

1964-2022 Yılları Arasında İşletme Ana Bilim Dalı'nda Hazırlanan Tezlerin Gizli Dirichlet Tahsisi Yöntemi ile Konu Modellemesi ¹

Mehmet ÖZÇALICI ²

Başvuru Tarihi: 22.08.2022

Kabul Tarihi: 27.01.2023

Makale Türü: Araştırma Makalesi

Öz

Bilimsel bilginin üretilmesini sağlayan en önemli araçlar üniversitelerde hazırlanan lisansüstü tezlerdir. Bilim dallarında hazırlanan lisansüstü tezlerin içerik, bibliyometrik veya konu analizine tabi tutulması, söz konusu bilim dalındaki gelişim düzeyinin tespit edilmesi açısından önem arz etmektedir. Bu çalışmanın amacı, değişen çevre koşullarına göre yeniden şekillenmesi kaçınılmaz olan, İşletme Ana Bilim Dalı'nda yayınlanan tezlerin konu modellemesini gerçekleştirmektir. Çalışmada, 1964-2022 yılları arasında Yüksek Öğretim Kurulu Başkanlığı Tez Merkezi'ne kayıtlı ve İşletme Ana Bilim Dalı'nda hazırlanmış olan 29.381 tane lisansüstü tezin özet kısmı kullanılmıştır. Özetlerin analiz edilmesi için Gizli Dirichlet Tahsisi analizinden yararlanılmıştır. Tez özetlerine ait tanımlayıcı istatistikler tablolar halinde sunulmuştur. Sonuçta, Sosyal Bilimler Enstitüsü'nde 2020 ve 2021 yıllarında yayınlanan tezlerin %15'inin İşletme Ana Bilim Dalı ile ilgili olduğu tespit edilmiştir. Analiz sonucunda tezlerin; muhasebe, pazarlama, finans ve yönetim olmak üzere dört farklı konuda incelenebileceği ve 2021 yılında tezlerde en sık kullanılan kelimelerin; analiz, sonuç, örgüt, veri ve ilişki olduğu belirlenmiştir. 1995 ile 2012 yılları arasında muhasebe ile ilgili, 2013 yılından 2022 yılına kadar ise yönetim konusunda yayınlanan tezlerin sayısının baskın olduğu sonucu ortaya çıkmaktadır.

Anahtar Kelimeler: Gizli Dirichlet Tahsisi, İşletme, Konu Modellemesi, Lisansüstü Tez, Kelime Bulutu

Atıf: Özçalıcı, M. (2023). 1964-2022 yılları arasında İşletme Ana Bilim Dalı'nda hazırlanan tezlerin Gizli Dirichlet Tahsisi yöntemi ile konu modellemesi. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 23(1), 163-188.

¹ Bu çalışma etik kurul izin belgesi gerektirmemektedir.

² Kilis 7 Aralık Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Uluslararası Ticaret ve Lojistik Bölümü, mozcatici@kilis.edu.tr, ORCID: 0000-0003-0384-6872

Topic Modeling of Business Administration Dissertations (1964-2022) with Latent Dirichlet Allocation

Mehmet ÖZÇALICI³

Submitted by: 22.08.2022

Accepted by: 27.01.2023

Article Type: Research Article

Abstract

The most important tools that enable the development of scientific knowledge and its dissemination to society are postgraduate theses and dissertations prepared at universities. Subjecting the dissertations prepared in the disciplines to content analysis, bibliometric analysis or topic modeling is important in determining the discipline's level of development. This study aims to perform the topic modeling of the dissertations published in the Department of Business Administration. In the study, the abstracts of 29,381 dissertations registered in the Thesis Center of the Council of Higher Education between 1964-2022 and prepared in the Department of Business Administration were used. Latent Dirichlet Allocation analysis was used to analyze the abstracts. Descriptive statistics of dissertations summaries are presented in tables. As a result, it was determined that 15% of the theses published in the Social Sciences Institute in 2020 and 2021 were related to the Department of Business Administration. Basic subjects are determined as Accounting, Marketing, Finance, and Management. It has been determined that words such as analysis, result, organization, data and relationship are being used more frequently in 2021. The most used words are visualized with the help of word clouds according to universities, thesis advisors, and subjects. It was determined that the theses prepared in the field of management were dominant from 2013 to 2021, and the theses published on accounting were dominant between 1995 and 2012.

Keywords: Latent Dirichlet Allocation, Business Administration, Topic Modeling, Dissertation, Word Cloud

³ Kilis 7 Aralık University, Faculty of Economics and Administrative Sciences, Department of International Trade and Logistics, mozcalici@kilis.edu.tr, ORCID: 0000-0003-0384-6872

Giriş

Üniversitelerde hazırlanan tezler yardımıyla, bilimsel bilginin üretilmesi ve topluma yayılması sağlanmaktadır. Bilim dallarında hazırlanan lisansüstü tezleri içerik, bibliyometrik veya konu analizine tabi tutmak suretiyle ilgili bilim dalındaki gelişim düzeyi takip edilebilmektedir.

Türkiye’de Sosyal Bilimler Enstitüsü’nde İşletme Ana Bilim Dalı ile ilgili 1964 yılında lisansüstü tezler hazırlanmaya başlamıştır. Her ne kadar fakültelerin kendilerine ait vizyon ve misyon tanımları olsa da genel anlamda, işletme eğitiminin amacı, özel ve kamu sektörünün ihtiyaç duyduğu, olay ve olguları sorgulayabilme, karar verme, sorun çözme, planlama yapabilme, ekip çalışmasına yatkın olabilme, örgütleyebilme, düzenleyebilme ve denetleyebilme gibi niteliklere sahip yönetici adaylarını yetiştirmektir (Düzakın ve Yılmaz, 2009; Özkul, 2012a, 2012b). Bu çalışmada yer alan Tablo 1’e göre Sosyal Bilimler Enstitüsü’nde 2020 ve 2021 yıllarında hazırlanan tezlerin %15’i işletme alanındadır ve bu bulgu, işletme eğitiminin enstitü içerisindeki ağırlığını ortaya çıkarmaktadır.

İşletme eğitimi ve işletme eğitiminin önemi hakkında araştırmalar gerçekleştirilmiştir. Aysan (2005), geleneksel işletme yönetimi tekniklerinin, işletme ve çevresi hakkında elde edilebilen sınırlı bilgilerle karar verilmesinin zorunlu olduğu ortamlara göre hazırlanmış bulunduğunu ifade etmektedir. Yazara göre günümüz bilgisayar teknolojisindeki gelişme işletme eğitiminin de değişimini zorunlu kılmaktadır. İşletmenin çevre ile yakın ilişki içinde olduğu ve işletmecilik eğitimi-öğretimi gerçekleştiren kurumların, söz konusu çevreye hızlı bir şekilde uyum sağlaması gerekmektedir (Aydınlı ve Halis, 2004). Değişen çevre koşulları işletme bölümü öğrencilerinin değişen çevre şartlarına göre yetiştirilmesini zorunlu kılmaktadır (Sargut, 2009). Türkiye’deki işletmecilik eğitime yönelik eleştiriler ve çevre şartlarının değişmesi ile işletme eğitiminin de bu şartlara uygun olarak yeniden yapılandırılmasına yönelik görüşler Bilginoglu (2018)’ nun çalışmasında yer almaktadır.

Özkul (2012b), tezinde Türkiye’de işletmecilik yükseköğretimi hakkında detaylı bir araştırma gerçekleştirmiştir. Benligiray ve Tez (2011) ise 2011 yılı içerisinde İşletme eğitimi veren programların veya bölümlerin web sayfalarında yer alan tanıtımları, misyon ve vizyon ifadeleri, değerleri, eğitim hedefleri ve amaçları, bölüm başkanlarının mesajları, bölüme özgü yeterlilikleri ve bölümün öğrenme çıktılarını incelemişlerdir. Sonuçta 2011 yılı için durum değerlendirmesi gerçekleştirilmiştir. Güvemli (2018) ise Türkiye’de işletmecilik eğitiminin tarihçesine değinmiş ve genel durum değerlendirmesi gerçekleştirmiştir. Üsdiken ve Erden (2001), 1970’li ve 1990’lı yıllarda yönetim ile ilgili dört tane akademik dergide (Amme İdaresi Dergisi, Eskişehir Anadolu Üniversitesi İİBF Dergisi, İstanbul Üniversitesi İşletme Fakültesi Dergisi ve ODTU Gelişme Dergisi) yayınlanan makele içeriklerini kıyaslamış ve Türkiye’deki yönetim yazınındaki değişikliği detaylı bir şekilde incelemiştir. Berkman (1987) ise, Amme İdaresi Dergisi’nde yirmi yıl içinde yayınlanan 503 tane makeleyi sınıflandırmış ve detaylı incelemeye tabi tutmuştur. Üsdiken ve Erçek (2009), Türkiye’de işletme konularına yönelik ilk dergi olarak kabul edilen İşletme – İşletme Ekonomisi ve Organizasyon Mecmuası adlı dergide 1943 – 1948 yılları arasında yayınlanan makaleleri incelemiştir.

Yukarıda özetlenen çalışmalar ancak uzman/insan desteği ile mümkün olabilmektedir. Bunun yanı sıra bilgisayar teknolojisindeki gelişmeleri kullanmak suretiyle, daha yüksek hacimdeki metinleri analiz edebilmek mümkün olabilmektedir. Gizli Dirichlet Tahsis (LDA) analizi popüler bir konu modelleme tekniğidir ve bir yığın metinde yer alan konuların tespiti hedeflenmektedir. Bu çalışmanın amacı, değişen çevre koşullarına göre yeniden şekillenmesi kaçınılmaz olan, İşletme Ana Bilim Dalı (İABD)’nda yayınlanan tezlerin konu modellemesini, Gizli Dirichlet Tahsisi (Latent Dirichlet Allocation – LDA) analizi yardımıyla gerçekleştirmektir. İşletme ile ilgili hazırlanan tezlerin konu modellemesi suretiyle, tezlerin hangi ana başlıklar altında hazırlandığı belirlenebilecektir. Bu durum gelecekte yazılacak tezlere de yol gösterici nitelikte olacaktır. Bu amaçla, 1964-2022 yılları arasında Yüksek Öğretim Kurulu Başkanlığı Tez Merkezi (www.tez.yok.gov.tr)’ne

kayıtlı ve İşletme Ana Bilim Dalı'nda yayınlanmış lisansüstü tezlerin özetleri bir araya getirilmiştir. Yayınlanan tezlerde popüler olan kelimeler ve konular belirlenmiş, yüksek hacimdeki veri seti tanımlayıcı analizler yardımıyla incelemeye tabi tutulmuştur.

Çalışmada cevap aranan araştırma soruları (AS) aşağıdaki gibidir:

- AS01: Yıllar itibariyle yayınlanan yüksek lisans ve doktora tez sayıları nedir? Söz konusu yıllar içinde Sosyal Bilimler Enstitüsü'nde yayınlanan toplam tez sayısı nedir?
- AS02: Lisansüstü tezlerin yayınlandıkları yıllara göre dağılımı nasıldır?
- AS03: Yüksek lisans ve doktora tezlerinde sayfa sayılarının frekans dağılımı nasıldır? Özetlerde geçen kelime sayılarının frekans dağılımı nasıldır?
- AS04: Yıllara göre üniversitelerde yayınlanan yüksek lisans ve doktora tezi sayıları nelerdir?
- AS05: Danışman sayılarına göre tezlerin dağılımı nasıldır?
- AS06: LDA ile belirlenen konularda yıllar itibariyle yayınlanan tez sayıları nedir?
- AS07: 2021 yılında tezlerde kullanılan kelimelerin frekans dağılımı nasıldır?

Sıralanan araştırma sorularının cevaplanması ile birlikte İşletme Ana Bilim Dalı (İABD)'nda hazırlanmış lisansüstü tezler hakkında tanımlayıcı bilgiler açığa çıkmaktadır. Yıllar itibariyle hangi konularda daha fazla tez yazıldığı da belirlenebilmektedir.

