

Veri madenciliği algoritmalarına dayalı olarak sosyal medya üzerinden mobilya seçimindeki yönelimlerin belirlenmesi: Twitter örneği

Selman Karayılmazlar^a, Timuçin Bardak^{b,*}, Özkan Avcı^c, Kadir Kayahan^b,
Atakan Süha Karayılmazlar^d, Yıldız Çabuk^a, Rifat Kurt^a, Erol İmren^a

Özet: İnternet kullanımındaki artışa paralel olarak dünyanın farklı noktalarındaki insanlar farklı konulardaki duygu ve düşüncelerini sosyal medyalar aracılığıyla kolay bir şekilde aktarabilmektedirler. Bu sosyal medyalar içerisinde önemli bir yere sahip olan Twitter vasıtasıyla küresel ölçekte her gün çeşitli konularda milyonlarca mesaj yazılmakta ve okunmaktadır. Firmaların rekabet gücünü artırmak açısından tüketici davranışlarını anlamak önemli bir konu iken Twitter gibi büyük veri kaynakları, davranışların analiz edilebilme yöntemlerini çok yönlü ele almaktadır. Aynı zamanda gelişmiş ülkeler güç sahibi olmak için veri madenciliği projelerine önemli kaynaklar ayırmaktadır. Çalışmada veri madenciliği algoritmaları kullanılarak mobilya tercihindeki eğilimleri belirlemek için bir sosyal medya ağı olan Twitter’da yapılan ilgili paylaşımlar değerlendirmeye alınmıştır. Bu kapsamda Rapidminer ve doğal dil işleme yazılımları kullanılarak içinde mobilya geçen popüler tweetler Mayıs 2018-Şubat 2019 tarihleri arasında on ay boyunca toplanmış ve doğal dil işleme yazılımları sayesinde tweetlerin duygu durumları (pozitif ve negatif) belirlenmiştir. Daha sonra pozitif ve negatif tweetlerde geçen anahtar kelimelerin morfolojik analizleri gerçekleştirilmiştir. Son olarak veri madenciliğinde kullanılan karar ağacı ve birliktelik algoritmalarından faydalanarak anlamlı bilgiler elde edilmiştir. Karar ağacı algoritmasına göre pozitif veya negatif duyguların oluşumunda *itiraz*, *kampanya*, *keşfetmek* ve *fikir* gibi kelimelerin baskın olduğu belirlenmiştir. Birliktelik algoritması sonucunda ise en pozitif duyguları uyandıran kelimelerin *sipariş ile yapılmış*, *fırsat* ve *ağaç* gibi ifadeler olduğu tespit edilmiştir. Aynı algoritmada en negatif duyguları uyandıran kelimeler ise *kasvet*, *keyifsiz*, *rahatsız* ve *kumaş* olarak sıralanmıştır.

Anahtar kelimeler: Mobilya, Sosyal medya, Twitter, Veri, Duygu analizi

Determining the orientation in choosing furniture based on social media based on data mining algorithms: Twitter example

Abstract: In parallel with the increase in internet usage, people from different parts of the world can easily convey their thoughts and feelings on social issues through social media. Millions of messages are written and read every day on various topics on a global scale through Twitter, which has an important place in these social media. While it is important to understand consumer behaviors in order to increase the competitiveness of firms, big data sources such as Twitter have multi-faceted the methods of analyzing behaviors. At the same time, developed countries allocate significant resources to data mining projects in order to have power. The use of Twitter and data mining as an alternative data source to identify trends in furniture choice has been proposed. The popular tweets with furniture using the Rapidminer and natural language processing software were gathered for ten months between May 2018 and February 2019, and natural language processing software enabled us to determine the mood of the tweets (positive and negative). Morphological analysis of the keywords in positive and negative tweets was then performed. Finally, meaningful information was obtained by utilizing the decision tree and association algorithms used in data mining. According to the decision tree algorithm, the most dominant words in the formation of positive or negative emotions were the challenge, campaign, discover and idea. As a result of the syntax of association, the most positive emotions were made with the order of words that awaken the emotions, and the opportunity was found as wood. In the same algorithm, the words that awaken the most negative emotions were listed as gloom, seedy, uncomfortable and fabric.

Keywords: Furniture, Social media, Twitter, Data, Sentiment analysis

1. Giriş

Teknolojideki gelişmelerle birlikte internet, güncel yaşantımızın vazgeçilmez bir parçası haline gelmiş ve yaygın bir kullanım alanı bulmaya başlamıştır. Sadece

iletişim aracı olmaktan çıkan internet, yaşam şekillerini, çalışma koşullarını, alışveriş alışkanlıklarını, sosyal ilişkileri, eğitim ve sağlık sektörlerini çok farklı bir noktaya getiren ve şekillendiren bir duruma gelmiştir. Eskiden tek taraflı olarak okuduğumuz, dinlediğimiz, izlediğimiz

- ✉ ^a Orman Endüstri Mühendisliği Bölümü, Orman Fakültesi, Bartın Üniversitesi, Bartın, Türkiye
^b Malzeme ve Malzeme İşleme Teknolojileri Bölümü, Meslek Yüksek Okulu, Bartın Üniversitesi, Bartın, Türkiye
^c Pazarlama ve Reklamcılık Bölümü, Meslek Yüksek Okulu, Bartın Üniversitesi, Bartın, Türkiye
^d Peyzaj Mimarlığı ABD, Fen Bilimleri Enstitüsü, Bartın Üniversitesi, Bartın, Türkiye

@ * **Corresponding author** (İletişim yazarı): timucinb@bartin.edu.tr

✓ **Received** (Geliş tarihi): 23.08.2019, **Accepted** (Kabul tarihi): 12.12.2019



Citation (Atıf): Karayılmazlar, S., Bardak, T., Avcı, Ö., Kayahan, K., Karayılmazlar, A.S., Çabuk, Y., Kurt, R., İmren, E., 2019. Veri madenciliği algoritmalarına dayalı olarak sosyal medya üzerinden mobilya seçimindeki yönelimlerin belirlenmesi: Twitter örneği. Turkish Journal of Forestry, 20(4): 447-457. DOI: [10.18182/tjf.609967](https://doi.org/10.18182/tjf.609967)

iletişim araçları günümüzde internet ve sosyal medya dediğimiz Twitter, Facebook, LinkedIn vb. sosyal ağlar vasıtasıyla iletişimi karşılıklı ve daha yaygın hale getirmiştir. Böyle yaygın bir iletişim ağının olduğu alanda firma ve kurumlar da elbette sosyal medyayı kendi lehlerine kullanabilecek yolları keşfetme arayışı içine girmişlerdir. Bu yollardan biri de veri madenciliğidir. Veri madenciliği ile anlık olarak yüz milyonlarca kullanıcı tarafından üretilen büyük miktarlarda veriler, ayırt edici veya tanımlayıcı yeni bilgilere dönüştürülebilmekte ve tüketicilerin davranış ve eğilimleri belirlenebilmektedir. Bu kapsamda orman ürünleri sanayisi içerisinde önemli bir yere sahip olan mobilya sektöründe de tüketici davranışlarının değerlendirilmesi önem arz etmektedir. Mobilya üzerine sosyal ağlar üzerinden paylaşılan verilerin incelenmesi ile mobilya kullanıcılarının etkileşimleri anlamlı olarak değerlendirilebilmektedir. Küresel ölçekte her gün çeşitli konularda milyonlarca mesaj yazılan ve okunan sosyal medyalardan biri olan Twitter, söz konusu bu verilerin toplanması için önemli bir potansiyel oluşturmaktadır. Bu çalışma ile tüketicilerin duygu durumları temel alınarak mobilya sektörüne yönelik anlamlı bilgiler elde edilmesi ve böylece firmaların tüketici davranışlarını daha doğru tahmin etmelerine, pazarlama faaliyetlerini etkin olarak düzenlemelerine ve farklı stratejiler geliştirmelerine yardımcı olunması hedeflenmektedir.

2. Sosyal medya ve duygu analizi

21. Yüzyılın en önemli özelliklerinden birisi olarak iletişim teknolojileri internet kullanımındaki artış ile birlikte çok ileri seviyelere gelmiştir. İletişim teknolojileri içerisinde internetin öneminin her geçen gün artarak her kesim tarafından çok yoğun bir şekilde kullanılmasından dolayı bireyler artık görüş, fikir ve duygularını sosyal paylaşım siteleri, forum, blog ve mikroblog benzeri birçok ortam aracılığı ile paylaşmaya başlamıştır.

Yeni medya içerisinde yer alan sosyal medya, söz hakkının kullanıcılarda olduğu, bireylerin bilgi, kişisel fikir ve deneyim paylaşımında bulunduğu etkileşimli dijital ortamlardan meydana gelmektedir. İletişimin yeni bir türü olarak öne çıkan sosyal medya dünya tarihinde en hızlı büyüyen iletişim ve paylaşım alanı olmuştur (Avcı ve Sürücü, 2018). Sosyal ilişkilerden sağlığa, alışveriş alışkanlıklarından iş yaşamına, eğitimden eğlenceye, dilden ülke savunmasına kadar hayatın her alanında sosyal medya etkisini göstermektedir (Kılınç vd., 2016).

