



Kıyafet Öneri Sistemi için Giyim Metaverilerine dayalı Temsil Öğrenimi

Ahmet Dundar^{1*}, Arzu Gorgulu Kakisim²

¹ İstanbul Ticaret Üniversitesi, Fen Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye (ORCID: 0000-0003-1498-0946)

² İstanbul Ticaret Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye (ORCID: 0000-0001-6169-3486)

(International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) 2021 – 21-23 October 2021)

(DOI: 10.31590/ejosat.1008736)

ATIF/REFERENCE: Dundar, A. & Gorgulu Kalkisim, A. (2021). Kıyafet Öneri Sistemi için Giyim Metaverilerine dayalı Temsil Öğrenimi. *Avrupa Bilim ve Teknoloji Dergisi*, (29), 105-110.

Öz

Son zamanlarda, ürün ve hizmetleri tüketiciler ile buluşturan e-ticaret platformları üzerinden çevrimiçi tüketici alışverişi giyim, ayakkabı, makyaj, ev eşyası gibi birçok kategoride önemli ölçüde artmıştır. Çok sayıda seçeneğin mevcut olduğu çevrimiçi alışveriş ortamlarında, ürün bulma ve seçme işlem maliyetlerini azaltmak ve verimli bir şekilde ilgili ürünleri kullanıcılara iletmek önemli bir problem haline gelmiştir. Alışveriş şirketleri, müşterilerin satın alma geçmişlerine ve davranışlarına göre müşterilere benzer ürün seçenekleri sunarak gelirlerini artırmak ve müşterileri memnun etmek için öneri sistemleri kullanmaktadır. Bu nedenle, müşterilerin alışveriş davranışlarına göre en uygun, kişiselleştirilmiş ve tercih edilen tarz zevklerini yansıtan önerilerde bulunan akıllı öneri sistemler geliştirilmektedir. Son zamanlarda, ürünlerin metaverilerinin analizi ile ürünler arasındaki benzerliği ve uyumluluğu modelleyen ve bu sayede müşterilerin ürün bulma ve seçme işlemlerindeki memnuniyetini artırmaya çalışan yeni yöntemler sunulmaktadır. Bu çalışmada, rastgele yürüyüş ve Skipgram yöntemleri kullanarak ürünlere ait malzeme, desen, renk ve stil gibi metaveriler üzerinden ürünler arasındaki benzerliği modelleyen yeni bir yaklaşım önerilmektedir. Önerilen yöntem, ürünlere ait özellikler arasındaki üst düzey korrelasyonları keşfederek, ürünler arasındaki yakınlığın yansıtıldığı düşük boyutlu vektörel temsiller öğrenir. Bu amaçla, ürünler her bir düğümün ürünleri ve her bir kenarın ürünler arasındaki ilişkileri temsil ettiği ağırlıklı bir graf yapısına dönüştürülmüştür. Graf üzerinden öğrenilen bu temsiller sayesinde, müşterilerin alışveriş davranışına göre onlara en uygun önerilerin ve kombin tamamlayıcı ürünlerin sunulmasını sağlayan bir öneri sistemi sunulmaktadır.

Anahtar Kelimeler: Kıyafet Öneri Sistemi, Ağırlıklı Rastgele Yürüyüş, Kelime Gömme, Giyim Temsil Öğrenimi.

Representation Learning based on Clothing MetaData for Outfit Recommendation System

Abstract

Recently, online consumer shopping through e-commerce platforms that bring products and services to consumers has increased significantly in many categories such as clothing, shoes, make-up, and household goods. In online shopping environments, where there are many options, it has become an important problem to reduce the transaction costs of finding and selecting products and to recommend the relevant products to the users. Online shopping companies use recommendation systems to increase revenue and satisfy customers by offering similar product options based on their purchasing history and behavior. For this reason, intelligent recommendation systems are developed that make recommendations reflecting the most appropriate, personalized and preferred style tastes according to the shopping behaviors of the customers. Recently, researchers focus on developing new methods that model the similarity and compatibility between products through the analysis of products's metadata, thereby trying to increase customer satisfaction in finding and choosing products. In this study, a new approach is proposed that models the similarity between products through metadata such as material, pattern, color and style using random walk and Skipgram methods. The proposed method explores the high-level correlations between the properties of the products and learns low-dimensional vector representations that reflect the affinity between the products. For this purpose, the products are transformed into a weighted graph structure where each node represents the products and each edge represents the relationships between the products. Thanks to these representations learned from outfit graph, a recommendation system framework is presented that allows the most appropriate recommendations to the customers and according to their shopping behavior, and recommends complementary items to the query outfits.