Çalışma beş bölümden oluşmaktadır. Bu giriş bölümünden sonra ikinci bölümde LDA analizi ve bibliyometrik analiz ile ilgili yayınlanan çalışmalar incelenmiş, üçüncü bölümde ise LDA'nın çalışma prensipleri özetlenmiştir. Dördüncü bölümde veri seti tanıtılmış, tanımlayıcı istatistiklere yer verilmiş ve LDA analiz sonuçları sunulmuştur. Beşinci bölümde bulgular tartışılmış, önerilen tekniğin avantajlı ve dezavantajlı tarafları sunulmuş ve sonraki çalışmalar için önerilere yer verilmiştir.

Literatür Taraması

Bibliyometrik analiz yardımıyla, akademik çalışmalar analize tabi tutulmaktadır. Bibliyometrik analizde, çalışma sayıları, referans sayıları, atıf sayıları, yayın niteliği, mali destek alıp almadıklarına göre incelenmekte ve araştırmacılar için faydalı olacak durum değerlendirilmekte gerçekleştirilmektedir. Bibliyometrik çalışmalarda en sık kullanılan yazılım VOSviewer' dır. Bunun haricinde CiteSpace (Coşkun ve Kazan, 2021) ve Bibexcel (Sipahi, 2020) gibi yazılımlar da kullanılmaktadır. Bibliyometrik çalışma örnekleri aşağıdaki gibi özetlenebilir.

Toplu Yaşlıoğlu ve arkadaşları (2018), Web of Science'da indekslenen ve Türkiye adresli stratejik yönetim ile ilgili makalelerin bibliyografik analizini gerçekleştirmiştir. Toplamda 126 tane makale kullanılmıştır. Durum değerlendirmesi yapılmış ve araştırmacılar için tavsiyelerde bulunulmuştur. Alkan ve Özkaya (2015), Türk finans ve muhasebe yazınının Web of Science'da endekslenen dergiler kapsamında genel bir görünümü ortaya koymak amacıyla, 75 farklı dergide toplamda 398 tane finans ve muhasebe alanında yayınlanmış akademik çalışmayı incelemeye tabi tutmuştur. Sakin (2008), ise çalışmasında, İstanbul Üniversitesi İşletme Fakültesi Dergisi'nde yayınlanan ve muhasebe ve finans alanındaki çalışmalar için içerik analizi gerçekleştirmiştir. Çalışma sonunda finans ile ilgili çalışmaların sayısının artmaya başladığı raporlanmıştır. Kıymetli Şen ve çalışma arkadaşları (2017), Türkiye'de yayınlanan muhasebe dergileri üzerinde bibliyometrik araştırma gerçekleştirmiştir. 852 tane akademik yayın çalışma kapsamında değerlendirilmiştir. Hotamışlı ve Erem (2014) ise, Muhasebe ve Finansman Dergisi'nde yayınlanan makalelerin bibliyografik analizini gerçekleştirmişlerdir. Alkan (2014), çalışmasında, Türkiye'deki üniversitelerde muhasebe alanında yapılmış 656 tane lisansüstü tezin; profilini, alana katkılarını, içerik ve konu açısından zaman içindeki yönelimini belirlemiştir.

Yukarıda özetlenen çalışmalar her ne kadar detaylı bir durum değerlendirmesine olanak sağlasa da, çalışmada yer alan konuların (içerik) değerlendirmesi hususunda eksik kalmaktadır. LDA analizi yardımıyla bu eksik olan taraf güçlendirilebilmektedir. Aşağıda LDA analizini konu ile ilgili alanlarda uygulayan çalışmalar özetlenmiştir.

Borah ve arkadaşları (2018), pazarlama kavramının etkisini araştırmışlardır. Bu amaçla Harvard Business Review dergisinin 94 yıl, Sloan Management Review dergisinin 46 yıl ve Management Science dergisinin 47 yılını kapsayan özetleri bir araya getirilmiş ve LDA analizine tabi tutulmuştur. Konu sayısını 6 olarak belirlemişlerdir. Sonuçta pazarlama kavramının çalışmalarda önemli bir paya sahip olduğu ve bu önemin de giderek artacağı raporlanmaktadır.

Ekinci ve arkadaşları (2020) Tıp veri kümesi için Gizli Dirichlet Ayrımı analizini kullanmışlardır. 11 yılı (2007-2017 yılları arası) kapsayan ve PubMed veri tabanında yayınlanan makaleler (152.225 tane) veri seti olarak kullanılmıştır. Önerilen sistemin, literatürü otomatik olarak taramada kullanılabilirliğini ifade etmektedirler. Konu sayısını 20 olarak belirlemişlerdir. Hangi yıllarda hangi konuların çalışıldığı belirlenmiştir.

Şimdi ve Garip (2021) ise çalışmalarında, 2000 ile 2019 yılları arasında Web of Science'da indekslenen ve uluslararası ticaret ile ilgili olan akademik çalışmaların özetlerini (5001 tane) LDA analizi ile birlikte kullanmışlardır. İncelenen dönemi kriz öncesi ve kriz sonrası olmak üzere iki bölüme ayırmışlardır. Gelecek yıllarda, ticaret ile ilgili dergilerde hakim olacak konular hakkında tahminlerde bulunmuşlardır.

Çallı ve arkadaşları (2021), yönetim bilişim sistemleri alanında yayınlanmış yüksek lisans tezlerinin, LDA analizi yardımıyla konu modellemesini gerçekleştirmişlerdir. 2002-2020 yılları arasında söz konusu alanda yayınlanmış 574 tane lisansüstü tez incelemeye tabi tutulmuştur. Yazarlar, tezlerin özetlerini 11 konuda toplamışlardır.

Omurca ve arkadaşları (2021), müşteri şikayetleri ile ilgili topladıkları veri setine LDA'yı da içeren üç farklı konu modelleme tekniğini uygulamışlardır. Konu sayısı 100'de sabitlenmiştir. 6963 şikayet incelenmiştir. Sonuçta, şikayetlerin gruplandırılması ile birlikte şikayetlerin daha tutarlı bir şekilde yönetilebileceği raporlanmıştır.

Ekin ve arkadaşları (2021), eğitim alanında yer alan akademik dergilerde COVID19 ile ilgili yayınlanan çalışmaların profilini ortaya çıkarmışlardır. Ocak 2020 ile Mayıs 2021 tarihleri arasında SCOPUS veri tabanı tarafından indekslenen 3039 makale LDA ve tanımlayıcı istatistikler ile değerlendirmeye tabi tutulmuştur. 10 tane konu belirlenmiştir.

Tokdemir (2021) çalışmasında, SCOPUS veri tarafından indekslenen ve deneysel yazılım mühendisliği alanında yayınlanan 10.658 makale özetini kullanmıştır. Konu sayısı 12 olarak belirlenmiştir. Maddi olarak desteklenen çalışmaların hangi konularda yayınlandığı araştırılmıştır.

Onan ve arkadaşları (2020), üniversite bilgi yönetim sistemine yapılan servis destek taleplerinin konu modelleme tabanlı analizini gerçekleştirmişlerdir. Konu sayısı 4 ve 8 olmak üzere iki model uygulanmıştır. Kullanılan veri setinde 17.831 kayıt yer almaktadır. Sonuçta ortaya çıkan konuların birbirleriyle ne düzeyde iç içe geçtikleri ve uyumlu oldukları belirlenmiştir.

Şahin ve Bayazıt (2020), Stack Overflow sitesinde Java programlama dili ile ilgili sorulmuş soruları ve cevapları veri seti olarak kullanmış ve LDA analizini gerçekleştirmişlerdir. Konu sayısı 7 olarak sabitlenmiştir. Veri setinde 1.606.056 tane gönderi (metin) yer almaktadır. Yıllar itibariyle konularda yayınlanan gönderilerin değişimi tartışılmıştır.

Gürcan ve Özyurt (2020), çalışmalarında e-öğrenme alanında yapılan çalışmaları LDA yöntemi ile konu modelleme analizine tabi tutmuştur. 27.735 tane makale üzerinde çalışmalarını gerçekleştirmişlerdir. 42 tane konu belirlenmiştir. Söz konusu 42 konu daha sonra 5 ana grupta toplanmıştır.

Güven ve arkadaşları (2019), çalışmalarında n-adımlı LDA analizini kullanmak suretiyle Türkçe yazılan tweetlerde duygu tespiti gerçekleştirmişlerdir. Sonuçta, %81,5' e varan performans değeri ile tweetlerdeki duyguları belirleyen modeller geliştirmişlerdir.

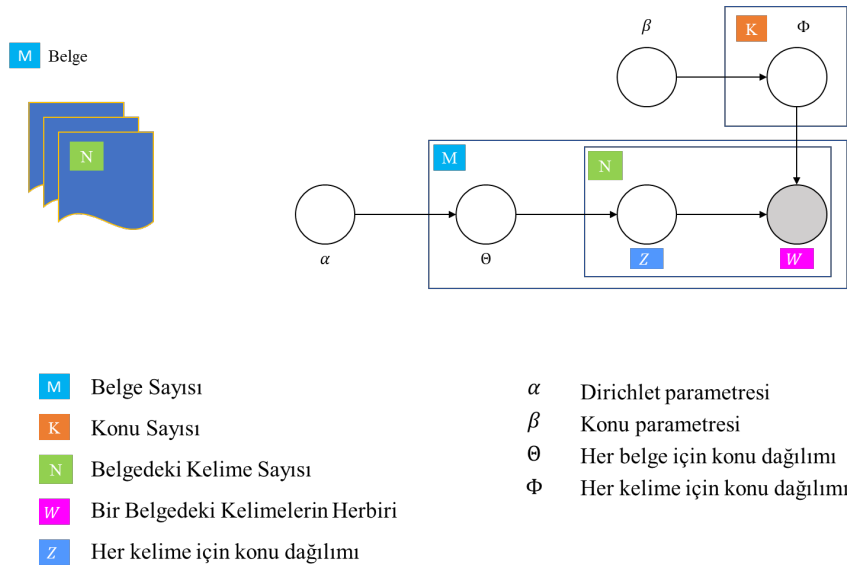
Wang ve arkadaşları (2018), benzer özelliklere sahip iki tane bilgisayar faresi ile ilgili yapılan müşteri yorumlarına ait metinlerin LDA ile konu modellemesini gerçekleştirmişlerdir. Perplexity değeri daha yüksek sayıdaki konulara işaret etse de, yazarlar konu sayısını 10'da sabitlemişlerdir. Bir ürün için 10.014 ve diğeri için de 9.520 tane müşteri yorumu yer almaktadır. Pazar konumlandırma için yöneticilere faydalı olabilecek sonuçların ortaya çıktığı raporlanmıştır.

Yukarıda özetlenen çalışmalar dikkate alındığında, LDA analizinin başarılı bir şekilde konu modellemesinde kullanıldığı belirlenmektedir. Ne var ki, henüz Türkçe ve İşletme Ana Bilim Dalı (İABD)'nda yayınlanan lisansüstü tezlerin konu modellemesi LDA yardımıyla gerçekleştirilmemiştir. Bu çalışma ile birlikte literatürde yer alan bu açığın kapanması hedeflenmektedir.

Gizli Dirichlet Tahsisi (Latent Dirichlet Allocation – LDA) Yöntemi

Gizli Dirichlet Tahsisi (Latent Dirichlet Allocation – LDA) yöntemi ilk olarak Blei ve arkadaşları tarafından 2003 yılında literatüre tanıtılmıştır (Blei vd., 2003). Bir dizi belgede gömülü bulunan soyut konuların belirlenmesinde kullanılan istatistiksel bir yöntemdir (Wang vd., 2018). Her belge için konu dağılımı ve her konu için de kelime dağılımı söz konusu olacaktır (Güven, 2021). LDA yöntemi, konu modelleme alanında en bilinen algoritmalarından bir tanesidir (Ekin vd., 2021). LDA analizinde belgeler, farklı konuların bir bileşeni olarak görülmektedir (Calvo vd., 2018).