Ortaya çıkan yeni teknolojiler ve sosyal medya alanları ile günümüzde internet blogları, e-ticaret, etkileşimli web siteleri, dijital gazeteler, Facebook, Twitter, çevrimiçi forumlar gibi elektronik platformları bireyler sadece okumakla kalmayıp aynı zamanda görüşlerini ve duygularını da bırakmaktadır (Gözükara ve Özel, 2016). Örneğin 2006 yılında kurularak kayıtlı kullanıcılarının en çok 140 karakterden oluşan kısa mesajlar yani diğer adıyla tweet (civildama) gönderebilmesine olanak tanıyan ve çok popüler bir sosyal ağ sitesi olan Twitter, diğer sosyal medya alanlarından bağımsız olarak düşünüldüğünde günümüzde bir ülke olsaydı eğer kullanıcı sayısı bakımından dünya üzerindeki 3. büyük ülke konumuna yerleşirdi. Ülkeler bazında bakıldığında ise 193 Birleşmiş Milletler üyesi ülkenin %83'ünde Twitter aktif olarak kullanılmaktadır. Ayrıca tüm internet kullanıcılarının %24'ü Twitter kullanırken, tüm internet kadın kullanıcılarının %21'i

Twitter'ı aktif olarak kullanmaktadır. Twitter kullanıcılarının %37'sinin yaşları 18 ile 29 yaşları arasında ve %25'i ise 30 ile 49 yaşları arasındadır (Aslam, 2019). Bu durum günümüzde insanlık tarihi boyunca hiç görmediğimiz bir veri üretimi ve paylaşımına tanıklık ettiğimizi göstermektedir. Veri kelime anlamı olarak, işlenmemiş gerçek veya bilgi parçacığına verilen isimdir. Veriler; ölçüm, sayım, deney, gözlem veya araştırma yoluyla elde edilmektedir. İnternetin hızla yayılması, sosyal mecraların sayısının artması ve mobil cihazların yaygınlaşması, veri miktarının artmasına neden olmuştur. Veri miktarının her 2 yılda bir iki katından fazla arttığı düşünülürse 2013 yılı istatistiklerine göre 4.4 trilyon gigabayt olan dijital evrenin, 2020 yılında 10 katına ulaşacağı (44 trilyon gigabayt) öngörülmektedir (Özmen, 2016).

Büyük veri, günlük hayatın birçok alanına uygulanabilmekte ve doğru şekilde kullanılabilirse her aşamada paraya ve prestije dönüşebilecek muazzam bir güç durumuna gelmektedir. Büyük verinin analiz edilerek işlenmesi, işletme veya kurumların iş uygulamalarını ve iş yapış şekillerini kökünden değiştirecek, mevcut yapılarını da bilgi tabanlı organik bir şekle dönüştürebilecektir. Yaşanılan bu büyük dönüşüm ile verinin farklı taraflarındaki tüm kullanıcılarına yeni perspektifler kazandırarak büyük bir vizyon sunacaktır (Ünal, 2015). Bu vizyon doğrultusunda farklı kurumların karşılaştığı sorunların çözümünde büyük veri, uygun yararlı olan verileri organize ederek, depolama ve analiz etmek için kullanılan bir teknoloji olarak ifade edilmektedir. Büyük veriler "3V" ile karakterize edilmiştir. Bunlar hacim (volume), çeşitlilik (variety) ve hız (velocity)'dir (Mishra, 2018);

Hacim; Verilerin boyutudur. Bugün çok sayıda veri vardır ve bu veriler yaratıcı tüketicilerle birlikte sağlık, kamu sektörü, perakende vb. gibi farklı alanlardan üretilmektedir. Dolayısıyla veri boyutu petabayt (PB) ve eksabayt'ı (EB) geçiyor ve günümüzde yaklaşık 8 zettabayt (ZB) dijital veri dünyasına ulaşılmıştır. Dolayısıyla doğru depolama, gerçek analizin ilk şartı olmuştur.

Çeşitlilik; Verilerin toplanması esnasındaki yayınlar, bloglar, mikrobloglar, sesli kitaplar, görüntüler vs. olarak çeşitli veriler yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış olarak üçe ayrılmaktadır. Örneğin veritabanı gibi tanımlanmış bir biçime sahip olan veriler yapılandırılmış verilerdir. Görüntüler, videolar, yapıları veya formları olmayan ses dosyaları ise yapılandırılmamış verilerdir.

Hız; Veri üretme hızının ölçümüdür. Gerçek zaman verilerinin akışı çok hızlı olduğu için bugünün hızında verilerin anlamını bulmak zorlaşmaktadır.

Ayrıca çok miktarda veriyi denetlemek zor olduğu için veri güvenliğinin hayati rolü ön plana çıkmaya başlamış ve büyük veriye iki farklı özellik daha eklenmiştir. Bunlar doğruluk (veracity) ve değer (value)'dir. Verilerin doğruluk kalitesi, istenilen gerçek bilgiyi içermesi ve aynı zamanda konu içeriklerinin değeri önem taşımaktadır (Mishra, 2018). Çünkü verilerin çoğu yanıltıcı olabilmekte ve aynı zamanda veriler gerektiği anda gerektiği zaman aralığında kullanılmazsa artık hiçbir değeri kalmamaktadır. Dolayısıyla değerini yitiren bir verinin depolanması da artık gereksiz olacaktır.

Her geçen gün daha da artan veri sayısı ve boyutu, bu verilerden manuel olarak anlamlı bilgiler çıkartılmasını çok zahmetli ve pahalı bir iş haline getirmektedir (Kaynar vd., 2017). Artık geleneksel yöntemler yardımıyla veri toplamak

ile dijital çağda veri toplamak arasında büyük farklılıklar olduğu kuşkusuzdur. Eskiden kahvehane ve kiraathaneler, pazaryerleri, camiler, çeşme başları, hamamlar, düğünler, şöenler, berber salonları, toplu çamaşır yıkama alanları, ortak fırınlar gibi ortamlar aslında sosyal mecranın ilkel örneklerini oluşturlardı ve veriler geleneksel iletişim araçları yöntemleri ile toplanmaktaydı (Tuncer, 2013). Ancak günümüzde geleneksel araçlar verilerin depolanması ve analiz edilmesi için yeterli değildir. O yüzden günümüzde sesler, videolar ve resimler gibi farklı kaynaklarla donatılmış büyük veriler, dijitalleşmenin etkisi ile mobil ortamlar ve sosyal ağlar vasıtasıyla elde edilmektedir (Mishra, 2018). Bu kapsamda günümüzde sosyal medya, halkın duygularını temsil etmenin hayati bir platformu olarak ortaya çıkmış ve duygu analizi alanındaki veri madenciliği araştırmalarını üst düzeye çıkarmıştır (Pandey vd., 2017).

Çağımızın yeni endüstriyel vizyonu olan dördüncü sanayi çağında (diğer adıyla Endüstri 4.0 olarak isimlendirilen dördüncü sanayi devriminde) veri toplamak için sosyal medya platformları daha çok kullanılmaktadır. Bireylerin sosyal medya platformlarından paylaştığı verilerin duygu içerip içermediği otomatik olarak saptanmakta ve paylaşılan duyguların pozitif veya negatif olma durumunun belirlenmesi ile duygu analizleri yapılmaktadır. Bu bağlamda birçok çağdaş otomasyon sistemini, veri alışverişlerini ve üretim teknolojilerini içeren kolektif bir bileşen olarak veri madenciliği sayesinde bireylerin duygu durumlarına göre hedef kitleleri tarafından hangi ürünlerinin daha çok beğenildiği, hangi konularda şirketlerin kendilerini geliştirmelerinin gerektiği Endüstri 4.0'ın sosyal medyadaki yaratıcı (üreten-paylaşım) bireyleri sayesinde takip edilebilmektedir.

Bu kapsamda bireyler yaşadıkları duygularına göre, deneyim yaşadıkları veya satın aldıkları ürünler hakkında olumlu ya da olumsuz olarak düşüncelerini ifade ederler. Burada önemli olan amaçlanan iletişim doğrultusunda duyguların doğru izlenerek gittiği yönün otomatik olarak belirlenmesidir. Çünkü bireylerin hangi ürünü neden tercih ettiği, rakip firmaların tüketicilerine yaşattıkları olumlu veya olumsuz duyguların neler olduğu, Pazar analizleri, itibar yönetimi, kriz yönetimi gibi konularda strateji belirlenebilmesi için Endüstri 4.0 çağının yeni medya etkileşiminin takip edilmesi ve verilerin anlamlandırılması şart olmuştur. Hatta duygu analizleri sayesinde tüketiciye sunulacak olan reklam ve satın alma kampanyalarında da dikkatli olunacak ve müşterinin yaşadığı olumsuz duygular esnasında firma hiçbir şekilde tanıtım unsurları uygulamayacaktır. Aksine eğer firma olumsuz duyguların yaşadığı anlarda tüketim olarak daha çok tercih edilen bir firma ise olumsuz duyguların tespit edildiği anlarda otomatik olarak devreye girebilecektir.