Keywords: Outfit Recommendation System, Weighted Random Walks, Word Embedding, Outfit Representation Learning.

* Sorumlu Yazar: İstanbul Ticaret Üniversitesi, Fen Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye, ORCID: 0000-0003-1498-9460, ahmetdundar47@hotmail.com

1. Giriş

E-ticaret ürün içeriğinin ve çeşitliliğinin artmasıyla kullanıcıların seçimini kolaylaştıran ve onlara en uygun şekilde kişiselleştirilmiş seçenekler sunan akıllı öneri sistemlerinin önemi gün geçtikçe artmaktadır. Öneri sistemleri, kullanıcıların ilgi alanlarını tahmin etmeyi ve onlar için oldukça ilginç olabilecek ürün öğelerini önermeyi amaçlar. Öneri sistemleri, müşterilerin satın alma geçmişlerine ve davranışlarına göre müşterilere uygun kişiselleştirilmiş ürün seçenekleri sunarak, ürün yelpazesini daraltmayı ve ürün seçme süreçlerini daha verimli bir hale getirmeyi amaçlar. Araştırmalara göre [1], çevrimiçi satın almaların %43'ü kişiselleştirilmiş önerilerden oluşmaktadır. Şirketlerin %94'ü kişiselleştirilmiş ürün sunan sistemlerin kazançlarını arttırmak için kritik bir strateji olarak algılanmaktadır.

Kişiselleştirilmiş öneri sistemleri, çevrimiçi perakendecilerin mevcut ve gelecekteki pazarlama başarısı için kullanılan en güçlü makine öğrenimi sistemleri arasındadırlar. Son yıllarda, gelişmiş yapay sinir ağı modelleri kullanarak, müşteri ve ürünler arasındaki yakınlığı modelleyen yeni yöntemler önerilmektedir. Literatürdeki yöntemler kullandıkları veri açısından ikiye ayrılmaktadır: ürünlere ait görsel veriyi kullanan yöntemler ve ürünlere ait içeriği ve bilgiyi temsil eden metaverileri kullanan yöntemler. Son zamanlarda ise her iki veriyi eş zamanlı olarak kullanan yeni modeller önerilmektedir. Fakat aynı zamanda özellikle moda endüstrisindeki büyüme ile birlikte, genel tüketim ürünlerinden farklı olarak giysi ve ayakkabı gibi ürünlerdeki çeşitlilikte günden güne artmaktadır. Ürünler desen yapısına, malzeme yapısına, rengine ve tarzına göre farklılaşmakta ve bu nedenle ürünler ve kullanıcı tercihleri arasındaki ilişkileri analiz etme daha zor bir problem haline gelmektedir. Son birkaç yılda, giyim ürünleri arasındaki yüksek seviyedeki ilişkileri ortaya çıkarmayı hedefleyen ve metaveriler üzerinde yerleştirme yaklaşımları kullanarak temsil öğrenimi gerçekleştiren yeni yöntemler önerilmektedir. Bu yöntemler doğal dil işlemede sıklıkla kullanılan bir belgedeki bir kelimenin bağlamını, anlamsal ve sözdizimsel benzerliğini, diğer kelimelerle ilişkisini ortaya çıkaran ve her kelimeyi ya da belgeyi düşük boyutlu bir vektörle temsil etmeyi amaçlayan Word2Vec [2] ve Doc2Vec [3] gibi gömme yaklaşımlarını temel almaktadır. Word2Vec'ten esinlenen Item2Vec [4] yöntemi yaklaşımı işbirlikçi filtreleme senaryolarındaki öğelerin gizli temsillerini öğrenmek için bir sinirsel öge gömme tabanlı yaklaşım kullanmaktadır. Item2Vec'in temel amacı, öğelerin benzerliğini ortaya çıkarmak amacıyla kullanıcının bilgisi olmadığında bile öge-ürün ilişkilerini öğrenebilmektir. Benzer bir başka model, satın alma dizilerinden ürün gömmelerini öğrenen ve gömme uzayındaki en benzer ürünleri temel alarak öneri gerçekleştiren Prod2vec [5] yöntemidir. Prod2vec, bir sinirsel dil modelini kullanarak ürünleri gerçek değerli, düşük boyutlu vektör uzayına yerleştirir. Sonuç olarak, benzer