LDA yönteminde modellenecek olan veri kümesi külliyyat (corpus), külliyyat içindeki her bir öge belge (document), belge içerisindeki her bir kelime ise terim (term) olarak adlandırılmaktadır (Çalli vd., 2021). LDA yöntemi, gizli içerikler (konu) yardımıyla külliyyatın kapsamlı bir temsilini ortaya çıkarmayı amaçlamaktadır. Konu sayısının araştırmacı tarafından belirlenmesi gerekmektedir. Analize başlamadan önce konular bilinmediğinden, başlangıçta konular ile ilgili herhangi bir gözlem yapılamamaktadır. LDA süreci Şekil 1'de özetlenmeye çalışılmıştır.



Şekil 1. LDA'nın genel yapısı (Ekin vd., 2021)

LDA yöntemi sonucunda kelime-konu (Φ) ve belge-konu (Θ) olmak üzere iki matris ortaya çıkmaktadır (Şekil 1) ve bu matrisler yardımıyla her bir belgenin hangi konuya ait olduğu belirlenebilmektedir. Modellemede kullanılmamış olan bir belgenin de hangi konuya ait olduğu yine bu matrisler yardımıyla atanabilir.

Analizin avantajlı ve dezavantajlı tarafları mevcuttur (Maier vd., 2018). LDA'nın avantajları; büyük hacimdeki veri setinde yer alan konuların hızlı bir şekilde belirlenmesini sağlamaktadır. Gerçek hayata uygun olarak, bir belge birden fazla konuyu içerebilme olanağı mevcuttur. Dezavantajlı tarafları ise, sonuçların deterministik olmaması, konu sayısının araştırmacı tarafından belirlenmesi zorunluluğu olarak sıralanabilir.

Analiz

Veri Seti

4 Mayıs 2022 tarihinde tez.yok.gov.tr adresinde "İşletme Ana Bilim Dalı"nda yayınlanan tezler Yüksek Öğretim Kurulu Başkanlığı Tez Merkezi (www.tez.yok.gov.tr) adresinden sorgulanmıştır. Site söz konusu tarihte tek seferde sadece 2.000 tane sonucun gösterimini gerçekleştirebilmektedir. Örneğin Marmara Üniversitesi'nde Sosyal Bilimler Enstitüsü'nde İşletme Ana Bilim Dalı'nda 2.843 tane tez bulunmaktadır ancak söz konusu sorgu gerçekleştirildiğinde sadece 2.000 tane sonuç görülebilmektedir. Geri kalan 843 tane sonuç listelenmemektedir. İlk olarak üniversiteler bazında tarama gerçekleştirilmiş ancak sınırlı sayıda sorgunun görülmesi problemi ile karşılaşmış, sorgu yarıda bırakılmış ve sorguya yıllar itibariyle tarama yapmak suretiyle yeniden başlanmıştır. 2018, 2019 ve 2021 yıllarında yine 2.000'den fazla tez yazıldığı belirlenmiştir. Bu durumda sadece söz konusu yıllar için ayrıca yüksek lisans ve doktora tezleri için iki farklı sorgu gerçekleştirilmiş ve böylelikle incelenen yıllar içerisinde yayınlanan bütün tezlerin özet kısımlarına erişim mümkün hale gelmiştir. Toplamda 31.603 tane lisansüstü teze ilişkin özetler siteden edinilmiştir.

Tanımlayıcı Bulgular

Tablo 1'de yıllar itibariyle Türkiye'de "İşletme Ana Bilim Dalı"nda gerçekleştirilen tezlerin sayılarının dağılımı yer almaktadır (AS01: *Yıllar itibariyle yayınlanan yüksek lisans ve doktora tez sayıları nedir? Söz konusu yıllar içinde Sosyal Bilimler Enstitüsü'nde yayınlanan toplam tez sayısı nedir?*). Yıllar itibariyle tez sayılarında bir artış gözlenmektedir. Yüksek lisans sayılarının yıllara göre dalgalanma sergilediği dikkat çeken bir unsurdur. Bunun yanı sıra doktora tez sayılarındaki dalgalanma yüksek lisans tez sayılarında olduğu kadar keskin değildir. Tablo

1’de, incelenen yıllar içinde Sosyal Bilimler Enstitüsünde (SBE) yayınlanan tez sayıları da yer almaktadır. “SBE’de yayınlanan tezlerin yüzde kaçını İşletme Ana Bilim Dalı (İABD)’na aittir?” sorusunun cevabı da yine tabloda yer almaktadır. Oranda dalgalanmalar mevcuttur. Örneğin 1971 yılında SBE’de 8 tane tez yayınlanmış ve bunun 7 tanesi İABD ile ilgilidir. Bu oran incelenen yıllar içindeki en yüksek orandır. 1985 yılından 1996 yılına kadar SBE’deki tezler içinde İABD’nin oranı tek hanelidir. 1997 yılından itibaren oran iki haneye çıkmıştır. Son iki yılda (2020-2021) tezlerin %15’i, İABD’na aittir.

Tablo 2

Yıllara Göre İABD’nda Yayınlanan Doktora ve Yüksek Lisans Tez Sayıları Dağılımı

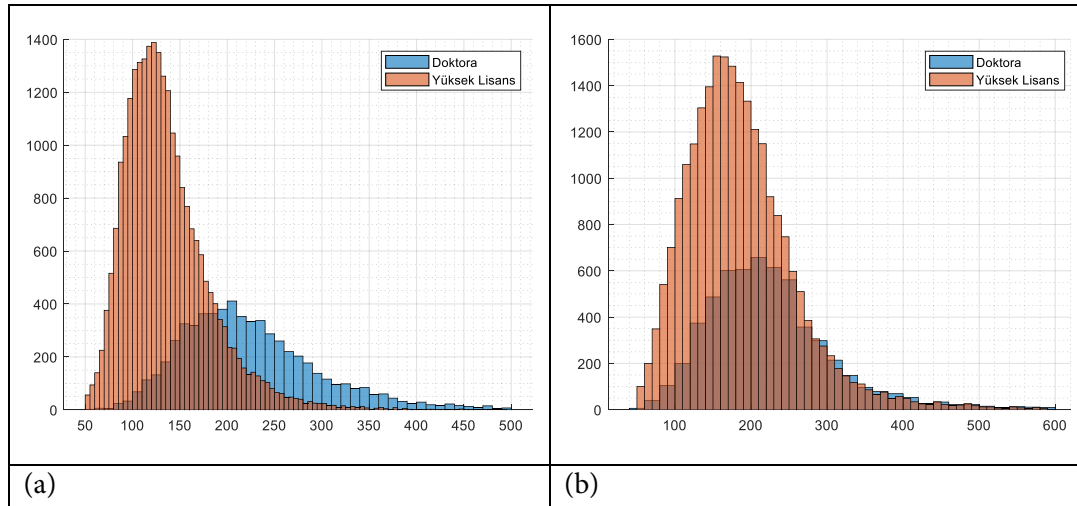
Yıllar	YL	Doktora	Toplam	SBE	SBE (%)	Yıllar	YL	Doktora	Toplam	SBE	SBE (%)
1964	1	0	1	4	25	1996	118	19	137	3799	3,61
1968	0	1	1	5	20	1997	347	52	399	3514	11,35
1971	6	1	7	8	87,50	1998	509	75	584	3896	14,99
1972	3	1	4	11	36,36	1999	451	100	551	3888	14,17
1973	8	8	16	36	44,44	2000	348	58	406	3315	12,25
1974	24	4	28	36	77,78	2001	477	59	536	4284	12,51
1975	11	3	14	36	38,89	2002	645	110	755	5162	14,63
1976	4	4	8	26	30,77	2003	655	122	777	4968	15,64
1977	1	2	3	18	16,67	2004	720	129	849	5233	16,22
1978	2	4	6	33	18,18	2005	633	130	763	5439	14,03
1979	1	2	3	34	8,82	2006	919	123	1042	7559	13,78
1980	0	4	4	33	12,12	2007	940	192	1132	7557	14,98
1981	0	9	9	64	14,06	2008	659	189	848	7111	11,93
1982	2	6	8	72	11,11	2009	718	196	914	7375	12,39
1983	4	4	8	89	8,99	2010	1095	199	1294	9937	13,02
1984	20	10	30	197	15,23	2011	697	223	920	7676	11,99
1985	21	18	39	427	9,13	2012	621	227	848	6881	12,32
1986	24	11	35	611	5,73	2013	818	229	1047	7846	13,34
1987	29	5	34	835	4,07	2014	1155	276	1431	9179	15,59
1988	47	5	52	1255	4,14	2015	1305	328	1633	10132	16,12
1989	47	7	54	1492	3,62	2016	1313	376	1689	10654	15,85
1990	53	14	67	1751	3,83	2017	1427	428	1855	12172	15,24
1991	25	4	29	1835	1,58	2018	1593	430	2023	15194	13,31
1992	27	7	34	1990	1,71	2019	3576	569	4145	29772	13,92
1993	79	12	91	2117	4,30	2020	1348	551	1899	12611	15,06
1994	71	18	89	2629	3,39	2021	1539	527	2066	13211	15,64
1995	58	5	63	3009	2,09	2022	243	80	323		

Veri setinde yer alan tezlerin dillerine ve tez türlerine göre dağılımları Tablo 2’de sunulmuştur (AS02: *Lisansüstü tezlerin yayınlandıkları dillere göre dağılımı nasıldır?*). İngilizce veya diğer yabancı dillerde hazırlanan tezler bulunsa da, tezlerin önemli bir kısmı Türkçe dilinde hazırlanmıştır. Türkçe hazırlanan 29.381 tane tez bu çalışma kapsamında değerlendirmeye alınmış geri kalan (yabancı dillerde hazırlanmış olan) 2.222 (=31.603-29.381) tane tez ise kapsam dışı bırakılmıştır.

Tablo 2
Tez Dili ve Tez Türü Dağılımı

Dil	Yüksek Lisans	Doktora	Toplam
Türkçe	23.644	5.737	29.381
İngilizce	1.704	426	2.130
Almanca	70	0	70
Fransızca	13	0	13
Kırgızca	6	3	9
Toplam	25.437	6.166	31.603

Şekil 2’de tezlerin sayfa sayıları (a) ve özetlerde yer alan kelime sayılarının (b) dağılımı yer almaktadır (AS03: *Yüksek lisans ve doktora tezlerinde sayfa sayılarının frekans dağılımı nasıldır? Özetlerde geçen kelime sayılarının frekans dağılımı nasıldır?*). Beklendiği gibi doktora tezlerinin sayfa sayıları yüksek lisans tezlerinden fazladır. Özetlerdeki kelime sayılarında ise yine yüksek lisans tezlerindeki kelime sayılarının daha az olduğu ortaya çıkmaktadır.