Bu durumda gerekli olan şey müşteri verilerinin iyi modellenmesi ile bilinmeyenlerin azaltılması ve müşterilerin genel eğilimlerinin ve duygu değişimlerinin erkenden öğrenilerek, şirketlerin de bu doğrultuda gerekli önlemleri almasının sağlanmasıdır. Sosyal mecralarda paylaşılan ve kurumların iç veri kaynakları sayesinde anlamlandırılan çok sayıda verinin doğal sonucu olarak, eski pazarlama evrelerinde popüler olan segmentasyon kavramı da önemini yitirmeye başlamış ve Pazarlama 4.0'ın kişiselleştirme, müşterilerin duyularına hitap edebilme özelliği, anlamlandırılan veriler sayesinde önem kazanmaya başlamıştır (Özmen, 2016).

Burada önemli olan diğer bir konu da mevcut olan büyük verinin kalitesini ve bunun sonucunda ortaya çıkabilecek problemleri daha derin bir anlayışa sahip olabilmektedir. Aksi halde yapılacak araştırmalar ve anlamlandırmalar, muhtemelen önyargılı veya yanlış sonuçlara neden olabilecektir (Liu vd., 2016). Çünkü konular genellikle tek başına anahtar kelimelerle tanımlanabilirken, duygular çok ince bir şekilde ifade edilebilmektedir (Tang vd., 2009). Ayrıca statik analizler yapılırken çok sayıda kelimenin olduğu normal web sayfası metinlerinden farklı olarak, sosyal medya ortamlarında paylaşılan mesajlar çok kısadır ve dil varyasyonları olarak çeşitli kültürel farklılıklardan ötürü kullanıcılar dil bilgisi kurallarına genel itibariyle uymayan düzensiz kelimeler kullanmaya daha çok eğilimlidir (Huang vd., 2017). Bu durum sosyal medya ortamlarının kendine ait bir jargonunun olmasından ve yazım hatalarıyla sıklıkla karşılaşılıyor olmasından da kaynaklanmaktadır. Dolayısıyla duygu analizlerinde morfolojik çözümler yapılırken dikkatli olunması gerekmektedir (Nizam ve Akın, 2014).

Tüm bunların analizi için sosyal medya aracılığında daha çok duyguların, haberlerin ve fikirlerin paylaşıldığı bir mikroblog ortamı olan Twitter ön plana çıkmaktadır. Çünkü Twitter daha çok bireylerin duygularının ifade edildiği bir yer olarak diğer sosyal medya platformlarından ayrılmaktadır. Günümüzde yaygın bir iletişim platformu haline gelen Twitter'daki zengin veri miktarı, kamuoyunu ve sosyal duyguyu analiz etmek ve yönetmek için yararlı bir ortam olarak dikkatleri üzerine çekmiştir. Geleneksel metinlerden farklı olarak mikroblogging verileri; ifadeler, görüntü vb. gibi çok yönlü verileri içeren çok modlu bir özelliğe sahiptir. Bu yüzden çoğu mevcut duyarlılık ve konu algılama yaklaşımları, benzersiz mikro verileri değerlendirmektedir. Ayrıca Twitter da yer alan fikir zenginliği ve konu odaklı kaynakların artması, hükümetler, işletmeler ve akademik kurumlarca da büyük ilgi görmüştür. Örneğin devlet daireleri, gerçek zamanlı olarak önemli siyasi veya ekonomik olayların kamuoyu görüşlerini izlemek ve hızlı bir şekilde takipçilerine rehberlik etmek istemektedir. Kurumsal satış karar vericileri ise ürün veya hizmetleri ile ilgilenen hedef kullanıcıları tarafından hassas pazarlama stratejisini etkin bir şekilde yürütmek istemektedir. Bu nedenle, duygu durumlarını otomatik olarak saptama yöntemi ile karakterize edilen mikroblogging verilerindeki hisleri keşfetmek değerli bir konuma gelmiştir (Huang vd., 2017). Özetle Twitter'da duygu analizi: potansiyel pozitif, negatif veya sinersel duyguları tanımlamak için tweet'leri sınıflandırmada ve kontrol etmede çok önemli bir rol oynamaktadır (Xiong vd., 2018).

Duygu analizi, veri madenciliği için zor bir çalışma alanıdır. Piyasa değerinin olması ve pratik sonuçlar alınabilmesi hem akademik çalışmaların hem de endüstrinin bu alana ciddi bir şekilde yönelmesini sağlamıştır. Ancak duygu bildiren kaynakları internet üzerinde bulmak ve onlara ulaşım işlemek hâlâ zorlu bir görev olarak karşımızda durmaktadır. Çünkü her biri büyük miktarda duygu barındıran geniş sayıda farklı kaynaklar mevcuttur ve bu kaynakların birçoğunda, duygu uzun metinler içerisinde gizli bir şekilde yer almaktadır. Bir insan için ilgili kaynakları bulmak, o kaynaklardan ilgili duygu içeren kısımları bulup onları özetlemek ve kullanılabilir bir biçimde organize etmek çok zor ve zahmetli bir iştir. Bundan dolayı, otomatik olarak duyguları keşfetmek, analiz

etmek ve özetlemek için özel doğal dil işleme (DDİ) sistemlerine ihtiyaç vardır. Duygu analizi de bu ihtiyaçtan doğmaktadır (Türkmenoğlu, 2015).

Duygu analizleri, sözlüğe dayalı modeller ve makine öğrenmesine dayalı modeller olmak üzere ikiye ayrılmaktadır. Sözlüğe dayalı modellerde, ilk olarak metinlerde hangi duygu durumlarının aranmak istendiği belirlenmektedir. Daha sonra belirlenen duygu durumlarını ifade eden kelimeler ve o kelimelerin anlamdaşları metin içerisinde aranarak her bir kelime için bir sözlük yardımıyla duygu durumu gösteren bir skor değeri elde edilmektedir. Son adımda ise istatistiksel yöntemler ile metnin hangi duygu durumunu ifade ettiği tahmin edilmektedir. Makine öğrenmesine dayalı yöntemlerde ise, ilk olarak metinler etiketlenmekte ve ardından bu metinler çeşitli madenciliği yöntemleri ile temizlenerek ön işlemeden geçirildikten sonra sınıflandırmaya uygun hale getirilmek üzere vektör uzay modelleri oluşturulmaktadır. Ardından bu vektör uzayları alt setlere bölünerek duygu durumu tahminleri yapılmaktadır (Kaynar vd., 2017).

Sonuç olarak duygu analizlerinin yapılacağı ortam olarak seçilen Twitter, kullanıcıların etkin olarak iletişim kurabildikleri, fikirlerini beyan ettikleri sosyal bir platformdur. Twitter ortamında paylaşılan metin mesajları hem araştırmacılar hem de uygulayıcılar için önemli bir işleve sahiptir. Twitter verileri, güncel olayların belirlenmesinde, yaygın hastalıkların önceden tahmin edilmesinde ve kriz yönetimi gibi birçok farklı alanda uygulama alanı bulabilmektedir. Duygu analizi, doğal dil işleme gibi alanlarda yöntem ve tekniklerin kullanılması ile bireylerin metin içerisinde belirttiği, duygu, görüş, tutum gibi öznel bilgilerin belirlenmesini amaçlayan güncel bir araştırma alanıdır. Twitter mesajları üzerinde duygu analizleri gerçekleştirilerek, pazarlama kararlarına ilişkin stratejik kararlar alınabilmesi, finansal oranların tahmin edilmesi, müşteri kayıp analizi, sektördeki fırsat ve tehdit faktörlerinin belirlenmesi, rakiplerin etkinliklerinin önceden belirlenmesi, pazarlama karar verme sürecinin etkin bir biçimde ele alınabilmesi gibi birçok farklı pazarlama uygulaması olanaklı hale gelmektedir (Onan, 2017).

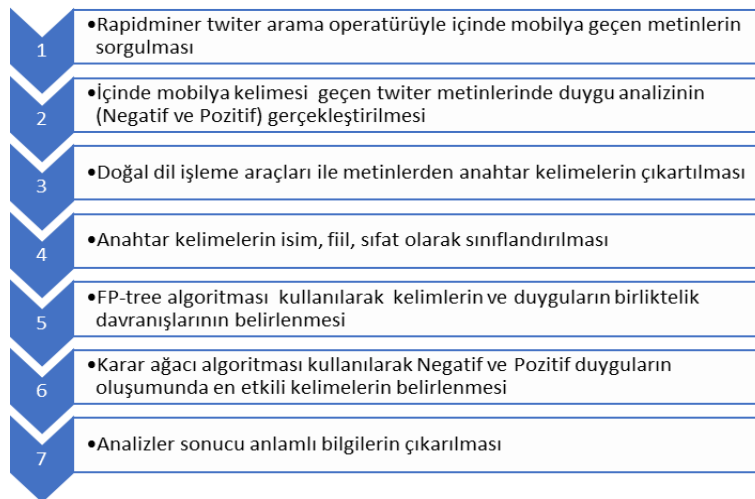
Bu kapsamda şirketler, duygu analizleri sayesinde hedef kitleleri ile çift yönlü iletişim kurarak doğrudan doğruya duygulara hitap eden bir pazarlama anlayışla hareket

edebilmektedir. Böylece tespit edilen duygu dönütleri ve fikirler ile şirketler, ürün ve hizmetlerinin hazırlanışından, tasarımına ve müşteriye ulaşmasında kadar tüm aşamalarda doğru bir şekilde hareket edebilecektir. Dolayısıyla bunların hepsi kurumların ve şirketlerin itibarını da doğrudan etkilemektedir.

