



BURDUR MEHMET AKİF ERSOY ÜNİVERSİTESİ

T.C

BURDUR MEHMET AKİF ERSOY ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

YÖNETİM BİLİŞİM SİSTEMLERİ ANA BİLİM DALI

**AKILLI SERALARDAN TOPLANAN VERİLERİN VERİ
MADENCİLİĞİ YÖNTEMLERİ İLE MODELLENMESİ**

KEVSER YEŞİLÇİMEN

YÜKSEK LİSANS TEZİ

DANIŞMAN

PROF. DR. ÖZLEM ÇETİNKAYA BOZKURT

BURDUR- 2023

T.C
BURDUR MEHMET AKİF ERSOY ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YÖNETİM BİLİŞİM SİSTEMLERİ ANA BİLİM DALI

AKILLI SERALARDAN TOPLANAN VERİLERİN VERİ
MADENCİLİĞİ YÖNTEMLERİ İLE MODELLENMESİ

KEVSER YEŞİLÇİMEN
YÜKSEK LİSANS TEZİ

DANIŞMAN

PROF. DR. ÖZLEM ÇETİNKAYA BOZKURT

JÜRİ ÜYELERİ

Üye: Doç. Dr. İsmail KIRBAŞ

Üye: Doç. Dr. Onur DOĞAN

BURDUR- 2023

T.C.
BURDUR MEHMET AKİF ERSOY ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

ETİK BEYANI

Burdur Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'ne göre hazırlamış olduğum “Akıllı Seralardan Toplanan Verilerin Veri Madenciliği Yöntemleri İle Modellenmesi” adlı tezin hazırlanması sürecinde akademik ve etik kuralları ihlal etmediğimi, tezimin özgün olduğunu, başvurduğum kaynakları metin içinde, dipnotlarda ve kaynakçada eksiksiz ve bilimsel kurallara uygun olarak gösterdiğimi taahhüt ederim.

Kevser YEŞİLÇİMEN

Burdur, 2023

TEŞEKKÜR METNİ

Bu çalışmanın başarıyla nihayetlenebilmesi için bana her zaman yol gösteren ve desteğini eksik etmeyen saygıdeğer danışmanım Prof. Dr. Özlem ÇETİNKAYA BOZKURT'a gönülden teşekkür ederim.

Beni yetiştiren kıymetli aileme ve beni her zaman destekleyen annem Hatice KOZAN'a, manevi desteğe ihtiyaç duyduğum her an yanımda bulduğum dostlarıma, her sorduğum sorunun cevabını alabildiğim canım arkadaşım Cemal Furkan ÖZBEK'e ve üretim verilerini paylaşan Sandıklı Tarım Firma sahiplerine sonsuz teşekkürlerimi sunarım.

Burdur, 2023

ÖZET

Dünya nüfusunun giderek artması, gıdaya olan ihtiyacın artmasını da beraberinde getirmektedir. Tüketicinin ihtiyaçlarına hızla cevap verebilmek, daha kaliteli ürün sunmak için diğer sektörlerde olduğu gibi tarım sektörünün de çağın getirdiği yeniliklere ayak uydurması gerekmektedir. Gelişen teknoloji akıllı tarım uygulamalarını da beraberinde getirmiştir. Bu sayede çiftçilere akıllı seralarda üretime devam edebilme olanağı sağlanmıştır.

Çalışmanın amacı akıllı serada veri madenciliği yöntemlerini kullanılarak domates üretimindeki sulama suyu ve ortalama iç sıcaklık arasındaki ilişkinin alınan tonaja etkisini incelemektir. Ayrıca son dört yıla ait veriler içinde oluşturulan modellerden doğruya en yakın sonucu veren algoritmalar belirlenmiştir. Belirtilen amaç doğrultusunda akıllı seradan elde edilen üretim verilerinden yararlanılarak veri madenciliği yöntemlerinden, rastgele orman, k-en yakın komşu, yapay sinir ağları ve karar ağaçları algoritmaları kullanılarak modelleme yapılmıştır.

Çalışmada kullanılan üretim verileri 2004 yılında Afyon'un Sandıklı ilçesinde üretime başlamış olan akıllı sera uygulamasını kullanan bir işletmeden elde edilmiştir. Söz konusu veriler Mart 2019 ile Aralık 2022 yıllarına ait domates üretim verileridir. Bu verilerden yararlanarak ayrıca korelasyon analizi yapıp, dağılım grafiğinde sulama ve ortalama iç sıcaklık değerleri incelenmiştir.

Çalışmanın sonucunda akıllı serada yetişen domateslerin verimliliğini etkileyen sulama suyu ve ortalama iç sıcaklık faktörlerinin doğru orantıda olmasının üretim verimliliği üzerindeki olumlu etkisi doğrulanmıştır. Ayrıca, her yıl için ayrı ayrı yapılan modelleme sonucunda farklı algoritmalar en az hata ile doğruya en yakın sonucu vermiştir.

Tarım alanında üretim yapan işletmeler üretim verilerini veri madenciliği yöntemleriyle analiz ederek ürün verimliliğini arttırabilirler.

Anahtar Kelimeler: Akıllı Sera, Veri Madenciliği, Üretim Verimliliği

(YEŞİLÇİMEN, Kevser, Modeling The Data Collected From Smart Greenhouses With Data Mining Methods, Master Thesis, Burdur, 2023)

ABSTRACT

The increasing population of the world causes the need for food to increase. In order to meet the needs of the people and to offer better quality products, those working in the agricultural sector keep up with the innovations. Developing technology has brought smart agriculture applications with it. In this way, farmers were provided with the opportunity to continue production in smart greenhouses.

The aim of the study is to examine the effect of the relationship between irrigation water and average internal temperature in tomato production on the tonnage taken in a smart greenhouse using data mining methods. In addition, algorithms that provide the closest results to reality were determined from the models created within the data of the last four years. In line with the stated purpose, modeling was done using data mining methods, random forest, k-nearest neighbors, artificial neural networks and decision trees algorithms, using the production data obtained from the smart greenhouse.

The production data used in the study was obtained from a business using the smart greenhouse application, which started production in Afyon's Sandıklı district in 2004. The data in question is tomato production data between March 2019 and December 2022. Correlation analysis was performed using these data, and irrigation and average internal temperature values were examined in the distribution graph.

As a result of the study, the positive effect of irrigation water and average internal temperature factors, which directly affect the yield of tomatoes grown in the smart greenhouse, on production efficiency was confirmed. In addition, as a result of the modeling done separately for each year, different algorithms gave the closest result to reality with the least error.

Enterprises engaged in production in the field of agriculture can increase product efficiency by analyzing production data with data mining methods.

Keywords: Smart Greenhouse, Data Mining, Production Efficiency

İÇİNDEKİLER

ETİK BEYANI.....	i
TEŞEKKÜR METNİ	ii
ÖZET.....	iii
ABSTRACT.....	iv
İÇİNDEKİLER.....	v
KISALTMALAR	vii
TABLolar DİZİNİ	viii
ŞEKİLLER DİZİNİ.....	ix
GİRİŞ.....	1

BİRİNCİ BÖLÜM

AKILLI TARIM VE AKILLI SERA

1.1.Akıllı Tarım.....	4
1.2. Akıllı Sera	5
1.2.1.Akıllı Seraların Avantaj ve Dezavantajları	6
1.2.2.Akıllı Seralarda Sulama Sistemleri.....	6
1.2.3.Akıllı Seralarda Havalandırma Sistemleri	7
1.2.4.Akıllı Seralarda Isıtma	9
1.2.5.Akıllı Seralarda İklimlendirme	10
1.2.6.Akıllı Seralarda Verimlilik	10
1.2.7.Dünyada ve Türkiye’de Akıllı Tarım	11
1.2.7.1.Tayvan’da Akıllı Tarım.....	11
1.2.7.2.Hollanda’da Akıllı Tarım	12
1.2.7.3.İsrail’de Akıllı Tarım.....	12
1.2.7.4.Türkiye’de Akıllı Tarım	13

İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

2.1. Veri Ambarı.....	15
2.2. Veri Madenciliği	15
2.3. Veri Madenciliğinde Veri Ambarı Kullanımı	16
2.4. Veri Madenciliği Modelleri.....	17
2.4.1. Niteliğine Göre Veri Madenciliği Modelleri.....	17
2.4.2. İşlevselliğine Göre Veri Madenciliği	18
2.5. Veri Madenciliğinin Tarım Alanında Kullanımı.....	23

ÜÇÜNCÜ BÖLÜM

AKILLI SERALARDAN TOPLANAN VERİLERİN VERİ MADENCİLİĞİ YÖNTEMLERİ İLE MODELLENMESİ

3.1. Araştırmanın Amacı ve Önemi	26
3.2. Araştırmanın Kısıtları.....	26
3.3. Araştırmanın Yöntemi.....	27
3.4. Verilerin Toplanması.....	27
3.5. Araştırma Bulguları.....	28
3.5.1. Korelasyon Analizi Bulguları.....	29
3.5.1.1. Pearson Korelasyon Analizi Bulguları	29
3.5.2. Dağılım Grafiği	33
3.5.3. Toplanan Verilerin Veri Madenciliği Modelleri İle İncelenmesi.....	37
SONUÇ.....	44
KAYNAKÇA.....	47
ÖZGEÇMİŞ.....	53

KISALTMALAR

AR-GE: Arařtırma Geliřtirme

BİT: Bilgi ve İletişim Teknolojileri

IOT: Nenelerin İnterneti

İHA: İnsansız Hava Araçları

R²: Belirleme Katsayısı

MSE: Ortalama Kare Hatası

PLC: Programlanabilir Mantıksal Denetçi

RMSE: Kök Ortalama Kare Hatası

MAE: Ortalama Mutlak Hata

VTYS: Veri Tabanı Yönetim Sistemleri

TABLÖLAR DİZİNİ

Tablo 1. 2019-2022 Yılı Verilerinin Belirleme Katsayısına (R^2) Göre Algoritma Değerleri.....	39
Tablo 2. 2019-2022 Yılı Verilerinin Ortalama Kare Hatasına (MSE) Göre Algoritma Değerleri.....	39
Tablo 3. 2019-2022 Yılı Verilerinin Kök Ortalama Kare Hatasına (RMSE) Göre Algoritma Değerleri	40
Tablo 4. 2019-2022 Yılı Verilerinin Ortalama Mutlak Hatasına (MAE) Göre Algoritma Değerleri.....	41
Tablo 5. 2019-2022 Yılı En Başarılı Sonucu Veren Algoritmalar	42



ŞEKİLLER DİZİNİ

Şekil 1. Endüstri 4.0 İle Tarım 4.0 Arasındaki Teorik Bağıın Açıklanması.....	5
Şekil 2. Damlama Sulama Otomasyon Sistemi	7
Şekil 3. Akıllı Sera Çatı Yapısı.....	8
Şekil 4. Sera Havalandırma Otomasyon Sistemli	8
Şekil 5. Jeotermal Enerji Kullanımlı Otomasyon Sistemi	10
Şekil 6. Veri Madenciliği Süreçleri	16
Şekil 7. Veri Ambarı İle Veri Madenciliği Arasındaki İlişki	17
Şekil 8. Veri Madenciliği Modellerinin Gruplanması	17
Şekil 9. Karar Ağacı Ağaç Yapısı Örneği.....	19
Şekil 10. Rastgele Orman Ağaç Sınıflamaları Örneği	20
Şekil 11. Rastgele Orman Ağaç Sınıflamaları Sonuç Sınıfı Örneği	20
Şekil 12. İleri Beslemeli Yapay Sinir Ağı	21
Şekil 13. Geri Beslemeli Yapay Sinir Ağı	21
Şekil 14. K-En Yakın Komşu Algoritması Sınıf Belirleme İşlemi.....	22
Şekil 15. Kümeleme Algoritması Örneği.....	23
Şekil 16. Tez Çalışmasında Kullanılan Veri Örneği.....	28
Şekil 17. Ortalama İç Sıcaklık, Sulama ve Tonaj Verilerinin Pearson Korelasyon Analizi Sonuçları.....	30
Şekil 18. 2019 Yılı Pearson Korelasyon Analiz Sonuçları	30
Şekil 19. 2020 Pearson Korelasyon Analiz Sonuçları	31
Şekil 20. 2021 Yılı Pearson Korelasyon Analiz Sonuçları	32
Şekil 21. 2022 Yılı Pearson Korelasyon Analiz Sonuçları	33
Şekil 22. Ortalama İç Sıcaklık ve Sulama İlişkisinin Dağılım Grafiği Ölçümlenmesi	34
Şekil 23. 2019 Yılı Dağılım Grafiği	34
Şekil 24. 2020 Yılı Dağılım Grafiği	35
Şekil 25. 2021 Yılı Dağılım Grafiği	36
Şekil 26. 2022 Yılı Dağılım Grafiği	37
Şekil 27. Ortalama İç Sıcaklık, Sulama ve Tonaj Verilerinin Veri Madenciliği Modelleri İle İncelenmesi	38

GİRİŞ

Tarım alanının modernleşmesi teknolojik aletlerin insan gücünün yerini alması, otomasyon sistemlerinin tarımda kullanılmasıyla olmuştur. Önceden tarım işlerinin fiziksel güç gerektirmesi sebebiyle üretim miktarı daha az ve daha fazla insan gücü gerektiriyordu. Elektriğin kullanılmaya başlanması ile seri üretime geçilmesi ve elektronik aletlerin üretim aşamalarında yaygın kullanılması ile birlikte üretim de makineleşmeye gidildi. Bilgisayarların yaygınlaşması, elektroniğin, otomasyon sistemlerinin üretimde kullanılmasıyla birlikte ise üretim otomatik hale geldi insan gücü yerini otomasyon sistemlerine devretti. Günümüz teknolojisinde ise akıllı robotlar, sürücüsü olmayan araçlar, sürücüsü olmayan tırlar, kaptansız gemiler ve sanal gerçeklik gibi örneklerinin çoğaltılabileceği pek çok alanı kapsayan dijital dönüşümler yer almaktadır (Kırkaya, 2020: 5).

Akıllı tarım, modern Bilgi ve İletişim Teknolojilerinin (BİT) tarıma uygulanmasını temsil etmektedir ve üçüncü yeşil devrim olarak da adlandırılmaktadır. Bitki ıslahı ve genetik devrimlerin ardından bahsedilen bu üçüncü yeşil devrim, hassas ekipman, nesnelerin interneti (IOT), sensörler ve aktüatörler, coğrafi konumlandırma sistemleri, büyük veri, insansız hava araçları (İHA), robotik vb. gibi BİT çözümlerinin bir araya getirilmesine dayalı olarak tarım endüstrisini ele almaktadır (Duman ve Özsoy, 2019: 541).

Forbes tarafından sunulan rapora göre, dünya nüfusunun 2050 yılına kadar 9.6 milyardan fazla olacağının tahmin edilmesi sonucu gıda tüketiminde %70'lik bir artışın olacağı öngörülmektedir. Ekilebilir arazilerin sınırlı olmasından dolayı artan bu ihtiyacın üstesinden gelmenin yolu akıllıca planlar yapmaktan geçmektedir. Akıllı tarıma yaklaşımda temel odak, tarımsal üretkenliği ve gelirleri artırmaktır. Akıllı tarım uygulamaları ile mahsul veriminden, gübre uygulamalarından, toprak haritalamadan, hava etkisinden vb. olgulardan veriler toplanarak ve uzmanlık kullanılarak üretkenlik, sürdürülebilirlik ve ekonomiklik amaçlanmaktadır (Duman ve Özsoy, 2019: 541).

Bu çalışmada tarım alanı olarak yüksek verim ve şartların kontrol edilebilirliği açısından sera ortamı seçilmiştir. İncelenen sera ortamı jeotermal enerji ile ısıtılmakta ve sera içerisinde topraksız tarım uygulanmaktadır. Seralar, bitki gelişimi ve üretimi

amacıyla uygun çevre koşullarının sağlanması için oluşturulan, cam, plastik, polikarbon gibi ışığı geçiren örtülü yapılardır. Sera yapılarında bitki büyüme etmenleri olan sıcaklık, karbondioksit, toprak ile hava nemi en uygun düzeylerde sürdürülür (Öztürk, 2003: 40-44). Sera içindeki sıcaklığı, nemi, sulama sistemini kolay ve anında yapabilmek için otomasyon sistemleri kullanılmaktadır. Bu sistem sayesinde bitkiye verilen su miktarı ölçülür, bitkinin ihtiyacı kadar su verilebilir. Sera içi nem, sıcaklık değerleri görülüp bitkinin ihtiyaç duyduğu oranda ayarlanır. Sera içinde bitkinin uygun koşullarda yetişmesi, seraların ısıtma, soğutma ve havalandırma sistemleri ile mümkün olmaktadır. Günümüzde İngiltere ve Hollanda gibi serin iklime sahip ülkeler otomatik ve tam kontrollü seralarda tarımsal faaliyetlerini gerçekleştirmektedir (Kalkaşım vd., 2019: 23).

Teknolojinin bu denli gelişmesi veri artışına da sebep olmuştur. Verilerin işlenmesi depolanması veri artışıyla birlikte oldukça güç duruma gelmiştir. Günümüz teknoloji şirketleri ve kurumlar büyük veriler üzerinde çalışmaktadırlar. Büyük veri yığınları içerisinde yararlı bilginin çekip çıkartılması oldukça zahmetlidir.

Veri madenciliği, veri ambarlarında depolanan çeşitli verileri inceleyerek daha önce keşfedilmemiş bilgileri ortaya çıkarma, bu bilgileri karar verme ve eylem planını oluşturmak için kullanma sürecidir. Bu nedenle veri madenciliği kendi başına bir çözüm değil, çözüme ulaşmak için verilecek kararı destekleyen, problemi çözmek için gerekli olan bilgileri bulmaya yarayan araçtır (Özmen, 2005: 2). Veriler ancak belirli bir amaç doğrultusunda işlendikleri takdirde anlam kazanmaktadırlar. Bu amaç ile günümüzde en çok kabul gören tekniğin veri madenciliği olduğu bilinmektedir. Dünyada veri madenciliği en çok biyoloji, tıp, genetik, finans, borsa, sigorta ve istihbarat alanlarında kullanılmakla birlikte eğitim alanında da sıkça kullanılmaya başlamıştır (Çeşmeli vd., 2015: 37). Türkiye’de son yıllarda veri madenciliği yöntemlerinin tarım alanında da kullanılmaya başlandığını gösteren araştırmalar (Özer vd. ,2022; Aydılek ve Sandıkçı, 2018) bulunmaktadır.

Bu çalışmanın amacı da akıllı serada veri madenciliği yöntemlerini kullanılarak domates üretimindeki sulama suyu ve ortalama iç sıcaklık arasındaki ilişkinin alınan tonaja etkisini incelemektir. Ayrıca son dört yıla ait veriler içinde oluşturulan modellerden doğruya en yakın sonucu veren algoritmalar belirlenmeye çalışılacaktır.

Çalışma sonucunda elde edilen bulgulara göre sera işletmelerinin üretim verimliliklerini incelemelerini sağlayacak modeller öne sürülecektir. Akıllı seralardaki üretimdeki verilerle veri madenciliği yöntemlerinin kullanılmasının bilinci ve verimliliği arttıracakı düşünülmektedir. İlgili literatürde akıllı seralarda üretim verileri ile veri madenciliği modellerinin kullanıldığı çalışmalara rastlanılmamıştır. Bu yönüyle çalışmanın ilgili literatüre önemli bir katkı sağlayacağı düşünülmektedir.