Şekil 2. Tezlerin sayfa sayılarının (a) ve özetlerdeki kelime sayılarının dağılımı (b)

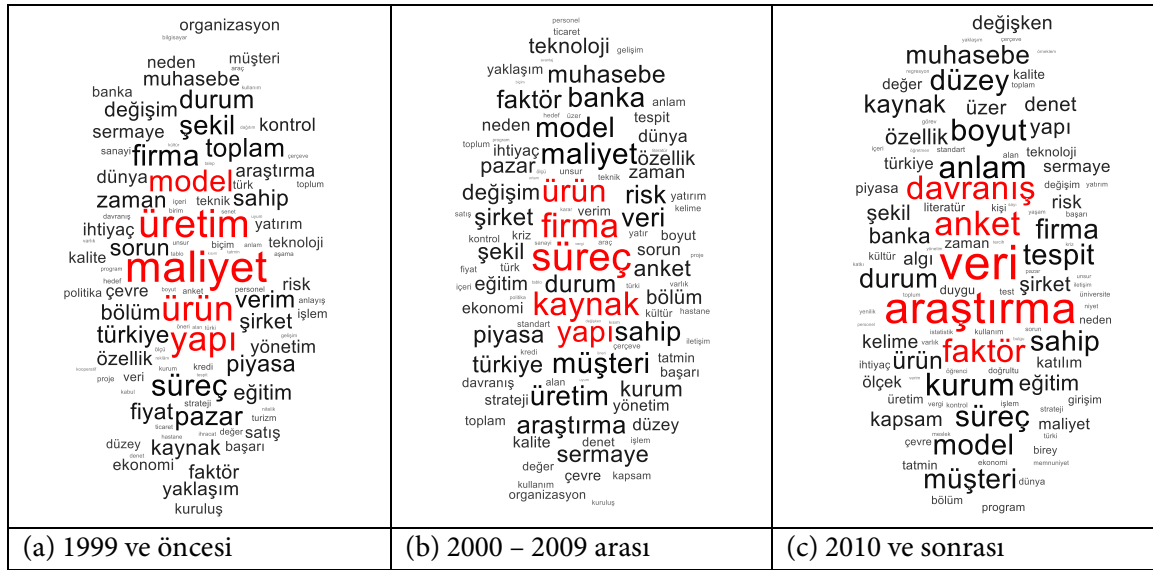
Lisansüstü tezlerin sayfa sayıları ve özetlerindeki kelime sayılarına ilişkin detaylı tanımlayıcı istatistikler Tablo 3’de sunulmuştur. Yapılan incelemede bazı yazarların tezlerini yanlışlıkla iki katı hacimde sisteme girdikleri belirlenmiştir. Örneğin sayfa sayısı normalde 80 olması gerekirken, sisteme yükledikleri pdf dosyasında tez 160 sayfa görünmektedir (aynı 80 sayfalık içerik tekrar edilmiştir). Buna benzer hataların olabileceğini varsayıp, tablodaki değerler hesaplanırken, sayfa sayısı 50’nin altında ve 500’ün üzerindeki tezler dikkate alınmamıştır. Kelime sayılarına ilişkin tanımlayıcı istatistikler hesaplanırken, benzer şekilde, özetlerde, 50’nin altında ve 600’ün üzerinde kelimesi olan tezler dikkate alınmamıştır.

Tablo 3
Tez Sayılarının Tanımlayıcı İstatistikleri

	Sayfa Sayıları		Özetteki Kelime Sayıları	
	Yüksek Lisans	Doktora	Yüksek Lisans	Doktora
Minimum	55,00	50,00	52,00	51,00
Maksimum	498,00	495,00	599,00	593,00
Ortalama	222,59	135,30	223,50	185,27
Standart Sapma	71,83	45,47	82,10	71,69
%25'lik değer	171,00	104,00	167,00	136,00
Medyan	212,00	127,00	213,00	176,00
%75'lik değer	261,00	157,00	261,00	222,00
Çarpıklık	0,83	1,39	1,15	1,22
Basıklık	3,82	6,58	5,26	5,95

Yıllara göre üniversitelerde yayınlanan tez sayılarında ekde yer alan tabloda sunulmuştur (AS04: *Yıllara göre üniversitelerde yayınlanan yüksek lisans ve doktora tezi sayıları nelerdir?*). İncelenen zaman dilimi 54 yıldır. Her bir yılda, her bir üniversitede yayınlanan tezleri tablo halinde sunabilmek takip edilmesi zor boyutlarda bir tablo ortaya çıkaracaktır. Bu nedenle de incelenen zaman dilimi (1) 1999 ve öncesi, (2) 2000-2009 yılları arası ve (3) 2010 yılı ve sonrası olarak üç alt bölüme ayrılmıştır. Her bir alt bölüm için ve toplamda da üniversitelerde gerçekleştirilen yüksek lisans ve doktora tez sayıları tabloda yer almaktadır. Tabloda ayrıca bütün dönemleri kapsayan tez sayıları da yer almaktadır. Tablonun son sütununda ise her bir üniversitede incelenen bütün dönemlerde yayınlanan toplam tez sayısı (yüksek lisans + doktora) yer almaktadır. Tabloda sadece Türkçe yazılan tez sayıları yer almaktadır. Çalışma kapsamında değerlendirmeye alınan 159 tane üniversitenin 36 tanesinde 1999 yılı ve öncesinde hazırlanmış lisansüstü tezler bulunmaktadır. Zaman içerisinde üniversite isimlerinde değişiklikler meydana gelmiştir. Örneğin Yıldız Üniversitesi, 1992 yılından itibaren Yıldız Teknik Üniversitesi olarak adlandırılmaktadır. Bu gibi isim değişiklikleri önceki tezleri de kapsayacak şekilde sistem tarafından otomatik olarak düzeltilmektedir. Başka bir ifade ile, tez.yok.gov.tr veri tabanında Yıldız Üniversitesi'ne ait bir kayıt bulunmamaktadır.

Yukarıda ifade edilen dönemler içerisinde en sık kullanılan kelimeler kelime bulutu haline getirilmiş ve Şekil 3'de sunulmuştur. 1999 ve öncesinde; maliyet, üretim, ürün, yapı ve model gibi kelimeler ön plana çıkmaktadır. 2000 ile 2009 yılları arasında ise yayınlanan tezlerde; süreç, kaynak, firma, ürün, yapı gibi kelimeler ön plana çıkmaktadır. Bu dönemde banka, teknoloji ve müşteri gibi kelimeler de önceki dönemde olmadığı kadar sıklıkla geçmeye başlamıştır. 2010 ve 2021 yılları arasında yayınlanan tezlerde ise, veri, araştırma, anket, davranış ve faktör gibi kelimeler sıklıkla kullanılmaya başlamıştır. Durum, algı, kurum ve anlam gibi kelimelerin de sıklıkla kullanılmaya başladığı belirlenmiştir.



Şekil 3. Üç farklı dönem için kelime bulutları

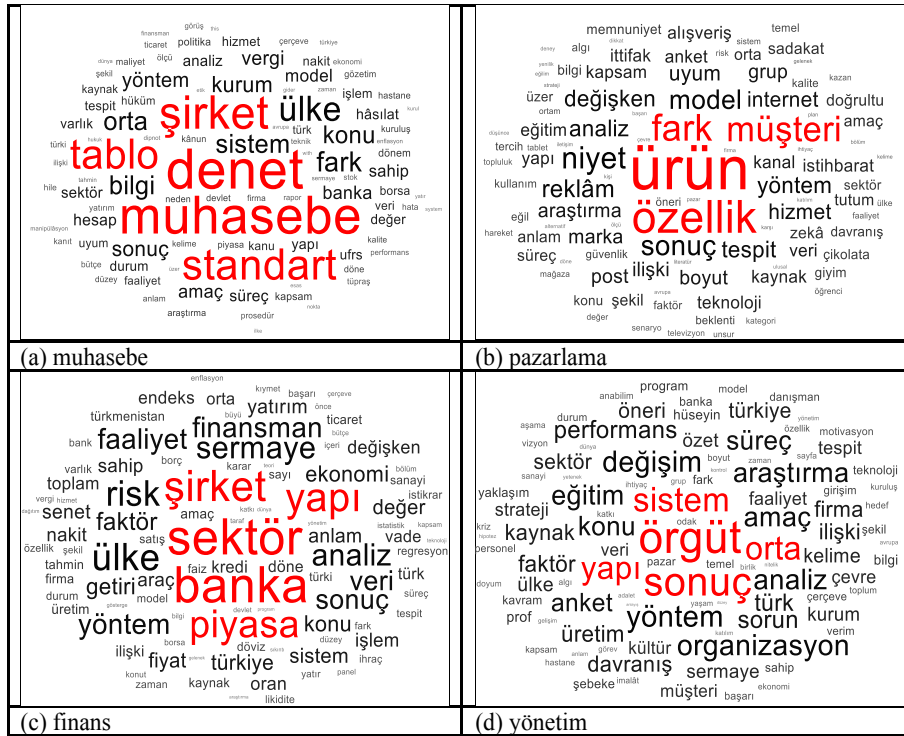
Tez yöneten danışmanların isimleri de çalışma kapsamında incelemeye tabi tutulmuştur. Ünvanlar temizlendikten sonra, toplamda 4.836 tane danışman ismi bir araya getirilmiştir. Ancak bazı isimlerin birden fazla yazıldığı belirlenmiştir. Örneğin iki tane adı olan danışmanın adı, bir tezde kısaltma olmadan yazılmışken, bir başka tezde adlardan birinde kısaltmaya gidilmiştir. Böylelikle bazı danışmanlar, daha fazla tez yönettiği halde, farklı bir isim altında daha az sayıda tez yönetmiş gibi görünmektedir. Soy isim değişikliği de söz konusu olabilmektedir. Bunlara ek olarak aynı isim ve soy isime sahip danışmanlar bulunabilmektedir. Bu durumun otomatik bir şekilde nasıl düzeltileceği bilinmemektedir. Bu nedenle bu çalışmada danışmanların isimleri konusunda detaylı bilgi sunmaktan kaçınılmaktadır. 80'in üzerinde tez yöneten dört tane danışman belirlenmiştir. 5'in altında tez yöneten danışman sayısı ise 3.303 olarak belirlenmiştir.

Danışman sayısına göre tezlerin dağılımı Tablo 4'de sunulmuştur. 29.381 tane Türkçe yazılmış tezin çok önemli bir bölümü (%98 = 28.754/29.381) tek danışmanlı olarak yönetilmiştir. Geri kalan %2'si ise iki veya üç danışmanlı olarak yönetilmiştir. Dört veya daha fazla danışman ile yönetilen teze rastlanmamıştır (AS05: *Danışman sayılarına göre tezlerin dağılımı nasıldır?*).

Tablo 4
Danışman Sayılarına Göre Tezler

	Yüksek Lisans	Doktora	Toplam
Tek Danışman	23.490	5.624	28.754
İki Danışman	111	103	214
Üç Danışman	43	10	53
Toplam	23.644	5.737	29.381

Bir araya getirilen veri seti ile danışman hocaların yönettiği tezlerde en sık geçen kelimeler analiz edilebilir. Bu amaçla dört farklı alandan (muhasebe, pazarlama, finans ve yönetim) seçilen danışmanların yönettiği tezlerde en sık geçen kelimeler Şekil 4'de yer almaktadır. Danışmanların isimleri saklı tutulmaktadır. Alan ile ilgili en sık tekrar eden kelimeler bariz bir şekilde diğer danışmanların yönettiği tezlerden farklıdır.



Şekil 4. Seçilen danışmanların yönettiği tezlerde en sık geçen kelimeler

Bunlara ek olarak üniversitelerde yönetilen tezlerde en sık geçen kelimeleri de incelemek mümkündür. Bu amaçla İşletme Ana Bilim Dalı'nda en fazla tez yönetilen iki üniversite olan Marmara Üniversitesi ve İstanbul Üniversitesi'nde yönetilen tezlerde en sık geçen kelimeler birer kelime bulutu halinde Şekil 5'de sunulmuştur. Marmara Üniversitesi'nde; süreç, bilgi, sektör, faaliyet, gibi kelimeler ön plana çıkarken, İstanbul Üniversitesi'nde bilgi, veri, sistem, model ve ülke gibi kelimeler ön plana çıkmaktadır. Benzer kelime bulutlarını diğer üniversiteler için de oluşturmak mümkündür.



Şekil 5. Üniversitelerde yönetilen tezlerde en sık geçen kelimeler

LDA ile Konu Modelleme

Veri seti analizde kullanılmadan önce ilk işlem olarak kelime kökleri bulunmuştur. Kelime kökleri bulunurken, Ahmet Akın tarafından geliştirilen Zemberek kütüphanesi kullanılmıştır (Akın, 2020). Söz konusu kütüphane Java programlama dilinde geliştirilmiştir ve Türkçe doğal dil işleme çalışmalarında kullanılmaktadır.

