüanslı ve dinamik hale getirme konusunda yeni fırsatlar sunmuştur. Bu çalışma, all-MiniLM-L6-v2 gömme yöntemleri ve vektör veritabanlarını entegre ederek, perakende platformları için gerçek zamanlı ve kişiselleştirilmiş "Sonraki En İyi Ürün" (Next-Best-Product, NBP) önerileri sunmayı amaçlamıştır.Bu raporlama sistemi ile, perakende sektöründe satış tahminini ve müşteri davranışlarına dayalı öneri sistemlerini geliştirmek amacıyla yapılan çalışmaların teknik fizibilitesini incelemektedir. Temel hedefler şunlardır:1. **Veri Ön İşleme:** 2022-2023 yıllarına ait 36 milyon satış kaydının temizlenmesi ve analiz için hazır hale getirilmesi. Bu süreçte işlem sayısı, toplam gelir ve son satın alma tarihinden bu yana geçen gün sayısı gibi temel metrikler hesaplanmıştır.
2. **Model Geliştirme ve Eğitimi:** Temizlenmiş veriler kullanılarak, müşteri tercihlerini tahmin eden ve ürün önerilerinde bulunan makine öğrenmesi modelleri geliştirilmiştir. Bellek yönetimi optimize edilerek veriler daha küçük parçalara ayrılarak işlenmiştir.
3. **Eksik Veri İçin Mantıksal Çıkarımlar:** Eksik müşteri-ürün çiftleri için mantıksal çıkarımlar yapılmış ve bu çıkarımlar, öneri sistemlerinin kritik bir bileşeni olarak kaydedilmiştir.
4. **Öneri Sistemlerinin Otomasyonu:** Kişiselleştirilmiş ürün önerilerinin otomatik olarak oluşturulması ve müşterilere sunulması süreçleri otomatikleştirilmiştir. Bu sayede müşteri deneyimi geliştirilmiş ve satış dönüşüm oranları artırılmıştır.

**YÖNTEM*** 1. **Proje Fikrinin Ortaya Çıkışı**

Bu çalışmanın temel amacı, perakende sektöründe tüketim malları (kozmetik, şampuan vb.) için bir öneri motoru geliştirmektir. Bu motor, kullanıcı ve ürün verilerinin vektörleştirilmesi için all-MiniLM-L6-v2 yöntemini, benzerlik aramaları için ise bir vektör veritabanını kullanarak, platformun kullanıcı deneyimini iyileştirmeyi hedeflemektedir. **1.2. Hedefler** * Kullanıcı ve ürün listelemeleri için, All-MiniLM-L6-v2 veya OpenAI Ada-002 kullanarak vektör gömmeleri oluşturmak.
* Vektör gömmelerini depolamak ve sorgulamak için uygun bir vektör veritabanı tasarlamak ve entegre etmek.
* Vektör veritabanını kullanarak, gerçek zamanlı NBP (Next-Best-Product) önerileri sağlayacak bir backend hizmeti geliştirmek.
* Öneri motorunun doğruluk, yanıt süresi ve kullanıcı memnuniyeti açısından performansını değerlendirmek.

**1.3. Problem Tanımı**Perakende platformları, kullanıcı tercihlerinin çeşitli ve dinamik doğası ile ürün listelerinin karmaşıklığı nedeniyle, kullanıcılara uygun ürünler önermekte zorluk çekmektedir. Mevcut öneri sistemleri, kullanıcı davranışları ve ürün özellikleri arasındaki ince nüansları yakalamakta yetersiz kalmakta; özellikle de marka sadakati ve kullanım sıklığının önemli olduğu tüketim malları bağlamında bu zorluklar daha da belirgin hale gelmektedir. (Bu çalışma, kullanıcıları ve ürünleri yüksek boyutlu vektörlere kodlamak için all-MiniLM-L6-v2'yi kullanarak ve vektör veritabanını verimli benzerlik aramaları için devreye sokarak bu zorlukların üstesinden gelmeyi amaçlamaktadır1. **LİTERATÜR ANALİZİ**