Araştırmanın birinci bölümünde, akıllı tarım ve akıllı sera kavramlarının tanımlarına yer verilmiştir. Akıllı seraların avantaj ve dezavantajları açıklanmıştır. Akıllı seralarda sulama sistemleri, havalandırma sistemleri, ısıtma yöntemleri, iklimlendirme ve verimliliğin öneminden bahsedilmiştir. Dünyada akıllı tarımda öncü olan ülkeler ve Türkiye'deki akıllı tarım uygulamalarından örnekler verilmiştir.

Araştırmanın ikinci bölümünde, veri ambarı, veri madenciliği, veri madenciliği modelleri kavramlarından bahsedilmiştir. Veri madenciliğinin tarım alanında kullanımına değinilmiştir.

Araştırmanın üçüncü bölümünde, araştırmanın amacı ve önemi, yöntemi, kapsamı ve kısıtları, bulgularına yer verilmiştir. Kullanılan verilerin nasıl toplandığına değinilmiştir. Araştırmada korelasyon analizine, dağılım grafiğine ve toplanan verilerin veri madenciliği modellerinden olan k-en yakın komşu, yapay sinir ağları, rastgele orman ve karar ağaçları algoritmaları ile incelemelerine yer verilmiştir. Kullanılan algoritmalar ile en az hatayı veren algoritmalar oluşturulmaya çalışılmıştır.

BİRİNCİ BÖLÜM

AKILLI TARIM VE AKILLI

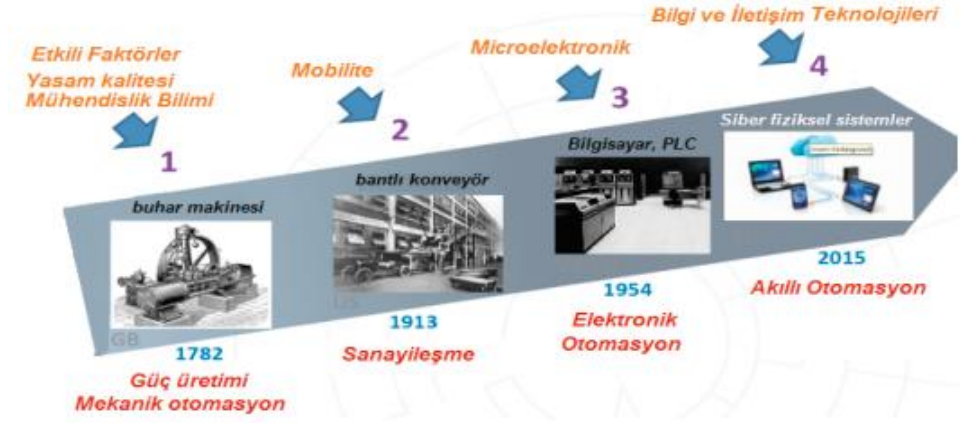
SERA

Araştırmanın temel kavramlarının daha iyi anlaşılması için bu bölümde sırasıyla akıllı tarım ve akıllı sera kavramları açıklanmıştır. Akıllı seralara ilişkin sistemler ile avantajları ve dezavantajlarından bahsedilmiştir. Dünyadaki ve Türkiye’deki akıllı sera örneklerine yer verilmiştir.

1.1. Akıllı Tarım

Tarım, bitkisel veya hayvansal ürün üretimi ve bunların kalitesinin, verimlerinin yükseltilmesi aynı zamanda üretilen ürünlerin uygun koşullarda işlenip, saklanması ve maddi değer elde edilmesini planlayan bilim dalıdır. Yani tarım, kırsal kesimlerin kalkınmasını amaçlayan aynı zamanda sanayi sektörüne hammadde sağlayan, her ülke için ihracat gelirlerinin artmasına olanak sağlayan bir sektördür (Keskinkılıç, 2013: 121). Tarımın bu denli önemli olması her geçen gün yeniliği ve rekabeti de beraberinde getirmiştir. Bitkisel tarım üretimde sık karşılaşılan risklerin başında yangın, sel, dolu, don ve kuraklık gibi doğal etmenlerin üretim verimliliğini etkilemesi çiftçiyi maddi ve manevi yönden sarsıcı bir şekilde etkilemesi akıllı tarımı gerekli kılmaktadır.

2011 yılında Almanya’ da Endüstri 4.0 olarak ortaya çıkan dördüncü sanayi devrimi imalat sanayinin üretim süreçleri başta olmak üzere tüm sektörlerin üretim faaliyetlerini etkilemiştir. Şekil 1’de Endüstri 4.0 ile tarım 4.0 arasındaki ilişki görülmekte olup dört endüstri devriminin, tarım sektörü üzerinde yarattığı değişiklikler özet olarak ifade edilmeye çalışılmıştır. Birinci sanayi devriminde suyun gücünden faydalanılan mekanik tezgâhlar icat edildi ve buhar gücünden gittikçe daha fazla yararlanılmaya başlandı. Daha sonra, 1800’lü yılların sonunda elektrikle çalışan ve toplu üretim yapan (montaj hattı) ilk üretim hattı kullanıldı ve bu da ikinci sanayi devrimi olarak adlandırıldı. Üçüncü sanayi devrimi ise 1969 yılında başladı. İlk Programlanabilir Mantıksal Denetçi (PLC) imalatın otomatikleştirilmesi için elektronik ve bilişim alanlarında kullanılmaya başlandı. Dördüncü sanayi devrimi ise sanayide, genel olarak makinelerin insan gücüne gerek kalmaksızın kendilerini ve üretim süreçlerini yönetmeye başlamalarıyla ortaya çıkmıştır (Kaya, 2019: 132).



Şekil 1. Endüstri 4.0 İle Tarım 4.0 Arasındaki Teorik Bağın Açıklanması

Kaynak: (Pakdemirli vd. 2021:80)

Dijital ve akıllı tarım çağı olarak da bilinen yeni döneme Tarım 4.0 adı verilmektedir. Tarımsal üretim sürecinde kullanılan tüm tarım makineleri sensörler ile donatılmakta olup Nesnelerin İnterneti (IoT) tarım sektörüne de girerek makinelerin tüm üretim süreci boyunca birbirleriyle iletişim halinde olmasını sağlamaktadır. Akıllı tarım teknolojileri, ürünlerin yalnızca gıda kalitesini veya üretimini iyileştirmek için değil, aynı zamanda çiftçilerin mahsullerin ihtiyaçlarını daha iyi görmesine ve beslenmelerini buna göre ayarlamasına olanak sağlamak için de kullanılmaktadır (Yıldız, 2023: 2-3). Akıllı tarım, iklim faktörleri, toprak özellikleri, toprak nemi vb. değişikliklerin izlenmesine izin verdiği için yetiştirilecek ürün tahmini ve daha sonra eşzamanlı sağlanan veri ile diğer analizler gibi ilgili birçok konuyu ele almaktadır (Özer vd., 2022: 1418).

Seralar plastik veya camla kaplı üretim alanları olup, günümüzde teknolojinin de dahil olduğu işletmeler haline dönüşmektedir. Birim alandan maksimum verim arayışı sonucunda geliştirilen seracılığın, gelecekte teknolojik tarımsal üretim teknikleri içerisinde daha öne çıkan ve daha stratejik bir konuma geleceği düşünülmektedir.

1.2. Akıllı Sera

Seralar, örtü altı yetiştiriciliği yapılan tarım alanları olarak da tanımlanmaktadır. Cam veya plastik gibi ışığı geçiren materyallerle üretim alanının kapatılmasıdır. Bu sayede bitkilerin üretim mevsimleri dışında üretilmesi için uygun koşullar sağlanmaktadır. 1990'lı yıllarda seralarda teknolojinin kullanılır hale gelmesiyle birlikte

modern seralar da kurulmaya başlanmıştır (Gale vd.,2014: 69). Akıllı seralar, tarım alanında yapılan yenilikler ile modernleşen tarımın getirdiği yeniliklere adapte olup maksimum verim almak için üretimde geçilen yöntemdir. Akıllı seralar minimum maliyet ile yüksek üretim verimliliğini amaçlayan işletmelerdir. Sulama sistemleri, havalandırma sistemi otomasyonu gibi teknolojik sistemleri tarımsal üretimde senkron bir şekilde kullanarak kaliteli ürün yetiştirmek, üretimde yararlanılan insan gücünü azaltmak, sulama maliyetlerini azaltmak ve olumsuz hava koşullarından en az hasarla üretim dönemini geçirmelerini sağlamıştır (Çağlar, 2018: 5).

1.2.1. Akıllı Seraların Avantaj ve Dezavantajları

Akıllı seraların bir takım avantaj ve dezavantajları bulunmaktadır. Kısaca avantajlarını sıralamak gerekirse:

- Üretim verimliliği artar.
- Ürün çeşitliliği artar.
- Dört mevsim taze ürün alınabilir.
- Olumsuz hava koşullarından minimum etkiyle ürünler korunabilir.
- İklimin uygun olmadığı ortamlarda bitkinin yaşaması için uygun ortam sağlanıp üretim yapılabilir (Çağlar, 2018: 1).

Dezavantajları ise:

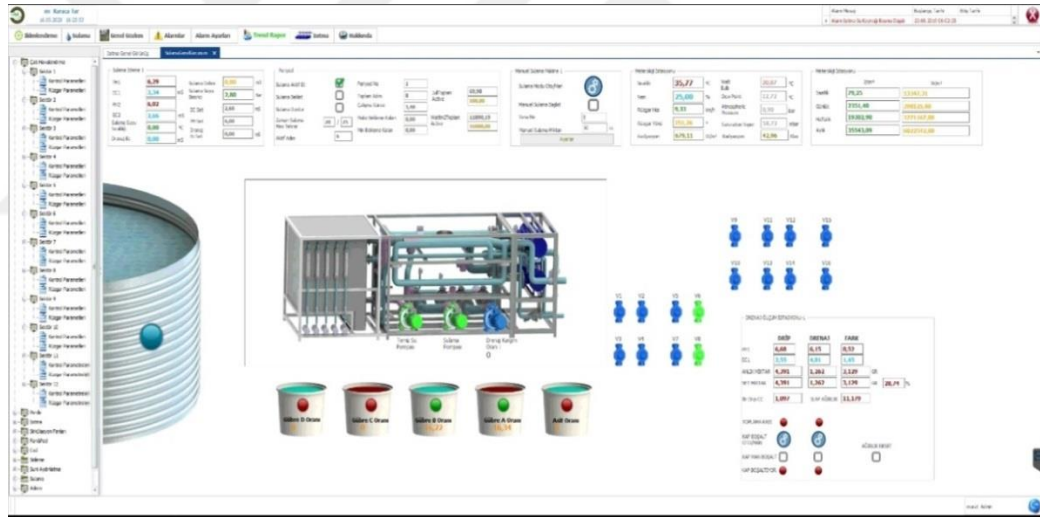
- Maliyetlidir.
- Teknolojiye gereksinim duyar.
- Kalifiyeli personel gereksinimi vardır.
- 7/24 internet bağlantısı gereklidir (Çağlar, 2018: 1).

Görüldüğü üzere akıllı seraların avantajları dezavantajlarına göre daha fazladır.

1.2.2. Akıllı Seralarda Sulama Sistemleri

Sulamanın manuel olarak yapılması her bitkinin ihtiyacı kadar su almasını engellemektedir. Önceden sulamanın tarlaya döşenen borular ile fiskiyeler yardımıyla

(yağmurlama yöntemiyle) yapılması tarlada bulunan bitkilerin hepsine yeterli miktarda gitmemesine neden olmaktadır. Çiftçilerin kendilerinin yaptığı damlama sulama ile belli saat aralarında suyu açıp kapatarak vermeleri vakit alıyordu, bazen sulamayı geç veya erken yapmaları, bitkinin zamanında su almasını engellemekteydi. Ancak akıllı tarım uygulamaları sayesinde sistem bitkinin ihtiyacı olan suyu girilen miktarda, girilen saat aralığında insana ihtiyaç duymadan otomatik olarak verebilmektedir. Bu yöntemle ayrıca iş gücünden tasarruf sağlanmasının yanında su tasarrufu da sağlanmaktadır. Şekil 2’de gösterilmiş olan damla sulama otomasyon sistemi sayesinde su, zaman, enerji ve iş gücü tasarrufu sağlanmaktadır. Damla sulama sisteminde; damlatıcılar aracılığıyla merkez birimden araziye uygulama yapılabilir veya yakın aralıklarla konuşlandırılmış damlatıcılardan ya da üzerinde eşit aralıklı sulama delikleri bulunan tüplerden su verilebilir. Buradaki amaç, bitkilerin terlemeyle su kaybını yeterli olarak karşılayabilmek ve topraktaki nem düzeyini bitkinin ihtiyacını karşılamaktır.



Şekil 2. Damlama Sulama Otomasyon Sistemi

Kaynak: (Karaca Otomasyon Sistemi)

1.2.3. Akıllı Seralarda Havalandırma Sistemleri

Seralarda üretimin uygun koşullarda yapılabilmesi için içerisinde havalandırılması gerekmektedir. Aynı zamanda bitkinin uygun sıcaklığa ulaşması ve dış etkenlerden korunması içinde havalandırma sistemi gereklidir. Modern seralarda havalandırma otomasyon sistemi ile yapılmaktadır. Sisteme girilen bilgiler



Şekil 3. Akıllı Sera Çatı Yapısı

Kaynak: (Verilerin Alındığı Sandıklı Tarım Serasının Çatı Yapısı)

Şekil 4'te otomasyon sistemli seraların havalandırma yöntemi gösterilmektedir. Sol üstte bulunan rüzgâr yönünü göstermekte olan pusula sistemi sayesinde çatılar rüzgâr yönüne göre girilen bilgiler doğrultusunda çatı sistemlerini otomatik açıp kapatmaktadır. Yağış algıladığı zaman ise çatılar sistem tarafından otomatik olarak kapatılmaktadır.



8

1.2.4. Akıllı Seralarda Isıtma

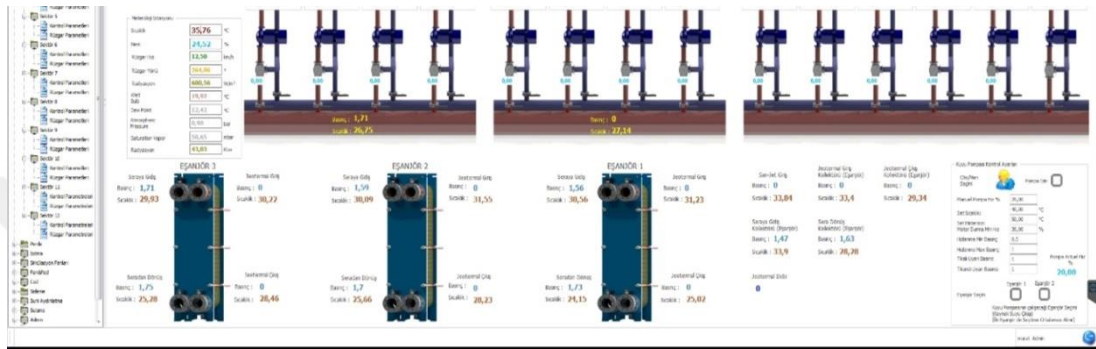
Seraların ısınma sorunları üretilecek ürünlerin maliyetini doğrudan etkilemektedir. Bu nedenle üreticiler özellikle az maliyetli ve enerji tasarrufu yapabilecekleri yöntemleri seçmektedirler. Seranın bulunduğu konum da ısıtma için seçilecek yöntemleri etkilemektedir. Sera ısıtmasında kömür ile ısıtma, elektrik enerjisi ile ısıtma, doğalgaz ile ısıtma ve yenilenebilir enerji kaynakları ile ısıtma yöntemlerinden birisi kullanılabilir. Yenilenebilir enerji kaynaklarının ucuz, ulaşımının kolay olması ve çevreye zararsız olması üreticilerin dikkatini çekmekte ve seracılığı da olumlu yönden etkilemektedir. Üreticilerin yenilenebilir enerji kaynaklarından faydalanıp enerji tasarrufu sağlayabilecekleri yöntemler güneş enerjisi, biyogaz ve jeotermal enerji gibi ısıtma yöntemlerinden seranın bulunduğu konuma göre yararlanabilirler (Yüksel, 2011: 93).

Jeotermal enerji tarımsal ürün üretim ve işletmesinde kullanılan önemli bir enerji kaynağıdır. Tarımsal üretimde jeotermal enerjinin doğrudan kullanım alanı seracılıktır. Jeotermal enerji, yeryüzünün derinliğinde birikmiş olan ısı ile oluşan, sıcaklığın sürekli bölgesel atmosfer ortalama sıcaklığının üzerinde ve çevresindeki yeraltı ve yerüstü sularına göre daha fazla erimiş mineral, çeşitli tuzlar ve gazlar içerebilen sıcak su ve buhar olarak tanımlanmaktadır (Çerçioğlu ve Şahin, 2016: 460).

Seraların bulunduğu bölgede jeotermal su sürdürülebilir tarım için çok yaygın olarak kullanılmaktadır. Seraların ısıtılmasında, sıcak su kuyuları ve yer altına döşenen çelik izolasyonlu borular yardımıyla sıcak su plakalı eşanjör sisteminden geçtikten sonra seraya verilir. Sera ısıtmasında kullanıldıktan sonra Re enjekte yapılarak jeotermal suyun alındığı bölgenin biraz ilerisine geri verilir. Isıtmanın en çok kullanıldığı zaman kış mevsimidir. Önceki araştırmalar, jeotermal enerjiyle ısıtılan seralar, sıvı ve gaz yakıt ile ısıtılan seralara oranla maliyetinin az olduğunu göstermiştir (Milivojevic ve Martinovic, 2003: 33).

Akıllı tarım uygulamalarının işletmede etkin bir şekilde kullanılmasıyla tarım ve kırsal kalkınmada önemli derecede gelişme sağlanmıştır. Teknolojik gelişmelerle birlikte üretken ve bilinçli bir işletme portföyü ortaya çıkmıştır (Uzun vd., 2018: 3). Ayrıca jeotermal enerji ile ısınan seralarda 12 ay üretim yapılmaktadır. Jeotermal suyun

sera içerisinde kullanılması açısından yapay zekâ teknolojileri ile geliştirilmiş olan ve tarım ve kırsal kalkınmada üretim verimliliğini artırmak için otomasyon program sistemini Şekil 5 içerisinde jeotermal suyun kullanımı gösterilmektedir. Jeotermal enerji kullanımlı otomasyon sisteminde Sol altta bulunan eşanjör göstergeleri sayesinde sera içerisine verilecek olan jeotermal ısının sıcaklık, basınç ve debi ölçümleri hakkında bilgi sağlanmaktadır. Üst kısımda bulunan materyal boru sistemleri seranın içerisinde dolaşan jeotermal suyun sıcaklığı hakkında bilgi sunmaktadır.



Şekil 5. Jeotermal Enerji Kullanımlı Otomasyon Sistemi

Kaynak: (Karaca Otomasyon Sistemi)

1.2.5. Akıllı Seralarda İklimlendirme

Yılın her günü bitkiden verim alabilmek iklimlendirme ile mümkün olmaktadır. Akıllı seralarda iklimlendirme işletme maliyetleri arasında önemli yere sahiptir. Yaz aylarında sera içerisini havalandırma, soğutma yapmak, kış aylarında ise bitkinin büyümesi, verimli ve kaliteli ürün alına bilmesi için havalandırma ve sera içi ısıtımların bitki için uygun düzeyde olması gerekmektedir. Sera içi iklimlendirmenin zamanında ve gerekli düzeyde yapılmaması üretilen üründe verim kaybına ve kalite düşüklüğüne neden olabilmektedir. Akıllı seralarda iklimlendirme, sera içerisine yerleştirilen sensörler ile otomasyon sisteminde girilen parametreler doğrultusunda yapıldığı için verim kaybını en aza indirmekte ve ürün kaybını engellemektedir (Şen, 2013: 49).