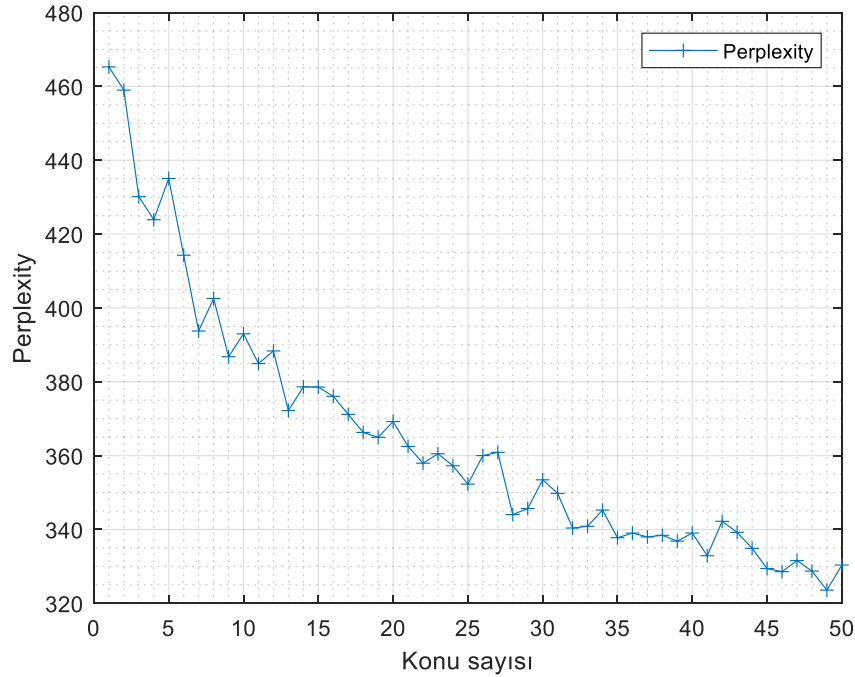
Veri seti analizde kullanılmadan önce ön işleme tabi tutulmuştur. İsim harici diğer kelime kökleri (sıfat, fiil, bağlaç gibi) analiz dışında tutulmuştur. Özetlerdeki her bir kelime için kelime köklerinin ne olduğu belirlendikten sonra iki veya daha az sayıda karakter içeren kökler veri setinden çıkarılmıştır. Ayrıca karakter sayısı 15 veya daha fazla olan kelime kökleri de analizden çıkarılmıştır. Noktalama işaretleri ve rakamlar silinmiş ve bütün kelimeler küçük harflere çevrilmiştir. Analizde kullanılan veri setinde 29.381 tane belge ve 3.178 tane tekil kelime kökü (isim) yer almaktadır.

LDA analizini uygulayabilmek için araştırmacının konu sayısını önceden belirlemiş olması gerekmektedir. Perplexity değeri LDA analizinde konu sayısının belirlenmesinde kullanılmaktadır. LDA'nın tahmin gücünü ölçmek için kullanılan standart bir ölçüdür (Blei vd., 2003; Güven, 2020).

$$perplexity = \exp\left(-\frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d}\right) \quad (1)$$

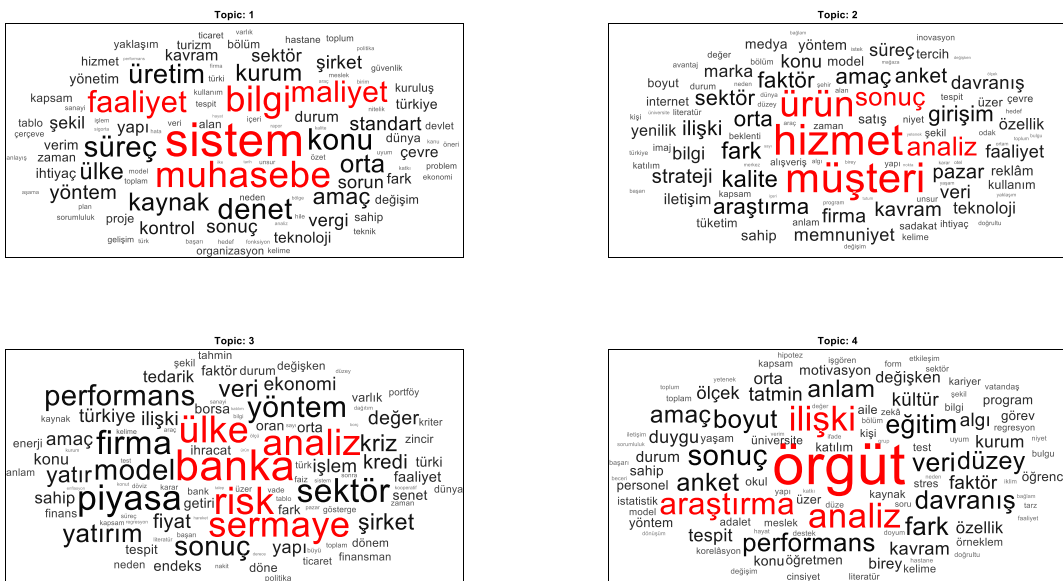
formülde N_d ifadesi d belgesinde ortaya çıkan toplam terim sayısının temsil etmektedir. w_d ifadesi ise belgedeki d sırasında yer alan terimi temsil etmektedir. Perplexity değerinin en küçük olması tercih edilmektedir. Bu değer en küçük olduğu konu sayısını kullanmak suretiyle LDA analizinin uygulanması önerilmektedir (Wang vd., 2018).

Konu sayısı 1'den 50'ye kadar LDA analizi çalıştırılmış ve her bir konu sayısında hesaplanan perplexity değerleri incelenmiştir. Veri setinin %10'u (2.938 tane tez) test amacı ile kullanılmış ve geri kalan %90'lık kısmı (26.443 tane tez) eğitim amacıyla kullanılmıştır. Her bir konu sayısı için eğitim seti ile model eğitilmiş ve eğitilen modelin test setindeki perplexity değerleri kaydedilmiştir. Sonuçlar Şekil 6'da sunulmuştur. Beklenildiği üzere, Şekil 6'da konu sayısı arttıkça perplexity değeri azalmaya devam etmektedir. Konu sayısının fazla olması analiz sonuçlarının yorumlanmasını zorlaştırmaktadır. Şekil 6'ya göre optimal konu sayısı dört olarak belirlenmiştir. Beş konu ile model çalıştırıldığında (435.06), dört konu ile erişilen perplexity değerinden (423.89) daha yüksek bir değer ortaya çıkmaktadır.



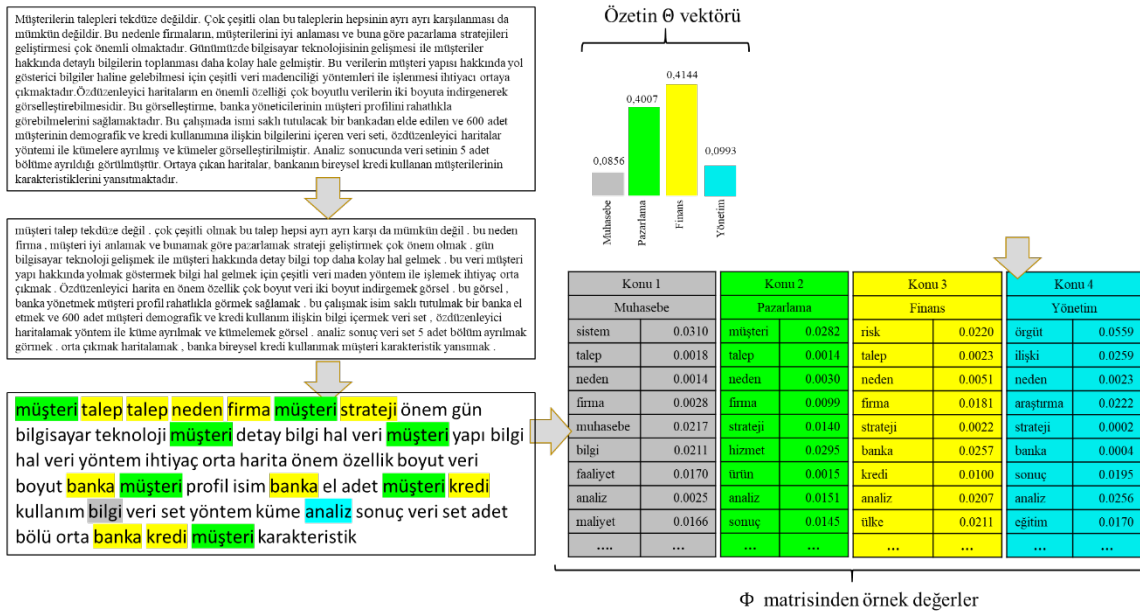
Şekil 6. Farklı konu sayılarında perplexity değerleri

Konu sayısı 4 olarak belirlenmiş ve LDA analizi veri setinin tamamını kullanmak suretiyle nihai olarak bir kere daha çalıştırılmıştır. Sonuçta her bir konuda en sık geçen kelimeler, kelime bulutu olarak Şekil 7'de gösterilmiştir. Bundan sonraki aşama şeklindeki kelime bulutlarının hangi konular ile ilgili olduğunun etiketlenmesi sürecidir. Birinci konuda muhasebe ve denetim ile ilgili kelimeler daha çok kullanıldığı için birinci konu muhasebe olarak etiketlenmiştir. İkinci konuda ise ürün, müşteri, gibi kelimeler daha sık kullanıldığı için pazarlama, üçüncü konu finans ve son olarak da dördüncü konu yönetim olarak belirlenmiştir.



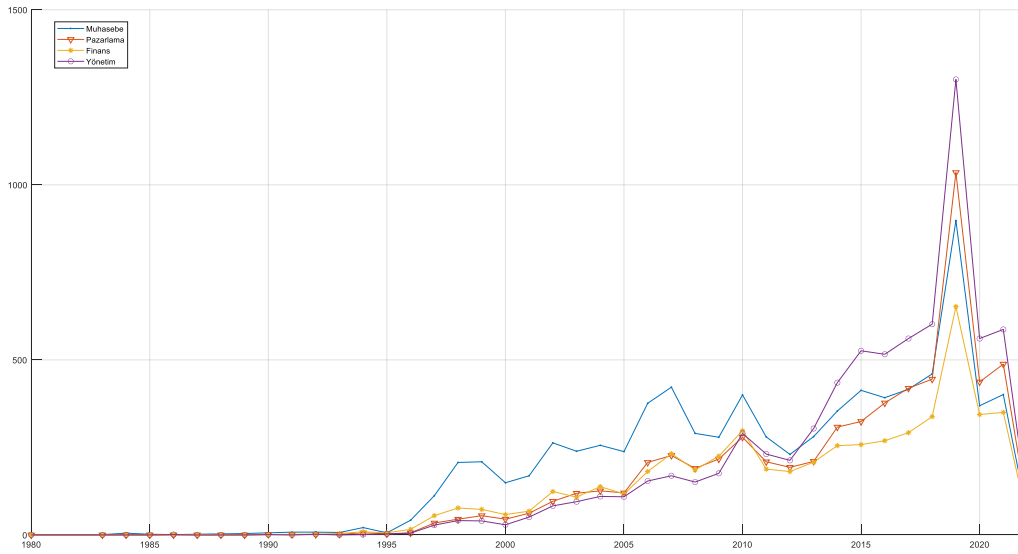
Şekil 7. Dört farklı konuya ilişkin kelime bulutları

LDA analizinin uygulama süreci, yazarın 2011 yılında Gaziantep Üniversitesi Sosyal Bilimler Enstitüsü'nde tamamladığı yüksek lisans tezinin özeti üzerinde detaylandırılmıştır. Şekil 8'de tezin özeti olduğu gibi yer almaktadır. Hemen altında ise her bir kelimenin köklerinin küçük harflerle yazılmış hali yer almaktadır. Onun da altında ön işleme sonrası kalan kelimeler yer almaktadır. Bütün tez özetleri için bu işlemler gerçekleştirilmiştir. Model, her bir kelimenin hangi konuya ait olduğu bilgisini içeren Φ matrisini ve Θ matrisini (Şekil 1) hesaplamaktadır. Bazı kelimelerin birden fazla konuya ait olabildiğine, bununla birlikte her konu için farklı olasılık değerlerine sahip olduklarına okuyucunun dikkat etmesi gerekir. En yüksek olasılık değerlerine sahip kelime kökleri işaretlenmiştir. Diğer kelime kökleri için de benzer işaretlemelerin gerçekleştirilmesi mümkündür.



