

## Çapraz Satışı Destekleyebilecek Transformer ile Geliştirilmiş Bir Öneri Sistemi

İbrahim Erdem KALKAN<sup>1</sup> ORCID 0000-0002-1997-5436  
Cenk ŞAHİN<sup>\*1</sup> ORCID 0000-0002-6076-7794

<sup>1</sup> Çukurova Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Adana

Geliş tarihi: 08.05.2023

Kabul tarihi: 23.06.2023

Atıf şekli/ How to cite: KALKAN, İ.E., ŞAHİN, C., (2023). Çapraz Satışı Destekleyebilecek Transformer ile Geliştirilmiş Bir Öneri Sistemi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 38(2), 571-584.

### Öz

Öneri sistemlerinin, perakende sektöründe çapraz satış bağlamında bir ürün grubunu satış için hedeflemek veya hangi müşterilerin diğerlerine göre daha pazarlanabilir olduğunu tahmin edebilen yeteneklere sahip olduğu düşünülmektedir. Bu sayede mevcut müşterilerin bir sonraki seferde hangi ürün veya hizmeti satın alacaklarına ilişkin bir tahmin oluşturularak çapraz satış etkinliği artırılabilir. Bu araştırmada temel amaç, çevrimiçi alışveriş endüstrisine, çapraz satış olanaklarını arttırabilmek bağlamında, belirli bir ürün ya da ürün grubu için, belli bir satın alma tarihçesi bulunan müşterilerinden hangilerinin diğerlerine göre daha uygun olduğunu tahmin etmek için bir öneri sistemi geliştirip sunmaktır. Bu kapsamda transformer kullanılarak probleme adapte edilmiş öneri sisteminin karşılaştırmalı bir çalışması yapılmış ve elde edilen sonuçlara göre önceki çalışmalarda sunulan modellere göre daha başarılı olduğu gözlenmiştir.

**Anahtar Kelimeler:** Transformer, Öneri sistemi, Çapraz satış, Derin öğrenme

### A Transformer-Improved Recommender System Supporting Cross-Selling

#### Abstract

It is believed that recommender systems have the ability to target a product group for sales in cross-selling context or predict which customers are more marketable than others in the retail sector. In this way, cross-selling efficiency can be increased by creating a forecast about which product or service current customers will purchase next time. The main purpose of this research is to develop and present a recommendation system to the online shopping industry to predict which customers with a certain purchasing history are more suitable than others for a particular product or product group, in order to increase cross-selling opportunities. A comparative application of the recommendation system adapted to the problem using transformers is presented in this study, and successful results were observed compared to previous studies.

**Keywords:** Transformer, Recommender system, Cross selling, Deep learning

---

\*Sorumlu yazar (Corresponding Author): Cenk ŞAHİN, [cenksahin@cu.edu.tr](mailto:cenksahin@cu.edu.tr)

## 1. GİRİŞ

İşletmelerdeki teknolojik gelişmeler ve dijitalleşme, pazarlamada hızlı dönüşümlerin ortaya çıkmasına neden olmuştur. Örneğin, biyometrik, akıllı kart ve e-ticaret uygulamaları; etkileşimli karar destek sistemleri sayesinde hızla yaygınlaşmıştır [1]. Bu dönüşümün bir sonucu olarak, etkili müşteri ilişkileri yönetimi (Customer Relationships Management, CRM) stratejilerinin geliştirilmesi, pazarlamanın önemli bir parçası haline gelmiştir. Bu bağlamda karar vericiler, mevcut tüm verileri değerlendirebilmeli ve büyüyen rekabetle başa çıkmak için sürdürülebilir veri odaklı modeller geliştirebilmelidirler. Geline nokta, yeni müşteri edinmenin çok daha maliyetli olduğu ve mevcut müşterilerle ilişkileri geliştirmenin daha iyi bir uygulama olacağı genel bir bilgi haline gelmiştir [2,3]. Bu noktada, “çapraz satış,” mevcut ilişkileri geliştirmeye odaklanmış etkili çapraz satış araçlarından biridir [2]. Mevcut bağlantıları derinleştirmek suretiyle satışların nasıl arttırılacağı ise bir iş problemi olarak karşımıza çıkar.

Çapraz satış, daha önce satın alınmış bir ürüne, bağlantılı veya bazen bağlantısız diğer ürünlerin ek satışı olarak tanımlanmıştır [4]. Çapraz satışın en önemli zorluğu, hangi ürünün hangi müşteriye hedeflenmesi gerektiğini tahmin edebilmektir. Optimal çapraz satış fırsatlarını belirleyebilmek için, pazarlama stratejistleri öncelikle, bir sonraki satın alımı yapacağı muhtemel mevcut bir müşterinin, karar aşamasında, hangi ürünü satın alacağını tahmin edebilmelidir [5]. Temel olarak, bir işletmenin birden fazla aday ürünü vardır. Gerek iletişim ve pazarlama giderleri, gerek çok zaman alıcı ve verimsiz olması nedeniyle, tüm ürünleri tüm müşterilere hedeflemek olanaklı değildir. Bununla beraber, müşteriyi çok fazla teklif ile karşı karşıya bırakarak onu rahatsız etmemek beklenmektedir [6]. Çapraz satışın en önemli noktası, hedeflenen bir ürün için uygun müşterileri araştırmak ve değerlendirmektir. Örneğin, bir çevrimiçi perakendeci belirli bir ürünlerdeki satışları arttırmayı hedefliyorsa, öncelikle mevcut portföyündeki pazarlanabilir müşterileri keşfetmeye çalışmalıdır.

Teknolojik dönüşümde, çapraz satış, müşterilerin geçmiş satın alma davranışlarını analiz etmeli, müşteri benzerliklerini belirlemeli ve müşteriyle her

etkileşimde çapraz satış fırsatlarını değerlendirebilmek için tüm analitik araçları kullanılmalıdır [4]. Bu anlamda, çok etiketli sınıflandırma algoritmaları gibi istatistiksel yöntemlerin kullanılması, çapraz satış fırsatlarının ortaya çıkarılması noktasında, alışveriş yapması beklenen müşterilerin sıradaki satın alacağı ürünün tahmin edilmesinde faydalı olacaktır [7]. Bu tarz çalışmaları yürütmek için müşteri demografik kayıtları gibi ilişkisel bir veri setinin hazırlanması gereklidir. Müşteri segmentasyonu, çapraz satış için faydalı kabul edilebilecek bir başka yöntemdir. Bu teknik, müşterileri segmentasyon sürecinde kullanılan müşteri özelliklerine bağlı olarak, bir grup içinde müşterilerin benzerliklerini ve gruplar arasındaki farklılıklarını maksimize edecek şekilde gruplara böler. Aynı grupta öteki müşterilerin satın aldığı ürünler, söz konusu müşteri için çapraz satış olarak değerlendirilebilir [8].

Çapraz satış destekleyebilecek, etkili, başka bir yöntem öneri sistemlerini (Recommender System, RS) kullanmaktır. RS'nin hedefi, genel olarak, geçmiş işlemlerden veya kullanıcı tercihlerinden elde edilen verileri kullanarak, kullanıcıların gelecekteki olası beğenileri veya ilgi alanları hakkında çıkarımda bulunmaktır [9]. RS, kullanıcılara zevklerine ve ihtiyaçlarına uygun kişiselleştirilmiş öneriler sunabilmek üzere geliştirilir. Bu sistemler, farklı türde kullanıcı geri bildirimlerinden yararlanırlar. En uygun olanı, ürünlere olan eğilimleri konusunda kullanıcıların açıkça verdiği, kaliteli, açık geri bildirimdir [10]. Puanlama verileri açık geri bildirim örneklerinin en yaygın olanlarındandır. Bu tür veri setini kullanan sistemler, çevrimiçi satış gibi bazı endüstrilerde, yeterli miktarda açık geri bildirim toplamak zor olduğundan, daha fazla kullanıcı çabası, zamanı ve dikkati gerektirir. Açık geri bildirim toplama ile ilgili çeşitli zorluklar vardır. İlk olarak, kullanıcılar genellikle tüketmeleri gerektiğini düşündükleri öğelere yüksek puan verme eğilimindedirler, bu durum sonuç olarak verinin yanlışlığının yüksek olmasına neden olur. İkinci olarak, kullanıcılar satın aldıkları ürünlerin sadece küçük bir bölümüne puan verme eğilimindedirler [11,12]. Bu ikinci problemden, açık geri bildirim her zaman kullanıcıların gerçek tercihini yansıtmayabileceğini göstermektedir. Bu zorluklardan dolayı ilgi,

giderek, satın alınan ürün, izlenen yayın, dinlenen şarkı, kiralanan veya okunan kitap, hatta tıklanan reklamlar gibi örtük geri bildirim kaymaktadır. Yalnızca pozitif veriden oluşan satın alma geçmiş verisi tipik bir örtük geri bildirim örneğidir [12].