Proje kapsamında yapılan çalışmalar, öneri sistemleri, doğal dil işleme (NLP), vektör tabanlı veritabanları ve ilgili teknolojiler üzerine yapılmış araştırmalarla temellendirilmiştir. Aşağıda, bu alanlara yönelik kapsamlı bir akademik literatür özeti sunulmaktadır.**2.1. Öneri Sistemleri ve Uygulamaları** Proje çerçevesinde kullanılan Öneri sistemleri, müşterilerin ilgisini çekebilecek ürün, hizmet veya içerikleri öneren algoritmalar ve tekniklerdir. Bu sistem, e-ticaret ce mağazalarda yaygın bir şekilde kullanılmaktadır. Genellikle öneri sistemleri; işbirlikçi filtreleme, içerik tabanlı filtreleme ve hibrit modeller gibi çeşitli yaklaşımlar kullanır. İşbirlikçi filtreleme, kullanıcıların geçmiş davranışlarını analiz ederek benzer kullanıcıların tercihlerini öngörürken, içerik tabanlı filtreleme, kullanıcının geçmişte ilgi gösterdiği öğelerin özelliklerine dayalı öneriler sunar.Öneri sistemlerinin kullanımı, müşteri bağlılığını artırmak ve satışları optimize etmek için kritik bir rol oynamaktadır. E-ticaret platformları, bu sistemleri kullanarak müşteri deneyimini kişiselleştirir ve böylece satışları artırır (Linden vd., 2003). Bu tür sistemler, kullanıcı davranışları ile ürün özellikleri arasındaki karmaşık ilişkileri modellemek için gelişmiş veri analitiği tekniklerinden faydalanır.**2.2. Doğal Dil İşleme (NLP) ve all-MiniLM-L6-v2** Doğal dil işleme (NLP) alanındaki gelişmeler, özellikle all-MiniLM-L6-v2 gibi teknikler, öneri sistemlerinin doğruluğunu artırmak amacıyla kullanılmaktadır. all-MiniLM-L6-v2, kelimeleri vektörlere dönüştürerek kelimeler arasındaki anlamsal ilişkileri modelleyip, böylece ürünler ve kullanıcılar için anlamlı temsiller oluşturur. Bu teknik, kullanıcının ürünlerle olan etkileşimlerini analiz ederek, kişiselleştirilmiş öneriler sunmak için bu vektörleri kullanır.Bu bağlamda, all-MiniLM-L6-v2 gibi vektör tabanlı teknikler, geleneksel öneri sistemlerinin ötesine geçerek kullanıcı davranışlarını ve ürün ilişkilerini daha derinlemesine analiz eder. Özellikle perakende sektöründe, kullanıcıların belirli markalara olan bağlılıkları ve alışveriş sıklıkları gibi dinamikleri modellemek için bu tür tekniklerin kullanılması önemlidir.**2.3. Vektör Tabanlı Veri Tabanları ve Benzerlik Aramaları** Vektör tabanlı veri tabanları, büyük vektör kümeleri üzerinde benzerlik aramaları gerçekleştirmek için optimize etmektedir. Bu tür veri tabanları, yüksek boyutlu veri setlerinde hızlı ve verimli aramalar yapabilme yetenekleriyle bilinir. Pinecone gibi modern vektör veri tabanları, yüksek performanslı ve ölçeklenebilir benzerlik arama hizmetleri sunar, bu da gerçek zamanlı öneri sistemlerinde kritik bir bileşen haline gelir. Vektör tabanlı veri tabanlarının entegrasyonu, öneri sistemlerinin kullanıcı etkileşimlerini daha hızlı ve verimli bir şekilde işlemesine olanak tanır. Bu, kullanıcı deneyimini optimize etmek ve öneri motorunun performansını artırmak için önemli bir stratejidir. **2.4. Performans Değerlendirmesi ve Kullanıcı Memnuniyeti** Öneri sistemlerinin başarısı, genellikle doğruluk, hassasiyet, geri çağırma ve kullanıcı memnuniyeti gibi performans ölçütleriyle değerlendirilir. Doğruluk, önerilen ürünlerin kullanıcılar için ne kadar uygun olduğunu gösterirken, geri çağırma, kullanıcının ilgisini çeken tüm ürünlerin ne kadarının önerildiğini ölçer (He vd., 2017). Kullanıcı memnuniyeti ise, önerilerin kullanıcı algısı üzerindeki etkisini ve kullanıcıların bu önerilerden ne kadar memnun kaldığını değerlendiren kritik bir göstergedir.Bu çalışma, öneri motorunun performansını bu ölçütler doğrultusunda değerlendirerek, motorun kullanıcılar üzerindeki etkisini ve potansiyel iyileştirme alanlarını belirlemeyi amaçlamaktadır.**2.5. Hibrit Öneri Sistemleri** Gelecekteki öneri sistemleri, hibrit yaklaşımlarla daha kişiselleştirilmiş ve bağlamsal farkındalığa sahip öneriler sunmayı hedeflemektedir. Hibrit modeller, işbirlikçi filtreleme ve içerik tabanlı filtreleme gibi farklı yöntemleri birleştirerek daha zengin ve çeşitli öneriler sunar (Burke, 2002). Bu tür sistemlerin, özellikle e-ticaret gibi dinamik ortamlarda daha etkili olduğu kanıtlanmıştır. Hibrit modellerin geliştirilmesi ve optimizasyonu, kullanıcıların beklentilerini karşılamak ve perakende platformlarının rekabet gücünü artırmak için önemli bir adım olacaktır. 1. **İHTİYAÇ ANALİZİ**