Şekil 8. LDA ile Konu Modelleme

Şekil 9'da LDA tarafından belirlenen dört tane konu ile ilgili yıllar itibariyle yayınlanan tez sayıları grafik halinde sunulmuştur (*AS06: LDA ile belirlenen konularda yıllar itibariyle yayınlanan tez sayıları nedir?*). Grafik oluşturulurken, LDA analizi sonucunda hesaplanan belge-konu (θ) matrisinden faydalanılmıştır. Bu matriste her bir dökümanın, hangi konuya ait olduğuna ilişkin olasılık değeri yer almaktadır. Örneğin Şekil 8'de incelenen tezin özeti %8,56 ihtimalle muhasebe, %40,07 ihtimalle pazarlama, %41,44 ihtimalle finans ve son olarak da %9,93 ihtimalle yönetim konusunda hazırlanmış bir tezdir (ihtimaller toplamı 1'e eşittir). Her bir döküman için en yüksek ihtimal değerine sahip konu o tezin konusu olarak kabul edilmiştir. Örneğin Şekil 8'deki özetin finans konusunda hazırlanmış bir teze ait olduğu varsayılmış ve tezlerin tamamı bu şekilde etiketlenmiştir. Bununla birlikte bazı tezlerin farklı konularda hazırlanmış olabileceği ve ayrımın ancak uzmanlar tarafından doğru bir şekilde gerçekleştirilebileceği unutulmamalıdır. Bununla birlikte 29.381 tane tezin hangi konuya ait olduğunun uzmanlar tarafından belirlenmesinin önemli düzeyde bir maliyetle mümkün olacağı bir gerçektir. Şekil 9'a göre 1995 yılları ile 2012 yılları arasında muhasebe alanı ile ilgili yayınlanmış tezlerin sayısının baskın olduğu ortaya çıkmaktadır. Söz konusu yıllarda yönetim ile ilgili hazırlanan tezlerin sayısı düşük düzeyde seyretmektedir. 2013 yılından itibaren ise yönetim ile ilgili hazırlanan tez sayısında ciddi bir artış olmakta ve konu tez sayısı ile ilgili liderliğini 2022 yılında dahi sürdürmeye devam etmektedir.



Şekil 9. Yıllar itibarıyla konularda yayınlanan tez sayıları

2021 yılında tezlerde en sık kullanılan kelimelerin özeti Tablo 5’de sunulmuştur (AS07: 2021 yılında tezlerde kullanılan kelimelerin frekans dağılımı nasıldır?). Tabloya göre analiz, sonuç, örgüt ve veri kelimeleri 2021 yılı için en sık kullanılan kelime kökleridir. Menşe, genç, yemin ve servis gibi kelimeler ise 11 kez kullanılmıştır. Benzer tabloyu diğer yıllar için oluşturmak da mümkündür. (AS11: 2021 yılında tezlerde kullanılan kelimelerin frekans dağılımı nasıldır?)

Tablo 5

2021 yılında tezlerde kullanılan kelimeler ve kullanım sıklıkları

Kelime	Frek.	Kelime	Frek.	Kelime	Frek.	Kelime	Frek.	Kelime	Frek.	Kelime	Frek.
analiz	3129	boyut	1121	değişken	792	kaos	13	beyan	12	taktik	11
sonuç	2401	anket	1112	şekil	736	suistimal	13	adaptasyon	12	şubat	11
örgüt	2149	süreç	1006	denet	728	robot	13	dışarı	12	regülasyon	11
veri	2101	sahip	987	kaynak	702	kahramanmaraş	13	kastamonu	12	güçlük	11
ilişki	1972	düzyey	973	algı	701	duyu	13	temerrüt	12	anlaşma	11
araştırma	1740	hizmet	956	özellik	695	özerklik	12	alfa	12	seyir	11
amaç	1736	sistem	917	yapı	689	envanter	12	youtube	12	anahtar	11
fark	1557	durum	905	teknoloji	685	eksen	12	salınım	12	kargo	11
yöntem	1484	ülke	903	ölçek	635	prensip	12	emisyon	12	hane	11
sektör	1397	müşteri	894	sermaye	631	sendika	12	eklenti	12	akım	11
anlam	1393	kavram	885	zaman	623	tanıt	12	eser	12	kazan	11
orta	1359	üzere	883	risk	613	denklem	12	ofis	12	şikâyet	11
davranış	1350	kurum	879	niyet	561	alman	12	wuhan	12	hariç	11
performans	1278	kapsam	864	türkiye	552	yüzde	12	montaj	12	kocaeli	11
faaliyet	1265	banka	860	eğitim	548	büyükbaş	12	hayes	12	kampanya	11
bilgi	1260	muhasebe	847	literatür	547	yazılım	12	trabzon	12	aşağı	11
tespit	1186	fırma	824	girişim	523	geri	12	özgüven	12	servis	11
konu	1157	ürün	807	birey	522	bitki	12	tokat	12	yemin	11
model	1124	katılım	802	duygu	516	hissedar	12	feminist	12	genç	11
faktör	1123	şirket	800	piyasa	514	mecra	12	sayaç	12	menşe	11

Sonuç

Çalışmada, son 58 yılda (1964-2022) tez.yok.gov.tr adresinde “İşletme Ana Bilim Dalı”nda yayınlanan 29.381 tane tezin özetleri bir araya getirilmiş ve bu veri seti tanımlayıcı istatistikler yardımıyla değerlendirmeye tabi tutulmuştur. Ayrıca, konu modellemesi gerçekleştirilmiştir. Çalışmada aşağıdaki sonuçlara erişilmiştir:

- İşletme Ana Bilim Dalı, Sosyal Bilimler Enstitüsü’nde önemli bir ağırlığa sahiptir.
- LDA analizi sonucunda tezlerin, başlıca dört konu altında toplandığı belirlenmiştir. Bu konular; muhasebe, pazarlama, finans ve yönetimdir.
- 1995 ile 2012 yılları arasında muhasebe, 2013-2022 yılları arasında ise yönetim ile ilgili hazırlanan tezlerin sayısı yüksektir.
- Üretim yönetimi, sayısal yöntemler ve ticaret hukuku gibi işletme ana bilim dalları hakkında yazılan tezlerin sayısı henüz, diğer konuların (muhasebe, pazarlama, finans ve yönetim) arasında belirgin düzeyde değildir. Başka bir ifade ile söz konusu ana bilim dallarında hazırlanan tezlerle ilgili kelimeler belirgin bir birliktelik oluşturamamaktadır (analiz tarafından fark edilememektedir). Bu nedenle eksik kalan ama önemli olan bu konular hakkında da yazılan tezlerin sayısının ve niteliğinin artması önerilmektedir.
- İşletme ile ilgili en fazla lisansüstü tez yayınlayan üniversiteler sırasıyla; Marmara (2843 tane tez), İstanbul (1370), Gazi (1281), Dokuz Eylül (1192) ve Dumlupınar (842) üniversiteleridir.
- 2021 yılında tezlerde; analiz, sonuç, örgüt, veri, ilişki, araştırma gibi kelimeler en sık kullanılan kelimeler arasındadır. Menşe, genç, yemin ve servis gibi kelimelerin daha az kullanıldığı belirlenmiştir.

Önerilen sistemin çeşitli faydaları mevcuttur. İşletme Ana Bilim Dalı’nda lisansüstü tez hazırlamayı planlayan öğrenciler için genel bir durum değerlendirmesi olarak kullanılabilir. Bu çalışma ile birlikte konu modellemesinin pratik ve akademik uygulamalarına yönelik çalışmaların sayısının ve niteliğinin artması beklenmektedir.

Çalışmanın çeşitli kısıtları bulunmaktadır. İlk olarak, sorgulamada sadece İşletme Ana Bilim Dalı’nda hazırlanan tezler kullanılmıştır. Sonraki çalışmalarda; İşletme (İngilizce), İşletme (Almanca), İşletme Mühendisliği, İşletme Finansı, İşletme Yönetimi, İşletme Eğitimi gibi ana bilim dallarını da kapsayacak sorgular ile veri setinin hacminin artması sağlanabilir. Kelime köklerine inerken Zemberek kütüphanesinin belirlediği kelime kökleri olduğu gibi kullanılmıştır. Kütüphane gelişmiş özelliklere sahip, çok güçlü bir araçtır. Kelime kökleri belirlenirken ek ayarlamalar kullanılmamış, herhangi bir kelimenin kökü ilk haliyle kullanılmıştır. Örneğin hisse senedindeki hisse kelimesinin kökü his olarak belirlenmektedir. İlerleyen çalışmalarda daha gelişmiş ayarlar (veya farklı yazılımlar) kullanmak suretiyle çalışma tekrarlanabilir. Çalışmada birden fazla kelimeyi içeren kavramlar göz ardı edilmiştir. “Sosyal medya” kavramı bir bütün olarak dikkate alınması gerekirken, “sosyal” ve “medya” olmak üzere iki farklı kelime olarak dikkate alınmıştır. Tezlerin sadece özet kısımlarının kullanılması çalışmanın bir başka kısıtıdır. Erişilebildiği durumda, kelime sayısı artacak olsa da tam metinlerin kullanılması daha isabetli değerlendirmelerin gerçekleştirilmesine olanak sağlayacaktır.

Gelecek çalışmalar için öneriler söz konusudur. Birden fazla konu modelleme algoritmalarını kullanmak suretiyle karşılaştırma gerçekleştirilebilir. Farklı ana bilim dallarında benzer çalışmalar gerçekleştirilebilir. Sadece belirli bir alt konu hakkında yayınlanmış tezlerin metinlerine odaklanabilir (örneğin sadece muhasebe veya sadece yönetim alanında yazılmış tezler) ve böylelikle detaylı bir analiz gerçekleştirilebilir.

