KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ENDÜSTRİ MÜHENDİSLİĞİ
ANABİLİM DALI

YÜKSEK LİSANS TEZİ

METİN MADENCİLİĞİ İLE ENDÜSTRİ MÜHENDİSLİĞİNDE KARİYER PLANLAMA

ZEYNEP AKSOY

KOCAELİ 2023

KOCAELİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

ENDÜSTRİ MÜHENDİSLİĞİ

ANABİLİM DALI

YÜKSEK LİSANS TEZİ

METİN MADENCİLİĞİ İLE ENDÜSTRİ MÜHENDİSLİĞİNDE KARİYER PLANLAMA

ZEYNEP AKSOY

Dr. Öğr. Üyesi Hatice ESEN		
Danışman,	**Kocaeli Üniv**	**.....**
Prof.Dr.Gülşen AKMAN		
Jüri Üyesi,	**Kocaeli Üniv**	**.....**
Doç.Dr. Berrin DENİZHAN		
Jüri Üyesi,	**Sakarya Üniv**	**.....**

Tezin Savunulduğu Tarih: 9.10.2023

ETİK BEYAN VE ARAŞTIRMA FONU DESTEĞİ

Kocaeli Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez/proje çalışmasında,

- Bu tezin/projenin bana ait, özgün bir çalışma olduğunu,
- Çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı,
- Bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi,
- Bu çalışmanın Kocaeli Üniversitesi'nin abone olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü'nün belirlemiş olduğu ölçütlere uygun olduğunu,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Tezin/Projenin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez/proje çalışması olarak sunmadığımı,

beyan ederim.

[X] Bu tez/proje çalışmasının herhangi bir aşaması hiçbir kurum/kuruluş tarafından maddi/alt yapı desteği ile desteklenmemiştir.

[ ] Bu tez/proje çalışması kapsamında üretilen veri ve bilgiler ..... tarafından ..... no'lu proje kapsamında maddi/alt yapı desteği alınarak gerçekleştirilmiştir.

Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

(İmza)
.....
Zeynep Aksoy

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI

Fen Bilimleri Enstitüsü tarafından onaylanan lisansüstü tezimin/projemin tamamını veya herhangi bir kısmını, basılı ve elektronik formatta arşivleme ve aşağıda belirtilen koşullarla kullanıma açma izninin Kocaeli Üniversitesi'ne verdiğimi beyan ederim. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet hakları bende kalacak, tezimin/projemin tamamının ya da bir bölümünün gelecekteki makale, kitap, tebliğ, lisans, patent gibi çalışmalarda kullanımı, danışmanımın isim hakkı saklı kalmak koşuluyla ve her iki tarafın bilgisi dâhilinde bana ait olacaktır.
Tezin/projenin kendi özgün çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin/projenin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.
Yükseköğretim kurulu tarafından yayınlanan **"Lisanüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"** kapsamında tezim aşağıda belirtilen koşullar haricinde YÖK Ulusal Tez Merkezi/ Kocaeli Üniversitesi Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

☐ Enstitü yönetim kurulu kararı ile tezimin/projemin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir.
☐ Enstitü yönetim kurulu gerekçeli kararı ile tezimin/projemin erişime açılması mezuniyet tarihinden itibaren 6 ay ertelenmiştir.
☒ Tezim/projem ile ilgili gizlilik kararı verilmemiştir.

.....
(İmza)
.....
Zeynep Aksoy

ÖNSÖZ VE TEŞEKKÜR

Endüstri mühendisinin çalışmadığı sektör hemen hemen yok gibidir. Birçok sektörle bağlantılı olup, dinamik yapısı gereği sürekli yenilenmektedir. İş ilanları içerisinde yüzlerce bilgi barındırmaktadır. Bu bilgiler işin niteliği, kişinin işe yatkınlığı, uyumu gibi, değişen Dünya ve ihtiyaçlarına uyum sağlamak için sürekli gelişim ile karşı karşıyayız.

Yapılan bu tez çalışmasında veriler Python 3.10 programı aracılığıyla alınıp, incelenmiştir. Endüstri mühendisi iş ilanlarının içerdiği bilgilerden yola çıkılarak, Endüstri mühendisinde aranan özellikler saptanmış olup Endüstri mühendisliği alımları için veri analizleri yapılarak veriler incelenmiştir.

İçerisinde veri ve metin madenciliğinin birçok tekniği bulunan bu çalışmamda tüm aşamalarında, yüksek lisans eğitimim süresince her konuda yardımcı olan, bilgilerini ve tecrübelerini paylaşan canım hocam Endüstri Mühendisliği Ana Bilim Dalı Öğretim Üyesi Dr. Öğr. Hatice ESEN hocama teşekkürü borç bilirim. Ayrıca tüm eğitim ve öğrenimim boyunca en başından en sonuna kadar desteklerini ve yardımlarını esirgemeyen başta babam Salih Aksoy' a ve tüm aileme teşekkürlerimi sunarım.

Aralık- 2023 Zeynep AKSOY

İÇİNDEKİLER

ETİK BEYAN VE ARAŞTIRMA FONU DESTEĞİ i
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI ii
ÖNSÖZ VE TEŞEKKÜR iii
İÇİNDEKİLER iv
ŞEKİLLER DİZİNİ vi
TABLOLAR DİZİNİ vii
SİMGELER VE KISALTMALAR DİZİNİ viii
ÖZET ix
ABSTRACT x
1. GİRİŞ 1
2. LİTERATÜR ARAŞTIRMASI 4
3. VERİ MADENCİLİĞİ 8
3.1. Veri Madenciliğine Giriş 8
3.2. Veri Madenciliğinin Gelişimi 8
3.3. Veri Madenciliği Modelleri 9
3.3.1. Veri Madenciliği Tahminleyici Modelleri 9
3.3.2. Veri Madenciliği Tanımlayıcı Modelleri 11
3.4. Veri Madenciliğinin Önemi 12
3.5. Veri Madenciliğinin Kullanım Alanları 12
3.6. Veri Madenciliğinde Karşılaşılan Sorunlar 14
3.7. Veri Madenciliği Süreci 15
3.7.1. Problemin Tanımlanması 15
3.7.2. Verilerin İşlenmeye Uygun Şekilde Hazırlanması 15
3.7.3. Probleme Uygun Modelin Kurulması ve Değerlendirilmesi 16
3.7.4. Modelin Çalıştırılması 16
3.7.5. Testlerin Yapılarak Sorunların İzlenmesi 16
3.8. Web Madenciliği Kavramı 17
3.8.1. Web Sayfa İçerik Madenciliği 18
3.8.2. Web Yapı Madenciliği 18
3.8.3. Web Kullanım Madenciliği 19
3.9. Metin Madenciliği 19
3.10. Belgelerdeki Kelimelerin Değeri 20
4. KARİYER PLANLAMA 23
4.1 Kariyer Kavramı 23
4.2 Kariyer Planlaması 23
4.3. Kariyer Planlamasında Veri ve Metin Madenciliğinin Önemi 23
4.4. Kariyer Planlamasında Endüstri Mühendisliği 24
5. METİN MADENCİLİĞİ İLE ENDÜSTRİ MÜHENDİSLİĞİNDE KARİYER PLANLAMA UYGULAMASI 25
5.1. Çalışmanın Amacı ve Önemi 25
5.2. Çalışmanın Sınırlılıkları 25
5.3. Çalışmanın Amacı Verilerinin Elde Edilmesi 26
5.4. İstatistiksel Analiz Uygulamaları 28
5.4.1. İş Fonksiyonlarının Saptanması 28
5.4.2. İş Fonksiyonlarının Ülkelere Göre Yüzdelik Dağılımları 29
5.4.3. İş Tanımlarının Dağılımları 35

5.4.4. Çalışma Şekillerinin Dağılımları 36
5.4.5. İş İlanlarının Ülkelere Göre Dağılımları 36
5.5. Metin Madenciliği Çalışmaları 38
5.5.1. Metin Madenciliği ile Türkiye'deki İş İlanlarının İncelenmesi 39
5.6. Metinlerin Kategorize Edilmesi 40
5.7. İş İlanlarının Sınıflandırılması 44
6. SONUÇ VE ÖNERİLER 45
KAYNAKLAR 48
KİŞİSEL YAYIN VE ESERLER 53
ÖZGEÇMİŞ 54

ŞEKİLLER DİZİNİ

Şekil 3.1. Veri madenciliği modelleri 9
Şekil 3.2. Veri madenciliği süreci 15
Şekil 3.3. Veri madenciliği aşamaları 16
Şekil 3.4. Web madenciliği alt dalları 18
Şekil 3.5. Metin madenciliğinin ilişkili olduğu alanlar 20
Şekil 5.1. İş akış şeması 26
Şekil 5.2. Elde edilen ikinci veri seti 28
Şekil 5.3. İş fonksiyonlarının dağılımı 29
Şekil 5.4. İş tanımlarının dağılımı 35
Şekil 5.5. İş ilanlarının ülkelere göre dağılımı 37
Şekil 5.6. En çok alım altı ülke 37
Şekil 5.7. Rapidminer programında kurulan model 38
Şekil 5.8. En sık kullanılan kelimelerin özeti 39
Şekil 5.9. En sık kullanılan kelimelerin özetinin pasta grafiği 39
Şekil 5.10. Türkiye'deki iş ilanlarında en çok kullanılan kelimeler 40
Şekil 5.11. İş ilanların Sınıflandırılmasından elde edilen veriler 44

TABLOLAR DİZİNİ

Tablo 5.1. Veri seti tablosu 27
Tablo 5.2. İş fonksiyonlarının ülkelere göre dağılımı..... 29
Tablo 5.3. Çalışma şekillerinin dağılımı 36
Tablo 5.4. İş fonksiyonlarının katagorize edilmesi..... 41
Tablo 5.5. İş ilanlarında aranan özelliklerin niteliği 41
Tablo 6.1. Bilgisayarların özellikleri 46

SİMGELER VE KISALTMALAR DİZİNİ

Kısaltmalar

A.B.D.	: Amerika Birleşik Devletleri
Bt	: Bilgi Teknolojileri
Max	: Maksimum
Ort.	: Ortalama
VM	: Veri Madenciliği

METİN MADENCİLİĞİ İLE ENDÜSTRİ MÜHENDİSLİĞİNDE KARİYER PLANLAMA

ÖZET

Bu çalışmada, Endüstri Mühendisliği ile ilgili iş ilanları dünya çapında araştırılmıştır. Çalışmanın verileri Linkedin'den elde edilmiştir. Verileri elde etmek için web madenciliği yapılmış ve Python 3.10 sürümü kullanılmıştır. Alınan veriler ışığında en çok hangi ülkelerde Endüstri Mühendisi alımı yapıldığı, iş fonksiyonlarının ülkelere göre dağılımı, iş tanımlarının dağılımı veri madenciliği teknikleriyle incelenip, Endüstri Mühendisinden istenen temel yetenekler için iş açıklamalarından metin madenciliği yapılmıştır.

Metin madenciliği çalışmasını derinleştirmek için ise RapidMiner programından yararlanılmıştır. En sık geçen kelime öbekleri tespit edilmiş ve yetenek analizi tablosu oluşturulmuştur. Yetenek analizi tablosu, anahtar kelimeler baz alınarak oluşturulmuş ve yapılan gruplama k-NN algoritmasıyla test edilmiştir. Sonuçlara göre %72,45 doğru bir sınıflandırma yapıldığı görülmüştür.

Temel olarak, dünyada Endüstri Mühendisliği alımlarını, ülkelere göre aranan özellikler araştırılmış ve buna uygun olarak ilerideki çalışmalara katkı sağlamak, üniversitelerin ders içerikleri ile iş dünyasında aranan özellikler arasında köprü kurması amaçlanmıştır.

Anahtar Kelimeler: Büyük Veri, Endüstri Mühendisliği, Metin Madenciliği, Web Madenciliği.

CAREER PLANNING IN INDUSTRIAL ENGINEERING WITH TEXT MINING

ABSTRACT

In this study, job postings related to Industrial Engineering were searched worldwide. The data of the study was obtained from Linkedin. Web mining was performed and Python 3.10 version was used to obtain the data. In the light of the data received, the countries that mostly recruit Industrial Engineers, the distribution of job functions by country, the distribution of job descriptions were examined with data mining techniques, and text mining was performed from job descriptions for the basic skills required from Industrial Engineers.

RapidMiner program was used to deepen the text mining study. The most frequently occurring word phrases were identified and a skill analysis table was created. The talent analysis table was created based on keywords and the grouping was tested with the k-NN algorithm. According to the results, it was seen that a 72.45% correct classification was made.

Basically, the qualifications sought for Industrial Engineering recruitment in the world by country have been researched and accordingly, it is aimed to contribute to future studies and to build a bridge between the course contents of universities and the qualifications sought in the business world.

Keywords: Big Data, Industrial Engineering, Text Minining, Web Mining.

1. GİRİŞ

Büyük miktarda verinin hem kamu hem özel sektördeki çeşitli şirketlerin bilgisayar sistemlerinde depolanması, veri yönetimi ve analizi konusunda yeni teknolojilerin hızla geliştirilmesine öncülük etmiştir. Bu bağlamda ortaya çıkan veri madenciliği teknikleri, farklı nitelikteki bilgiler arasındaki gizli ve önemli ilişkileri keşfetme amacı taşımaktadır. Bu yöntemler, eğitim ortamları dahil olmak üzere çeşitli sektörlerde kullanılmaktadır ve geleneksel veri analizi yöntemlerinden ayrılarak büyük veri setlerini etkili bir şekilde işleyebilme kabiliyeti ile öne çıkmaktadır (Campagni ve diğ., 2015).

Veri Analitiğinin hızla genişlemesi, şirketleri İnsan Kaynakları ihtiyaçlarını ihtiyacında değişime sebep olmuştur. Bu anlamda hangi tür iş rolleri ve becerilerin gerektiği konusunu belirsizleştirmektedir. Çevrimiçi yayınlanan çok sayıda gerçek iş ilanını analiz ederek, Büyük Veri mesleklerinde gerekli olan becerilerin heterojen doğasına açıklık getirebilmektedir (Mauro ve diğ., 2018).

İş ilanları birçok bilgiler içermektedir. Birçok çevrim içi site bulunmakta ve bu sitelerde iş tanımı, aranan nitelikler ile ilgili birçok bilgi bulunmaktadır. Bu bilgiler iş tanımı, şirket kültürü, işin için gereken nitelikler gibi birçok bilgi barındırmaktadır.

Çalışma ortamındaki verilerin birikmesi ve otomasyon seviyesinin iyileştirilmesi ile birlikte, endüstri mühendisliğinin çözmesi gereken sorunlar giderek daha karmaşık hale gelmektedir. Endüstri mühendisliği, bilgisayar ve sistem mühendisliği alanlarına birçok katkı sağlamaktadır. Veri analizi, genetik algoritmalar, yapay zeka vb. gibi konular bu katkılara örnek olarak gösterilebilir.

Endüstri Mühendisliği 20. yüzyılın başında Amerika Birleşik Devletleri'nde ortaya çıkmış ve modern arka plan olarak endüstriyel üretim, araştırma olarak büyük ölçekli endüstriyel üretim ve sosyal ekonomik sistem mühendislik dalıdır. Endüstri mühendisliği, aşamalı olarak inşa edilen ve geliştirilen disiplinler arası bir mühendislik disiplinidir.

Endüstri mühendisliği, insanlardan oluşan bütünleşik sistemleri tasarlayan, geliştiren ve uygulayan bir disiplindir. Matematiğin özel bilgi ve teknolojisini bütünleştirir, fizik ve sosyal bilimler ve doğrulamak için mühendislik analizi ve tasarımının ilke ve

yöntemlerini birleştirir, insan, mali, malzeme, bilgi, zaman ve diğer yönetim kaynaklarını bilimsel bir şekilde Endüstri mühendisliğinin gelişimi birçok aşamadan geçmiştir.

Asıl hedef maliyetleri düşürme, yükleme ve boşaltma sürelerini kısaltma ve en uygun çözümleri sağlama, en ekonomik, etkin ve emek tasarrufu sağlayan çalışma yöntemleri ve standartları kullanılmıştır.

Ardından, Gantt Şeması, Matematiksel İstatistik yöntemi, Değer Akış Haritalama, Stratejik Yerleşim Planlaması, Mühendislik Ekonomik Analizi, vb., Yöneylem Araştırması, Faktör Analizi, Ergonomi gibi birçok yöntem dahil edilmiştir.

Endüstri mühendisliği ABD de ve İngiltere'de Endüstri mühendisliği, Endüstri mühendisliği vizyonu ile tutarlıyken Hollanda da imalat mühendisliğine yakın, İtalyada ise yönetim ve mühendisliği esas almaktadır. Dünyada oldukça farklı alan ve sektörlere hitap eden Endüstri mühendisliği, iş imkanlarını, sektörleri de tahmin etmenin zorluğunu beraberinde getirmiştir.

Endüstri mühendisliğinin uzmanlığına duyulan ihtiyacın önemli derecede kaynak yönetimini içerdiği, eğitim yoluyla mühendis oluşumu için sadece teknik bilginin edinilmesi ve güçlendirilmesinin değil, aynı zamanda hem işverenlerin hem de işverenlerin ihtiyaç duyduğu yeterliliklerin geliştirilmesinin de gerektirmektedir (Palmaa, Ríos ve Guerrero, 2012).

Türkçedeki "kariyer" kelimesi Fransızca "carriere", İngilizce "career" kelimesinden türemiştir. Kariyer, mesleği, yaşam boyu ilerlemeyi ve kişinin yaşamının işini ifade etmek için kullanılır (Akkaya ve diğ., 2015).

Meslek planlaması, bir işçinin sahip olduğu bilgi, yetenek, beceri ve motivasyonlarının geliştirilmesiyle birlikte, çalıştığı organizasyon içindeki ilerleme veya yükselme sürecinin planlanmasıdır (Vergiliel Tüz, 2003).

Kariyer planlama sürecinde bireyin öncelikle kendi potansiyelini tanıması ve güçlü- zayıf yönlerini fark etmesi kritik bir adımdır. Yeteneklerini ve ilgi alanlarını objektif bir şekilde değerlendirerek, hangi mesleğin kişiye daha uygun olduğuna karar vermesi gerekmektedir. Meslek seçimi, bireyin yetenekleri ve ilgileriyle uyumlu olmalı, ancak

aynı zamanda seçilen mesleğin avantajları ve dezavantajları da göz önünde bulundurulmalıdır. Bu süreçte, işin getirdiği fırsatlar, bireyin yetenekleri ile örtüşüyorsa daha başarılı bir kariyer planlaması ortaya çıkabilir. Aynı zamanda, bu seçimin aile ve sosyal faktörlerle de uyumlu olması, kariyerin sürdürülebilirliği açısından önemlidir. Kariyer planlamasının başlangıç noktası olan meslek seçimi, bireyin kendini tanıması, yeteneklerini değerlendirmesi ve çevresel faktörleri dikkate almasıyla sağlam bir temele oturmalıdır (Akkaya ve diğ., 2015).