Offline ve online müşterilerin tek bir platformdan yönetilmesi, kişiselleştirilmiş kampanyaların ve puan sisteminin yönetilmesi ve KVKK'ya uygun şekilde veri tabanında tutulması gerekmektedir.Dijital dönüşümün hız kazandığı çağımızda, işletmelerin müşteri verileri üzerindeki hakimiyeti, onların pazarlama stratejilerinin başarısında kritik bir rol oynamaktadır. Ancak bu verilerin yönetilmesi ve optimize edilmesi, karşılaşılan zorluklar ve mevzuat gereklilikleri nedeniyle kompleks bir süreç haline gelmiştir. Tek Platformda Müşteri Yönetimi: İşletmeler, müşterileriyle farklı kanallar üzerinden etkileşimde bulunmaktadırlar. Bu kanalların başında fiziksel (offline) mağazalar, online mağazalar, sosyal medya platformları ve diğer dijital kanallar gelmektedir. Her kanal, kendi içerisinde farklı müşteri veri setlerini üretmektedir. Bu verilerin tek bir platformda toplanması, işletmelere müşterileri hakkında bütünsel bir görüş sunar. Bu, pazarlama stratejilerinin daha etkili bir şekilde kurgulanmasına olanak tanır. Kişiselleştirilmiş Kampanyaların Yönetimi: Müşterilerin alışkanlıkları, tercihleri ve ihtiyaçları farklılık göstermektedir. Bu nedenle, genel ve tek tip kampanyaların yerine, müşteri segmentlerine özel olarak tasarlanmış kişiselleştirilmiş kampanyaların yürütülmesi büyük önem taşımaktadır. Bu sayede, müşteri memnuniyeti artar, marka sadakati güçlenir ve dönüşüm oranları yükselir. Puan Sisteminin Yönetimi: Müşteri sadakati programları, tüketicilere markaya olan bağlılıklarını sürdürmeleri için teşvikler sunar. Ancak bu programların yönetimi, özellikle çoklu kanal kullanımı söz konusu olduğunda karmaşıklık arz edebilir. Müşterilerin kazandığı puanların, kullandığı avantajların ve özel tekliflerin tek bir platformda yönetilmesi hem işletmelere operasyonel kolaylık sağlar hem de müşterilere kesintisiz bir deneyim sunar.  KVKK Uyumlu Veri Yönetimi: Türkiye'de Kişisel Verilerin Korunması Kanunu (KVKK) gereği, müşteri verilerinin korunması ve uygun şekilde yönetilmesi zorunludur. İşletmeler, bu kanuna uygun olarak müşteri verilerini toplamalı, saklamalı ve gerektiğinde silmelidir. Ayrıca, bu verilerin güvende tutulması ve yetkisiz erişimlere karşı korunması da büyük önem taşımaktadır.  Sonuç olarak, Analitik CRM projesi, işletmenin yukarıda belirtilen temel ihtiyaçlarına yanıt vermek üzere tasarlanmıştır. Bu ihtiyaçlar, müşteri veri yönetiminin karmaşıklığını ve işletmeler için taşıdığı stratejik önemi net bir şekilde ortaya koymaktadır.**Proje Kapsamında Yapılan Faaliyetler** 1. Müşteri Verilerinin Toplanması: Proje, offline ve online kanallar aracılığıyla müşteri verilerinin toplanmasıyla başlamıştır. Bu süreçte, müşterilerden alınan SMS ve ETK izinleri KVKK'ya uygun olarak kaydedilmiştir. 2. Veri Tekilleştirme: Toplanan veriler tek bir platformda toplanarak tekilleştirilmiştir. Bu sayede, aynı müşteriye ait farklı kanallardan gelen veriler birleştirilerek müşteri hakkında daha bütünsel bir görüş elde edilmiştir. 3. RFM Analizi: Müşteri segmentasyonu için RFM (Son Satın Alma, Sıklık, Maliyet) analizi gerçekleştirilmiştir. Bu analiz ile müşteriler, alışveriş sıklıklarına, harcadıkları toplam tutara ve son alışveriş tarihlerine göre segmentlere ayrılmıştır. 4. Kampanya Yönetimi: Müşteri segmentasyonu sonrası, her segmente uygun kampanya senaryoları kurgulanmıştır. Bu kampanyaların etkinliği ve dönüşüm oranları sürekli olarak izlenmiştir. 5. KVKK Uyumu: Tüm bu faaliyetler sırasında müşteri verilerinin korunması ve yönetilmesi KVKK'ya uygun olarak gerçekleştirilmiştir. Bu kapsamda gerekli güvenlik önlemleri alınmış, yetkisiz erişimlere karşı koruma sağlanmıştır. 6. Yazılım Geliştirme: Proje geliştirme süreci Eve Mağazacılık A.Ş. tarafından yürütülmektedir. 7. Sürekli İzleme ve Değerlendirme: Proje süresince, elde edilen sonuçlar ve kurgulanan kampanyalar sürekli olarak izlenmiş ve değerlendirilmiştir. Böylece projenin başarısı ölçülmüş ve gerektiğinde rotalar güncellenmiştir. * **Yenilikçi Yaklaşım: Tek Platformda Müşteri Verisi Yönetimi**

Çoğu CRM sistemi ya offline ya da online müşteri verileri üzerinde yoğunlaşırken, bu proje hem offline hem de online kanallardan elde edilen müşteri verilerini tek bir platformda toplama ve yönetme yeteneğine sahiptir. Bu bütünsel yaklaşım, işletmelere müşteri davranışları hakkında daha derinlemesine ve kapsamlı bir görüş sunar. * **Ar-Ge Niteliği: RFM Analizi ve Segmentasyon**

Proje, müşteri verilerini sadece toplamakla kalmaz, aynı zamanda bu veriler üzerinde ileri düzey analizler gerçekleştirir. RFM analizi, Ar-Ge süreçlerinin bir çıktısı olarak, müşterilerin satın alma davranışlarını değerlendirerek segmentlere ayırmaktadır. Bu segmentasyon, müşterilere daha özelleştirilmiş hizmetler ve kampanyalar sunma potansiyeli taşır. * **KVKK Uyumuyla Güncel Mevzuata Uyum**

Müşteri verilerinin KVKK'ya uygun olarak toplanması, saklanması ve işlenmesi, proje için kritik bir öneme sahiptir. Bu yaklaşım, projenin sadece teknik ve stratejik yönleriyle değil, aynı zamanda güncel yasal gerekliliklerle de uyumlu olduğunu göstermektedir. **4. METODOLOJİ** Çalışmanın metodolojisi aşağıdaki adımları içermektedir: 4.1. Veri Toplama: ● Kullanıcı Etkileşim Verileri: Kullanıcıların perakende platformunda yaptığı tıklamalar, görüntülemeler, satın almalar ve incelemeler gibi etkileşim verileri toplanmaktadır. ● Ürün Verileri: Platformda bulunan ürünlerin ayrıntılı açıklamaları ve özellikleri toplanmaktadır.4.2. Veri Ön İşleme: ●Metin İşleme: Ürün açıklamaları temizlenip ön işlemden geçirilmektedir. ●Etkileşim Kayıtları: Kullanıcı etkileşim verileri normalleştirilip, kullanıcı profilleri oluşturulacaktır. 4.3. all-MiniLM-L6-v2 Gömme: ●Model Eğitimi: İşlenmiş ürün açıklamaları üzerinde bir all-MiniLM-L6-v2 modeli eğitilecektir. ●Kullanıcı Gömme: Kullanıcının etkileşimde bulunduğu ürünlerin gömmeleri birleştirilerek kullanıcı gömmeleri oluşturulacaktır. 4.4. Vektör Veritabanı: ●  Veritabanı Seçimi: Vektör gömmelerini depolamak ve sorgulamak için uygun bir vektör veritabanı (Pinecone) seçilmiştir.●  Entegrasyon: Vektör veritabanı, backend hizmetiyle entegre edilecektir. 4.5. Tahmin Algoritması: ●  Algoritma Seçimi: Vektör gömmelerine dayanarak NBP önerilerini belirlemek için bir tahmin algoritması (LGBM, LSTM/GRU) seçilecektir. ●  Entegrasyon: Tahmin algoritması, gerçek zamanlı NBP önerileri sağlamak için backend hizmetiyle entegre edilecektir. 4.6. Backend Hizmet Geliştirme: ●  Hizmet Mimarisi: Öneri motoru için mikro hizmet tabanlı bir backend mimarisi tasarlanacaktır. ●  API Geliştirme: Kullanıcı isteklerini işlemek ve gerçek zamanlı NBP önerileri sağlamak için API'ler geliştirilecektir. 4.7. Değerlendirme: ●  Performans Ölçütleri: Öneri motoru, kesinlik, geri çağırma, F1 skoru ve ortalama ters sıra (MRR) gibi ölçütler kullanılarak değerlendirilecektir (He vd., 2017). ●  Kullanıcı Memnuniyeti: Önerilerin algılanan geçerliliği ve kullanıcı memnuniyetini değerlendirmek için kullanıcı anketleri yapılacaktır. 1. **PROJENİN YENİLİKÇİ YÖNÜ**