1.2.6. Akıllı Seralarda Verimlilik

Seralarda verimlilik sulama, havalandırma, ısıtma sistemleri, üretimi etkileyecek çevresel ve insani faktörlerden en az maliyetle fazla ürün almayla sağlanmaktadır. Akıllı seralarda üretime geçilmesiyle birlikte sulama maliyetlerinde azalma ve su tasarrufu sağlandığı, bitkilere zamanında ve ihtiyaç duyduğu oranda su verilmesiyle verimli ürün

alınması, iklimlendirme metotlarının doğru ve zamanında kullanılmasıyla ısınma veya soğutmadan kaynaklanan verimlilik kayıpları en aza indirilip, kaliteli ürünler alınmaktadır. Basit fakat zaman alan işlerin birçok personel ile yapılması yerine otomasyon sistemleri ile yapılması sayesinde ise iş gücü verimi arttırılmıştır (Güzey, 2020: 12-17).

1.2.7. Dünyada ve Türkiye’de Akıllı Tarım

Günümüzde büyük işletmeler tarım teknolojilerini takip ederek sonuçları iyi değerlendiren firmalardan biri olmayı amaçlarlar. Artan nüfusla birlikte gıda ihtiyacının 2050 yılına kadar 705 ile 10,5 milyar kişi gibi önemli bir kişi sayısına ulaşması beklenmektedir (Güzey, 2020: 5). Aynı zamana küresel ısınmanın artması gıda ya olan talebe cevap veremez hale gelmesi, fiyatlarda artışa neden olacaktır. Çevre kirliliğinin artması, iklim değişikliği de tarım üretiminde olumsuz etkiye sahiptir. Fakat yenilenebilir enerji kaynaklarının kullanılması az da olsa çevre kirliliğini ve maliyetleri azaltmaktadır.

Seraların yaygınlaşmasının birçok sebebi vardır. İlk sırada artan nüfus ve toprak kaynaklı sebeplerdir. Dünya nüfusunun artması tarım için kullanılan toprakların gıda ihtiyacını karşılayamayacağı düşünülmektedir. Toprak ve su kaynakları yetersizliği nedeniyle bazı ülkelerin dışa bağımlılıktan kurtulmak istemeleri, su ve gübre girdi maliyetlerini azaltarak daha az girdiyle yüksek verim almak için akıllı seralar ve topraksız tarım üreticiler için önemli bir alternatif olarak ortaya çıkmaktadır (Şahin, 2020).

1.2.7.1. Tayvan’da Akıllı Tarım

36.193 km² alan ve 23,38 milyon nüfusu (2023) olan Tayvan hızlı gelişen ada ülkesidir. Tayvan’da yüzde 25’lik bir tarımsal alan aktif kullanılmaktadır. Tarımda makineleşmeyle birlikte yeniliklere önem vermiş ve tarımda kullanılacak cihazların üretilmesi ve akıllı tarım proje örnekleriyle öne çıkmaktadırlar. Dronlar, robotlar, güneş panelleri gibi akıllı tarımda ihtiyaç duyulan birçok teknolojinin üreticisidir. Ada, tarımsal üretimde, tarım 4.0’a merkez olmaya adaydır. Tayvan ihtiyaç duyulan mühendislik alt yapısıyla akıllı tarım için ihtiyaç duyulan bileşenlerle birlikte nitelikli iş gücüne sahiptir (Kılavuz ve Erdem, 2019: 139).

Tayvan sera tarzı kapalı alanlarda üretim yapmaya önem vermektedir. İklim değişikliklerinden dolayı ürün kayıplarını azaltmak için. Tayvan kapalı alanda led teknolojisiyle üretim yapmaktadır. Tayvan led teknolojisinde ikinci büyük led sektörüne sahiptir. Özel led ışıklarla güneş ışığını taklit edebilen uygun ayarlarla bitkinin yılın her gününde büyümesi amaçlanmaktadır. Ayrıca damla sulama yöntemiyle iş gücü ve su tasarrufu da sağlanmaktadır. Akıllı tarım uygulamalarıyla sulama ile gübrelemede yüksek verime ulaşmak hedeflenmektedir (Kılavuz ve Erdem, 2019: 140).

1.2.7.2. Hollanda’da Akıllı Tarım

41.543 km² alan ve 17 milyon 850 bin nüfusa (2023) sahip Hollanda’nın başarılı olduğu alanlar teknolojik hazırlık, eğitim ve yükseköğretim, inovasyondur. Gelirin büyük bir kısmı ticaret ve sanayi sektöründen karşılanmaktadır. Akıllı tarım ile etkin su ve gübre kullanımı yapmaktadır (Kılavuz ve Erdem, 2019: 142).

Hollanda’da arazilerin kooperatiflere kayıtlı olması gerekmektedir. Kooperatifler, hangi bitkinin, hangi dönemde üretileceğinin kararını verip çiftçileri bu konularda uzman kişilere yönlendirmektedir. Uzmanlar tarlada gerekli incelemeleri yapıp o dönem ne üretileceği konusunda çiftçileri bilgilendirmektedirler. Amaç aynı ürünün fazla üretilmesinin önüne geçmek ve verimliliği üst düzeye çıkarmaktır (Donat, 2020).

Hollanda’da tarımın gelişmesinde etkili olan bir diğer faktör tarım üniversitelerinin olmasıdır. Üniversitelerin hedefi az maliyetle kaliteli ürün üretilmesini araştırmaktır. Bu tarımsal bilgi birikimini kullanıp, danışmanlık yapan firmalar ile başka ülkelere pazarlamaktadırlar. Kısaca Hollanda akıllı tarım uygulamalarını kullanıp hayata geçirebilmiştir (Kılavuz ve Erdem, 2019: 142-143).

1.2.7.3. İsrail’de Akıllı Tarım

27.817 km² alan ve 9.053 milyon nüfusa (2023) bulunmaktadır. İsrail büyük önem arz eden sektörlerde devlet kontrolü altında olan karma ekonomik sisteme sahip bir ülkedir. Söz konusu sektörler, teknolojik ürünler, tarımsal teknoloji, iletişim, kimya sektörü gibi dünyada adından söz ettiren sektörlerdir (Kılavuz ve Erdem, 2019: 145). Bulunduğu coğrafyada 27.817 km² yüz ölçümü ile tarımsal üretimde üstünlüğü olamayacak bir ülke gibi görülmektedir. Fakat araştırma geliştirme (Ar-Ge) çalışmaları ve makineleşmeyi yakından takip ederek tarım alanında öne çıkan ülke konumundadır

(Kılavuz ve Erdem, 2019: 145). Dünyaya damla sulama sistemini tanıtan, tarım üretiminde kendine yetebilen, tohumlar üzerinde oynayabilen, verimli olmayan arazilerde üretim yapabilen ve bunları ihraç edebilen ülkedir. İleri teknolojiler yardımıyla toprak ve su kıtlığı içerisinde akıllı teknolojiler yardımıyla üretim yapmaktadır. Tatlı kiraz domatesler, kuraklığa dayanıklı tohumlar, hasat sonrası dayanıklı ürün gibi alanlarda başarıları vardır. İsrail'in tarımda teknolojik ilerlemeye ve bilgiye dayanan üstünlüğü vardır. Bu başarının sırrı İsrail tarım endüstrisinin ve çiftçilerin teknolojik araştırmalarda işbirliği yapmasına bağlıdır. İsrail bütçesinin yüzde 17'sini tarım alanı için ar-ge ye ayırmıştır (Kılavuz ve Erdem, 2019: 145-146).

1.2.7.4.Türkiye’de Akıllı Tarım

Tarım alanlarında kontrol edilemeyen hava koşulları ürün kayıplarına neden olmuştur. Tahmin edilemeyen hava sıcaklığı, yağmur ve dolu yağması tarım ürünlerinin deforme olmasını bitkilerin zarar görmesine neden olmaktadır. Bunların önüne geçebilmek için seralarda üretim yapılmaya başlanmıştır. Seralar tarım alanı üzerinin plastik veya cam yapı malzemesi ile kapatılmasıdır. Toprak yapısının dikkate alınması fidenin dikim metodu, bitki koruma uygulamaları, sulama sistemleri ve hasat teknikleri, gibi konularda geleneksel izleme yöntemlerinin, profesyonel bir şekilde izleme, denetleme ve uygulama olanağı sunmuştur akıllı tarım. Seralar içerisine yerleştirilmiş sensörler sayesinde içerinin sıcaklığını ölçmesi girilen sıcaklık değerlerine göre bitki için gerekli olan sıcaklığın ayarlanması amaçlanmıştır (Kılavuz ve Erdem, 2019: 148).

Türkiye’deki tarımsal üretim işletmelerinin bir kısmı büyük aile işletmesi şeklinde olduğu için, tarımın dönemsel yapılmasına, verimin istenilen oranda olmamasına ve satış üretiminin az olmasına neden olmaktadır. Aynı zamanda girdi maliyetlerinin yüksek olmasıyla üretimin azalması sonucu fiyatlar artmakta ve üretimin tüketimi karşılayamaması sonucunda dış alım yapılmaktadır. Avrupa Komisyonu tarım politikasında AB ile üye ülkeler arasında bütçenin etkin ve daha verimli kullanılmasını sağlamak için değişikliklerin arasında sulama suyunun kullanılmasında dijital tarımın yaygınlaşması konusunda karar alınmıştır. Otomasyonla sulama sistemlerinin yaygınlaştırılmasıyla birlikte su tasarrufu sağlamak, geri dönüşüm için yapılacak çalışmalarla tarım daha da sürdürülebilir hale gelecektir (Kılavuz ve Erdem, 2019: 149).

İsrail gibi bir lke olumsuz kořullarda tarımsal retim yapabiliyorsa, verimli ve byk toprakları olan Trkiye’de kapsamlı olarak yapılacak ar-ge alıřmalarıyla birlikte tarımsal eēitimin artması ve tarımda makine kullanımına gidilerek sebze ve meyve ihracatısı konumuna gelinebilir.



İKİNCİ BÖLÜM

VERİ MADENCİLİĞİ

Araştırmanın bu bölümünde veri madenciliği kavramı ve modelleri açıklanmıştır.

2.1. Veri Ambarı

Veri ambarı, 1980’li yıllarda bilgisayarların ve bilgisayar kullanım alanlarının genişlemesiyle birlikte ilk olarak William Inmon tarafından kullanılmıştır (Dunham 2003: 35). Birbirleri ile ilişkili verilerin tarama analizlerinin yapılabileceği depolardır. Veri ambarı farklı kaynaklardan toplanan verilerin içerisinden kolay arama yapılmasını sağlamaktadır. Veri ambarları isminden de anlaşılacağı üzere karmaşık, ham verilerin depolandığı ambarlardır.

Veri ambarlarında depolanan birbirleri arasında ilişkisi bulunmayan yığınlardan, sağlık, coğrafi sistemler, pazarlama, üretim gibi birçok bölüm geleceğe yönelik sonuçlar çıkartarak, yönetim stratejilerini planlamada kullanabilecekleri sistemdir. İlk başta işletmeler için maliyetli görünse bile ilerleyen safhalarda getirisinin fazla olduğu anlaşılmaktadır (Bilekdemir, 2010: 6)

2.2. Veri Madenciliği

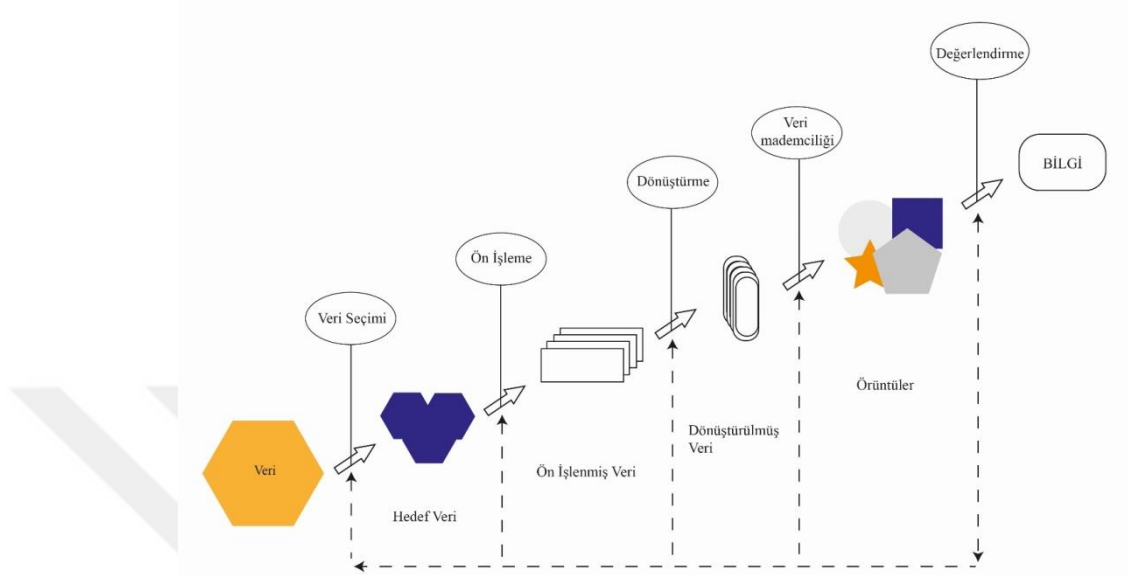
Veri ambarlarında keşfedilmemiş bilgileri, veriler arasındaki kavramları ve ilişkileri çeşitli modeller ve analizler yardımıyla mantıklı tahminler yapılarak ortaya çıkarılması olarak tanımlanabilen veri madenciliğinin literatürde farklı tanımları bulunmaktadır. Bunlardan birkaçına örnek vermek gerekirse;

Büchner ve arkadaşları (1997: 1), büyük veri yığınları arasında önemli olan yararlı bilgilerin ve örüntülerin keşfedilmesi olarak tanımlanmıştır

Thuraisingham (1999), büyük miktardaki verilerin saklandığı veri tabanlarına bilgileri ayırt edici sorularla önceden farkında olunmayan faydalı bilgilerin gün ışığına çıkarılması işlemi olarak ifade etmektedir.

Yalçın (2008: 134) ise, otomatik araçlar yardımıyla büyük miktardaki verilerin içerisinde saklı olan desenlerin ortaya çıkartılması için araştırmacılar tarafından gerçekleştirilen keşif sel analizler olarak tanımlamaktadır.

Özkul ve Pektekin (2009: 71), büyük verilerin içerisinde gizlenen desenler arası ilişkilerin, veriler içerisinde bulunan değişimlerin ve düzensiz verilerin istatistiksel olarak önemli olan bilgilerin keşfedilmesi olarak ifade etmektedir.



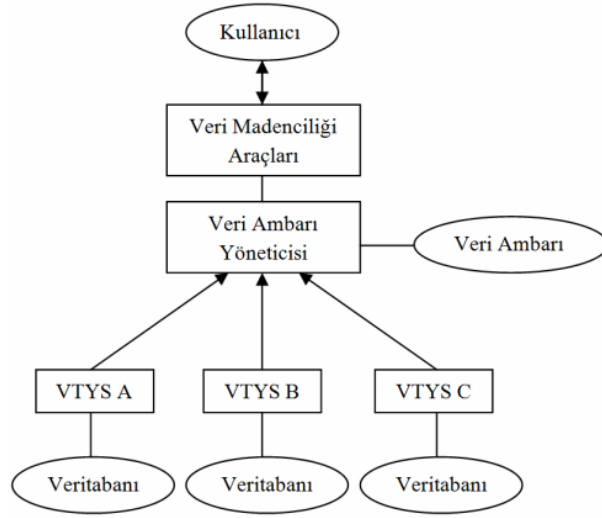
Şekil 6. Veri Madenciliği Süreçleri

Kaynak: (Sandıkçı, 2022:14)

Veri ambarlarında karmaşık, yığınlar halinde depolanan birbirinden farklı verileri analiz ederek keşfedilmemiş bilgileri çıkarma ve bu bilgileri karar verme, plan oluşturmak için kullanılan süreç veri madenciliğidir (Şekil 6). Buradan anlaşılacağı üzere veri madenciliği tek başına sonuç üreten değil, istenilen çözüme ulaşmak için gerekli bilginin bulunmasını sağlayan ve bu bilgiler doğrultusunda en uygun kararın verilmesini destekleyen araçtır (Özmen, 2005: 2).

2.3. Veri Madenciliğinde Veri Ambarı Kullanımı

Veri ambarları karmaşık fakat birbirleri ile ilişkili verilerin analiz edilebildiği bir depodur. Ambarda yığınlar halinde bulunan verilerin ayırt edici özelliği, karar almaya yardımcı bilgiler içermesidir. Veri ambarlarında veriler organize edilirken veri madenciliği, yararlı bilgileri bulmaya ve bunları hayatta kullanılabilir hale getirmeyi amaçlar. Şekil 7’de veri ambarı ile veri madenciliği arasındaki ilişki görülmektedir (Köktürk ve Dirsehan, 2012: 3).

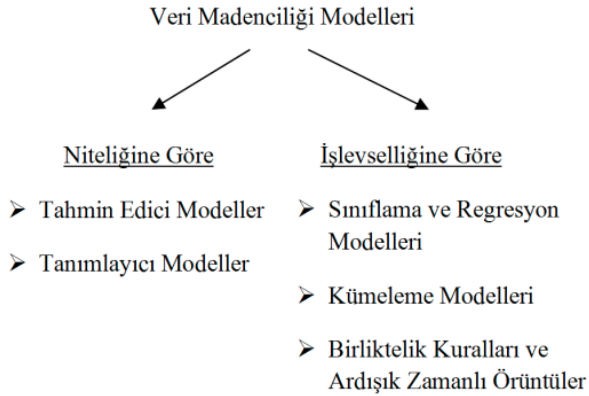


Şekil 7. Veri Ambarı İle Veri Madenciliği Arasındaki İlişki

Kaynak: Thuraisingham, 2003: 50

2.4. Veri Madenciliği Modelleri

Veri madenciliğinde kullanılan modeller Şekil 8’de gösterildiği gibi iki yaklaşıma göre gruplanabilir: Niteliklerine göre ve işlevselliklerine göre



Şekil 8. Veri Madenciliği Modellerinin Gruplanması

Kaynak: Köktürk ve Dirsehan, 2012:6

2.4.1. Niteliğine Göre Veri Madenciliği Modelleri

Tahmin edici modeller, sonuçları bilinen veri modellerinden yararlanarak yeni modeller geliştirip, yeni veri kümelerine uygulama yaparak sonucun tahmin edilmesini amaçlar (Özekes, 2003: 67). Bir başka ifadeyle tahmin edici modellerde, mevcut veri üzerinden yaratılan modelin, sonuçları bilinmeyen başka bir veri kümesi için tahmin edilmesi amaçlanmaktadır (Köktürk ve Dirsehan, 2012: 4). Örneğin bir turizm işletmesi

yirmi yıldır gelecek turist sayısının tahminleri yapıyor ve her yıl yapılan tahminlerin sapmalarını hesaplamaktadır. Önceki yılın raporlarının verileri kullanılarak gelecek yılın turist sayısı tahmini olarak bulunmaya çalışılır.

Tanımlayıcı modelde karar verme aşamasında öncü olabilecek mevcut verilerden elde edilen örneklerin çözümlenmesini sağlanmaktadır (Akpınar, 2000: 5). Örneğin bir mağaza müşteri veri tabanı analiz edilerek, alışveriş sıklıklarına göre müşterilerini çeşitli kümelere ayırabilir. Bu kümelere göre de farklı promosyon uygulamaları yapılabilir.

2.4.2. İşlevselliğine Göre Veri Madenciliği

İşlevselliğine göre veri madenciliği modelleri üç başlıkta sınıflandırılabilir (Akpınar, 2000: 2).

