

ULAKBİM İndeksinde Taranan Sosyal Bilimler Alanındaki Dergilerde Öne Çıkan Konu ve Yöntemlerin Metin Madenciliği Yaklaşımı ile Belirlenmesi

Serkan ETİ¹

Özet

Bu çalışmada ULAKBİM indeksinde taranan sosyal bilimler alanındaki dergilerin içeriği analiz edilmiştir. Bu çerçevede, 537 dergi içerisinde dergipark sistemine kayıtlı olan 308 derginin 2019 yılına ait son sayıları inceleme kapsamına alınmıştır. Bununla birlikte, analiz sürecinde metin madenciliği yönteminden faydalanılarak en fazla geçen kelime ve kelime grupları belirlenmiştir. Sonuç olarak, uygulamalı çalışmaların ön plana çıktığı tespit edilmiştir. Belirtilen bu çalışmalarda AHP, DEMATEL, VIKOR ve TOPSIS gibi çok kriterli karar verme yöntemlerinin tercih edildiği anlaşılmaktadır. Dolayısıyla, yapılacak yeni çalışmalarda bahsi geçen bu yöntemlerin kullanılmasının ilgili çalışmanın kabul edilme oranını artıracak olduğu düşünülmektedir. İlgili çalışmada elde edilen bir diğer sonuç ise, finans alanında yapılan çalışmaların çoğunlukta olmasıdır. Öte yandan, çalışmaların önemli bir bölümünde Türkiye üzerine analiz yapıldığı anlaşılmaktadır. Buradan yola çıkarak, yapılacak yeni çalışmalarda farklı ülke gruplarının dikkate alınarak panel veri analiz yapılması söz konusu makalelerin özgünlüğünün artmasına katkı sağlayacaktır.

Anahtar Kelimeler: ULAKBİM İndeksi, Metin madenciliği, Sosyal bilimler

Determination of the Featured Topics and Methods in the Social Sciences Journals Indexed in ULAKBİM with Text Mining Approach

Abstract

In this study, the content of journals in the field of social sciences indexed in ULAKBİM is evaluated. Within this framework, the latest issues of 308 journals in 2019 out of 537 journals are included in the scope of the analysis. On the other side, the most common words and word groups were identified with the help of text mining method in the analysis process. The findings show that applied studies are mostly preferred in these journals. Additionally, multi-criteria decision making methods such as AHP, DEMATEL, VIKOR and TOPSIS are taken into consideration. Therefore, it is thought that the use of these approaches in new studies will increase the acceptance possibility. Another result is that the majority of the studies are carried out in the field of finance. Furthermore, it is understood that a significant proportion of the studies focused on Turkey. From this point of view, it can be said that panel data analysis by considering different country groups in the new studies will contribute the originality of the articles.

Keywords: ULAKBİM Index, Text mining, Social sciences

1. Giriş

Akademide bilim insanlarının performansını değerlendirmede makale, bildiri, kitap ve atf gibi kriterlere dikkat edilir. Bununla birlikte yapılan yayınların nitelikleri de önemli bir kriterdir. Yayınların kalitesini belirleyen en önemli unsur ise yayınlandığı indekstir. Akademik çalışmalarda birçok farklı indeks söz konusudur. En çok kabul gören indeks Social Sciences Citation Index (SSCI) ve Science Citation Index (SCI) indeksleridir. Birincisi Sosyal bilimlerde, ikincisi ise fen bilimlerinde ön plana çıkan indekslerdir (Eti, 2016). Bundan dolayı, akademik yayın yapan hakemli dergiler bu indekslere girmeye çalışırlar. Fakat bu indekslere girmek zordur ve bu yüzden de buradaki dergiler seçilmiş dergiler olarak literatürde kabul görmektedir. Dolayısıyla, bu indekslerde yapılan yayınlar da daha kaliteli çalışmalar olarak kabul edilmektedir. Öte yandan, söz konusu indeksler dışında da kabul gören indeksler bulunmaktadır. Scopus ve Econlit indekslerinde uluslararası kabul gören diğer indekslerdir.