Çok çeşitli tekniklerin kullanılabilirdiği RS uygulamaları için oldukça geniş bir akademik literatür bulunduğu söylenebilir. Adomavicius ve Tuzhilin [13], belli başlı teknikler hakkında oldukça kapsamlı bir çerçeve sunabilmektedir. Bu çalışmaya göre, bu teknikler sezgisel tabanlı ve model tabanlı olmak üzere iki ana başlık altında incelenebilmektedir. Sezgisel tabanlı teknikler, farklı benzerlik ölçüleriyle yakın komşu yaklaşımlarını içerirken, model tabanlı yöntemler genellikle yapay sinir ağları (YSA), Bayes sınıflandırıcılar ve kümeleme gibi veri madenciliği tekniklerini içerir. Tüm yöntemler arasında, YSA'nın kullanımı son zamanlarda hız kazanmıştır. Derin öğrenme (Deep Learning, DL) tabanlı öneri sistemleri (DL'nin sinir ağlarıyla ilgilenen makine öğrenmesinin bir alt kümesi olması nedeniyle, bu çalışmada YSA ve DL terimleri eş anlamlı olarak kullanılmıştır.) geleneksel modellerin zayıflıklarını aşarak yüksek öneri kalitesi elde etmesi nedeniyle ilgi çekmeye devam etmektedir [14]. Özyinelemeli sinir ağları (Recurrent Neural Networks, RNN) ve onun özel bir mimarisi olan uzun kısa-sürelilik bellek (Long Short-term Memory, LSTM) gibi teknikler, standart en yakın komşu ve matris faktörizasyonu (MF) gibi yöntemlere kıyasla oldukça etkili sonuçlar ortaya çıkarabilmektedir [15]. Buna ilaveten, RNN'nin yerini alabilecek ve doğal dil işlemedeki en ileri teknik olarak kabul edilen transformer [16] mimarisinin yapısal veri ile çalışan örneklerinin daha etkili sonuçlar ortaya çıkarabilmesi beklenmektedir.

Bu çalışmada temel amaç, çevrimiçi alışveriş endüstrisine, çapraz satış olanaklarını arttırabilmek bağlamında, belirli bir ürün ya da ürün grubu için, belli bir satın alma tarihçesi bulunan müşterilerinden hangilerinin diğerlerine göre daha uygun olduğunu tahmin etmek için bir çözüm çerçevesi önerilebilir. Bu çalışma, özellikle örtük geri bildirim verisini işleyen, transformer ile geliştirilmiş bir RS modelinin açık kaynaklı bir veri seti ile sınanmasını ve sonuçlarının RNN ile

geliştirilmiş modeller ile karşılaştırılmasını içermektedir. Bu anlamda, bu çalışmanın bilimsel katkısı iki yönlü olarak düşünülebilir: (1) çevrimiçi alışverişte çapraz satışa analitik bir yaklaşım sunması nedeniyle pazarlama yönetimi bilgi birikimine veri temelli bir çerçeve sunması ve (2) bir öneri sisteminin karşılaştırmalı olarak test etmesi ve sonuçlarını ortaya koyması. Bu çalışma, ürüne ve kullanıcıya ait ek özelliklerin birleştirilmesinin modellere faydalı bilgi sağladığını ve değerlendirme sonuçlarını iyileştirdiğini göstermektedir.

Çalışmanın takip eden bölümü, farklı veri madenciliği araçlarını kullanan çapraz satış uygulamaları ve derin öğrenme ile geliştirilmiş öneri sistemleri çalışmaları hakkında bilgi vermektedir. Öneri sistemi ve transformer yapısının teorik arka planı üçüncü bölümde ele alınacak olup dördüncü bölümde açık kaynaklı veri seti üzerindeki karşılaştırmalı sonuçları sunulacaktır. Son bölümde ise araştırma bulguları ve çıkarımlar sunulacaktır.

## **2. ÖNCEKİ ÇALIŞMALAR**

Bu çalışma ile ilgili olabilecek önceki çalışmalar iki alt kategoride incelenmektedir. İlk olarak, literatür, çapraz satış uygulamaları açısından incelenmekte ve ardından farklı yöntemleriyle RS uygulamaları sunulmaktadır.

Çapraz satış açısından bakıldığında, 2000'lerin başında veri odaklı modellerin kullanıma eğilimi göze çarpmaktadır. Kamakura ve arkadaşları [2] çalışması, müşteri alışveriş verilerine dayalı, çapraz satış için geliştirilmiş bir modeli ve onun genişletilmiş versiyonlarını sunan, bu alanda ilk olmamakla birlikte oldukça ilham verici bir çalışmadır. Çalışmalarını, finansal işlemlere, 22 farklı finansal ürüne ve müşterilere ait öznitelik verilerini kullanarak, 5.550 farklı müşteri için uygulamışlardır. Knott ve arkadaşları [6], çapraz satış bir sonraki satın alma (next-product-to-buy, NPTB) modeli olarak ele alır. Belirli bir müşterinin en olası satın alma ürününü tahmin etmek için; veri toplama, model seçimi, model değerlendirmesi ve müşterilerin puanlanması ve hedeflenmesi olmak

üzere dört ana adımdan oluşan bir NPTB çerçevesi sunarlar. Çalışmasını, 14 farklı perakende finansal ürün ve 270.842 müşteriden oluşan bankacılık alanından bir veri seti ile geliştirilen dört istatistiksel modeli bir sezgisel saha testi ile karşılaştırarak uygular. Thuring ve arkadaşları [17], Avrupa'dan bir sigorta şirketi tarafından sağlanan geçmiş satın alma davranışı veri kümesini kullanarak hedef müşteri seçimi için bir yöntem uygularlar. Geuens ve arkadaşları [11], çalışması, RS ve sınıflandırma modellerinin etkililiklerini karşılaştırma yaklaşımıyla farklı bir bakış açısına sahiptir. Çalışmalarında, 83 farklı finansal ürün ve 96.602 benzersiz müşteriden oluşan veri setleri üzerinde geliştirdikleri, Adaboost ve rastgele orman gibi "state-of-the-art" çok-etiketli sınıflandırma tekniklerinin RS yaklaşımından biraz daha iyi olduğu sonucuna varırlar.

Önceki çalışmalara bakıldığında, yeniden satın alma tahmini modellerinin sadece CRM'nin ayrılmaz bir parçası değil, aynı zamanda çapraz satış uygulamalarıyla da yüksek derecede ilişkili olduğu görülebilir. Bu tür modeller de bir sonraki satın alma işlemi tahmin etmek için önceki müşteri alışveriş verilerinden özellik çıkarımı yapmak suretiyle yararlanır. Martinez ve arkadaşları [18], bir zaman serisi verisini, 276 adet tahmin edici oluşturarak makine öğrenmesi algoritmalarını değerlendirmek üzere kullanırlar ve en iyi sonuçlarını "gradient tree boosting" algoritması ile elde ederler. Chou ve arkadaşları [19] ise, benzer bir veri setinden yaklaşık 100 tahmin edici elde ederek bir RNN mimarisinden daha iyi sonuçlar üreten yeni bir yaklaşım önermektedir.