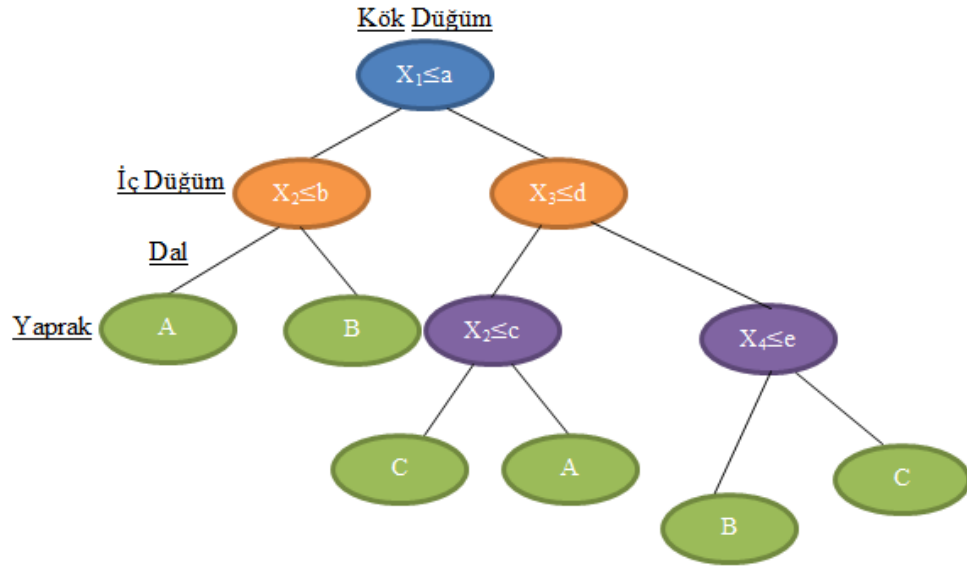
Sınıflandırma ve Regresyon: Sınıflama ve regresyon modeli, değerli olan veri sınıflarını belirleyen veya gerçekleşmesi mümkün olan veri eğilimlerini tahmin edebilecek veri modellerini kuran veri analiz yöntemleridir. Sınıflama modeli kategorik (sıralanmış) fonksiyonları tahmin eder, regresyon ise süreklilik sağlayan fonksiyonların tahmininde kullanılır (Han vd. 2006: 15).

Sınıflama ve regresyon modelinde kullanılan teknikler şunlardır: (Akpınar, 2000: 6).

1. Karar Ağaçları
2. Rastgele Orman
3. Yapay Sinir Ağları
4. Genetik Algoritmalar
5. K-En Yakın Komşu
6. Naive Bayes Metodu

Karar Ağaçları: Adından anlaşılacağı üzere ağaç görünümünde olup, tahmin edici bir yöntemdir. Düğümler, dallar ve yapraklardan meydana gelir. Karar ağaçları yapıları itibarıyla ters çevrilmiş ağaca benzemektedir. Karar ağacı algoritmasında işlem kök düğümünden başlar, test edilen değişkenler sonucuna göre dallara ayrılarak yeni düğümler oluşturulur ve yukarıdan aşağı doğru düğümleri takip ederek yaprağa (sonuca)

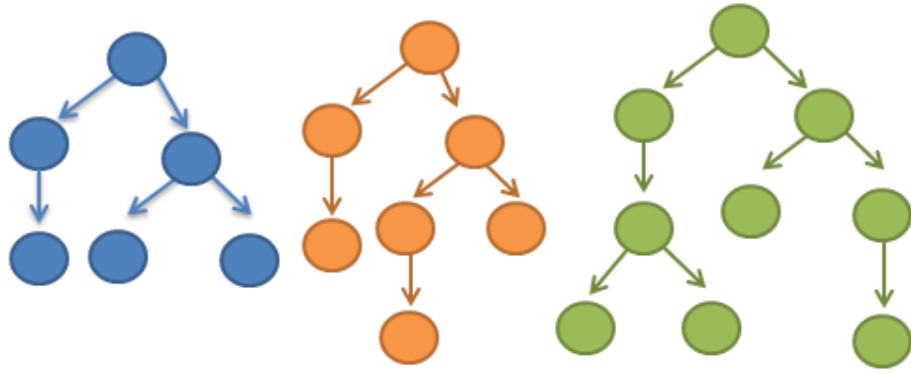
ulaşana kadar devam eder (Bilekdemir, 2010: 24). Şekil 9’da bir karar ağacı yapısı örneği görülmektedir.



Şekil 9. Karar Ağacı Ağaç Yapısı Örneği

Kaynak: (Kavzoğlu ve Çölkesen 2010:40)

Rastgele Orman: Denetimli bir model olan rastgele orman algoritması, rastgele seçilen veri kümelerinden oluşur (Gültepe, 2019: 12). Rastgele orman algoritmasında bir çok karar ağacı birbirlerinden farklı olarak oluşmaktadır, her bir düğümde “mtry” olarak isimlendirilen rastgele seçim yapılır. *Mtry* değeri veride bulunan özelliklerin toplam sayısının karekökü alınarak hesaplaması yapılır. İşlem için gerekli olan her bir ağacın oluşturulması için bu işlem tekrarlanır. Bütün işlemler tamamlandıktan sonra orman oluşturulduktan sonra çözüm için sınıflandırılan ağaçların çoğunluk değeri alınarak sonuç sınıfı elde edilir (Gültepe, 2019: 12). Şekil 10’da rastgele orman algoritması ile oluşturulan karar ağacı algoritma örnekleri yer almaktadır. Şekil 11’de ise rastgele orman algoritması ile ulaşılan sonuç sınıfı örneği gösterilmektedir.



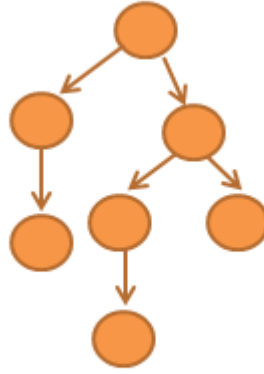
Sınıf 1

Sınıf 2

Sınıf 3

Şekil 10. Rastgele Orman Ağaç Sınıflamaları Örneği

Kaynak: (Yazar tarafından oluşturulmuştur.)



Sonuç Sınıfı

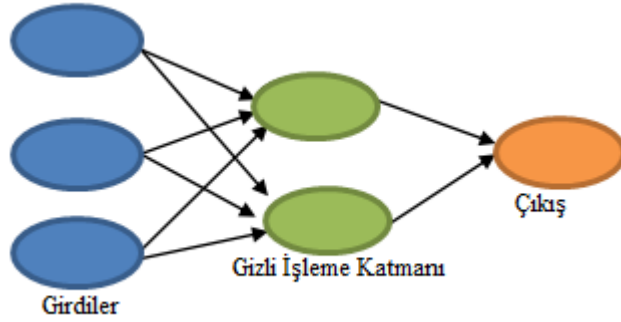
Şekil 11. Rastgele Orman Ağaç Sınıflamaları Sonuç Sınıfı Örneği

Kaynak: (Yazar tarafından oluşturulmuştur.)

Yapay Sinir Ağları: Yapay sinir ağları, özellik bakımından birbirleriyle bağlantılı ve bağımsız değişkenlerin arasındaki karmaşık ilişkilerin modellenenebilmesi için tercih edilen algoritmadır. Yapay sinir ağları modeli kullanılarak oluşturulan çözümlerlerin yorumlanması, karar ağacı ile oluşturulan algoritma yorumlarına göre daha zor olabilmektedir (Seyrek ve Ata, 2010: 72). Yapay sinir ağları, insan beyin sinir

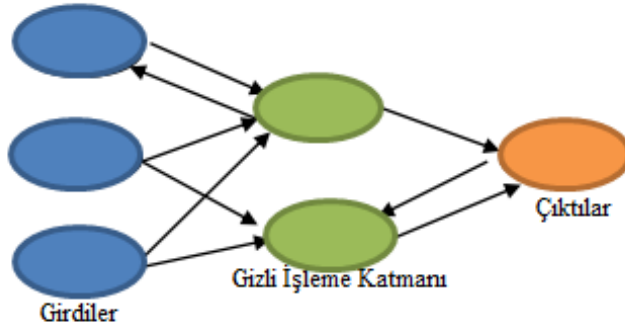
sistemi model alınarak benzer yapıda geliştirilmiştir (Jackson, 2002: 273). Girdi değerleriyle modelleme yapar ve öğrendiği bilginin çıktısını verir (Erdem, 2009: 41).

Yapay sinir ağları üç katmandan oluşmaktadır. Bunlar girdiler, gizli katman ve çıkış katmanıdır. Aynı zamanda yapay sinir ağları iki ana yapıdan oluşmaktadır. Bu yapılardan biri ileri beslemeli yapay sinir ağı iken diğeri geri beslemeli yapay sinir ağıdır (Deliismail 2019:28). Şekil 12’de ileri beslemeli yapay sinir ağı yapısı, Şekil 13’de geri beslemeli yapay sinir ağı yapısı örnekleri görülmektedir.



Şekil 12. İleri Beslemeli Yapay Sinir Ağı

Kaynak: (Deliismail 2019:28)



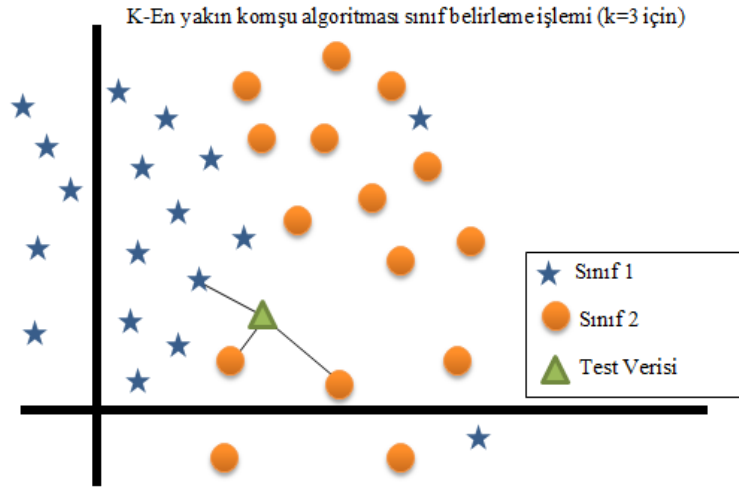
Şekil 13. Geri Beslemeli Yapay Sinir Ağı

Kaynak: (Deliismail 2019:28)

Genetik Algoritmalar: Karışık, geleneksel yöntemlerle çözümü imkânsız olabilecek problemlerin çözüm algoritması olarak görülmektedir. Genetik algoritmalar, yapıları itibarıyla evrimsel yapıya benzetilmektedir (Alagöz, vd. 2014: 9-10). Problem popülasyonu (bireyler) ile ilgili bilgiler girildiğinde, evrimsel bilgi işlem, girilen popülasyon için potansiyel olarak yeni çözümlerle işlem alanını genişletiyor. Evrimsel

süreçte en iyi veya en uygun bireyler üretilir. Genetik Algoritmaların evrimsel sürece benzetilmesindeki neden de en uygun modeli bulma çalışmasıdır. Genetik algoritmalar raststaldır buda hangi adımdan başlanıp, hangi adımda biteceğini bilmediğimiz anlamına gelir. Bu sebeple her çalışmada farklı sonuç elde etme olasılığı yüksektir (Dunham, 2003: 67). Genetik algoritmalar ile probleme ait uygun çözüm bulunamayabilir fakat çözüme en hızlı ve verimli şekilde ulaşılması mümkündür (Bolat, 2006: 35).

K-En Yakın Komşu: Sınıfları belli olan kümede, kümeye yeni katılacak verinin hangi sınıfa dahil edileceğini belirlemek için uygulanan yöntemdir (Kolyiğit vd: 2012:424). Yani sisteme eklenen yeni verinin k adet komşusuna bakılır. Veri hangi sınıfa yakınsa o sınıfa kaydedilir. Şekil 14’de sınıfları belirli olan çözüm kümesinin içerisinde test verisi yerleştiriliyor ve k-en yakın komşu algoritması ile k adet komşusuna bakılıp hangi sınıfa yakınsa o sınıfa kaydedildiği görülmekte.

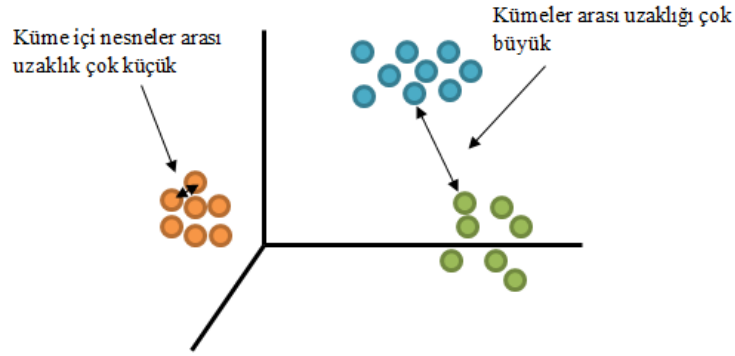


Şekil 14. K-En Yakın Komşu Algoritması Sınıf Belirleme İşlemi

Kaynak: (Çavuşoğlu ve Kaçar, 2018:210)

Naive Bayes: Naive bayes algoritmasında girilen değişken ile bağımsız değişkenler arasındaki bağlantıyı inceleyerek analiz etmemizi sağlayan tahmin edici aynı zamanda tanımlayıcı yöntemler algoritmasıdır. Burada anlatılandan hareketle, içinde bulunduğu durum hakkında inceleme yapılabilen ve çıkan sonuca göre gelecek hakkında tahmin yapılabilir (Deliismail, 2019: 11). Naive bayes algoritmasında bir sorunun çözümü için o sorunu etkileyen tüm teoremlerin o sorunu sağlama olasılıkları ile çarpımıdır (Bilekdemir, 2010: 34).

- **Kümeleme:** Kümeleme algoritmasında, girilen verinin ortak sınıflara veya kümelere ayrılması işlemidir (Özekes, 2003: 71). Birbirine benzeyen fakat özellik olarak incelendiğinde farklı verilerin, ayrı ayrı kümelere yerleştirilmesidir. Şekil 15’de gösterilen veri kümelerinin de birbirlerine benzediği fakat nesneler arası uzaklıklarına ve renklerine göre kümelendikleri görülmektedir.



Şekil 15. Kümeleme Algoritması Örneği

Kaynak: (Alagöz vd, 2014:11)

Kümeleme modelleri veri tabanındaki kayıtlarda farklı kümelerin bulunmasıdır (Köktürk ve Dirsehan, 2012: 7). Örneğin, bir havayolu işletmesi müşteri veri tabanını analiz ederek, yolculuk sıklıklarına göre müşteriler çeşitli kümeler oluşturabilir. İşletme her bir küme için farklı kampanyalar düzenleyebilir.

- **Birliktelik Kuralı:** Büyük verilerin arasında birliktelik ilişkilerini bulur. Büyük miktardaki verilerin ilginç birliktelik ilişkilerini keşfederek karar almayı daha verimli hale getirmeyi amaçlar. En bilinen örneği market sepet uygulamasıdır (Özekes, 2003: 76). İşletme alışverişte müşterinin hangi ürünleri birlikte satın aldığını belirler. Örneğin, bir spor mağazası birliktelik analiziyle hangi ürünlerin birlikte satıldığını saptayabilir, böylece daha etkin promosyonlar oluşturabilir.

2.5. Veri Madenciliğinin Tarım Alanında Kullanımı

Tarım alanında da veri kullanımı yeni bir oluşum değildir. Yenilik sektörün dijitalleşmesinden kaynaklanmaktadır. Geçmiş dönemlere bakıldığında çiftlik ortamında bu verilerin toplanması, depolanması ve istenildiği zaman ulaşıp işlenmesi zaman alabilmekteydi. Ayrıca ciddi veri kayıplarının yaşanması da olası durumlar arasında yer

almaktaydı. Sensör teknolojilerinin gelişmesi tarım alanlarının da dijitalleşmesiyle birlikte, çiftçilerin belirli parametreleri gerçek zamanlı olarak izlemesi ve depolayabilmesiyle insan gücü kullanımının azalmasına yardımcı olmaktadır (Güzel ve Okatan, 2022: 202).

Veri madenciliği tarım alanıyla ilgili çeşitli konularda karar vermede önemli bir rol oynamaktadır. Örneğin veri madenciliği modellerinin kullanımıyla birlikte sulama ve gübreleme sistemlerinin optimizasyonu gibi gerekli tarımsal veriler ile sulama zamanı ve miktarı tahmin edilebilmektedir. Hava olayları insan kontrolü dışında gelişir bu sebeple önlem alınması güçleşmektedir. Ancak doğru tahminle zamanında müdahaleyle olumsuz hava koşullarının etkilerini azaltarak gereken önlemlerin alınması mümkündür. Ayrıca tarımdaki faaliyetleri verimli bir şekilde planlayarak, maliyetleri en aza indirip, verimi arttırabilmektedir. Yerli ve yabancı literatürde tarım alanında veri madenciliği tekniklerinin kullanıldığı çok sayıda araştırma yer almaktadır. Bunlardan bazılarına örnekler verilmiştir

Jagielska ve diğerleri (1999), veri madenciliğinin tarımla ilgili alanlara yönelik uygulamaları açıklamışlardır. Verim tahmini çok önemli bir tarımsal sorundur. Her çiftçi ne kadar verim beklendiğini bilmek isteyebilir. Araştırmada olasılık teorisi ve bulanık küme teorisindeki verilerle verimlilik tahmini yapılmaya çalışılmıştır.

Rajagopalan ve Lal (1999), tarımsal veri kümeleri için veri madenciliği tekniklerinin uygulanmasına ilişkin bir dizi çalışma yürütmüşlerdir. Örneğin, araştırmada günlük yağışları ve diğer hava durumu değişkenlerini simüle etmek için K-En Yakın Komşu uygulanmıştır.

Somvanshi ve diğerleri (2006), yapay sinir ağlarını ve Box Jenkins yöntemini kullanarak yağışın modellenmesini ve tahmin edilmesini tartışmışlardır.

Veenadhari (2007) 'nin araştırmasının bulguları, karar ağacı analizinin, soya fasulyesi mahsulünün verimliliğinin çoğunlukla nemden, ardından sıcaklık ve yağıştan etkilendiğini ortaya koymuştur. Araştırmada karar ağacından oluşturulan kuralların, yüksek veya düşük mahsul verimliliğine yönelik koşulların belirlenmesinde faydalı olduğu kanıtlanmıştır.

Tellaecche ve arkadaşları (2007: 72-79), hassas tarımda yabancı ot tespitini bilgisayarlı görme sisteminden yararlanarak yapılışı anlatılmaktadır. Tahıl ürünlerine zarar veren ve zehirli olan avena sterilis otunun tespiti için otomatik bilgisayarlı görme sistemi araştırma konusudur. Bayes apriori olasılığının k-en yakın komşu algoritması ile sağlandığı k-en yakın komşu algoritmasına dayalı hibrit karar verme sistemi tasarlanmıştır.

Urtubia ve arkadaşları (2007: 1512-1517) şarap fermantasyon tanklarında tıkanıklık yaşanabileceği gibi üretimin yavaşlamaması için araştırma gerçekleştirmişlerdir. Üretim sürecinde problemlerin tahmini için k-en yakın komşu algoritmasını kullanarak iyi bir fermantasyon süreci gerçekleştirilebileceği öngörülmüştür.

Aydilek ve Sandıkçı (2018: 1-7), toprak ekiminden başlayarak hasat yapılincaya kadar tarımsal deneylerden elde edilmiş veri kümeleri ile veri madenciliği analizlerinin yapılması, tarımsal üretimde kalite ve verimlilik artışının sağlanması için zeki karar verme sistemlerinin oluşturulmasını önermişlerdir.