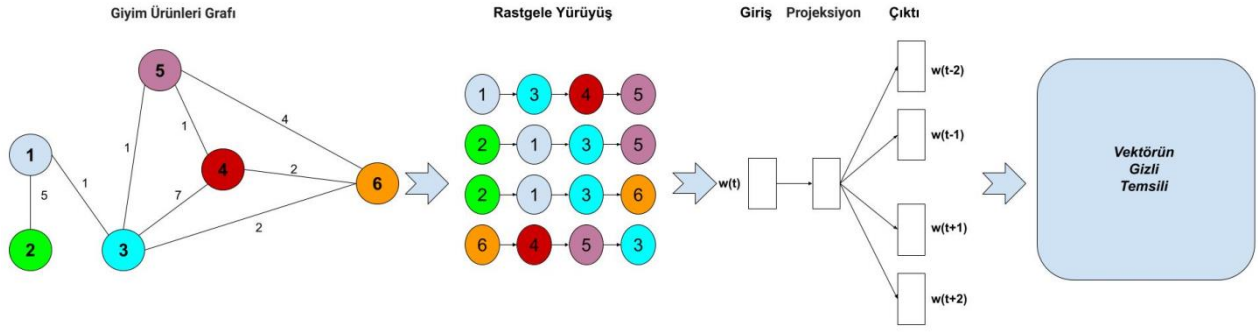
bağlamlara sahip ürünlerin vektörleri gömme uzayında birbirine yakın konumlanmış olmaktadır. Bir sonraki satın alınacak ürünle ilgili anlamlı önerilerde bulunabilmek için ürün vektörlerini kümeleyerek gömme alanında en yakın ürünleri kullanmak için kullanılır. Meta-Prod2Vec [6], öğelerin düşük boyutlu yerleştirmelerini hesaplarken yan bilgi olarak öğelerin nitelikleri ve kullanıcıların geçmiş etkileşimlerini içeren metaverileri kullanır. Outfit2vec [7] yöntemi ise hiyerarşik olarak oluşturulmuş moda kıyafetlerinin temsili vektörlerini öğrenmek için Doc2vec modelini temel almıştır. Giyim ürünlerinin metaverileri üzerinden ürünlere ait gömme vektörlerini elde etmektedirler. Fakat ürünleri tanımlayan metaveriler genellikle kısa metinlerden oluşur. Bu nedenle öznitelik uzayı oldukça seyreklik göstermektedir ve bahsedilen gömme yaklaşımları da kısa metinlerdeki bu seyreklik probleminden dolayı ürünler arasındaki yakınlığı modellemede istenen başarıyı göstermeyebilmektedir. Dahası, çoğu kelime gömme yöntemi, sahip oldukları kelimelerle (bilinmeyen veya sözlük dışı (OOV) kelimeler) karşılaştıklarında, bizi ideal olmaktan uzak vektörler üretmeye veya yeni gelen girdilerle [8] öğrenme sürecini yeniden başlatmaya zorlayabilir. Bu yöntemlerin diğer eksikliği, farklı ve yeni dillerle karşılaşıldığında ortaya çıkmaktadır [9]. Bu yöntemlerin yeni diller için yeni temsiller oluşturması gerekir çünkü bir dil için oluşturulmuş bir model diğer dillere ölçeklendirmek mümkün değildir.

Bu çalışmada, yukarıda bahsedilen problemleri adresleyen, ürünlere ait malzeme, desen, renk ve stil gibi metaveriler üzerinden rastgele yürüyüş [10] ve Skipgram [2] yöntemleri kullanarak ürünler arasındaki benzerliği modelleyen yeni bir yaklaşım önerilmektedir. Önerilen yöntem, her bir düğümün ürünleri ve her bir ağırlıklı kenarın ürünler arasındaki ortak özelliği ve sayısını temsil ettiği bir graf yapısına dönüştürülmüştür. Deepwalk [11] mimarisinden esinlenerek, graf gömme yaklaşımı ile oluşturulan graf üzerinden her bir ürüne ait düşük boyutlu vektörler temsiller elde edilmektedir. Önerilen sistem hem kıyafet hem de kıyafet tamamlayıcı ürün önerisi için ürünler arasındaki yakınlığın yansıtıldığı bu vektörel temsilleri kullanmaktadır.