Teknolojinin gelişmesiyle birlikte istenen yetkinliklerde değişmektedir ve bu nitelikleri takip etmek oldukça zordur. Bu çalışmada Endüstri mühendisliğine genel bakış için veri madenciliği, metin madenciliği çalışmaları yapılarak Kariyer planlamasına katkı sağlaması hedeflenmiştir.

2. LİTERATÜR ARAŞTIRMASI

Bu bölümde uluslararası makalelerin incelenmesi, araştırmaların katkıları üzerinde inceleme yapılmıştır. İş ilanlarının veri ve metin madenciliği ile incelenmesi konusunda ne yazık ki çok fazla bir çalışma bulunmamaktadır.

Bu araştırmanın amacı endüstri 4.0 kuruluşların ve iş ilanların incelenmesi, bu alanda bilgi ve becerilerin saptanmasına yönelik metin madenciliği çalışmasıdır (Pejic-Bach ve diğ., 2020). Linkedin üzerinden endüstri 4.0 ile alakalı iş ilanları kümeleme yöntemi ile sınıflandırılmıştır.

İnsan kaynakları uzmanları, yüksek öğretim kurumları ve endüstri 4.0 kuruluşlarının yararlanmasına yönelik bir çalışmadır.

Farari ve diğ. (2020), çok uluslu bir şirket olan Whirlpool'a ait iş verilerini analiz etmek amacıyla bir çerçeve sunarak, iş tanımlarını, teknolojileri ve becerileri temel alarak her bir iş tanımının Endüstri 4.0'a uyumunu incelemiştir. Bu çalışma, Endüstri 4.0'a hazır ve hazır olmayan her iş modelini değerlendirmek ve firmanın insan sermayesinin Endüstri 4.0 uygunluğunu tahmin etmek için bir temel sağlamaktadır.

Spanda ve diğ. (2022), pazarlama sektöründeki iş bilgileri ile pazarlama kurslarının ders içeriklerini kullanarak metin madenciliği çalışması gerçekleştirmiştir. Çalışmanın temel amacı, iş gücü piyasasındaki açık pozisyonlarda talep edilen becerileri belirlemek ve eğitim sisteminin sektörle uyumunu sağlamaktır. Bu, sektör ile eğitim sistemi arasındaki boşlukları keşfetmeye yönelik bir katkı sağlamıştır.

Li ve diğ. (2021) ABD'deki iş ilanlarının eğilimin saptanması, imalat sektöründeki kritik becerilerin özetlenmesi, beceri ve iş arzı arasındaki boşlukları tespit etme ve yeni nesil üretim sistemi gelişmesine katkı sağlamaktır.

Sodi ve diğ. (2008) iş verenlerin ameliyathane mezunlarında aradıkları özellikler belirlenmek istenmiştir. Çevrimiçi yöntem ile 1000'den fazla toplanan iş ilanı toplamayı ve ameliyathane ile ilgili kelime ve kelime öbeklerinden hiyerarşi oluşturmayı ve hiyerarşi ile iş analizi yapmayı amaçlamıştır.

Lane ve diğ. (2019) ABD'nin dört farklı eyaletini baz alarak gerçekleştirilen çalışma, iş ve veri analitiği programlarının müfredat geliştirmesine ve literatüre katkı sağlamayı amaçlamaktadır. Araştırma, farklı iş kategorilerine yönelik talepler üzerinde odaklanarak dört farklı iş pozisyonunu belirlemiş ve her bir pozisyon için içerik analizi yöntemini kullanmıştır.

Gardiner ve diğ. (2017) büyük veri alanındaki iş ilanları Indeed sitesinden alınarak eksik bilgilerden temizlenmiş ve analiz için Weka programı kullanılmıştır. En sık kullanılan kelimelerin ve sıklıklarının belirlenmesi amacıyla ise 3 aşamalı kümeleme yöntemi uygulanmıştır.

Febriani ve diğ. (2021) üretim ile endüstri ihtiyaçları arasında bir sinerji oluşturmayı hedefler. Akademik hizmetlerin kalitesi ile endüstrinin uyumunu analiz etmek amacıyla önem performans analizi yöntemi kullanılmıştır. Çalışma, öncelikli iyileştirilmesi gereken yetenekleri belirlemeyi amaçlamaktadır.

Karakatsanis ve diğ. (2017) aynı coğrafyada yer alan farklı sektörleri incelenmiştir. Katar, Suudi Arabistan, Umman, Birleşik Arap Emirlikleri ve Bahreyn'de yapılan incelemede, ilk olarak petrol ve gaz endüstrisinde en çok aranan meslekler ele alınmış, daha sonra ise finans alanında en çok tercih edilen meslekler incelenmiştir. Ayrıca, 5 farklı ABD eyaletindeki petrol ve gaz endüstrisi ile bu bölgelerdeki finans çalışmaları da analize dahil edilmiştir. Çalışmanın verileri online iş sitelerinden elde edilmiş olup, Python ve doğal dil işleme teknikleri kullanılmıştır. Çalışmanın temel amacı, farklı sektörlerde ve coğrafyalarda bulunan mesleklerin önemini karşılaştırmak ve daha geniş bir perspektif sunmaktır.

Usabiaga ve diğ. (2022) Polonya'da 2017-2019 yılları arasında çevrim içi sitelerden toplanan iş tekliflerini incelemiştir. Polonya'daki iş gücü piyasasında heterojenlik bulunmakta; bunun sebebinin ise kaliteli mezun sayısındaki düşüş olduğu ön görülmektedir. Metin ve web madenciliği tekniklerinin kullanıldığı bu çalışmada, kümeleme yaklaşımı önerilmektedir. Meslek grupları ve bölgeler arasında ikili bir kümeleme haritası çıkarılmıştır. En çok aranan becerilerin tespit edildiği bu çalışmada, kişilerarası beceriler, organizasyon, ofis ve teknik beceriler en çok aranan nitelikler arasında yer almıştır.

Schedlbauer ve diğ. (2021) bir Alman iş platformu olan STEPSTONE dan veriler R dili kullanılarak alınmıştır. Web madenciliği ve metin madenciliğinin kullanıldığı bu çalışmada lisans derecesi gerektiren işler tespit edilmiş, 147 iş ilanı incelenmiş ve bu işler için gereken yetkinlikler araştırılmıştır. Kullanılan terimler altı farklı gruba ayrılmıştır. Bunlar, Mesleki uzmanlık, sosyal beceriler, takım çalışması, süreçler, öğrenme ve problem çözme yetenekleridir. Sonuçlara göre aranan niteliklerin %45 i mesleki uzmanlık ile ilgili %55 i ise sosyal beceriler ile alakalı olduğu görülmüştür ve bu da sosyal becerilerin ne kadar önemli olduğunu göstermektedir.

Kino ve diğ. (2017) iş eşleştirmede kaliteyi artırmaya yönelik olarak iş ilanlarındaki bilgiler detaylı bir şekilde incelenmiştir. İnceleme kapsamında işe gidip gelme süresi, saatlik ücret, adayın becerileri, iş türü gibi önemli faktörler üzerinde durulmuştur. Metin madenciliği teknikleri kullanılarak gerçekleştirilen çalışmada, pozitif ve negatif anahtar kelimeler belirlenmiş ve bu kelimelerin sıklıkları hiyerarşik küme analizi ile incelenmiştir. Çalışmanın elde ettiği sonuçlara göre, kurumlar personellerine gereken niteliklerde eğitim vererek çalışanlarının ve şirketin performansını artırabilir.

Asano ve diğ. (2015) Japonya'da yürütülen çalışmada, kamu istihdam ofisinden elde edilen veriler metin madenciliği ile birlikte doğal dil işleme teknikleri kullanılarak edilmiştir. İşe alım kararlarını anlamak ve etkileyen faktörleri tespit etmek amacı güdülmüş, aynı zamanda olumsuz iş görüşmelerinin nedenleri üzerinde durulmuştur. Olumsuz durumların nedenleri, yorumlar için ayrılan metin alanlarından alınmıştır. Elde edilen sonuçlara göre, olumsuz durumların başında deneyim, yaş, çalışma koşulları, işe gidip gelme süresi, izlenim ve bilgi gibi faktörler bulunmuştur.

Jackson (2009) Birleşik Krallık, ABD ve Avustralya'daki lisans üstü programlar, beceri açığını tespit etmeyi amaçlamaktadır. Yeni literatür incelenmiş ve bulgular, mezunların yetkinliklerini tanımalarını ve istihdam edilebilirliklerine ilişkin fikir edinmelerini amaçlamıştır. Yazılı ve sözlü iletişim becerileri en önemli yetkinlikler arasında yer alırken, iş ahlakı, karar alma, ekip içinde çalışabilme, geri bildirim alma da diğer arananlardandır.

Aljohani ve diğ. (2022) tarafından yapılan çalışmada, 2010-2021 yılları arasında Scopus'ta yer alan makaleler incelenerek müfredat uyumu ile beceriler ele alınmıştır.

Analizlerde VOSviewer sürüm 16.16 programı kullanılarak terim haritaları oluşturulmuştur. Bu metin tabanlı çalışmanın temel veri kaynağı anahtar kelimelerdir. Yapılan yayınlar üniversiteler, alıntılar ve iş birlikleri açısından tartışılmıştır. Elde edilen sonuçlara göre, "Kovid-19" ifadesi sadece 2020 yılında ortaya çıkmışken, "Endüstri 4.0" terimi 2017 yılında ortaya çıkmış ve 2020 yılına kadar hızlı bir şekilde artış göstermiştir. En çok talep gören beceriler arasında veri bilimi, siber güvenlik, doğal dil işleme ve makine öğrenmesi bulunmaktadır.

Faryna ve diğ. (2022) Bir milyondan fazla çevrim içi benzersiz iş ilanına ait veriler incelenmiştir. Ücret ile işsizlik oranı arasında bir ilişkinin olup olmadığı araştırılmış ve Phillips eğrilerinden elde edilen sonuçlara göre, negatif bir korelasyon olduğu tespit edilmiştir.

Boselli ve diğ. (2018) her ülke için iş piyasasında önemli olan web siteleri uzman bir ekip tarafından belirlenmiştir. Coğrafi bölge, firma, beceri ve meslek gibi girdilere odaklanılmıştır. Becerilere göre en çok aranan meslekler belirlenmiş, coğrafi bölgelere özgü meslek becerileri detaylı bir şekilde incelenmiş ve her meslek için elde edilen ayrıntılardan yararlanılmıştır. Elde edilen sonuçlara göre, üniversiteler ile işletmeler arasındaki ilişkinin güçlendirilmesi amaçlanmıştır.

3. VERİ MADENCİLİĞİ

3.1. Veri Madenciliğine Giriş

Veri madenciliği, büyük miktarlardaki veriyi analiz ederek anlamlı ve yararlı bağlantılar ile gelecekteki olayları tahmin etmeye yardımcı olan bir bilgi çıkarma sürecidir. Bu yöntem, bilgisayar programlarının aracılığıyla veri setleri içinde derinlemesine bir araştırma yaparak, içerdikleri ilişkileri ve kuralları ortaya çıkarır. Veri madenciliği, özellikle çok büyük miktardaki veri setlerindeki karmaşık ilişkileri inceleyerek bu veriler arasındaki bağlantıları bulmada etkili olan bir veri analizi tekniğidir (Kalikov, 2006).

Günümüzde globalleşme ile birlikte birçok bilgi bulunmaktadır. Bu bilgilerin teknoloji sayesinde işlenerek anlamı ve değerli bilgiler olması son derece önemlidir.

Sorunun saptanması ve sorunun çözümü ile geçen dönemde birçok belirsizlik bulunmakta ve ham bilgiye ulaşmak oldukça zordur. Bu durum araştırmacıları ve kurumları veri madenciliğine doğru iteklemektedir.

Veri madenciliği, veriler arasındaki ilişkiyi ve ilişki düzeyini ortaya çıkarır ve gelecek için tahminlerde bulunmayı sağlamaktadır. Veri madenciliği birçok yöntem barındırmaktadır. Yapay sinir ağları, veri tabanı yönetimi, istatistiksel analizleri içermektedir.

3.2. Veri Madenciliğinin Gelişimi

Bilgisayarların kullanımı ile birlikte veri depolanması başlamaktadır. Kullanıcıların isteklerine göre veri depolayan bilgisayarlar ile birlikte veri tabanı ortaya çıkmıştır (Çetin, 2009).

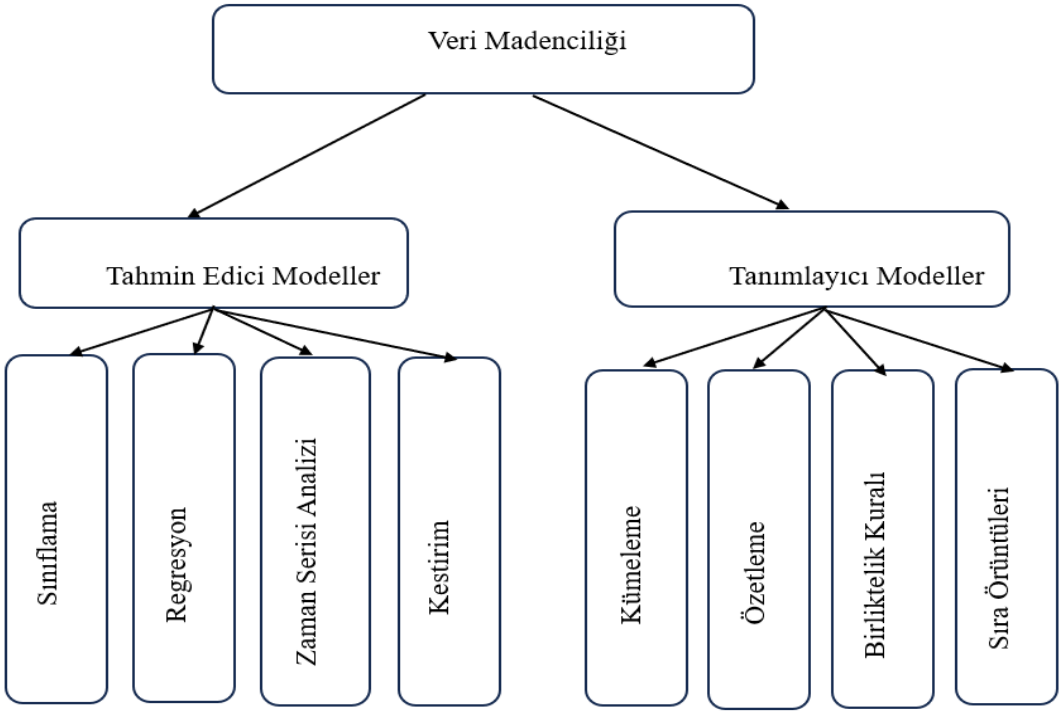
Veri tabanlarının giderek artan bir eğilimde olması sonucu donanımsal olarak veri tabanlarının gelişmesine sebep olurken bu dönmede veri ambarı kavramı ortaya çıkmıştır. Saklanmak istenen veriler fiziksel sürürücülerde saklanmaktaydı. Gittikçe artan veri tabanlarının yöntemi, organizasyonu zamanla güç bir hal almaktadır.

Veri madenciliği terimi ilk defa 1960'lı yıllarda gündeme gelmiş ve amacı, veri analizindeki sorunları gidermektir. Bu dönemde uzun zaman harcanarak büyük veriye

ulaşılabileceği düşünülüyordu. 20. yüzyılda bilgisayar mühendisleri tarafından veri madenciliği adıyla yeni bir akım ortaya çıkarıldı. Veri madenciliğin asıl amacı, geleneksel istatistik yöntemlerinden kurtulup, veri analizinin bilgisayar algoritmik modülleri ile yapılmasını sağlamaktır. Veri madenciliğinin temelinde istatistik, veri tabanı, makine öğrenmesi, araştırma gibi birçok teknik bulunmaktadır.

3.3. Veri Madenciliği Modelleri

Veri madenciliğini, sonuçlarına dair öngörülerde bulunmak, tahminlemek ya da sonucu tanımlamak için kullanılmaktadır (Şekil 3.1).



Şekil 3.1. Veri madenciliği modelleri

3.3.1. Veri Madenciliği Tahminleyici Modelleri

Geçmiş verilerden yola çıkarak, gelecek ile ilgili öngörülerde bulunmayı amaçlamaktadır.

Tahmin edici modeller genellikle sınıflandırma, regresyon, zaman serisi analizi ve tahmin olmak üzere dört temel grupta kategorize edilebilir. Sınıflama modelleri, verileri belirli kategorilere ayırmak için kullanılırken, regresyon modelleri bir değişkenin sayısal değerini tahmin etmek amacıyla kullanılır. Zaman serisi analizi, zamanla değişen verilerin modellenmesi ve kestirim modelleri ise gelecekteki olayları öngörmek için kullanılır. Her bir model türü, belirli bir tahmin görevine uygun olarak kullanılabilir ve veri madenciliği süreçlerinde farklı uygulama alanlarına sahiptir.

3.3.1.1. Sınıflandırma

Temel veri madenciliği işlevlerinden biri, kategorik sonuçları öngörmek amacıyla kullanılan modellerdir. Sınıflama modelleri, öğrenme algoritmaları kullanarak eğitim verilerinden öğrenilen desenleri kullanarak yeni, bilinmeyen verileri belirli kategorilere sınıflandırmak için kullanılır.

Kategorik sonuçları öngörmek amacıyla kullanılan modeller, temel veri madenciliği görevlerinden birini oluşturur. Sınıflama modelleri, öğrenme algoritmaları kullanarak eğitim verilerinden öğrenilen desenleri kullanarak yeni, bilinmeyen verileri belirli kategorilere sınıflandırmak için kullanılır.

Sınıflama, belirli kriterlere göre daha önceden belirlenmiş sınıfların veya kategorilerin atandığı bir veri analizi yöntemidir. Bu kriterler genellikle yaş, cinsiyet, gelir durumu, eğitim düzeyi gibi özellikler olabilir.

Bu sınıflandırma yöntemi, veri madenciliği ve makine öğrenimi uygulamalarında yaygın olarak kullanılır. Örneğin, bir banka müşterisinin kredi başvurusunu değerlendirirken, müşterinin gelir durumu, kredi geçmişi, yaş gibi faktörlere dayalı olarak "riskli" veya "riskli değil" gibi sınıflar atayabilir. Bu sınıflama, karar vericilere müşterileri belirli gruplara ayırmak ve daha iyi stratejiler oluşturmak konusunda yardımcı olabilir. Sınıflama modelleri, belirli özelliklere dayalı olarak veri setindeki desenleri tanımlamak ve öngörmek için kullanılarak bilgi çıkarma süreçlerini destekler (Çetin, 2009).

3.3.1.2. Regresyon

Regresyon, veri madenciliğinde kullanılan önemli bir tekniktir. Basit doğrusal regresyon modeli, Y bağımlı değişkeni ile X bağımsız değişkeni arasındaki ilişkiyi tahmin etmeye odaklanan bir modeldir (Tüzüntürk, 2010).

Gösterimi ise Denklem (3.1)'deki gibi ifade edilmiştir.

$$Y = \alpha + \beta X + u \tag{3.1}$$

3.3.2. Veri Madenciliği Tanımlayıcı Modelleri

Verilerdeki desenleri veya ilişkileri belirleyen tanımlayıcı modeller, analiz edilen verilerin özelliklerini incelemek için kullanılırlar. Bu modeller, tahmin edici modellerin aksine, örneğin, sigorta poliçesini yenilememiş müşterilerin benzer özelliklerini tespit etmek amacıyla yapılan bir kümeleme çalışması gibi, özel analizler için kullanılırlar. Kümeleme, özetleme, birliktelik kuralları ve sıra örüntüleri keşfi modelleri, tanımlayıcı modeller olarak adlandırılır (Aydın ve Özkul, 2015).

3.3.2.1. Birliktelik Kuralı

Veri tabanlarındaki değişkenler arasındaki farklı ilişkileri keşfetmek için kullanılan teknikleri içermektedir (Durakşahin,2017).

3.3.2.2. Kümeleme

Kümeleme analizinin temel hedefi, benzerliklerine göre sınıflandırılmamış verileri gruplamak ve araştırmacıya özetleyici bilgiler sunmaktır. Bu analiz, veri setindeki benzer özelliklere sahip olan gözlemleri bir araya getirerek homojen kümeler oluşturur. Bu sayede, veri setindeki desenleri ve ilişkileri anlamak, benzer özelliklere sahip olan veri noktalarını belirlemek mümkün olur.

Kümeleme analizi genellikle verilerin normal dağılıma sahip olduğu varsayımına dayanır. Ancak, bu varsayım teorik düzeyde kalmakta ve uygulamalarda genellikle ihmal edilmektedir. Pratikte ise genellikle sadece uzaklık değerlerinin normal dağılıma uygunluğuyla yetinmek daha yaygındır. Yani, gözlemler arasındaki uzaklık değerlerinin

belirli bir düzeni takip etmesi ve kümeleme analizi için uygun olması önemlidir, ancak bu değerlerin tam olarak normal dağılıma uyması zorunlu değildir.

Bu yaklaşım, kümeleme analizinin pratik uygulamalarında daha esnek bir çerçeve sağlar ve çeşitli veri türlerinde, normal dağılım varsayımının katı bir şekilde yerine getirilemediği durumlarda da kullanılabilirliğini sürdürmesine imkân tanır (Coşlu, 2013).

3.4. Veri Madenciliğinin Önemi

Günümüzde ekonomik koşulların belirsizliği ve iş dünyasındaki hızlı değişim, karar alıcılar için büyük zorluklar doğurmaktadır. İş deneyimi ve önsezi artık tek başına yeterli değildir, çünkü bu tür kararlar büyük risk taşımaktadır. Yanlış bir karar, bir işletmeyi olumsuz etkileyebilir ve rekabet avantajını kaybetmeye neden olabilir. Bu riskleri azaltmanın ve doğru kararlar almanın etkili bir yolu, bilgiye dayalı yönetimi benimsemektir. İş dünyasında başarılı olmak için bilgi teknolojilerinden yararlanmak kaçınılmazdır. Veri madenciliği yöntemleri, pratik anlamda etkili bir karar destek sistemini kurmak için essiz araçlar sağlar. Bu teknikler, büyük veri setlerini analiz ederek gizli bilgileri ortaya çıkarır ve karar alıcıların daha bilinçli ve doğru kararlar vermesine yardımcı olur (Savaş ve diğ., 2012).

İnternette birçok iş ilanı ve buna bağlı olarak aranan özellikler, işin kişiye uygunluğu gibi birçok veri bulunmaktadır. Artan bilgiler nedeniyle ham bilgiye ulaşmak hayli karmaşık bir durum almaktadır. Ancak firmalar kendilerine uygun personel sıkıntısı çekmekte, iş arayan kişiler ise iş ve niteliğinin hakkında detaylı bilgiye sahip değillerdir.

Birçok işin uzaktan ve hibrid gibi özelliğinin son yıllarda artması nedeniyle iş arama Dünya çapında olmaya başlamıştır. İş veren açısından doğru işe doğru personeli atamak kadar, iş arayan kişinin de kendisine uygun işi seçmesi son derece önemlidir. Veri tabanında yer alan verilerin analizciler tarafından analiz edilmesiyle, istenen bilgilere ulaşıla bilinir, gelecek eğilimleri hakkında öngörülerde bulunabiliniz.

3.5. Veri Madenciliğinin Kullanım Alanları

Günümüzde verilerin saklanmasındaki kolaylık, işlenme maliyetini düşürmesi, gelecek trendlerin öngörülebilmesi acısından veri madenciliğine olan ilgi artmaktadır.

Birçok sektörde veri madenciliği uygulanmaktadır (Baykar, 2006).

Bankacılık ve Finans Alanında,

- Kredi kartı sahtekarlıklarının tespiti
- En karlı müşteri segmentlerinin saptanması
- Pazarlama stratejilerini harcama profillerine göre optimize etme
- Kart değişikliği olasılığı yüksek müşteri analizi
- Müşteri kredi risk değerlendirmesi
- Vergi kaçakçılığı vakalarının belirlenmesi

Pazarlama alanında,

- Pazarlama sektöründe yeni stratejilerin oluşturulmasında
- Pazar sepeti analizinde,
- Müşteri etkileşimlerinin yönetimi,
- Satış tahminlerinde,
- Müşteri alışveriş davranışlarının analizi

Sigortacılık alanında,

- Yeni sigorta poliçesi talep edecek müşterilerin öngörülmesi
- Riskli müşteri profilinin tespitinde
- Dolandırıcılık analiziyle sigorta dolandırıcılıklarının önceden algılanması
- Sigorta risk analiziyle riskli müşteri örüntülerinin çözümlenmesi

Perakendecilik alanında,

- Satış verilerinin değerlendirilmesinde
- Alış-veriş sepeti analizlerinde

Üretim alanında

-Üretim süreçlerinin analizi ve optimizasyonunda
-Boşa harcanan zamanların tespitinde

Kalite alanında

-Hatalı ürün tespitinde
-Müşteriden geri bildirimler alarak, kalite faaliyetlerinin iyileştirilmesinde

3.6. Veri Madenciliğinde Karşılaşılan Sorunlar

Veri tabanından veri alarak analiz yapılan veri madenciliği yönteminde, büyük miktarda veri ile çalışıldığından dolayı birçok sorun ile karşılaşıla bilinmektedir. Bilginin doğru olması, eksik bilgi içermemesi ve boş verilerin olmaması son derece önemlidir.

Bilginin sınırlı olması, verinin eğitilmesini zorlaştırmakta ve doğruluğunu düşürmektedir. Eksik bilgiler ise çoğu paket programlar tarafından analiz edilmeden dikkate alınmamaktadır bu sebeple tam net sonuçlar alınamamaktadır. Alınan verilerin doğru olmaması ise sonuçların yanlış çıkmasına neden olur.

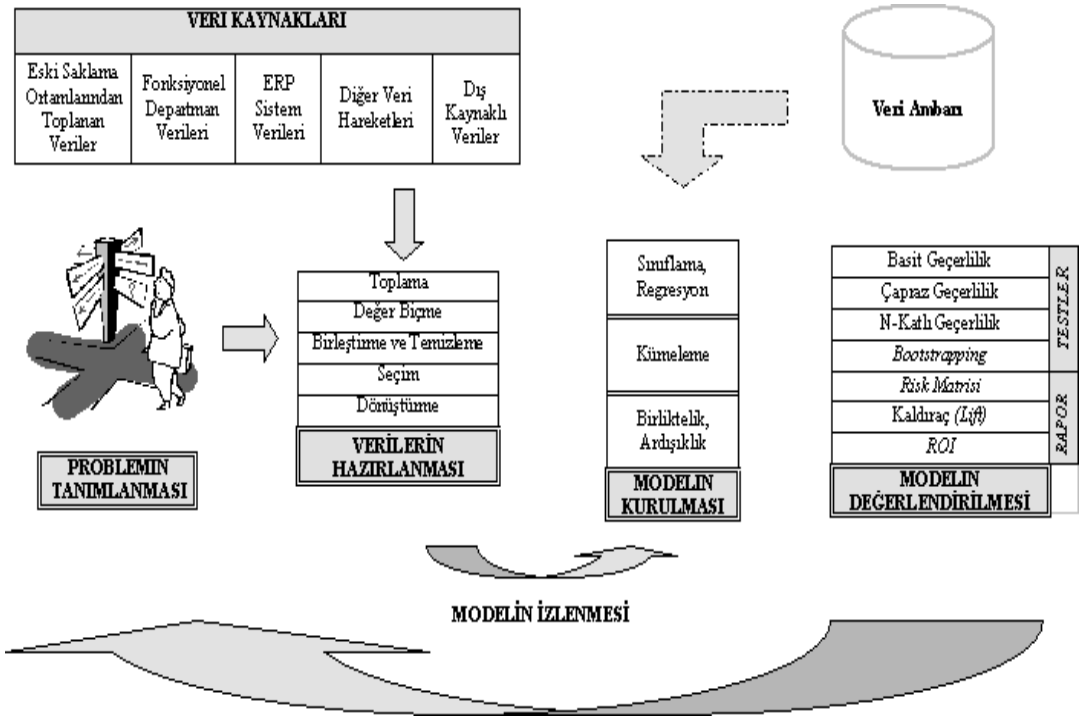
Veri madenciliğinde karşılaşılabilecek sorunlardan bazıları şöyledir (Tozak, 2021),

-Sınırlı bilgi, genellikle veri madenciliği için özel olarak tasarlanmamış veri tabanlarında ortaya çıkabilir. Bu tür veri tabanları, öğrenmeyi kolaylaştırmak amacıyla geliştirilmemiş özelliklere sahip olabilir.
-Belirsizlik, verideki gürültü düzeyi ve yanlışlıkların şiddeti ile ilgilidir. Veri sınıflarındaki veya özelliklerindeki hatalar, gürültü olarak adlandırılır ve bu, veri madenciliği sürecini etkileyebilir. Veri tabanlarındaki bilgi hataları, öznel yaklaşım veya veri toplama hatalarından kaynaklanabilir. Artık veri, çeşitli işlemlerde karşımıza çıkabilen ve genellikle problem çözme sürecinde istenen sonuca ulaşmak için gereksiz niteliklere sahip olan verilerdir. Eksik veri, istatistiksel analizlerde problemlere neden olan bir diğer faktördür. Bu durum, genellikle veri kümesinin büyüklüğünden kaynaklanır. Eksik verilerle başa çıkmak için çeşitli yöntemler kullanılabilir, örneğin eksik veriyi içeren kayıtların çıkarılması, değişken ortalamalarının kullanılması veya var olan verilerden yola çıkarak optimum değerin belirlenmesi vb.
-Boş veri, kendisi de dahil olmak üzere eşsiz veriyi temsil eder. Bu tür veri, analiz ve modelleme süreçlerinde dikkate alınması gereken bir diğer önemli faktördür.

3.7. Veri Madenciliği Süreci

Veri yığınlarından öz bilgiye ulaşabilmek ve faydalı verileri ortaya çıkarmak için veri madenciliği çalışmasının uygun bir yöntem uygulanmasına bağlıdır. Temel olarak 5 aşamadan oluşmaktadır (Şekil 3.2).

Veri madenciliği süreci şekil Şekil 3.2'de görüldüğü gibidir.



Şekil 3. 2. Veri madenciliği süreci

3.7.1. Problemin Tanımlanması

Veri madenciliği çalışmalarında etkili olmanın en temel koşulu, projenin amacının net bir şekilde belirlenmesidir. İlgili hedef, konuya özgü bir problemi net bir şekilde ifade etmeli ve projenin başarı düzeyini ölçmenin yolları belirtilmelidir (Turgut, 2012).

3.7.2. Verilerin İşlenmeye Uygun Şekilde Hazırlanması

Modelin oluşturulması aşamasında karşılaşılan zorluklar, genellikle bu aşamaya geri dönülmesine ve verilerin yeniden düzenlenmesine neden olabilir. Bu durum, veri keşfi sürecinin genelinde, veri hazırlığı ve modelin oluşturulması aşamalarına odaklanan bir analistin zamanının ve eforunun %50 ila %85'ini harcamasına yol açar. Verilerin

hazırlanması aşaması, toplama, değerlendirme, toplama ve temizleme, seçme ve dönüştürme aşamalarını içerir (Çetin, 2009).

3.7.3. Probleme Uygun Modelin Kurulması ve Değerlendirilmesi

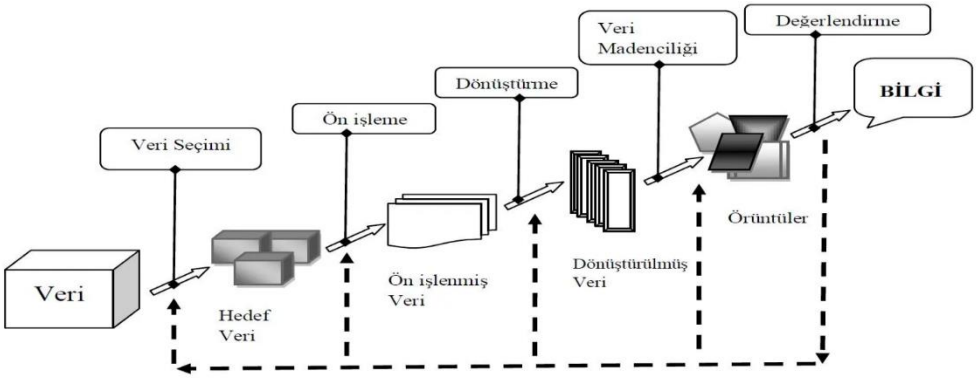
Belirlenmiş problem için optimum modelin belirlenmesi, mümkün olan en fazla sayıda modelin kurularak test edilmesiyle elde edilebilir. Bu nedenle, veri hazırlama ve model kurma aşamaları, en iyi olduğuna inanılan modele ulaşılıncaya kadar sürekli olarak tekrarlanan bir süreçtir (Savaş ve diğ., 2012).

3.7.4. Modelin Çalıştırılması

Oluşturulan ve güvenirliliği onaylanan model, sadece direkt bir uygulama olarak kullanılmakla kalmaz, aynı zamanda başka bir uygulamanın bir parçası olarak da entegre edilebilir. İşletme uygulamalarında, kredi değerlendirmesi, dolandırıcılık tespiti, risk analizi gibi direkt olarak kullanılabilirken, Kampanya planlaması simülasyonuna dahil edilebilir veya beklenen stok seviyelerinin otomatik siparişle karşılanmasın sağlayacak bir uygulamanın içine gömülebilir (Dumouchel, 1999).

3.7.5. Testlerin Yapılarak Sorunların İzlenmesi

Zaman içinde, tüm sistemlerin genel özelliklerinde ve bu nedenle oluşturdukları verilerde meydana gelen değişiklikler, oluşturulan modellerin sürekli olarak izlenmesini ve ihtiyaç duyulması halinde yeniden düzenlenmesini gerektirecektir. Tahmin edilen ve gözlenen değişkenler arasındaki farklılıkları gösteren grafikler, model sonuçlarını etkili bir şekilde takip etmek için kullanılabilecek önemli bir yöntemdir (Dumouchel, 1999).



Şekil 3. 3. Veri madenciliği aşamaları

3.8. Web Madenciliği Kavramı

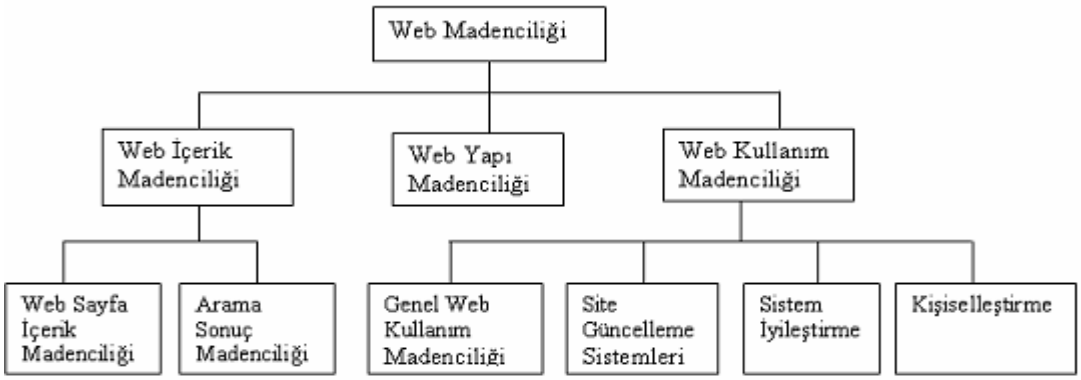
Günümüzde, arama motoruna yazdığımız herhangi bir kelimenin sonrasında sosyal medyada reklam olarak karşımıza çıkması veya sosyal ağlarda daha önce profilini incelediğimiz biri için sonradan arkadaş eklenmesi gibi birçok kullanımın temelinde web madenciliği bulunmaktadır. İnternet üzerinde herhangi bir siteye erişim sağladığımızda, giriş log kayıtlarının tutulduğunu ve bununla birlikte kullanıcı kayıtları, oturum bilgileri ve site içeriklerinin veri yığını halinde olduğunu bilmekteyiz. Bu veri yığınlarından etkili bir şekilde faydalanmak için bilgi çıkarımı yapılması gerekmekte ve bu da web madenciliği ile mümkün olmaktadır. Yapılan analizlerle siber güvenlik, ekonometri, arama motorları optimizasyonu ve hatta açık kaynak istihbaratı gibi birçok alanda geniş bir bilgi çıkarımı yapılabilir (Ateş, 2021).

Veri madenciliğinin çeşitli uygulama alanları olmakla birlikte web alanına uygulanan ve web verilerini girdi olarak kullanan türüne web madenciliği adı verilmektedir (Canay ve Kocabıçak, 2023).

Web madenciliği, veri madenciliği teknikleri kullanılarak web sitelerinden anlamlı ve yararlı bilgiler çekmektir. Çıkarılan bilgiler ile site sahibine maddi katkılar sağlayabileceği gibi internet sayfalarının farklı düzenleyerek siteyi kullanan kişilerinde memnuniyetini artırmaktadır. Kullanıcıların sayfaya kayıt olma sırasında verdiği bilgiler, müşteri yorumları, beğenilen ürünler, satış miktarları gibi birçok bilgi edinile bilinir ve bu bilgilerden ise veri analizi çalışmaları yapıla bilinilir.

Web sitesinin bağlı bulunduğu sunucular ile kullanıcının her tıkladığı alan birden fazla alan biriktir. Ancak bu kayıtlar düzenli olarak temizlenir. Ancak bu veriler sitenin içeriği ve kullanıcının profiliyle bağdaştırıldığında yararlı birçok bilgi sağlanabilmektedir.

Web madenciliğini 3 ana alanda sınıflandırmak mümkündür. Şekil 3.4'de gösterilmiştir.



Şekil 3.4. Web madenciliği alt dalları

3.8.1. Web Sayfa İçerik Madenciliği

Web içerik madenciliği, web sayfalarının içeriğinden, çeşitli türlerde bilgiler elde etmek için kullanılır. Bu bilgiler metin formatında mümkün olduğu gibi, sayfalardan ses, resim, video gibi farklı biçimlerde bilgilerin de çıkarımını sağlar (Kuşak, 2010).

Verinin farklı tiplerde oluşu ve yapısal olmayışı, bu alandaki tekniklere daha karmaşık bir yaklaşım getirir. Web içerik madenciliği, metin madenciliği ve veri madenciliği kavramları birbirine benzer olabilir, ancak aralarında bazı farklılıklar bulunmaktadır. Web içerik madenciliği, veri madenciliği ile bağlantılı bir kavramdır çünkü web dokümanları içindeki verileri çıkarmak için veri madenciliği tekniklerini kullanır. Veri madenciliğinde genellikle yapısal veriler kullanılırken, web verileri kısmen yapısal ve yapısal olmayan veriler içerir. Benzer şekilde, web içerik madenciliği, metin madenciliği ile ilintilidir çünkü web üzerindeki bilgilerin büyük bir kısmı metin tabanlıdır. Ancak, web içerik madenciliği ile metin madenciliği arasındaki fark, metin madenciliğinin tamamen yapısal olmayan verilere odaklanmış olmasıdır.

3.8.2. Web Yapı Madenciliği

Web yapı madenciliği, bir web sitesinin düzenini geliştirmek amacıyla kullanılır ve odak noktasını siteler arasındaki bağlantılara yönlendirir. Bu, web sayfaları arasındaki bağlantıları ortaya çıkarmakla ilgilenir. Başka bir deyişle, HTML kodlarındaki <a href> etiketleri arasındaki veriyi analiz eder. İçerik madenciliği, bir belgenin içeriğini esas alırken, yapı madenciliği belgeler arası ilişkilere odaklanır (Arslan, 2008).

3.8.3. Web Kullanım Madenciliği

Web kaynakları üzerindeki kullanım günlükleri üzerinden, kullanıcılara dair kaynağa erişim modellerinin keşifsel analizi, web sitesini kullanan kişilerin ziyaretçi eylemlerine dair kayıtların ilgili web sunucularında web günlüğü olarak kaydedilmesini içerir. İncelenen kullanım içerikleri bireysel kullanıcıları veya kullanıcı gruplarını kapsayabilir. Bu kayıtlar üzerinden hareketle, web sitelerinin hızlı ve kolay kullanımını sağlamak için yeniden düzenlenmesi, bağlantıların kullanım ve eğilimlere göre optimize edilmesi gibi çalışmalarla web sitelerinin verimliliği artırılır. Analizde, kullanıcı profillemesi ve sık kullanılan öğelerin belirlenmesi için kümeleme ve birliktelik kuralları aktif olarak kullanılmaktadır (Ateş, 2021).

3.9. Metin Madenciliği

Metin madenciliği, metin verilerinden yararlı ve daha önce tespit edilmemiş bilgilerin çıkarılması için bilgi işlem yöntemleri ve teknikleri kullanılarak gerçekleştirilen bir süreçtir (Beyhan, 2014).

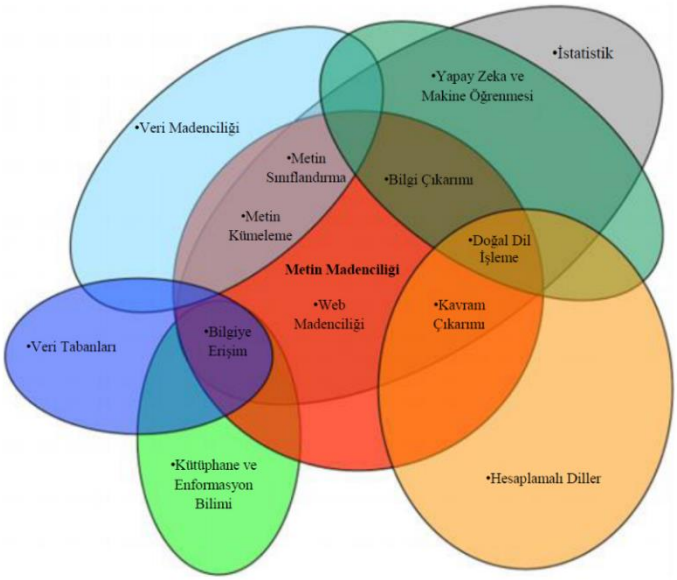
Metin madenciliği, büyük miktardaki metinden anlamlı bilgi ve verilerin çıkarılmasıyla ilgilenir. Bu alandaki yöntemler, yapılandırılmamış veya metinsel bilgileri analiz etmek ve araştırmak için kullanılır. Metin madenciliği, metni verilere dönüştürmek ve daha sonra diğer veri madenciliği teknikleri ile analiz etmek için önemlidir. Çoğu zaman büyük veri setlerini manuel olarak analiz edemeyecek kadar büyük olduğundan, metin madenciliği tekniklerine başvurulur. Belirli bir metindeki bilgileri tanımlamak ve çıkarmak için metin madenciliği teknikleri veya yöntemleri kullanılır. Bu yöntemler aynı zamanda verileri belirli kriterlere göre kümeler halinde gruplandırarak analiz etmeye yardımcı olur. Kapsamlı verileri belirli kriterlere göre analiz etmek ve doğrulamak gerektiğinde, metin madenciliği yöntemlerine başvurulur. Metin madenciliği, metinsel bilgilerin incelenmesi ve analiz edilmesi için istatistiksel, dilbilimsel ve makine öğrenimi tekniklerini içeren daha derin veri analizi, araştırma ve inceleme için gereklidir (Haşimi ve diğ., 2015).

Metin madenciliği, kurumsal dokümanlar, müşteri değerlendirmeleri, web sayfaları ve XML dosyaları gibi yapısal olmayan verilerden önceden bilinmeyen ve potansiyel olarak

faydalı bilgilerin keşfedilmesini sağlar. Elde edilen bilgilerle, analiz edilecek metin kaynaklarında açıkça görülmeyen ilişkiler, hipotezler veya eğilimler ortaya çıkarılabilir. Veri madenciliği büyük verilerin toplanması, işlenmesi ve bu bilgilerden önemli sonuçlar çıkarılmasıdır. Aralarında gruplandırıla bilinen veriler yapısal verilerdir ve kendi içlerinde arama yapıla bilinir (Ye ve diğ., 2023).

SQL. Oraclle ve Access en çok kullanılan veri tabanlarıdır. Metin dosyaları, e-mailler, pdf dokümanları yapısal olmayan verilerdir ve bu verilerin tanımları yoktur. Metin madenciliği, veri madenciliği alanına ait bir alt dal olarak kabul edilir ve doğal dil ile yazılmış metinlerin analizini içerir. Metin madenciliği, metinlerin sınıflandırılması, metinlerden konu çıkarılması, metinin kümelenmesi gibi birçok konuyu hedefler.

Metin madenciliği temelde 3 aşamadan oluşur. İlk aşamada özetleme yapılır. Metinler kategorilerine göre ayırmak için kütüphanelerden yararlanılır. İkinci aşama ise bilgi edinimidir. Üçüncü aşama ise metin verilerinden, anlamlı ve yararlı bilgilerin çıkarılmasıdır, Metin madenciliği birçok alan ile ilişkilidir (Şekil 3.5).



Şekil 3. 5. Metin madenciliğinin ilişkili olduğu alanlar

3.10. Belgelerdeki Kelimelerin Değeri

Birçok veri madenciliği uygulamasında, belgelerin ana başlıklara göre gruplandırılması gerekir. Belgeleri konulara göre belirlemek için kullanılan anahtar kelimeler vardır. Belge

üzerinde en çok geçen kelime o metin için en önemli konu olarak görülmektedir. Ancak bazı durumlarda kullanılan bağlaçlar ile ve vb. gibi bu kelimelerin dikkate alınması doğru değildir. ‘Kelimelerin belge içerisinde az kullanılması önemsiz olduğu anlamına gelmez. Kelimelerin bir veya birden fazla belgede bulunma sıklığına bakılmalıdır.

Metin madenciliği uygulamalarına örnek vermek gerekirse,

1. Konu izleme, kullanıcı profilleri ve kullanıcı görüşleri temel alınarak, kullanıcının ilgisini çekebilecek diğer belgelerin tahmin edilmesidir. Örneğin, sosyal ağlarda kullanıcı profillerine dayalı olarak farklı kullanıcılara farklı reklamlar gösterilmesi, bu konunun bir örneğidir.
2. Özetleme, Yazarların çalışmalarının özetlenmesi ve anahtar sözcük çıkarımı, metin madenciliği uygulamalarının bir parçasıdır. Bu yöntemler, bir yazarın metnini özetleyerek içerisindeki örüntüleri tespit etmeyi ve çeşitli çıkarımlarda bulunmayı mümkün kılar (Abdala ve Angın, 2022).
3. Sınıflandırma, Temel hedefi, belirli bir nesnenin özelliklerini açıklamak ve bu yeni nesnenin daha önce tanımlanmış sınıf kümesinden birine atanmasını gerçekleştirmek, metin sınıflandırma alanında günümüze kadar birçok farklı yöntemin kullanılmasına neden olmuştur. Bu yöntemler arasında şunlar bulunmaktadır:

Bu yöntemler arasında şunlar bulunmaktadır (Kaşıkçı ve Gökçen, 2014).

-K-En Yakın Komşu Sınıflandırması (k-NN): Bir örneğin, komşularının sınıflarına dayanarak sınıflandırılması prensibine dayanır.
-Naive Bayes Olasılıklı Sınıflandırma: Bayes teoremi temel alınarak çalışan bir olasılık tabanlı sınıflandırma yöntemidir.
-Karar Ağaçları: Veri kümesini belirli bir hedef değişkenine göre sınıflandırmak için kullanılan ağaç yapılarıdır.
-Yapay Sinir Ağları: Beyin yapısını taklit eden bir model kullanarak öğrenen algoritmalar içerir.
-Destek Vektörleri (SVM): Sınıflar arasında bir karar sınırı çizerek, veri noktalarını sınıflandıran bir yöntemdir. Ayrıca, birçok çalışmada birden fazla yöntemin bir arada kullanılıp sonuçlarının karşılaştırılması da yapılmıştır. Bu, sınıflandırma performansını

artırmak veya belirli bir metin sınıflandırma görevine en uygun yöntemi belirlemek amacıyla gerçekleştirilebilir.

Kümeleme, metin madenciliğinde önemli bir yöntemdir. Kümeleme, belirli bir kategori kümesine sahip olmadan benzer belgelerin bir araya getirilmesidir. Bu noktada karar ağaçları, makine öğrenmesi, istatistik gibi çeşitli teknikler kullanılır. Karar ağaçları, yapay sinir ağları, bulanık mantık, yaklaşık kümeleme ve içerik tabanlı öğrenme gibi teknikler, kümeleme için yaygın olarak kullanılan yöntemler arasında yer alır. Metin kümeleme, doküman koleksiyonlarının doküman benzerliklerine bağlı olarak kümelere ayrıştırılması işlemidir. Bu süreç sonucunda, bir küme içindeki dokümanların genellikle benzer bir konuda olmaları beklenir. Kümeleme hipotezine göre, birbiriyle ilgili dokümanlar, ilgisizlere göre birbirine daha çok benzerdir. Bu hipotezden yola çıkarak, doküman kümeleme yöntemleri kullanılarak geniş doküman koleksiyonları içinde dolaşım ve arama işlemlerinin iyileştirilmesi amaçlanmaktadır (Tunalı, 2011).

4. KARİYER PLANLAMA

4.1 Kariyer Kavramı

Kariyer , Latince ‘Carrus’ (at arabası) ve ‘Carrera ‘ (yol) , Fransızca ‘Carrierre’ (Yarış Yolu) , İngilizce ‘Carreer’ (Meslek) kelimelerinden gelmektedir. Geçmiş zamanlarda atlı arabaların kat ettikleri yol manasında kullanılmaktaydı. Şu an ise bir personelin bir organizasyondaki kişisel gelişimi anlamında kullanılmaktadır.

4.2 Kariyer Planlaması

Kişilerin Kariyer kavramı ile tanışmaları eğitim öğretim hayatlarının başlaması ile birlikte başlamayıp iş hayatlarının sonuna kadar devam etmedir. Kişiler nasıl bir hayat süreceklerinin planlamasını yapıp bu doğrultuda kariyer planlamasına başlayıp bu planlar ışığında geleceklerinin mimarisini tasarlamaktadırlar. İstedikleri hayat ve mesleki standartlarla ile buluştuklarında ise kariyer planları gerçekleşmiş olacaktır.

4.3. Kariyer Planlamasında Veri ve Metin Madenciliğinin Önemi

Etkili bir kariyer sistemi, bireyin kariyer çapası ile örgütün iş gereksinimleri arasında iyi bir eşleştirme sürecini gerektirir (Schein ve Maanen, 2016).

Gelişen endüstriyel süreçler ve kaynakların azalmasıyla birlikte endüstriyel çevre daha çok kıt kaynaklara itilmiş ve bu sebeple her geçen saniye insan etkinliği ve verimliliği daha önemli noktaya gelmiştir. Bu doğrultuda kişisel gelişim ve donanım her geçen gün daha önemli bir noktaya gelmiştir. Kişinin kendini tanıması ve sektördeki diğer rakiplerinden ayıran pozitif ve negatif özelliklerinin fark ederek bu kapsamda daha verimli olabileceği iş pozisyonlarını araştırması ve iş hayatında kendini huzurlu hissedebileceği veya verimsiz kılacağı işleri elemine etmesine olanak sağlanabilir. Kişinin istenen nitelikleri saptaması bu doğrultuda akademik ve kişisel bilgileri edinmesi gerekmektedir. Nitekim kişisel niteliklerin yanı sıra aynı doğrultuda gelişen sektörel ihtiyaçlar, talepler, personel donanımları ve özellikler de değerlendirme kriterleri içerisine girdiğinde çok büyük bir bilgi yığınını oluşturmaktadır. Verilerin özetlenmesi, yorumlanması ve anlamlı birer veri haline gelmesi için veri madenciliği teknikleri kullanılabilir. Oluşan büyük metin yığınlarının düzenlenmesi, kategorileştirilmesi ve

anlamlı bir veri haline getirilmesinde metin madenciliği bu konuda önemli bir yöntem olarak karşımıza çıkar.

Veri madenciliği bu noktada sektörel ve Kişisel kariyer planlama da hassas bir öneme sahip olup bu verilerin bir okunan kaynak noktası haline getirip kişilere ve bu personelleri yetiştirecek olan üniversitelere bir yol gösterici olması hususunda derin bir öneme sahiptir.

4.4. Kariyer Planlamasında Endüstri Mühendisliği

Endüstri mühendisleri, üretim, bankacılık, finans, planlama, üretim geliştirme, insan kaynakları yönetimi, süreç yönetimi, yazılım, sistem analizi, ergonomi gibi birçok alanda görev alabilmektedir. Endüstri mühendisleri, sistem yaklaşımı, yaratıcılık, analitik düşünme, etkili iletişim, yönetebilme kabiliyeti gibi birçok farklı özelliğe sahiptir.

Çalışma sektörlerini incelediğimizde, demir çelik, hazır giyim, metal, otomotiv, inşaat, yapı, Ar-Ge, gıda, lojistik, hazır giyim, otomotiv ve benzeri farklı alanlarda çalışabilen endüstri mühendisleri bulunmaktadır. Endüstri mühendisliği eğitimi sürecinde ise programlama, istatistik, süreç iyileştirme, optimizasyon, yöneylem, sistem analizi, benzetim modelleri, dizayn programları, mühendislik ekonomisi, lojistik, tesis planlama, kalite yönetimi, termodinamik başta olmak üzere birçok konuda bilgi sahibi olmaktadırlar.

İş imkanları oldukça geniş bir yelpazeye sahip olan Endüstri Mühendisliği, kişinin yetenekleri doğrultusunda kendisini ve işini geliştirebileceği, mutlu olabileceği bir seçim yapması açısından büyük öneme sahiptir. Kariyer planlaması, bu noktada Endüstri Mühendisliği için hayati bir öneme sahiptir.

5. METİN MADENCİLİĞİ İLE ENDÜSTRİ MÜHENDİSLİĞİNDE KARİYER PLANLAMA UYGULAMASI

5.1. Çalışmanın Amacı ve Önemi

Değişen dünya koşulları,globelleşme ve teknolojinin hızla gelişmesi sonucunda, şirketler için en önemli faktör nitelikli ve doğru kişinin seçimidir.

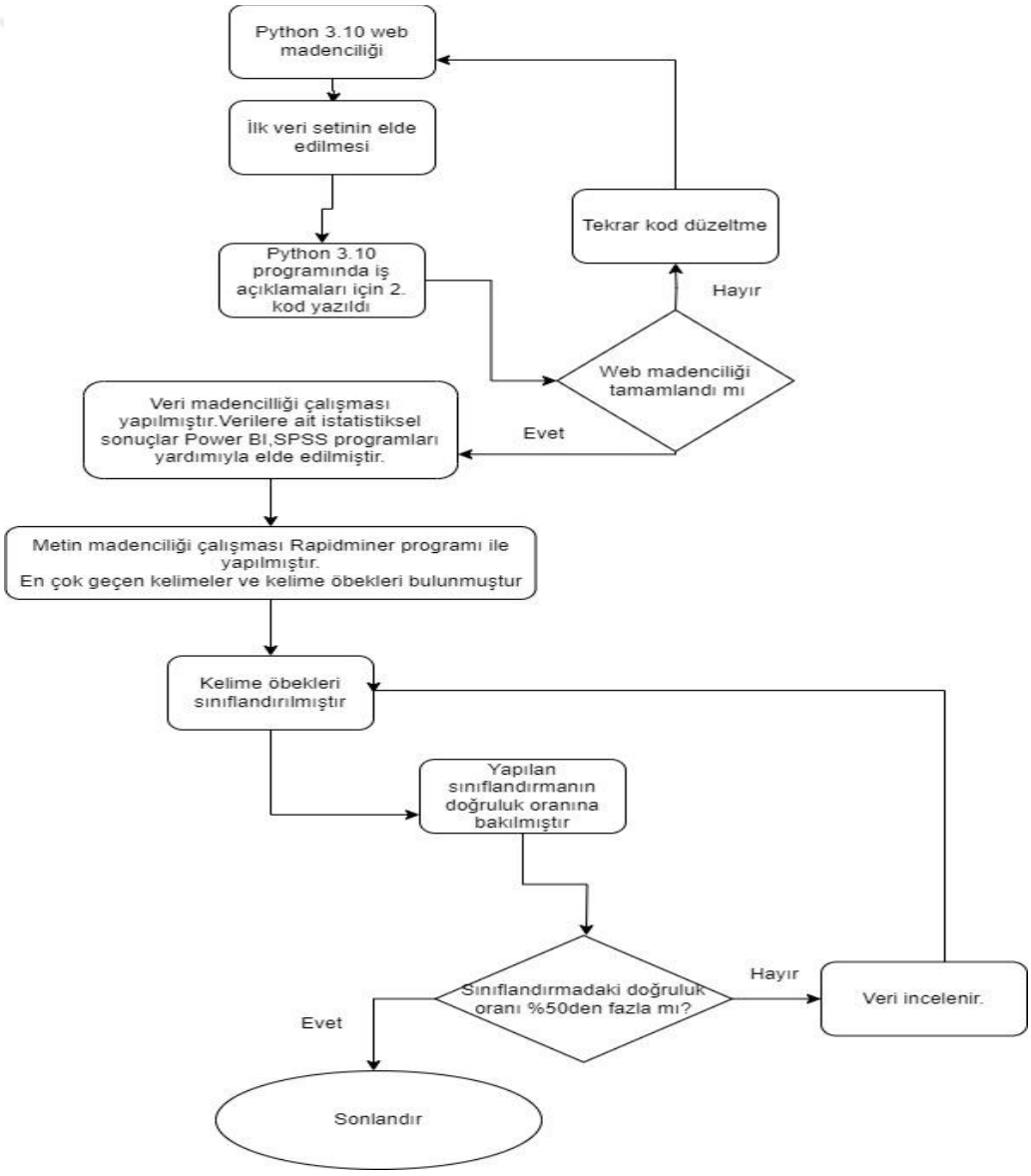
Endüstri mühendislerinin kariyer seçeneklerini, diğer mühendislikler ile karşılaştırdığımızda karşımıza birçok sektör ve iş tanımı çıkmaktadır. Her farklı mühendislik dalının bilmesi gereken bazı donanım ve programlar var iken endüstri mühendisliğinde, sektöre, firmaya, pozisyona bağlı olarak değişmektedir. Endüstri mühendisleri birçok farklı disiplin ile ile ilgili eğitim almaktadır ve bu bilgilerin derinlemesine verilmemesi Endüstri mühendisleri için her alanda bilgisi bulunmakta ancak hiçbir işi tam olarak yapamadıklarına dair bir düşünceye sebep olmaktadır. Endüstri mühendisliği hem öğrenciler hm de şirketler tarafında halen anlaşılabilmiş değildir. Ne yazık ki günümüzde halen Endüstri veya Makine mühendisi gibi diğer mühendisliklerin yerine geçebilecek ilanlar görmekteyiz. Sektörde Endüstri mühendisinden ne gibi nitelikler bekleniyor? Tüm Dünyaya baktığımızda Endüstri mühendisliği hangi ülkelerde daha çok iş imkanına sahip? Çalışma şekilleri nelerdir? Hangi pozisyon için en çok alım yapılıyor? gibi birçok soruya yanıt bulmak ve sektörde Endüstri mühendisliğinin önemi saptamak ve aday öğrencilere yol göstermesi amaçlanmıştır.

5.2. Çalışmanın Sınırlılıkları

Çalışma tüm Dünyadaki Endüstri mühendisliği iş ilanlarını baz almaktadır ve araştırma İngilizce olarak aratılmıştır. Araştırmanın verileri Linkedin programından alınmıştır. Ancak bazı ülkeler Linkedin dışında başka kariyer sitelerini de kullanmaktadır. Her ülke için diğer kariyer sitelerini de birleştirerek bir veri tabanı oluşturmak, zor olduğundan sadece en çok kullanılan ve popüler olan Linkedin programı örnek uzay için seçilmiştir.

5.3. Çalışmanın Amacı Verilerinin Elde Edilmesi

Yapılan bu çalışmada Endüstri Mühendisliği ile ilgili Dünya çapında bir araştırma yapılmıştır. İlk olarak 23.04 ve 23.05 tarihleri arasında "Industrial engineer" olan sadece İngilizce olarak arama yapılmıştır. Veri setini elde etmek için Linkedin kullanılmıştır. Linkedinden Pyhton 3.10 programı ile 24593 adet veri çekilmiştir. Verileri elde etmek için iki aşamalı bir kod yazıldı. Kodumuzda ilk olarak İş ilanın adı, ilana ait link,geliri,maaşı.tarih alanı, firmanın adı,maaş düzeyi ve çalışma şekli gibi bilgiler alınmıştır. Daha sonra ikinci bir kod yazılmış ve iş açıklamaları kısmı alınmıştır. Genel olarak iş akış şeması verilmiştir (Şekil 5.1).



Şekil 5.1. İş akış şeması

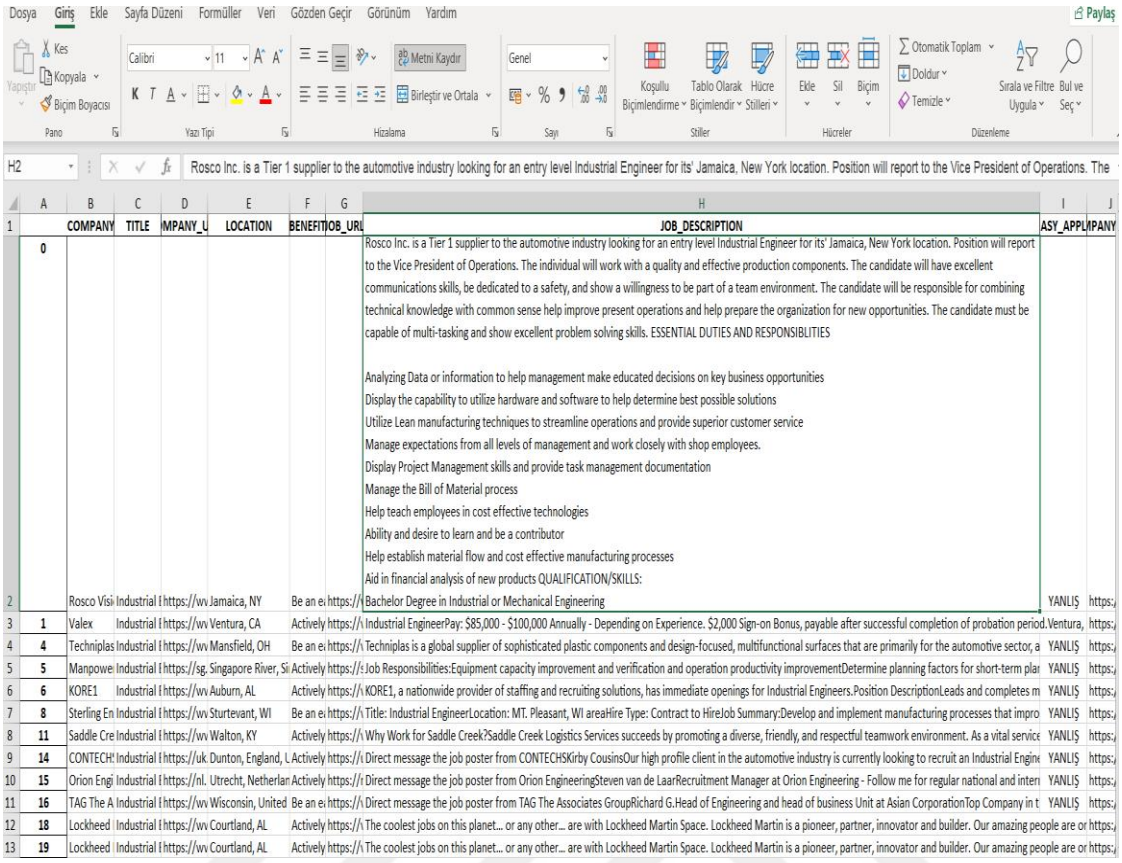
Tablo 5.1. Veri seti tablosu

İş Adı	İlanın Linki	Firma adı	Konum	Maaş	Başlangıç	Bitiş	Çalışma şekli
Manufacturing Process Engineer	https://uk.linkedin.com/jobs/view/manufacturing-process-engineer-at-manpower-uk-3592119604?refId=TS2SwsUgsXXZQWboKw4Z3A%3D%3D&trackingId=aAbUldYDwaNJR24lNmODdQ%3D%3D&position=2&pageNum=0&trk=public_jobs_jserp-result_search-card	Manpower UK	Cambridge, İngiltere, Birleşik Krallık	40.000,00 GBP - 42.000,00 GBP	2023-04-30 11:04:30	2023-05-02 20:11:46	On-site
Manufacturing Engineer	https://www.linkedin.com/jobs/view/manufacturing-engineer-at-tata-technologies-3594552104?refId=TS2SwsUgsXXZQWboKw4Z3A%3D%3D&trackingId=ZqX8GAPeNhUumcnPoFHzDw%3D%3D&position=4&pageNum=0&trk=public_jobs_jserp-result_search-card	Tata Technologies	Lafayette, Indiana Metropol Bölgesi		2023-04-18 00:00:00	2023-05-02 20:11:46	On-site
Manufacturing Process Engineer	https://ie.linkedin.com/jobs/view/manufacturing-process-engineer-at-celestica-3589380264?refId=TS2SwsUgsXXZQWboKw4Z3A%3D%3D&trackingId=sXfGbWoo7vhqM%2F1Wj5B0Xg%3D%3D&position=5&pageNum=0&trk=public_jobs_jserp-result_search-card	Celestica	Galway, Galway, İrlanda		2023-04-26 12:20:08	2023-05-02 20:11:46	On-site
Manufacturing Process Engineer	https://uk.linkedin.com/jobs/view/manufacturing-process-engineer-at-premier-group-recruitment-3593001156?refId=TS2SwsUgsXXZQWboKw4Z3A%3D%3D&trackingId=X1CvJ3HOxpKSs4rxdSkzxw%3D%3D&position=6&pageNum=0&trk=public_jobs_jserp-result_search-card	Premier Group Recruitment	Marlow, İngiltere, Birleşik Krallık	50.000,00 GBP - 55.000,00 GBP	2023-04-29 14:45:49	2023-05-02 20:11:46	On-site
Industrial Engineer	https://www.linkedin.com/jobs/view/%C4%B1ndustrial-engineer-at-sterling-engineering-3594951140?refId=TS2SwsUgsXXZQWboKw4Z3A%3D%3D&trackingId=BxMKGFzZmCFdaCLjzLIfcQ%3D%3D&position=11&pageNum=0&trk=public_jobs_jserp-result_search-card	Sterling Engineering	Sturtevant, Wisconsin, Birleşik Devletler	30,00 USD - 45,00 USD	2023-04-19	2023-05-02 20:11:46	On-site

Tablo 5.1'de görüldüğü gibi "Industrian engineer" dışında, "Manufactoring engineer" yada "Proses engineer" gibi iş ilanları da gelmiştir. Bu ilanlar incelendiğinde ise Endüstri mühendisliği, makine, malzeme mühendisliği gibi diğer alanlarında başvurabildiği görülmektedir. Bu yüzden bu veriler tablodan çıkarılmıştır.

Sadece Endüstri mühendisliği iş ilanları geldiğinde toplam 5331 adet ilan elde edilmiştir.

Python 3.10 programında iki aşamalı yazılan bu kod ilk aşamada iş ilanının ismini, çalışma şekli, Firma ismi gibi detayları almıştır. İkinci aşama ise iş açıklamaları olan kısmı almıştır burada veriler Şekil 5.2'de gösterilmiştir. Bu şeklide yapılmasının amacı ise çok büyük bir veri ile çalışılmıştır.



		COMPANY	TITLE	MPANY_U	LOCATION	BENEFIT	JOB_URL	JOB_DESCRIPTION	ASY_APPL	IPANY
	0							Rosco Inc. is a Tier 1 supplier to the automotive industry looking for an entry level Industrial Engineer for its' Jamaica, New York location. Position will report to the Vice President of Operations. The individual will work with a quality and effective production components. The candidate will have excellent communications skills, be dedicated to a safety, and show a willingness to be part of a team environment. The candidate will be responsible for combining technical knowledge with common sense help improve present operations and help prepare the organization for new opportunities. The candidate must be capable of multi-tasking and show excellent problem solving skills. ESSENTIAL DUTIES AND RESPONSIBLITIES Analyzing Data or information to help management make educated decisions on key business opportunities Display the capability to utilize hardware and software to help determine best possible solutions Utilize Lean manufacturing techniques to streamline operations and provide superior customer service Manage expectations from all levels of management and work closely with shop employees. Display Project Management skills and provide task management documentation Manage the Bill of Material process Help teach employees in cost effective technologies Ability and desire to learn and be a contributor Help establish material flow and cost effective manufacturing processes Aid in financial analysis of new products QUALIFICATION/SKILLS:		
2		Rosco Visi	Industrial I	https://ww	Jamaica, NY	Be an e	https://\	Bachelor Degree in Industrial or Mechanical Engineering	YANLIŞ	https:/
3	1	Valex	Industrial I	https://ww	Ventura, CA	Actively	https://\	Industrial EngineerPay: $85,000 - $100,000 Annually - Depending on Experience. $2,000 Sign-on Bonus, payable after successful completion of probation period.	Ventura,	https:/
4	4	Techniplas	Industrial I	https://ww	Mansfield, OH	Be an e	https://\	Techniplas is a global supplier of sophisticated plastic components and design-focused, multifunctional surfaces that are primarily for the automotive sector, a	YANLIŞ	https:/
5	5	Manpowe	Industrial I	https://sg.	Singapore River, Si	Actively	https://:	Job Responsibilities:Equipment capacity improvement and verification and operation productivity improvementDetermine planning factors for short-term plar	YANLIŞ	https:/
6	6	KORE1	Industrial I	https://ww	Auburn, AL	Actively	https://\	KORE1, a nationwide provider of staffing and recruiting solutions, has immediate openings for Industrial Engineers.Position DescriptionLeads and completes m	YANLIŞ	https:/
7	8	Sterling En	Industrial I	https://ww	Sturtevant, WI	Be an e	https://\	Title: Industrial EngineerLocation: MT. Pleasant, WI areaHire Type: Contract to HireJob Summary:Develop and implement manufacturing processes that impro	YANLIŞ	https:/
8	11	Saddle Cre	Industrial I	https://ww	Walton, KY	Actively	https://\	Why Work for Saddle Creek?Saddle Creek Logistics Services succeeds by promoting a diverse, friendly, and respectful teamwork environment. As a vital service	YANLIŞ	https:/
9	14	CONTECH!	Industrial I	https://uk.	Dunton, England, L	Actively	https://\	Direct message the job poster from CONTECHSKirby CousinsOur high profile client in the automotive industry is currently looking to recruit an Industrial Engine	YANLIŞ	https:/
10	15	Orion Engi	Industrial I	https://nl.	Utrecht, Netherlan	Actively	https://\	Direct message the job poster from Orion EngineeringSteven van de LaarRecruitment Manager at Orion Engineering - Follow me for regular national and interi	YANLIŞ	https:/
11	16	TAG The A	Industrial I	https://ww	Wisconsin, United	Be an e	https://\	Direct message the job poster from TAG The Associates GroupRichard G.Head of Engineering and head of business Unit at Asian CorporationTop Company in t	YANLIŞ	https:/
12	18	Lockheed	Industrial I	https://ww	Courtland, AL	Actively	https://\	The coolest jobs on this planet... or any other... are with Lockheed Martin Space. Lockheed Martin is a pioneer, partner, innovator and builder. Our amazing people are or		https:/
13	19	Lockheed	Industrial I	https://ww	Courtland, AL	Actively	https://\	The coolest jobs on this planet... or any other... are with Lockheed Martin Space. Lockheed Martin is a pioneer, partner, innovator and builder. Our amazing people are or		https:/

Şekil 5.2. Elde edilen ikinci veri seti

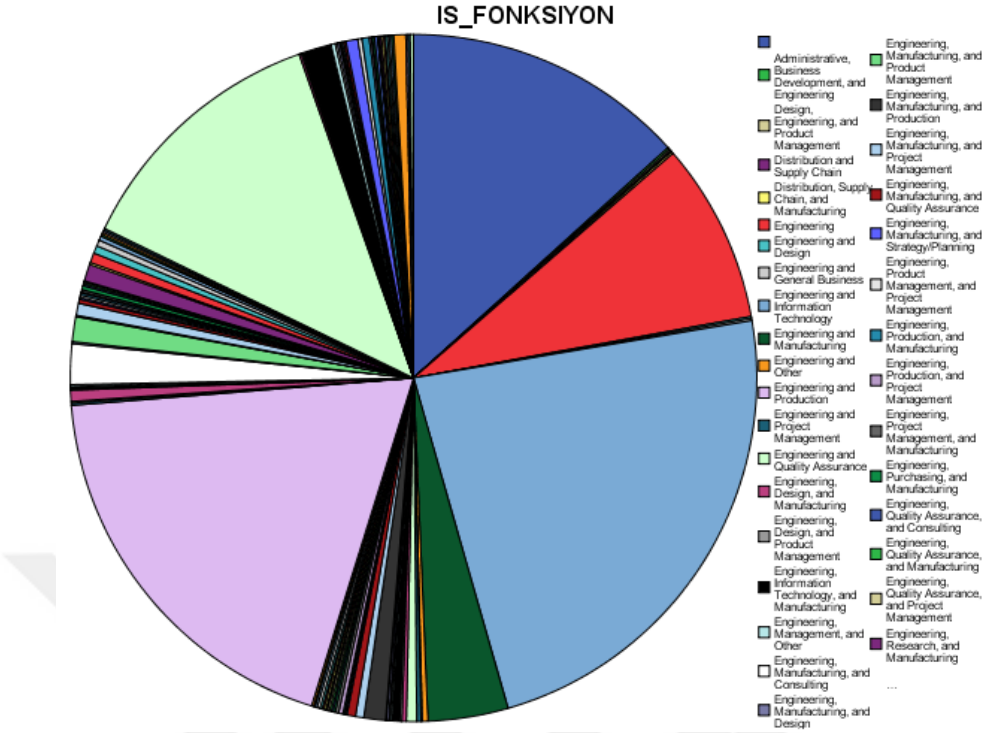
Şekilden anlaşıldığı gibi bu alan birçok metin içeren bir alandır. 24593 veri iş açıklamaları bu şekilde elde edilmiştir. Bu bilgiler 3. Aşamada kullanılacaktır.

5.4. İstatistiksel Analiz Uygulamaları

Bu bölümde veriler üzerinde istatistiksel analizler yapılmış ve IBM SPSS 24.0 programı kullanılmıştır. Data setimiz büyük veri kapsamındadır ve bu nedenle genel olarak veriler ile ilgili fikir elde edilmek istemiştir.

5.4.1. İş Fonksiyonlarının Saptanması

Endüstri mühendisliği birçok alanda çalışabilmektedir. Bu anlamda hangi sektörlerde etkin olarak çalıştığı saptanmak istenmiştir. İlgili alanlar ile ilgili frekans analizi yapılmıştır.



Şekil 5.3. İş fonksiyonlarının dağılımı

İş fonksiyonlarının verileri incelediğinde Dünyada en çok İşletmede en çok teknoloji bilgi alanında alım yapıldığı, ikinci olarak mühendislik proje yönetimi 3. Olarak ise Üretim ve strateji planlama olduğu görülmektedir (Şekil 5.3).

5.4.2. İş Fonksiyonlarının Ülkelere Göre Yüzdelik Dağılımları

İş fonksiyonlarının verileri incelediğinde en çok alım yapılan alan tespit edilmiş ancak hangi ülke için hangi fonksiyonlar önemli olduğu hakkında net bir bilgiye ulaşılamamıştır. Bu yüzden ülke bazında iş fonksiyonları düzenlenmiş olup Tablo 5.2'de gösterilmiştir.

Tablo 5.2. İş fonksiyonlarının ülkelere göre dağılımı

Ülke	Yüzdelik dağılım
AVUSTURALYA	
Mühendislik	21,686
Mühendislik ve Bilgi Teknolojisi	24,096
Yönetim ve Üretim	54,216
AVUSTURYA	
Genel İş, Tedarik Zinciri ve İmalat	31,25
Üretim ve Mühendislik	68,75

Tablo 5.2. (Devam) İş fonksiyonlarının ülkelere göre dağılımı

BELÇİKA	
Mühendislik	23,333
Yönetim, Üretim ve Mühendislik	19,583
Kalite Güvence ve Analist	16,25
Kalite Güvencesi ve İş Geliştirme	7,5
Kalite Güvencesi, Mühendislik ve Üretim	22,916
Bilim, Kalite Güvencesi ve Araştırma	10,416
BOTSVANA	
Yönetim ve Üretim	54,716
Kalite güvencesi	33,962
BULGARİSTAN	
Mühendislik ve Bilgi Teknolojisi	76,923
Kalite Güvencesi, Ürün Yönetimi ve Mühendislik	23,076
ÇEKYA	
Yönetim ve Üretim	29,268
Üretim ve Kalite Güvencesi	70,731
ÇİN	
Mühendislik ve Bilgi Teknolojisi	25,123
Yönetim ve Üretim	12,807
İmalat,Üretim,Mühendislik	8,374
Kalite güvencesi	22,167
Kalite Güvencesi, Ürün Yönetimi ve Üretim	8,866
Bilim, Bilgi Teknolojisi ve Mühendislik	12,315
ENDONEZYA	
Mühendislik ve Bilgi Teknolojisi	14,529
Yönetim ve Üretim	22,222
Üretme	5,982
FİLİPİNLER	
Mühendislik ve Bilgi Teknolojisi	42,857
Yönetim ve Üretim	57,142
FİNLANDİYA	
Yönetim ve Üretim	100
FRANSA	
Mühendislik	47,916
Üretim, Mühendislik ve Yönetim	52,083

Tablo 5.2. (Devam) İş fonksiyonlarının ülkelere göre dağılımı

GÜNEY AFRİKA	
Mühendislik ve Bilgi Teknolojisi	6,666
Mühendislik, Yönetim ve Diğer	9,333
Yönetim ve Üretim	60,444
Kalite güvencesi	11,111
GÜRCİSTAN	
Mühendislik ve Bilgi Teknolojisi	100
HİNDİSTAN	
Mühendislik ve Bilgi Teknolojisi	16,111
Yönetim ve Üretim	58,189
Üretme	5,193
Proje Yönetimi, İmalat ve Mühendislik	2,929
Kalite güvencesi	17,576
HOLLANDA	
Mühendislik ve Bilgi Teknolojisi	21,828
Mühendislik, İmalat ve Proje Yönetimi	7,079
Mühendislik, Proje Yönetimi ve İmalat	7,374
Mühendislik, Araştırma ve Üretim	5,014
Yönetim ve Üretim	6,489
Kalite Güvencesi ve Mühendislik	18,879
Kalite Güvencesi, Mühendislik ve Üretim	7,669
Kalite Güvencesi, Üretim ve Mühendislik	20,353
Kalite Güvencesi, Araştırma ve Ürün Yönetimi	5,309
İNGİLTERE	
Mühendislik	11,478
Mühendislik ve Bilgi Teknolojisi	26,751
Mühendislik ve İmalat	5,729
Mühendislik, İmalat ve Üretim	2,705
Yönetim ve Üretim	25,868
İmalat ve Mühendislik	6,800
Kalite güvencesi	13,638
Kalite Güvencesi, Mühendislik ve Üretim	5,260
Kalite Güvencesi, Proje Yönetimi ve Tasarım	1,765
İRLANDA	
Mühendislik	12,5
Mühendislik ve Bilgi Teknolojisi	17,452
Mühendislik ve İmalat	8,726
Mühendislik ve Kalite Güvencesi	15,566
Mühendislik, Üretim ve Ürün Yönetimi	8,254
Üretim ve Kalite Güvencesi	3,066

Tablo 5.2. (Devam) İş fonksiyonlarının ülkelere göre dağılımı

Üretim ve Tedarik Zinciri	5,896
Üretim, Kalite Güvencesi ve Üretim	8,726
Üretim ve İmalat	8,018
Kalite Güvencesi ve Üretim	5,896
Kalite Güvencesi, Mühendislik ve Tasarım	5,896
İSPANYA	
Mühendislik	5,863
Mühendislik, İmalat ve Üretim	16,286
Mühendislik, Üretim ve Kalite Güvencesi	8,143
Yönetim ve Üretim	40,065
Proje Yönetimi ve Bilgi Teknolojileri	8,143
Kalite Güvencesi, Müşteri Hizmetleri ve Üretim	6,188
Kalite Güvencesi, Mühendislik ve Satın Alma	15,309
İSRAİL	
Kalite güvencesi	100
İSVİÇRE	
Mühendislik, Kalite Güvence ve Danışmanlık	19,266
Yasal	15,596
Kalite Güvencesi ve Proje Yönetimi	33,944
Kalite Güvencesi, Mühendislik ve Ürün Yönetimi	31,192
İTALYA	
Mühendislik	19,109
Mühendislik ve Bilgi Teknolojisi	4,081
Mühendislik, İmalat ve Üretim	4,823
Mühendislik, İmalat ve Proje Yönetimi	9,090
Yönetim ve Üretim	17,254
Üretim, Mühendislik ve Strateji/Planlama	9,276
Proje Yönetimi, Danışmanlık ve Mühendislik	4,638
Proje Yönetimi, Kalite Güvencesi ve Mühendislik	6,864
Kalite güvencesi	20,593
Araştırma	4,267
JAPONYA	100
Diğer	
KANADA	
Mühendislik ve Bilgi Teknolojisi	59,487
Yönetim ve Üretim	30,769
Üretme	9,743
MACARİSTAN	

Tablo 5.2. (Devam) İş fonksiyonlarının ülkelere göre dağılımı

Yönetim ve Üretim	71,929
Kalite güvencesi	28,070
MALEZYA	
Mühendislik ve Bilgi Teknolojisi	51,086
Üretim ve Mühendislik	21,739
Kalite güvencesi	27,173
MEKSİKA	
Mühendislik	31,779
Mühendislik ve Bilgi Teknolojisi	37,711
Mühendislik, Kalite Güvence ve Üretim	19,915
Yönetim ve Üretim	10,593
NORVEÇ	
Mühendislik, Ürün Yönetimi ve Proje Yönetimi	100
POLONYA	
Mühendislik	60,975
Kalite güvencesi	39,024
PORTA RİKO	
Kalite güvencesi	100
PORTEKİZ	
Mühendislik ve Bilgi Teknolojisi	30,120
Mühendislik ve Kalite Güvencesi	45,783
Yönetim ve Üretim	24,096
ROMANYA	
Mühendislik	15,625
Mühendislik ve Genel İşletme	7,5
Mühendislik ve Bilgi Teknolojisi	40
Üretim, Mühendislik ve Yönetim	23,75
Kalite güvencesi	13,125
SIRBİSTAN	
Mühendislik	29,166
Mühendislik ve Bilgi Teknolojisi	25
Mühendislik, Üretim ve Proje Yönetimi	25
Üretme	20,833
SİNGAPUR	
Mühendislik ve Bilgi Teknolojisi	35,294
Mühendislik, Satın Alma ve İmalat	12,443

Tablo 5.2. (Devam) İş fonksiyonlarının ülkelere göre dağılımı

Yönetim ve Üretim	41,855
Kalite güvencesi	10,407
SLOVAKYA	
Mühendislik	48,979
Tedarik Zinciri ve Dağıtım	51,020
SUUDİ ARABİSTAN	
Mühendislik ve Bilgi Teknolojisi	52,380
Kalite güvencesi	47,619
TAYLAND	
Mühendislik ve Bilgi Teknolojisi	61,061
Kalite güvencesi	38,938
TÜRKİYE	
Mühendislik	28,504
Mühendislik ve Genel İşletme	6,074
Mühendislik ve Bilgi Teknolojisi	9,345
Mühendislik, İmalat ve Üretim	36,915
Yönetim ve Üretim	8,878
Kalite güvencesi	10,280
UGANDA	
Yönetim ve Üretim	100
VİETNAM	
Dağıtım, Tedarik Zinciri ve Üretim	58,139
Kalite Güvencesi, Mühendislik ve Üretim	41,860
YENİ ZELLANDA	
Mühendislik	100
YUNANİSTAN	
Kalite Güvencesi ve Üretim	100
AMERİKA	
İdari, İş Geliştirme ve Mühendislik	0,258
Tasarım, Mühendislik ve Ürün Yönetimi	0,227
Mühendislik ve Bilgi Teknolojisi	34,32
Mühendislik ve İmalat	18,04
İnsan kaynakları	0,258
Yönetim ve Üretim	21,77
Yönetim, Proje Yönetimi ve Mühendislik	0,247
Kalite güvencesi	22,35
Strateji/Planlama ve Bilgi Teknolojileri	1,425

Tablo 5.2. (Devam) İş fonksiyonlarının ülkelere göre dağılımı

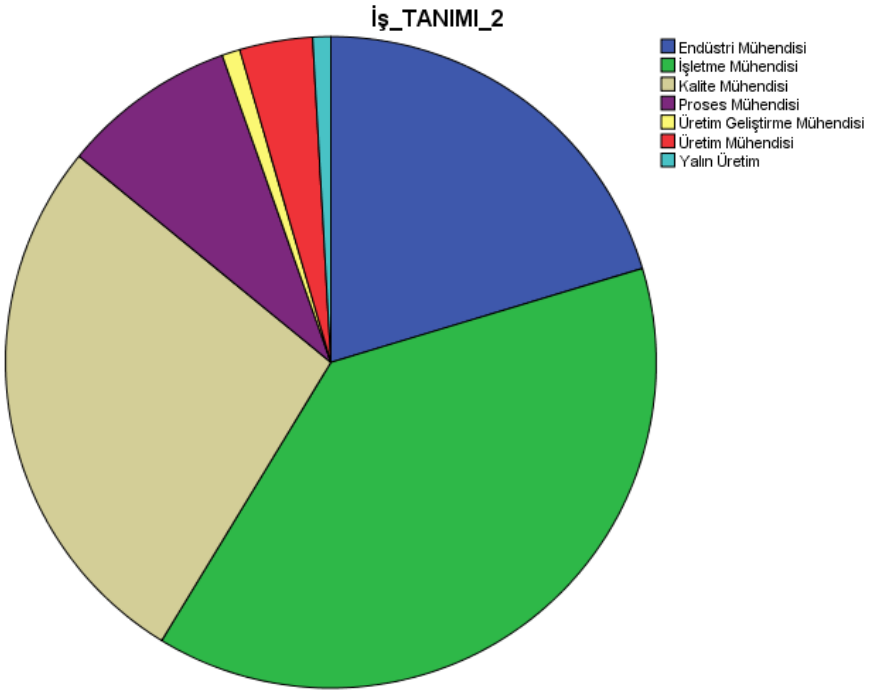
Tedarik Zinciri, Dağıtım ve Mühendislik	1,074
ALMANYA	
Mühendislik	37,7
Mühendislik ve Bilgi Teknolojisi	15,4
Mühendislik, Üretim ve Proje Yönetimi	15,4
Mühendislik, Tedarik Zinciri ve İmalat	15,4
Yönetim ve Üretim	16

5.4.3. İş Tanımlarının Dağılımları

İş arama için endüstri mühendisi olarak arama yapıldığında birçok farklı seçenek çıkmaktadır

İş Tanımlarında birçok farklı ifade bulunmaktadır. Örnek verecek olursak "Manufactoring engineer mechanical" mekanik ile ilgili sektörü açıklamış ancak yapılan iş işletme mühendisi ile aynı işe sahiptir.

Bu yüzden veriler üzerinde veri dönüşümü yapılmıştır. Veri dönüşümünden sonra veriler üzerinde kümele yapılmıştır.



Şekil 5.4. İş tanımlarının dağılımı

Şekil 5.4'de görüldüğü gibi endüstri mühendisleri en çok İşletme mühendisi, Kalite Mühendisi ve Endüstri mühendisi olarak alım yapılmaktadır. Yalın üretim mühendisi olarak ise çok küçük bir kısımda çalışmaktadır.

Genel olarak ülkemizde Yalın Üretim için özel pozisyonlar açılmakta ancak birçok ülkede ise iş ilanın incelendiğinde Endüstri mühendisi iş ilanlarının detayında Yalın üretim sertifikası istenmektedir. Üretim Geliştirme için ise genelde Makine mühendislerinin alım yapıldığı görülmüştür.

5.4.4. Çalışma Şekillerinin Dağılımları

Dünya üzerinde Kovid ile birlikte birçok sektörde çalışma şekilleri değişmiştir. Bu anlamda çalışma şekilleri ile alakalı olarak istatistiksel analiz yapılmıştır. Tablo 5.3'de işlerin çalışma şekline göre dağılımı gösterilmiştir.

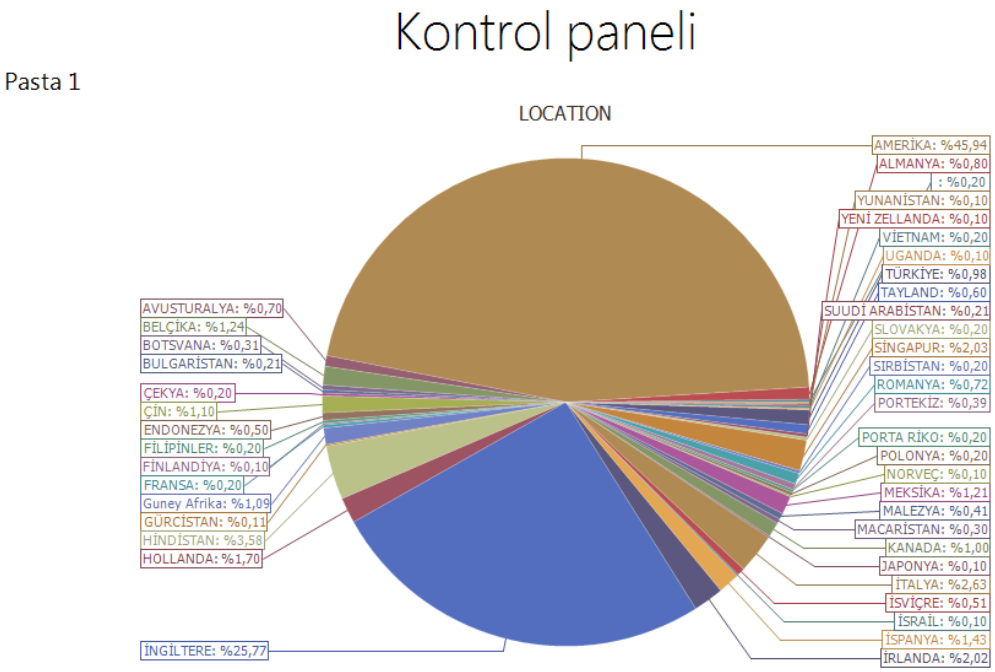
Tablo 5.3. Çalışma şekillerinin dağılımı

Çalışma şekli	Frekans	Yüzde Frekans	Kümülatif Frekans
Sözleşmeli	1522	6,2	13,3
Tam zamanlı	19203	78,1	19,5
Staj	118	0,5	97,6
Uzaktan	122	0,5	98,0
Part time	206	0,8	98,5
Geçici	153	0,6	99,4
Toplam	24594	100,0	100,0

Endüstri mühendisliği iş ilanlarının %78,1 tam zamanlıdır.

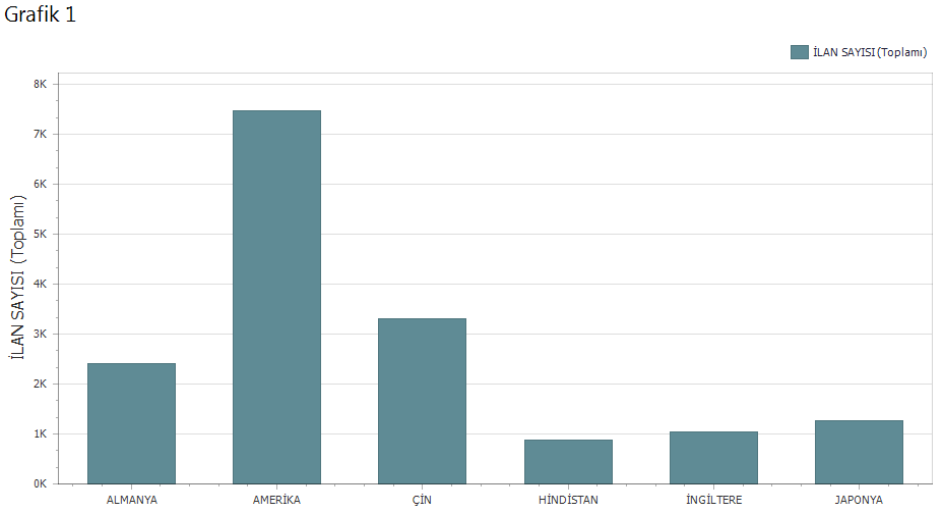
5.4.5. İş İlanlarının Ülkelere Göre Dağılımları

Toplam 24593 Endüstri Mühendisi iş ilanı analiz edilmiştir. İşler İngilizce olarak Dünya çapında araştırılmıştır. Araştırma sonuçlarına bakıldığında Almanya ve Japonya ya endüstri mühendisliği oranı oldukça düşük kalmaktadır. Bu sebeple daha detaylı analiz yapılması gerekmektedir. Almanya'daki iş ilanları Almancada verilebileceğinden dolayı araştırma genişletilmiştir.



Şekil 5.5. İş ilanlarının ülkelere göre dağılımı

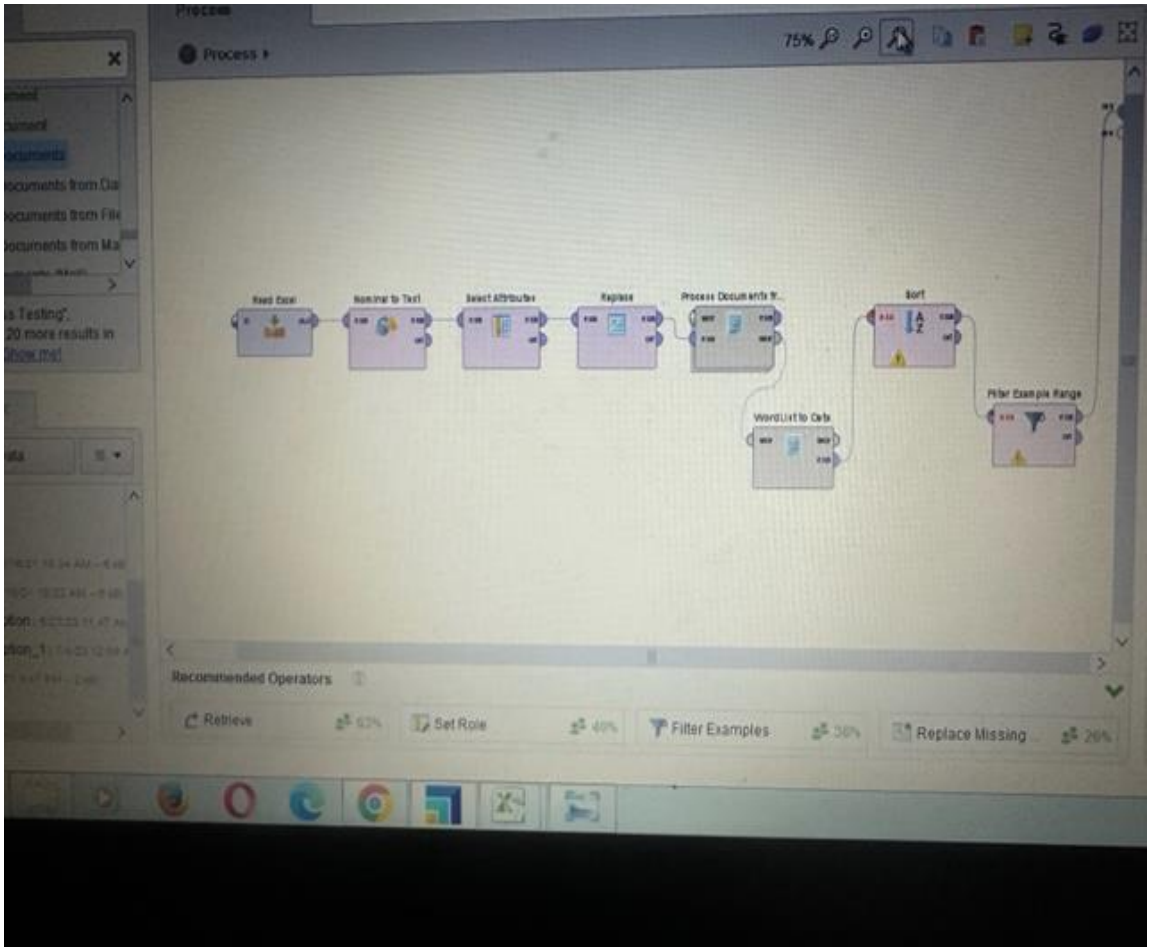
Şekil 5.5'de görüldüğü gibi İngiltere'deki ilan oranı çok fazladır. Asıl sebebi her ülke kendi dilinde ilan vermesidir. Bu yüzden her ülke için kendi dil ve konumunda araştırma yapılmış ve en yüksek alıma sahip 6 ülke belirlenmiştir. Sonuçlar asılda en çok neden bilgi ve teknoloji alanında alım yapıldığını doğrular niteliktedir. Amerika ve Çin başta olmak üzere en çok alım yapan ülkeler bu alanda endüstri mühendisi aramaktadır (Şekil 5.6).



Şekil 5.6. En çok alım altı ülke

5.5. Metin Madenciliği Çalışmaları

Veri setimizde yüzbinlerce veri bulunmaktadır. Bu verilerin anlamlı hale gelmesi için veriler Rapidminer programında işlenmiştir. Amacımız bir iş ilanında en fazla geçen kelimeleri bulmaktır. İki aşamalı bir ön elemeden geçmiştir. Kurulan model Şekil 5.8' de gösterilmiştir.

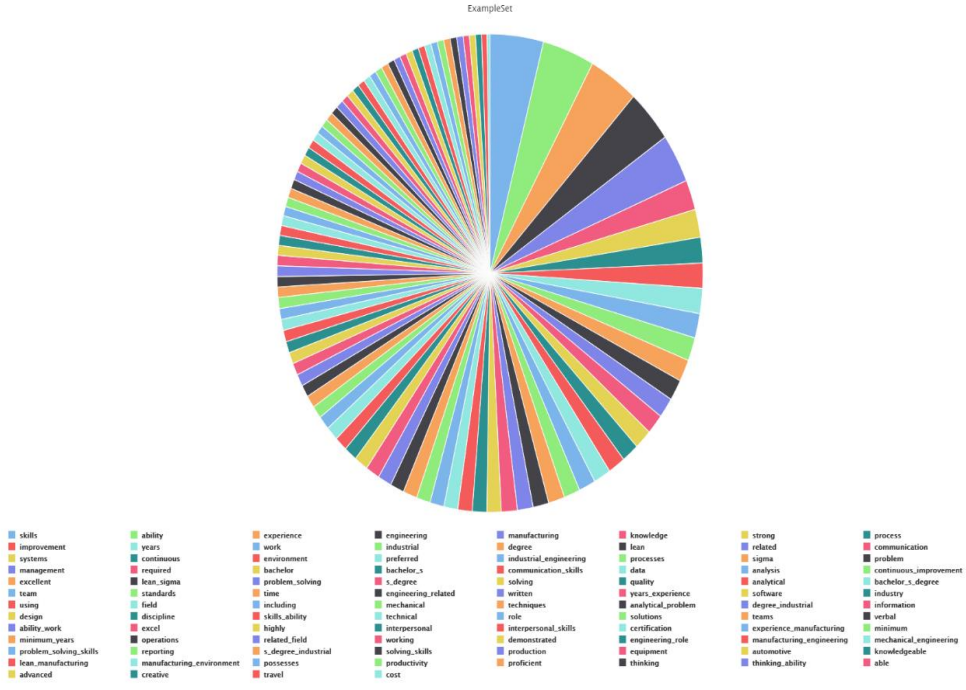


Şekil 5.7. Rapidminer programında kurulan model

Programı çalıştırdığımızda çıkan ilk sonuçlar Şekil 5.9'daki gibi olmaktadır. Sonuçlara ait pasta grafiği ise Şekil 5.10'daki gibidir.



Şekil 5.8. En sık kullanılan kelimelerin özeti



Şekil 5.9. En sık kullanılan kelimelerin özetinin pasta grafiği

5.5.1. Metin Madenciliği ile Türkiye'deki İş İlanlarının İncelenmesi

Linkedinden alınan veriler ile kariyer netten alınan veriler birleştirilmiştir. Birleştirme yapmamızdaki asıl amaç daha fazla veri ile çalışıp, Endüstri mühendisinde aranan en çok

özellikleri bulmaktır. İlk olarak en çok geçen kelimelere odaklanılmıştır.256 adet iş ilanı elde edilmiş ancak ön eleme yapılmıştır ve 122 adet ilan dikkate alınmıştır. Eleme yaparken, diğer mühendislikleri içeren iş ilanlarını temizlendi. Örneğin "İnşaat veya endüstri mühendisi " gibi ifadeler bulunmaktaydı ancak iş ilanında Endüstri mühendisi ile ilgili temel nitelikler bulunmamaktaydı. Türkiye'deki sonuçlar için kelime bulutu Şekil 5.11'de gösterilmiştir.

Dünyadaki ve Türkiye'deki Endüstri mühendisliği iş ilanlarındaki kelime bulutlarını incelediğimizde en temel fark askerlik durumu ve planlama yeteneğidir. Bunun dışında ise çok fazla benzer özellik görülmektedir. İletişim, Ms Office kullanımı, Üretim ve yönetim becerisi ve analiz ve analitik düşünme yeteneği ortak özelliktedir.



Şekil 5.10. Türkiye'deki iş ilanlarında en çok kullanılan kelimeler

5.6. Metinlerin Kategorize Edilmesi

Bu aşamada ise kelime ve kelime öbekleri üzerinde durulmuş ve nihaini kelimeler yeteneklere göre ayrıştırılmıştır. Daha önceki çalışmalarda yeteneklerin ayrılmasında iş verenler, öğretim üyeleri vs. ile görüşülerek hazırlandığı tespit edilmiştir ancak bu çalışmada çıkan kelime analizlerinden elde edilmiştir. Anahtar kelime öbekleri için ilk olarak n-Gram oluşturulmuş daha sonra oluşan kelime öbekleri kategorize edilmiştir. Tablo 5.4'de verilmiştir.

Tablo 5.4. İş fonksiyonlarının katagorize edilmesi

Kelime	Kelime kavramları
Analiz ve analitik yeteneği	Analitik düşünme yeteneği, Matematik kabiliyetleri ,Problem çözme yeteneği, Karar alma yeteneği
İletişim becerisi	Yazılı ve sözlü iletişim beceriler
Eğitim derecesi	Endüstri Mühendisliği, Bilgisayar bilimi, İşletme ve makine mühendisliği
Kalite Yönetimi	Kaizen,Altı Sigma,Kalite Yönetimi
Takım	Takım oluşturma, ekip çalışması ,ekipler arası yönetimi
Bt ve kurumsal kaynak yönetimi	ERP,Excel,Ms Office,SAP
İş planlama,Analiz,Geliştirme	Ürün geliştirme, Yeni ürün, Süreç kabiliyeti ve geliştirilmesi
Dizayn ve Çizim	Autocad,Solid,Tasarım,teknik resim
Bakım Yönetim	Saha servis raporlama,maliyet,zaman yönetimi, Liderlik
Malzeme Bilimi, yönetimi ve ekipman	Geometrik Boyutlandırma ve Toleranslandırma,Isıl işlem,Ölçümler,Hammaddele
Mühendislik yöntemleri ve uygulamaları	Değer akışı haritalama,5s,Ergonomi,Kök neden analizi
Operasyon yönetimi	Tedarik zinciri, Üretim planlama, Değişiklik yönetimi, Maliyet azaltma.denetim,Belgelendirme
Proje Yönetimi	Proje planlama, Proje yönetimi, Proje yönetimi
Kişisel yetenekler	Zaman yönetimi, sunum becerileri, gözlem yeteneği, Zaman yönetimi ve organizasyon yönetimi
İş sağlığı ve güvenliği	İş güvenliği, sağlığı ve yönetimi, Üretim güvenliği
Programlama	C++,Python,Programala becerisi
Veri Analizi ve Veri Yönetimi	SQL,Power BI,İstatistik,Minitab,Veri yönetimi
Yalın Üretim, Sürekli İyileştirme ve Optimizasyon	Süreç iyileştirme, Süreç optimizasyonu, Görsel yönetim sistemleri, Yalın üretim

Genel olarak en çok kullanılan ifadeler Tablo 5.5'de verilmiştir. Amacımız her bir katogori için önem değerini saptamaktır. İş ilanlarının neredeyse tamamında geçen ifade ise iletişim yeteneğidir.

Tablo 5.5. İş ilanlarında aranan özelliklerin niteliği

Analiz ve analitik yeteneği		Bt ve kurumsal kaynak yönetimi	
Karar Verme	6,140		
Problem ve Sorun Çözme	57,017	ERP	29,292
Analitik beceriler	34,210	Microsoft Excel	19,191
		Microsoft Powerpoint	12,121
		Visio	3,030
		Ms Office	36,363
Bakım Yönetimi		Derece	

Tablo 5.5. (Devam) İş ilanlarında aranan özelliklerin niteliği

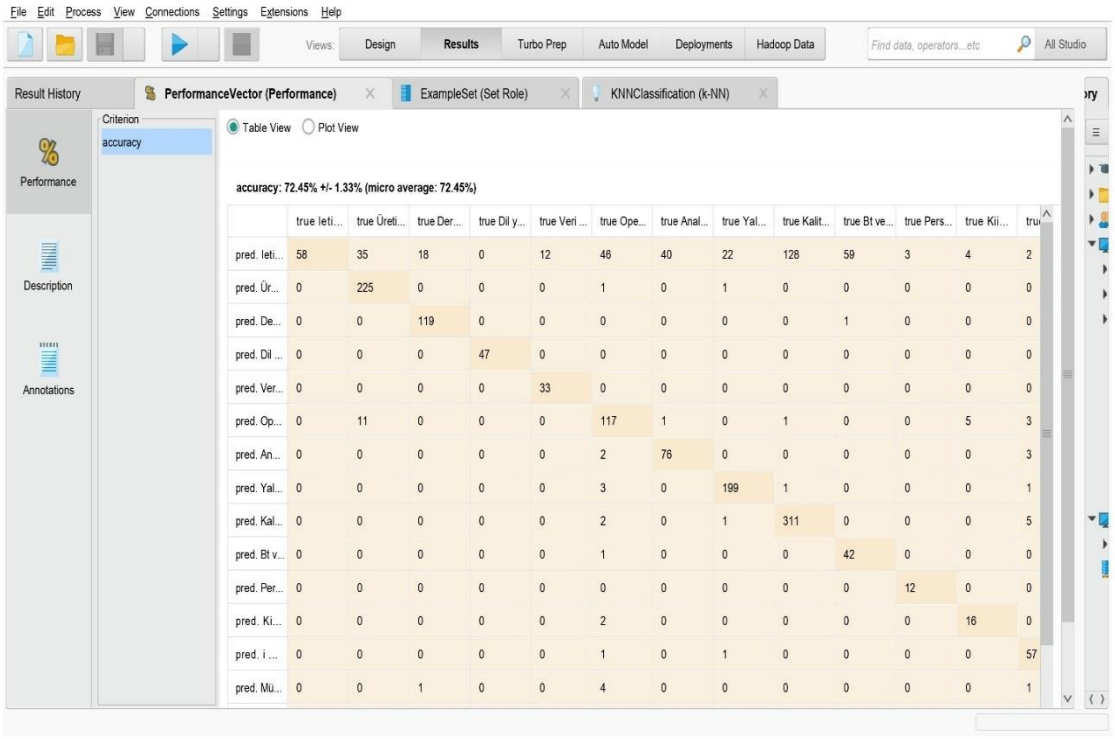
Saha Servis Mühendisliği	33,333	Endüstri Mühendisliği	42,222
Önleyici Bakım	33,333	Makine Mühendisliği	23,703
Ekipman Bakımı	33,333	Proje ve Tasarım Mühendisliği	4,444
		Süreç mühendisliği	5,185
		Üretim ve İmalat Mühendisliği	24,444
Dil yeteneği		Dizayn ve çizim	
İngilizce	90,909	CAD (Bilgisayar Destekli Tasarım)	52,272
İspanyolca	9,090	SolidWorks	34,090
		Teknik Çizim	20,454
		Tasarım için kullanılan araçlar	22,727
		AutoCAD	38,636
		İmalat İçin Tasarım	45,454
İletişim yeteneği		İş planlama,Analiz,Geliştirme	
İletişim	89,655	İş gereksinimlerinin tanımlanması	6,779
Sözlü İletişim	11,538	Yeni ürün sunumu	11,864
		Süreç Kabliyeti ve Geliştirmesi	45,762
		Ürün Hizmeti	20,338
		Ürün Geliştirme	15,254
İş sağlığı ve güvenliği		Kalite Yönetimi	
İş Sağlık ve Güvenlik Yönetimi (HSE)	64,285	Altı Sigma	21,391
Üretim Güvenliği	35,714	FMEA (Hata Türleri ve EtkileriAnalizi)	6,958
		ISO 16949,13485,14001,9001	10,309
		Kaizen	7,731
		Kalite Güvencesi	7,216
		Kalite Kontrol	4,123
		Kalite Yönetimi	10,051
		PPAP	4,639
		PFMEA Analizi	6,958
		Kalite Mühendisliği	4,896
		Ürün Kalitesi	11,082
		APQP (İleri Ürün Kalite Planlaması)	4,639
Kişisel yetenekler		Malzeme Bilimi, yönetimi ve ekipman	
Gözetim Becerileri	11,111	Geometrik Boyutlandırma ve Toleranslandırma	20
Güvenilirlik	16,666	Isıl İşlem	7,5
		Malzeme Şartnamesi	20
Organizasyon Yetenekleri	16,666	Mastarlar	12,5
Sunum Becerileri	38,888	Malzeme İnceleme Panoları	17,5
Zaman Yönetimi	16,666	OEE	7,5
		Ölçümler	7,5
		Ham Maddeler	7,5

Tablo 5.5. (devam) İş ilanlarında aranan özelliklerin niteliği

Liderlik		Mühendislik yöntemleri ve uygulamalar	
Görevler Arası Ekip Liderliği	52,941	Değer Akışı Eşlemesi	13,218
Liderlik	47,058	Ergonomi	4,597
		Tesis Yerleşimi	2,298
		GMP	4,597
		Kanban	5,747
		Kök Neden Analizi	14,942
		MES	2,298
		Mühendislik Yönetimi	20,689
		Tesis Yerleşimi	5,747
		Zaman Etüdü	8,045
		5S	14,942
		Montaj süreçleri	2,873
Operasyon yönetimi		Yalın Üretim,Sürekli İyileştirme ve Optimizasyon	
Belgelendirme	3,921	Süreç İyileştirme ve Süreç kontrol	12,796
Denetim	6,53	Süreç Optimizasyonu	4,265
Düzeltici Eylemler	13,725	Sürekli İyileştirme	22,748
Maliyet Azaltma	10,457	Görsel Yönetim Sistemleri	3,791
Müşteri Hizmetleri	3,921	Yalın Prensip ve süreçler	56,398
Risk Yönetimi	3,921		
Satın Alma Emirleri	2,614		
Sermaye Projeleri	3,921		
Tedarik Zinciri Optimizasyonu ve Yönetimi	19,607		
Üretim Planlama	8,496		
Değişiklik Yönetimi	10,457		
Depo Yönetimi	3,92		
Düzeltici Eylemler	8,496		
Proje Yönetimi		Takım	
Proje Planlama	23,809	Takım Oluşturma	36,363
Proje yönetimi	76,190	Ekip Çalışması	54,545
		Ekip Yönetimi	9,090
Üretim yöntemleri ve yönetimi		Veri Analizi ve Veri yönetimi	
İmalat Prosesleri	69,49807	Minitab	17,073
İş Talimatları	6,177606	SPC	31,707
Üretim Planları	3,861004	Veri Analizi	26,829
Bilgisayarlı Sayısal Denetim (CNC)	6,949807	İstatistik	14,634
Üretim Hattı ve Üretim akışı Programlama	13,51351	Veri yönetimi	9,756
C++	50		
Programlama	50		

5.7. İş İlanlarının Sınıflandırılması

İş açıklamaları kısımları uzun cümleler içermektedir. Hangi niteliğin bizim için önemli olduğunu belirlemek için işler sınıflandırılmalı ve daha sonra detaylandırılmalıdır. İş açıklamalarında geçen anahtar kelimeler gruplandırılmıştır. Sınıflandırma yaparken metin madenciliği tekniklerinden yararlanılmıştır en yakın komşu algoritmasından yararlanılmıştır. Verilerin %90 eğitim %10 u ise test verisi olarak kullanılmıştır. Çıkan sonuçlara göre %72,45 doğruluk oranı ile doğru bir sınıflandırma yaptığımız görünmektedir (Şekil 5.11).



Şekil 5.11. İş ilanların Sınıflandırılmasından elde edilen veriler

6. SONUÇ VE ÖNERİLER

Endüstri mühendisinin çalışmadığı sektör hemen hemen yok gibidir. Diğer mühendislik bölümlerine göre daha fazla seçenek içerdiğinden iş ilanlarında birçok pozisyon çıkmaktadır. Pozisyonlar lokasyonlara göre farklılık göstermektedir. Pozisyon, departman maaş gibi birçok etkileyen faktör arasından seçim yapabilmek neredeyse imkansızdır. Bu konuya ülke sınırları dışını da kattığımızda büyük bir veri karşımızdadır. Günümüzün ve geleceğinin en büyük sorunlarından olan büyük veriyi yönetmek, çıkarımlardan bulunmak için veri analizi araçlarından yararlanılmaktadır. Veriyi yönetmek, görselleştirmek ve saf bilgi haline getirebilmek daha anlaşılır olduğu gibi veriler hakkında çıkarımlarda bulunmayı gelecek hakkında yorumlarda bulunmayı sağlamaktadır. Veri madenciliği; eğitimde, savunma sanayi, bankacılık, tarım, imalat, ülke stratejilerine kadar birçok alanda göre daha fazla seçenek içerdiğinden iş ilanlarında birçok pozisyon çıkmaktadır. Pozisyonlar lokasyonlara göre farklılık göstermektedir. Pozisyon, departman maaş gibi birçok etkileyen faktör arasından seçim yapabilmek neredeyse imkansızdır. Bu konuya ülke sınırları dışını da kattığımızda büyük bir veri karşımızdadır. Günümüzün ve geleceğinin en büyük sorunlarından olan büyük veriyi yönetmek, çıkarımlardan bulunmak için veri analizi araçlarından yararlanılmaktadır. Veriyi yönetmek, görselleştirmek ve saf bilgi haline getirebilmek daha anlaşılır olduğu gibi veriler hakkında çıkarımlarda bulunmayı gelecek hakkında yorumlarda bulunmayı sağlamaktadır. Veri madenciliği; eğitimde, savunma sanayi, bankacılık, tarım, imalat, ülke stratejilerine kadar birçok alanda karşımıza çıkmaktadır. Çalışma temel olarak web madenciliği, veri madenciliği ve metin madenciliği çalışmalarını içerir.

Linkedin üzerinden alınan veriler kullanılmıştır. Verileri elde etmek için ise Python 3.10 programında 2 aşamalı bir kod yazılmıştır. Koddan Linkedindeki tüm veriler çekilmiştir. İş açıklamaları, iş tanımları, sektör maaş düzeyi. Ülkeler ve eyaletler, başvuru sayısı gibi.

Veri setindeki verileri değerlendirmek için ise IBM SPSS 24.00 programı kullanılmıştır. Bu veriler ile çalışma şekillerinin dağılımı, sektör dağılımları bulunmuştur. Daha sonra ülkelerdeki en fazla Endüstri mühendisi alımları ve bu ülkelerin aradıkları departmanlar tespit edilmiştir.

Çalışmaya ikinci bir boyut kazandırmak ve kullanılan kelimeleri daha net inceleyebilmek için ikinci bir metin madenciliği çalışması Rapidminer programında yapılmıştır. İlk olarak en çok geçen kelime ve kelime öbekleri bulunmuş daha sonra ise yetenek analizi yapılmıştır. Daha sonra kelime öbeklerinden oluşan cümleler elde edilmiştir. Bu işlemlerde n-Gram kullanılmıştır. Elde edilen kelimeler ile belirli kategoriler elde edilmiştir. İş açıklamalarının tamamı içinde geçen kelimelere göre gruplanmıştır. Bu gruplamalarda sorgulama dillerinden yararlanılmıştır. Amacımız tek tek her iş ilanında yetenek aramak değil tüm veriyi tahminleme yöntemiyle sınıflamaktır. Elde edilen sınıfların güvenirliği için k-NN algoritması kullanılmıştır. Sonuçlara göre yapılan sınıflandırmanın doğruluğu 72,45 olarak bulunmuştur. Katagorilerde geçen sözcükler incelenmiş ve her kategori için yüzdelik olarak önem dereceleri belirlenmiştir. Bu anlamda Endüstri mühendisliği için en fazla öneme sahip yetenekler saptanmıştır.

Tablo 6.1. Bilgisayarların özellikleri

Nitelikler	**Nitelik Türleri**
İşlemci	Intel® Core ™ i7
Ram	32 GB
Sistem Türü	64 bit İşletim Sistemi
Python Sürümü	Python 3.10 ,IBM SPSS 24.00

Uygulamamalarda kullanılan bilgisayar özelliği Tablo 6.1 gösterilmiştir.

Yapılan bu çalışma Türkiye yapılmış Endüstri Mühendisliğini Dünya çapında inceleyen ilk çalışmadır. Tezin asıl amacı, Endüstri mühendisliği çalışma alanları ve geleceği hakkında bilgi sahibi olmaktır.

Çalışma Endüstri mühendisliğine genel bir bakış sunarken aynı zamanda ileriki çalışmalara zemin hazırlamayı amaçlamaktadır.

Bu çalışmadan yola çıkarak, 3 farklı alanda katkı sağlanabilmektedir. Bunlar,

1) Üniversitelerde verilen ders eğitimlerinin açıklamaları ve özetleri çıkarıla bilinir ve bu açıklama ve yeteneklerin iş dünyasında aranan nitelikler ile aralarındaki ilişkiler incelene bilinir ve dolayısıyla bu anlamda akademik camiaya katkı sağlanması amaçlana bilinir.

2) Çalışma sonuçlarında Endüstri mühendisinde olması gereken yetenekler verilmiş ve Üniversiteye hazırlanan öğrencilerine fikir sağlana bilinir. Ve Endüstri mühendisliğine ait kariyer planlaması derinleştirilebilir.

3) İş ilanı yayınlayan ülkelere ve Üniversite mezun sayısına bakılarak kontenjanlar hakkında fikir sahibi oluna bilinir.

KAYNAKLAR

Abdala, S., Angın, Z. (2022). Angela Carter'ın Reflections Adlı Kısa Hikâyesinin Metin Madenciliği ile Analizi ve Simgesel Bağlamda Değerlendirilmesi. *Sosyal, Beşerî ve İdari Bilimler Dergisi, 5*(6), 696-707.

Akkaya, G., Turanoğlu, B., Öztaş, S. (2015). An İntegrated Fuzzy AHP and Fuzzy MOORA Approach to The Problem of İndustrial Engineering Sector Choosing. Expert Systems with Applications. *Expert Systems with Applications, 42*(24), 9565-9573.

Aljohani, N., Aslam, A., Khadidos, A., Hasan, S.-U. (2022). Bridging The Skill Gap Between The Acquired University Curriculum and The Requirements of The Job Market: A Data-Driven Analysis of Scientific Literature. *Journal of Innovation & Knowledge,* 100190.

Arendt, L., Gałecka-Burdziak, E., Núñez, F., Pater, R., Usabiaga, C. (2023). Skills Requirements Across Task-Content Groups İn Poland: What Online Job Offers Tell Us. *Technological Forecasting and Social Change, 187,* 122245.

Arslan, H. (2008). Sakarya Üniversitesi Web Sitesi Erişim Kayıtlarının Web Madenciliği ile Analizi, Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 216041.

Asano, H., Tanaka, K., Fujita, Y., Tsuda, K. (2015). Study on Hiring Decision: Analyzing Rejected Applicants By Mining. *Procedia Computer Science,* 1156-1163.

Ateş, E. (2021). Web Madenciliği. *Jandarma ve Sahil Güvenlik Akademisi*, 2-7.

Aydın, S., Özkul, A. (2015). Veri Madenciliği ve Anadolu Üniversitesi Açıköğretim Sisteminde Bir Uygulama. *Eğitim ve Öğretim Araştırmaları Dergisi*, 2146-9199.

Baykar, A. (2006). Veri Madenciliği Uygulama Alanları. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 95-107.

Beyhan. H. (2014). Sosyal Medya Üzerinden Metin Madenciliği ve Duygu Analizi İle Pazar Değerlendirme, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 353755.

Boselli, R., Cesarini, M., Mercorio, F., Mezzanzanica, M. (2018). Classifying Online Job Advertisements Through Machine Learning. *Future Generation Computer Systems*, 319-328.

Campagni, R., Merlini, D., Sprugnoli, R., Verri, M. (2015). Data Mining Models For Student Careers. *Expert Systems with Applications*, 5508-5521.

Canay, Ö., Kocabıçak, Ü. (2023). An İnnovative Data Collection Method to Eliminate The Preprocessing Phase in Web Usage Mining. *Engineering Science and*

Technology, An International Journal, 101360.

Chiarello, F., Trivelli, F., Bonaccorsi, A., Fanto, G. (2018). Extracting and Mapping İndustry 4.0 Technologies Using Wikipedia. C*omputers in Industry, 100*, 244-257.

Coşlu, E. (2013). Veri Madenciliği. Akademik Bilişim 2013 – XV. *Akademik Bilişim Konferansı Bildirileri*, (2). Antalya.

Çetin, M. (2009). Bir Üretim İşletmesinde Veri Madenciliği Uygulaması. Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 245259.

Dumouchel, W. (1999). Bayesian Data Mining İn Large Frequency Tables, With An Application To The FDA Spontaneous. *American Statistician. Statistical Practice,* 190-196.

Durakşahin, F. (2017). Büyük Veri Yığını Analizi: Yalın Üretim Literatürü Üzerine Bir Uygulama. Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 484481.

Elsafty, A., Riedl, M., Biemann, C. (2018). Document-based recommender system for job postings using dense representations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* (Volume 3) (pp. 216-224).

Fareri, S., Fantoni, G., Chiarello, F., Coli, E., Binda, A. (2020). Estimating Industry 4.0 İmpact on Job Profiles and Skills Using Text Mining. *Computers in Industry*, *118*, 103222.

Faryna, O., Pham, T., Talavera, O., Tsapin, A. (2022). Wage and Unemployment: Evidence From Online Job Vacancy Data. *Journal of Comparative Economics,* 52-70.

Frey, C., Osborne, C. (2017). The Future of Employment: How Susceptible Are Jobs To Computerisation?. *Technological forecasting and social change*, *114*, 254-280.

Gardiner, A., Aasheim, C., Rutner, P. (2018). Williams,S., Skill Requirements İn Big Data: A Content Analysis of Job Advertisements. *Journal of Computer Information Systems*, *58*(4), 374-384.

Gardiner,A., Aasheim,C.,Susan,R., Williams.,T. (2017). Skill Requirements in Big Data: A Content Analysis of Job Advertisements. *Journal of Computer Information Systems*, *58*(4), 374-384.

Haşimi, H., Hafız, A., Mathkour, H. (2015). Selection Criteria For Text Mining Approaches. *Computers İn Human Behavior,* 729-733.

Jackson, D. (2009). An İnternational Profile of İndustry-Relevant Competencies. *International Journal of Management Education*, 3-8.

Kalıkov, A. (2006). Veri Madenciliği ve Bir E-Ticaret Uygulaması. Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara, 180400.

Karakatsanis, I., El-Kadir, K., Maccrory, F., Alibasiç, A., Atıf Ömer, M., Aung, Z., Woon, W. (2017). Data Mining Approach to Monitoring The Requirements of The Job Market: A Case Study. *Information Systems*, 1-6.

Kaşıkçı, T., Gökçen, H. (2014). Metin Madenciliği ile E-Ticaret Sitelerinin Belirlenmesi. *Bilişim Teknolojileri Dergisi*, 25-33.

Kino, Y., Kuroki, H., Machida, T., Furuya, N., Takano, K. (2017). Text Analysis for Job Matching Quality Improvement. *Procedia Computer Science*, 1523-1530.

Kuşak, L. (2010). Web Ortamında Mekansal Veri Kazanımı. Doktora Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 350628.

Li, G., Yuan,C., Kamarthi, S., Mohsen, M., Jin, X. (2021). Data Science Skills and Domain Knowledge Requirements İn The Manufacturing İndustry: A Gap Analysis. *Journal of Manufacturing Systems*, *60*, 692-706.

Li,X., Liu,Dç., Sun,J., Zhu, Z. (2022). Challenges of Industrial Engineering in Big Data Environment and Its new Directions on Extension Intelligence. *Procedia Computer Science, 214*, 1561-1567.

Matsumoto, Yuji. (2000). Easy to Use Practical Freeware for Natural Language Processing:Morphological Analysis System ChaSen. *Journal of Information Processing*, 1208-1214.

Mauro, A., Greco, M., Grimaldi, M., Ritala, P. (2018). Human Resources For Big Data Professions: A Systematic Classification of Job Roles and Required Skill Sets. *Information Processing & Management*, 807-817.

Miyazaki S. (1981). One Machine Scheduling Problem With Dual Criteria. *Journal of Operations Research Society of Japan, 24*(1), 37-50.

Mohadab,M., Bouikhalene, B., Safi, S. (2017). Automatic CV Processing for Scientific Research Using Data Mining Algorithm. *Journal of King Saud University-Computer and Information Sciences, 32* (5), 523–1530.

Palmaa, M., Ríos, I., Guerrero, D. (2012). Higher Education in İndustrial Engineering İn Peru: Towards A New. *Procedia - Social and Behavioral Sciences, 46*, 1570 – 1580.

Pearce,A., Pons,D. (2019). Advancing Lean Management: The Missing Quantitative Approach, *Operations Research Perspectives, 6*, 100114.

Pejic-Bach, M., Bertoncel, T., Meško, M., Krstić, Ž. (2020). Text mining of industry 4.0 job advertisements. *International Journal of İnformation Management*, *50*, 416-431.

Restuccia D, Taska B. (2018). Different Skills, Different Gaps: Measuring and Closing The Skills Gap. *Dev Ski A Chang World Work*, 207–26.

Savaş, S., Topaloğlu, N., Yılmaz, M. (2012). Veri Madenciliği ve Türkiyede Bir Uygulama. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi,* 1-23.

Schedlbauer, J., Raptis, G., Ludwig, B. (2021). Medical İnformatics Labor Market Analysis Using Web Crawling, Web Scraping, and Text Mining, *International Journal of Medical Informatics, 150*, 104453.

Schedlbauer, J., Raptis, G., Ludwig, B. (2021). Medical İnformatics Labor Market Analysis Using Web Crawling, Web Scraping, and Text Mining. *International Journal of Medical Informatics*, 104453.

Schein, E., Maanen, J. (2016). Career Anchors and Job/Role Planning: Tools For Career and Talent Management. *Organizational Dynamics, 45*(3), 165-173.

Spada,I., Chiarello,F., Barandoni,S., Ruggi,G., Martini,A,. Fantoni,G. (2022). Are Universities Ready To Deliver Digital Skills and Competences? A Text Mining-Based Case Study of Marketing Courses in Italy. *Technological Forecasting and Social Change, 182,* 121869.

Tozak, E. (2021). Veri Madenciliği Programları Kullanılarak Bir Tekstil Firmasının Satış Verilerinin Değerlendirilmesi. Yüksek Lisans Tezi, Karabük Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 664677.

Tunalı, V. (2011). Metin Madenciliği İçin İyileştirilmiş Bir Kümeleme Yapısının Tasarımı Ve Uygulaması. Doktora Tezi, Marmara Üniversitesi. Fen Bilimleri Enstitüsü, İstanbul, 2011, 304604.

Turgut, H. (2012). Veri Madenciliği Süreci Kullanılarak Alzheimer Hastalığı Teşhisine Yönelik Bir Uygulama, Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Isparta, 330849.

Tüzüntürk, S. (2010). Veri Madenciliği ve İstatistik. *Uludağ Üniversitesi İktisadi Bilimler Fakültesi Dergisi*, 65-90.

Usabiaga, C., Núñez, F., Arendt, L., Gałecka-Burdziak, E., Pater, R. (2022). Skill Requirements and Labour Polarisation: An Association Analysis Based on Polish Online Job Offers. *Economic Modelling*, 105963.

Vergiliel Tüz, M. (2003). Kariyer Planlamasında Yeni Yaklaşımlar. *Uludağ Üniversitesi Fen-Edebiyat Fakültesi,Sosyal Bilimler Dergisi*, 4.

Ye, L., Yimeng, Y., Wei, C. (2023). Analyzing Public Perception of Educational Books Via Text Mining of Online Reviews. *Procedia Computer Science,* 617-625.

Yıldız, M. ve Şeker, Ş. E. (2016). Veri Madenciliği Araçlar. *YBS Ansiklopedi, 3*(4), 10-

19.

KİŞİSEL YAYIN VE ESERLER

Özcan B., Yıldırak E., **Aksoy, Z.** (2021). Attitudes Towards Vaccines and Intention to Vaccinate Against Covid-19: A Statistical Analysis, *Naturengs*, *2*(1), 44-59.

ÖZGEÇMİŞ

İlköğretimini ve Ortaöğrenimini İnönü İlköğretim Okulu'nda, Lise eğitimini 2011 yılında Oruçreis Anadolu Lisesi'nde tamamladı. Lisans eğitimini 2012-2016 yıllarında Dumlupınar Üniversitesi, Endüstri Mühendisliği Bölümünde tamamladı. 2017- 2022 yılları arasında Er-El Makine A.Ş.'de Endüstri Mühendisi ve ERP proje yöneticisi olarak çalışmıştır.