Kaynakça

- Akın, A. (2020). *Zemberek-NLP*. Erişim adresi: <https://github.com/ahmetaa/zemberek-nlp>.
- Alkan, G. (2014). Türkiye’de muhasebe alanında yapılan lisansüstü tez çalışmaları üzerine bir araştırma (1984-2012). *Muhasebe ve Finansman Dergisi*, 61, 41–52. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/427486>
- Alkan, G. ve Özkaya, H. (2015). Türk muhasebe ve finans yazınının SSCI’da 20 yılı. *Muhasebe ve Finansman Dergisi*, 66, 175–192. Erişim adresi: <http://journal.mufad.org.tr/attachments/article/794/10.pdf>
- Benligiray, S. ve Tez, H. Ö. (2011). Üniversitelerin işletme eğitimi veren bölümleri öğrencilerine genel işletmecilik değerlerini kazandırmayı amaçlıyorlar mı? *Afyon Kocatepe Üniversitesi, İİBF Dergisi*, 13(2), 49–70. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/18915>
- Berkman, Ü.A. (1987). Amme İdaresi Dergisi’nde yayınlanan makaleler ve Türk yönetim bilimi. *Amme İdaresi Dergisi*, 20(4), 19–42. Erişim adresi: <https://dergipark.org.tr/tr/pub/oguiibf/issue/56499/785684>
- Bilginoğlu, F. (2018). İşletme eğitimi üzerine düşünceler. *Muhasebe Enstitüsü Dergisi*, 16(59), 3–6. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/541883>
- Blei, D. M., Ng, A. Y. ve Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022. Erişim adresi: <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Borah, A., Wang, X. S. ve Ryoo, J.H. (2018). Understanding influence of marketing thought on practice: an analysis of business journals using textual and Latent Dirichlet Allocation (LDA) analysis. *Customer Needs and Solutions*, 5(3), 146–161. <https://doi.org/10.1007/S40547-018-0089-Z>
- Calvo, H., Hernández-Castañeda, Á. ve García-Flores, J. (2018, Ekim). *Author identification using latent dirichlet allocation*. 18th International Conference, CICLing 2017, konferansında sunulan bildiri, Springer, Budapest. https://doi.org/10.1007/978-3-319-77116-8_22/TABLES/4
- Coşkun, S. S. ve Kazan, H. (2021). Bibliyometrik analiz yoluyla küresel tedarik zinciri araştırma alanının görselleştirilmesi. *Journal of Transportation and Logistics*, 6(1), 125-136. doi: 10.26650/JTL.2021.943712.
- Çalli, L., Çalli, F., Büşra, V. ve Çalli, A. (2021). Yönetim bilişim sistemleri disiplinde hazırlanan lisansüstü tezlerin Gizli Dirichlet Ayrımı algoritmasıyla konu modellemesi. *MANAS Sosyal Araştırmalar Dergisi*, 10(4), 2355–2372. <https://doi.org/10.33206/MJSS.894809>
- Düzakın, E. ve Yılmaz, Ö. (2009). İşletme mezunlarının iş hayatındaki yeri ve işletme eğitimi: 1000 büyük sanayi kuruluşunun işletme mezunlarından beklentileri üzerine araştırma. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 18(1), 149–164. <https://dergipark.org.tr/en/pub/cusosbil/issue/4381/60101>
- Ekin, C. Ç., Çakıcı, M., Şener, E., Türker, S. ve Altanlar, S. (2021). Research trends analysis in educational journal publications on Covid-19 using descriptive and text mining methods: Ppreliminary analysis. *Avrupa Bilim ve Teknoloji Dergisi*, 29(29), 432–437. <https://doi.org/10.31590/EJOSAT.1036109>

- Ekinci, E., Omurca, S. İ., Kırık, E. ve Taşçı, Ş. (2020). Tıp veri kümesi için Gizli Dirichlet Ayrımı. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 22(64), 67–80. <https://doi.org/10.21205/DEUFMD.2020226408>
- Gürcan, F. ve Özyurt, Ö. (2020). E-öğrenme araştırmalarındaki temel eğilimler ve bilgi alanları: 2008-2018 yılları arasında yayımlanan makalelerle konu modelleme analizi. *Journal of Computer and Education Research*, 8(16), 738–756. <https://doi.org/10.18009/JCER.769349>
- Güvemli, O. (2018). Türkiye’de işletme eğitiminin tarihçesi ile ilgili bir görüş. *Accounting and Financial History Research Journal*, 14(14), 238–241. Erişim adresi: <https://dergipark.org.tr/tr/download/issue-file/9467>
- Güven, Z.A., Diri, B. ve Çakaloğlu, T. (2019). Emotion detection with n-stage Latent Dirichlet Allocation for Turkish Tweets. *Academic Platform Journal of Engineering and Science*, 7(3), 467-472. doi: 10.21541/apjes.459447
- Güven, Z.A., Diri, B. ve Çakaloğlu, T. (2020). Duygu analizi için n-aşamalı Gizli Dirichlet Ayrımı ile diğer konu modelleme yöntemlerinin karşılaştırılması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 35(4), 2135-2145. doi: 10.17341/gazimmfd.556104.
- Güven, Z.A. (2021, Eylül). *Türkçe tweetlerde duygu analizi için BERT modelleri ve makine öğrenme yöntemlerinin karşılaştırılması*. UBMK’2021 6th International Conference on Computer Science and Engineering–98 konferansında sunulan bildiri, IEEE, Ankara. doi: 10.1109/UBMK52708.2021.9559014.
- Hotamışlı, M. ve Erem, I. (2014). Muhasebe ve Finansman Dergisi’nde yayımlanan makalelerin bibliyometrik analizi. *Muhasebe ve Finansman Dergisi*, 63, 1–20. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/427519>
- Kıymetli Şen, İ., Hatunoğlu, Z. ve Terzi, S. (2017). Muhasebe araştırmalarında muhasebe eğitiminin yeri ve önemi: Muhasebe dergileri üzerinde bibliyometrik bir araştırma. *Muhasebe Bilim Dünyası Dergisi*, 19(1), 247–291. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/1127967>
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H. ve Adam, S. (2018). Applying lda topic modeling in communication research: toward a valid and reliable methodology. *Communication Methods and Measures*, 12(3), 93–118. <https://doi.org/10.1080/19312458.2018.1430754>
- Omurca, S. İ., Ekinci, E., Yakupoğlu, E., Arslan, E. ve Çapar, B. (2021). Automatic detection of the topics in customer complaints with artificial intelligence. *Balkan Journal of Electrical and Computer Engineering*, 9(3), 268–277. <https://doi.org/10.17694/BAJECE.832274>
- Onan, A., Yalçın, A. ve Atik, E. (2020). Üniversite bilgi yönetim sistemi servis destek taleplerinin konu modelleme tabanlı analizi. *Avrupa Bilim ve Teknoloji Dergisi*, EJOSAT(Özel Sayı 2020), 389–397. <https://doi.org/10.31590/EJOSAT.780642>
- Özkul, A. S. (2012a). 9. Yüzyıl Türk Yükseköğretiminde İşletme Eğitimi. *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 16, 223-241. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/215339>

- Özkul, A. S. (2012b). *Türk yükseköğretiminde işletme eğitimi* (Yayınlanmamış Doktora Tezi). Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü, Isparta
- Sakin, T. (2008). A content analysis of papers published in the Journal of School of Business Administration: Accounting and finance (1972-2007). *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 37(1), 13–21. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/98129>
- Sargut, S. A. (2009). Türkiye’de işletme yönetimi eğitiminin kurumsal çerçevesi: çeşitlilikten eşbiçimliliğe. *Eskişehir Osmangazi Üniversitesi İİBF Dergisi*, 4(1), 51–63. Erişim adresi: <https://dergipark.org.tr/tr/pub/oguiibf/issue/56499/785685>
- Sipahi, G.A. (2020). 2008-2018 yılları arasında perakendecilik ile ilgili çalışmaların bibliyometrik analizi. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 29(4), 54-69. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/936596>.
- Şahin, A. S. ve Güler Bayazıt, N. (2020). What Java Developers have talked about? An empirical study on Stack Overflow. *Avrupa Bilim ve Teknoloji Dergisi*, 19, 354–365. <https://doi.org/10.31590/EJOSAT.702949>
- Şimdi, H. ve Garip, B. (2021). Global financial crisis and trade papers: topic analysis via Latent Dirichlet Allocation model. *Current Research in Social Sciences*, 7(2), 76–94. <https://doi.org/10.30613/CURESOSC.931149>
- Tokdemir, G. (2021). Using text mining for research trends in empirical software engineering. *Politeknik Dergisi*, 24(3), 1227–1235. <https://doi.org/10.2339/POLITEKNIK.831391>
- Toplu Yaşlıoğlu, D., İnan, A. ve Sözüer, A. (2018). Stratejik yönetim araştırmalarında Türkiye kapsamı: 2000-2015 döneminde yayınlanan makalelerin bibliyometrik incelemesi. *İstanbul Management Journal*, 29(84), 79–92. <https://doi.org/10.26650/imj.2018.29.84.0005>
- Üsdiken, B. ve Erçek, M. (2009). Türkiye’de iş dünyası için ilk işletme dergisi: İşletme-İşletme Ekonomisi ve Organizasyon mecmuası. *Yönetim Araştırmaları Dergisi*, 9(1), 53–90. Erişim adresi: http://yad.baskent.edu.tr/files/2009_cilt_9_mart_3.pdf
- Üsdiken, B. ve Erden, Z. (2001). Örnek alma, mecbur tutulma ve geçmişe bağımlılık: Türkiye’deki yönetim yazınında değişim. *Amme İdaresi Dergisi*, 34(4), 1–31. Erişim adresi: https://research.sabanciuniv.edu/id/eprint/26712/1/1bf154f04b19782_ek.pdf
- Wang, W., Feng, Y. ve Dai, W. (2018). Topic analysis of online reviews for two competitive products using Latent Dirichlet Allocation. *Electronic Commerce Research and Applications*, 29, 142–156. <https://doi.org/10.1016/J.ELERAP.2018.04.003>

Ekler

Tablo Ek-1

Yıllara Göre Üniversitelerde Yayınlanan Tez Sayıları

Üniversite	1999 ve öncesi		2000-2009		2010 ve sonrası		Toplam		
	YL	DR	YL	DR	YL	DR	YL	DR	TOPLAM
Abant İzzet Baysal	4	0	57	0	67	5	128	5	133
Adnan Menderes	5	1	39	8	71	35	115	44	159
Adıyaman	0	0	0	0	28	0	28	0	28
Afyon Kocatepe	9	4	113	54	134	59	256	117	373
Akdeniz Karpaz	0	0	0	0	1	0	1	0	1
Akdeniz	9	0	87	24	137	50	233	74	307
Aksaray	0	0	0	0	104	22	104	22	126
Alanya Alaaddin									
Keykubat	0	0	0	0	27	0	27	0	27
Altınbaş	0	0	0	0	46	0	46	0	46
Anadolu	95	31	190	60	215	106	500	197	697
Ankara Hacı Bayram									
Veli	0	0	0	0	84	10	84	10	94
Ankara Yıldırım Beyazıt	0	0	0	0	10	3	10	3	13
Ankara	56	19	167	78	132	80	355	177	532
Antalya Bilim	0	0	0	0	9	0	9	0	9
Artvin Çoruh	0	0	0	0	4	0	4	0	4
Atatürk	79	56	130	50	176	155	385	261	646
Atılım	0	0	24	0	134	5	158	5	163
Avrasya	0	0	0	0	66	5	66	5	71
Aydın Adnan Menderes	0	0	0	0	50	16	50	16	66
Ağrı İbrahim Çeçen	0	0	0	0	25	0	25	0	25
Bahçeşehir	0	0	0	0	317	0	317	0	317
Balıkesir	5	0	43	0	126	29	174	29	203
Bandırma Onyediy Eylül	0	0	0	0	20	3	20	3	23
Bartın	0	0	0	0	58	1	58	1	59
Batman	0	0	0	0	25	0	25	0	25
Bayburt	0	0	0	0	33	0	33	0	33
Başkent	13	0	66	12	97	91	176	103	279
Beykent	0	0	37	0	325	75	362	75	437
Bilecik Şeyh Edebali	0	0	0	0	76	2	76	2	78
Bingöl	0	0	0	0	29	6	29	6	35
Bolu Abant İzzet Baysal	0	0	0	0	49	5	49	5	54
Bozok	0	0	1	0	23	0	24	0	24
Burdur Mehmet Akif									
Ersoy	0	0	0	0	66	20	66	20	86
Bursa Orhangazi	0	0	0	0	1	0	1	0	1
Bursa Teknik	0	0	0	0	6	0	6	0	6
Bursa Uludağ	0	0	0	0	109	25	109	25	134
Bülent Ecevit	0	0	0	0	33	29	33	29	62
Celal Bayar	13	0	81	28	89	48	183	76	259
Cumhuriyet	31	8	68	19	90	24	189	51	240
Çanakkale Onsekiz									
Mart	14	0	29	0	91	37	134	37	171
Çankaya	0	0	0	0	40	3	40	3	43
Çankırı Karatekin	0	0	0	0	62	0	62	0	62