RS çalışmaları açısından, matris faktörizasyonu (MF) en yaygın yöntemlerden biri olmaktadır [20]. Gizli faktör analizinden türetilen popüler bir teknik olan MF, temel olarak kullanıcıların öğe tercihlerinden oluşan seyrek bir matrisin düşük boyutlu çözümlenmesine dayanır [21]. MF'nin, negatif olmayan MF [22], olasılıksal MF ve Bayes versiyonları [23] gibi farklı versiyonları hakkında kapsamlı bir literatür bulunmaktadır. He ve

arkadaşları [17] gömülmüş katmanlar ile MF'nin DL çıkarımını tanımlar ve RS için üç yaklaşım önerirler: (1) genelleştirilmiş MF (GMF), (2) çok katmanlı sinir ağı (Multi Layer Perceptron, MLP) yaklaşımı ve (3) sinir ağı ile işbirliği filtresi (Neural Collaborative Filtering, NCF). Daha geniş kapsamlı DL tabanlı RS geliştirmeleri için Zhang ve Yao [25] kapsamlı bir araştırma sunarlar. Kısıtlanmış Boltzmann makinelerinin [26] kullanımı en ilham verici çalışma olarak kabul edilmekle beraber, MLP'lerden daha karmaşık yapı taşlarına kadar, etkili bir RS elde etmek için çeşitli sinir ağları uygulanabilmektedir.

RNN ile yapılmış RS uygulamalarının çoğu, kullanıcı-ürün etkileşimleriyle zaman farkındalığına vurgu yapar. Bundan dolayı, puanlama ve alışveriş gibi kullanıcı etkileşimleri, zaman sırasına göre bir dizi haline gelir. Bu uygulamalar genellikle, bir hedef kullanıcının muhtemelen tüketeceği bir sonraki öğeyi tahmin etmeyi amaçlayan tüketilmiş öğelerin sıralı bir listesi gibi önceki kullanıcı-ürün etkileşimlerini dikkate alır [27]. Özel bir RNN mimarisi olan LSTM ile uygulanmış modeller, en yakın komşu ve MF yöntemleri ile karşılaştırıldığında oldukça etkili sonuçlar ortaya koyabilirler [15]. Oturum temelli zaman sıralı verileri ele almak açısından, RNN mükemmel bir araç olarak kabul edilebilir. Oturum, kullanıcı ve sistem arasındaki belirli bir zaman aralığındaki etkileşimler olarak kabul edildiğinden, oturum temelli RS kullanıcı ile gerçek zamanlı etkileşimleri yönetebilmek üzere geliştirilir [20,28, 29,30]. Bu kapsamdaki örneklerden Smirnova ve Vasile [31], oturum temelli tıklama verisi üzerine bağlamsal özelliklerin de ilave edilmesiyle elde edilen bir vektör üzerine geliştirilmiş bir model önermektedirler. LSTM ve zincir veri kullanan modeller arasında Kalkan ve Şahin [32], önceden filtrelemek suretiyle elde edilmiş hedef ürünlerin olasılıklarının hesaplanması yerine tüm ürünlerin olasılık değerlerinin hesaplanmasını içeren, örtük geri bildirim, ürüne ve kullanıcıya ait özelliklerin de ilave edilmesini içeren bir model tanıtmışlardır. Geniş ve derin ağlar olarak tanıtılan bir diğer karmaşık sinir ağı modelinde [33], gömülü

vektörleri girdi kabul eden derinleşen bir MLP ile kullanıcı örtük geri bildirimini işleyen bir geniş ağ yapısı bütünleşmiş şekilde kullanılmıştır. Bu çalışmaya eleştirel yaklaşan bir başka çalışma, doğal dil işleme alanındaki gelişmelerin adapte edilmesi yoluyla, çok başlı öz-dikkat mekanizmasını uygulayan davranışsal zincir transformer (Behavioral Sequence Transformer, BST) modelini tanıtmıştır [34]. Bu model, 200 milyondan fazla kullanıcıya sahip çevrimiçi alışveriş ürün tıklama veri seti üzerinde eğitilmiş ve tatmin edici sonuçlar üretmiştir.

### 3. TEORİK ÇERÇEVE

Bu bölümde, çalışmanın teknik arka planı ve teorik bileşenleri açıklanacaktır; ancak, daha öncesinde 3.1 numaralı bölümde ele alınan problemin matematiksel tanımının sunulması gerekmektedir. 3.2 numaralı bölüm ise, önerilen modelin bileşenlerini açıklamakta ve her biriyle ilgili kısaca teorik bilgi vermektedir.

#### 3.1. Problemin Tanımı

Önceki çalışmalarda (örneğin [7]), RS'nin etkili bir pazarlama çerçevesinde faydalı bir araç olarak kabul edilebileceği vurgulanmıştır. Şekil 1'in ifade ettiği gibi, RS modelinin her müşteri için, daha önce satın alınmış olunsun veya olunmasın, önerilen ürünleri sunması, seçilen müşteriler için hedef ürünler kümesi veya belirli bir ürün için hedef müşteriler kümesi olarak yorumlanabilir.

Bu çalışmada, müşterilerin bir sonraki sefer hangi ürünleri veya hizmetleri satın alacaklarına dair olasılıklarını tahmin ederek çapraz satış etkinliğini arttırmak şeklinde yaygın bir iş problemi üzerinde durulmaktadır. Bu bağlamda, önerilen sistem Çizelge 1'de soyut bir temsil olarak gösterildiği gibi bir temel alışveriş veri kümesini kullanmaktadır. Bu problemin açıklanmasındaki matematiksel temsil büyük ölçüde Martinez ve arkadaşları [18] çalışmasından uyarlanmıştır. Her müşterinin her alışverişi, bir zaman damgasıyla bir satır olarak işlemler veri setine kaydedilir.