Sosyal medya ve veri madenciliği üzerine yapılan bazı bilimsel çalışmalar bu kısımda özetlenmiştir. Ürün ve hizmet üreten şirketlerin promosyonlarını daha büyük kitlelere ulaştırmak için gerçekleştirilen bir bilimsel çalışmada sosyal medya ve veri madenciliği teknikleri kullanılmıştır. Çalışmada Naive Bayes algoritması ile ürün ve hizmetleri takip eden kullanıcılar sınıflandırılmıştır. Aynı zamanda trend kelimeler belirlenmiştir. Çalışma sonucunda sosyal medya madenciliği ile takipçi sayısı %69 oranında artmış olduğu gösterilmiştir (Luke ve Suhajito, 2015). Yapılan bir başka yayında sosyal medya madenciliği ile bisiklet kullanıcılarının duyguları ve motivasyonları anlaşılmasına çalışılmıştır. Bulgular, bisiklet kullanımının hava durumu ve mevsimsel düzen ile ilişkili olduğunu gösterilmiştir. Aynı zamanda bisiklet sürmeye ilişkin genel duyarlılığın olumlu olduğu belirtilmiştir (Das vd., 2019). Halkın aşı konusuna yönelik eğilimlerini belirlemek için sosyal medya madenciliğinden faydalanılmıştır. 10 ay süren bir izleme sonucunda aşlamaya karar verme konusunda kamuoyunu izlemek için sosyal medya veri madenciliğinin düşük maliyetli, gerçek zamanlı ve hızlı bir şekilde kullanılabilmesi gösterilmiştir (D'Andrea vd., 2019).

3. Materyal ve metod

Yapılan çalışmada Rapidminer paket programı kullanılmış ve Twitter API'ı üzerinden bağlantı sağlanarak içinde mobilya geçen popüler tweetler toplanmıştır. Rapidminer; metin incelemesi, makine öğrenimi ve verilerin analizi için ortam sağlayan bir yazılımdır. Rapidminer içindeki doğal dil işleme araçlarından faydalanılarak metinlerin analizi gerçekleştirilmiştir. Son olarak metin analizleri ve veri madenciliği algoritmalarından faydalanarak mobilya seçimindeki yönelimler belirlenmiştir. Şekil 1'de mobilya seçimindeki yönelimleri Twitter datalarına dayalı olarak belirlemek için kullanılan adımlar grafikte gösterilmiştir.



Şekil 1. Mobilya seçimindeki yönelimleri Twitter verilerine bağlı olarak belirlemek için kullanılan adımlar

Rapidminer farklı senaryoların oluşturulabileceği bir veri madenciliği programıdır. İçinde mobilya geçen tweetler, rapidminer programı yardımı ile çekilmiştir. Çalışmada Mayıs 2018-Şubat 2019 tarihleri arasında içinde mobilya kelimesi geçen “14145” adet İngilizce tweet toplanmıştır. Şekil 2’de içinde mobilya geçen popüler tweetlerin çekilmesi gösterilmiştir. Toplanan tweetlerden “29177” anahtar kelime çıkartılmıştır.

Rapidminer içindeki doğal dil işleme eklentileri kullanılarak (Rosette Text Analytics) tweetlerin pozitif veya negatif duygu durumları belirlenmiştir (RTA, 2019). Eklentiler tweeti atan kişinin duygularının pozitif ya da negatif olduğu belirlemede kullanılmaktadır. Şekil 3’te Rapidminer programı ve doğal dil işleme eklentileri ile içinde mobilya geçen tweetlerin toplanması ve duygu durumlarının belirlenmesi gösterilmiştir.

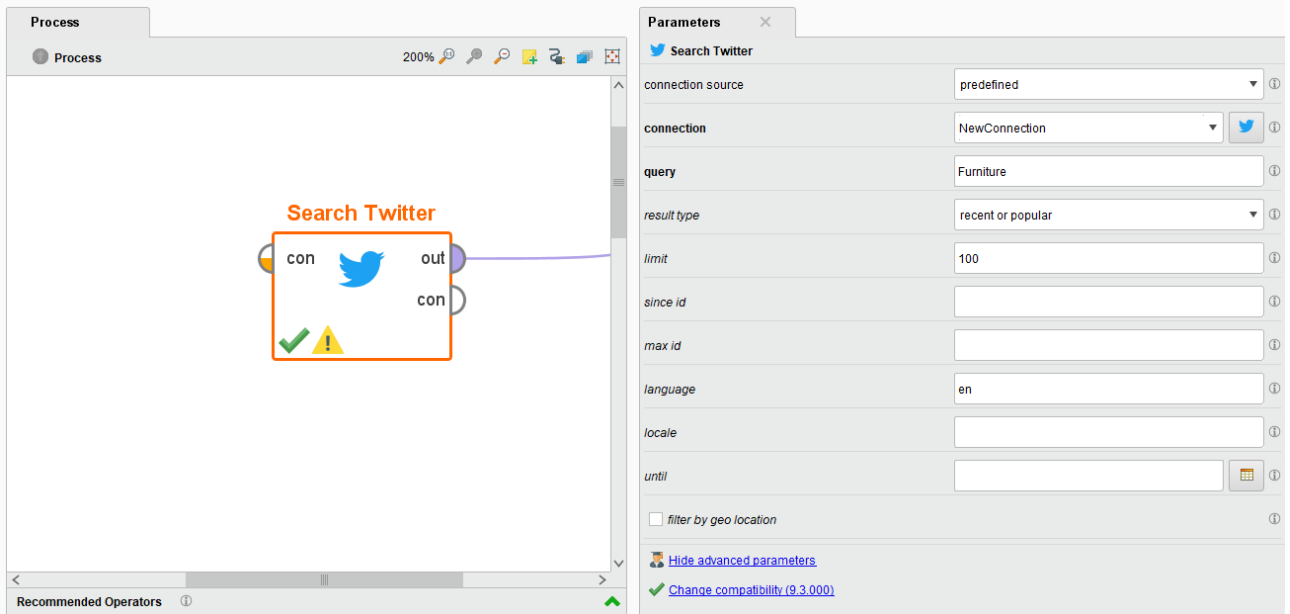
Twitter’da paylaşılan metinler içerisinde kişiler, yerler, ürünler ve kuruluşlar gibi birçok varlıklardan bahsedilmektedir. Buna ek olarak bağlantılar, telefon numaraları, e-posta adresleri, para birimi tutarları, yüzdeler gibi belirli değerler veya önemli sıfatlar, fiiller ve isimler gibi öğeleri de içerebilmektedir. Bu varlıkları ve öğeleri bir metin parçasından çıkarmak için, Rapidminer programında bulunan doğal dil işleme varlık ayıklayıcı operatörü

kullanılmıştır. Bu sayede içinde mobilya geçen tweet metinlerinden anahtar kelimeler çıkartılmıştır. Bu amaç için Aylie Text Analysis (ATA) eklentisi kullanılmıştır (ATA, 2019). Hem önceki adımdaki duygu durumu hem de anahtar kelimeler veri olarak saklanmıştır. Şekil 4’te Rapidminer yazılımı ile anahtar kelimelerin çıkartılması için kullanılan süreç gösterilmiştir.

İçinde mobilya kelimesi geçen bir tweet için rapidminer ile elde edilmiş duygu durumu analizi ve anahtar kelime örneği aşağıda gösterilmiştir;

- This flexible furniture is the ultimate way to save space (Bu esnek mobilyalar yer kazanmak için en mükemmel yoldur) (sentiment (duygu) = positive (pozitif))
- Keywords = flexible, space, ultimate / Anahtar kelimeler = esnek, alan, en mükemmel

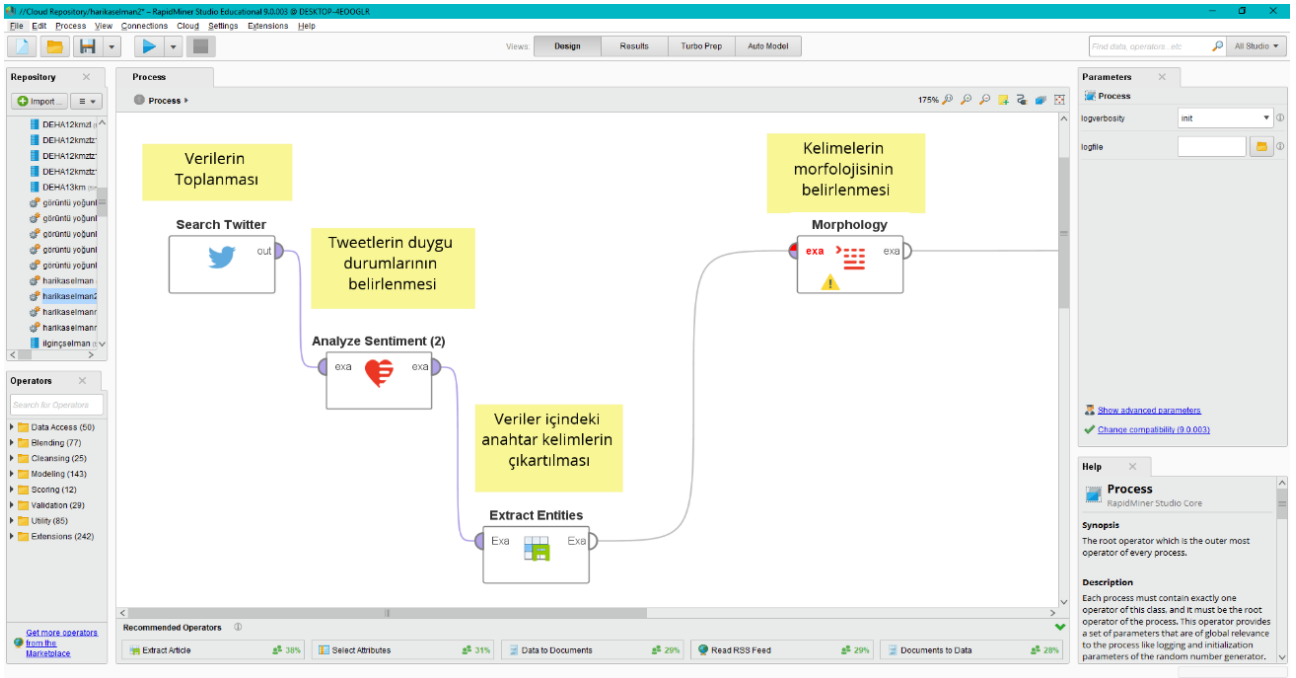
Doğal dil işleme eklentisi ile söz dizimsel işlevlerine uygun olarak bir önceki adımda elde edilen anahtar kelimeler isim, fiil, sıfat ve zarf olarak ayrılmıştır. Veriler toplandıktan sonra veri kümesinde gerekli düzenlemeler ve filtreler yapılmıştır. Şekil 5’te verilerin hazırlanması için oluşturulan süreç gösterilmiştir.



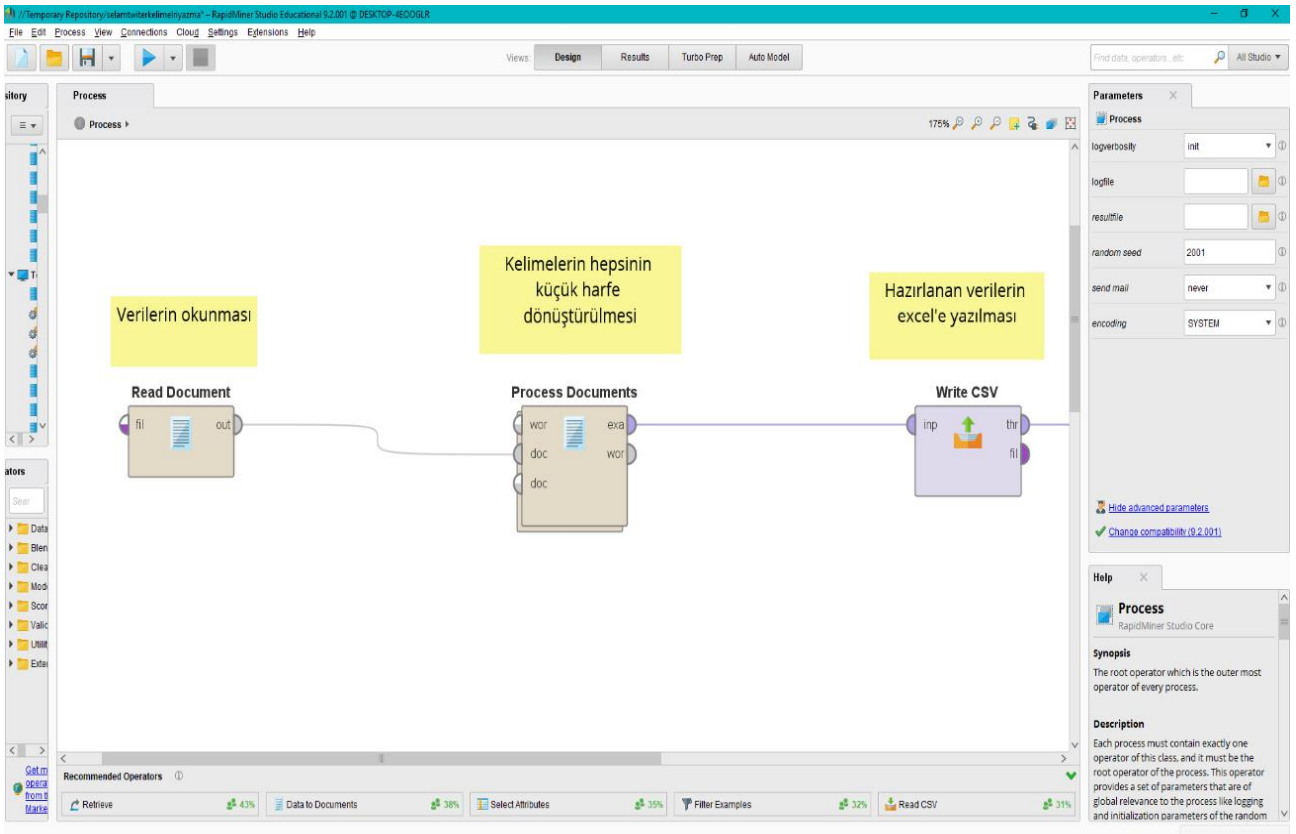
Şekil 2. Rapidminer programı ile içinde mobilya geçen popüler tweetlerin çekilmesi

Row No.	Text	polarity
1	Say hello to Sprinkle! She loves cute furniture and frolicking in the snow. Is she a perfect fit for your campsite?...	positive
2	NO ONE moves furniture quite like @BraunStrowman	negative
3	Federal Customs Officials Seize \$22M Cocaine Shipment Hidden in Furniture - https://t.co/RqFUwzskUc #OA...	neutral
4	If you have ever been curious about image transfers for your painted furniture makeovers or DIY proj https://t.c...	neutral

Şekil 3. Rapidminer programı ve doğal dil işleme eklentileri ile içinde mobilya geçen tweetlerin toplanması ve duygu durumlarının belirlenmesi



Şekil 4. Rapidminer yazılımı ile anahtar kelimelerin çıkartılması



Şekil 5. Verilerin hazırlanması için oluşturulan proses

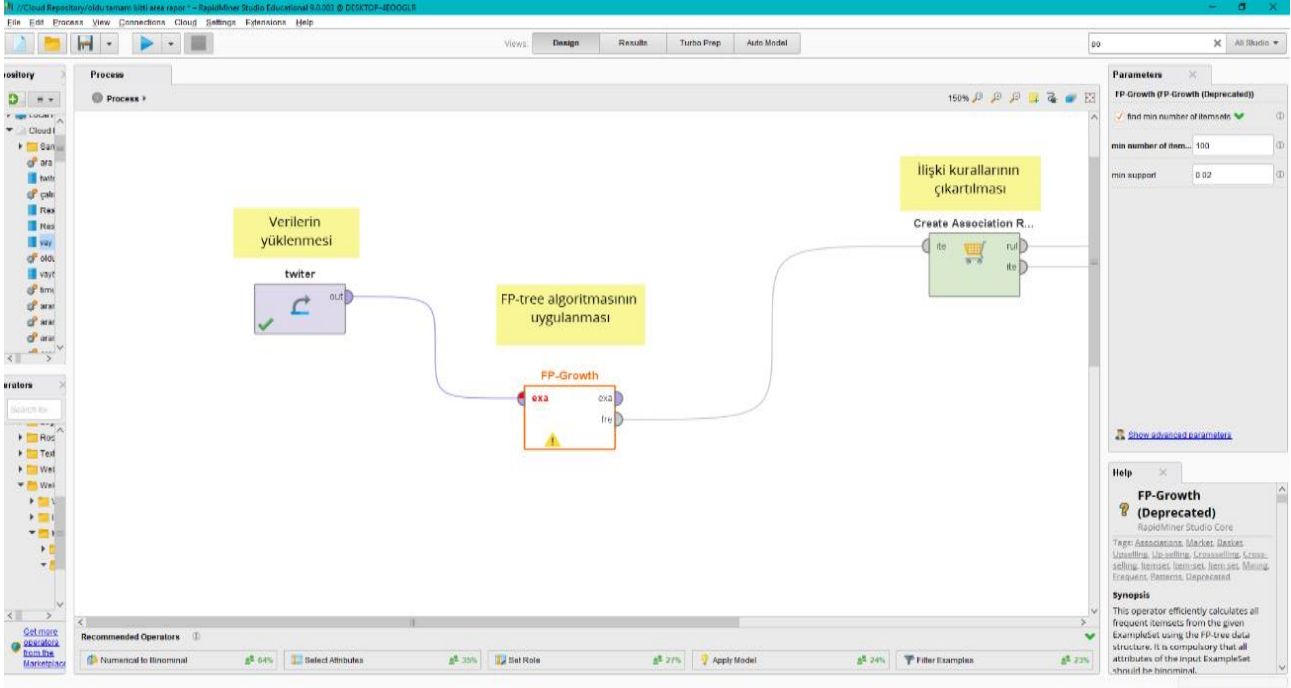
Veri madenciliğinde kullanılan ilk metotlardan birisi birliktelik kurallarıdır. Birliktelik kuralı, geçmiş verilerin analiz edilerek bu veriler içindeki birliktelik davranışlarının tespiti ile geleceğe yönelik çalışmalar yapılmasına yardımcı olmaktadır (Sözen vd., 2017). Duygu durumlarının belirlenmiş kelimelere dönüşümleri uygulanarak veriler hazırlanmıştır. Nominal tipteki veriler binominal (ikili) tipe dönüştürülmüştür. Çalışmada, Twitter verileri ve FP-

Growth algoritması kullanılarak kelimelerin ve duyguların birliktelik davranışları belirlenmiştir. Toplamda 29177 adet veriden faydalanılmıştır. Şekil 6'da Birliktelik kurallarının çıkartılması için hazırlanan proses gösterilmiştir. FP-Growth algoritması büyük veri kümeleri için kullanılan bir birliktelik algoritmasıdır. Veri tabanını algoritma toplamda iki kez taramaktadır. İlk taramada bütün öğelerin destek

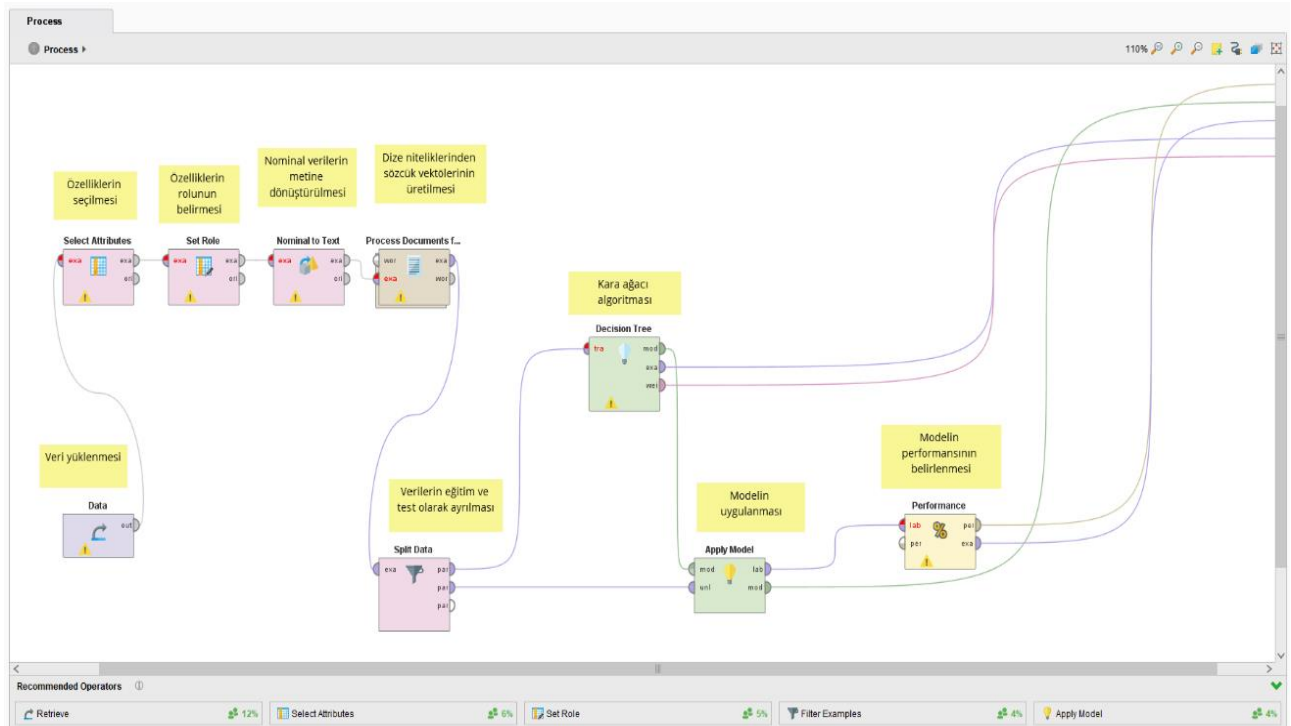
değerini bulunmaktadır. İkinci taramada ise ağaç veri yapısını oluşturulmaktadır.

Karar ağaçları, sınıflandırma için kullanılan parametrik olmayan bir öğrenme yöntemidir. Karar ağaçlarını anlamak ve yorumlamak kolaydır. Bu algoritmalar analizlerde oldukça iyi sonuçlar vermektedir. Aynı zamanda karar ağaçları aktif olarak araştırılan dinamik bir yöntemdir (Wang vd., 2017). Karar ağaçları veri setindeki en ayırt

edici nitelikleri belirlemektedir. Çalışmada önceki adımlardan elde edilen anahtar kelimeler ve duygu analizlerine dayalı karar ağacı algoritması kullanılarak pozitif ve negatif duyguların oluşumda baskın kelimelerin olup olmadığı belirlenmiştir. Bu sayede mobilya satın alınırken sergilenen müşteri davranışları anlaşılabilir çalışılmıştır. Verilerin analizi için hazırlanan proses Şekil 7’de gösterilmiştir.



Şekil 6. Birliktelik kurallarının çıkartılması için hazırlanan proses



Şekil 7. Verilerin analizi için hazırlanan proses

Modelde kullanılan verilerin %80 eğitim ve %20 ise test amacı ile kullanılmıştır. Çalışma sonucunda algoritmalar sayesinde mobilya ürününde hangi kelimelerin olumlu ya da olumsuz algılamada baskın olduğu ve hangi kelimelerin daha çok birlikte kullanıldıkları bilgisi elde edilmiştir.

4. Bulgular

Çalışmada duygu durumları, kelime yapıları ve kelimeler ile bir veri kümesi oluşturulmuştur. Daha sonra algoritmalar ile birliktelikler belirlenmiştir. Çalışmada FP-Growth Algoritması sayesinde güvenilirliği yüzde 10 ile 100 arasında değişen 208 adet kural çıkartılmıştır. Pozitif duygular için kuralların grafik halinde gösterimi Şekil 8'de gösterilmiştir.

Şekil 8'de yer alan kurallardan mobilya ile ilgili öne çıkanlar aşağıda gösterilmiştir:

- [kelime = house] --> [duygu = pos, yapı = noun] (confidence: 0.675)
- [kelime = kids] --> [duygu = pos, yapı = noun] (confidence: 0.786)
- [kelime = family] --> [duygu = pos] (confidence: 0.788)
- [kelime = design] --> [duygu = pos] (confidence: 0.818)

- [kelime = wood] --> [duygu = pos, yapı = noun] (confidence: 0.957)
- [kelime = chance] --> [duygu = pos] (confidence: 0.983)
- [duygu = pos, kelime = bespoke] --> [yapı = verb] (confidence: 1.000)

FP-Growth Algoritması kullanılarak elde edilen kurallar incelendiğinde pozitif duygular uyandıran kelimeler güven değerine göre bespoke (sipariş ile yapılmış), chance (fırsat), wood (ahşap), design (dizayn), family (aile), kids (çocuklar), house (ev) olarak sıralanmıştır. Mobilya kelimesi geçen tweetlerde en pozitif duyguları uyandıran kelimelerin "sipariş ile yapılmış ve fırsat" gibi ifadeler olduğu görülmüştür. Çalışma sonuçlarına göre insanlar özel yapım mobilyaları ve fiyatlarda fırsatlar sunan ürünleri daha pozitif bulmaktadır. Aynı zamanda, tüketicilerin mobilyaları ile özdeşleştiklerini ve ailelerin özellikle çocuk mobilyaları ile ilgili konularda daha fazla paylaşım yaptıkları görülmüştür. Bulgular literatür ile uyumludur. Yapılan önceki çalışmalarda insanların mobilya ürününü tercih ederken aile üyelerinden etkilendikleri vurgulanmıştır (Çabuk vd., 2012; Belch vd., 1985).

Son olarak ahşap gibi sıcak malzemelere ve tasarıma insanların önem verdiği ve bunların pozitif duygular hissettirdiği düşünülmektedir. Negatif duygular için kuralların grafik halinde gösterimi ise Şekil 9'da gösterilmiştir.



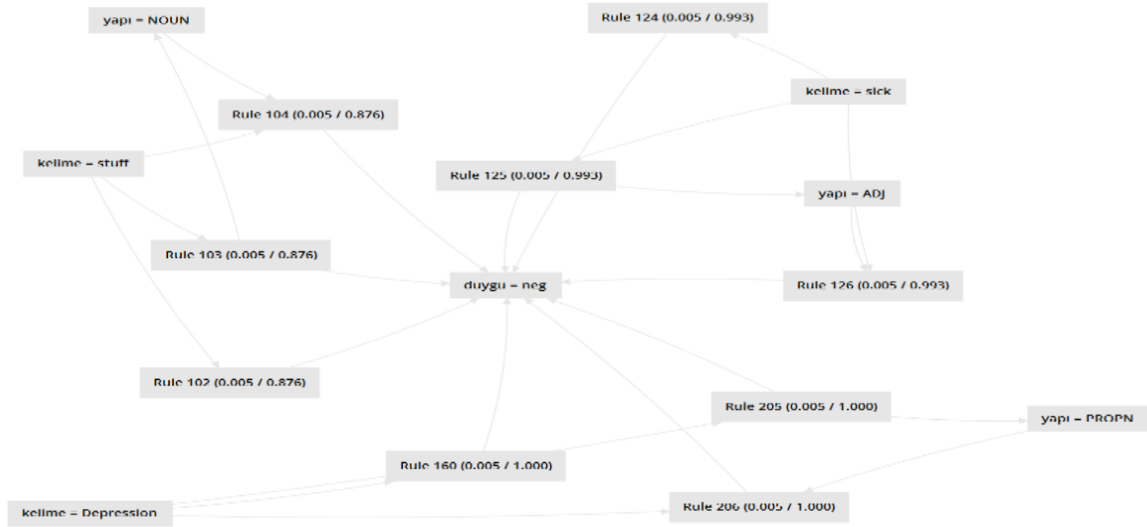
Şekil 8. Pozitif duygular için kuralların grafik halinde gösterimi

Şekil 9'da yer alan FP-Growth Algoritması kullanılarak elde edilen kurallar incelendiğinde negatif duygular uyandıran kelimeler ve sistem tarafından atanan yapıları güven değerine göre aşağıda gösterilmiştir:

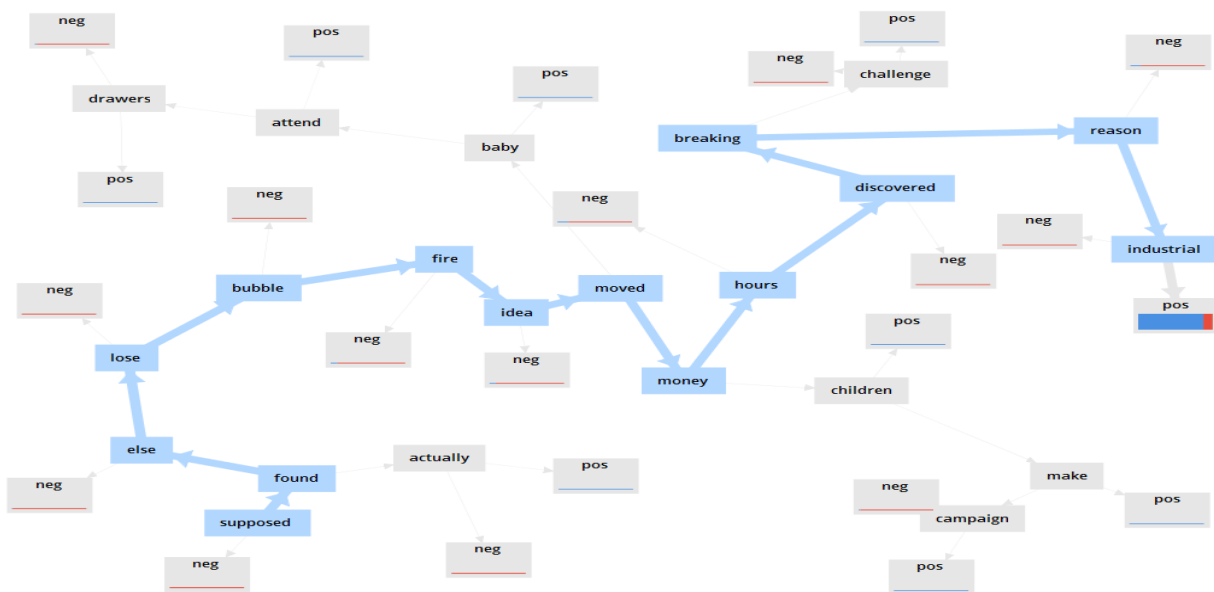
- [kelime = stuff] --> [duygu = neg, yapı = noun] (confidence: 0.876)
- [kelime = sick] --> [duygu = neg, yapı = adj] (confidence: 0.993)
- [kelime = depression] --> [duygu = neg, yapı = verb] (confidence: 1.000)

FP-Growth Algoritması kullanılarak elde edilen kurallar incelendiğinde negatif duygular uyandıran kelimeler güven değerine göre "depression (kasvet), sick (keyifsiz, rahatsız), stuff (kumaş)" olarak sıralanmıştır. Sonuçlar incelendiğinde mobilya ürünü ve insan psikolojisi arasında bir bağlantı kurulabilmektedir. Örneğin mobilyalarda kullanılan kumaşların uygun olmamasının genel anlamda negatif duyguları uyandırdığı belirlenmiştir.

Karar ağaçları iki aşamadan oluşmaktadır. Bunlardan biri öğrenme aşaması diğeri ise sınıflandırma aşamasıdır. Bu iki aşama sonucunda bir ağaç oluşturulmakta ve veriler görselleştirilmektedir (Healey ve Ramaswamy, 2017). Bu sayede daha kolay anlaşılır bir şekle dönüşmüş olmaktadır (Lan vd., 2018). Şekil 10'da Twitter verilerinin analizi sonucu oluşturulan karar ağacı gösterilmiştir.



Şekil 9. Negatif duygular için kuralların grafik halindeki gösterimi



Şekil 10. Twitter verilerinin analizi sonucu oluşturulan karar ağacı

Karmaşık verilerden anlamlı bilgilerin çıkartılmasında karar ağaçları bize büyük kolaylıklar sağlamaktadır. Kurulan modelin performansı incelendiğinde %87.65 oranında başarı göstermiştir. Bu sonuç birçok çalışmada kabul edilebilir düzeyde ve tatminkârdır (Onan, 2015; Elarabi ve Abdelgalil, 2014; Wadie vd., 2006). Ayrıca model bir duygu (pozitif, negatif) durumunun oluşumunda etkili olan kelimeleri önem derecesine göre sıralayabilmektedir. Çizelge 1’de modeldeki kelimeler ve ağırlıkları gösterilmiştir. Modelde ağırlığı en yüksek olan kelime duygu durumunun oluşmasında en önemli kelimedir.

Sonuçlar incelendiğinde, pozitif ya da negatif duyguların oluşumunda challenge (itiraz), actually (sahi, aslında), campaign (kampanya) gibi kelimelerin daha sık olarak kullanıldığı görülmüştür. Bu durum eldeki verilere göre tüketicilerin yaşadıkları problemleri sosyal medyadan diğer kullanıcılarla paylaştıklarını, vurgulu kelimelere önem verdiklerini ve mobilya pazarlamasında kampanyalara önem verilmesi gerektiğini göstermektedir. Aynı zamanda önem sıralamasında discovered (keşfetmek) ve idea (fikir) kelimelerinin bulunması insanların mobilya seçerken orijinal tasarıma, ağızdan ağıza iletişime ve tavsiyelere önem verdiğini, children (çocuklar) ve baby (bebek) kelimelerinin ağırlıklı olarak bulunması ise ailelerin çocuklarına yönelik mobilya ihtiyaçlarında ve ilgili pazara yönelik kullanıcı görüşlerinde diğer ebeveynlere daha çok geri bildirimde bulduklarını göstermektedir. Bu durum çalışmada önceki algoritmalar ile elde edilen anlamlı bilgilerle de uyumludur. Son olarak money (para), breaking (kıрма) gibi kelimelerin önem listesinde bulunması bize insanların mobilya fiyatlarından, kırılmalardan daha çok şikâyetçi olduklarını göstermektedir. Sonuçlar literatür ile uyumludur. Yapılan bir çalışmada tüketicilerin ürün tercihlerinin kabaca %50’si fiyat, %20’si ürün özelleştirme, %20’si teslimat süresi ve %10’u ürünü müşteri isteği ile değiştirmekten kaynaklandığı belirtilmiştir (Lihra vd., 2012). Karar ağacı algoritmasında öne çıkan kelimeler literatürde belirtilen fiyat ve özgünlük unsurlarını kapsamaktadır.

Çizelge 1. Modeldeki en önemli kelimeler ve ağırlıkları

Öznelik	Ağırlık
Challenge (itiraz)	0.0518
Actually (sahi)	0.0515
Campaign (Kampanya)	0.0508
Drawers (çekmeceler)	0.0504
Make (Yapmak)	0.0488
Attend (hazır bulunmak)	0.0479
Industrial(Sanayi)	0.0453
Reason (neden)	0.0451
Breaking (kıрма)	0.0449
Discovered (keşfetmek)	0.0447
Children (çocuklar)	0.0446
Baby (bebek)	0.0446
Hours (saatler)	0.0444
Money (para)	0.0440
Moved (etkilenmiş)	0.0437
Idea (fikir)	0.0435
Fire (ateş)	0.0432
Bubble (hayal)	0.0430
Lose (kaybetmek)	0.0426
Else (başka)	0.0421
Found (kurmak)	0.0417
Supposed (farzedilmiş)	0.0412

5. Tartışma ve sonuç

Çalışmada mobilya seçimindeki yönelimleri belirlemek için geleneksel veri toplama yöntemlerine alternatif olarak tweet verilerinden faydalanılmıştır. Aynı zamanda bu yöntem ile zaman ve mekân sorunu ortadan kalkmıştır. Veri madenciliği tekniklerinden faydalanarak mobilya endüstrisinde beğenilerin ve memnuniyetsizliklerin oluşmasında etkili olan faktörlere ışık tutulmuştur. Bu kapsamda 14145 adet tweet ve 29177 adet anahtar kelime kullanılmıştır. Elde edilen büyük miktarda veriden anlamlı bilgiler çıkarmak için FP-Growth ve Karar ağacı algoritmalarından faydalanılmıştır. Verilerden elde edilen bilgilere göre mobilya ve aile arasında insanların bir bağlantı kurdukları söylenebilir. Aynı zamanda insan psikolojisi üzerine mobilya ürününün etkili olduğu düşünülmektedir. Veri madenciliği algoritmalarından elde edilen bilgiye göre mobilya ürününde olumlu duyguları tetikleyen en önemli unsurların kampanya, fikirler ve orijinal tasarımlar olduğu belirlenmiştir. Sonuçlar, literatürde geleneksel yöntemler ile gerçekleştirilen benzer çalışmalar ile uyumludur. Bu durum daha düşük maliyet ve daha az zaman harcayarak sosyal medya madenciliği ile müşterilerin izlenebileceğini göstermektedir. Çalışma, sosyal medya madenciliğinin mobilya ürününde negatif ya da pozitif duyguların oluşum nedenlerinin anlaşılması konusunda bir çerçeve geliştirmekte ve gelecekteki araştırmalar için çeşitli fırsatlar sunmaktadır. Çalışmadan elde edilen bulgular, mobilya üzerine eğilimleri değerlendirmek amacıyla üreticiler için önemli bir kaynak olarak kullanılabilir. Mobilya üreticilerine sosyal medyadaki geri bildirimlerin değerlendirilmesi önerilmektedir. Bu sayede müşterilerin beklentileri daha iyi anlaşılabilir ve doğru müşteriye hızlı bir şekilde ulaşılabilir. Sonuç olarak, bu çalışma Twitter verilerinin müşteri beklentilerini ve şikâyetlerini gerçek zamanlı olarak takip etmek için kullanılabilirliğini göstermiştir.

Açıklama

Bu çalışma Bartın Üniversitesi, Bilimsel Araştırma Projeleri (BAP) Koordinatörlüğü tarafından desteklenmiştir (Proje Numarası: 2018-FEN-A-015). Yazarlar, desteklerinden dolayı Bilimsel Araştırma Projeleri Koordinatörlüğü’ne teşekkür eder.

Kaynaklar

- Aslam, S., 2019 Twitter by the Numbers: Stats, Demographics & Fun Facts. <https://www.omnicoreagency.com/Twitter-statistics/>. Erişim: 15.01.2019.
- ATA, 2019. Aylien Text Analysis, <https://aylien.com/>, Erişim: 05.11.2019.
- Avcı, Ö., Sürücü, Ç., 2018. Üniversite öğrencilerinin sosyal medya etkileşimleri: Bartın üniversitesi örneği. Karadeniz Uluslararası Bilimsel Dergisi, 40: 500-514.
- Belch, G.E., Belch, M.A., Ceresino, G., 1985. Parental and teenage child influences in family decision making. Journal of Business Research, 13(2): 163-176.
- Çabuk, Y., Karayılmazlar, S., Türedi, H., 2012. Mobilya tercihinde tüketici davranışlarının demografik faktörler bakımından incelenmesi (Zonguldak ili örneği). Bartın Orman Fakültesi Dergisi, 14(21): 1-10.

- D'Andrea, E., Ducange, P., Bechini, A., Renda, A., Marcelloni, F., 2019. Monitoring the public opinion about the vaccination topic from tweets analysis. *Expert Systems with Applications*, 116: 209–226.
- Das, S., Dutta, A., Medina, G., Minjares-Kyle, L., Elgart, Z., 2019. Extracting patterns from Twitter to promote biking. *IATSS Research*, 43(1): 51–59.
- Elarabi, H., Abdelgalil, S.A., 2014. Application of artificial neural network for prediction of Sudan soil profile. *American Journal of Engineering, Technology and Society*, 1(2): 7-10.
- Gözükara, F., Özel, S.A., 2016. Türkçe ve İngilizce yorumların duygu analizinde doküman vektörü hesaplama yöntemleri için bir deneysel inceleme. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 31(2): 464-482.
- Healey, C., Ramaswamy, S. 2017. Visualizing Twitter Sentiment. https://www.csc2.ncsu.edu/faculty/healey/tweet_viz/, Erişim: 09.11.2017.
- Huang, F., Zhang, S., Zhang, J., Yu, G., 2017. Multimodal learning for topic sentiment analysis in microblogging. *Neurocomputing*, 253: 144-153.
- Kaynar, O., Aydın, Z., Görmez, Y., 2017. Sentiment analizinde öznelik düşürme yöntemlerinin oto kodlayıcı derin öğrenme makinaları ile karşılaştırılması. *Bilişim Teknolojileri Dergisi*, 10(3): 319-326.
- Kılınç, B., Eriş, U., Tarkan, M., 2016. İletişim Bilgisi. Anadolu Üniversitesi Yayınları, Eskişehir, Türkiye.
- Lan, T., Zhang, Y., Jiang, C., Yang, G, Zhao, Z., 2018. A Automatic identification of Spread F using decision trees. *Journal of Atmospheric and Solar-Terrestrial Physics*, 179: 389-395.
- Lihra, T., Buehlmann, U., Graf, R., 2012. Customer preferences for customized household furniture. *Journal of Forest Economics*, 18(2): 94–112.
- Liu, J., Li, J., Li, W., Wu, Y., 2016. Rethinking big data: A review on the data quality and usage issues. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115: 134–142.
- Luke, J., Suharjito, S., 2015. Data mining of automatically promotion Tweet for products and services using Naïve Bayes Algorithm to increase Twitter engagement followers AtPT. Bobobobo. In *Procedia Computer Science*, 59: 254–61.
- Mishra, S., 2018. A review on big data analytics in medical imaging. *International Journal of Computer Engineering and Applications*, 7(1): 31-37.
- Nizam, H., Akın, S.S., 2014. Sosyal medyada makine öğrenmesi ile duygu analizinde dengeli ve dengesiz veri setlerinin performanslarının karşılaştırılması. XIX. Türkiye'de İnternet Konferansı'nda sunuldu, 27-29 Kasım, İzmir, s. 129-136.
- Onan, A., 2015. Şirket iflaslarının tahmin edilmesinde karar ağacı algoritmalarının karşılaştırmalı başarımların analizi. *Bilişim Teknolojileri Dergisi*, 8(1): 9-19.
- Onan, A., 2017. Twitter mesajları üzerinde makine öğrenmesi yöntemlerine dayalı duygu analizi. *Yönetim Bilişim Sistemleri Dergisi*, 3(2): 1-14.
- Özmen, U., 2016. İnternet ve Mobil Pazarlamada Veri ve Veri Madenciliği, (Ed., Eriş, U.), Anadolu Üniversitesi Yayınları, Eskişehir, Türkiye. s. 104-121.
- Pandey, A.C., Rajpoot, D.S., Saraswat, M., 2017. Twitter sentiment analysis using hybrid cuckoo search method. *Information Processing and Management*, 53: 764-779.
- RTA, 2019. Rosette Text Analytics, <https://www.rosette.com/>, Erişim: 04.11.2019.
- Sözen, E., Bardak, T., Peker, H., Bardak, S., 2017. Apriori algoritması kullanılarak mobilya seçiminde etkili olan faktörlerin analizi. *İleri Teknoloji Bilimleri Dergisi*, 6(3): 679-684.
- Tang, H., Tan, S., Cheng, X., 2009. A survey on sentiment detection of reviews. *Expert Systems with Applications*, 36: 10760–10773.
- Tuncer, A.S., 2013. Sosyal Medyanın Gelişimi, (Ed. Özata, F., Z.), Anadolu Üniversitesi Yayınları, Eskişehir, Türkiye. s. 2-25.
- Türkmenoğlu, C., 2015. Türkçe metinlerde duygu analizi. Yüksek lisans tezi, Fen Bilimleri Enstitüsü, İstanbul Teknik Üniversitesi, İstanbul.
- Ünal, F., 2015. Büyük Veri ve Semantik. Abaküs Yayınevi, İstanbul, Türkiye.
- Wadie, B.S., Badawi, A.M., Abdelwahed, M., Elemabay, S.M., 2006. Application of artificial neural network in prediction of bladder outlet obstruction: A model based on objective, noninvasive parameters. *Urology*, 68(6): 1211–1214.
- Wang, Y., Xia, S.T., Wu, J., 2017. A less-greedy two-term Tsallis Entropy Information Metric approach for decision tree classification. *Knowledge-Based Systems*, 120: 34-42.
- Xiong, S., Lv, H., Zhao, W., Ji, D., 2018. Towards Twitter sentiment classification by multi-level sentiment-enriched word embeddings. *Neurocomputing*, 275: 2459-2466.