Üniversite	1999 ve öncesi		2000-2009		2010 ve sonrası		Toplam		
	YL	DR	YL	DR	YL	DR	YL	DR	TOPLAM
Çağ	0	0	6	0	3	0	9	0	9
Çukurova	61	27	105	41	141	95	307	163	470
Dicle	0	0	0	0	31	30	31	30	61
Dokuz Eylül	98	45	428	74	466	81	992	200	1192
Doğuş	0	0	4	0	40	1	44	1	45
Dumlupınar	74	3	366	32	297	70	737	105	842
Düzce	0	0	0	0	67	29	67	29	96
Ege	3	2	46	8	97	34	146	44	190
Erciyes	24	2	94	25	87	72	205	99	304
Erzincan Binali Yıldırım	0	0	0	0	25	8	25	8	33
Erzincan	0	0	0	0	20	1	20	1	21
Erzurum Teknik	0	0	0	0	2	0	2	0	2
Eskişehir Osmangazi	1	0	79	0	121	36	201	36	237
European University of									
Lefke	0	0	0	0	0	1	0	1	1
Fatih	0	0	7	0	30	0	37	0	37
Fırat	0	0	0	0	108	0	108	0	108
Galatasaray	0	0	0	0	88	9	88	9	97
Gazi	102	9	476	66	475	153	1053	228	1281
Gaziantep	0	0	50	0	156	41	206	41	247
Gaziosmanpaşa	0	0	18	0	60	28	78	28	106
Gebze Teknik	0	0	0	0	86	62	86	62	148
Gebze Yüksek Tek. Enst.	56	0	164	58	77	44	297	102	399
Gedik	0	0	0	0	4	0	4	0	4
Gediz	0	0	0	0	34	0	34	0	34
Giresun	0	0	0	0	66	6	66	6	72
Gümüşhane	0	0	0	0	100	16	100	16	116
Hacettepe	69	11	139	57	218	85	426	153	579
Haliç	0	0	29	0	112	64	141	64	205
Harran	0	0	4	0	41	1	45	1	46
Hasan Kalyoncu	0	0	0	0	239	56	239	56	295
Hatay Mustafa Kemal	0	0	0	0	29	6	29	6	35
Hitit	0	0	0	0	89	15	89	15	104
Işık	0	0	0	0	28	0	28	0	28
İbn Haldun	0	0	0	0	21	0	21	0	21
İnönü	95	16	91	30	142	97	328	143	471
İstanbul Arel	0	0	0	0	237	35	237	35	272
İstanbul Aydın	0	0	0	0	437	43	437	43	480
İstanbul Bilgi	0	0	0	0	5	0	5	0	5
İstanbul Esenyurt	0	0	0	0	33	0	33	0	33
İstanbul Gelişim	0	0	0	0	335	55	335	55	390
İstanbul Kemerburgaz	0	0	0	0	10	0	10	0	10
İstanbul Kent	0	0	0	0	4	0	4	0	4
İstanbul Kültür	0	0	15	0	87	24	102	24	126
İstanbul Medeniyet	0	0	0	0	9	0	9	0	9
İstanbul Medipol	0	0	0	0	3	0	3	0	3
İstanbul Okan	0	0	0	0	218	16	218	16	234
İstanbul Rumeli	0	0	0	0	6	0	6	0	6

Üniversite	1999 ve öncesi		2000-2009		2010 ve sonrası		Toplam		
	YL	DR	YL	DR	YL	DR	YL	DR	TOPLAM
İstanbul Sabahattin									
Zaim	0	0	0	0	74	5	74	5	79
İstanbul Teknik	34	5	35	12	24	4	93	21	114
İstanbul Ticaret	0	0	38	0	182	42	220	42	262
İstanbul Yeni Yüzyıl	0	0	0	0	7	0	7	0	7
İstanbul	90	37	336	143	448	316	874	496	1370
İstinye	0	0	0	0	17	0	17	0	17
İzmir Demokrasi	0	0	0	0	6	0	6	0	6
İzmir Katip Çelebi	0	0	0	0	62	24	62	24	86
KTO Karatay	0	0	0	0	161	4	161	4	165
Kadir Has	0	0	57	1	27	8	84	9	93
Kafkas	0	0	24	0	107	8	131	8	139
Kahramanmaraş Sütçü									
İmam	10	0	49	0	156	44	215	44	259
Karabük	0	0	0	0	230	37	230	37	267
Karadeniz Teknik	59	7	90	26	171	64	320	97	417
Karamanoğlu									
Mehmetbey	0	0	8	0	96	6	104	6	110
Kastamonu	0	0	0	0	126	33	126	33	159
Kilis 7 Aralık	0	0	0	0	18	4	18	4	22
Kocaeli	30	0	214	28	121	43	365	71	436
Kütahya Dumlupınar	0	0	0	0	114	12	114	12	126
Kırgızistan-Türkiye									
Manas	0	0	12	0	22	9	34	9	43
Kırklareli	0	0	0	0	32	0	32	0	32
Kırıkkale	17	3	40	3	74	47	131	53	184
Kırşehir Ahi Evran	0	0	0	0	20	0	20	0	20
Maltepe	0	0	40	0	84	27	124	27	151
Manisa Celal Bayar	0	0	0	0	96	29	96	29	125
Marmara	290	82	1123	148	993	207	2406	437	2843
Mehmet Akif Ersoy	0	0	0	0	38	6	38	6	44
Mersin	0	0	10	0	60	29	70	29	99
Mevlana	0	0	0	0	2	0	2	0	2
Munzur	0	0	0	0	27	0	27	0	27
Mustafa Kemal	7	0	26	0	34	1	67	1	68
Muğla Sıtkı Koçman	0	0	0	0	129	42	129	42	171
Namık Kemal	0	0	0	0	20	3	20	3	23
Necmettin Erbakan	0	0	0	0	63	9	63	9	72
Nevşehir Hacı Bektaş									
Veli	0	0	0	0	84	21	84	21	105
Niğde Ömer Halisdemir	0	0	0	0	73	27	73	27	100
Nişantaşı	0	0	0	0	137	0	137	0	137
Nuh Naci Yazgan	0	0	0	0	57	0	57	0	57
Okan	0	0	0	0	474	19	474	19	493
Ondokuz Mayıs	0	0	0	0	80	14	80	14	94
Osmaniye Korkut Ata	0	0	0	0	84	28	84	28	112
Pamukkale	12	0	51	0	176	47	239	47	286
Recep Tayyip Erdoğan	0	0	0	0	56	5	56	5	61
Sakarya	69	8	307	51	230	173	606	232	838

Üniversite	1999 ve öncesi		2000-2009		2010 ve sonrası		Toplam			
	YL	DR	YL	DR	YL	DR	YL	DR	TOPLAM	
Selçuk		51	14	190	61	249	159	490	234	724
Sivas Cumhuriyet		0	0	0	0	63	25	63	25	88
Süleyman Şah		0	0	0	0	1	0	1	0	1
TOBB Ekonomi ve Teknoloji		0	0	2	0	101	0	103	0	103
Tekirdağ Namık Kemal		0	0	0	0	30	6	30	6	36
Tokat Gaziosmanpaşa		0	0	0	0	43	24	43	24	67
Toros		0	0	0	0	128	3	128	3	131
Trakya		14	0	64	0	156	51	234	51	285
Tunceli		0	0	0	0	3	0	3	0	3
Turgut Özal		0	0	0	0	2	4	2	4	6
Türk Hava Kurumu		0	0	0	0	405	29	405	29	434
Ufuk		0	0	4	0	197	4	201	4	205
Uludağ		86	11	135	39	74	18	295	68	363
Uluslararası Kıbrıs		0	0	0	0	2	0	2	0	2
Uşak		0	0	2	0	88	6	90	6	96
Van Yüzüncü Yıl		0	0	0	0	36	7	36	7	43
Yalova		0	0	0	0	27	0	27	0	27
Yaşar		0	0	4	0	63	33	67	33	100
Yeditepe		0	0	4	0	3	0	7	0	7
Yeni Yüzyıl		0	0	0	0	1	0	1	0	1
Yozgat Bozok		0	0	0	0	23	3	23	3	26
Yüzüncü Yıl		0	0	0	0	37	1	37	1	38
Yıldırım Beyazıt		0	0	0	0	1	1	1	1	2
Yıldız Teknik		13	0	196	0	219	39	428	39	467
Zonguldak Bülent Ecevit		0	0	0	0	42	20	42	20	62
Zonguldak Karaelmas		0	0	56	0	21	7	77	7	84

Extended Abstract

Purpose

This study, it is aimed to analyze the abstracts of the theses prepared in the Department of Business Administration. In the study, the abstracts of 29,381 graduate theses registered in the Thesis Center of the Council of Higher Education between 1964 and 2022 and prepared in the Department of Business Administration was used. Latent Dirichlet Allocation (LDA) was used as the analysis method. In determining the number of topics, the perplexity value were used. The study seeks answers to the research questions such as “How many master's and doctoral theses have been published in which language?” “What was the number of master's and doctoral theses published over the years?” “If the abstracts in dissertations are modeled under four different topics with LDA analysis, what is the word cloud of the most frequently used words in each model?”

Design and Methodology

In the study, sampling was not used and all accessible thesis summaries were used. The study is an exploratory and applied study. The summaries brought together were first analyzed with the help of descriptive statistics. The distribution of theses according to their languages, the number of pages of master's and doctoral theses, the number of words in the abstracts, the number of theses published in the department of business administration by years, the number of theses published in the social sciences institute in the same years, the number of master's and doctoral theses published in universities by years, The most frequently repeated words in the theses published in selected universities, the most frequently repeated words in the theses directed by selected advisors, the most frequently used words in the theses by years, how many theses were published on selected topics by years were examined. After examining the data set descriptively. Subject modeling was carried out with the help of LDA analysis. For this purpose, firstly, data preprocessing was carried out. First, word roots were found with the help of the Zemberek-NLP Library, then the roots of less than 2 characters and more than 15 characters were deleted from the data set. It was subjected to LDA analysis and perplexity values were calculated to determine the number of subjects. The perplexity scores indicates that the optimal number of subject is four.

Findings

One of the important findings reached in the study is that approximately 15% of the theses published in the institute of social sciences in 2020 and 2021 were prepared in the department of business administration. This finding reveals the weight of the business department in the field of social sciences. The number of pages of doctoral theses is more than the number of pages of master's theses. The majority of assertions are conducted with a single advisor. About 1% of assertions are conducted with two or more advisors. The most frequently repeated words in the theses administered at two selected universities were turned into a word cloud. The most frequently repeated words in the theses directed by four selected advisors were turned into a word cloud. The number of doctoral theses written on the topics is presented in a table. At the end of the study, in order of the subjects on which the most thesis was written; accounting, marketing, finance, and management. The number of theses prepared between 1995 and 2012 on accounting and between 2013-2021 on management is high.

Research Limitations

Subject modeling was tried to be applied to the entire data set that could be accessed in the research. However, the perplexity value decreases as the number of subjects increases. For this reason, the volume of the data set used in the study was reduced. The summary of doctoral theses written on the first five topics on which most theses were written was used in the study. Using only the abstracts of the theses is another limitation of the study. Although the number of words will increase, the use of full texts will allow more accurate evaluations.

Implications (Theoretical, Practical, and Social)

The proposed system has several benefits. Students who are considering doing postgraduate studies can determine their advisors or universities with the help of the model. They can identify thesis topics by using topics and words that have gained popularity over time. A situation assessment has been made regarding the subjects on which theses have been published. Over the years, it has been determined that various topics dominate business theses.

Originality Value

In the study, LDA analysis was applied to perform the subject modeling of the theses prepared in the department of business administration. The literature review for the study revealed that expert knowledge was used or bibliometric analysis was carried out for the analysis of the theses prepared in the department of business administration. However, LDA analysis was not carried out for the subject modeling of the theses prepared in the department of business administration. LDA analysis, on the other hand, has been used in the literature for modeling different subjects.

Arařtırmacı Katkısı: Mehmet ÖZÇALICI (%100).

Teřekkür: alıřmanın önceki versiyonunu detaylı bir řekilde inceleyip, geliřtirilmesi için yapıcı eleřtiriler ile katkıda bulunan hakemlere teřekkürlerimi sunarım.