Günümüzde verinin önemli bir varlık haline gelmesi, işletmeleri müşteri verilerini daha etkin ve stratejik bir şekilde yönetmeye yönlendirmiştir. Analitik CRM projesi, bu ihtiyaca yanıt vermek amacıyla kurgulanmıştır ve birçok yenilikçi yönüyle öne çıkar. Proje kapsamında üzerinde çalışılan başlıklar;  Sinirsel İşbirlikçi Filtreleme (NCF) kullanılarak geliştirilen Ürün Tabanlı İşbirlikçi Filtreleme modeli, kullanıcı ve ürün etkileşimlerini optimize etmeyi amaçlamaktadır. Model, gömme katmanları, yoğun katmanlar ve doğrusal bir çıkış katmanı içerir. Modelin eğitimi, Adam optimizasyonu ve ortalama kare hata (MSE) kayıp fonksiyonu ile gerçekleştirilmiş, erken durdurma yöntemi kullanılarak aşırı uyum önlenmiştir. Model, eğitim seti üzerinde eğitilmiş ve test seti üzerinde doğrulanmıştır. Kullanıcı tabanlı öneri sistemi de Sinirsel İşbirlikçi Filtreleme (NCF) kullanılarak geliştirilmiştir. Model, kullanıcılar ve ürünler için gömme katmanları ile yoğun katmanları entegre etmektedir. Modelin eğitimi 20 epoch boyunca gerçekleştirilmiş, sonuçlar doğrultusunda modelin performans metrikleri değerlendirilmiştir. Tahmin fonksiyonu, kullanıcı gömümleri ile ürün gömümleri arasındaki nokta çarpımını kullanarak her bir müşteri için öğelerin sıralanmasını ve önerilmesini sağlamaktadır. FP-Growth algoritması kullanılarak, müşteri sepetlerine eklenebilecek ek ürünlerin belirlenmesi amacıyla sıkça birlikte satın alınan ürün kombinasyonları tespit edilmiştir. Bu süreçte, veriler temizlenmiş, işlemler müşteri kimliği ve satın alma tarihine göre gruplandırılmış ve FP-Growth algoritması ile sık ürün kümeleri oluşturulmuştur. Birliktelik kuralları oluşturularak, hangi ürünlerin birlikte satın alınma olasılığının yüksek olduğu belirlenmiştir. Bu çalışma, öneri sistemlerinin geliştirilmesi sürecinde kullanılan yöntemleri detaylı bir şekilde açıklamakta ve elde edilen sonuçları akademik bir perspektiften değerlendirmektedir. Bu çalışmanın amacı, perakende sektöründe faaliyet gösteren bir işletme için müşterilerin alışveriş deneyimini iyileştirmek amacıyla öneri sistemleri geliştirmektir. Öneri sistemleri, müşterilere geçmiş alışveriş alışkanlıkları ve ürün tercihlerine dayalı olarak kişiselleştirilmiş ürün önerileri sunarak hem müşteri memnuniyetini artırmayı hem de satışları artırarak sürdürülebilir büyüme sağlamayı hedeflemektedir. Bu çalışmada, makine öğrenimi ve veri madenciliği teknikleri kullanılarak etkin ve verimli bir öneri sistemi tasarlanması amaçlanmaktadır. Tek Platformda Müşteri Verisi Yönetimi Çoğu CRM sistemi ya offline ya da online müşteri verileri üzerinde yoğunlaşırken, bu proje hem offline hem de online kanallardan elde edilen müşteri verilerini tek bir platformda toplama ve yönetme yeteneğine sahiptir. Bu bütünsel yaklaşım, işletmelere müşteri davranışları hakkında daha derinlemesine ve kapsamlı bir görüş sunar.  Ar-Ge Niteliği: RFM Analizi ve Segmentasyon Proje, müşteri verilerini sadece toplamakla kalmaz, aynı zamanda bu veriler üzerinde ileri düzey analizler gerçekleştirir. RFM analizi, Ar-Ge' nin bir ürünü olarak, müşterilerin satın alma davranışlarını değerlendirerek segmentlere ayırmaktadır. Bu segmentasyon, müşterilere daha özelleştirilmiş hizmetler ve kampanyalar sunma potansiyeli taşır.  KVKK Uyumuyla Güncel Mevzuata Uyum Müşteri verilerinin KVKK'ya uygun olarak toplanması, saklanması ve işlenmesi, proje için kritik bir öneme sahiptir. Bu yaklaşım, projenin sadece teknik ve stratejik yönleriyle değil, aynı zamanda güncel yasal gerekliliklerle de uyumlu olduğunu göstermektedir.  **VERİ SETİNİN HAZILANMASI** Veri Ön İşleme Faaliyetleri 2022-2023 yıllarına ait 36 milyon satış kaydını içeren büyük bir veri setini işleyerek toplu satış metrikleri oluşturduk. Bu süreç, verilerin temizlenmesi, işlem sayısı, toplam gelir ve son satın alma tarihinden bu yana geçen gün sayısı gibi temel metriklerin hesaplanmasını ve sonuçların bir CSV dosyasına kaydedilmesini içeriyordu. Veriler, belleği verimli bir şekilde yönetmek amacıyla parçalar halinde işlendi, böylece nihai çıktı, müşteri davranışları ve ürün performansı hakkında değerli içgörüler sağlamaktadır. Veri İşleme İş Akışı ●  Veri Yükleme: * + Veriler, CSV dosyasından parçalar halinde yüklendi.
	+ Veriler temizlendi ve biçimlendirildi (örneğin, tarihler dönüştürüldü, eksik değerler