Ülkemizde ise Ulusal Akademik Ağ ve Bilgi Merkezi (ULAKBİM) indeksi öne çıkan yerel indekslerdendir. TÜBİTAK arama tabanını kullanan ULAKBİM indeksine her dergi girememektedir. Adı geçen indeks sosyal ve

¹Öğ. Gör., İstanbul Medipol Üniversitesi, seti@medipol.edu.tr

fen olmak üzere iki ayrı kategoride dergileri sınıflandırmaktadır. Türkiye’de ULAKBİM’de taranan dergiler diğer yerel dergilere göre daha kalite kabul edilmektedir. Özellikle son yıllarda doçentlik başvuru sürecinde ULAKBİM’de taranan dergilerde yayın yapma kriteri ön koşul haline getirilmiştir. Bundan dolayı da hem dergiler ULAKBİM indeksine girmeye çalışmakta hem de akademisyenler bu indekste yayın yapmak istemektedirler. Bu talep durumu, ULAKBİM indeksinin kalitesini artırmaktadır.

Özellikle doçentliğe hazırlanan akademisyenler, ön koşul olan ULAKBİM indeksindeki dergilere ciddi talep göstermektedir. Bu yüzden birçok dergi daha yayınlanmamış sayılarını doldurduklarını ilan etmişlerdir. Gelen makale taleplerine, önlerindeki bir iki sayı için dolu olduklarını ve yayın alamadıklarını bildirmişlerdir. Bu yüzden, bu dergilerde talep artığından ret oranlarında yükselme görülmüştür. Ret oranlarının yükselmesinden dolayı, artık ULAKBİM indeksindeki dergilerin hangi yayınları kabul ettiği konusu önem arz etmektedir. Bu sayede yazarların bu indekste dergilere makale gönderirken nelere dikkat etmesi gerektiğinin belirlenmesi gerekmektedir.

Bu çalışmanın amacı da ULAKBİM indeksindeki sosyal bilimler alanında yer alan dergilerin kabul ettiği makalelerin detaylarının analiz edilmesidir. Bu çerçevede 2019 Temmuz ayı itibari ile bu kapsamdaki 308 adet derginin son sayılarındaki yayınlarının özetleri analiz edilmiştir. Analiz sürecinde metin madenciliği kullanılarak makalelerde en çok geçen kelime ve kelime grupları belirlenmiştir. Bu sayede özellikle ULAKBİM’deki makalelere yayın göndermek isteyen akademisyenlerin nelere dikkat etmesi gerektiğinin ön plana çıkarılması hedeflenmiştir. Bu sayede hem literatüre katkı bulunulması hem de yazarlara yol gösterilmesi hedeflenmektedir.

Çalışmamız 4 bölümden oluşacaktır. İlk bölüm olan bu kısımda konu hakkında genel bilgiler verilmiştir. İkinci bölümde metin madenciliği hakkında teorik bilgi verilip bu konu hakkında yapılan çalışmalardan bahsedilecektir. Çalışmamızın üçüncü kısmında ise konuyla ilgili yapılan uygulamaya yer verilecektir. Son bölümde ise, analiz sonuçları ve buna yönelik tarafımızca verilen önerilerin altı çizilecektir.

2. Metodoloji

Birçok çalışmanın temel taşı genellikle veridir. Araştırma için farklı tipte veriler toplanabilmektedir. Bazı çalışmalarda sayısal veriler elde edilebilirken (Dinçer et al., 2019a,b), bazı durumlar ise sadece metin içerikli veriler olabilmektedir. Sayısal olmayan ya da sayısallaştırılamayan veriler için nitel araştırmalar en tercih edilen araştırma türlerinden biridir. Verilerin sözel ifadelerden olduğu durumlarda, bir diğer kullanılabilecek yöntem ise veri madenciliği yöntemlerinden birisi olan metin madenciliğidir.