Bu makalenin geri kalanı aşağıdaki gibi düzenlenmiştir. Bölüm 2'de önerilen yöntem anlatılmaktadır. Uygulama detayları ve deneysel sonuçlar Bölüm 3'te verilmektedir. Son olarak Bölüm 4'de makaleye ait sonuçlar özetlenmekte ve gelecek çalışmalardan bahsedilmektedir.

2. Önerilen Yöntem

Şekil 1'de önerilen sistemin mimarisi gösterilmektedir. Önerilen sistem dört aşamayı içerir. Bunlar, sırasıyla metaveriler için ön işleme süreci, ürün grafi oluşturma, rastgele yürüyüşler ile graftan sınırlı uzunlukta yollar elde etme, yolları girdi olarak alan bir gömme yöntemi ile temsil öğrenme ve temsiller üzerinden kıyafet ürünü önerme süreçleridir.



Şekil 1. Önerilen sistem mimarisi

Ürün verilerinin yer aldığı yapılarda, katalog görselleri ile birlikte sağlanan ve mevcut öğelerin kısaca açıklamalarını içeren metaveriler bulunmaktadır. Metaveriler, ürünlere ait malzeme, desen, renk, marka adları ve etiketler olmak üzere ürün hakkında bilgi veren birçok detayı içerebilmektedir. Ürün açıklamaları metinsel veriler olduğu için metinleri analiz edilebilir bir forma getirmek amacı ile bir ön işleme süreci uygulanmaktadır. Bu doğrultuda, çalışmamızın ön işleme sürecinde, noktalama işaretleri, nümerik değerler, etkisiz terimler, özel karakterler temizlenmekte ve tüm harfler küçük harfe dönüştürülüp var olan bazı yazım hataları düzeltilmektedir. Daha sonra metaveri içerisindeki her bir kelime kelimenin kökü ile temsil edilmekte ve böylece aynı anlama gelen kelimeler öznitelik uzayında tek bir kelime ile temsil edilmektedir. Bazı ürünlerin metaverilerinde ürünlere ait renk bilgisi bulunmadığından, önerilen yöntem ürünlere ait katalog görsellerinden renk bilgisi elde ederek, her

bir ürünün en azından tek bir renk ile temsil edilmesini sağlamaktadır. Ürünlerin rengini elde etmek için katalog görseli içinde bulunan beyaz renkler çıkarılıp resim saydam hale getirilmekte ve sonrasında resimde ağırlıklı olarak kullanılan renk kodu seçilmektedir. Renk sabitleri dosyasında en yakın renk kodu bulunarak renk adı ataması yapılmaktadır [12]. Verinin hazırlanması ve temizlenmesi sürecinden sonra, önerilen yöntem, her bir ürünün sahip olduğu öznitelikleri temsil eden bir $F = \{f_1, f_2, \dots, f_n\}$ öznitelik kümesi oluşturmaktadır. Burada, f_i i. ürüne ait öznitelikleri içeren bir alt öznitelik kümesidir. Tablo 1'de örnek her bir ürüne ait metaveriler sağlanmaktadır. Örneğin; ilk ürün için ürünü tanımlayan ürün açıklaması, ürünün kategorisi ile ilgili açıklama, ürünün kategorisi, ürünün içerisinde bulunduğu önceden hazırlanmış kombin grubu ve renk bilgisi ile birlikte oluşan öznitelik kümesinde dokuz farklı öznitelik elde edilmiştir.