Sandıkçı (2022: 1-54), pamuk üretiminde rekolte tahmini için veri madenciliği yöntemlerinden yararlanmış ve Şanlıurfa ilinde örnek çalışma gerçekleştirmiştir. Toprak hazırlığı, ekim, gübreleme ilaçlama, sulama, hasat gibi değişen veri kümesine veri madenciliği yöntemlerini uygulamıştır. Elde edilen tüm verilerle gerçekleştirilen yöntemler ile yapay sinir ağları algoritması ve destek vektör makinesi algoritması en başarılı (hatasız) sonucu üretmiştir.

Özer, Kuş ve Yıldız (2022: 1417 – 1429), akıllı tarımda veri analizi ile ürün tahmini üzerine çalışmışlardır. Veri madenciliği algoritmaları ile farklı modeller oluşturmuşlardır. Rastgele Orman algoritmasıyla geliştirilen model sayesinde Hindistan'ın tarımsal ürün haritası çıkarılmıştır.

ÜÇÜNCÜ BÖLÜM

AKILLI SERALARDAN TOPLANAN VERİLERİN VERİ MADENCİLİĞİ YÖNTEMLERİ İLE MODELLENMESİ

3.1. Araştırmanın Amacı ve Önemi

Çalışmanın amacı akıllı serada veri madenciliği yöntemlerini kullanarak domates üretimindeki sulama suyu ve ortalama iç sıcaklık arasındaki ilişkinin alınan tonaja ya da verime etkisini incelemektir. Belirtilen amaç doğrultusunda korelasyon analizi yapıp, dağılım grafiğinde domates üretimine ait sulama ve ortalama iç sıcaklık değerleri ele alınmıştır.

Ayrıca son dört yıla ait veriler içinde oluşturulan modellerden doğruya en yakın sonucu veren algoritmaları belirlemeye çalışmaktır. Bir başka ifadeyle sulama, ortalama iç sıcaklık ve tonaj verileri, veri madenciliği yöntemlerinden rastgele orman algoritması, k-en yakın komşu, yapay sinir ağları ve karar ağaçları algoritmaları kullanılarak üretim verimliliğine ilişkin modeller oluşturulmuştur.

Çalışmanın sonucunda elde edilen bulgular ile akıllı seralarda üretim yapan işletmelerin, veri madenciliği yöntemlerini kullanarak üretim verimliliklerini inceleyebilecekleri ispatlanmış olacaktır. Aynı zamanda bir sonraki üretim yılında firma bünyesinde bulunan verilerden yararlanarak, akıllı sera işletmelerinin daha sağlıklı ve doğru modellemelerle üretim verimliliklerini incelemelerine imkân sağlamaktadır.

3.2. Araştırmanın Kısıtları

Her araştırmada olduğu gibi bu araştırmanın da birtakım sınırlılıkları bulunmaktadır. Bunlardan ilki araştırmaya katkı sağlayacak olan firmanın verebileceği veriler ile sınırlı olmasıdır. Afyonkarahisar ili Sandıklı ilçesindeki sera işletmesi inceleneceği için coğrafi kısıt söz konusudur. Ayrıca verilerin toplanacağı işletme domates üretimi yaptığı için sadece domates fidesinin verimliliği incelenecektir. Geneli yansıtabilmesi açısından sorun teşkil edebileceğinden dolayı ilerdeki çalışmalar için uygulamanın farklı bölgelerde farklı bitki çeşitlerine uygulanarak karşılaştırma yapılabilmesi açısından yararlı olacağı düşünülmektedir.

3.3. Araştırmanın Yöntemi

Araştırmada tarım işletmesinin üretim verilerini incelemek üzere veri madenciliği yöntemleri kullanılacaktır. Veri madenciliği yöntemlerinden rastgele orman, k-en yakın komşu, yapay sinir ağları ve karar ağaçları algoritmaları kullanılarak üretim verimliliğine ilişkin algoritmalar oluşturularak incelenmiştir.

Elde edilen veriler yıllar bazında ele alınarak son dört yıla ait veriler içinde oluşturulan modellerden doğruya en yakın sonucu veren algoritmalar belirlenmiştir. Ayrıca söz konusu verilerden sulama ve ortalama iç sıcaklık değerleri arasındaki korelasyon analizi yapıp, dağılım grafiği incelenmiştir. Domates ürününün sulama ve ortalama iç sıcaklık değerlerinin tonaja etkisi de belirlenmeye çalışılmıştır.

3.4. Verilerin Toplanması

Araştırmada kullanılacak olan veriler bir tarım işletmesinden elde edilmiştir. Seçilen firma 2004 yılında Afyonkarahisar ilinin Sandıklı ilçesinde kurulmuştur. Yetkililer tarafından Sandıklı ilçesinin ısıtılmasından dönen jeotermal suyun değerlendirilmesi için düşünülmüş ve araştırmalar sonucu bu bölgeye sera kurulabileceği kanaatine varılmıştır. İşletme bölgedeki kurulan ilk seradır. Hollanda ve İsrail modeli temel alınarak sera mimarisi geliştirilmiştir. İlk olarak 2004 yılında 10.000 m² üzerine kurulmuş ve büyümeye devam etmiştir. 2007 yılında 7.000 m² alan eklenmiş, üretim ve istihdam arttırılmıştır. 2010 yılında 10.000 m² alan daha kurularak üretime devam etmiştir. 2016 yılında 25.000 m² alan daha kurularak üretim ve istihdam arttırılmıştır. İşletme 2013 yılında bölgenin jeotermal zenginliğinden yararlanmak için sera çevresinde araştırma yapmış ve 1115 m² derinliğinde sıcak su kuyusunu bünyesine dahil etmiştir. Kış dönemi boyunca şehir ısıtmasından dönen suyu kullanmaya devam etmekte, yaz aylarında gece sıcaklığı düştüğü zaman bünyesindeki jeotermal enerjiden yararlanmaktadır.

Örtü altı yetiştiriciliği yani topraksız tarım seçilmesinin nedeni ise daha kaliteli üretim yapabilmek ve jeotermal kullanımına uygun olmasıdır. Sıra arası raylardan ve katır altlarından geçen borulardan ısıtma sağlanmaktadır. Bu sistemi toprak üzerine uygulamak, boruların toprak üzerine gelecek kısımlarının kısa zaman içerisinde çürümesine neden olacaktır ve toprak zemini ısıtmak kolay olmayacak ısı kaybı yaşanacaktır. Sera ısıtma sisteminde jeotermal kullanılmasının en önemli özelliği

olumsuz hava koşullarında 12 ay üretim sağlanabilmesidir. Dikilen fidenin cinsine çiçek yapısına göre 9-10 ay içinde ürün alınabilmektedir.

Sandıklı Tarım işletmesinden Mart 2019-Aralık 2022 yılları 4534 adet üretim verisi firma sahibini izniyle excel dosyasıyla alınmıştır. Şekil 16’da araştırmada kullanılan veri örneğinin görüntüsü yer almaktadır. Alınan üretim bilgilerinde beş farklı veri bulunmaktadır. Bunlar; hasatın yapılp sevk edildiği tarihler, sevk gerçekleşen tarihlerdeki tonaj miktarı, bitkinin büyümesi ve meyvenin olgunlaşması için geçen süreçte gerekli sulamanın ne kadar yapıldığı, bitkinin yetiştiği ortamdaki ortalama iç sıcaklığın derecesi ve domates cinsidir.

GÜN	TOPLAM TONAJ	SULAMA LT	İÇ SICAKLIK ORTALAMA	DOMATES CİNSİ	GÜN	TOPLAM TONAJ	SULAMA LT	İÇ SICAKLIK ORTALAMA	DOMATES CİNSİ	GÜN	TOPLAM TONAJ	SULAMA LT	İÇ SICAKLIK ORTALAMA	DOMATES CİNSİ	GÜN	TOPLAM TONAJ	SULAMA LT	İÇ SICAKLIK ORTALAMA	DOMATES CİNSİ
8.04.2019	0	812	19	CLIMBO	8.04.2020	13185	812	19	CLIMBO	8.04.2021	19211	812	19	SIRANZO	8.04.2022	5635	812	19	SIRANZO
9.04.2019	7108	1708	20	CLIMBO	9.04.2020	0	1708	20	CLIMBO	9.04.2021	1701	1708	20	SIRANZO	9.04.2022	0	1708	20	SIRANZO
10.04.2019	0	2226	19	CLIMBO	10.04.2020	0	2226	19	CLIMBO	10.04.2021	0	2226	19	SIRANZO	10.04.2022	0	2226	19	SIRANZO
11.04.2019	0	378	17	CLIMBO	11.04.2020	0	378	17	CLIMBO	11.04.2021	0	378	17	SIRANZO	11.04.2022	0	378	17	SIRANZO
12.04.2019	0	1652	19	CLIMBO	12.04.2020	0	1652	19	CLIMBO	12.04.2021	0	1652	19	SIRANZO	12.04.2022	0	1652	19	SIRANZO
13.04.2019	0	2016	19	CLIMBO	13.04.2020	0	2016	19	CLIMBO	13.04.2021	0	2016	19	SIRANZO	13.04.2022	1040	2016	19	SIRANZO
14.04.2019	10747	1974	20	CLIMBO	14.04.2020	0	1974	20	CLIMBO	14.04.2021	0	1974	20	SIRANZO	14.04.2022	5400	1974	20	SIRANZO
15.04.2019	0	1988	21	CLIMBO	15.04.2020	26096	1988	20	CLIMBO	15.04.2021	17912	1988	21	SIRANZO	15.04.2022	2080	1988	20	SIRANZO
16.04.2019	27703	1764	20	CLIMBO	16.04.2020	8955	1764	20	CLIMBO	16.04.2021	31763	1764	20	SIRANZO	16.04.2022	0	1764	20	SIRANZO
17.04.2019	17722	266	17	CLIMBO	17.04.2020	11719	266	17	CLIMBO	17.04.2021	420	266	17	SIRANZO	17.04.2022	0	266	17	SIRANZO
18.04.2019	0	1666	20	CLIMBO	18.04.2020	0	1666	20	CLIMBO	18.04.2021	0	1666	20	SIRANZO	18.04.2022	0	1666	20	SIRANZO
19.04.2019	0	294	17	CLIMBO	19.04.2020	0	294	17	CLIMBO	19.04.2021	0	294	17	SIRANZO	19.04.2022	0	294	17	SIRANZO
20.04.2019	19820	1988	19	CLIMBO	20.04.2020	0	1988	19	CLIMBO	20.04.2021	0	1988	19	SIRANZO	20.04.2022	0	1988	19	SIRANZO
21.04.2019	10093	2310	20	CLIMBO	21.04.2020	34157	2310	20	CLIMBO	21.04.2021	0	2310	20	SIRANZO	21.04.2022	0	2310	20	SIRANZO
22.04.2019	0	2184	21	CLIMBO	22.04.2020	9177	2184	21	CLIMBO	22.04.2021	19023	2184	21	SIRANZO	22.04.2022	4200	2184	21	SIRANZO
23.04.2019	10705	1680	20	CLIMBO	23.04.2020	0	1680	21	CLIMBO	23.04.2021	16661	1680	20	SIRANZO	23.04.2022	5373	1680	21	SIRANZO
24.04.2019	0	2016	21	CLIMBO	24.04.2020	0	2016	21	CLIMBO	24.04.2021	0	2016	21	SIRANZO	24.04.2022	0	2016	21	SIRANZO
25.04.2019	0	2170	21	CLIMBO	25.04.2020	0	2170	21	CLIMBO	25.04.2021	0	2170	21	SIRANZO	25.04.2022	0	2170	21	SIRANZO
26.04.2019	0	1848	21	CLIMBO	26.04.2020	0	1848	21	CLIMBO	26.04.2021	0	1848	21	SIRANZO	26.04.2022	0	1848	21	SIRANZO
27.04.2019	0	1792	21	CLIMBO	27.04.2020	11398	1792	21	CLIMBO	27.04.2021	0	1792	21	SIRANZO	27.04.2022	19225	1792	21	SIRANZO
28.04.2019	32568	1974	22	CLIMBO	28.04.2020	13445	1974	22	CLIMBO	28.04.2021	0	1974	22	SIRANZO	28.04.2022	0	1974	22	SIRANZO
29.04.2019	0	2170	22	CLIMBO	29.04.2020	65863	2170	22	CLIMBO	29.04.2021	34769	2170	22	SIRANZO	29.04.2022	0	2170	22	SIRANZO
30.04.2019	0	728	18	CLIMBO	30.04.2020	0	728	18	CLIMBO	30.04.2021	26622	728	18	SIRANZO	30.04.2022	19357	728	18	SIRANZO
1.05.2019	13394	7960	20	CLIMBO	1.05.2020	0	7960	20	CLIMBO	1.05.2021	0	7960	20	SIRANZO	1.05.2022	0	7960	20	SIRANZO
2.05.2019	0	742	17	CLIMBO	2.05.2020	0	742	18	CLIMBO	2.05.2021	0	742	17	SIRANZO	2.05.2022	0	742	18	SIRANZO
3.05.2019	0	1554	20	CLIMBO	3.05.2020	0	1554	22	CLIMBO	3.05.2021	0	1554	20	SIRANZO	3.05.2022	0	1554	22	SIRANZO
4.05.2019	24052	1232	19	CLIMBO	4.05.2020	0	1232	19	CLIMBO	4.05.2021	0	1232	19	SIRANZO	4.05.2022	0	1232	19	SIRANZO
5.05.2019	24148	1596	20	CLIMBO	5.05.2020	32482	1596	18	CLIMBO	5.05.2021	0	1596	20	SIRANZO	5.05.2022	31523	1596	18	SIRANZO
6.05.2019	0	1988	20	CLIMBO	6.05.2020	13291	1988	19	CLIMBO	6.05.2021	0	1988	20	SIRANZO	6.05.2022	19553	1988	19	SIRANZO
7.05.2019	0	1344	18	CLIMBO	7.05.2020	6012	1344	20	CLIMBO	7.05.2021	18102	1344	18	SIRANZO	7.05.2022	0	1344	20	SIRANZO

Şekil 16. Tez Çalışmasında Kullanılan Veri Örneği

3.5.Araştırma Bulguları

Araştırmamanın amacına uygun olarak verilerin orange machine learning programında analiz edilmesi uygun bulunmuştur. Sulama, ortalama iç sıcaklık verilerinin tonaj verimliliğini test etmek için pearson korelasyon analiz sonuçlarına bakılmıştır. Dağılım grafiği ile sulama, ortalama iç sıcaklık ve tonaj değerleri arasında bir ilişkinin olup olmadığı incelenmiştir. Söz konusu değişkenlere ait regresyon doğrusuna yakınlığı ve lineer modelin ne kadar doğru olduğuna ilişkin değerler gösterilmiştir. Çalışmanın son aşamasında ise veri madenciliği modellerinden k-en yakın komşu, rastgele orman, yapay sinir ağları ve karar ağaçları algoritmalarının oluşturulmasına yer verilmiştir.

Mart 2019- Aralık 2022 yılları arasındaki veri görsellerinin anlaşılabilir olması için veriler yıl bazlı olacak şekilde incelenmiştir.

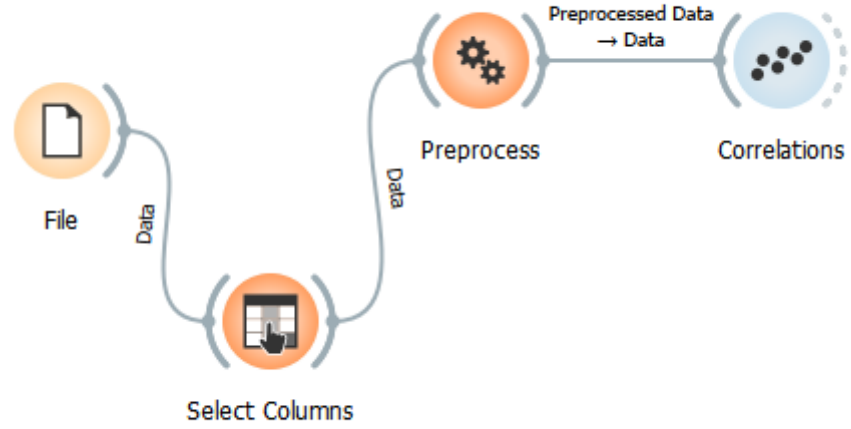
3.5.1. Korelasyon Analizi Bulguları

İki değişken arasındaki doğrusal ilişki veya bir değişkenin birden fazla değişkenle arasında olan ilişkinin kontrol edildiği, derecelerinin ölçüldüğü model korelasyon analiz modelidir. Değişkenlerin normal dağılım göstermesi gereken modelde değerler “r” ile gösterilir. Veriler -1 ile +1 değerleri arasında değer alırlar. r değeri -1 ise ters yönlü doğrusal ilişki, r değeri +1 ise aynı yönde ilerleyen doğrusal ilişki ve r değerinin 0 olması durumunda değişkenler arasında bir ilişki olmadığı anlaşılır (Özbek, 2021:83).

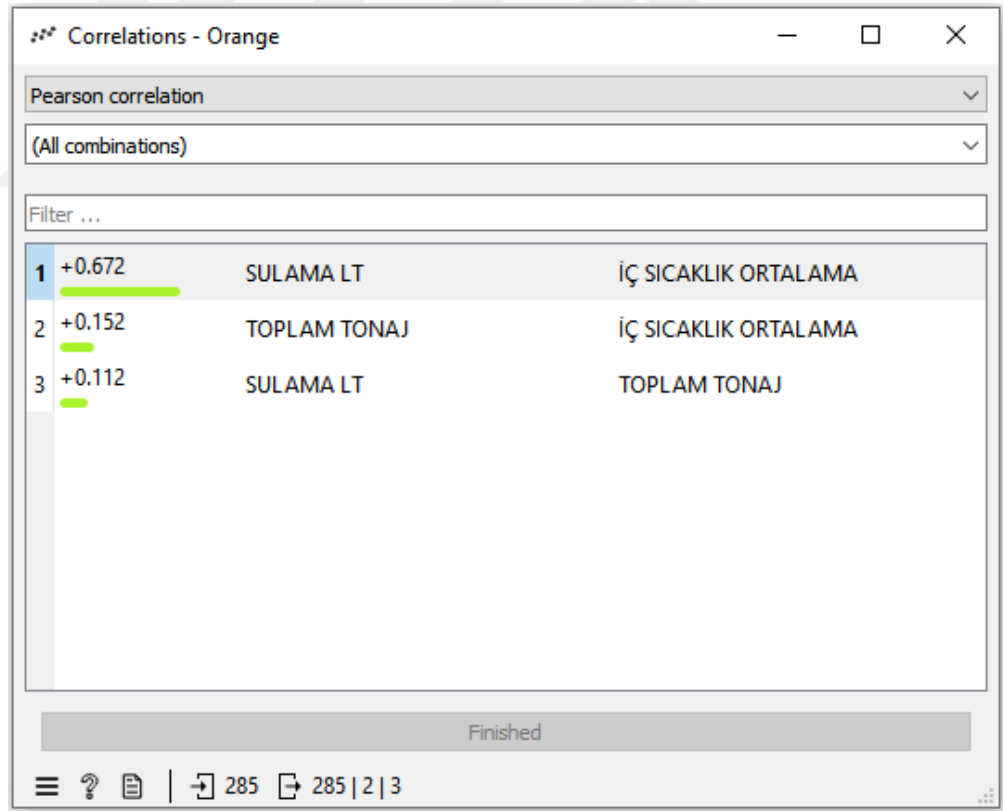
3.5.1.1. Pearson Korelasyon Analizi Bulguları

İki sürekli değişken arasında anlamlı bir ilişkinin olup olmadığı sorusunun cevabı aranır. Pearson korelasyon katsayısı “r” ile gösterilir. Değerler -1 ile +1 arasında değişir. $r=-1$ ise tam negatif doğrusal ilişki olduğu, $r=+1$ ise tam pozitif doğrusal ilişki olduğu $r=0$ ise değişkenler arasında ilişki olmadığını gösterir (Kalaycı ve Sungur, 2018: 116).