Metin madenciliği yöntemi, metinlerden oluşan veri ambarlarında anlamlı sonuç çıkarma yöntemlerinden birisidir. Söz konusu yöntemde, kelime ve kelime gruplarının özelliklerine dayanarak sonuç çıkarma temeline dayalıdır (Dolgun vd., 2009). Söz konusu yöntemde, terimlerin özelliklerine göre kişi, zaman, organizasyon, ürün gibi sınıflandırmalar sayesinde analiz yapılabilmektedir. Metin madenciliği analizi, metinlerin sınıflandırılması, konu/anahtar kelime çıkarımı, duygu analizi, metin özetleme ve ilişki modellemesi gibi amaçlar için kullanılmaktadır (Şeker, 2015). Metin madenciliği, ilk olarak 1960’larda işlenmemiş verilerden anahtar kelime arama süreciyle başlamıştır. Yapay zekanın gelişmesi ile birlikte doğal dil işleme süreç (NLP) algoritmaları geliştirilmiştir. Günümüzde mevcut metin madenciliği algoritmaları bu süreci temel almaktadır (Cambria & White, 2014). Bu amaçla, doğal dil işleme kullanılması ile özellik çıkarımı yapılabilmektedir. Hece analizi, kelime frekansı dağılımı, örüntü tanımlama, etiketleme, görselleştirme de bu amaçla yapılan bazı metin madenciliği analiz teknikleridir (Tan, 1999 & Mattingly, 1972). Metin madenciliği, müşteri ilişkilerinde, pazarlamada, sahtekarlık tespitinde, güvenlik araştırmalarında ve bilimsel araştırmalar olmak üzere birçok alanda kullanılmaktadır.

Literatüre incelendiğinde, web analizi başta olmak üzere birçok alanda metin madenciliği kullanılmaktadır. Shen vd. (2018) çalışmasında sosyal medyayı dikkate alarak iki dilli metin madenciliği üzerine bir araştırma gerçekleştirmişlerdir. Söz konusu çalışmada online to offline (O2O) ticaretinin trendi üzerinde durulmuştur. Farklı dil bölgelerinde O2O gelişim ile ilgili trend analizleri tartışılmıştır. Ayrıca, Ding vd. (2018) enerji tasarrufu üzerine bir çalışma gerçekleştirmişlerdir. Çalışmalarında 1600 makale ve rapor metin madenciliği ile analiz edilmiştir. Çalışmanın sonucu olarak yeşil binalar, akıllı binalar ve düşük karbonlu binalar inşaat sektöründe eğilim olunan konular olarak belirlenmiştir. Eskici ve Koçak (2018) ise çalışmalarında metin madenciliği finans alanında kullanmışlardır. İlgili çalışmada merkez bankasının aylık fiyat geliştirme raporları metin madenciliği yardımıyla analiz edilmiş ve sonuç olarak, aylık fiyat geliştirme raporlarının enflasyon değerlendirilmesinde çekirdek yoğunlaştığı görülmüştür. Ayrıca kümelerin yıllık Tüketici Fiyatları Endeksi (TÜFE) enflasyon rakamları ile anlamlı olarak uygunluk gösterdiği rapor edilmiştir.

Öte yandan, Tobback vd. (2018) çalışmasında Belçika ekonomisindeki belirsizlik endeksinde metin madenciliğinden yararlanmıştır. Söz konusu amaç için destek vektör makinası tekniğinden yararlanarak sınıflandırma yoluna gidilmiştir. Sonuç olarak, belirsizliğin yüksek olduğu dönemlerde devlet tahvilinin getirisi ve yayılımı, krediler ve tüketici güvenceleri öngörülebildiği saptanmıştır. Kim ve Kang (2018) kozmetik ürünlerin ayırt edici özellikleri üzerine yaptıkları çalışmada metin madenciliğinden yararlanmışlardır. Çalışmanın yöntem kısmında duygu analizi ve doğal dil işleme algoritmaları kullanılmıştır. Böylelikle, pazar araştırmaların alternatif olacak bir yaklaşım önermişlerdir. Pınarbaşı ve Canbolat (2018) çalışmasında mobil uygulamalara gelen yorumları metin madenciliği ile incelemiştir. Adı geçen yazarlar, ngram yöntemi ile en çok geçen ikili ve üçlü kelime gruplarını analiz etmişlerdir. Yuksel vd. (2018) çalışmalarında Türk bankacılık sektörünü metin madenciliği ile performans değerlendirmesi yapmışlardır.