ele alındı). ●  Metriğin Hesaplanması: ○ Müşteri-ürün çiftleri başına metrikler hesaplandı: * İşlem sayısı
* Toplam gelir
* Son satın alma tarihinden bu yana geçen gün sayısı
* Satın alma hacmi

●  Veri Toplama: * + Parçalar arasında metrikler toplandı.
	+ Müşteri-ürün çiftleri başına ortalama fiyat hesaplandı.

●  Çıktı: ○ Toplanan veriler, kademeli olarak bir CSV dosyasına kaydedildi. Satış verileri ön işleme süreci tamamlandıktan sonra, veri setinde eksik olan müşteri-ürün çiftleri için derecelendirmeler mantıksal olarak çıkarıldı. Bu derecelendirmeler, öneri sistemleri için kritik öneme sahiptir. Derecelendirme Çıkarımı İş Akışı ● Toplanan Verilerin Yüklenmesi:○ Toplu satış metriklerini içeren önceden işlenmiş veri seti yüklendi. ● Ağırlık Atama:○ Her bir metriğe belirli ağırlıklar uygulanarak birleşik bir derecelendirme hesaplandı: * İşlem sayısı: 10
* Toplam gelir: 3
* Ortalama fiyat: 3
* Son satın alma tarihinden bu yana geçen gün sayısı: 3
* Satın alma hacmi: 2

● Birleşik Derecelendirmenin Hesaplanması:○ Metriğin ağırlıklı toplamı kullanılarak birleşik derecelendirme hesaplandı. ● Logaritmik Dönüşüm:○ Normallik sağlamak için birleşik derecelendirmeye logaritmik dönüşüm uygulandı. * Derecelendirme Binning:
	+ Hassas kategorilendirme sağlamak için ince sınıflar oluşturuldu.
	+ Her bir müşteri-ürün çifti, 1 ile 5 arasında bir ölçek üzerinde nihai bir derecelendirme aldı.
* Eksik Derecelendirmelerin Ele Alınması:

○ Eksik derecelendirmeler kontrol edildi ve ortalama derecelendirme ile dolduruldu. Nihai derecelendirmeler, güncellenmiş veri setine kaydedildi ve bu sayede daha doğru bir öneri sistemi sağlanmış oldu. page6image13522624Şekil 1: Veri setinin genel görünümü Derin Öğrenme Ağı Kullanarak Ürün Tabanlı İşbirlikçi Filtreleme (IBCF) Özet Kullanıcı ve ürün etkileşimlerine dayalı önerileri optimize etmek amacıyla, derin öğrenme kullanarak bir Ürün Tabanlı İşbirlikçi Filtreleme (IBCF) sistemi oluşturulmuştur. Bu amaçla, bir sinir ağı modeli geliştirilmiştir. Derin Öğrenmeye Dayalı IBCF İş Akışı ●  Veri Hazırlığı: ○  Veri Yükleme: Ön işlenmiş veri seti Google Drive'dan yüklendi. ○  Kodlama: Müşteri kimlikleri (CUST\_ID) ve ürün kodları (PRODUCT\_CODE) sayısal indekslere dönüştürüldü. ○  Eğitim-Test Ayrımı: Veriler, eğitim ve test setlerine ayrıldı. ●  Model Mimarisi: ○  Gömme Katmanları: Kullanıcılar ve ürünler için gömme katmanları oluşturuldu. ○  Yoğun Katmanlar: Daha iyi genelleme sağlamak için yoğun katmanlar, batch normalization ve dropout katmanlarıyla istiflendi. ○ Çıkış Katmanı: Derecelendirmeleri tahmin etmek için doğrusal bir katman kullanıldı. * Model Eğitimi:
	+ Derleme: Model, Adam optimizasyonu ve ortalama kare hata (mean squared error) kaybı ile derlendi.
	+ Erken Durdurma: Aşırı uyumu önlemek için erken durdurma uygulandı.
	+ Eğitim: Model, eğitim seti üzerinde eğitildi ve test seti üzerinde doğrulama yapıldı.
* Model Değerlendirmesi:

○ Model, ilk epoch'ta 0.9292'lik bir eğitim kaybı ve 0.4258'lik bir doğrulama kaybı elde etti, bu da etkili bir öğrenmeye işaret etmektedir. Model Testi ve Değerlendirme Derin öğrenme kullanılarak eğitilen Ürün Tabanlı İşbirlikçi Filtreleme (IBCF) modelinin, ayrılmış test seti üzerinde performansı değerlendirildi. Test İş Akışı ● Eğitilmiş Modelin Yüklenmesi:○ Optimize edilmiş öneri modeli, kaydedilen dosyadan yüklendi. ● Tahminlerin Yapılması:○ Eğitilmiş model kullanılarak test seti için derecelendirme tahminleri yapıldı. * Performans Değerlendirmesi:
	+ Ortalama Kare Hata (MSE): Test setinde MSE hesaplandı ve 0.4258 olarak bulundu.
	+ Kök Ortalama Kare Hata (RMSE): Daha iyi yorumlanabilirlik sağlamak amacıyla RMSE de hesaplandı ve 0.6526 değeri elde edildi.
* Örnek Tahminler:

Modelin doğruluğunu göstermek için bir dizi test örneği için tahmin edilen derecelendirmeler ile gerçek derecelendirmeler karşılaştırıldı. Örnek tahminler ve gerçek derecelendirmeler: * Tahmin: 4.92, Gerçek: 5.00
* Tahmin: 4.44, Gerçek: 4.00
* Tahmin: 4.04, Gerçek: 5.00
* Tahmin: 3.90, Gerçek: 5.00
* Tahmin: 1.51, Gerçek: 1.00
* Tahmin: 3.25, Gerçek: 4.00
* Tahmin: 3.81, Gerçek: 3.00
* Tahmin: 3.08, Gerçek: 2.00
* Tahmin: 4.85, Gerçek: 4.00

● Tahmin: 1.47, Gerçek: 2.00 Model Ağırlıkları Kullanarak Ürün Önerilerini Otomatikleştirme Bu aşamada, rastgele ürün kimliklerinin alınması ve eğitilmiş modelin ağırlıklarına dayalı olarak ürün önerilerinin otomatikleştirilmesi süreci gerçekleştirilmiştir. Önerilerin Otomatikleştirilmesi İş Akışı ● Rastgele Ürün Seçimi:○ Rastgele Ürün Kimliklerini Getirme: Öneriler oluşturmak üzere test setinden rastgele 50 ürün kimliği seçildi.  ●  Benzer Ürünlerin Oluşturulması: ○  Ürün Gömümlerini Çıkarma: Eğitilmiş modelden ürün gömümleri alındı. ○  Kosinüs Benzerliği Hesaplama: Seçilen ürünlerle diğer tüm ürünler arasındaki kosinüs benzerliği hesaplanarak en benzer N ürün belirlendi. ○  Öneri Oluşturma: Her seçilen ürün kimliği için en çok önerilen 10 ürün oluşturuldu. ●  Önerilerin Genişletilmesi: ○  Ürün İsimleri ile Birleştirme: Önerilen ürün kimlikleri, dış bir Excel dosyasından alınan ilgili ürün isimleri ile birleştirildi. ○  Önerileri Genişletme: Her önerilen ürünü ismiyle eşleştirerek okunabilir bir format oluşturuldu. ●  Sonuçların Kaydedilmesi ve Görüntülenmesi: ○  CSV'ye Kaydetme: Genişletilmiş öneriler bir CSV dosyasına kaydedildi. ○  Etkileşimli Görüntüleme: Öneriler, etkileşimli bir DataFrame içinde görüntülendi. Kullanıcı Tabanlı Öneri Sisteminin Geliştirilmesi ve Eğitimi Bu bölümde, Sinirsel İşbirlikçi Filtreleme (NCF) kullanarak bir Kullanıcı Tabanlı Öneri Sistemi oluşturulmuş ve eğitilmiştir. Sürecin özeti aşağıdaki gibidir: İş Akışı ● Veri Yükleme ve Kodlama: * Veri Yükleme: Önceden işlenmiş satış metriği veri seti Google Drive'dan yüklendi.
* Kodlama: Müşteri (CUST\_ID) ve ürün (PRODUCT\_CODE) kimlikleri, model uyumluluğu için sayısal indekslere dönüştürüldü.
* Model Mimarisi:
	+ Kullanıcı ve Ürün Gömümleri: Sinir ağı kullanılarak kullanıcılar ve ürünler için gömme katmanları oluşturuldu.
	+ Yoğun Katmanlar: Daha iyi genelleme sağlamak amacıyla batch normalization ve dropout ile entegre yoğun katmanlar eklendi.
	+ Çıkış Katmanı: Bir kullanıcı-ürün çifti için derecelendirmeyi tahmin etmek üzere doğrusal bir çıkış katmanı eklendi.
* Model Derlenmesi:

○ Optimizasyon ve Kayıp Fonksiyonu: Model, Adam optimizasyonu ve ortalama kare hata (MSE) kayıp fonksiyonu kullanılarak, ortalama mutlak hata (MAE) metriği ile derlendi. * Model Eğitimi:
	+ Eğitim-Test Ayrımı: Veriler, %80 eğitim ve %20 test seti olarak ayrıldı.
	+ Eğitim Süreci: Model, 1024 batch boyutu ile 20 epoch boyunca eğitim verileri üzerinde eğitildi.
	+ Doğrulama: Modelin genelleme yeteneğini izlemek için, eğitim sırasında test seti üzerinde performansı doğrulandı.
* Model Kaydetme:

○ Modelin Kalıcı Hale Getirilmesi: Eğitilmiş model, ncf\_model.h5 olarak gelecekteki kullanım için kaydedildi. Modelin 20 epoch boyunca eğitilmesinin sonuçları aşağıdaki gibi özetlenmiştir: Sonuçlar: * Eğitim Kaybı: 0.1919
* Eğitim MAE: 0.2889
* Doğrulama Kaybı: 0.4812
* Doğrulama MAE: 0.4818

Model, hem eğitim kaybında hem de MAE'de azalan bir eğilim gösterdi ve doğrulama metrikleri son epochlara doğru stabil hale geldi. Tahmin Fonksiyonu ve Puanların Belirlenmesi Sinirsel İşbirlikçi Filtreleme (NCF) kullanılarak oluşturulan Kullanıcı Tabanlı Öneri Sistemi'nde, tahmin fonksiyonu, her bir müşteri için öğelerin sıralanması ve önerilmesi sürecinde önemli bir rol oynar. Tahmin fonksiyonunun ve puanların nasıl hesaplandığı aşağıda açıklanmıştır: Tahmin Fonksiyonu İş Akışı ● Gömümleri Çıkarma: * Kullanıcı Gömümleri: Eğitilmiş modelin user\_embedding katmanı, her kullanıcı için yoğun bir vektör temsili sağlar.
* Ürün Gömümleri: Benzer şekilde, item\_embedding katmanı, her ürün için gömümler sağlar.