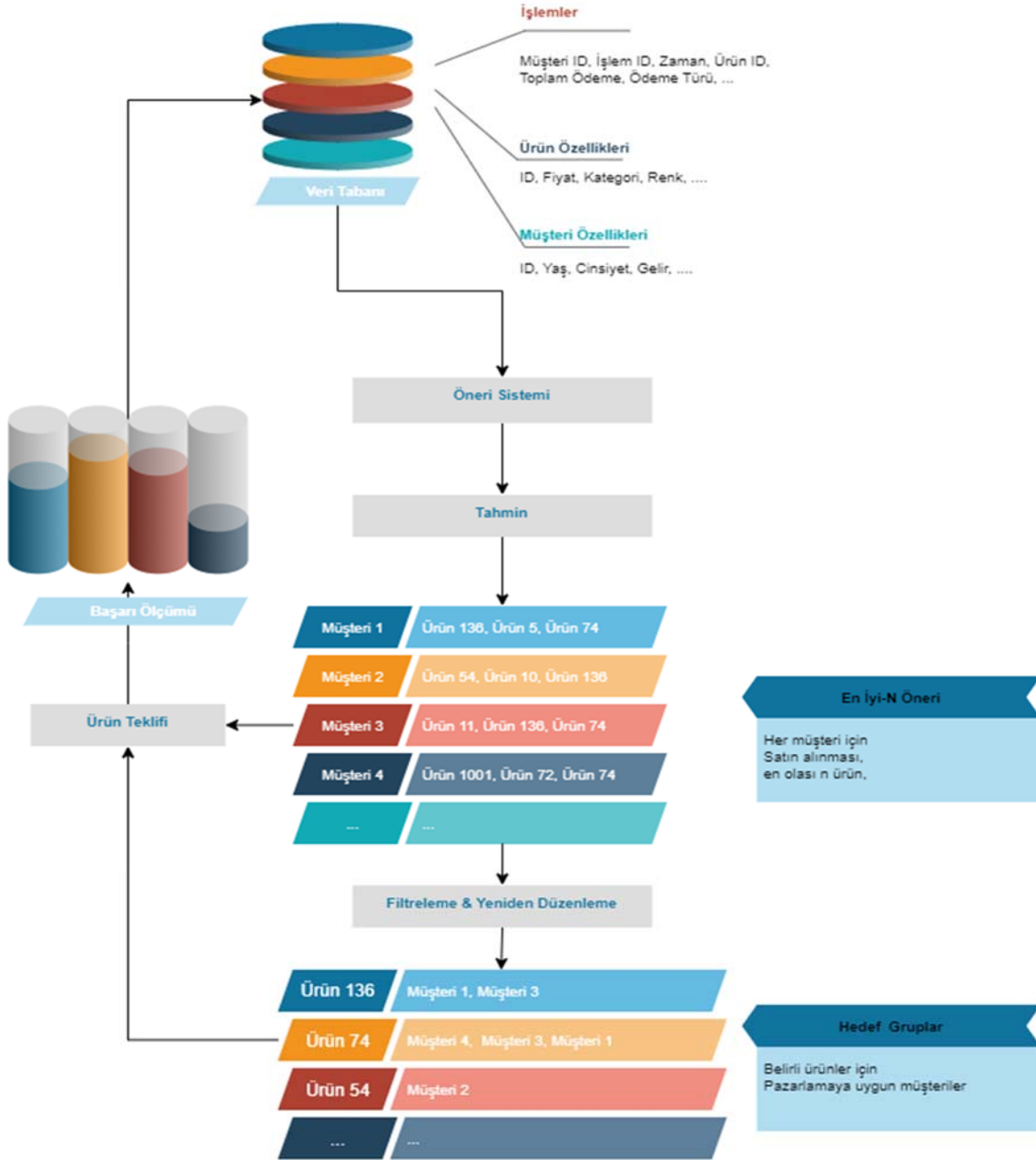
Bir çevrimiçi perakendecinin portföyündeki toplam müşteri sayısı  $U$  olarak gösterilirse, ID ile kayıtlı her bir müşteri,  $u$ , ( $u \in \{1, \dots, U\}$ ),  $N_u$  işlemleri göstermek üzere, belli bir periyotta kaydedilmiş bir satın alma geçmişine,  $i_u, N_u$ , sahip olacaktır. Her satın alma davranışında bir zaman damgası  $t_{u,n}$ , ve bir ürün,  $i_{u,n}$ , bulunmaktadır. Tüm işlemlerin her bir müşteri için kronolojik olarak sıralandığı varsayılmaktadır. Bir başka ifadeyle; her bir müşterinin,  $u \in \{1, \dots, U\}$  satın aldığı her ürün  $i_{u,1} \leq i_{u,2}$  için  $t_{u,1} \leq t_{u,2}$  olacaktır. Bu durumda verilen bir tarihsel veri için:  $\{(t_{u,n}, i_{u,n}) | n = 1, \dots, (N-1)_u\}$  problem, söz konusu müşteri için, gelecek satın almadaki  $N_u$ , ürünün önceden belirli bir katalog,  $i \in \{i_{U,N_u}\}$  içerisinde hangisi olacağının tahmini şeklinde tanımlanabilir.

**Çizelge 1.** İşlemler veri seti temsili gösterimi

Müşteri ID	İşlem ID	Zaman damgası	Ürün
$u$	$n$	$t_{u,n}$	$i_{u,n}$
1	1	$t_{1,1}$	$i_{1,1}$
...	...	...	...
1	$N_1$	$t_{1,N_1}$	$i_{1,N_1}$
2	1	$t_{2,1}$	$i_{2,1}$
...	...	...	...
$U$	1	$t_{U,1}$	$i_{U,1}$
...	...	...	...
$U$	$N_u$	$t_{U,N_u}$	$i_{U,N_u}$

Öte yandan bu şekilde tanımlanmış bir problem, zamansal sınırlamalar eklenmek suretiyle genişletilebilir. Başka bir ifadeyle herhangi bir müşterinin  $t_a$  ve  $t_b$  zaman aralığındaki alışverişleri veri olmak suretiyle  $t_b$ 'den belirli bir süre sonraki satın almada hangi ürünün tercih edileceğinin tahmini şekline getirilebilir.

Şu ana kadar bahsedilen tüm veri öğeleri kronolojik olarak sıralanabildiği için zincir olarak ele alınabilir. Ancak, müşterilere özgü değişkenleri içeren verilerin herhangi bir zaman özelliği olmadığından, bunlar statik özellikler olarak kabul edilir. Çizelge 1'e, müşteriye ait değişkenler ve ürüne ait değişkenler gibi yan bilgi kaynaklarından oluşan farklı veri kümeleri ilave etmek de mümkündür.

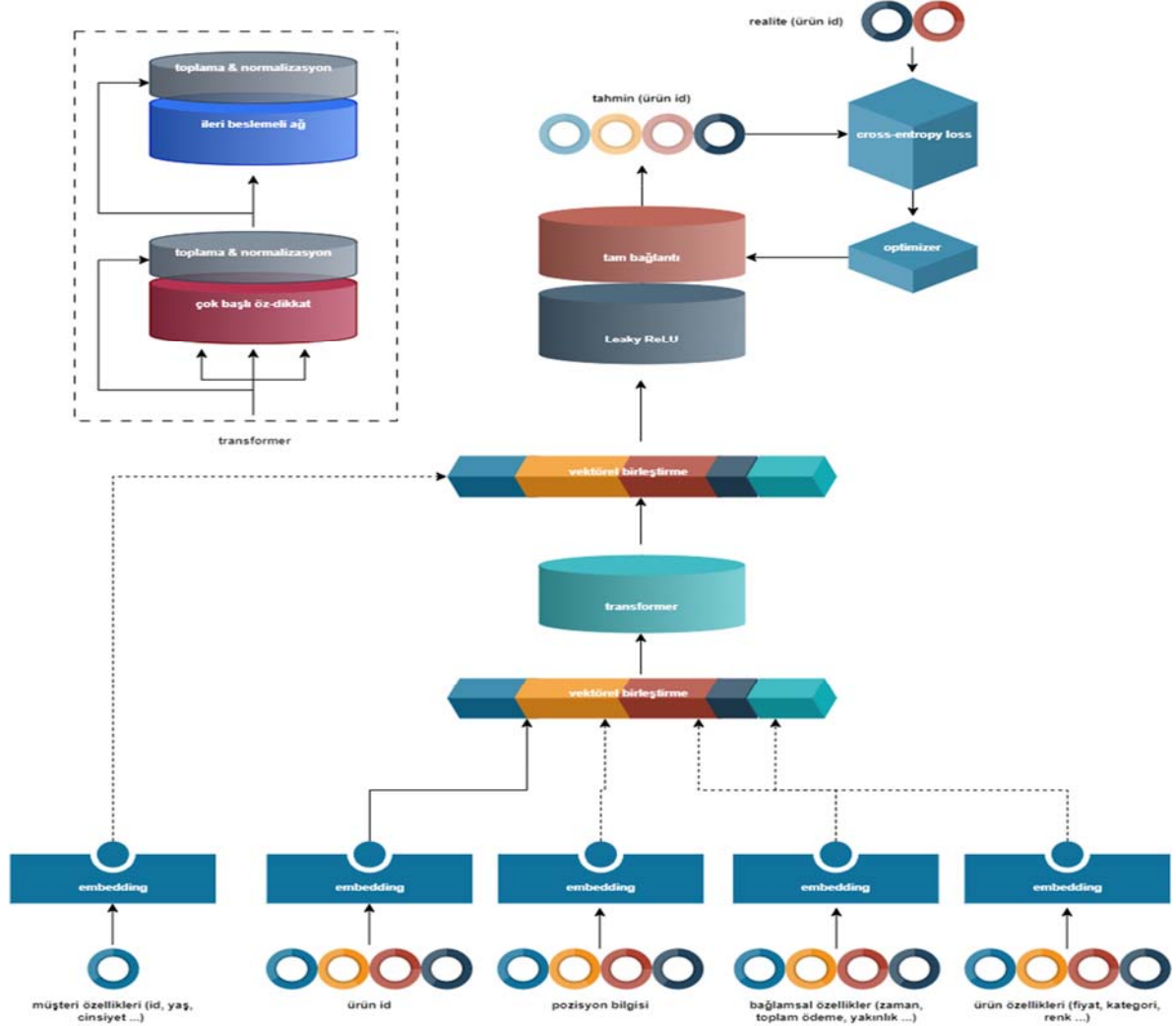


Şekil 1. Çapraz satış sisteminin genel çerçevesi

### 3.2. Önerilen RS Modeli

Bu çalışmada, farklı yan bilgi kaynaklarının da eklenmesi ve BST modelinin çevrimiçi alışverişte satın alma verisine uyarlanmasıyla elde edilecek,

orijinal çalışmadan [34] farklı olarak ön filtrelemeye tabi tutulmaksızın tüm katalog ürünleri üzerinde olasılık dağılımını üretebilecek bir RS modeli önerilerek test edilecektir. Şekil 2, modelin bileşenlerinin tam bağlantı yapısını göstermektedir.