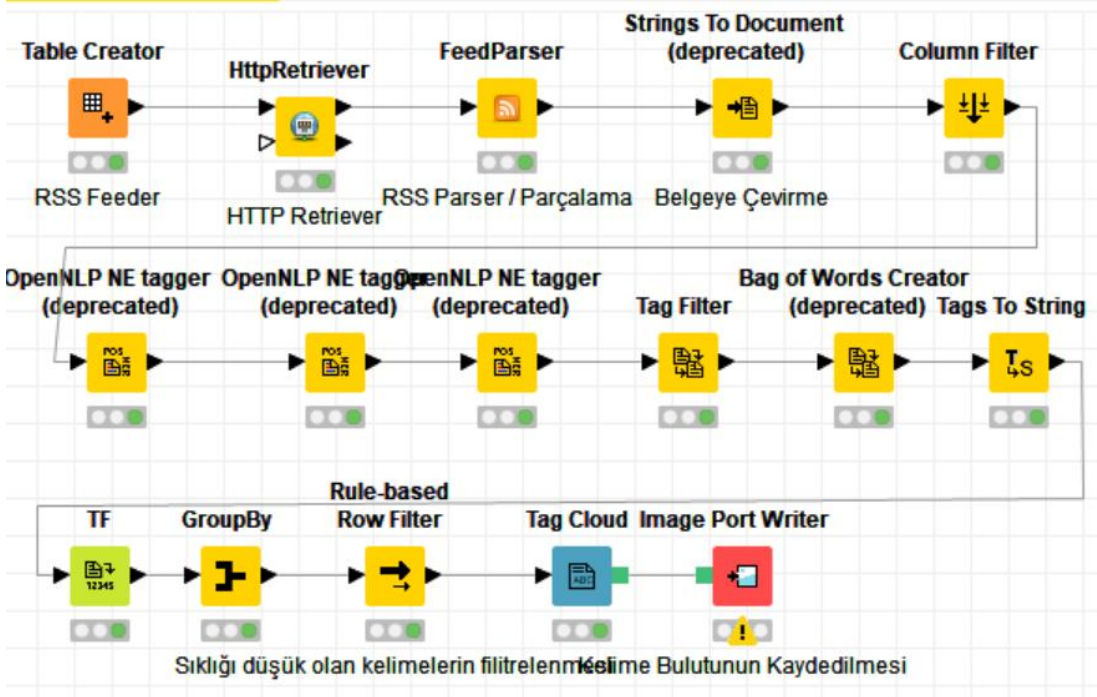
Belirtilen çalışmalara ek olarak, Hu ve Trivedi (2020) metin madenciliği ile otel kullanıcılarının 111,986 yorumunu değerlendirmişlerdir. Söz konusu değerlendirmeye göre otel marka konumlandırması ve rekabet üzerine görselleştirme çalışmasında bulunmuşlardır. Greco ve Polli (2019) duygu metin madenciliği üzerine çalışmışlardır. Söz konusu çalışmada marka yönetimi için müşteri profilleri oluşturulmuştur. İlgili çalışmanın örnekleme spor markasına gelen Twitter kullanıcı yorumları olarak alınmıştır. Loureiro vd. (2019) sanal gerçekliğin pazarlamadaki kullanımı üzerine yaptıkları bir çalışmada metin madenciliğinden yararlanmışlardır. Web of Science indeksinde yayınlanan 115 derginin 150 makalesi değerlendirilmiştir. Yapılan araştırmanın sonucunda en çok atf yapılan yazarlar belirlenmiştir.