Tablo 1. Örnek ürünler ve ürünlere ait metaveriler

Ürün No	Ürün Açıklaması	Alt Kategori Ürün Açıklaması	Kategori	Kombin	Renk Bilgisi	Alt öznitelik Kümesi
1	mock neck embroidery suede sweatshirt	casual	k1	kombin 1	gray	{ mock, neck, embroidery, suede, sweatshirt, casual, k1, kombin 1, gray }
2	mock neck embroidery suede brown sweatshirt	camel meanswear	k2	kombin 2	brown	{ mock, neck, embroidery, suede, brown, sweatshirt, camel, meanswear, k2, kombin 2 }
3	citizens humanity high rise rocket jean	casual	k3	kombin 1	blue	{ citizen, humanity, high, rise, rocket, jean, casual, k3, kombin 1, blue }
4	citizens humanity high rise jeans	skinny denim	k3	kombin 3	blue	{ citizen, humanity, high, rise, jean, skinny, denim, k3, kombin 3, blue }
5	joy denim jacket	chanel fashion	k4	kombin 4	blue	{ joy, denim, jacket, chanel, k4, kombin 4, blue }
6	pre-owned chanel denim jacket jean	miranda kerr	k4	kombin 5	blue	{ preowned, chanel, denim, jacket, jean, k4, kombin 5, blue }

Ön işleme adımından sonra, önerilen yöntem, ürünler arasındaki yakınlığın temsil edildiği bir ürün grafı yapıdır. Graflar, düğüm adı verilen varlıklar arasındaki ilişkileri temsil etmede kullanılan etkili veri yapılarıdır. Graf oluşturma aşamasında, her bir ürün bir düğüm olarak temsil edilmekte ve grafın kenarları ise ürünler arasındaki ortak özellikler tarafından belirlenmektedir. Matematiksel olarak ifade etmek gerekirse, bir ürün grafı $V = \{v_1, v_2, \dots, v_n\}$ düğümler kümesi, $E = (V \times V)$ kenarlar kümesi olmak üzere $G = (V, E)$ grafı olarak tanımlanmaktadır. Düğüm $v_i \in G$ verisetindeki i . ürünü temsil etmektedir. Eğer iki ürünün ortak bir niteliği var ise bu iki düğüm arasında bir kenar var demektir. Graftaki kenarların ağırlığı ise iki düğümün ortak nitelik sayısı ile belirlenir. Bu durumda, verisetinde bulunan i . ürün ve j . ürün

eğer k adet ortak niteliğe sahip ise grafta karşılıklı geldiği v_i ve v_j düğümleri arasındaki e_{ij} kenarın ağırlığı k olacaktır. Tablo 1'deki alt öznitelik kümeleri dikkate alındığında, 1. ürün ve 2. ürünün ortak öznitelik sayısı $|f_1 \cap f_2| = 5$ olduğu görülmektedir. Bu durumda, v_1 ve v_2 düğümleri arasındaki kenar ağırlığı $e_{12} = 5$ olacaktır. Şekil 1'de Tablo 1'de verilen ürünlerin ortak öznitelik sayısı hesaplanarak elde edilen graf yapısı gösterilmektedir.

Graf oluşturma aşamasından sonra, önerilen yöntem, düğümler arasındaki yakınlığını modellemek için oluşturulan graf üzerinde kısa rastgele yürüyüşler gerçekleştirir. Rastgele yürüyüşler, yollar ve sık alt yapılar gibi düğümler arasında daha karmaşık ilişkiler yansıtan daha yüksek dereceli yakınlığı

keşfetmemizi sağlamaktadır. Bir t anında düğüm v_i başlangıç noktası olarak seçildiğinde, $t + 1$ anı için v_i ile bağlantılı düğümlerden biri rastgele seçilir. Bu rastgele yürüyüş, önceden belirlenen adım sayısı boyunca devam eder. Graftaki her düğüm için, L uzunluğunda M sayıda rastgele yürüyüşler elde edilir. Böylece her bir düğüm sonlu sayıda adımda gidebileceği düğümleri içeren birden fazla yol ile temsil edilir. Burada, herhangi bir düğümden elde edilmiş rastgele yürüyüş y olarak, graftan elde edilmiş tüm rastgele yürüyüşleri ifade eden küme ise Y olarak tanımlanır. Örneğin, Şekil 1'de verilen graf yapısı üzerinden örneklendirilmiş 4 uzunluğunda ve 1. düğümden başlayan rastgele bir yürüyüş şu şekilde olabilir: $y = (1, 3, 4, 5)$. Fakat oluşturulan ürün grafi ağırlıklı bir graf olduğundan rastgele yürüyüş yönteminin ağırlıklı versiyonu gerçekleştirilir. Ağırlıklı rastgele yürüyüşte, mevcut düğüme bağlı düğümler, kenar olasılığı dikkate alınarak seçilir. Kenar olasılığı, o kenarla ilişkili kenar ağırlığı ile orantılıdır. Özetle, kenar ağırlığı büyük bağlantıların seçilme olasılığı daha yüksektir.

Algoritma 1: Önerilen yöntemin algoritması

Giriş: $G(V, E)$

Pencere boyutu w
Temsil boyutu d
Yürüyüş sayısı M
Yürüyüş uzunluğu L

Çıktı: Ürün temsillerinin matrisi $\Phi \in \mathbb{R}^{|V| \times d}$

```
1:  $G(V, E) \leftarrow$  Ürün metaverilerinden ürün grafi oluştur
2: for  $i = 0$  to  $n$  do
3:    $y = \text{RasgeleYürüyüş}(G, v_i, L, M)$ 
4:    $Y = Y \cup y$ 
5: end for
6:  $\Phi \leftarrow \text{SkipGram}(R, d, w)$ 
7:  $R = \text{EnYakınKomşu}([v_i, v_j], k=5)$ 
```

Önerilen yöntem, her bir ürün Φ_i vektörü ile temsil edildiği, ürünler arasındaki yakınlığı yansıtan bir Φ ürün temsilleri matrisini kullanarak, ürün önerisi gerçekleştirmektedir. Her bir ürün için sorgulama yapılarak, ürüne en yakın k adet komşu yakınlık dereceleri ile birlikte öğrenilebilmektedir. Benzer şekilde, birden fazla ürün ya da bir kombin aynı anda sorgulanarak, bu kombine en yakın tamamlayıcı ürünler elde edilebilmektedir. En yakın komşular için vektörler arasındaki mesafe Kosinüs benzerliği kullanarak elde edilir.

3. Araştırma Sonuçları ve Tartışma

3.1. Veri Seti

Deneyler için Polyvore [13] veri seti kullanılmıştır. Polyvore kullanıcıların kıyafet verilerini oluşturup yükledikleri popüler bir moda sitesidir. Bu moda kıyafetleri, moda öğelerinin resimleri, açıklamaları ve kıyafetin beğeni sayısı gibi zengin çok modlu bilgiler içerir. Polyvore veri kümesi 164.379 öge içerir. Ayrıca, 21.889 adet giyim ürünlerine ait kombin bulunmaktadır. Bu çalışmada, Polyvore veri setinde var olan giyim kategorisi dışında kalan sandalye, masa, çiçek gibi ürünler veri setinden çıkarılmıştır. Veri setimiz bu nedenle 17316 tane kombin ve

Düğüm dizilerini içeren rastgele yürüyüşler elde edildikten sonra, önerilen yöntem, hem düğüme ait düşük boyutlu gömme vektörlerini Φ öğrenmek için SkipGram modelini uygular. SkipGram, bağlamsal bir pencere kullanarak bir cümlede görünen kelime çiftlerinin birlikte bulunma olasılıklarını maksimize etmeyi amaçlayan bir doğal dil işleme modelidir. Önerilen yöntemde, SkipGram modeli, her düğüm dizisini bir cümle olarak değerlendirilir ve girdi olarak alır. Her giriş düğümü, bir ikili vektör olan bir-sıcak kodlama ile 0 ve 1'lerden oluşan bir ikili vektör ile temsil edilir. SkipGram, girdi olarak bir hedef düğüm kullanır ve w 'nin bağlam penceresi boyutu olduğu bir pencerede bağlam düğümlerinin $2w$ çok terimli dağılımlarını çıkarır. Başka bir deyişle, hedef düğümün solundaki ve sağındaki çevreleyen düğümleri tahmin etmek için bir hedef düğüm kullanılır. Skip-Gram çıktı olarak her bir ürün için d uzunluğu düşük boyutlu bir gömme vektörü üretir. Önerilen modele ait algoritmanın detayları Algoritma 1'de açıklanmaktadır.

85883 ürün içermektedir. Veri setinde 159 adet farklı kategori bulunmaktadır. Veri setindeki ürünlerin için renk tahsisi yapılmış ve 378 farklı renk ile temsil edilmişlerdir.

3.2. Deneysel Sonuçlar

Bu bölümde, önerilen yöntemin başarısını değerlendirmek için birkaç çevrimdışı deney gerçekleştirdik. Bu deneylerimizde, yürüyüş sayısı $M = 20$, yürüyüş uzunluğu $L = 20$, pencere boyutu $w = 20$, temsil boyutu $d = 128$ olarak ayarlanmıştır. Şekil 2'de sorgulanan ürüne ve bu ürüne en benzer olarak önerilen en yüksek puanlı ürünlere ait görseller sunulmuştur.

Dört farklı giyim ürünü için önerilen ürünler incelendiğinde, ürünler aynı kategoriden gelen benzer ürünler olduğu görülmektedir. Örneğin; sorgulanan ürün mavi bir kot pantolonu olduğu durumda ise önerilen ürünlerin yine aynı renk koduna sahip mavi kot pantolonları olduğu görülmektedir. Veri setinde çok farklı renkte kot pantolonları olduğu için önerilen ilk beş ürünün aynı kategoride ve aynı renk aralığında olması yöntemin ürünlerin arasındaki yakınlığı modelleme de anlamlı sonuçlar ürettiği gözlemlenmektedir.



Şekil 2. Sorgulanan ürünü ve önerilen yöntemin tavsiye ettiği en yakın benzerliğe sahip ilk beş ürüne ait görseller



Şekil 3. Önerilen yöntemin sorgulanan kombinler için önerdiği tamamlayıcı ürünlere ait görseller

Önerilen yöntem, kombin önerileri için de test edilmiş ve örnek sorgular üzerinden önerilen ürünler Şekil 3’de verilmiştir. İki örnek için iki farklı üst giyim ürünü kombinlenmiştir. Önerilen sistem ile bu kombini tamamlayıcı ürünler olarak bu iki ürüne ayakkabı, çanta ve takı ürünü önermiştir. Önerilen ürünler en bu iki ürüne en yakın uzaklıkta olan ürünlerdir.

Tablo 2. Yöntemlerin Sınıflandırma Başarısı

Yöntem	F-ölçüm (Mikro)	F-ölçüm (Makro)
Word2Vec + RF	%62.92	%41.23
Doc2Vec + RF	%60.74	%30.51
Önerilen Yöntem + RF	%76.72	%35.31

Önerilen yöntemin ürünler arasındaki yakınlığı modellemedeki başarısını analiz etmek için ürün sınıflandırma deneyi gerçekleştirilmiştir. Bu nedenle, önerilen yöntemin ürünlere ait düşük boyutlu temsilleri üretilirken kategori bilgisi ürünlerin öznitelik kümelerinden çıkarılmıştır. Yöntemin sınıflandırma başarısı için Rastgele Orman (RF) sınıflandırıcısı kullanılmıştır. Yöntem Word2Vec ve Doc2Vec yöntemleri ile karşılaştırılmıştır. Tablo 2’de yöntemlere ait F-ölçüm değerleri verilmektedir. Elde edilen sınıflandırma sonuçları incelendiğinde, önerilen yöntemin ürün kategorisi tespitinde Mikro F-ölçüm diğer gömme yönteminden daha iyi sonuçlar elde ettiği görülmektedir. Makro F-ölçüm değerine göre en iyi sonuç Word2Vec yöntemi ile edilmiş, fakat önerilen yöntem Doc2Vec yönteminden daha iyi bir sonuç vermiştir. Makro F-ölçüm değeri her sınıf için bağımsız olarak hesaplanır ve ardından ortalaması alınarak elde edilir. Fakat çok sınıflı bir sınıflandırma sürecinde, sınıflara ait örnek sayısı dengesizliği

olabileceğinden ortalama hesabında tüm sınıfların katkılarını tek tek ele aldığı için mikro F-ölçüm bize daha anlamlı bir sonuç sunmaktadır. Bu nedenle, önerilen yöntemin ürünler arasındaki yakınlığı modelleme de diğer iki yöntemden daha iyi olduğu sonucuna varılmaktadır.