● Tahmin Edilen Puanların Hesaplanması:○ Her müşteri (kullanıcı) için, tahmin edilen puan, kullanıcının gömümü ile ürün gömümleri arasındaki nokta çarpımı alınarak hesaplanır: \text{predicted\_scores} = \text{item\_embeddings} \cdot \text{user\_embedding} Bu puan, modelin kullanıcının her bir ürünü ne kadar beğeneceğine dair tahminini yansıtır. * Aday Öğeleri Filtreleme:
	+ Satın Alınmış Öğelerin Hariç Tutulması: Kullanıcının daha önce satın aldığı ürünler, öneri havuzundan filtrelenir.
	+ Aday Öğeler: Geriye kalan ürünler öneri için değerlendirilir.
* En İyi-N Önerilerini Seçme:
	+ Aday öğeler, tahmin edilen puanlarına göre azalan sırayla sıralanır.
	+ En yüksek puana sahip N öğe, kullanıcıya önerilmek üzere seçilir.
* Çıktı:

○ Her önerilen öğe için, ürün kimliği, ürün adı ve öneri puanı döndürülür. Bu puanlar, önerinin gücünü gösterir ve yüksek puanlar, kullanıcının tercihlerine ürünün daha iyi uyduğunu gösterir. Örnek Uygulama Verilen kodda, recommend\_and\_score\_ncf fonksiyonu, bu süreci birden fazla müşteri kimliği için otomatikleştirir. Fonksiyon: ●  Her bir müşteri kimliğini iteratif olarak ele alır. ●  Tüm öğeler için tahmin edilen puanları hesaplar. ●  Kullanıcının zaten sahip olduğu öğeleri filtreler. ●  Puanlara dayalı olarak en iyi-N öğeyi seçer ve döndürür. FP-Growth Algoritması Kullanarak Sepet Önerisi Bu aşamada, öneri sistemi kapsamında FP-Growth algoritması kullanılarak sepet önerileri oluşturulmuştur. Amacımız, sıkça birlikte satın alınan ürün kombinasyonlarını belirleyerek, müşterinin mevcut sepetine eklenebilecek ürünleri önermektir. Veri Ön İşleme ● İşlemleri Filtreleme: Öncelikle veri seti temizlenmiş, müşteri kimliği 0 olan işlemler çıkarılmıştır. Bu sayede yalnızca geçerli işlemler analiz edilmiştir. ● Zaman Damgası Çıkarma: Sonraki adımda, işlem zaman damgaları işlenerek her satın almanın tam tarihi belirlenmiştir. Bu, bir müşterinin aynı gün satın aldığı ürünleri tek bir listeye gruplamamıza olanak tanımış ve tipik bir alışveriş sepetinin içeriğini simüle etmiştir. ●  Ürünleri Gruplama: İşlemler, müşteri kimliği ve satın alma tarihine göre gruplandırılmış, her işlemde satın alınan tüm ürünler bir listeye eklenmiştir. Bu gruplama, FP-Growth algoritmasının uygulanması ve ortak ürün kombinasyonlarının belirlenmesi için temel oluşturmuştur. Birliktelik Kuralı Madenciliği ●  Sık Ürün Kümesi Oluşturma: FP-Growth algoritması kullanılarak, gruplandırılmış işlemler analiz edilmiş ve sıkça birlikte satın alınan ürün kümeleri belirlenmiştir. Bu kümeler, genellikle birlikte satın alınan ürün kombinasyonlarını temsil eder. ●  Birliktelik Kurallarının Oluşturulması: Sık ürün kümelerinden birliktelik kuralları oluşturulmuştur. Bu kurallar, ürünler arasındaki ilişkileri vurgular ve hangi ürünlerin birlikte satın alınma olasılığının yüksek olduğunu gösterir. Örneğin, bir müşteri A ve B ürünlerini satın aldığında, büyük olasılıkla C ürününü de satın alabilir. Sepet Önerisi FP-Growth tabanlı birliktelik kuralları, müşteri satın alma davranışlarına dair uygulanabilir içgörüler sunar. Bu desenleri anlayarak, sistem müşterilerin sepetlerine ekleyebileceği ek ürünler önerebilir, böylece genel alışveriş deneyimini zenginleştirebilir. **BULGULAR**Proje bulguları; 24.366 benzersiz öğe yönetimi için, ChromaDB kullanıldı ve gömme işlevi olarak Hugging Face’in all-MiniLM-L6-v2 varsayılan fonksiyonu kullanıldı. Veriler, binlik gruplar halinde işlenip yüklendi. Veritabanına verilerin tamamının yüklenmesi Google Colab üzerinde 51,8 dakika sürdü. Ayrıca, veritabanına eklenmesi gereken 10 satırlık (yeni olabilecek) küçük bir veri kümesini yüklemek 1.632 saniye sürdü. **Simülasyon Açıklaması ve Sonuçlar:** Veritabanının performansını değerlendirmek amacıyla, veritabanındaki belgelerin içeriklerinden 1.000 rastgele sorgu üreten bir simülasyon gerçekleştirildi. Simülasyon şu adımları içeriyordu: 1. **Sorgu Üretimi:** Her bir sorgu, veritabanındaki rastgele seçilmiş bir belgeden oluşturuldu ve bu belgeden rastgele bir kelime alt kümesi seçildi. Her sorgu, en fazla 10 kelimeden oluşan bir alt kümeden oluşturuldu.
2. **Performans Ölçümü:** Her sorgunun işlenme süresi kaydedildi. Bu işlem, veritabanında ilgili sonuçları bulmak için sorgulama yapılmasını ve sonuçların alınması için geçen süreyi ölçmeyi içeriyordu.

**Simülasyon Sonuçları:** * **Mod Sorgu Süresi:** 0.4491 saniye
* **Medyan Sorgu Süresi:** 0.3833 saniye
* **Sorgu Süresi Standart Sapması:** 0.1294 saniye
* **Mod Sorgu Uzunluğu:** 1 kelime
* **Medyan Sorgu Uzunluğu:** 5 kelime
* **Sorgu Uzunluğu Standart Sapması:** 2.9452 kelime
* **Ortalama Sorgu Uzunluğu:** 5.4080 kelime
* **Ortalama Sorgu Süresi:** 0.4388 saniye

Bu metrikler, ChromaDB'nin sorgu performansına dair içgörüler sunar. Medyan sorgu süreleri yaklaşık 0.38 saniye, ortalama sorgu süresi ise yaklaşık 0.44 saniye olarak gözlemlenmiştir. Standart sapma, sorgu sürelerinde bazı değişkenliklerin olduğunu, bu değişkenliğin sorguların karmaşıklığı ve uzunluğuna bağlı olabileceğini gösterir.**PROJE ÇIKTILARI** Bu projenin çıktıları, hedeflenen işlevsellikler ve yenilikçi yaklaşımlar doğrultusunda elde edilen somut sonuçları kapsayarak, projenin başarıyla tamamlanmasını ve sağladığı katma değeri özetlemektedir.* **Müşteri Segmentasyonu ile Ciro Artışı:** Müşteri segmentasyonuna dayalı stratejiler kullanılarak, müşterilere doğru zamanda alışveriş fırsatları sunulmuş ve bu sayede şirketin cirosunda **%5 (4.86) oranında** önemli bir artış sağlanmıştır.
* **Müşteri Memnuniyeti ve Sadakati:** Proje kapsamında geliştirilen sistemlerle müşteri memnuniyeti ve marka sadakati kayda değer şekilde artırılmıştır, bu da tekrar eden satış oranlarına olumlu yansımıştır. Şu anki müşteri sayısı **7 milyon** üzerinde rakamlara ulaşılmıştır.
* **Yenilikçi Öneri Sistemi Yaklaşımı:** Perakende platformlarına yönelik olarak, all-MiniLM-L6-v2 ve vektör veritabanlarının kullanan bir öneri sistemi geliştirilmiştir. Bu sistem, **gerçek zamanlı öneriler** sunabilen verimli bir backend hizmet mimarisi ile desteklenmektedir.
* **Öneri Motoru Performansı ve Kullanıcı Memnuniyeti:** Öneri motorunun performansı, kullanıcı memnuniyeti üzerindeki etkisi de dahil olmak üzere, ampirik olarak değerlendirilmiştir. Bu analizler, öneri sisteminin müşteri deneyimine olan katkısını objektif verilerle ortaya koymaktadır. Anketler ile müşterilerden gelen geri bildirimler ile çalışmanın faydası net olarak görülmektedir.

 1. Müşteri Segmentasyonu: Projede gerçekleştirilen RFM analizi sayesinde, müşteri segmentasyonu başarılı bir şekilde yapılmıştır. Bu segmentasyon, müşteri davranışlarını ve alışkanlıklarını daha derinlemesine anlamayı mümkün kılar. Her segment, farklı ihtiyaçlara, beklentilere ve davranışlara sahip olduğundan, şirketlere onları daha etkin bir şekilde hedefleme ve hizmet sunma imkânı verir.  2. Doğru Zamanlama ile Alışveriş Fırsatları: Segmentasyonun getirdiği bu derinlemesine anlayış, müşterilere doğru zamanda özel alışveriş fırsatları ve kampanyalar sunma olanağı sağlar. Bu doğru zamanlamayla, müşterilerin alışveriş yapma olasılığı artar, dolayısıyla şirketin cirosunda gözle görülür bir artış meydana gelir.  3. Müşteri Memnuniyeti ve Sadakatinin Artışı: Müşterilere özelleştirilmiş kampanyalar ve teklifler sunma yeteneği, onların markayla olan ilişkisini güçlendirir. Müşterilerin ihtiyaçlarını ve beklentilerini doğru bir şekilde karşılayan bir şirket, müşteri memnuniyetini önemli ölçüde artırabilir. Bu projenin uygulanmasının bir sonucu olarak, müşteri memnuniyetinde ve sadakatinde belirgin bir artış gözlemlenmiştir.  4. Daha Etkin Müşteri İlişkileri Yönetimi: Analitik CRM' in sağladığı yenilikçi yaklaşım, müşteri ilişkileri yönetimini (CRM) sadece bir veri toplama ve saklama aracı olmaktan çıkarıp, stratejik bir pazarlama ve satış aracına dönüştürülmüştür. Şirket, müşteri bilgilerini sadece saklamakla kalmaz, aynı zamanda bu bilgileri işleyerek satış ve pazarlama faaliyetlerini daha etkin bir şekilde yürütmektedir. Sonuç olarak, Analitik CRM projesi, şirkete müşteri ilişkileri yönetiminde yenilikçi ve etkili bir yaklaşım sunmaktadır. Projenin beklenen çıktıları ve sağladığı faydalar, şirketin hem finansal performansını (%5 ciro artışı) hem de müşteri ilişkilerini (7 Milyon üzerinde müşteri) olumlu bir şekilde etkilemektedir. Özellikle müşteri memnuniyeti ve sadakatinin artırılması, uzun vadede şirketin rekabetçiliğini ve pazardaki konumunu güçlendirmektedir. |