3. Uygulama

Çalışmanın amacı, ULAKBİM indeksinde yer alan sosyal bilimler alanındaki dergilerde yayınlanan çalışmaların hangi yöntem, konu ve çerçevelerden oluştuğunu araştırmaktır. Bu bağlamda söz konusu dergilerin son sayılarındaki çalışmalar incelenmiştir. Bu inceleme sonuçlarına göre ULAKBİM indeksinde taranan dergilerde yayın yapmayı hedefleyen akademisyenlere rehberlik edilmesi amaçlanmıştır. Söz konusu amaç doğrultusunda, bu çalışmada ULAKBİM indeksinde taranan sosyal bilimler alanındaki dergilerin son sayıları araştırma kapsamına alınmıştır. ULAKBİM indeksinde yer alan 537 sosyal bilim bulunmaktadır. Bu dergilerden 308 adedinin Dergipark sisteminde aktif olarak yer aldığı tespit edilmiştir. Söz konusu çalışma için bahsi geçen 308 derginin Rich Site Summary (RSS) linkleri temin edilmiştir.

Çalışmanın analiz sürecinde KNIME programından yararlanılmıştır. RSS linkleri üzerinden yayınlanan makalelerin özet kısımları, başlıkları, yayınlama tarihleri ve linkleri program aracılığıyla temin edilmiştir. Daha sonrasında ise özet kısımları dikkate alınarak kelimeler etiketlenmiş ve bu kelimelerin frekans dağılımları elde edilmiştir. KNIME programı kullanılarak oluşturulan iş akışına Şekil 1'de yer verilmiştir.

Şekil 1: KNIME Programının İş Akışı



“Table Creator” yardımıyla RSS linkleri üzerinden “Httpretriever” ve “Feedpaser” modülleri ile programa çekilmiştir. Böylelikle, 308 derginin son sayısında yer alan toplam 4,480 makalenin bilgileri elde edilmiştir. Daha sonrasında ise analize hazır hale getirilmesi için “String To Document” modülü yardımıyla özetler belge formatına çevrilmiştir. Daha sonra bu veriler doğal dil işleme modülü olan “Tagger” yardımıyla, metinlerdeki kelimelerin özelliklerine göre etiketlenmiştir. Daha sonrasında “Bag of Words Creator” modülü ile kelime özellikleri hazırlanmıştır. “TF” modülü ile kelimelerin metinlerde kaçar defa geçtikleri sayılmıştır. “TAG Cloud” yardımıyla da kelime bulutu oluşturulmuştur. Bütün kelimeler analizin anlaşılmasını zorlaştırılacağı için öncesinde az geçen kelimeler filtrelenmiştir. Kelime bulutunun çıktısına Şekil 2’de yer verilmiştir.

Şekil 2: Kelime Bulutu



Öte yandan kelime bulutunda yer alan bu kelimelerin sıklıkları Tablo 1’de paylaşılmıştır. Söz konusu tablodan görülebileceği üzere korelasyon, fark testi, çok kriterli karar verme teknikleri en çok kullanılan kelimelerdir. Buradan yola çıkarak, ULAKBİM indeksindeki sosyal bilimler alanındaki dergilerin son sayılarında araştırma makalelerinin çoğunlukta olduğu yorumu yapılabilmektedir.