Alınan üretim verilerinin Mart 2019- Aralık 2022 yılları arasındaki veriler değerlendirilmiş olup, sulama, ortalama iç sıcaklık ve tonaja olan etkileri incelenmiştir. Mart 2019- Aralık 2022 yıllarına ait tüm veriler kullanılmış olup, veriler arası normalizasyon işlemi de yapılmıştır. Veriler arasında anlamlı bir ilişki olduğu için Pearson korelasyon katsayısı yöntemi kullanılmıştır. Şekil 17’de verilerin analiz edildiği Orange Machine Learning programında Pearson korelasyon katsayısı analizi için kullanılan parametreler gösterilmektedir



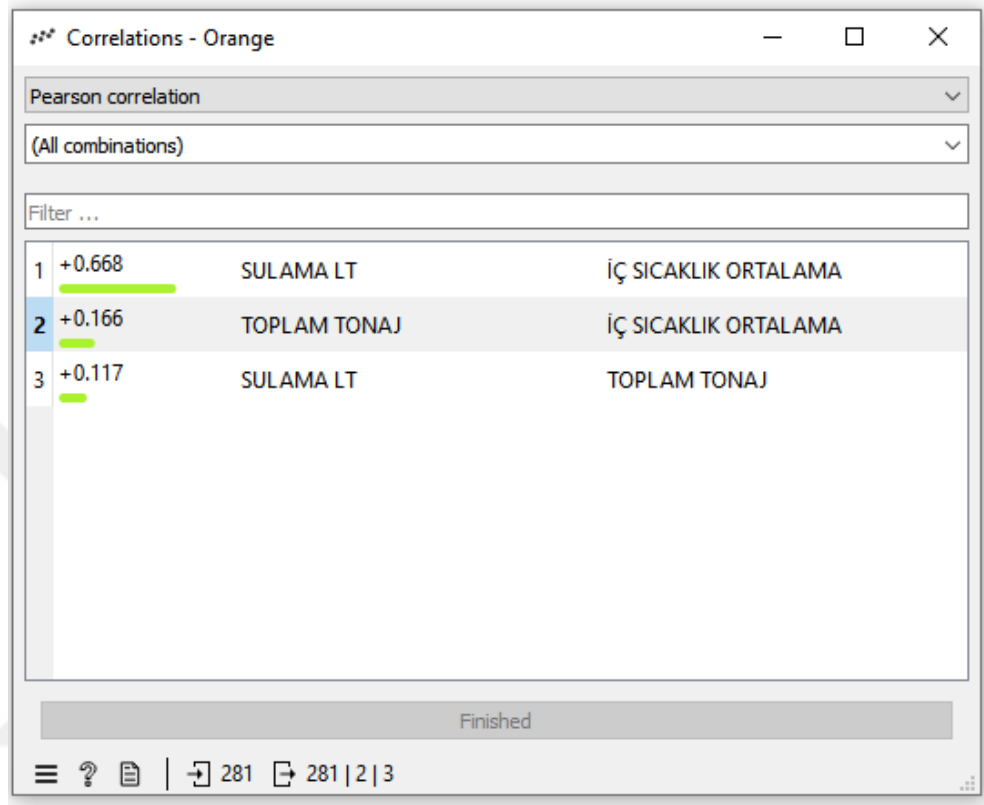
Şekil 17. Ortalama İç Sıcaklık, Sulama ve Tonaj Verilerinin Pearson Korelasyon Analizi Sonuçları



Şekil 18. 2019 Yılı Pearson Korelasyon Analiz Sonuçları

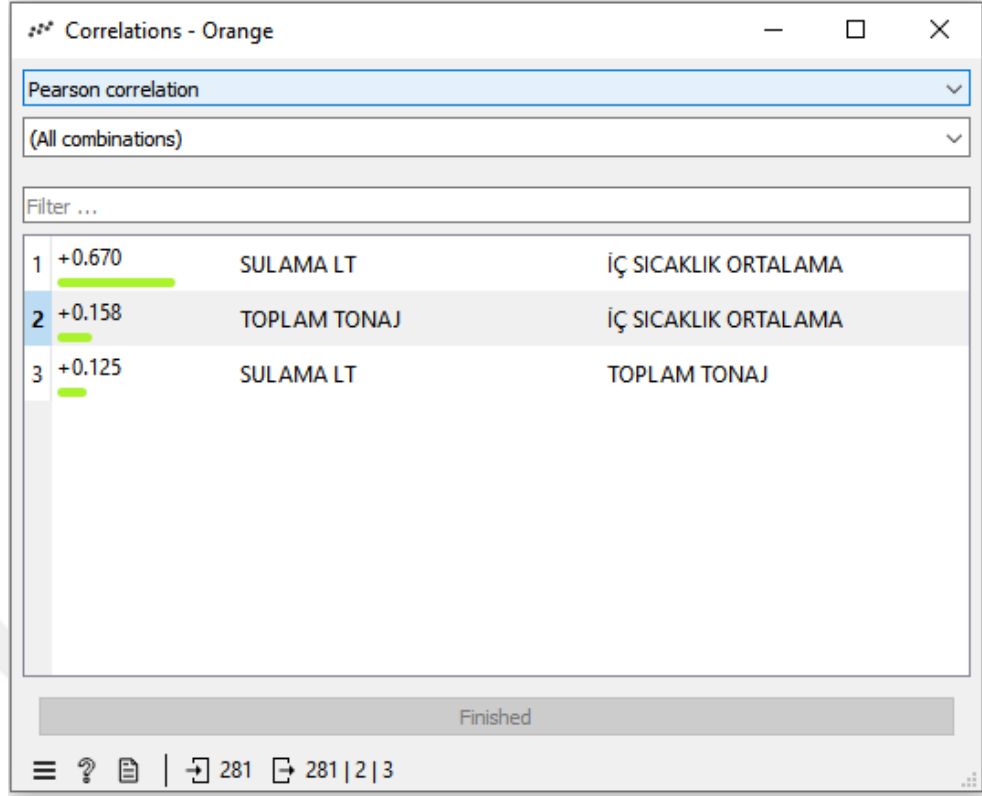
Yapılan analiz sonucu, sulama ve iç sıcaklık ortalama arasında orta düzeyde pozitif yönlü korelasyon, toplam tonaj ve iç sıcaklık arasında zayıf düzeyde pozitif

yönlü korelasyon, sulama ve toplam tonaj arasında zayıf düzeyde pozitif yönlü korelasyon olduğu görülmektedir (Şekil 18).



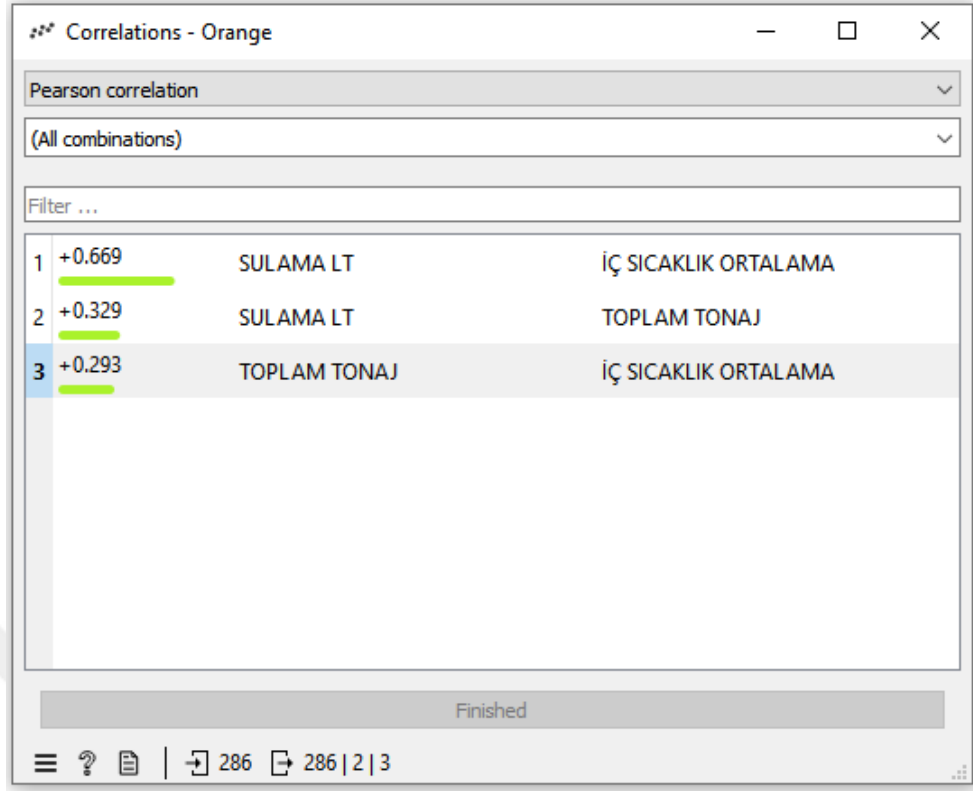
Şekil 19. 2020 Pearson Korelasyon Analiz Sonuçları

Yapılan analiz sonucu, sulama ve iç sıcaklık ortalama arasında orta düzeyde pozitif yönlü korelasyon, toplam tonaj ve iç sıcaklık arasında zayıf düzeyde pozitif yönlü korelasyon, sulama ve toplam tonaj arasında zayıf düzeyde pozitif yönlü korelasyon olduğu görülmektedir (Şekil 19).



Şekil 20. 2021 Yılı Pearson Korelesyon Analiz Sonuçları

Yapılan analiz sonucu, sulama ve iç sıcaklık ortalama arasında orta düzeyde pozitif yönlü korelasyon, toplam tonaj ve iç sıcaklık arasında zayıf düzeyde pozitif yönlü korelasyon, sulama ve toplam tonaj arasında zayıf düzeyde pozitif yönlü korelasyon olduğu görülmektedir (Şekil 20).

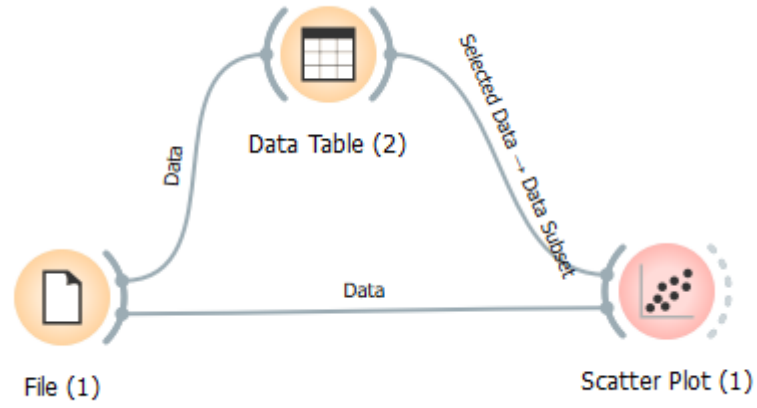


Şekil 21. 2022 Yılı Pearson Korelasyon Analiz Sonuçları

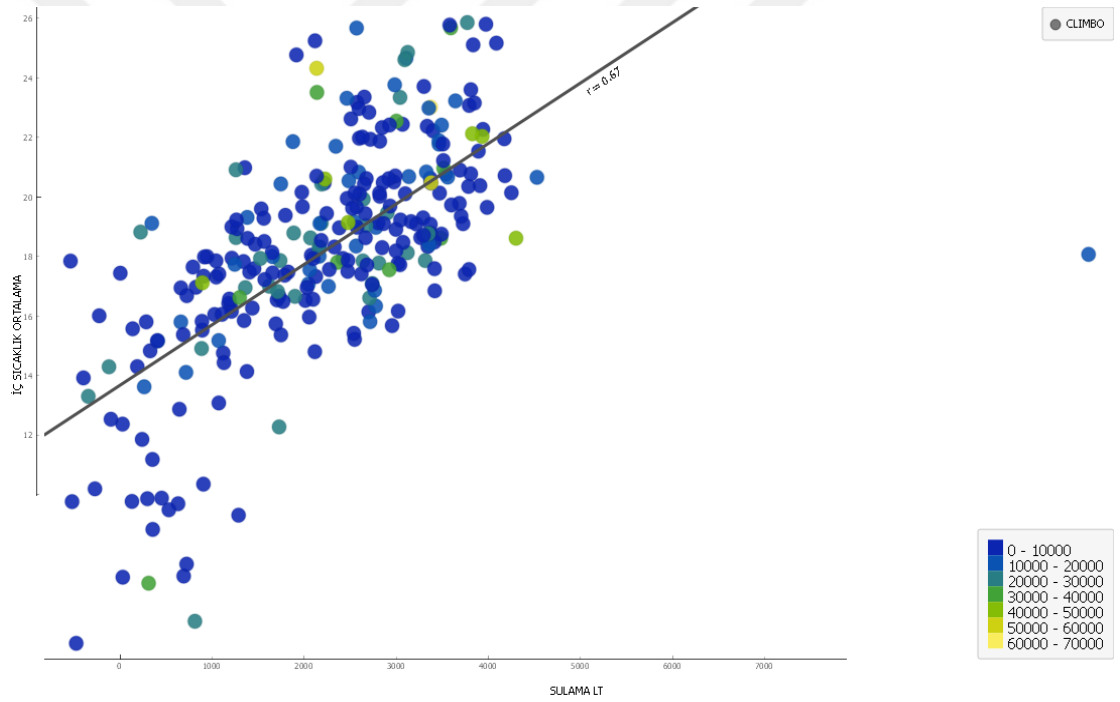
Yapılan analiz sonucu, sulama ve iç sıcaklık ortalama arasında orta düzeyde pozitif yönlü korelasyon, sulama ve toplam tonaj arasında zayıf düzeyde pozitif yönlü korelasyon, toplam tonaj ve iç sıcaklık arasında zayıf düzeyde pozitif yönlü korelasyon olduğu görülmektedir (Şekil 21).

3.5.2. Dağılım Grafiği

Dağılım grafiğinin orange machine learning veri madenciliği uygulamasında analiz edilme sürecinin ekran görüntüsü Şekil 22’de yer almaktadır. Mart 2019- Aralık 2022 yıllarına ait verilerin dağılım grafikleri şekillerde gösterilmiştir.



Şekil 22. Ortalama İç Sıcaklık ve Sulama İlişkisinin Dağılım Grafiği Ölçümlenmesi

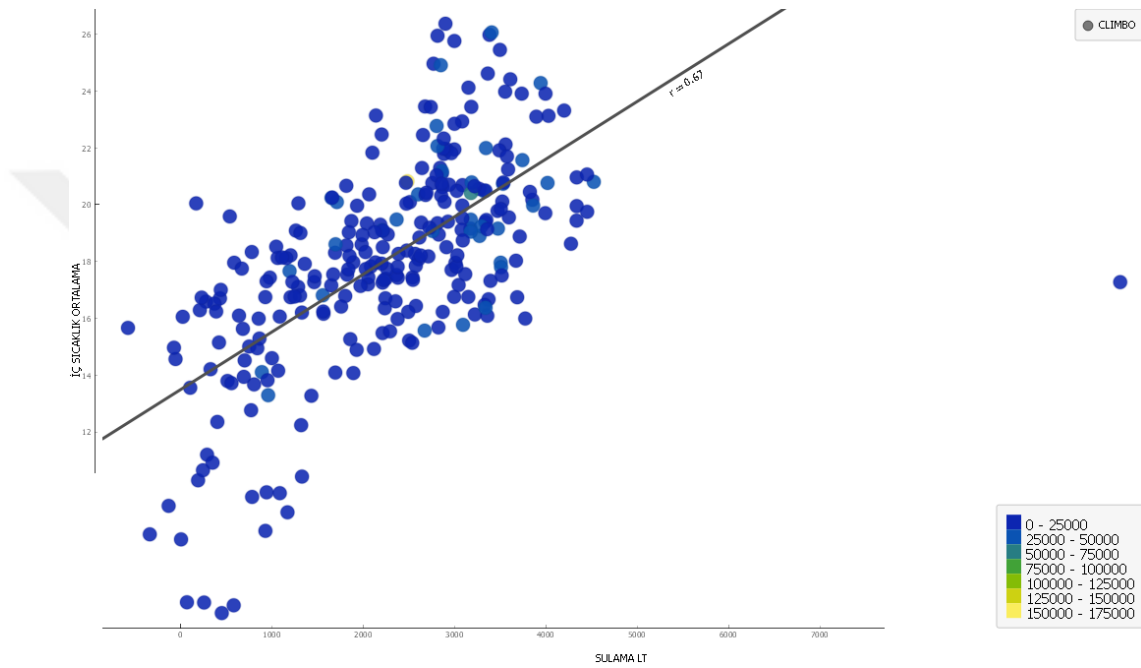


Şekil 23. 2019 Yılı Dağılım Grafiği

2019 yılı dağılım grafiğine bakıldığında sulama - iç sıcaklık ortalama değerlerinin yıllık tonaj üzerindeki etkisi incelendiğinde genel olarak regresyon doğrusu üzerine yakın olduğu görülmektedir (Şekil 23). Korelasyon katsayısı $r = 0,67$ olarak hesaplanmıştır. Noktaların lineer doğru etrafından uzaklaşıp regresyon doğrusuna yaklaşması lineer modelin o kadar kuvvetli olduğunu gösterir. Korelasyon değerinin

1'e yakın olması sulama ve ortalama iç sıcaklık değerleri arasında orta düzeyde pozitif yönlü bir ilişkinin olduğunu göstermektedir.

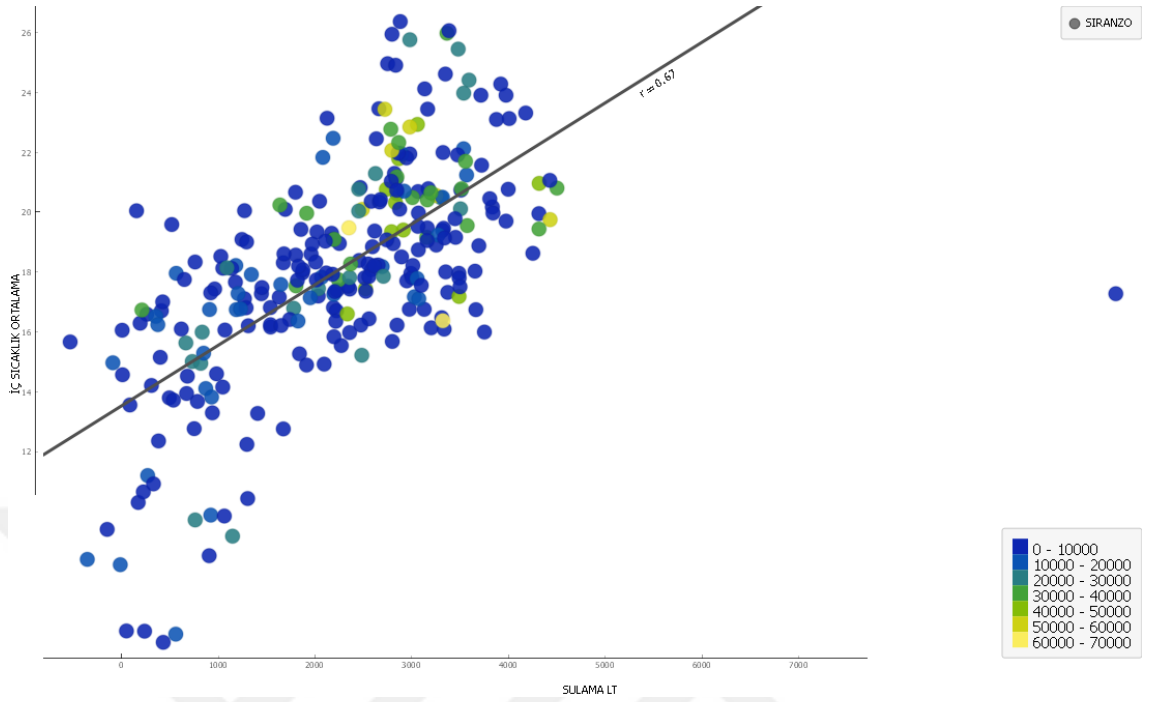
Sulama ve ortalama iç sıcaklık arasında orta düzeyde pozitif ($r=0,67$) ve anlamlı ($p<0,01$) bir ilişki bulunmuştur. Yani sulama ve iç sıcaklık değerleri orta düzeyde pozitif yönlü bir ilişki ile istatistiksel olarak artmaktadır.