Şekil 2. Transformer yapısını da kapsayan, modelin şematik gösterimi

### 3.2.1. Girdi Yapısı

Şekil 2'de görülen girdiler içerisinde "ürün id" müşterinin satın alma geçmişine atıfta bulunurken, ürüne bağlı diğer girdiler ayrı olarak gösterilmiştir. Satın alma geçmişi ise sadece pozitif olarak temsil edilebilen örtük geri bildirimdir. Buna ilaveten, ürünlerin gömülü pozisyonel konumlarının da girdilere eklenmesi önerilmiştir [16,34]. Pozisyonel konumlar her bir olay için sıra bilgisini tutmakta olup bu modelde bir seçenek olarak dikkate alınmıştır. Alışveriş olayının bağlamına vurgu

yapan ödeme biçimi, toplama ödeme miktarı, olay günü gibi girdilerin ve müşteriye özgü zamana göre değişmediği varsayım olarak kabul edilen, cinsiyet, yaş, gelir gibi girdilerin modele beslenmesi ve model içindeki yayılımı Şekil 2 de gösterildiği gibidir. Daha önce bahsedildiği gibi (bkz. Bölüm 1) bu çeşit bir verinin toplanabilmesindeki zorluklara binaen model girdileri arasında bir seçenek olarak yer almıştır.

Zamana bağlı olarak anlam ifade ettiği düşünülen tüm girdiler bir zincir vektörü haline getirilerek

kullanılmaktadır. Orijinal çalışmanın [34] aksine bu çalışmada ürüne bağlı özellikler de her satın alma olayında ürünün değişebileceği varsayımı ile sıralı olarak düşünülebilir kullanılmıştır. Bununla ilaveten girdilerin kodlanarak daha düşük boyutlarda temsil edilebilmesi amacıyla Şekil 2’de görülen ilk katman gömülü (embedding) katmandır. Modelin bu kısmında bir seçenek olarak girdilerin öncelikle kesikli hale getirilmesi ve 1 ve 0 gösterimi (one-hot-encoding) ile kodlanması da mümkündür.

Kodlanan tüm girdilerin transformer katmanına aktarılabilmesi için birbirlerini eklenerek birleştirilmesi ve tek bir vektör olarak temsil edilebilmesi gereklidir.

### 3.2.2. Transformer Katmanı

Şekil 2’de görülen kodlanmış sıralı verinin yayıldığı ilk yer transformer katmanıdır. Bu bölüm modelin ürünlerin birbirleri ile ilişkilerini içerebilecek şekilde yeni bir gösterimini öğrendiği kısımdır. Şekil 2’deki detay gösteriminde fark edildiği gibi çok başlı öz-dikkat mekanizmasını bünyesinde barındırmaktadır. Genel olarak dikkat mekanizması, model eğitimi sırasında belirli özellik bilgilerine odaklanabilme amacıyla görüntü işleme alanında önerilmiştir. Öz dikkat veya iç-dikkat mekanizması ise girdilerin içsel bağlantılarına odaklanarak, girdiler arası bir iç bağlantı vektörü hesaplar bu sayede harici veri bağımlılığını azaltan bir gelişme olarak kabul edilir [16,35]. Genel dikkat mekanizması modelin çıktısıyla ilgili önemli bölümleri keşfetmeye çalışırken, öz dikkat mekanizması, girdilerin birbirleriyle ilgili önemli kısımları bulmaya çalışmaktadır [36].

Orijinal çalışmalarında [16] ölçeklendirilmiş noktasal-çarpım dikkat (Scaled Dot-Product Attention) olarak adlandırdıkları hesaplama çok başlı öz dikkat mekanizmasının giriş katmanını oluşturmaktadır. Bu sistemde girdi,  $d_k$  boyutlu sorgu ve anahtarlar (Query and Keys,  $Q, K$ ) ve  $d_v$  boyutlu değerlerden (Values,  $V$ ) oluşur. Sorgular tüm anahtarlarla noktasal çarpıma tabi tutulmakta ve her biri  $\sqrt{d_k}$  ile bölünmektedir, değerler üzerinde ağırlıkları elde etmek için *softmax* fonksiyonunu uygulanmaktadır. Ölçeklendirilmiş

noktasal-çarpım dikkat 3 adet çıktı matrisi hesaplamakta ve çok başlı bir öz dikkat mekanizmasına iletmektedir [16,36]:

$$\begin{aligned} \text{Attention}(Q, K, V) &= \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \\ S &= MH(E) \\ &= \text{Birleştire}(head_1, head_2, \dots, head_n)W^H, \\ head_i &= \text{Attention}(EW^Q, EW^K, EW^V) \end{aligned}$$

Bu ifadelerde  $W^Q, W^K, W^V$  üç adet projeksiyon matrisini göstermekte iken  $E$ , gömülü girdi matrislerini göstermektedir.  $T$ , toplam pozisyon sayısını,  $H$ , toplam *head* (baş) sayısını ifade etmektedir.

### 3.2.3. Tam Bağlantılı Katmanlar

Bu bölüm modelin tam bağlı sinir ağı katmanlarını içermektedir (bkz: Şekil 2). Modelin girdi ve çıktıları arasındaki ilişki bu bölümde açıkça görülür. Eğitimde modele beslenen çıktı, ürün girdi dizisinin bir zaman adımı kaydırılması sonucu elde edilen dizinin son ögesidir. Bu öge müşterinin satın aldığı son ürünü temsil etmekle beraber öncesinde satın aldığı ürünler modele bir girdi olarak sunulmaktadır. Bir tahmin modeli olarak her bir müşteri için sadece ürünlerden oluşan girdi:  $\{(t_{u,n}, i_{u,n}) | n = 1, \dots, (N-1)_u\}$  iken; çıktı:  $\{t_{u,N}, i_{u,N}\}$  olacaktır. Modelin son katmanı, eğitimde kullanılan tüm ürünlerin katalog boyutu şeklinde tahminleri içerir. Yüksekten düşüğe doğru sıralanmış olasılıklardan, önceden belirli bir sayı olarak, ilk  $k$  tanesi, söz konusu müşteri için öneri listesini oluşturacaktır.

Farklı bir bakış açısından modelin çıktıları bazı yönlerden bir sınıflandırma temelli çıktılar olarak kabul edilebilecektir [20]. Bu nedenle, böyle bir sınıflandırma hedefi için minimize edilecek bir kategorik çapraz entropi (Categorical Cross-Entropy) kayıp fonksiyonu uygun olacaktır:

$$L = \sum_{m=1}^M y_m \log p(m)$$

$M$ , eğitim verisindeki farklı ürün sayısını,  $y_m$ , olması gereken ürünü,  $p(m)$  ise tüm ürünler üzerinde olasılık dağılımını göstermektedir.



## 4. DENEYSEL ÇALIŞMA

Bölüm 3'te tanıtılan RS yapısı, açık kaynak Python kütüphaneleri sayesinde uygulanmış ve test edilmiştir. Çok başlı dikkat, doğal dil işleme yöntemleri gibi bazı uygulamalarda Keras, [37] ve Tensorflow [38] kütüphanelerinin sunduğu hazır katmanlar ve yapılar kullanılmıştır. Tüm GPU tabanlı deneyler için, Kaggle [39] platformunun sunduğu bir adet ücretsiz Tesla P100-PCIE-16 GB GPU kullanılmıştır. Bu bölümde elde edilen sonuçlar [32]'de elde sonuçlar ile kıyaslanmıştır. Bu nedenle veri seti seçimi ve üzerinde yapılan tüm işlemler kıyaslama yapılabilmesi açısından çalışmada tarif edildiği gibidir.