Tablo 1: Kelime Frekans Dağılımı

<i>Kelime</i>	<i>Sıklığı</i>
Pearson	80
Mann Whitney U	51
VIKOR	25
TOPSIS	22
FLA	21
Korelasyon	21
OECD	16
TÜİK	14
TÜFE	12
AFA	10
İslam	10
AR	9
CDS	9
DİBS	9
VAR	9
YSA	8
KDV	7
PISA	7
TMK	7

AHP	6
Brand Finance	6
DEMATEL	6
Dickey Fuller	6
EURO	6
FED	6
FRS	6
GSS	6
GSYİH	6
Kendall	6
OI	6
ROA	6
ROK	6

4. Sonuç ve Değerlendirme

Son düzenleme ile doçentlik kriterlerinin zorunlu koşullarından biri ULAKBİM indeksinde taranan dergilerde yayın yapmak olmuştur. Akademisyenlerin de bu amaçla yayınlarını bu dergilere çevirmeleri, söz konusu dergilerin önemini artırmıştır. Bundan dolayı, ULAKBİM indeksinde yer alan dergilerde hangi anahtar kelimeler ön plana çıktığı önem kazanmıştır. Metin madenciliği ile sosyal bilimlerde ULAKBİM indeksinde yer yapan dergilerin son sayıları incelenmiştir. Bu çerçevede bahsi geçen dergilerde yer alan makalelerde en çok tekrarlanan kelimeler dikkate alınmıştır. Analiz sonucunda, ULAKBİM’de taranan dergilerde yayın yapmayı hedefleyen akademisyenlere yol gösterecek sonuçlar elde edilmiştir.

Elde edilen analiz sonuçlarına göre, korelasyon, fark testleri gibi istatistiki yöntemlerin ön plana çıktığı görülmektedir. Bunun yanı sıra, çok kriterli karar verme tekniklerinin bu indekste yayınlanan makalelerde kullanıldığı tespit edilmiştir. Bu çerçevede, VIKOR, TOPSIS ve DEMATEL yöntemlerinin en fazla dikkate alındığı anlaşılmıştır. Söz konusu yöntemler SSCI ve SCI gibi önemli indekslerde de oldukça tercih edilmektedir (Dinçer vd., 2019c,d). Öte yandan, finans ve ilahiyatın en fazla odaklanılan konu olduğu belirlenmiştir. ULAKBİM indeksinde yayın yapan dergilerin belli bir oranının İlahiyat dergileri olmasından dolayı, İslam, Kur’an gibi kelimelerin sık kullanılmış olması beklenen bir sonuçtur. Bununla birlikte finans alanında yapılan çalışmaların sıklığı bu alanın popüleritesinin yüksek olduğu bilgisini vermektedir. Bu çerçevede CDS, DİBS, Brand Finance, EURO, FED, ROA ve ROK gibi kelimeler en fazla geçen kelime listesinde yer almaktadır. Belirtilen bu sonuç dikkate alındığında yapılacak yeni çalışmalarda bu kapsamdaki araştırmalarının şansının yüksek olacağı düşünülmektedir.

Çalışma alanı olarak, genellikle Türkiye çalışmaları olduğu tespit edilmiştir. Türk indeksi olması ve genellikle Türkçe yayın yapıyor olmasından dolayı, çalışmaların uygulama alanı genellikle Türkiye üzerine olmuştur. Panel çalışmalar yapılması bu indekslerde orijinallik sağlayacağı düşünülmektedir. Bu bağlamda, gelişmiş, gelişmekte olan, yükselen ve az gelişmiş ülke grupları üzerinde analizler yapılması yerinde olacaktır. Örnek olarak, E7 ve G7 ülkeleri üzerinde yapılacak yeni bir çalışmanın farklılık yaratacağı düşünülmektedir.

Bununla birlikte ölçek/metot olarak, eğitim çalışmalarında başta olmak üzere çok sık kullanılan PISA ölçeği ön plana çıkmıştır. Veri seti olarak da çalışmalarda TÜİK ve OECD verileri tercih edilen veri setleri olarak ön plana çıkmıştır. Belirtilen bu çalışmada, sadece 2019 yılının son sayıları dikkate alınmıştır. Bundan dolayı, yapılacak yeni bir çalışmada bu kapsamın artırılması ve fen bilimlerinin de dikkate alınması literatüre önemli bir katkı sağlayacaktır.

Kaynakça

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. IEEE Computational intelligence magazine, 9(2), 48-57.

Dinçer, H., Yüksel, S., & Adalı, Z. (2019b). Economic Effects in Islamic Stock Development of the European Countries: Policy Recommendations for Ethical Behaviors. In Handbook of Research on Managerial Thinking in Global Business Economics (pp. 58-78). IGI Global.

- Dinçer, H., Yüksel, S., & Martínez, L. (2019c). Interval type 2-based hybrid fuzzy evaluation of financial services in E7 economies with DEMATEL-ANP and MOORA methods. *Applied Soft Computing*, 79, 186-202.
- Dinçer, H., Yüksel, S., & Martínez, L. (2019d). Analysis of balanced scorecard-based SERVQUAL criteria based on hesitant decision-making approaches. *Computers & Industrial Engineering*, 131, 1-12.
- Dinçer, H., Yüksel, S., Eti, S., & Tula, A. (2019a). Effects of Demographic Characteristics on Business Success: An Evidence From Turkish Banking Sector. In *Handbook of Research on Business Models in Modern Competitive Scenarios* (pp. 304-324). IGI Global.
- Ding, Z., Li, Z., & Fan, C. (2018). Building energy savings: Analysis of research trends based on text mining. *Automation in Construction*, 96, 398-410.
- Dolgun, M. Ö., Özdemir, T. G., & Oğuz, D. (2009). Veri madenciliğinde yapısal olmayan verinin analizi: Metin ve web madenciliği. *İstatistikçiler Dergisi: İstatistik ve Aktüerya*, 2(2), 48-58.
- Eskici, H. B., & Koçak, N. A. (2018). A text mining application on monthly price developments reports. *Central Bank Review*, 18(2), 51-60.
- Eti, S. (2016). Üniversitelerdeki Akademik Üretkenliğe Etki Eden Faktörlerin İncelenmesi. *İş'te Davranış Dergisi*, 1(1), 67-73.
- Greco, F., & Polli, A. (2019). Emotional Text Mining: Customer profiling in brand management. *International Journal of Information Management*.
- Hu, F., & Trivedi, R. H. (2020). Mapping hotel brand positioning and competitive landscapes by text-mining user-generated content. *International Journal of Hospitality Management*, 84, 102317.
- Kim, S. G., & Kang, J. (2018). Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews. *Information Processing & Management*, 54(6), 938-957.
- Loureiro, S. M. C., Guerreiro, J., Eloy, S., Langaro, D., & Panchapakesan, P. (2019). Understanding the use of Virtual Reality in Marketing: A text mining-based review. *Journal of Business Research*, 100, 514-530.
- Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness.
- Pınarbaşı, F., & Canbolat, Z. N. (2018). Evaluation Of Augmented Reality Mobile Applications In Turkey Market: A Data Mining Approach To Consumer Reviews. *Changing Organizations*, 187.
- Seker, S. E. (2015). Metin Madenciliği (Text Mining). *YBS Ansiklopedi*, 2(3), 30-32.
- Shen, C. W., Chen, M., & Wang, C. C. (2018). Analyzing the trend of O2O commerce by bilingual text mining on social media. *Computers in Human Behavior*.
- Tan, A. H. (1999, April). Text mining: The state of the art and the challenges. In *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases* (Vol. 8, pp. 65-70). sn.
- Tobback, E., Naudts, H., Daelemans, W., de Fortuny, E. J., & Martens, D. (2018). Belgian economic policy uncertainty index: Improvement through text mining. *International journal of forecasting*, 34(2), 355-365.
- Yüksel, S., Dinçer, H., & Emir, S. (2018). Analysis of Service Innovation Performance in Turkish Banking Sector Using a Combining Method of Fuzzy MCDM and Text Mining. *MANAS Sosyal Araştırmalar Dergisi*, 7(3), 479-504.