Şekil 24. 2020 Yılı Dağılım Grafiği

2020 yılı dağılım grafiğine bakıldığında sulama - iç sıcaklık ortalama değerlerinin yıllık tonaj üzerindeki etkisi incelendiğinde genel olarak regresyon doğrusu üzerine yakın olduğu görülmektedir (Şekil 24). Korelasyon katsayısı $r = 0,67$ olarak hesaplanmıştır. Noktaların lineer doğru etrafından uzaklaşıp regresyon doğrusuna yaklaşması lineer modelin o kadar kuvvetli olduğunu gösterir. Korelasyon değerinin 1'e yakın olması sulama ve ortalama iç sıcaklık değerleri arasında orta düzeyde pozitif yönlü bir ilişkinin olduğunu göstermektedir.

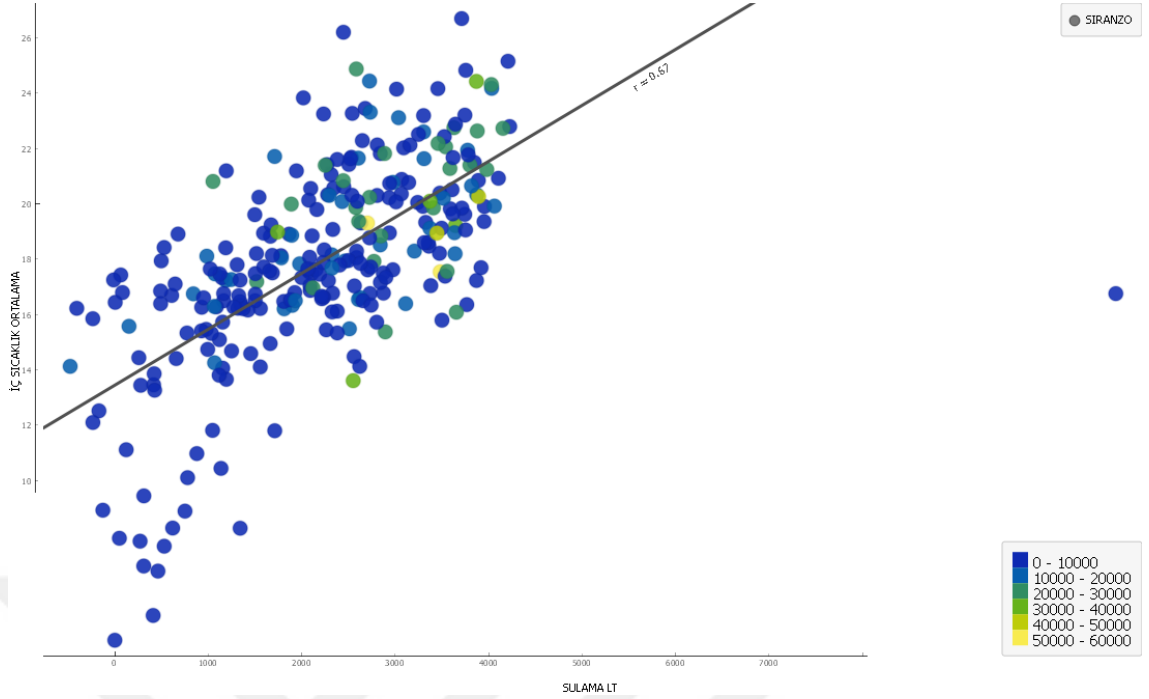
Sulama ve ortalama iç sıcaklık arasında orta düzeyde pozitif ($r=0,67$) ve anlamlı ($p<0,01$) bir ilişki bulunmuştur. Yani sulama ve iç sıcaklık değerleri orta düzeyde pozitif yönlü bir ilişki ile istatistiksel olarak artmaktadır.



Şekil 25. 2021 Yılı Dağılım Grafiği

2021 yılı dağılım grafiğine bakıldığında sulama - iç sıcaklık ortalama değerlerinin yıllık tonaj üzerindeki etkisi incelendiğinde genel olarak regresyon doğrusu üzerine yakın olduğu görülmektedir (Şekil 25). Korelasyon katsayısı $r = 0,67$ olarak hesaplanmıştır. Noktaların lineer doğru etrafından uzaklaşıp regresyon doğrusuna yaklaşması lineer modelin o kadar kuvvetli olduğunu gösterir. Korelasyon değerinin 1'e yakın olması sulama ve ortalama iç sıcaklık değerleri arasında orta düzeyde pozitif yönlü bir ilişkinin olduğunu göstermektedir.

Sulama ve ortalama iç sıcaklık arasında orta düzeyde pozitif ($r=0,67$) ve anlamlı ($p<0,01$) bir ilişki bulunmuştur. Yani sulama ve iç sıcaklık değerleri orta düzeyde pozitif yönlü bir ilişki ile istatistiksel olarak artmaktadır.



Şekil 26. 2022 Yılı Dağılım Grafiği

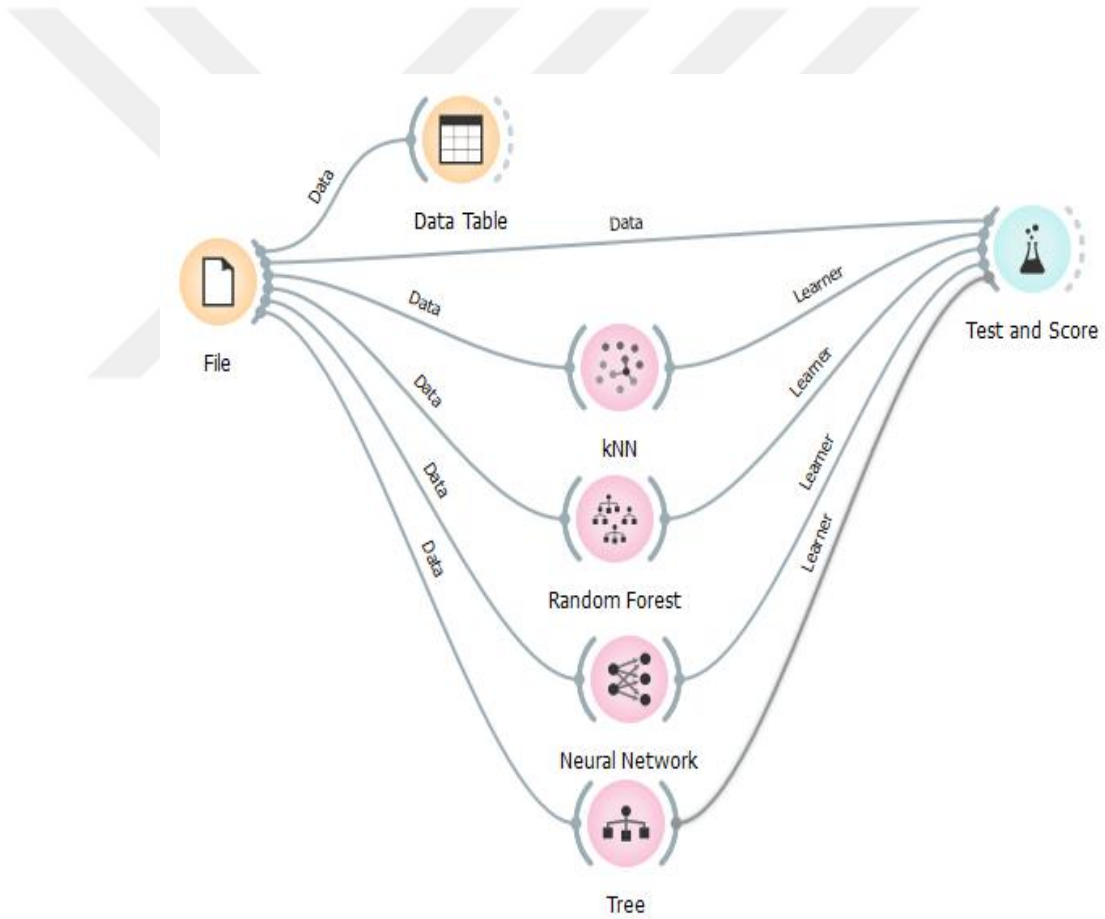
2022 yılı dağılım grafiğine bakıldığında sulama - iç sıcaklık ortalama değerlerinin yıllık tonaj üzerindeki etkisi incelendiğinde genel olarak regresyon doğrusu üzerine yakın olduğu görülmektedir (şekil 26). Korelasyon katsayısı $r = 0,67$ olarak hesaplanmıştır. Noktaların lineer doğru etrafından uzaklaşıp regresyon doğrusuna yakınlaşması lineer modelin o kadar kuvvetli olduğunu gösterir. Korelasyon değerinin 1'e yakın olması sulama ve ortalama iç sıcaklık değerleri arasında orta düzeyde pozitif yönlü bir ilişkinin olduğunu göstermektedir.

Sulama ve ortalama iç sıcaklık arasında orta düzeyde pozitif ($r=0,67$) ve anlamlı ($p<0,01$) bir ilişki bulunmuştur. Yani sulama ve iç sıcaklık değerleri orta düzeyde pozitif yönlü bir ilişki ile istatistiksel olarak artmaktadır.

3.5.3. Toplanan Verilerin Veri Madenciliği Modelleri İle İncelenmesi

Firmadan alınan 2019-2022 yılları arasındaki sulama, ortalama iç sıcaklık ve tonaj verileri Lineer Regresyon, Rastgele Orman modelleri ve tahmin edici veri madenciliği yöntemlerinden olan K-En Yakın komşu, Yapay Sinir Ağları, Karar Ağaçları Modelleri kullanılarak incelenmiştir model örneği Şekil 27'de gösterilmektedir.

Toplanan veriler öncelikle analiz için hazır hale getirildi. Algoritmaların inceleme performans parametreleri arasında R^2 , MSE, RMSE ve MAE sayılabilir. Bu parametrelerden R^2 , kullanılan modelin doğruluk oranının, belirleme katsayısını gösterir. Ölçümlenen katsayı değerinin yüksek olması tahmin ilişkisinin yüksek olmasını ve tahmin ilişkisinin anlamlı olduğunu ifade eder. MSE, RMSE ve MAE parametreleri birer hata ölçütü olması sebebiyle düşük değerler, performans ile ters orantılı olarak yüksek performans gösteren değerlere göre daha az hata oranı vermesi sebebiyle düşük değerler modelin başarılı olduğunu gösterir. MSE, RMSE ve MAE değerlerinin sıfıra eşit veya yakın olması modelin ne kadar başarılı olduğunu açıklar (Gültepe, 2019: 12-13).



Şekil 27. Ortalama İç Sıcaklık, Sulama ve Tonaj Verilerinin Veri Madenciliği Modelleri İle İncelenmesi

Tablo 1. 2019-2022 Yılı Verilerinin Belirleme Katsayısına (R^2) Göre Algoritma Değerleri

	K-En Yakın Komşu	Rastgele Orman	Yapay Sinir Ağları	Karar Ağaçları
2019	0.599	0.567	0.433	0.722
2020	0.882	0.560	0.711	0.440
2021	0.932	0.987	0.598	0.402
2022	0.973	0.686	0.431	0.569

Tablo 1’deki verilerin belirleme katsayısı yöntemine göre;

- 2019 yılı incelendiğinde $R^2 = 0.722$ ölçüm değeriyle karar ağaçları algoritması doğruya en yakın sonucu vermektedir.
- 2020 yılı incelendiğinde $R^2 = 0.882$ ölçüm değeriyle k-en yakın komşu algoritması doğruya en yakın sonucu vermektedir.
- 2021 yılı incelendiğinde $R^2 = 0.987$ ölçüm değeriyle rastgele orman algoritması doğruya en yakın sonucu vermektedir.
- 2022 yılı incelendiğinde $R^2 = 0.973$ ölçüm değeriyle karar ağaçları algoritması doğruya en yakın sonucu vermektedir.

Tablo 2. 2019-2022 Yılı Verilerinin Ortalama Kare Hatasına (MSE) Göre Algoritma Değerleri

	K-En Yakın Komşu	Rastgele Orman	Yapay Sinir Ağları	Karar Ağaçları
2019	0.175	0.135	0.813	0.187
2020	0.064	0.271	0.438	0.562
2021	0.020	0.015	0.701	0.299
2022	0.138	0.321	0.599	0.401

Tablo 2’deki verilerin ortalama kare hatası yöntemine göre;

- 2019 yılı incelendiğinde MSE= 0.135 ölçüm değeriyle rastgele orman algoritması anlamlı şekilde başarılı olmuştur.
- 2020 yılı incelendiğinde MSE= 0.064 ölçüm değeriyle k-en yakın komşu algoritması anlamlı şekilde başarılı olmuştur.
- 2021 yılı incelendiğinde MSE= 0.015 ölçüm değeriyle rastgele orman algoritması anlamlı şekilde başarılı olmuştur.
- 2022 yılı incelendiğinde MSE= 0.138 ölçüm değeriyle k-en yakın komşu algoritması anlamlı şekilde başarılı olmuştur.

Tablo 3. 2019-2022 Yılı Verilerinin Kök Ortalama Kare Hatasına (RMSE) Göre Algoritma Değerleri

	K-En Yakın Komşu	Rastgele Orman	Yapay Sinir Ağları	Karar Ağaçları
2019	0.255	0.187	0.745	0.216
2020	0.076	0.267	0.376	0.624
2021	0.061	0.014	0.605	0.395
2022	0.088	0.319	0.593	0.407

Tablo 3’deki verilerin kök ortalama kare hatası yöntemine göre;

- 2019 yılı incelendiğinde RMSE= 0.187 ölçüm değeriyle rastgele orman algoritması anlamlı şekilde başarılı olmuştur.
- 2020 yılı incelendiğinde RMSE= 0.076 ölçüm değeriyle k-en yakın komşu algoritması anlamlı şekilde başarılı olmuştur.
- 2021 yılı incelendiğinde RMSE= 0.014 ölçüm değeriyle rastgele orman algoritması anlamlı şekilde başarılı olmuştur.
- 2022 yılı incelendiğinde RMSE= 0.088 ölçüm değeriyle k-en yakın komşu algoritması anlamlı şekilde başarılı olmuştur.

Tablo 4. 2019-2022 Yılı Verilerinin Ortalama Mutlak Hatasına (MAE) Göre Algoritma Değerleri

	K-En Yakın Komşu	Rastgele Orman	Yapay Sinir Ağları	Karar Ağaçları
2019	0.757	0.243	0.015	0.117
2020	0.252	0.409	0.006	0.591
2021	0.376	0.152	0.001	0.624
2022	0.325	0.505	0.172	0.495

Tablo 4’deki verilerin ortalama mutlak hata yöntemine göre;

- 2019 yılı incelendiğinde MAE= 0.015 ölçüm değeriyle yapay sinir ağları algoritması anlamlı şekilde başarılı olmuştur.
- 2020 yılı incelendiğinde MAE = 0.006 ölçüm değeriyle yapay sinir ağları algoritması anlamlı şekilde başarılı olmuştur.
- 2021 yılı incelendiğinde MAE = 0.001 ölçüm değeriyle yapay sinir ağları algoritması anlamlı şekilde başarılı olmuştur.
- 2022 yılı incelendiğinde MAE = 0.172 ölçüm değeriyle yapay sinir ağları algoritması anlamlı şekilde başarılı olmuştur.

Yukarıda yıllara göre incelenen modellerin en başarılı olduğu algoritma değerleri tablo’5 de yer verilmiştir. 2019-2022 yıllarına ait verilerin karşılaştırmalarının yapıldığı özet tablo 5’de görüldüğü gibi;

2019 yılında belirleme katsayısı yöntemine göre karşılaştırmasında $R^2=0.722$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde yüksek ölçülmesiyle karar ağaçları algoritmasıdır. 2020 yılında belirleme katsayısı yöntemine göre karşılaştırmasında $R^2=0.882$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde yüksek ölçülmesiyle k-en yakın komşu algoritmasıdır. 2021 yılında belirleme katsayısı yöntemine göre karşılaştırmasında $R^2=0.987$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde yüksek ölçülmesiyle rastgele orman algoritmasıdır. 2022 yılında belirleme katsayısı yöntemine göre karşılaştırmasında $R^2=0.973$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde yüksek ölçülmesiyle k-en yakın komşu algoritmasıdır.

Tablo 5. 2019-2022 Yılı En Başarılı Sonucu Veren Algoritmalar

	Belirleme Katsayısı (R2)	Ortalama Kare Hatası (MSE)	Kök Ortalama Kare Hatası (RMSE)	Ortalama Mutlak Hata (MAE)
2019	R2=0.722 - Karar Ağaçları	MSE=0.135 - Rastgele Orman	RMSE=0.187 - Rastgele Orman	MAE=0.015 - Yapay Sinir Ağları
2020	R2=0.882 - K-en Yakın Komşu	MSE=0.064 - Karar Ağaçları	RMSE=0.076 - K- en Yakın Komşu	MAE=0.006 - Yapay Sinir Ağları
2021	R2=0.987 – Rastgele Orman	MSE=0.015 - Rastgele Orman	RMSE=0.014 - Rastgele Orman	MAE=0.001 - Yapay Sinir Ağları
2022	R2=0.973 - K-en Yakın Komşu	MSE=0.138 - K-en Yakın Komşu	RMSE=0.088 - K- en Yakın Komşu	MAE=0.172 - Yapay Sinir Ağları

2019 yılında ortalama kare hatası yöntemine göre karşılaştırmasında MSE=0.135 ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma rastgele orman algoritmasıdır.

2020 yılında ortalama kare hatası yöntemine göre karşılaştırmasında MSE=0.064 ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma karar ağaçları algoritmasıdır.

2021 yılında ortalama kare hatası yöntemine göre karşılaştırmasında MSE=0.015 ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma rastgele orman algoritmasıdır.

2022 yılında ortalama kare hatası yöntemine göre karşılaştırmasında MSE=0.138 ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma k-en yakın komşu algoritmasıdır.

2019 yılında kök ortalama kare hatası yöntemine göre karşılaştırmasında RMSE=0.187 ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma rastgele orman algoritmasıdır.

2020 yılında kök ortalama kare hatası yöntemine göre karşılaştırmasında RMSE=0.076 ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma k-en yakın komşu algoritmasıdır.

2021 yılında ortalama kök ortalama kare hatası yöntemine göre karşılaştırmasında $RMSE=0.014$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma rastgele orman algoritmasıdır.

2022 yılında kök ortalama kare hatası yöntemine göre karşılaştırmasında $RMSE=0.088$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma k-en yakın komşu algoritmasıdır.

2019 yılında ortalama mutlak hata yöntemine göre karşılaştırmasında $MAE=0.015$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma yapay sinir ağları algoritmasıdır.

2020 yılında ortalama mutlak hata yöntemine göre karşılaştırmasında $MAE=0.006$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma yapay sinir ağları algoritmasıdır.