### 4.1. Veri Seti İşlemleri

Deneysel çalışmanın yürütülebilmesi için sıralı ve statik türden değişkenleri aynı anda içerebilen bir veri setine ihtiyaç bulunmaktaydı. Bu türden bir veri seti alışveriş tarihçesini ürüne bağlı özellikleri ve müşterilere özgü değişkenleri içermelidir. Bu nedenle açık kaynaklı bir çevrimiçi alışveriş veri seti tercih edilmiştir: Pakistan's Largest E-Commerce Dataset, [40]. Veri seti Temmuz 2016'dan Ağustos 2018'e kadar zaman damgalarıyla birlikte, yaklaşık yarım milyon sipariş kaydı, ilgili ürünlerin kategorileri, fiyatları ve sipariş edilen miktarları ve diğer özellikleri içermektedir. Her müşteri işlemi için bir kayıt vardır ve her müşteri için satın aldığı ürünlere bağlı olarak birden çok kayıt bulunabilir. Veri seti sadece tamamlanmış siparişler filtrelenerek kullanılmış olup bu açıdan 233.685 gözlem içerir.

**Çizelge 2.** İlgili olabilecek uygun kolonların filtrelenmesiyle elde edilen veri seti

Kolon Adı	Açıklama
t_id	Her bir işlem için anahtar
t_date	İşlem tarihi
c_id	Her bir müşteri için anahtar
c_since	Müşterinin sisteme kayıt tarihi
item_name	Ürün ismi
item_category	Ürün kategorisi
price	Her bir ürünün fiyatı
amount	Her bir işlem için işlem miktarı

Gerek içerdiği kayıtların kalitesi ve gerekse modelde kullanmaya uygunluk açısından kullanılabilir olan tüm özellikler kolon başlıklarında filtrelenerek ayrılmıştır. Bu kolonların isimleri ve kısa açıklamaları Çizelge 2'de sunulmuştur.

Çizelge 3, her bir müşterinin yaptığı satın alma sayısının dağılımına ilişkin özet istatistikleri içerirken, Çizelge 4'te her bir ürünün satın alma sayısının dağılımına ilişkin özet bilgiler sunulmuştur.

**Çizelge 3.** Müşterilerin satın alma sayılarına göre özet istatistik değerler

<b>Minimum</b>	1,0
<b>1. Çeyrek</b>	1,0
<b>Ortalama</b>	5,2
<b>Standart sapma</b>	26,95
<b>3. Çeyrek</b>	3,0
<b>Maksimum</b>	2863,0

**Çizelge 4.** Ürünlerin satın alınma sayılarına göre özet istatistik değerler

<b>Minimum</b>	1,0
<b>1. Çeyrek</b>	1,0
<b>Ortalama</b>	3,48
<b>Standart sapma</b>	16,82
<b>3. Çeyrek</b>	3,0
<b>Maksimum</b>	1657,0

Çizelge 3 ve 4'te özetlendiği gibi, veri setindeki müşterilerin çoğunluğunun sadece bir satın alma işlemi yaptığı ve bazı ürünlerin nadir gözlemlendiği tespit edilmiştir. Veri setinde gözlenen bu tarz bir seyrekliliğin üstesinden gelmek için, Bogaert ve arkadaşlarının [7] çalışmasında tarif edildiğine benzer bir yaklaşım kullanılmıştır. Bununla beraber çalışmada müşterilerin satın alma geçmişi önemli yer tutmaktadır, bu tarz bir geçmişten söz edilemeyen müşteri kayıtları veri setinden çıkarılmıştır. İlk filtreleme işleminde, iki satın alma işleminden daha az işleme sahip olan müşteriler çıkarılmıştır. İkinci filtrelemede ise iki defadan az satın alınmış nadir ürünler veri setinden elenmiştir. Çok nadir gözlenen ürün kategorileri için ise "diğer" adı altında bir kategori oluşturulmuş ve 2.500'den az gözlenen kategorilerdeki tüm ürünler

bu yeni kategori altına yerleştirilmiştir. Bu işlemlerden sonraki veri bileşimi Çizelge 5'te özetlendiği gibi olmuştur. Çizelge 5, tüm veri önileme adımları tamamlandıktan sonra elde edilen veri setinin içeriğini göstermektedir.

**Çizelge 5.** Önilemeden sonra elde edilen veri seti görünümü

İşlem sayısı	150.075
Müşteri sayısı	25.562
Ürün sayısı	12.861
Ürün kategorisi sayısı	13

Dizi olarak ele alınabilecek öznitelikler, bağlamsal özellikler olarak hazırlanmıştır. Türetilen bağlamsal özellikler aşağıda açıklanmıştır:

*recency*: Satın alınan her ürün için satın alma tarihi ile modelleme tarihi arasındaki gün sayısıdır. Değerleri 1 ile 100 arasında kategorize edilmiştir.

*month*: Her satın alınan ürünün ay numarası, 1 ile 12 arasında bir tamsayı olarak ifade edilir.

*dayofweek*: Her satın alınan ürünün haftalık gün numarası, 1 ile 7 arasında bir tamsayı olarak ifade edilir.

*payment*: Her işlem için ödenen toplam para miktarıdır. Fiyat ve miktarın çarpımıyla elde edilir. Değerleri 1 ile 1000 arasında kategorize edilmiştir.

Anahtar numaralı hariç ürüne bağlı özellikler de dönüştürülüp kullanılmıştır:

*price*: Ürün fiyatından oluşturulmuş olup değerleri 1 ile 500 arasında kategorize edilmiştir.

*category*: Her bir ürünün dahil olduğu kategoriyi ifade etmektedir. 1 ile 13 arasında bir tamsayı olarak ifade edilir. Bununla beraber, kullanıcı profiline uygun öneri oluşturma fırsatlarını yakalamak için statik veri özelliklerinin entegrasyonu da kullanılmaktadır. Bu amaçla, sisteme kaydolma tarihi ile modelleme tarihi arasındaki gün sayısı başka bir özellik olarak kullanılmaktadır (*customer\_since*). Değerleri 1 ile 100 arasında kategorize edilmiştir.

Tüm müşteri satın alma olayları zaman sırasına konulmuştur. Ürün kimlikleri ve ürünle ilgili ek özellikler kodlanarak sıralı diziler oluşturulmuştur. Tüm müşteriler eşit miktarda işlem yapmadığından, değeri deneysel olarak belirlene bir sayı kadar işlem sıralanmış eksik satın alma olayları 0 yazılarak doldurulmuştur. Model eğitimi aşaması için, her müşterinin son satın alma etkinliği realite olarak tutulmuş ve önceki etkinlikler girdi olarak saklanmıştır. Son etkinlik hariç tüm satın alma etkinliklerine ait diziler, modele girdi olarak verilmektedir.

## 4.2. Modelleme

Modeller, iki ayrı senaryoya göre eğitime tabi tutulup kıyaslanmıştır. İlk senaryo Monte Carlo Çapraz Doğrulama (MCCV) [41] tekniğidir. İkincisi eğitim öncesinde ayrılmış bir test veri seti üzerinde performans ölçümüdür. Yaklaşık olarak %10'luk bir kayıt rastgele seçilmiş ve test amaçlı ayrılmıştır. Geri kalanı üzerinde ise,  $k$  katlı alt örnekler halinde bölmek yerine, tekrarlı rastgele yeniden örnekleme ve doğrulama tekniği uygulanmıştır. MCCV, modelleme kısmının yaklaşık %80'inin eğitim ve %20'sinin doğrulama için rastgele bölündüğü şekilde beş kez gerçekleştirilir. Tüm doğrulama denemeleri maksimum 200 iterasyon (epoch) ve erken durdurma sistemi ile çalıştırılmıştır. Denemeler, her iterasyonda, doğrulama setinde bir performans metriği hesaplanması ve ardışık üç iterasyonda herhangi bir iyileşme elde edilemezse, çalışmayı durduracak şekilde ayarlanmıştır.

Hiper-parametrelerin en uygun değerlerinin ayarlanabilmesi için, grid arama yöntemi uygulanmış olup, veri setinin her yeniden örneklemeinden bir hiper-parametre birleşimi denenmiştir. Kombinasyon sayısı yüksek olacağından bazı hiper-parametreler ön deneyleri ile sabitlenmiştir. Öğrenme oranı ve optimizasyon değeri değiştirilerek anlamlı bir iyileşme gözlenememesinden dolayı bunlar, sırasıyla 0.0001 ve "adaptive momentum optimizer" (ADAM) olarak sabitlenmiştir. Öte yandan, dropout oranı %20 olarak sabitlenmiş, amaç fonksiyonu değiştirilmemiş ve ekstra gizli katmanlar eklemenin de anlamlı bir iyileştirme sağlamadığı görülmüştür.

Üç başlı öz dikkat sistemi kullanılmış olup giriş katmanının boyutu içi modelin tüm girdileri birleştirildiğinde elde edilen boyut miktarı ve bunun iki katı denenmiştir. Tam bağlantılı çıktı katmanı ve gömülü ürün vektörlerinin çıktı boyutu için 512 ve 1024 sayıları denenmiştir. Sıralı dizi pozisyon değerleri için 10 ve 12 sayısı üzerinden denemeler yapılmıştır. Bu sayının giderek yükselmesi olumlu sonuçlar verirken modelin hafıza kullanımını giderek arttırmaktadır.

#### 4.2.1. Performans Ölçümü

Modellerinin performansını değerlendirmek için temel bir değerlendirme metriği olarak hit rasyosu (HR) kullanılmaktadır. Validasyon setindeki bir ürünün önerilen ilk  $k$  içinde yer alması durumunda, bir "hit" olarak sayılır [42]. Böylece,  $HR(k)$  şu şekilde hesaplanır:

$$HR(k) = \text{toplam hit} / \text{validasyon seti büyüklüğü}$$

Diğer bir popüler metrik olarak kabul edilen, üst sıralarda elde edilen sonuçlara daha fazla önem veren marjinal oransal fayda ile başarılı sonuçları art arda düşük sıralara yerleştiren normalize edilmiş, indirilmiş kümülatif kazanç (NDCG) [42], bu çalışma için adapte edilmiştir:

$$NDCG(k) = Z_k \sum_{n=1}^k (2^{r_i} - 1) / \log_2(i + 1)$$

$Z_k$ , en iyi sonucun 1 olmasını garanti etmek üzere bir normalize edici,  $r_i$ ,  $i$ 'nci pozisyonundaki ürün için derecelendirilmiş isabet oranıdır. Eğer önerilen bir ürün validasyon setinde ise  $r_i$  1 değerini, değilse 0 değerini almaktadır.

Metrikler, satın alınan ürün numarasını doğru etiket olarak kabul etmek suretiyle hesaplanır. Gerçek etiketler, son satın alınan ürün numarasıdır. Ayrıca, model çıktısındaki sıralanmış olasılıklardan en üst  $k=20$  olasılığa karşılık gelen ürünlerden öneri listesi oluşturulur.

#### 4.3. Sonuçlar

Çapraz validasyon ortalama sonuçları, en iyi hiperparametre birleşimleri itibariyle Çizelge 6'da

özetlendiği gibidir. Bu çalışmadan elde edilen sonuçlar aynı yaklaşımla modellenmiş daha önceki çalışmayla [32], bu çalışmada elde edilmiş en iyi iki sonuçla kıyaslanmıştır.

**Çizelge 6.** MCCV ortalama değerleri itibariyle performans metrikleri (en iyi değerler koyu olarak gösterilmiştir)

	Bu çalışma	Kalkan ve Şahin (2022) [32]	
	Transformer	Model 3	Model 4
HR(20)	<b>0,3204</b>	0,3169	0,3184
NDCG(20)	<b>0,2606</b>	0,2332	0,2330

**Çizelge 7.** Test seti üzerinde en iyi iterasyondaki performans metrikleri (en iyi değerler koyu olarak gösterilmiştir)

	Bu çalışma	Kalkan ve Şahin (2022)[32]	
	Transformer	Model 3	Model 4
HR(20)	<b>0,3109</b>	0,3007	0,2944
NDCG(20)	<b>0,2332</b>	0,2123	0,2150

Test seti, modelleme sürecinin başında ayırdığımız veri kümesinin yaklaşık %10'u kadarlık kısmıdır. Test seti modelin önceden görmediği gözlemlerden oluştuğu için, bu sonuçlar modelin genelleme performansının bir göstergesi olarak değerlendirilebilir. Modeller, çapraz validasyon bölümünde elde edilen minimum kayıp değeri üreten birleşimleri ile test seti üzerinde çalıştırılmışlardır. Her modelin performansı test setinde 100 iterasyon boyunca takip edilmiş ve en iyi sonucu üreten iterasyon değerleri Çizelge 7'de özetlenmiştir.

Bu sonuçlara göre transformer modeli daha önce önerilmiş özyinelemeli sinir ağları ile geliştirilmiş öneri sistemlerden daha iyi sonuçlar üretmiştir. Özellikle NDCG(20) metriği açısından oldukça belirgin bir iyileşme gözlenmiştir.

## 5. BULGULAR

Bu çalışmada, mevcut müşteri portföyüyle ilerlemenin nasıl mümkün olabileceği şeklinde özetlenebilen bir iş problemi ortaya atılmıştır. Çapraz satış açısından, analitik tahmin araçlarının etkili çözümler olarak dikkate alınması

önerilmektedir. Doğal dil işleme alanında önemli bir araç olarak son zamanlarda popüler olan transformer ile RS uygulamaları bu alanda halen yeni sayılabilir. Bununla beraber, bu çalışmada elde edilen sonuçlara göre, daha önce önerilmiş yöntemler üzerinde önerilen değişiklikler ile açıkça bir iyileştirme sağlanabileceği gösterilmiştir.

CRM uygulamaları açısından müşterinin satın alabileceği bir sonraki ürünün ne olabileceğini tahmin etmek ve karar aşamasında bu ürünleri müşteriye önermek, genellikle zor bir görevdir. Bu çalışmadan elde edilebilecek bir diğer bulgu ise, RS uygulaması kullanmanın daha derin bir kataloğa sahip bir çevrimiçi perakendeci için karlı görüldüğüdür. Örneğin, bu çalışmada önerilen model, test setinin 2.636 müşterisinden, önceki modellere göre 50 daha fazla pazarlanabilir müşteri tahmini oluşturur. Bir öneri listesi, her müşterinin bir sonraki satın alacağı en muhtemel ürünleri bulmak amacıyla kullanılabilir ve pazarlanabilir ürün tahminleri belirli bir ürünü filtreleyerek kolayca pazarlanabilir müşterilere dönüştürülebilir. Benzer RS uygulamaları gibi, transformer kullanımı da ürünlerin ve müşterilerin seyrekliği gibi daha önce sözü edilen bazı sınırlamaları doğrudan aşmamaktadır. Bu modellerin bu tarz bir seyreklikle iyi çalışmadığı gözlenmiştir. Bu nedenle, yeterli satın alma geçmişi olmayan müşterilerin bu sisteme dahil edilebilmesi için sezgisel yöntemlerin de dikkate alınmasını gerektirir.

Bu çalışmada kullanılan transformer temelli model 262 milyonun üzerinde parametrenin optimum değerini araştırmıştır. Verinin non-lineer fonksiyonlardan yayıldığı da dikkate alındığında yüksek sayılabilecek bir hesaplama maliyetinin ortaya çıkışı da bir kısıtlılık unsuru olarak dikkate alınmalıdır.

Transformer, kendi başına bu kapsamda oldukça değerli bir araçtır. Bunun yanı sıra, gelecek çalışma alanları açısından daha fazla doğal dil işleme araçlarının bu alana adapte edilmesi dikkate değer bir çaba olarak görülmektedir. Gelecek çalışmalarda, daha önce önerilmiş geniş ve derin ağ [33] mimarisini kullanan modellerin de bir

kıyaslama unsuru olarak dahil edildiği, daha geniş kapsamlı uygulamalar düşünülmektedir.

## 6. KAYNAKLAR

1. Akçura, M.T., Srinivasan, K., 2005. Customer Intimacy and Crossselling Strategy. *Manage Sci*, 51(6), 1007-1012.
2. Kamakura, W.A., Wedel, M., Rosa, F., Mazzon, J.A., 2003. Crossselling Through Database Marketing: A Mixed Data Factor Analyzer for Data Augmentation and Prediction. *Int J Res Mark*, 20, 45-65.
3. Reinartz, W.J., Kumar, V., 2003. The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *J Mark*, 67, 77-99.
4. Kamakura, W.A., 2008. Cross-Selling: Offering the Right Product to the Right Customer at the Right Time. *J Relationship Market*, 6(3-4), 41-58.
5. Ansel, J., Archibald, T., 2007. Identifying Cross-Selling Opportunities, Using Lifestyle Segmentation and Survival Analysis. *Mark Intell Plan*, 25(4), 394-410.
6. Knott, A., Hayes, A., Neslin, S.A., 2002. Next-Product-to-Buy Models for Cross-Selling Applications. *J Interact Mark*, 16(3), 59-75.
7. Bogaert, M., Lootens, J., Van den Poel, D., Ballings, M., 2019. Evaluating Multi-Label Classifiers and Recommender Systems in the Financial Service Sector. *Eur J Oper Res*, 279(2), 620-634.
8. Zhang, L., Priestley, J., De Maio, J., Ni, S., Tian, X., 2021. Measuring Customer Similarity and Identifying Cross-Selling Products by Community Detection. *Big Data*, 9(2), 132-143.
9. Lü, L., Medo, M., Yeung, C., Zhang, Y., Zhang, Z., Zhou, T., 2012. Recommender Systems. *Phys Rep*, 519, 1-49.
10. Hu, Y., Koren, Y., Volinsky, C., 2008. Collaborative Filtering for Implicit Feedback Datasets. *Eighth IEEE International Conference on Data Mining*, 263-272.
11. Geuens, S., Coussement, K., De Bock, K.W., 2017. A Framework for Configuring Collaborative Filtering-Based Recommendations Derived from Purchase Data. *Eur J Oper Res*, 265(1), 208-218.

12. Verstrepen, K., Bhaduriy, K., Cule, B., Goethals, B., 2017. Collaborative Filtering for Binary Positiveonly Data ACM SIGKDD Explor Newsllett, 19(1), 1-21.
13. Adomavicius, G., Tuzhilin, A., 2005. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Trans Knowl Data Eng, 17(6), 734-749.
14. Zhang, S, Yao, L., 2019. Deep Learning Based Recommender System: A Survey and New Perspectives. ACM Comput Surv, 52(1), 1-38.
15. Devoght, R., Bersini H., 2017. Collaborative Filtering with Recurrent Neural Networks, 1-9.
16. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N. 2017. Attention is All You Need Adv. Neural Inf. Process. Syst. Nips, 5999-6009.
17. Thuring, F., Nielsen, J.P., Guillen, M., Bolance, C., 2012. Selecting Prospects for Cross-Selling Financial Products Using Multivariate Credibility. Expert Syst Appl, 39, 8809-8816.
18. Martinez, A., Schmuck, C., Pereverzyev, S., Pirker, C., Haltmeier, M., 2020. A Machine Learning Framework for Customer Purchase Prediction in the Non-Contractual Setting. Eur J Oper Res, 281(3), 588-596.
19. Chou, P., Chuang, H-C., Chou, Y-C., Liang, T-P., 2021. Predictive Analytics for Customer Repurchase: Interdisciplinary Integration of Buy Till You Die Modeling and Machine Learning. Eur J Oper Res, 296(2022), 635-651.
20. Tan, Y.K., Xu, X., Liu, Y., 2016. Improved Recurrent Neural Networks for Session-Based Recommendations. Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, 17-22.
21. Li, S., Kawale, J., Fu, Y., 2015. Deep Collaborative Filtering via Marginalized Denoising Auto-Encoder. Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM '15), 811-820.
22. Lee, D.D., Seung, H.S., 2000. Algorithms for Non-Negative Matrix Factorization. Proceedings of the 13th International Conference on Neural Information Processing Systems, 535-541.
23. Salakhutdinov, R., Mnih, A., 2008. Bayesian Probabilistic Matrix Factorization Using Markov Chain Monte Carlo. Proceedings of the 25th International Conference on Machine Learning, 880-887.
24. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T., 2017. Neural Collaborative Filtering. Proceedings of the 26th International Conference on World Wide Web, 173-182.
25. Zhang, S., Yao, L., 2019. Deep Learning Based Recommender System: A Survey and New Perspectives. ACM Comput Surv, 52(1), 1-38.
26. Salakhutdinov, R., Mnih, A., Hinton, G., 2007. Restricted Boltzmann Machines for Collaborative Filtering. Proceedings of the 24th International Conference on Machine Learning, 791-798.
27. Donkers, T., Benedikt, L., Ziegler, J., 2017. Sequential User-Based Recurrent Neural Network Recommendations. Proceedings of RecSys, 17, 27-31.
28. Hidasi, B., Karatzoglou, A., 2018. Recurrent Neural Networks with Top-k Gains for Session-Based Recommendations. The 27th ACM International Conference on Information and Knowledge Management (CIKM'18), 843-852.
29. Hidasi, B., Karatzoglou, A., Baltrunas, L., Tikk, D., 2016. Session-Based Recommendations with Recurrent Neural Networks, 4th International Conference on Learning Representations (ICLR), 1-10.
30. Wu, S., Ren, W., Yu, C., Chen, G., Zhang, D., Zu, J., 2016. Personal Recommendation Using Deep Recurrent Neural Networks in NetEase. IEEE 32nd International Conference on Data Engineering (ICDE), 1218-1229.
31. Smirnova, E., Vasile, F., 2017. Contextual Sequence Modeling for Recommendation with Recurrent Neural Networks. Proceedings of the 2nd Workshop on Deep Learning for Recommender Systems, 2-9.
32. Kalkan, İ.E., Şahin, C., 2022. Evaluating Cross-Selling Opportunities with Recurrent Neural Networks on Retail Marketing. Neural Computing and Applications, 35(8), 6247-6263.
33. Cheng, H.T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., Ispir, M., 2016. Wide and Deep Learning for Recommender Systems.

- Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, 7-10.
34. Chen, Q., Zhao, H., Li, W., Huang, P., Ou, W., 2019. Behavior Sequence Transformer for e-commerce Recommendation in Alibaba. Proceedings of the 1st International Workshop on Deep Learning Practice for High-Dimensional Sparse Data, 1-4.
  35. Li, W., Qi, F., Tang, M., Yu, Z., 2020. Bidirectional LSTM with Self-Attention Mechanism and Multi-Channel Features for Sentiment Classification. Neurocomputing, 387, 63-77.
  36. Katrompas, A., Metsis, V., 2022. Enhancing LSTM Models with Self-Attention and Stateful Training. Intell Syst Appl, 217-235.
  37. Keras. <https://keras.io/>, Erişim Tarihi: 06.09.2022
  38. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., Zheng, X., 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Tensorflow. <https://tensorflow.org/>, Erişim Tarihi: 06.09.2022.
  39. Kaggle. Your Machine Learning and Data Science Community. <https://www.kaggle.com/>, Erişim Tarihi: 06.09.2022.
  40. Pakistan's Largest E-Commerce Dataset, Version 2, <https://www.kaggle.com/datasets/zusmani/pakistans-largest-e-commerce-dataset/> Erişim Tarihi: 06.9.2022.
  41. Xu, Q-S., Liang, Y-Z., 2001. Monte Carlo Cross Validation. Chemom Intell Lab Syst, 56(1), 1-11.
  42. He, X., Chen, T., Kan, M.Y., Chen, X., 2015. TriRank: Review-Aware Explainable Recommendation by Modeling Aspects. CIKM'15: Proceedings of the 24th ACM International Conference on Information and Knowledge Management, 1661-1670.