4. Sonuç

Bu çalışmada, giyim ürünlerine ait düşük boyutlu temsilleri içeren gömme matrisi üzerinden sorgulanan ürün ve kombinler için öneriler üreten bir ürün öneri sistemi sunulmuştur. Önerilen sistem, giyim ürünlerine ait açıklayıcı metinlerden oluşan metaverilerini kullanır. Metaveriler kısa ve seyrek metinler içerdiği için seyreklik problemini ele almak için önerilen yöntem ürünleri bir graf yapısı ile temsil etmektedir. Önerilen yöntem, ürünlere ait özellikler arasındaki yakınlığın yansıtıldığı düşük boyutlu vektörel temsiller öğrenmeyi amaçlamaktadır. Yöntem, her bir ürün için birden fazla ürün önerilebilmekte, ayrıca her bir kombin için kombini tamamlayıcı ürünler önerilebilmektedir. Gelecekte, modelimizin performansını farklı ve daha büyük ölçekli veri setlerinde test etmeyi planlıyoruz.

Kaynakça

1. Orendorff A (2019) The state of the ecommerce fashion industry: statistics, trends and strategy. <https://www.shopify.com/enterprise/ecommerce-fashion-industry>. Erişim tarihi 13 Ocak 2020.
2. Mikolov, Tomas & Corrado, G.s & Chen, Kai & Dean, Jeffrey. (2013). Efficient Estimation of Word Representations in Vector Space. 1-12.
3. Li, J., Huang, G., Fan, C., Sun, Z., & Zhu, H. (2019). Key word extraction for short text via word2vec, doc2vec, and textrank. Turkish Journal of Electrical Engineering & Computer Sciences, 27(3), 1794-1805.
4. Barkan O, Koenigstein N (2016) Item2vec: neural item embedding for collaborative filtering. In: 2016 IEEE 26th international workshop on machine learning for signal processing (MLSP). IEEE, pp 1–6.
5. Grbovic M, Radosavljevic V, Djuric N, Bhamidipati N, Savla J, Bhagwan V, Sharp D (2015) E-commerce in your inbox: product recommendations at scale. In: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 1809–1818.
6. Vasile F, Smirnova E, Conneau A (2016) Meta-prod2vec: product embeddings using sideinformation for recommendation. In: Proceedings of the 10th ACM conference on recommender systems. ACM, pp 225–232.
7. Jaradat, Shatha & Dokoohaki, Nima & Matskin, Mihhail. (2020). Outfit2Vec: Incorporating Clothing Hierarchical MetaData into Outfits' Recommendation. 10.1007/978-3-030-55218-3_5.
8. Won MS, Lee JH.: Embedding for Out of Vocabulary Words Considering Contextual and Morphosyntactic Information. In 2018 International Conference on Fuzzy Theory and Its Applications (iFUZZY) 2018 Nov 14 (pp. 212-215). IEEE.
9. Al-Matham RN, Al-Khalifa HS. Synoextractor: a novel pipeline for Arabic synonym extraction using Word2Vec word embeddings. Complexity. Erişim tarihi 27 Şubat 2021.
10. Gori, Maria & Pucci, Augusto. (2006). Research Paper Recommender Systems: A Random-Walk Based Approach.

- Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence. 778-781. 10.1109/WI.2006.149.
11. Perozzi, Bryan & Al-Rfou, Rami & Skiena, Steven. (2014). DeepWalk: Online Learning of Social Representations. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 10.1145/2623330.2623732.
12. Nat Dunn, Python Color Constants Module, <https://www.webucator.com/article/python-color-constants-module>. Erişim tarihi 27 Temmuz 2021.
13. Han, Xintong & Wu, Zuxuan & Jiang, Yu-Gang & Davis, Larry. (2017). Learning Fashion Compatibility with Bidirectional LSTMs. 10.1145/3123266.3123394.
14. Chen W, Huang P, Xu J, Guo X, Guo C, Sun F, Li C, Pfadler A, Zhao H, Zhao B (2019) POG: personalized outfit generation for fashion recommendation at Alibaba iFashion. In: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 25 Jul 2019, pp 2662–2670.
15. Laenen K., Moens MF. (2020) Attention-Based Fusion for Outfit Recommendation. In: Dokoohaki N. (eds) Fashion Recommender Systems. Lecture Notes in Social Networks. Springer, Cham.