2021 yılında ortalama mutlak hata yöntemine göre karşılaştırmasında $MAE=0.001$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma yapay sinir ağları algoritmasıdır.

2022 yılında ortalama mutlak hata yöntemine göre karşılaştırmasında $MAE=0.172$ ölçüm değeriyle tahmin ilişkisinin anlamlı bir şekilde başarılı algoritma yapay sinir ağları algoritmasıdır.

SONUÇ

Veri madenciliği modelleri birçok alanda olduğu gibi tarım alanıyla ilgili konularda da işletme sahiplerinin ya da yöneticilerinin karar verme süreçlerinin şekillenmesinde önemli rol oynamaktadır. Tarım alanında veri madenciliği yöntemlerinin kullanılmasına ilişkin araştırmanın ikinci bölümünde bahsedildiği üzere birçok çalışma bulunmaktadır. Bu çalışmada ise literatürdeki diğer çalışmalardan farklı olarak akıllı serada veri madenciliği yöntemlerini kullanılarak domates üretimindeki sulama suyu ve ortalama iç sıcaklık arasındaki ilişkinin alınan tonaja etkisi incelenmiştir. Ayrıca son dört yıla ait veriler içinde rastgele orman, k-en yakın komşu, yapay sinir ağları ve karar ağaçları algoritmaları kullanılarak oluşturulan modellerden doğruya en yakın sonucu veren algoritmalar belirlenmiştir.

Araştırma sonuçlarına göre akıllı seralarda veya tarım alanında üretim yapan çiftçiler veri madenciliği uygulamalarını kullanarak analizler yaptıklarında daha sağlıklı ve doğru sonuçlar elde edebilecekleri kanıtlanmıştır. Üretimde verimliliğin artırılması için çiftçilerin üretim verilerini doğru ve eksiksiz olarak kayıt altına alması ve elde olan verilerle kullanılabilir bilgiye dönüşmesi ile mümkün olacaktır.

Araştırma sonuçlarının tarımda sulama, sıcaklık ve tonaj bilgilerinin veri madenciliği yöntemleri kullanarak üretim verimliliğini arttırmak isteyen çiftçilere yol göstereceği düşünülmektedir.

Mart 2019 tarihinden Aralık 2022 tarihine kadar olan sulama, ortalama iç sıcaklık ve tonaj verilerinden hareketle R^2 parametresiyle anlamlılık düzeyi, MSE, RMSE ve MAE parametreleri ile kullanılan modellerin hata oranlarını ölçümleyerek anlamlı sonuçlar ve tarım alanında kullanılabilecek veri madenciliği yöntemleri bulunmuştur.

2019 yılı verilerinin ölçüm sonuçlarına göre belirtme katsayısı $R^2=0.722$ ölçüm değeriyle tahmin ilişkisi karar ağaçları algoritmasına anlamlı bir şekilde yüksektir. Ortalama kare hatası incelendiğinde $MSE=0.135$ ölçüm değeriyle rastgele orman algoritması en az hata değeriyle en başarılı algoritmadır. Kök ortalama kare hatası incelendiğinde $RMSE= 0.187$ ölçüm değeriyle rastgele orman algoritması en az hata değeriyle en başarılı modeldir. Ortalama mutlak hatası incelendiğinde $MAE=0.015$ ölçüm değeriyle yapay sinir ağları en az hata değeriyle en başarılı algoritmadır.

2020 yılı verilerinin ölçüm sonuçlarına göre belirtme katsayısı $R^2=0.882$ ölçüm değeriyle tahmin ilişkisi k-en yakın komşu algoritması anlamlı bir şekilde yüksektir. Ortalama kare hatası incelendiğinde $MSE=0.064$ ölçüm değeriyle karar ağaçları en az hata değeriyle en başarılı modeldir. Kök ortalama kare hatası incelendiğinde $RMSE=0.076$ ölçüm değeriyle k-en yakın komşu algoritması en az hata değeriyle en başarılı modeldir. Ortalama mutlak hatası incelendiğinde $MAE=0.006$ ölçüm değeriyle yapay sinir ağları en az hata değeriyle en başarılı algoritmadır.

2021 yılı verilerinin ölçüm sonuçlarına göre belirtme katsayısı $R^2=0.987$ ölçüm değeriyle tahmin ilişkisi rastgele orman algoritması anlamlı bir şekilde yüksektir. Ortalama kare hatası incelendiğinde $MSE=0.015$ ölçüm değeriyle rastgele orman algoritması en az hata değeriyle en başarılı modeldir. Kök ortalama kare hatası incelendiğinde $RMSE=0.014$ ölçüm değeriyle rastgele orman algoritması en az hata değeriyle en başarılı modeldir. Ortalama mutlak hatası incelendiğinde $MAE=0.001$ ölçüm değeriyle yapay sinir ağları en az hata değeriyle en başarılı algoritmadır.

2022 yılı verilerinin ölçüm sonuçlarına göre belirtme katsayısı $R^2=0.973$ ölçüm değeriyle tahmin ilişkisi k-en yakın komşu algoritması anlamlı bir şekilde yüksektir. Ortalama kare hatası incelendiğinde $MSE=0.138$ ölçüm değeriyle k-en yakın komşu algoritması en az hata değeriyle en başarılı modeldir. Kök ortalama kare hatası incelendiğinde $RMSE=0.088$ ölçüm değeriyle k-en yakın komşu algoritması en az hata değeriyle en başarılı modeldir. Ortalama mutlak hatası incelendiğinde $MAE=0.172$ ölçüm değeriyle yapay sinir ağları en az hata değeriyle en başarılı algoritmadır.

Sonuç olarak Mart 2019 Aralık 2022 tarihleri geneli incelendiğinde, belirtme katsayı değeri 2022 yılında k-en yakın komşu algoritmasıyla ölçülmüştür. 2019, 2020, 2021 ve 2022 yıllarında ise algoritmaların hata ortalamaları incelendiğinde 2019, 2020 ve algoritması, 2022 yılında ise k-en yakın komşu algoritması en az hata ile anlamlı şekilde başarılı sonuçlar vermiştir.

Araştırma sonuçlarına dayanarak tarım sektöründeki uygulamacılara yönelik öneri olarak elde ettikleri verileri düzenli ve eksiksiz olarak kayıt altına almalarıdır. Bu sayede veri madenciliği yöntemlerini kullanarak Orange Data Mining programı ile üretim verimlilik incelemelerini gerçekleştirebilirler. Analiz edilen veri örneğine göre

de yapay sinir ağıları algoritmasının en az hata ile doğru sonuç verdiği görülmektedir. Dolayısıyla yapay sinir ağılarını kullanmaları tavsiye edilebilir.

Konuyla ilgili araştırmacılara yönelik öneriler ise; bu araştırmada kullanılmayan diğer veri madenciliği yöntemleri (genetik algoritmalar, bellek temelli nedenleme ve Naive Bayes) ile modelleme yapabilirler.



KAYNAKÇA

- Akpınar, H. (2000). Veri tabanlarında bilgi keşfi ve veri madenciliği. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 29(1), 1-15.
- Alagöz, A. Öge, S. ve Ortakarpuz, M. (2014). M. Bir Kurumsal Zekâ Teknolojisi Olarak Veri Madenciliği İle Muhasebe Bilgi Sistemi İlişkisi. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 31-1(1), 1-21.
- Aydilek, İ. ve Sandıkçı, Y. (2018). Tarımsal veri analizlerinin veri madenciliği ile yapılması. *Harran Üniversitesi Mühendislik Dergisi*, 3 (2), 1-7.
- Bilekdemir, G. (2010). *Veri madenciliği tekniklerini kullanarak üretim süresi tahmini ve bir uygulama*. [Yayımlanmamış Yüksek Lisans Tezi]. Dokuz Eylül Üniversitesi.
- Bolat B. (2006). *Asansör kontrol sistemlerinin genetik algoritma ile simülasyonu*. [Yayımlanmamış Doktora Tezi]. Yıldız Teknik Üniversitesi.
- Büchner, A. G., Sarabjot S. A. ve John G. H. (1997). *Data Mining in manufacturing environments: goals, techniques and applications*, Studies in Informatics and Control.
- Çağlar, K. (2018). *Nesnelerin interneti ile akıllı sera uygulaması*. [Yayımlanmamış Yüksek Lisans Tezi], Eskişehir Osmangazi Üniversitesi.
- Çavuşoğlu, Ü. ve Kaçar, S. (2019). Anormal trafik tespiti için veri madenciliği algoritmalarının performans analizi. *Akademik Platform Mühendislik ve Bilim Dergisi*, 7(2), 205-216.
- Çerçioğlu, M. ve Şahin, H. (2016). Simav'daki seraların ısıtılmasında jeotermal enerji kullanımı olanakları. *Akademik Sosyal Bilimler Araştırmaları Dergisi*, 47(1), 459-475. doi: 10.9761/JASSS3477
- Çeşmeli, M.Ş., Bozkurt, Ö.Ç., Kalkan, A., Pençe, İ., 2015. Yönetim bilişim sistemleri bölümü öğrencilerinin yönetim ve bilişim derslerindeki başarılarının veri madenciliği yöntemleri ile incelenmesi. *Yönetim Bilişim Sistemleri Dergisi* 1(1), 36-47.

Erdem M., B. (2009). *İnternet ortamındaki tüketicilerin demografik tabanlı modellenmesi ve ürün profillerinin oluşturulması*. [Yayımlanmamış Yüksek Lisans Tezi]. Sakarya Üniversitesi.

Deliismail, İ. (2019). *Veri madenciliği teknikleri kullanılarak internetten alışveriş yapan tüketicilerin analizi*. [Yayımlanmamış Yüksek Lisans Tezi], Sakarya Üniversitesi.

Donat, İ., (22 Mart 2020). *90 milyar dolarlık tarım mucizesi*. <http://www.businessht.com.tr/haber/haber/1072193-tarimda-hollandamucizesi>.

Dunham, M. H., (2003). *Data mining Introductory and advanced topics*, Prentice Hall.

Duman, B. ve Özsoy, K. (2019). Endüstri 4.0 perspektifinde akıllı tarım. 4TH International Congress On 3d Printing (Additive Manufacturing) Technologies And Dıgital Industry Bildiri Kitabı, Antalya, 11-14.04.2019, s.540-555.

Gale, U., Tüzel, Y., Öztekin, G. B. (2014). Antalya'nın Kepez ilçesinde geleneksel sera üretiminin özellikleri. *Türkiye Tarımsal Araştırmalar Dergisi*, 1(1), 68-77.

Gültepe, Y. (2019). Makine öğrenmesi algoritmaları ile hava kirliliği tahmini üzerine karşılaştırmalı bir değerlendirme. *Avrupa Bilim ve Teknoloji Dergisi*, 1(16), 8-15. doi: 10.31590/ejosat.530347.

Güzel, B., Okatan, E. (2022). *Yapay zekanın değiştirdiği dinamikler* (1. Baskı). Eğitim Yayınevi.

Güzey, A. (2020). *Otonom insansız kara ve hava araçları ile akıllı tarım: hasat optimizasyonu üzerine bir uygulama*. [Yayımlanmamış Yüksek Lisans Tezi]. Gazi Üniversitesi.

Han, J. ve Micheline K. (2006). *Data mining concepts and techniques* (2. Baskı). Morgan Kaufmann, San Francisco.

- Jackson, J. (2002), Data mining: a conceptual overview, *Communications of the Association for Information Systems*, 8(19), 267-296. doi: 1017705/1CAIS.00819
- Jagielska, I., C. Mattheews, T. Whitfort, (1999). An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems, *Neurocomputing*, (24), 37-54.
- Kalaycı, Ş. Ve Sungur, A. (Eds.) (2018) Spss uygulamalı çok değişkenli istatistik teknikleri (9. Baskı). Dinamik Akademi Yayınları.
- Kalkaşım, A.E. ve Akkaş, M.A. Yücedağ,İ., (2019). Nesnelerin interneti ve nesnelerin internetinin tarım alanında kullanımı. III. Uluslararası Bilimsel ve Mesleki Çalışmalar Kongresi,
- Kaya, M. (2019). Ağrı'nın kalkınması için akıllı tarım (tarım 4.0) önerisi. *Akademik Bakış Dergisi*, 1(75), 130-156.
- Kavzoğlu, T., Çölkesen, İ. (2010). Karar ağaçları ile uydu görüntülerinin sınıflandırılması: kocaali örneği. *Harita Teknolojileri Elektronik Dergisi*, 2(1), 36-45.
- Keskinkılınç, K., Alemdar, T. (2013). Tarım sigortacılığı: dünya ve Türkiye'deki uygulamaların değerlendirilmesi. *Ç.Ü Fen ve Mühendislik bilimleri Dergisi* 29(3), 114-123
- Kılavuz E. ve Erdem, İ. (2019). Dünyada tarım 4.0 uygulamaları ve Türk tarımının dönüşümü. *Social Sciences (NWSASOS)*, 14(4), 133-157. doi: 10.12739/NWSA.2019.14.4.3C0189
- Kırkaya, A. (2020). Akıllı tarım teknolojileri uygulamaları. III. Uluslararası Farkındalık Konferansı, Çanakkale Onsekiz Mart Üniversitesi.
- Kolyiğit, Ö. Aşlıyan, R. ve Günel, K. (2012). Türkçe Dokümanlar İçin Yazar Tanıma. Uşak Üniversitesi Akademik Bilişim Konferansı, Uşak Üniversitesi.

- Köktürk, M.S. ve Dirsehan, T. (2012). Veri madenciliği ile pazarlama etkileşimi, Nobel Akademik Yayıncılık.
- Milivojevic, M. ve Martinovic, M. (2003). Utilization of geothermal energy in serbia. *International Geothermal Conference IGC Reykjavik, September, 10(37)*, 32-38.
- Özekes, S. (2003) Veri madenciliği modelleri ve uygulama alanları. *İstanbul Ticaret Üniversitesi Dergisi*, 2(3), 65-82.
- Özer, B., Kuş, S. Ve Yıldız, O., (2022). Veri madenciliği yöntemleri ile tarımsal veri analizi: bir akıllı tarım sistemi önerisi. *Mühendislik Bilimleri ve Tasarım Dergisi*, 10(4), 1417-1429. doi: 10.21923/jesd.1081814.
- Özbek, C.F. (2021), *Arama motoru (google) reklamlarının veri madenciliği yöntemleri ile incelenmesi: bir e-ticaret sitesi örneği*. [Yayımlanmamış Yüksek Lisans Tezi]. Onyediy Eylül Üniversitesi.
- Özkul, F.U. ve Pektekin, P. (2009). Muhasebe yolsuzluklarının tespitinde adli muhasebecinin rolü ve veri madenciliği tekniklerinin kullanılması. *Muhasebe Bilim Dünyası Dergisi*, 11(4), 57-32.
- Özmen, Ş. (2005). İş hayatı veri madenciliği ile istatistik uygulamalarını yenidenkeşfediyor.
https://www.suleozmen.com/teblig_sunumlar/3is_hayati_veri_madenciligi_istatistik.pdf
- Öztürk, H. H. (2003). İklim koşullarının sera tasarımına etkisi. *Alatarım Dergisi*, 2(2), 40-44.
- Pakdemirli, B., Birişik, N., Aslan, İ., Sönmez, B. ve Gezici, M. (2021). Türk Tarımında Dijital Teknolojilerin Kullanımı ve Tarım-Gıda Zincirinde Tarım 4.0. *Toprak Su Dergisi*, 10(1), 78-87. doi: 1021657/topraksu.898774.
- Rajagopalan, B. and Lal, U. (1999). A K-nearest neighbor simulator for daily precipitation and other weather variable, *Water Resources*, (35), 3089-3101.

- Sandıkçı, Y. (2022). *Pamuk üretim rekolte tahmini için veri madenciliği yöntemlerinin kullanılması: Şanlıurfa örneği*. [Yayımlanmamış Yüksek Lisans Tezi]. Harran Üniversitesi.
- Seyrek, İ., H. ve Ata, A., H. (2010). Veri Zarflama analizi ve veri madenciliği ile mevduat bankalarında etkinlik ölçümü. *BDDK Bankacılık ve Finansal Piyasalar Dergisi*, 4(2), 67-84.
- Somvanshi, V. K., Pandey, O. P., Agrawal, P. K., Kalanker, N. V., Prakash, M. R., & Chand, R. (2006). Modeling and prediction of rainfall using artificial neural network and Arima Techniques. *J. Ind. Geophys. Union*, 10(2), 141-151.
- Şahin, U. A. (03 Ocak 2020). Artan nüfus ihtiyaçlarını ve toprak sorunlarına alternatif çözüm: topraksız tarım. *Tarım Ve Orman Dergisi*. Erişim tarihi: 12 Eylül 2023, <http://www.turktarim.gov.tr/Haber/398/artan-nufus-ihtiyaclari-ve-toprak-sorunlarina-alternatif-cozum-topraksiz-tarim>
- Şen, B.(2013). *Serasim: 3 boyutlu akıllı sera tasarım programı*. [Yayımlanmamış Yüksek Lisans Tezi]. Çukurova Üniversitesi.
- Tellaache, A., Burgosartızzu, X. P., Pajares, G., ve Ribeiro, A. (2007). A vision-based hybrid classifier for weeds detection in precision agriculture through the bayesian and fuzzy k-means paradigms. in innovations in hybrid intelligent systems. Springer, Berlin, Heidelberg.
- Thuraisingham, B., (1999), Data mining technologies, techniques, tools and trends CRC Press LLC, USA, s.1.DUNHAO CROWS CORPORATION), Introduction to Data Mining and Knowledge Discovery.
- Thuraisingham, B (2003). Web data mining and applications in business intelligence and counter-terrorism. CRC Press.
- Urtubia, A., Pérez-Correa, J. R., Soto, A., ve Pszczolkowski, P. (2007). Using data mining techniques to predict industrial wine problem fermentations. *Food Control*, 18(12).

- Uzun, Y. Bilban, M. & Arıkan, H. (2018). Tarım ve kırsal kalkınmada yapay zeka kullanımı. 6. *Uluslararası KOP bölgesel kalkınma sempozyumu* (26-27 Ekim). Konya.
- Veenadhari, S. (2007), “Crop productivity mapping based on decision tree and Bayesian classification”. Unpublished M.Tech Thesis submitted to Makhanlal Chaturvedi National University of Journalism and Communication, Bhopal
- Yalçın, E. C. (2008). *Proses iyileştirme çalışmalarında veri madenciliği yaklaşımının kullanılması üzerine bir çalışma*. [Yayınlanmamış Yüksek Lisans Tezi]. Dokuz Eylül Üniversitesi.
- Yıldız, M. (2023). *Smart agriculture applications; asmart greenhouse design*. [Yayınlanmamış Yüksek Lisans Tezi]. Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi.
- Yüksel, T. (2011). *Biyogaz, güneş ve toprak enerjisi kaynaklı sera ısıtmasının araştırılması*. (Yayın Nu. 99219202). [Doktora Tezi, Fırat Üniversitesi].

ÖZGEÇMİŞ

Kişisel Bilgiler:

Adı ve Soyadı : Kevser YEŞİLÇİMEN

Eğitim Durumu:

Lisans Eğitimi : Bilecik Şeyh Edebali Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri.

Yüksek Lisans Eğitimi : Burdur Mehmet Akif Ersoy Üniversitesi, Sosyal Bilimler Enstitüsü, Yönetim Bilişim Sistemleri Anabilim Dalı, Tezli Yüksek Lisans.

Yabancı Dil ve Düzeyi : İngilizce, Orta

İş Deneyimi : 2020-2021 Sandıklı Tarım Gıda ve Bitki LTD. ŞTİ., Ön Muhasebe.
2022-2023 Mergen Gıda SAN. VE TİC. LTD. ŞTİ., Kalite Kontrol Sorumlusu.

Bilimsel Yayınlar ve Çalışmalar: