**T.C.**
**HARRAN ÜNİVERSİTESİ**
**FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS TEZİ**

# MAKİNE ÖĞRENMESİ İLE NOHUTTA VERİM VE TÜR TAHMİNİ

**Fırat KESKİNBIÇAK**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**ŞANLIURFA**
**2023**

# İÇİNDEKİLER

**Sayfa No**


ÖZET ..... i
ABSTRACT ..... ii
TEŞEKKÜR ..... iii
ŞEKİLLER DİZİNİ ..... iv
ÇİZELGELER DİZİNİ ..... v
SİMGELER VE KISALTMALAR DİZİNİ ..... vi
1. GİRİŞ…… ..... 1
1.1. Tarım ve Önemi ..... 2
1.2. Tarımda Bitki ..... 3
1.3. Baklagiller ..... 3
1.3.1. Nohut ..... 4
1.3.2. Nohut yetiştiriciliği ..... 7
1.3.3. Verim ..... 9
2. ÖNCEKİ ÇALIŞMALAR ..... 11
3. MATERYAL ve YÖNTEM ..... 26
3.1. Materyal ..... 26
3.1.1. Veri seti özellikleri ..... 26
3.2. Yöntem ..... 28
3.2.1. Verinin hazırlanması ..... 28
3.2.2. Sentetik veri üretme ..... 30
3.2.3. Normalizasyon ..... 31
3.2.3.1. D min. max. normalizasyonu ..... 32
3.2.4. Veri setlerinin eğitim ve test için ayrıştırılması ..... 32
3.2.4.1. Çapraz doğrulama 5 (ÇD-5) ..... 33
3.2.4.2. %50 eğitim %50 test (%50-%50) ..... 34
3.2.5. Makine öğrenmesi ..... 34
3.2.5.1. Karar ağaçları (KA) ..... 35
3.2.5.2. Destek vektör makineleri (DVM) ..... 37
3.2.5.3. K en yakın komşu (KNN) ..... 39
3.2.6. Performans metrikleri ..... 41
3.2.7. Başarı ölçütü ..... 41
4. ARAŞTIRMA BULGULARI ve TARTIŞMA ..... 43
4.1. Verilerin Grafiksel Gösterimi ..... 43
4.2. Çalışmada Sınıflandırma Sonucu Elde Edilen Sonuçların Karşılaştırılması ..... 45
4.3. Çalışmada Regresyon Sonucu Elde Edilen Sonuçların Karşılaştırılması ..... 66
5. SONUÇLAR ve ÖNERİLER ..... 70
KAYNAKLAR ..... 71
ÖZGEÇMİŞ ..... **Hata! Yer işareti tanımlanmamış.**

# ÖZET

**Yüksek Lisans Tezi**

## MAKİNE ÖĞRENMESİ İLE NOHUTTA VERİM VE TÜR TAHMİNİ

**Fırat KESKİNBIÇAK**

**Harran Üniversitesi**
**Fen Bilimleri Enstitüsü**
**Elektrik-Elektronik Mühendisliği Anabilim Dalı**

**Danışman: Dr. Öğretim Üyesi Kerim KARADAĞ**
**Yıl: 2023, Sayfa: 76**

Tarım sektöründe üretimin artışı; ekilebilen tarım alanlarının genişletilmesi bakımın, sulamanın uygun yöntemlerle zamanında yapılması ve uygun tohum türlerinin secimi gibi parametrelere bağlıdır. Artmakta olan insan nüfusuna sağlıklı ve yeterli gıda üretilebilmesi için tarım alanları yetersiz kalmaktadır. Sağlıklı ve yeterli gıda ihtiyacının karşılanabilmesi için birim alandan sağlıklı ve çok ürün elde etmenin gerektiği anlaşılmıştır. Verimliği artırmanın ön koşulu ise uygun tohum türlerin seçilmesi olduğu görülmüştür. Yapılan bu çalışma iki aşamada gerçekleşmiştir. Birinci aşamada nohut bitkisine ait özelliklerden makine öğrenme yöntemleri ile sınıflandırma işlemi yapılarak nohut tür tahmini yapıldı ikinci aşamada türlere ait verim değerleri regresyon yöntemi ile tahmin edilmiştir. Verilerin test ve eğitim için ayrıştırılmasında çapraz doğrulama 5 ve %50 eğitim %50 test şeklinde gerçekleşmiştir. Makine öğrenmesi sınıflandırma yöntemlerinden karar ağaçları (KA), destek vektör makineleri (DVM) ve k en yakın komşuluk (KNN) yöntemleri uygun görülmüş ve kullanılmıştır. Başarı ölçütü olarak doğruluk kullanılmıştır. Ham veri seti, ham + sentetik veri seti ve sadece sentetik veri setinden oluşan üç sınıfta işlem yapılmıştır. Elde edilen sonuçların başarı oranları karşılaştırıldığında yalnızca ham veriler kullanılarak yapılan sınıflandırmada en yüksek başarı oranı %90.6 doğruluk oranı ile DVM olmuştur. Aynı şekilde ham veriler ile birlikte oluşturulan sentetik verilerin kullanıldığı veri setinin sınıflandırma başarı oranı en yüksek %100 doğruluk oranı ile KA olmuştur. Sadece sentetik verilerden oluşan veri seti kullanıldığında sınıflandırmada en yüksek başarı oranı %95.4 doğruluk oranı ile KNN olmuştur. Başarı performanslarının yüksek olması makine öğrenme yöntemlerinin bu alanda kullanılabileceğini göstermiştir.

**ANAHTAR KELİMELER:** Nohut, verim, Makine Öğrenmesi, Sınıflandırma, Doğruluk

## ABSTRACT

**MSc Thesis**

**CHICKPEA YIELD AND SPECIES PREDICTION WITH MACHINE LEARNING**

**Fırat KESKİNBIÇAK**

**Harran University**
**Graduate School of Natural and Applied Sciences**
**Department of Electrical and Electronics Engineering**

**Supervisor: Dr. Instructor Member Kerim KARADAĞ**
**Year: 2023, Page: 76**

Increase in production in the agricultural sector; The expansion of cultivated agricultural areas depends on parameters such as timely maintenance, irrigation with appropriate methods and selection of appropriate seed types. Agricultural areas are insufficient to produce healthy and sufficient food for the increasing human population. It has been understood that in order to meet the need for healthy and sufficient food, it is necessary to obtain healthy and many products from the unit area. It has been seen that the prerequisite for increasing productivity is the selection of suitable seed species. This study was carried out in two stages. In the first stage, chickpea species were estimated by classifying the characteristics of the chickpea plant with machine learning methods, and in the second stage, the yield values of the species were estimated by regression method. In the separation of the data for testing and training, cross validation was 5 and 50% training was 50% testing. Decision trees (CA), support vector machines (SVM) and k nearest neighbor (KNN) methods from machine learning classification methods were found suitable and used. Accuracy was used as a measure of success. Three classes of raw data set, raw + synthetic data set and only synthetic data set were processed. When the success rates of the results obtained were compared, the highest success rate in the classification using only raw data was SVM with an accuracy rate of 90.6%. Likewise, the classification success rate of the dataset, in which synthetic data created together with the raw data, was used, was KA with the highest 100% accuracy rate. When the dataset consisting of only synthetic data was used, the highest success rate in classification was KNN with an accuracy rate of 95.4%. The high success performances have shown that machine learning methods can be used in this field.



**KEY WORDS:** Chickpea, Yield, Machine learning, Classication, Accuracy

# TEŞEKKÜR

Tez konusunun seçiminde ve uygulamasında yardımlarını benden esirgemeyen danışmanım sayın Dr. Öğretim Üyesi Kerim KARADAĞ hocama, çalışmada kullanılan veri setinin elde edilmesinde her türlü desteğini sağlayan Doç. Dr. Abdullah Kahriman hocamıza ve her zaman yanımda olan desteğini esirgemeyen ailem ve arkadaşlarıma teşekkür ederim.

# ŞEKİLLER DİZİNİ

**Sayfa No**

Şekil 1.1. Nohut ..... 5
Şekil 1.2. Dünyadaki nohudun ekim alanı ve üretimi ..... 5
Şekil 1.3. Türkiye'deki nohudun ekim alanı ve üretimi ..... 6
Şekil 3.1. Çalışma akış diyagramı ..... 30
Şekil 3.2. SMOTE algoritmasının çalışma sistemi ..... 31
Şekil 3.3. Çapraz doğrulama 5 uygulanması ..... 33
Şekil 3.4..Karar Ağacı Yapısı ..... 35
Şekil 3.5. A) Doğrusal ayrılabilen veriler B) Doğrusal olarak ayrılmayan veriler ..... 37
Şekil 3.6. Doğrusal olarak ayrılabilen veri setlerinde hiper-düzlemin belirlenmesi ..... 38
Şekil 3.7. KNN Çalışma şeması ..... 40
Şekil 4.1. C. ech. ve C. reti. grafiksel gösterimi ..... 45
Şekil 4.2. Ham veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 46
Şekil 4.3. Ham veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 47
Şekil 4.4. Ham veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 48
Şekil 4.5. Ham veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 49
Şekil 4.6. Ham veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 50
Şekil 4.7. Ham veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 51
Şekil 4.8. Ham + sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 53
Şekil 4.9. Ham + sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminle. ..... 54
Şekil 4.10. Ham + sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler 55
Şekil 4.11. Ham + sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 56
Şekil 4.12. Ham + sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler 57
Şekil 4.13. Ham + sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 58
Şekil 4.14. Sadece sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler .. 60
Şekil 4.15. Sadece sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 61
Şekil 4.16. Sadece sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler 62
Şekil 4.17. Sadece sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 63
Şekil 4.18. Sadece sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler 64
Şekil 4.19. Sadece sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 65

# ÇİZELGELER DİZİNİ

**Sayfa No**


Çizelge 3.1. Veri setine ait özellikler ..... 28
Çizelge 4.1. Ham veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 46
Çizelge 4.2. Ham veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 47
Çizelge 4.3. Ham veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler..... 48
Çizelge 4.4. Ham veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 49
Çizelge 4.5. Ham veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 50
Çizelge 4.6. Ham veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler ..... 51
Çizelge 4.7. Ham veri (%) doğruluk oranları..... 52
Çizelge 4.8. Ham + sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler 53
Çizelge 4.9. Ham + sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 54
Çizelge 4.10. Ham + sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler..... 55
Çizelge 4.11. Ham + sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 56
Çizelge 4.12. Ham + sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 57
Çizelge 4.13. Ham + sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 58
Çizelge 4.14. Ham + sentetik veri (%) doğruluk oranları ..... 59
Çizelge 4.15. Sadece sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler ..... 60
Çizelge 4.16. Sadece sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 61
Çizelge 4.17. Sadece sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler..... 62
Çizelge 4.18. Sadece sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 63
Çizelge 4.19. Sadece sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler..... 64
Çizelge 4.20. Sadece sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler..... 65
Çizelge 4.21. Sadece sentetik veri (%) doğruluk oranları..... 66
Çizelge 4.22. Ham veri rmse başarı oranları..... 67
Çizelge 4.23. Ham veri+ sentetik veri rmse başarı oranları ..... 67
Çizelge 4.24. Yalnızca sentetik veri rmse başarı oranları ..... 68

# SİMGELER ve KISALTMALAR DİZİNİ

| | |
|---|---|
| % | Yüzde |
| %50-%50 | %50 eğitim %50 test |
| C. ech. | *Cicer echinospermum* |
| C. reti. | *Cicer reticulatum* |
| cm | Santimetre |
| $cm^2$ | Santimetrekare |
| $cm^3$ | Santimetreküp |
| ÇD-5 | Çapraz doğruluk 5 |
| DVM | Destek vektör makinaları |
| g | Gram |
| $g/cm^3$ | Gram/santimetre küp |
| g/parsel | Gram/parsel |
| KA | Karar ağaçları |
| kg/da | Kilogram/dekar |
| KNN | K en yakın komşuluk |
| Max | Maximum, |
| Min | Minimum |
| SMOTE | synthetic minority oversampling technique |

# 1. GİRİŞ

Canlılar sağlıklı büyüyen ve gelişen, üretken bir birey olarak uzun süre yaşaması ve yaşamını devam ettirebilmesi için beslenmesi gerekir. Besin öğeleri yeterli düzeyde alınamazsa enerji ve vücut dokuları oluşmaz. Aynı şekilde gereğinden fazla yemek tüketmek, sağlıksız gıdalar tüketmek ve yiyecekleri fazla pişirmek gibi etkenler düzensiz ve sağlıksız beslenme olarak isimlendirilir (Ünsal, 2019). İnsanların, hayvanların ve küçük canlıların devamını sağlayan besin maddeleri tarım ile sağlanmakta tarımın insanlar ve canlılar için ne kadar önemli olduğu anlaşılmaktadır. Tarım sektöründe üretimin artışı; ekilebilen tarım alanlarının genişletilmesi, sulama faaliyetlerinin zamanında ve uygun yöntemlerle yapılması, zararlılar ile mücadelenin yapılması ve uygun tohum türlerinin secimi gibi parametrelere bağlıdır. Artmakta olan insan nüfusuna sağlıklı ve yeterli gıda üretilebilmesi için tarım alanları yetersiz kalmaktadır. Sağlıklı ve yeterli gıda ihtiyacının karşılanabilmesi için birim alandan sağlıklı ve çok ürün elde etmenin gerektiği anlaşılmıştır. Verimliği artırmanın ön koşulu ise uygun tohum türlerin seçilmesi olduğu görülmüştür. Gelişen teknoloji ve artan dünya nüfusu ile veri yığınları meydana gelmiştir. Düzgün veri seçimi ile kayıt edilen veriler tahmin ve karar vermede kullanılmaktadır. Tarımın en önemli etkenlerinden olan verimli türlerin tespiti ile artmakta olan nüfus oranlarına yeterli ve sağlıklı ürün elde edilme hedeflenmiştir.

Yapılan tez çalışmasında, nohut türlerinden elde edilen özelliklere ait veri setleri kullanılmıştır. Nohut türlerine ait durum tespiti iki evreden oluşmuştur. Birinci evrede özelliklerin seçimi ve normalizasyon işlemi, ikinci evrede ise verilere ait özellikler kullanılarak sınıflandırma ve dekar başına elde edilen verimlilik tahmini regresyon yöntemleri ile yapılmıştır. Sınıflandırma ve Regresyon yöntemleri olarak, Karar Ağaçları (KA), Destek Vektör Makinaları (DVM) ve K en yakın komşuluk (KNN) yöntemleri kullanılmıştır. Çalışma iki aşamada ele alınmış. Birinci aşamada, nohut ile ilgili tür tespiti yapılmış. İkinci aşamada ise nohut türlerine ait dekar başı

verimlilik tahminde bulunulmuştur. Sonuç olarak bitkiler ile ilgili olarak tür veya verimlilik hesaplamalarının yapılmasında geleneksel yöntemlerin yerine bilgisayar teknikleri kullanılarak başarılı sonuçlar elde edilip edilemeyeceği araştırılmıştır. Daha özel anlamda ise nohuda ait özellikler kullanılarak tür ve verimlilik tahmini yapılmaya çalışılmıştır.

Giriş bölümünde, konunun daha iyi anlaşılması için kısaca, Tarım ve önemi, baklagiller, nohut, verimlilik, tezin hedefi ve çalışmada kullanılan veriler hakkında bilgiler verilmiştir.

## 1.1. Tarım ve Önemi

İnsanın ve hayvanların beslenmesi ve devamı için gerekli bitkisel ürünler elde edilme işlemine tarım denir (Hekimoğlu ve Altındeğer, 2006). Ürün elde etmek, ekonomik kazanç sağlamak amacıyla toprağın işlenmesi ekilecek bitki için toprağın hazırlanması, tohumun ekilmesi veya fide fidanların dikilmesi, sulama, gübreleme ve ilaçlama gibi bakım işlemlerinin yapılması hasat olgunluğuna gelen ürünün hasat edilmesi yapılan bu işlemlerin hepsine tarımsal faaliyet denir (Özçağlar, 2000). Tarım alanı ise tarımsal faaliyetlerin yapıldığı alanlara denilmektedir.

Günümüzde dünyanın en önemli stratejik sektörlerinden biriside tarımdır. İnsanların gıda ihtiyacından fazla ürün üretmesi ve pazarlaması, bazı ürünlerin sanayide ham madde olarak kullanılması ile tarıma ekonomik boyut kazandırılmıştır. Gıda üretimi ve ekonomik kazanç için yapılan tarım Türkiye gibi ekonomide payının büyük olduğu ülkeler için oldukça önemlidir (Bayar, 2004). Teknolojinin gelişmesi ile ulaşım ve iletişimde kolaylıklar oluşmuş artan insan nüfusunun beslenme alışkanlıklarında değişmelerin oluşacağını göstermiştir. İnsanlara mevcut gıdaların yetersiz gelmesi sağlıksız gıdalara yönlendirecektir. Gelişmiş ülkelerde ise sağlıklı olan doğal gıdalara dönüş olacağı beklenmektedir (Hekimoğlu ve Altındeğer, 2006). Türkiye de 1949-2000 yılları arasında genel arazi varlığı içerisinde, çayır ve mera alanlarının düştüğü, tarım alanlarının payının %20'den %34'e yükseldiği tarımın geçmişten günümüze öneminin her geçen gün artarak devam ettiği görülmüştür (Aydın, 2012).

## 1.2. Tarımda Bitki

Bitkiler bağlı oldukları topraktan su ve besin maddelerini alarak metabolizmasında elverişli hale getirerek insanlar için gerekli olan temel besin ögelerine dönüştürürler. Bitkilerin çoğu çok hücreli canlılardır. Yeşil bitkiler güneş ışığı ile fotosentez yoluyla, enerjilerini elde ederler. Bitkiler eşeyli ve eşeysiz üreme olarak ikiye ayrılır. Eşeysiz olarak üreyen bitki sayısı, eşeyli olarak üreyen bitkilerden sayıca daha fazladır. Eşeyli üreme ile karakteristik özellikleri yani nesillerin değişimi görülür. Dünyada büyük bir popülasyona sahip olan bitkiler yaklaşık 320 000 türe sahiptir. Bu türlerin çoğunluğunu oluşturan 250 000 ila 300 000 kadarı tohum üretir (IUCN, 2010). Yeşil bitkiler dünyadaki oksijenin önemli bir kısmını üreten canlılar olup dünyadaki canlıların çoğunluğunu oluştururlar (Field ve ark., 1998). Tahıl, baklagiller ve sebze üreten bitkileri ile meyve ağaçları insan beslenmesinde geçmişten günümüze devam etmiştir. Tarımın önemli bir kısmını oluşturan baklagiller insanların ve hayvan beslenmesinde, kurağa dayanıklı bitki olması ve toprağa havadaki serbest azotu bağlaması ile önemini daha da artmıştır.

## 1.3. Baklagiller

Tarımın önemli bir payını oluşturan baklagiller insanların beslenmesinde geçmişten günümüze her zaman yer almıştır. Baklagiller yüksek oranda ham protein içermeleri, vitamin ve mineral maddeleri açısından zengin oluşları tahıllara göre sap ve samanının yüksek düzeyde olması hem insan beslenmesinde hem de hayvan besiciliğinde tercih edilmelerini sağlamış kullanımını yaygınlaştırmıştır (Eyidogan ve Öz, 2007).

Baklagiller (Fabaceae), Fabales takımında yer almaktadır fabales takımı otsu bitkilerin daha çok olduğu ağaç ve çalı türlerini de içeren bir familyadır. 10 000 yakın tür içeren bu familya fasulye, nohut, mercimek ve bezelye gibi insan ve hayvanların beslenmesinde kullanılan türler bu familyadandır (wikipedia, 2022). Bitkisel olarak beslenmede ilk sırayı tahıllar alırken ikinci sırada yemeklik tane baklagiller almaktadır (Sepetoğlu, 1996). Yemeklik tane baklagiller uzun yıldır önce antik dönemlerde insanlar tarafından beslenmede kullanıldığı, nohut ve mercimeğin insan beslenmesinde çok eski olan 8 ila 10 bin yıl öncesine dayandığı kanıtlarla

ortaya çıkmıştır. Günümüzde de insan beslenmesinde önemli bir yer tutmuşlardır. Baklagillerin köklerinde yaşamını sürdüren Rhizobium bakterileri havada serbest halde bulunan azotu toprağa bağlama özellikleri ile köklerinin dağıldığı toprak katmanlarını organik azotça zenginleştirirler. Diğer canlılar tarafından direkt yararlanılamayan havanın serbest azotunu toprağa bağlama özellikleri ile yetiştirildikleri toprağa olumlu etkilerde bulunmakta, çevrecilik ve sürdürülebilir tarımın gelişmesi ile bu bitkilerin önemini daha da artırmıştır. Baklagiller ile topağa bağlanan azot miktarı iklim, çeşit ve çevre şartlarına göre değişmektedir (Adak ve ark., 2015). Yemeklik tane baklagiller yapısında yüksek protein bulunduran, ülkelerin beslenmesinde daha ekonomik ve etkin olan bitki grubunu oluşturmaktadırlar (Çakmak, 2019). Baklagiller dünyada insan beslenmesinin %22 bitkisel proteinlerini, %7 karbonhidratlarını oluşturmaktadır. Hayvan beslenmesinde ise %38 protein ve %5 karbonhidrat ile sağlamaktadır. Baklagillerin mutlak gerekli aminoasitler yüksek oranda içermeleri, yağ oranı düşük, kolesterol içermeyen, vitaminlerin yüksek olması gibi özelliklerinin belirgin olarak ortaya çıkması tercih edilmelerinde önemli etkenler olmuştur. Baklagiller, dünyada önemli bir kısmı olan yaklaşık iki milyar insanın protein kaynağı haline gelmiştir (Adak ve ark., 2015). Baklagillerin büyük çoğunluğunu nohut oluşturmaktadır.

### 1.3.1. Nohut

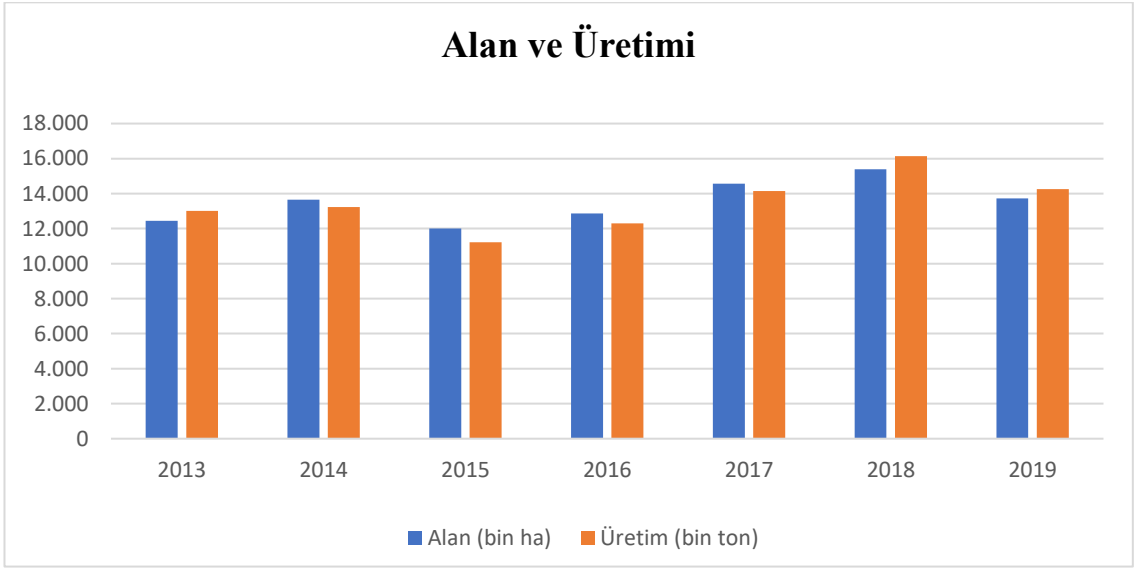
Nohut, Plantae âleminin kapalı tohumlular bölümden olup İki çeneklileri sınıfının baklagiller (Fabaceae) familyasına ait Cicer cinsinden bir baklagil türüdür. Tohum ile çoğalan bir yıllık olan nohut tohumunun rengi bej, koyu ve siyah olabiliyor (wikipedia, 2022). Baklagiller içerisinde kurağa en dayanıklı bitki nohut olduğundan, kurak alanlarda ekim nöbetine girerek nadas alanlarını önemli ölçüde azaltmaktadır (Burucu, 2017). Su kıtlığına ve kurağa dayanıklı bir bitkidir. Baklagiller, kazık köklüdürler ve köklerindeki ortak yaşama giren bakteriler ile havanın serbest azotunu toprağa bağlamaları ve hasadından sonra kalan bitki artıkları toprağa karışarak verimliliği üzerinde olumlu etkilere sahiptir (Karakullukçu, 2008). Ayrıca gübrenin az kullanılması ile girdi maliyetlerini azaltmakta ve çevre kirlenmesine engel olur. Yapılan araştırmalara göre çeşit ve çevre koşullarına göre

ortalama toprağa dekar başına 8 kg kadar azot bağlama özellikleri vardır (Aydoğan ve ark., 2011).



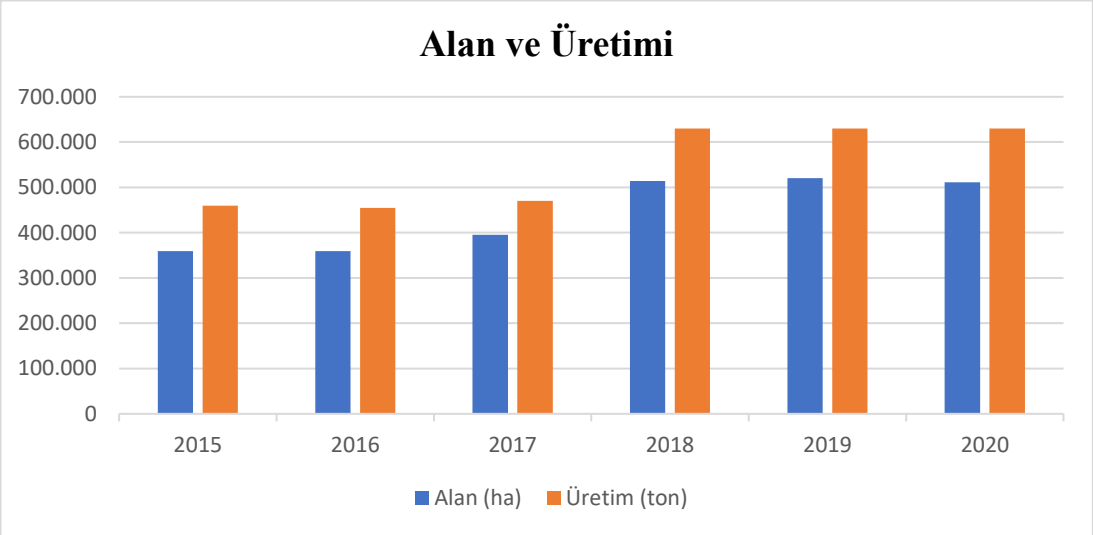
Şekil 1.1 Nohut (Olgunlar Tohum Nohut, 2022)

Gıda ve Tarım Örgütü 2019 yılı verisine göre dünyada nohudun 13.72 milyon hektar alandan 14.25 milyon ton üretilmiştir. Dünya verim ortalaması dekarda 104 kg olarak gerçekleşmiştir. Ülkemiz ise 2019 yılına ait FAO verilerinden alınan bilgilere göre dünya sıralamasında nohut ekimi alanında 4. Sırada olurken, üretimde 2. Sırada yer almaktadır (Tarım ve Orman Bakanlığı, 2021).




Şekil 1.2 Dünyadaki nohudun ekim alanı ve üretimi

TÜİK verilerine göre 2019 yılında Türkiye dünya sıralamasında ithalatta 5. ihracatta ise 3. sırada yer almaktadır. Ülkemiz 2020 yılı itibarıyla ihracatçı konumuna geçmiştir. Genel olarak pazar talebi iri taneli doğrultusunda olmuştur (Tarım ve Orman Bakanlığı, 2021).




Şekil 1.3 Türkiye'deki nohudun ekim alanı ve üretimi

Nohuttun arkeolojik kazılarda çıkartılan en eski kalıntıları Burdur yakınlarında olduğu bildirilmişti (Karaman ve ark., 2022). Nohudun dünyadaki yayılışı Akdeniz'den batıya doğru Hindistan tarafından ise doğuya doğru yayılmış olması gerektiği. Türkiye'de nohuttun tek ve çok yıllık olduğu ve 16 nohut türünün bulunduğu birçok araştırmacı tarafından bildirilmiştir (Erol, 2019).

Nohut yapısında, %12.6 ile %29 arasında protein, %3.4 ile %8.8 arasında yağ ve %54 ile %71 arasında karbonhidrat olduğunu belirtmiştir (Wood ve Grusak, 2007). Çalışmada kullanılan nohut türelerinden *Cicer reticulatum* L. (C. reti.) tek yıllık, sadece Güneydoğu Anadolu bölgesinde yetişen endemik bir türdür. Yetiştikleri yerlere bakıldığı zaman terk edilmiş arazilerde 650 ila 1100 m yüksekliklerde yaşamını sürdürmektedirler. Nohudun şu anda kültürü yapılan nohut bitkisinin ilk atası olmasına rağmen günümüzde kültürü yapılmamaktadır. Benzer özelliklerde olan *Cicer echinospermum* L. (C. ech.) endemik bir tür olan nohut günümüzde kültürü yapılmamaktadır. Eğimli dağlık yerler, terkedilmiş arazilerde ve

mera alanlarında 700 ila 1000 m yüksekliklerde yaşamını sürdürmekte ve kültürü yapılmakta olan nohudun ikinci ata türü olarak kabul edilmektedir (Öztürk, 2011).

### 1.3.2. Nohut yetiştiriciliği

Nohut nadas alanlarını daraltmakta, kırsal kalkınmada kullanılmaları ve yayım projeleri sayesinde üretim alanları gün geçtikçe artmaktadır. Nohut bitkisi kışları ılık geçen Türkiye'nin güney ve batı bölgesinde kışlık, iç Anadolu ve diğer bölgelerde ise yazlık olarak tarımı yapılmaktadır. İlk kültüre alınan bitkilerden biri olan nohut, ülkemizde yemeklik baklagiller içerisinde ekilen alan ve üretim bakımından en fazla yapılan kurak bölge bitkisidir (Akçin, 1988). Nohut bitki olarak çok hassas olmayan toprak istekleri bakımından geliş alanlara sahip olması havadaki serbest azotu bağlamaları ile tahıl yemeklik baklagiller ile ekim nöbetine girmesi ile çok önemli bir bitkidir. Ülke genelinde tarımın çoğunluğunu tahıl üretimi oluşturmaktadır. Baklagiller de tahıl ile ekim nöbetine girmesi ve nadas alanlarını azaltması ile oldukça önemli hale gelmiştir.

Nohudun ekim zamanı iklime bağlı ve tane verimini etkileyen önemli faktörlerden olduğu için çok dikkat edilmesi gerekir. Ülkemizde iç bölgelerde karasal iklime sahip olduğundan ekimler şubat ayından mayıs ayına kadar yazlık olacak şekilde yapılmaktadır. Bu tarihlerden daha erken ekilen bitkiler kışın sert geçmesi ile toprakta fazla su dolasıyla fazla nem bulunur. Aşırı vejetatif büyüme ile hastalıklar yaygınlaşmada, tane verimi düşmektedir. Buna karşılık geç ekimde sıcaklıkların artması ile topraktaki nem oranı düşmektedir. Az olan nem çimlenmeyi olumsuz yönde etkilemekte ve birim alanda olması gereken bitki sayısından daha az bitki çıktığı için, generatif ve vejetatif dönemlerin kısalmasıyla birlikte verim düşmektedir. Bu sebeplerden dolayı ekim tarihinin doğru yapılması verim açısından önemlidir.

Ülkemizin Doğu Anadolu gibi bölgelerine bakıldığı zaman kışı sert geçen bölgelerde yıllık yağışın düşük olduğu erken ekim olmadığı gibi geç ekimlerin yapılması ile topraktaki nemin azalması ve çiçeklenme döneminin sıcak aylara denk gelmesi verimi önemli ölçüde düşürmektedir. Yetiştiriciler nohutta önemli kayıplara neden olan nohut antraknozu hastalığından kaçınmak için geç ekim olan mayıs ayına

ekimi tercih ederler. Bu sebepten nohut bitkisi yeterli çiçeklenme ve bakla dönemi büyüme yapmadan tane dönemi devreye girmekte ve verimde önemli düşmelere neden olmaktadır (Açıkgöz ve ark., 1987).

Nohudun iyi yetişebileceği toprak yapısı kumlu, drenajının iyi olduğu, organik madde bakımından zengin olan verimli ve pH değeri 5.5 ile 8.6 arasında değişen topraklar oldukça uygundur. Nohut ülkemizde bazı kıyı bölgelerde sulanarak, Orta Anadolu'da ise kuru koşullarda yetiştirilmektedir (Demirci ve Bildirici, 2020).

Çalışmada kullanılan nohut bitkisinin yetiştirildiği deneme alanın toprağı killi bünyeli, pH değeri nötre yakın olan yapısı ise kireçlidir. Besin maddelerince azot, fosfor ve potasyum bakımında fakir, organik maddesi düşüktür.

Verilerini kullandığımız nohut bitkisinin ekildiği Şanlıurfa ilinin meteoroloji Genel Müdürlüğünden alınan verilere göre iklim özellikleri bakımından karasal iklim özelliklerini göstermektedir. Ortalama sıcaklıklara bakıldığında en soğuk ay ocak, en sıcak ay ise temmuz olduğu anlaşılmıştır. En fazla yağış ocak aylarında düştüğü, en düşük yağış ise temmuz aylarında düştüğünü gözlemlenmiştir (Çakmak, 2019).

Nohutta hastalık ve zararlı olan organizmaların, çevreyle olan ilişkilerini, beslenme şekillerini, çiftleşme dönemlerini ve yaşamsal döngülerini dikkate alarak, uygun zamanda ve doğru bir şekilde mücadele ile zararlı olan organizmaların sayılarını tamamen yok etmeden ekonomik zarar seviyesinin altında tutarak zararlı yönetim sistemi olan entegre mücadele ile yapılmalıdır. Nohut bitkisine zarar veren ve verimi düşüren zararlılar; Nohut yaprak sineği, nohut antraknozu hastalığı, nohut yeşil kurdu, bozkurt Thrips spp., nohutta pas hastalığı, nohutta solgunluk, nohutta kök çürüklüğü hastalığı olarak görülmektedir. Yabancı ot olarak yabani hardal, tarla sarmaşığı olarak görülmektedir. Nohut bitkisine zarar veren hastalıklar, böcekler ve yabancı otlarla mücadele ilk olarak kültürel önlemler ve biyolojik mücadele yöntemlerini kullanılarak yapılır. Bu yöntemler ile mücadele yapıldığı halde zararlı hala devam ediyorsa kimyasal Mücadele yapılmalıdır (Tarım Gıda ve Hayvancılık Bakanlığı, 2017).

Çalışmada kullanılan nohut bitkisinin ekimi sıra üzeri mesafe 10 cm, sıranın uzunluğu 50 cm ve sıralar arası mesafe 60 cm olacak şekilde deneme kurulmuştur. Bloklar arasında 100 cm yürüme ve kontrol mesafesi bırakılmıştır. Tohum elle ekilerek yapılmıştır. Yabancı ot mücadelesi kontroller sonucu gerekli görülmesi halinde el ile tarladan uzaklaştırılmıştır (Çakmak, 2019).

### 1.3.3. Verim

Tarımda en önemli parametrelerden birisi verimdir. Nohut bitkisi kurağa dayanıklı bitki olması ile yağışın az olduğu alanlara ekilerek tarlanın boş kalmasını önler ancak yağışın az olması verim değerlerini etkilemektedir. Birim alandan daha çok ürün elde etmek için verimi yüksek türler tercih edilmelidir (Ermen ve ark., 1997).

Dünya nüfusunun hızla artması gıda tedariki ve güvenliği açısından risk oluşturmaktadır. Yeterli ve sağlıklı gıda maddelerinin sağlanması için gelişmiş ülkeler başta olmak üzere yakın zamanda çalışmalar başlatmıştır. Ancak her iki risk için önlem sağlamak biraz zor olmuştur. Yeterli gıda sağlandığında sağlıksız olabilmektedir. Dünya ülkeleri başta gelişmiş ülkeler olmak üzere hızla artan nüfus için verimli çeşitlerin geliştirilmesi yoluna gitmişlerdir. Teknolojinin gelişmesi ile birlikte verim kapasitelerinin artırılması bitkilerin ıslah edilmesi ve makine öğrenmesi ile doğru türlerin seçilmesi ile tarımda verimlilik artışı sağlanabilmektedir (Bayramoğlu, 2010)

Yapılan tez çalışmasında veri setleri; yetiştirilen nohuttan ölçülerek kayıt altına alınan değerlerden oluşmuştur. Elde edilen nohut bitkisine ait değerlerden sınıflandırma işlemi yapılarak tür tahmini yapılmıştır (Akay ve ark., 2018). Bilgisayarların öğrenmesini sağlayan makine öğrenmesi ile bilgisayarlar matematiksel ve istatistik yöntemler kullanarak kendine bir model çıkarır ve öğrendiği veriyi sınıflandırabilir ya da yeni veri tahmin edebilir. Verilerin tiplerine göre makine öğrenme yöntemlerinin başarı oranları değişmektedir. Bu nedenle, bir makine öğrenmesi yönteminin diğerlerinden üstün olduğunu söylemek mümkün değildir. Çok sayıya sahip verinin elle işlenerek kayıt edilmesi ve analizinin yapılması oldukça zordur. Daha önce kullanılan verileri kullanarak gelecek durumlar

için tahminlerde bulunarak işlem kolaylığı ve zaman tasarrufu sağlamaları makine öğrenmesinin önemini her geçen gün artırmaktadır. Yapılan çalışmada, nohuda ait özelliklerden yararlanarak, laboratuvar olanakları kullanılmadan tür tahmini ve dekar başı verimlilik tahminlerinin bilgisayar teknikleri kullanılarak anlaşılması ve yeni bir yaklaşım olan bu tekniğin zaman ve maliyet açısından da uygunluğunun gösterilmesi amaçlanmıştır.

# 2. ÖNCEKİ ÇALIŞMALAR

Makine öğrenme yöntemlerinin tarım alanında uygulamaları incelendiğinde; Pourdarbani ve ark. (2019), yaptıkları çalışmada İran'da farklı özellik ve fiyatlara sahip en popüler beş türü olan Adel, Arman, Azad, Bevanij ve Hashem nohutları sınıflandırmıştır. Benzer morfolojik özelliklere sahip olan nohut türleri elle ayırt etmek zordur. Nohut çeşitlerinin sınıflandırılması için iki farklı yöntem önerilmiş ve karşılaştırılmıştır. Görüntüler kamera ile çekilmiştir. İlk yöntem, renk ve doku özniteliklerinin çıkarılmasına, ardından en etkili özniteliklerin seçilmesine ve yapay sinir ağları ve parçacık sürüsü optimizasyonu (ANN-PSO) melezi ile sınıflandırmaya dayanmaktadır. İkinci yöntem, özniteliklerin açık bir şekilde çıkarılmasına dayalı değildir; bunun yerine, görüntü yamaları (RGB piksel değerleri) doğrudan üç katmanlı bir geri yayılım ANN için girdi olarak kullanılır. İlk yöntem ile %97 doğruluk oranına ulaşırken, ikinci yaklaşım %99.3 doğruluk oranına ulaşmıştır. İkinci yöntem ile diğer mahsul türlerine daha kolay uygulanabilir.

Pourdarbani ve ark. (2020), yaptıkları diğer bir çalışma olarak Adel, Arman ve Azad üç nohut çeşidini otomatik sınıflandırmıştır. Bu çalışmada nohut çeşitlerinin otomatik sınıflandırılması için bir bilgisayarlı görme sistemi kullanılmıştır. İlk olarak renk alanı eşiği kullanılarak segmentasyon yapıldı. Daha sonra, nohut örnek görüntülerinden renk ve dokusal özellikler çıkarılmıştır. Bu ön işlemlerden sonra, hibrit yapay sinir ağı-kültürel algoritması (ANN-CA) kullanılarak, en etkili beş özelliğin ayırt edici özellikler olarak seçilmiştir. Son olarak, hibrit yapay sinir ağı-parçacık sürü optimizasyonu (ANN-PSO), hibrit yapay sinir ağı-karınca kolonisi optimizasyonu (ANN-) olmak üzere üç farklı sınıflandırıcı çıktısını birleştiren bir ANN-PSO/ACO/HS çoğunluk oylama (MV) topluluk metodolojisi. ACO) ve hibrit yapay sinir ağı-harmonik arama (ANN-HS) kullanılmıştır. Sonuçlar, toplu sınıflandırıcı olarak ANN-PSO/ACO/HS-MV olarak ortalama 1000 rastgele yinelemeden sonra, test seti üzerinden %99.10 sınıflandırma doğruluğuna ulaştığını göstermiştir.

Kini ve ark. (2021), Hint nohudu için akıllı sınıflandırma modeli adlı çalışmasında 60 adet nohut görüntüsünden oluşmaktadır gri mod sonrada yeniden boyutlandırılarak ön işlemlere tabi tutulmuştur. Sekiz farklı özellik ile veriler kayıt edildi. Özellikler, baklagillerin büyüklüğü, çiçek, çeşitlilik, dağcı, cüce, düz tohum tipi, sert tohum, nominal kromozom değeri, meyve rengi ve baklagil grubu. Verilerden eğitim için %80 test için %20'lik dilimler ayrılmıştır. Bu işlemlerden sonra doğrulama performanslarına bakılmıştır. Toplamda sekiz adet algoritma kullanılmıştır. Destek vektör makineleri %86.07 yüksek doğruluk sağlamıştır.

Çetin ve ark. (2023), yaptıkları çalışmada nohut çeşitlerini kütlesine göre tahmin edip ayırma işlemlerini makine öğrenmesi ile gerçekleştirilmiştir. Çalışmada kullanılan veri setleri Erciyes Üniversitesi deneme alanlarında 2018 ve 2019 yıllarında deneme desenine 26 farklı nohut çeşidinden elde edilmiştir. Veri setindeki özellikler boyut, şekil ve alan özellikleri görüntü işleme yöntemi ile renk özellikleri ise kromametrik ve makine öğrenmesi yöntemleri ile belirlenmiştir. RF ve ML, tren testi ayrımı ile her iki yöntemde %80 ve k-katlı çapraz doğrulama ile yaklaşık olarak iki yöntemde %82 değerinde başarılı sonuçlar vermiştir.

Çetin ve ark. (2022), yaptıkları çalışmada AugeLab Studio yazılımında CNN yapısını kullanan gerçek zamanlı bir nesne algılama algoritması olan YOLO kullanılarak tohum sayısı ve fasulye ile nohut olan türleri algılama algoritması geliştirilmiştir makine öğrenmesi yöntemlerini kullanarak görüntü işleme ile tarımın kolaylaştırılması problemlere çözüm üretmesi ile önemli hale gelmiştir. Fasulye sınıflandırma doğruluğunun %97 ile %100 arasında değiştiği, nohut sınıflandırma doğruluğunun ise %91 ile %100 arasında değiştiği tespit edilmiştir. Sonuç olarak tohum üretim şirketleri, tarımsal biyoteknoloji laboratuvarları ve tohum sertifikasyon kuruluşları tarafından tohumların sayım ve sınıflandırılması için kullanılan bir yazılımdır. Ekilecek ürünlerin sağlam ve çimlenme değeri yüksek tohumları ayırması, yabancı maddelerin tespiti ve ayrıştırılması işlemlerinde kullanılarak daha kısa sürede ve daha az maliyetle kullanılabilir.

Kartal ve ark. (2021), 3D kanopi taramalarını kullanarak sayma görevlerini otomatikleştirmek için makine öğrenimi tabanlı tesis algoritmaları geliştirmişlerdir. Araştırmada nohut ve fasulye bitkilerini ekip çevre sıcaklığı, nem ve rüzgâr değerleri

kontrol altında tutulmuştur. Spetrumetre ile lazer renk yansıtılıp kamera ile yansıyan ışıkları kayıt edilmiştir. Elde edilen görüntü verilerinden %80 eğitim %20 test için kullanılmıştır. Sistem nohut ve fasulye bitkilerini çok iyi tanıyan bir model geliştirilmiş oldu. Fasulye bitkisi ile nohut bitkisini sırasıyla %93.18 ve %92.87 doğrulukla tahmin edilmiştir.

Ayele ve Tamiru (2020), makine öğrenmesi yöntemleri ile nohut türlerini sınıflandırmıştır. Sınıflamada 8 özellik kullanmışlardır. Verileri ön işlemden sonra %80 eğitim %20 test için verileri ayırmıştır. En iyi performans sonucunu veren makine öğrenmesi algoritması %97.5 ile karar ağaçları olmuştur. Bu uygulama ekonomik olarak daha uygun olmakla birlikte zaman ve işçilikten de tasarruf sağlanmış olundu.

Khatri ve ark. (2022), çalışmalarında buğday tohum sınıflandırmasında buğdayın 7 fiziksel özelliğine bakmışlardır. Veri seti 210 tane buğday tanesi olup UCI kütüphanesinden alınmıştır. Çalışmalarında makine öğrenmesi yöntemlerinden karar ağaçları, k en yakın komşu ve Naive Bayes algoritmaları kullanmışlardır. Sonuçları Naive Bayes, karar ağaçları ve k en yakın komşu sınıflandırıcıları için hesaplanan doğrulukların sırasıyla %92, %94 ve %92 olduğunu ortaya koymaktadır.

Belay ve ark. (2022), yaptıkları çalışmada, bir nohut hastalığı tespit modeli geliştirmişler. Görüntü alınıp işlendikten sonra görüntü yeniden boyutlandırma, normalleştirme ve gürültü filtreleme gibi iyileştirme işlemi yapılmış diğer gizli özellikler için CNN kullanılmıştır hastalığın tespitinde bu gizli özellikler daha çok etkili olmuştur. Veriler düzenlenip yeni özellikler eklenerek eğitim daha sonra da test için uygun hale getirilmiştir. Sonuç için çok sayıda görüntü ve deney yapılmıştır en iyi performans gösteren model %92.55 doğruluk ile CNN ile LSTM olmuştur.

Bayar (2021), yaptığı çalışmada, bir nohut yığını, hacmini ve ağırlık tahmin aracının ölçümde hatalarını azaltıp doğru değerlerini artırmak için makine öğrenmesi yöntemleri ve nohutların Moore-komşu izleme algoritmalarını kullanmıştır. Uygulamada kamera ve düşük maliyetli lazer sensörü kullanılmıştır. Daha önceki çalışmalarda en az %5 hata payı olurken bu uygulama ile hata payını %0.57 kadar düşürmüştür.

Kosmowski ve ark. (2018), yaptıkları çalışmalarında, Etiyopyada arpa, nohut ve sorgum çeşitlerini ayırt edilmesinde makine öğrenmesi algoritmalarından yararlanmışlar. 24 arpa, 19 nohut ve 10 sorgum çeşitlerinden toplamda 2650 tane arpa, nohut ve sorgum adedi Kuru taneleri, Etiyopya Tarımsal Araştırma Enstitüsü'nden 2017 yılında toplanmıştır. Sırasıyla arpa, nohut ve sorgum %89, %96 ve %87 doğruluk sağlamıştır. Destek vektör makinesi ve kısmi en küçük kareler analizi (PLS-DA) algoritmaları sürekli olarak diğer algoritmalardan daha iyi performans göstermiştir.

Kirtis ve ark. (2022), yaptıkları çalışmada nohutta ait bitki rejenerasyon ile model tahmini oluşturmuşlar. Araştırmada kullanılan nohut tohumları Diyarbakır Dicle üniversitesinden alınmıştır. Steril edilen tohumlar farklı konsantrasyonlarda bekletilip kültüre alındı. Yetiştirildiği oda sıcaklığı 22-26 °C bağıl nem %60 ve farklı zamanlarda farklı ışıklandırmalar ile bitki yetiştirilip sürgünlerin sayısı ve uzunlukları gözlemlendi.

Taheri-Garavand ve ark. (2022), yaptıkları çalışmalarında, tohum görüntülerini kullanarak nohut çeşitlerinin otomatik olarak tanımlamışlardır. Görüntü için iki cihaz kullanılmıştır. Aydınlatma ve görüntü alma için yoğun, normal ve bırakma katmanları bulunmaktadır. Beş kat çapraz doğrulama ile %94 ün üzerinde bir başarı performansı elde edilmiştir.

Singh ve ark. (2023), yaptıkları çalışmada hastalıkları doğru tespit edip iyileştirmek verim kayıplarını en aza indirir fikrinden yola çıkarak nohut solgunluk şiddeti iyileştirilmesi makine öğrenimi kullanılarak tahmin etmişlerdir. Çalışmadaki veriler Hindistan Tarımsal Araştırma Enstitüsü'nde üç yıl boyunca (2017-2022) saha deneyleri yapıldı. Termal görüntüler ile makine öğrenme metotlarını birleştirerek kombinasyon şeklinde kullanılmıştır. Tarladan elde edilen veri seti dört farklı sınıftan oluşmaktadır. Sonuçlar Kübist'in en iyi makine öğrenimi modeli olduğunu gösterdi. KNN modeli ise tarla koşullarında nohut solgunluk şiddetinin en zayıf tahmin edicisiydi. Kombinasyon edilen makine öğrenme metotları bireysel modellere göre solgunluk şiddetinin tahmin doğruluğunu iyileştirdiği görülmüştür.

Sinha ve ark. (2021), önceki raporlardan ve hava faktörlerinden alınan nohutta öldürücü bir hastalık olan kuru kök çürüklüğü DRR insidansı verilerini dikkate alarak DRR insidansının yapay sinir ağı tabanlı bir tahminini gerçekleştirmişler. DRR insidansı suyun az olması ve sıcaklığın çok olması ile artar. Toprak yapısı ile ilişkili olduğu tespit edilmiştir. Potasyum, organik karbon ve kil içeriği, DRR insidansı ile pozitif ilişkiliyken, toprak silt içeriği negatif olarak ilişkilendirilmiştir. Bu ilişkilerin bilinmesi ile daha doğru kararlar alınarak nohut yetiştiriciliğinde daha verimli sonuçlar alınabilir.

Podolny ve ark. (2020), yaptıkları çalışmada nohutta çiçek açan genlerin kontrol edilmiş makine öğrenimi yöntemleri ile analizi yapmışlardır. Çalışmada beş çiçek açan nohut bitkisine ait genin ifade dinamiklerini ile yapılmıştır. Kısa gün ve uzun gün bitkisi olarak yetiştirme koşulları altında on çiçek açan gen için daha önce yayınlanmış ifade verileri üzerinde rastgele orman yöntemi kullanılarak eğitildi. Veri olarak gen adı, çeşit adı ve büyüme koşulları özelliklerinden yararlanılmıştır.

Azimi ve ark. (2020), yaptıkları çalışmada nohut bitkisinin görüntüleriyle su stresini geleneksel makine yöntemleri ile derin öğrenme yöntemlerini karşılaştırılmıştır. Araştırmada kullanılan veri seti 3 farklı sulama koşulunda strese dayanıklı ve strese duyarlı iki nohut bitkisi incelenmiştir bitkinin dış görünüşünü kayıt eden kamera kullanılmıştır. Makine öğrenmesi metotlarından DVM, KNN ve Karar ağaçlarını kullanmışlardır en iyi performans %72 doğrulukla ile DVM göstermiştir. Derin öğrenme tekniklerinden ise CNN ve ResNet-18'i kullanmıştır. Tüm Derin Öğrenme yöntemleri arasında ResNet-18, %86 ve %84 doğruluk elde ederek geleneksel CNN'den daha iyi sınıflandırma performansı göstermiştir.

Saha (2022), yaptığı çalışmada nohudun örüntü tabanlı kalite değerlendirmesi için makine öğrenimi teknikleri ile değerlendirmiştir. Tüketiciler ve üreticiler için kalite en önemli etkenlerden biridir fikrinden yola çıkarak. Çalışmada 8 adet nohut çeşidi kullanılıp değerler Saskatchewan üniversitesinden alınmıştır. Kalite sonuçlarına göre pişirmesi zor ve pişirilmesi kolay olarak iki sınıfa ayrıldı. %100 ile en yüksek genel sınıflandırma doğruluğu ResNet50 ve MobileNetV2 ile sağlanmıştır.

Öziç ve ark. (2020), görüntüleri işleme ile nesne koordinat özelliklerini bakılarak bakliyat sayma işlemine yapmışlardır. Bant üzerinde hareket eden bakliyatların sayımı için hareketli eğimli ve eğimi kullanıcı tarafından ayarlanabilen bir bant tasarlanmıştır. Görüntülerin alınması için kamera kullanılmıştır ayrıca sonraki çalışmalar için de telesentrik bir lens temin edilmiştir. Kameranın çalışması Python yazılımı içinden de kontrol edilip değiştirilebilir. Sayma işlemi dışında renk, şekil ve çürüklük değerlerine de bakıldığı için arka plana ışık konulmuştur. Sistemin etrafını siyah bir naylon ile kapatılıp dışarıdan etkilenmemesi ve doğru sonuçlar alınabilmesi sağlanmıştır. Video çekimi sırasında görüntüsü üzerinde oluşan gürültü parçacıkları bir parça gibi görmemesi için iki iterasyon erozyon işlemi olan dikdörtgen sınırlayıcı metodu ile nesneleri en küçük dikdörtgen kutuya sınırlayıcı ile her bir görüntü karesinden hepsini bulmuştur. Çalışma prensibi olarak video görüntüleri üzerine sanal olan tek ve çift çizgi çekilmiş ve performansları ölçülmüştür. Çalışma sonucunda çok küçük boyutlarda olan nesnelerin paketlenmesi için kamera sistemi kullanışlı olmaktadır. Ancak ürünlerin çok küçük olması, malzemenin ve nesnelerin kameranın önünden çok hızlı geçmesi ile kameranın kayıt yapamaması, pozlama süresi uzun olmaması, çözünürlüğün düşük olması, malzemelerin kamera önünden geçerken arka arkaya geçmesi veya üst üste yapışması, aydınlatma sorunları gibi birçok sorunun oluşması ile iyi sonuçlar elde edilememiştir.

Ermeydan (2019), yaptığı çalışmada makine öğrenmesi yöntemleriyle çeltiğinin çekilmesi sonucu oluşan pirinç tanelerinin sınıflandırılması işlemini gerçekleştirilmiştir. Günümüzde pirinç çeltiğinin çekilmesi sonucu kabuğu soyulmaktadır taneleri ayırma işlemi el ile yapılmaktadır tanelerin sınıflandırma işlemi için görüntü ses gibi nesneler kullanılarak nesne tanıma işlemi gerçekleştirilip zaman ve işçilikten kazanç sağlanma hedeflenmiştir. Çalışmada 2 farklı sınıftaki 5 farklı pirinç türünden oluşan 10 sınıf ile yapılmıştır. Makine öğrenme yöntemleri olarak destek vektör makineleri, naive bayes, k en yakın komşu ve karar ağaçları algoritmaları kullanılmıştır. Pirincin kırık, olgunlaşmamış veya çürük olarak sınıflandırılmasında %97 üzerinde başarı oranı ile DVM uygulaması ile elde etmiştir.

Arı ve ark. (2019), makine öğrenme ile kayısı yapraklarını sınıflandırmışlardır. Bu çalışmada, 357 kayısı yaprağından oluşan Malatya gıda bakanlığına ait kayısı araştırma enstitüsü bahçesinde yetişmekte olan kayısı ağaçlarından elde edilmiştir. Derin Evrişimli Sinir Ağlarına (DCNN) bağlı olan önceden eğitilmiş bir DCNN modeli olan VGG19 modeli ile otomatik kayısı türlerinin sınıflandırılması yapılmıştır. 3 farklı tam bağlantılı katmana sahip öznitelikler farklı kombinasyonlarda birleştirilerek 7 farklı öznitelik ile çekirge iz, tokat oğlu Erzincan, aprikoz şalak, çataloğlu, paviot, hacıkızı, hırmanlı kayısı türlerini sınıflandırılmıştır. En iy performans fc8 katmanından doğruluk oranı %98.8 olarak elde edilmiştir.

Ataş (2016), yaptığı bu çalışmada Siirt fıstığını çok iri, iri, orta, küçük ve kabuğu çatlamamış olanlar sınıflandırmıştır. Siirt fıstığının TSE kalite standartlarına göre sınıflandırılması işlemi sırasında bir huniye toplanan fıstıklar borudan geçirilerek kamera ile görüntüleri alınıp en ve boyundan yarıçapları ölçülerek alanları hesaplanmıştır. Görüntüleri alınan Siirt fıstığının çapların oranı, fıstığın alanı ve kenarların alan oranı öznitelikleri çıkarılmıştır. 6 sınıftan oluşan ve her sınıfta 100 adet olan toplamda 600 adet fıstıktan görüntüler elde edilmiştir. Makine öğrenmesi sınıflandırmada algoritmaları olarak naive bayes, yapay sinir ağları ve destek vektör makineleri kullanılmış Genel sınıflandırmaya bakıldığı zaman en iyi başarı yüzdesi ile yapay sinir ağları doğruluk oranı %83.33 olarak elde etmiştir.

Koyuncu ve Gök (2019), yaptıkları çalışmada ıslah çalışmalarında verim değerini regresyon ile tahmin etmiştir. Bitki ıslahında dış görünüş ve genotip özellikleri verim ve çevre ile olan ilişkileri üzerinde çalışır. Elde edilen değerlerin karşılaştırılmasında bazı istatistik programları kullanılmaktadır. Bu paket programlar gerekli genotip seçimleri ve analizleri için yetersiz kalmaktadır. Çalışmada 2018 yılında arpa bitkisine ait 1153 adet verim değerini tahmini yapılmıştır. Tahminleme için bitkilerin ıslah çalışmalarında doğrusal regresyonun kullanılmaktadır. Ayrıca makine öğrenmesi metotlarından olan SMO, KNN ve RO algoritmaları kullanılmıştır. Başarı oranları için mutlak hata metriklerine ile ortalama karesel hatanın karekökü göre yapılmıştır. Rastgele orman, diğer yöntemlere göre daha yüksek başarı göstermiş

Tunca ve Köksal (2021), yaptıkları çalışmada güvenli gıda ve tarım alanlarının yönetimi için önemli bilgiler ve kolaylıklar sağlayan bitkisel ürünlerin sınıflandırılması adlı çalışmasında uydu görüntülerinden yararlanmıştır. Yapılan çalışmada sentinel 2 olarak bilinen uydu görüntüleri ile bitki türlerini makine öğrenmesi yöntemleri ile belirlemiştir. Bitki türleri uydudan tahmin edilebiliyor fakat bitkilerin birbirine benzemesi başarı oranını düşürmektedir. Bu nedenden dolayı uydudan görüntülenen bitki türleri ile sınıflandırılması hassas ve başarı oranının yüksek olması yeni bir yaklaşıma ihtiyaç duyulmaktadır. Çalışma Amasya ilinin Tarım Müdürlüğü tarımsal üretim arazilerinde 2020 yılında gerçekleştirilmiştir. İklim özellikleri, yıllık yağış ve sıcaklık farklarına bakıldığı zaman yarı kurak bir bölge olarak sınıflandırılıyor. Sulama derin kuyulardan ve mevcut yakındaki sulama kanallarından sağlanmaktadır. Bölgede yetişen ürünler; buğday, nohut, yonca, fiğ, kenevir ve silajlık mısırdır. 2020 yılına ait Sentinel-2 uydudan farklı tarihlerde 7 adet görüntü kullanılmıştır. Veri seti, uydu ile elde edilen görüntüler her bir bitki sınıfı için ayrı olarak etiketlenmiş ve tarihe göre sıralanmıştır. Çalışmada toplamda elde edilen 712 026 adet görüntü hücresinden oluşmaktadır. İlk olarak veri normalize işleminden sonra %70 eğitim, %30 test Scikit-learn kütüphanesi tarafından otomatik olarak seçilip uygun hale getirilmiştir. Makine öğrenmesi yöntemlerinden RO, DVM ve K-NN algoritmalarını kullanmışlardır. Çalışmadan elde edilen başarı RO %95.3, K-NN %91.8 ve DVM %75.9 olmuştur.

Altun ve Türker (2021), yaptıkları çalışmada çoklu zamanlı sentinel-2 görüntüleri ile Mardin Kızıltepe bölgesindeki tarımsal ürün tespiti yapmışlardır. Bu çalışmada, tarımsal arazilerinden oluşan 2018 yılına ait Sentinel-2 uydu görüntülerinden oluşmaktadır. Görüntüler Earthexplorer web sitesinden ücretsiz olarak alınmıştır. Elde edilen tarım ürünleri buğday, mısır, pamuk, nohut, mercimek ve diğerleridir. Görüntüler nisan ayından ekim ayına kadar çekilmiş 6 görüntüden oluşmaktadır. Sınıflandırmada 10m mesafeden konumsal çözünürlüklü olan kırmızı, yeşil, mavi, kızıl ötesi ve NFBİ olan 5 bantta kullanılmıştır. Rastgele orman algoritması ile sınıflandırma işlemi, toplam 30 bantlı görüntü yığınının tek seferde yapılmıştır. Eğitim test ve sonuçların değerlendirilmesi mevcut Çiftçi Kayıt Sistemi (ÇKS) verilerinden yararlanılmıştır. ÇKS verisinde parsel büyüklükleri ile parsele ait özellikler olan veri tipi, toplam parsel yüz ölçümü, ekilmiş alanın yüz ölçümü ve

ekili ürün cinsi veya adı gibi kayıt bilgileri yer almaktadır. Toplam 41 968 adet parselden oluşmaktadır. ÇKS sisteminde var olan öznitelik verileri ile gerçekte olan veriler ile aynı olmadığı arasında farkların olduğu görülmüş gerekli ekleme, çıkarma ve düzeltmeler işlemleri yapılmıştır. Gerekli düzeltmelerden sonra kalan 5 803 adet parsel sayısı ile çalışmada verisi olarak kullanılmıştır. Sınıflandırma neticesinde %96.35 genel doğruluk değerlerine ulaşılmıştır.

Üstüner ve ark. (2019), yaptıkları çalışmada tarımsal ürünlerin sınıflandırılması işlemini çok zamanlı polarimetrik SAR verileri ile gerçekleştirilmiştir. Çalışmada kullanılan tarım ürünleri buğday, mısır, patates, ayçiçeği ve yem bitkileri ürünlerinin sınıflandırılması yapılmıştır. Makine öğrenmesi olarak LightGBM, destek vektör makineleri ve rastgele orman algoritmaları kullanılmıştır. Çalışmanın yapıldığı alan Konya ilinin kuzeyinde yer alan ve yaklaşık 42 km²'lik bir alan seçilmiştir. Sınıflandırmalar sonucunda başarı performansları ve genel doğruluk değerleri için k-katlamalı çapraz doğrulama yöntemi kullanılıp değerleri LightGBM, RO 0.857 ve DVM 0.834 olarak elde edilmiştir.

Teke ve ark. (2019), yaptıkları çalışmada tarımsal faaliyetlerin etkin olarak yürütülmesi için Güneydoğu Anadolu Bölgesi'nde GAP BKİ ve TÜBİTAK UZAY "GAP Bölgesi'nde Hassas Tarım Uygulamalarının Yaygınlaştırılması çalışmasını yürütmüşlerdir. İlk olarak uzaktan algılama ile veri kaynakları alınıp yersel spektral ölçümler ile analizler gerçekleşmiştir. Ayrıca hassas tarım analizleri için yazılımı ve bunların sunulacağı bir web arayüzü geliştirilmiştir. Çalışmanın ikinci aşamasında, mısır bitkisinde farklı oranlarda gübreleme yapılarak verim artışının ölçülmesi hedeflenmektedir. Proje önce GAP Bölgesi'ne daha sonra Türkiye geneline yaygınlaştırılıp uygulanması çiftçilere ve devlete kalkınma planlarının belirlenmesi amaçlanmaktadır.

Karadag ve ark. (2018), su tarımda ürün yetiştirip kaliteli ve verimli ürünler elde etmek için bitkinin en çok ihtiyaç duyduğu faktörlerden birisidir. Fikrinden yola çıkarak yapay öğrenme yöntemleri ile spektral yansımalardan biberde su stresinin belirlemiştir. Bitkiler yetişme dönemlerine göre farklı oranlarda suya ihtiyaç duyar. Çalışmada spektral yansımalar ile biber bitkisinin su stresinde olup olmadığı belirlenmeye çalışılmıştır. Çalışmada kullanılan veri seti 3 gruba ayrılmıştır. A grubu

%100 suya maruz bırakılan biber bitkisi yapraklarından alınan yansıma değerlerini içerir. B grubu %75 suya maruz kalan biber bitkisi yapraklarından alınan yansıma değerlerini içerir. C grubu %50 suya maruz kalan biber bitkisi yapraklarından alınan yansıma değerlerini içerir. Makine öğrenmesi sınıflandırma algoritması olarak k en yakın komşular ve yapay sinir ağları yöntemleri kullanılmışlardır.

Bayrakçı ve ark. (2021), yaptıkları çalışmada makine öğrenmesi ile tarım arazilerinde harcanan su miktarını belirlemişlerdir. Yapılan çalışmada bitki su tüketimleri üzerinde çalışılmıştır. Veriler Tarım ve Orman Bakanlığının sitesinde bitkilerin su tüketimi adlı çalışmasından alınmış ve veriler düzenlenmiştir. Türkiye'de sulanan bitkilerin Toplamda 3 203 adet veri kullanılmıştır. Veri setinde tarih, il, bitki, etc, gölge ve yağış sınıflarından oluşmaktadır. Çalışmada tahmin işlemi, Python uygulamasından olan bir arayüz oluşturulmuştur. XGBoost Regresyon algoritması %99.95 doğruluk oranı ile en iyi performansı göstermiştir. Verideki özellikler incelendiğinde genel parametrelerin aynı derecede önemde olduğu en önemli parametreye bakıldığında bitki herhangi bir hastalık veya bitki besin elementi eksikliği çekmediği su tüketimi olarak bilinen etc olmuştur. Bitkinin ihtiyaç duyduğu su miktarı dönemlere göre değişmekle birlikte ekilecek bitki için su parametrelerine bakıldığında yağış, etc değeri ve tarih değerlerine göre değişmektedir bu parametrelere dikkat etmek gerekir.

Karadağ ve Taşaltın (2016), yaptıkları bu çalışmada, işlem yapılmamış biber bitkisinin yapraklarından alınan spektral yansımalar ile makine öğrenmesi ile sınıflandırılma yapılmıştır. Veri olarak kullanılan toprağın özellikleri, nem ve ortam sıcaklığı gibi değerlerin sabit olduğu bir ortam olan Şanlıurfa GAPTEAM'da bitki yetiştirme laboratuvarında biber bitkisine ait örneklerden alınmıştır. Veriler laboratuvarda yetiştirilen biber bitkisi yapraklarından spektroradyometre cihazıyla alınmıştır. Zararlı funguslar bibere topraktan verilmiştir. Hastalıklı ve sağlıklı biber olarak iki sınıf kullanılmıştır. Bu çalışmanın amacını biber bitkisine ait yansıma değerleri ile hastalıklı bitkiyi tespit etmektir. İki grup olan çalışmada (A) sağlıklı biber, (B) ise hastalıklı biber bitkisinden oluşmaktadır. Yansıma değerlerinden uygun dönüşümler sağlanarak Matlab ortamına geçirilmiştir. Normalizasyon ve min. max.

değerlerden yararlanarak sınıflandırma için uygun hale getirilmiştir. Sınıflandırma yöntemi olarak Yapay Sinir Ağları yöntemi kullanılmıştır.

Yaman ve Tuncer (2022), makine öğrenmesi yöntemleriyle bitkilerdeki yaprak hastalığı tespiti ve özellik çıkarımı gerçekleştirilmiştir. Yaptıkları bu çalışmada ceviz yapraklarından elde edilen 726 görüntüden oluşmaktadır. Sağlıklı ve hastalıklı olmak üzere iki sınıfa ayrılmıştır. Çoğu makine öğrenmesi denemesi test yapılmış en iyi sonuç veren DarkNet53 ve ResNet101 modelleri seçilmiştir. İki modeldeki özellik birleştirilip hibrit yapılmıştır. Özellik arasında en önemli özellikler seçilmiştir. Başarı performansları karşılaştırıldığında en iyi başarı oranı destek vektör makinesi algoritması ile %99.58 doğruluk oranı hesaplanmıştır.

Sevli (2022), yaptığı çalışmada üretimde kayıpların azaltıp kaliteli ve verimli ürünler elde etmek için hastalıklarla mücadele etmek gerekir. Bitki hastalıkları çeşitli sebeplerden dolayı meydana gelmektedir. Hastalığın tespit edilmesi, mücadelesi ve uygulamada mühendislerin çabaları yetersiz kalabilmekte, yanlış uygulama yapılma olasılığı artmaktadır. Bu sebepten bitki hastalıklarının tespitinde uzmanlara ve ziraat mühendislerine destek olacak elma bitkisi hastalıklarının derin öğrenme ile tespiti çalışmasını yapmıştır. Çalışmada makine öğrenmesinin yapay sinir ağları algoritmasının çok katmanlı bir formu olan derin öğrenme ve evrişimsel sinir ağları (ESA) kullanılıp 1821 adet örnekten oluşmaktadır. Veri seti sağlıklı ve üç farklı hastalığa sahip elma ağacı yapraklarından oluşmaktadır. Modelin başarısını arttırmak için ve dengesiz verilerden arındırılmış eşit veri kümesine sahip sınıflar oluşturmak için veri artırımı uygulanmıştır. Modelin başarı oranına bakıldığı zaman %98.76 doğruluk yüzdesi değerine ulaşılmıştır.

Türkoğlu ark. (2021), yaptıkları çalışmada kayısı hastalıklarını sınıflandırmak için derin evrişimsel sinir ağı kullanmışlardır. DESA modelinde Normalizasyon, Havuzlama gibi katmanlardan oluşmaktadır. Çalışmada Bingöl ve İnönü Üniversitelerinin Ziraat Fakültelerinde yetişen kayısı hastalıklarından oluşan görüntülerden İHA ve fotoğraf makinesi ile elde edilmiştir. Elde edilen görüntülerden 467 çil, 139 kanser belirtisi, 195 kuruma ve 85 adet monilya hastalıklarından oluşan 4 hastalık ve toplam 886 görüntüden oluşmaktadır. Veri rastgele bir şekilde %70 eğitim, %30 test için ayrılmıştır. Gerçekleştirilen deneysel

sonuçlara göre, yüksek başarıya sahip model derin ağ modeli olmuştur. En başarılı model ile 9×9 boyutlarında evrişim filtresi kullanılarak başarı oranı %98.20 en yüksek doğruluk yüzdesi elde edilmiştir. Ayrıca önerilen model ile kayısı meyvesinde kanser belirtilerini tanıma %100 başarı oranı elde edilmiştir.

Jahanbakhshi ve ark. (2021), görüntü işleme tekniğine ve derin öğrenme ile zencefil tozundaki sahtekarlığı tespit etmişlerdir. Zencefil, gıda ve ilaç endüstrilerinde kullanılan ekonomik kazancı yüksek baharatlardan biridir. Bu çalışmanın amacı zencefil tozuna ekonomik değeri düşük tozlar karıştırılması gıda kalitesini ölçmek için tahribatsız yöntemlere olan talep olmuştur. Sınıflandırma sonuçlarını iyileştirmek için Toplu normalleştirme (BN) tekniği kullanıldı. 7 kategori ve 3 360 adet zencefil tozu görüntü örneği hazırlanmış sahtekarlığı tespit etmek için zencefil tozu görüntülerini sınıflandırılmıştır. Önerilen CNN'nin zencefil tozu görüntülerini diğer sınıflandırıcılara kıyasla %99.70 doğrulukla derecelendirebildiğini gösterdi.

Yıldız ve ark. (2015), yaptıkları bu çalışmada muz meyvesine ait 0.5 cm, 1 cm ve 1.5 cm dilim kalınlığında 50°, 60° ve 70° sıcaklıklarda kurutma havası ile kurutulmuştur. Çalışmada kullanılan muzlar ithal muz olup yerel pazardan alınmıştır. Hata oluşmaması için tedarik edilen kişiden toplu olarak alınıp nem değerlerine etüvde ilk nem tayini 50 gram ve 1 cm kalınlığında olan muz kabukları kesilerek 6 örnek kullanılmıştır. 70 °C sıcaklıkta kabuktaki ağırlık kaybı sabitleninceye kadar bekletilmiştir. Kurutma işlemi etüvde yapılmıştır. Örnekler 100 gram olarak hazırlanmış yaş ürün nemini %10 ile %13 seviyesine düşürmek amaçlanmıştır. Araştırmada ileri beslemeli yapay sinir ağları modeli kullanılmıştır giriş katmanında üç katmanlı sıcaklık, dilim kalınlığı ve zaman girişleri çıkış için tek nöron olan nem oranı değeri kullanılmıştır. Tüm veriden elde edilen değerlerden rastgele %60 eğitim, %20 doğrulama ve %20 test için ayrılmıştır. En kısa kurutma 70 °C sıcaklıkta 0.5cm kalınlıktaki dilim ile 6 saat kadar bekletilen uygulama olurken en uzun kurutma işlemi 50 °C sıcaklıkta 1.5 cm kalınlıktaki dilim ile 42 saat bekletilerek yapılan uygulama olmuştur. Araştırma sonuçlarına göre kuruma performansı, renk analizi ve suda çözünebilir madde analizleri yapıldı. Çalışma sonucunda genel değerlendirme

yapıldığında muz meyvesinin 60 °C sıcaklıkta kurutulması en uygun olarak belirlendi.

Sabancı ve ark. (2012), çalışmada değişik boyutlarda patatesler CCD sensörlü bir webcam kullanılarak alınmıştır. Matlab yazılımıyla görüntü işleme ve yapay sinir ağlarının kullanılmıştır. İşlemlerin olası hataların önüne geçmek ve hızlı yapılabilmesi için arka plan koyu siyah yapıldı. Gürültü ve parazitlerin giderilmesi için resimlere filtreleme işlemi yapılmıştır. Çalışmada kullanılan makine öğrenmesi algoritmalarından yapay sinir ağlarından çok katmanlı ileri beslemeli modeli kullanılmıştır. Çalışmadan önce ezik, çürük ve bozuk şekilli patatesler tespit edilip atılmış sınıflamaya dahil edilmemiştir. Çalışma sonucunda patateslerin büyük boy, orta boy ve küçük boy patatesler şeklinde üç gurupta sınıflandırma işlemine tutulmuştur. Çalışma ayrıca şekil bozukluğu olan patatesler tespit edilip kaliteli patatesleri belirlemiştir. Bu çalışma ile yürüyen bant ve kamera ile patatesler sınıflandırılabilir ezik, çürük ve şekil bakımından bozuk olan patatesler ayırt edilebilir. Bu çalışma ile satıcılar ve tüketiciler için kolaylık sağlanmış olur.

Timur ve ark. (2011), yaptıkları çalışmada İstanbul' un Göztepe bölgesinde rüzgâr tahminleri için makine öğrenmesi yöntemi kullanılmıştır. Rüzgâr ölçümleri ve gelecekteki rüzgâr tahmini birçok alan için gereklidir. Fikrinden yola çıkarak gelecekte rüzgâr hızının ne olacağını tahmin edilmeye çalışılmıştır. Makine öğrenmesi yöntemlerini uygulanıp sonuçlar elde etmek için giriş değerleri sıcaklık, tarih ve basınç seçilmiştir. Ölçümler bir gün 24 saat olacak şekilde 24 adet yapılmış ve kayıt edilmiştir. Eğitim verisi için 7 400 adet ölçüm yapılmıştır. Sonuçlar incelendiğinde en iyi başarı performansı Bagging sınıflandırı ile doğruluk yüzdesi CC değeri %91.14 olarak bulunmuştur. Bu değer gerçek değer ile tahmin edilen değerin birbirine çok yakın olduğunu gösterir. Ayrıca hata değerinin çok küçük sıfıra yakın olması önemli hataların olmadığını göstermiştir.

Alpay ve Erdem (2018), yaptıkları bu çalışmada, makine öğrenme yöntemlerinden olan yapay sinir ağlarını ile Elâzığ ilinde bulunan bir serada ısıtma ihtiyacını iklim koşullarına bakılarak yapmıştır. Çalışma 2017 yılında yapılmıştır. Isıtma ihtiyacı ilkbahar mevsiminde mart ve nisan sonbahar mevsiminde ekim ve kasım ayları kış mevsiminde aralık, ocak ve şubat aylarında olduğu önceki

senelerden belirlenmiştir. Serada plastik örtü kullanılarak taban ve yüzey alan ölçülerine belirlenmiştir. İklim verileri için seranın kurulduğu Elâzığ bölgenin en yakın bölgesindeki 2017 yılına ait sıcaklık ve nem değerleri Meteoroloji Genel Müdürlüğünden temin edilmiştir. Sıcaklık, nem ve ısıtma derece gün verileri ayrıca enlem, boylam ve rakım kullanılan verilerdir. 2017 yılında gece gündüz ve ortalamadan oluşan toplam 5 065 ölçüm yapılmıştır. Bu ölçümler 3 sınıfa ayrılmıştır. Yapay sinir ağları yöntemi ile yapılan çalışmada 75 veri eğitim, 12 veri test toplamda 87 veri kullanılmıştır. Araştırma sonucu yapay sinir ağları ile ısıtılan seraların %20 daha az enerji kullanılarak enerji tasarrufu sağladığı ve geleneksel olarak ısıtılan seralara göre daha başarılı sonuçlar verdiği görülmüştür.

McAllister ve ark. (2018), farklı gıda görüntü veri kümelerini sınıflandırmak için derin artık sinir ağı özelliklerini denetimli makine öğrenimi algoritmalarıyla birleştirmek adlı çalışmalarında obezitenin önüne geçmek sağlıklı besinler tüketerek sağlıklı yaşamak için bu araştırmayı yapmışlardır. Obezite günümüzde diyabet (şeker) kalp hastalığı vb. çoğu hastalığın temelini oluşturuyor bunun sebebi de sağlıksız besinler tüketmemiz yediklerimize dikkat etmememizden kaynaklanıyor. Çalışmada görüntü verisi ile makine öğretilmiştir. Naive bayes, yapay sinir ağları ve destek vektör makineleri yöntemleri kullanılmıştır. En iyi doğruluk performans gösteren sınıflandırıcı %99.4 ile yapay sinir ağları olmuştur.

Kar ve ark. (2021), yüksek verimli fenotipleme verilerinden evapotranspirasyon değerlendirmesi için optimum örnekleme süresinin belirlenmesi için bir toplu makine öğrenimi yaklaşımı çalışmalarında açık ortamda günde her 2 saatlik aralıklarla yaklaşık 5000 saksı bitkisinin 3 boyutlu (3B) lazer taramasını gerçekleşmiştir. Genotipleri sınıflandırmadan önce, yanlış özelliklerin performansı düşürdüğü fikrinden yola çıkarak yüksek düzeyde ilişkili olan özellikler seçildi. Yapay sinir ağları, rastgele orman ve destek vektör makineleri algoritması kullanılmıştır. Eğitim verileri için %70 test için %30 kullanılmıştır. Yapılan çalışma sonucunda F1'den F4'e kadar olan çalışmada RMSE değerlerinin tutarlı olduğu bulundu. F1 ile F12 arasında başarı oranları değişmekle birlikte SVM tüm ölçeklerde sürekli olarak RF ve ANN ‘den daha iyi performans gösterdi

Mahmood ve ark. (2022), tarımda çağdaş makine öğrenimi uygulamaları adlı çalışmalarında makine öğrenimi uygulamaları ile tarım ve gıda sistemlerini son 10 yılda yayınlanan 81 adet araştırma değerlerini incelemişlerdir. Araştırma sonucunda 4 ana kategori olan 1. Toprak gözlemi ve su izleme 2. Mahsul yönetimi 3. Hayvancılık yönetimi 4. Tür yönetimi ve alt dallarını inceleyip sınıflandırmışlardır. Araştırma sonuçları tarımın hangi alana yöneldiğini göstermektedir. Bu araştırma araştırmacılar, akademisyenler, mühendisler ve üreticiler olan çiftçinin hangi alanlara yönelmesi gerektiğini yol göstermiş oluyor.

Demir ve ark. (2018), makine öğrenmesi ile arı türlerini sınıflandırmıştır. Makine öğrenmesi ile ilgili araştırmalar incelendiğinde tarım zararlıları olan böcekler ve hayvanlarla ilgili yeterli kaynak ve araştırmanın olmadığı anlaşılmıştır. Bu çalışmada beş farklı ilden alınan Bozcaada, Gökçeada, Kırklareli, Saray ve Muğla bal arına ait kanat fotoğraflarından elde edilen veriler incelendiğinde üzerinde kavşak noktaları belirlenmiştir. Kanat yapıları üzerinde bulunan kavşak noktaları ile 27 morfolojik özellik ve 150 farklı arı kanatları üzerinde sınıflandırma yapılmıştır. Veriler toplanıp bütün özellikler çıkarıldıktan sonra ön işlem olarak normalizasyon işlemi uygulayarak verideki değerler 0 ile 1 arasına indirgenmiştir. Bu işlem ile verilerin işlenmesi kolaylaştırılmıştır. Yapay sinir ağları, k en yakın komşu ve destek vektör makineleri yöntemleri kullanılmıştır. Başarı performansları değerlendirildiğinde en başarılı sonuç veren yöntem yapay sinir ağları olmuştur.

# 3. MATERYAL ve YÖNTEM

Yapılan bu çalışmada, Harran Üniversitesi Ziraat Fakültesi deneme arazisi kullanılarak nohut türlerinden elde edilen veri setleri oluşturmaktadır. Veri setleri öncelikli olarak ham veriler, daha sonra ham + sentetik veriler ve en son olarak sadece sentetik veriler olacak şekilde oluşturulmuştur. Veri setlerinin ayrıştırma işlemi, çapraz doğruluk 5 (ÇD-5) yöntemi ve verilerin %50'si eğitim %50'si test (%50-%50) olacak şekilde yapılmıştır. Sınıflandırma yöntemleri olarak KA, DVM ve KNN yöntemleri kullanılarak nohut tohum tür tahmini yapılmıştır. Regresyon yöntemleri olarak ise KA ve DVM yöntemleri kullanılarak nohut tohum türleri de özelliklere dahil edilerek dekar başı verimlilik tahmini yapılmıştır. Performans ölçütü olarak sınıflandırma için doğruluk, regresyon için rmse kullanılmıştır. Tezin bu bölümünde materyal alt başlığında çalışmada kullanılan verilerin özellikleri ve nasıl elde edildiği detaylı olarak açıklanacak. Yöntem alt başlığında ise çalışmada uygulanan işlemler ve kullanılan makine öğrenme yöntemleri anlatılacaktır.

## 3.1. Materyal

### 3.1.1. Veri seti özellikleri

Çalışmada veri seti olarak kullanılan nohut türlerinden C. ech. ait 41 örnek, C. reti. ait 23 örnek olmak üzere toplam 64 örnek kullanılmıştır. Çalışmada veri olarak kullanılan nohut türleri hakkında bilgiler incelendiğinde C. reti. tek yıllık, kazık köklü bir baklagil olup, çiçeklenme dönemi mayıs ile haziran aylarında olup çiçek pembe renge sahiptir. Tohum rengi de gri, açık ve koyu kahverengi renk tonlarındadır. C. ech. tek yıllık kazık köklü bir baklagil türüdür. Çiçeklenme dönemi mayıs ayına denk gelmektedir. Çiçek pembe renge sahiptir. Tohum yüzeyi belirgin uzun dikenli tohum rengi koyu kahverengiden siyaha doğru olan tonlardadır (Öztürk, 2011).

Bitkinin ekildiği toprak kireçli olup pH değeri nötre yakın killi bünyelidir. Bitkinin temel besin değerleri olan azot, fosfor ve potasyum bakımından fakir organik maddece yetersiz özelliktedir. İklim koşulları tipik karasal iklimdir yazları

sıcak ve kurak kışları ılık geçer. Tohum gübresiz olarak el ile ekilmiş sulama gerekli görülmediği için yapılmamış. Yabancı otlar gözlem sonucu ihtiyaç olduğunda elle topraktan uzaklaştırılmıştır.

Hasat olgunluğuna gelen bitkiler el ile toplanmış sap ve çatlamış olan kuru bakla kabuklarından arındırılmıştır. Saç kurutma makinesi ile belli bir mesafede tutarak tohumlar tamamen diğer küçük kabuklardan temizlenmiştir.

Nohut türlerine ait özelliklere bakıldığı zaman; 100 tane ağırlığı, tartılarak gram cinsinden hesaplanmıştır. Tohumun eni, boyu ve genişlikleri, kumpas yardımı ile mm cinsinden hesaplanmıştır. Tohum hacmi; nohut tohumu ait en, boy ve genişlikler incelendiğinde şekil ve değerleri olarak dikdörtgen prizmaya benzediği için en, boy ve genişlik değerlerinin çarpımı ile hesaplanmıştır. Tohum yoğunluğu; 100 tane ağırlık olan değer 100'e bölünerek 1 nohudun ağırlık değeri bulunmuş ve daha önce bulunan tohum hacmine oranlanması ile elde edilmiştir. İlk çiçeklenme gün sayısı, ekim tarihinden ilk çiçeğin görüldüğü zamana kadar geçen gün sayısı. İlk bakla bağlama gün sayısı, ekim tarihinden ilk bakla bağlamanın görüldüğü zamana kadar geçen gün sayısı. Olgunlaşma gün sayısı, nohut ekimi ile bitkilerin tüm yeşil aksamları olan dal, yaprak ve tanelerin sarardığı tarihe kadar geçen gün sayısı. Bitki taç genişliği, bitkinin düşey ve yatay uzunluk toplanıp ikiye bölünerek çap hesaplanmış daire alanından ($\pi r^2$) bulunmuştur. Parsel verimliliği, hasat edilmiş bitkilerin tamamı sap ve tane verimlerinin birlikte tartılması ile bulunmuştur. Siyah noktalar, nohut tohumunun üzerindeki siyah lekeler. Dekar verimliliği, Parselin alanından 1000 $m^2$ olan 1 dekara oranlanarak hesaplanmış. Nohut türü, sonuç olarak bulunmaya çalışılan *Cicer echinospermum* ve *Cicer reticulatum* nohut türü değer olarak 0 ve 1 verilmiştir. Veri setine ait özellikler Çizelge 3.1.'de verilmiştir (Çakmak, 2019).

Çizelge 3.1 Veri setine ait özellikler

| Veri setinde Kullanılan Özellikler | Özellik açıklamaları | Özellik değer aralığı | Norm. Değer aralığı |
|---|---|---|---|
| 100 tane tohum ağırlık (g) | Tartılarak hesaplanmış | 8.62- 17.24 | |
| Tohum Genişlik (mm) | Kumpas Yardımı ile ölçülmüş | 4.17- 6.48 | |
| Tohum En (mm) | Kumpas Yardımı ile ölçülmüş | 4.52- 6.75 | |
| Tohum Boy (mm) | Kumpas Yardımı ile ölçülmüş | 5.95- 9.01 | |
| Tohum hacmi ($cm^3$) | Tohumun en, boy ve genişlik değerlerinden bulunmuştur. | 0.10- 0.30 | |
| Yoğunluk ($g/cm^3$) | Tohum ağırlığının hacmine oranlanmasıyla bulunmuştur | 0.39- 1.24 | |
| İlk Çiçeklenme gün sayısı | Ekim tarihi ile ilk çiçeğin görüldüğü gün | 121- 143 | |
| İlk Bakla bağlama gün sayısı | Ekim tarihi ile ilk bakla bağlamanın görüldüğü gün | 132.7- 149.5 | |
| Olgunlaşma gün sayısı | Nohut ekimi ile bitkilerin tüm yeşil aksamları olan dal, yaprak ve tanelerin sarardığı tarihe kadar geçen gün sayısı | 174- 223.7 | 0.1-0.9 |
| Bitki taç genişliği ($cm^2$) | Düşey ve yatay uzunluk toplanıp ikiye bölünerek çap hesaplanmış daire alanından bulunmuştur. | 1014-2950 | |
| Parsel verimi (g/parsel) | Hasat edilmiş bitkilerin tamamı sap ve tane verimlerinin birlikte tartılması ile bulunmuştur. | 15.6- 108.1 | |
| Siyah noktalar | Var- yok | 0-1 | |
| Dekar verimi (kg/da) | Parselden oranlanarak hesaplanmış | 52- 360.2 | |
| Cins tür | *Cicer echinospermum ve Cicer reticulatum* | 0- 1 | |

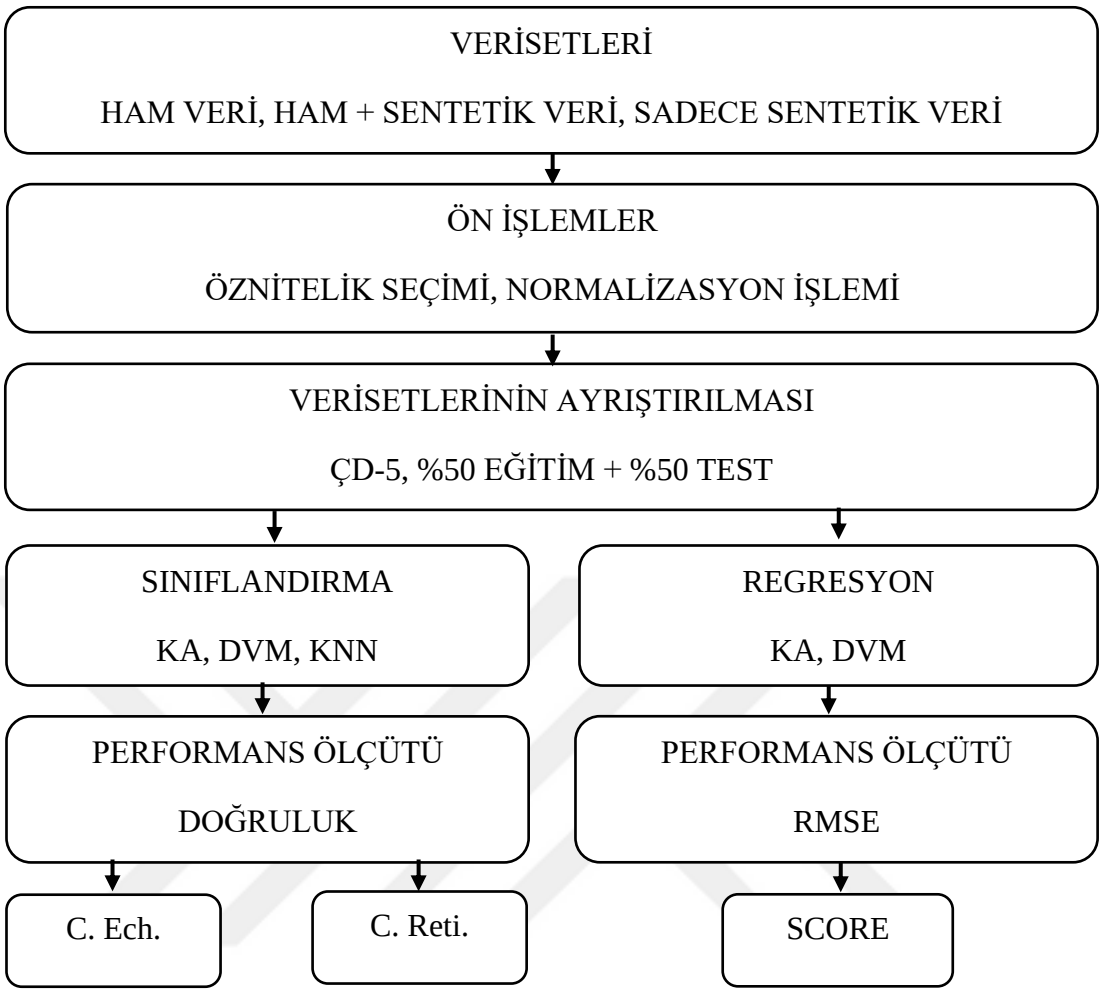
## 3.2. Yöntem

### 3.2.1. Verinin hazırlanması

Veri seti sınıflandırma ve regresyon yöntemlerinde işleme alınmadan önce özellik seçimi için nohut türlerini ayırt etmede ve verim değerleri için etkili olan özellikler seçilmiş geriye kalan özellikler yok sayılmıştır. Veri setinin normalize edilmesi ile değerler 0.1-0.9 arasına getirilerek daha küçük boyutlara getirilmiştir.

Normalizasyon işlemi için D min. max. normalizasyon yöntemi kullanılmıştır. Ön işlemler ile kullanılan yöntemin başarı oranı yükseltilmiş ve ayrıca gereksiz bilgiler ayıklanarak değerlendirmede olası hatalar önlenerek verilerin işlenmesi kolaylaştırılmıştır.

Elde edilen veri setinde 64 örneğin; 41 tanesi C. ech. ve 23 tanesi ise C. reti. nohut türünden oluşmuştur. İlk olarak ham verilerden 64 örnek kullanılarak sınıflandırma işlemi yapılmış ve nohut tohum tür tahmini yapılmıştır. Denetimli bir öğrenme algoritması olan sınıflandırma, örnekleri önceden gruplandırılmış bir eğitim veri kümesine göre kategorilere ayırma sürecine denir. Daha sonra aynı veri setine regresyon işlemi uygulanmıştır. Regresyon iki veya daha çok değişken arasındaki ilişkinin doğrusal olarak fonksiyon şeklinde gösterimi ve istatiksel olarak ifade edilmesine denir (Koyuncu ve Gök, 2019). Bağımlı değişkenler ile bağımsız değişkenler arasındaki ilişkiyi belirleyen regresyon, basit doğrusal, çoklu doğrusal ve doğrusal olmayan gibi çeşitli varyasyonları içerir. Her iki gruba ait nohut tohum türlerinin örnek sayıları eşit olmadığı için değerlendirmede olası hataların önüne geçmek ve başarı oranını yükseltmek için az olan C. reti. tohum türü için sentetik veriler oluşturulmuştur. C. ech. örnek sayılarına yakın sayıda olacak şekilde 23 tane C. reti. sentetik verisi üretilerek toplam 87 örnek değerlendirmeye alınmış tekrar sınıflandırma ve regresyon işlemleri yapılmıştır. Bir başka veri seti deneme işlemi de sadece sentetik verilerden oluşan, C. ech. için 41 yeni örnek, C. reti. için 46 yeni örnek olacak şekilde toplam 87 örnek işleme alınmıştır. Sentetik veriler için synthetic minority oversampling technique (SMOTE) yöntemi kullanılmıştır. Verilerin eğitim ve test için ayrıştırılmasında ÇD-5 ve %50-%50 kullanılmış sınıflandırma yöntemleri için KA, DVM ve KNN yöntemleri kullanılmıştır. Başarı ölçütü olarak ise doğruluk kullanılmıştır. Sonuç olarak makine öğrenmesi yöntemleri ile sınıflandırma işlemi yapılarak C. ech. ve C. reti. nohut türleri ayrılmıştır. Regresyon işlemi ile dekar başı verimlilik tahmini yapılmıştır. Çalışma akış diyagramı Şekil 3.1.'de verilmiştir.



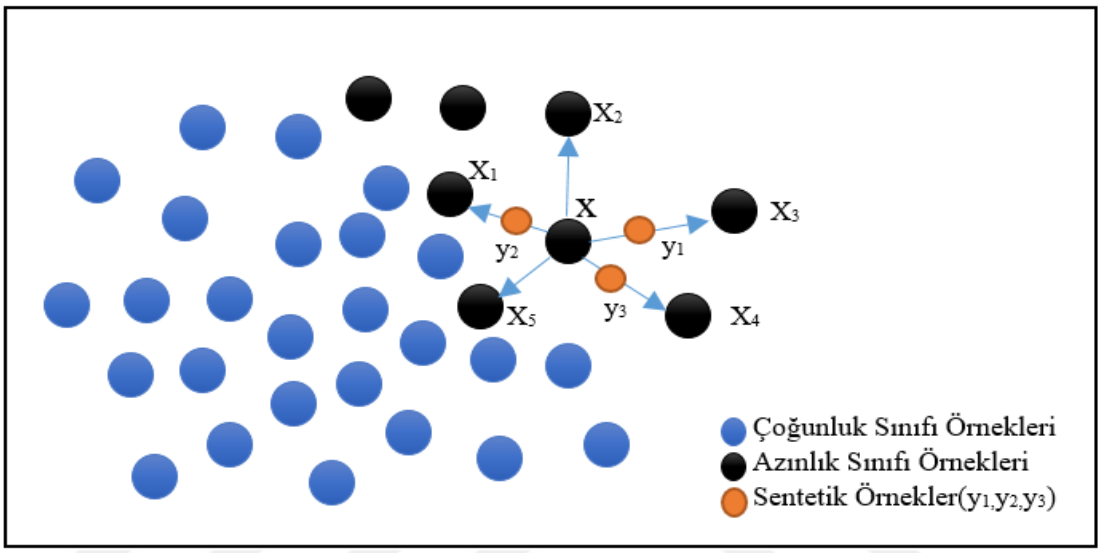

Şekil 3.1. Çalışma akış diyagramı

### 3.2.2. Sentetik veri üretme

Dengesiz veri problemleri için yeniden örnekleme yöntemleri kullanılır. Bu yaklaşım ile aşırı örnek sayısı veya az örnek sayısına sahip verilerde kullanılır. Bazı işlemeler ile fazla olan örneklerden azaltılarak, az olan örneklerden de artırılarak yeni ve dengeli sınıflardan oluşması sağlanır. Sentetik veri oluşturmak için SMOTE yöntemi kullanılmıştır.

Gerçek veriler üzerinde belirli işlemleri gerçekleştirerek yeni veri üretilmesine sentetik veri üretme denir. En yakın komşulardan rastgele seçim yapılır. Sentetik örnekler şu şekilde oluşturulur: İncelenen öznitelik vektörü ile en yakın komşusu

arasındaki farkı alınıp 0 ile 1 arasında herhangi bir sayı ile çarpılması ile elde edilen sayı incelenen özellik vektörüne eklenir. Elde edilen bu değer seçilen iki değer arasında çizgi parçası boyunca rastgele bir noktanın seçilmesini sağlar. Bu yöntem az olan sınıfın değerlerinin gelen olmasını, aykırı veri oluşmamasını sağlar. Dezavantaj olarak gürültü oluşmasına sebep olur. Gürültü, sistemde iç içe geçmiş örnekler ile sistemin kafasını karıştırır (Sağlam ve Cengiz, 2019)




Şekil 3.2. SMOTE algoritmasının çalışma sistemi (Yavaş ve ark., 2020)

Sentetik örneklerin üretilmesi denklem (3.1)'de verilmiştir.

1. İncelenen öznitelik vektörü (Ei) ile en yakın komşusu (Ej) arasındaki farkı alınır
2. Bu farkı 0 ile 1 arasında rastgele bir sayı ($\delta$) ile çarpılır
3. Çıkan sonuç incelenen özellik vektörüne eklenir ve yeni örnek (Eyeni) oluşur (Yavaş ve ark., 2020).

$$E(\text{yeni}) = Ei + (Ei - Ej)x \quad (3.1)$$

### 3.2.3. Normalizasyon

Makine öğrenmesi algoritmalarını uygulamadan önce veri üzerinde belirli ön işlemler uygulanır bu işlemler ile veriler daha verimli ve uygun hale getirilebilir.

Normalizasyon işlemi, makine öğrenmesi algoritmalarını uygulamadan önce veriyi düzenleyip uygulamaya hazırlayan bir tekniktir. Normalleştirmenin amacı, veri kümesindeki sayısal değerlerini her değer aralığını bozmadan farklılıkları ortak bir ölçeğe uygun bir şekilde değiştirmektir. Literatürde ondalık ölçeklendirme, min.-max., sigmoid ve z-skor normalizasyon çeşitleri bulunmaktadır (Jayalakshmi ve Santhakumaran, 2011). Verilerin normalizasyon işlemi için D min. max. modeli kullanılmıştır.

### 3.2.3.1. D min. max. normalizasyonu

Bu modelde kullanılan (3.2) eşitliğin uygulanması ile verilerin tamamı 0,1 ile 0,9 arasına alınarak normalize edilmiştir. Elimizdeki veriler bu işlem ile çok küçük boyutlara getirilmiş olur. Bu yöntem için (3.2) numaralı eşitlik kullanılır (Yavuz ve Deveci, 2012).

$$X' = 0.8 * (Xi - Xmin)/(Xmax - Xmin) + 0.1 \quad (3.2)$$

X´ = Normalize edilecek veriyi,

Xi = Girdi değerini,

Xmin = Veri seti içerisinde yer alan bölümün en küçük sayısı,

Xmax = Veri seti içerisinde yer alan bölümün en büyük sayısını, göstermektedir.

Veri sınıflandırma işlemlerinde birçok farklı yöntem kullanılmaktadır. Yaptığımız çalışmada verilerin sınıflandırması için verilerin özelliklerine, değerlerine ve performans ölçülerine bakıldığı zaman KA, DVM ve KNN yöntemlerinin daha uygun olduğu ve başarı oranlarının yüksek olduğu anlaşılmış ve bu yöntemler kullanılmıştır.

### 3.2.4. Veri setlerinin eğitim ve test için ayrıştırılması

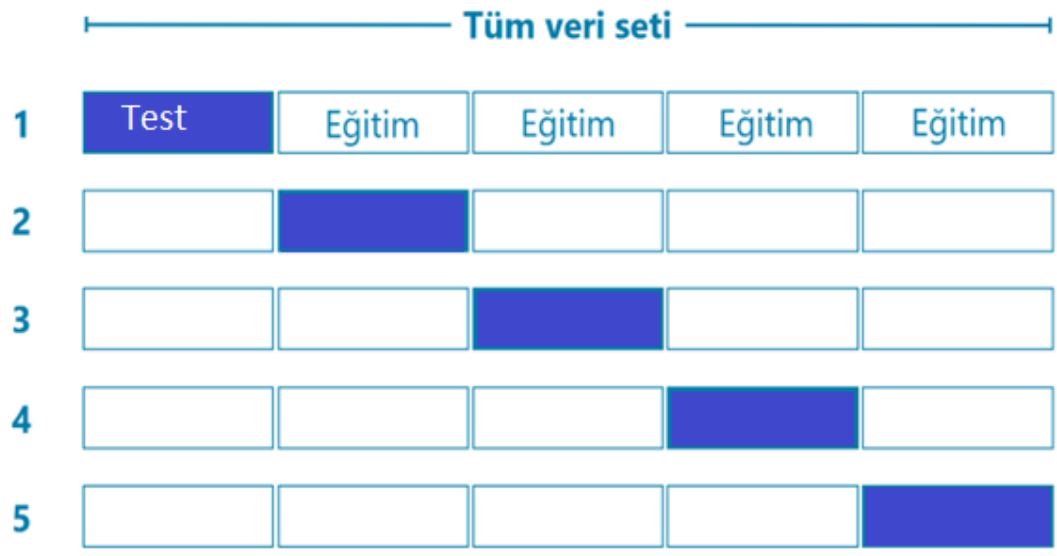
Modelin eğitildiği veri kümesine eğitim, eğitim kümesinde geliştirilen modeli değerlendirmek için kullanılan veri kümesine test kümesi denir. Veri seti eğitim ve test için ayrıştırılır uygun olan makine öğrenme yöntemin bulunabilmesi için eğitime ayrılmış veri seti üzerinde olabildiğince çok yöntem kurulup denenmelidir. Veri

setine uygun makine öğrenme yöntemi bulunduktan sonra doğrulama işlemi ile yöntemin uygunluğu karar verilir. İlk aşama veri ayıklama ve özellik seçimi ikinci aşama ilk aşamadaki veri seti için uygun olan makine öğrenme yöntemlerinin hangisi olduğunun tespit edilmesidir.

Uygun makine öğrenme yöntemi tespit edildikten sonra eğitim için ayrılan veriler dışında kalan test kümesi için ayrılmış veriler üzerinde denenir ve performans değerlerine bakılır. Test işleminde doğru ve yanlış tahmin olan doğruluk oranları için doğru pozitif ile yanlış pozitif ve doğru negatif ile yanlış negatif örnek sayılarının değerlerine bakılarak performans ölçütü olan doğruluk yüzdeleri ölçülür.

### 3.2.4.1. Çapraz doğrulama 5 (ÇD-5)

Sınıflandırma modellerinin değerlendirilmesi ve eğitilmesi için veri setini parçalara ayırma yöntemlerinden biridir.




Şekil 3.3. Çapraz doğrulama 5 uygulanması

Veri setinin %80'ini eğitim için, %20'sini de test için ayırarak başarısını değerlendirmektedir. Veriyi 5 parçaya bölerek, her bir parçanın hem eğitim hem de test için kullanılmasını sağlar. Bu işlem ile veri setinin hepsi hem eğitim hem de test verisinde kullanılır.

#### 3.2.4.2. %50 eğitim %50 test (%50-%50)

İşleminde ise veriyi iki parçaya ayırarak yarısını eğitim yarısını test için ayrılarak uygulanır başarısı değerlendirir.

### 3.2.5. Makine öğrenmesi

Makine öğrenmesi, elde bulunana verileri kullanarak sonucun ne olduğunu tahmin eden ve sonucu belli olmayan veriler için tahminde bulunan bu işlemleri matematiksel ve istatistik yöntemlerle yapan değerler dizisidir (Akay ve ark., 2018).

Makine öğrenmesi farklı algoritmalar ile bilgisayarların öğrenmesini sağlarlar. Bilgisayarlar, öğrendikleri verileri bazı işlemler ile yeni gelecek veriler için sonuç tahmini yapabilir verileri sınıflandırabilir bunu çeşitli fonksiyon ve istatistik programlar ile yaparlar. Verinin tipine göre metotların başarı oranları değişmektedir. Bu sebepten bir yöntemin diğerlerinden daha başarılı olduğunu söylenemez. Verinin tipine göre makine öğrenmesi yöntemlerinin başarı değerleri değişebilmektedir. Makine öğrenmesi ile büyük verinin işlenmesi, kayıt edilmesi ve analizinin yapılması ile işlem kolaylığı ve zamandan tasarruf sağlarlar. Sonuç değerlerini tahmin etmesi ve karar vermemize yardımcı olmaları makine öğrenmesinin değerini her geçen gün artırmaktadır.

Makine Öğrenmesi genel olarak denetimli, denetimsiz ve pekiştirmeli öğrenme olmak üzere üç sınıfa ayrılmaktadır (Şenol ve Timor, 2013).

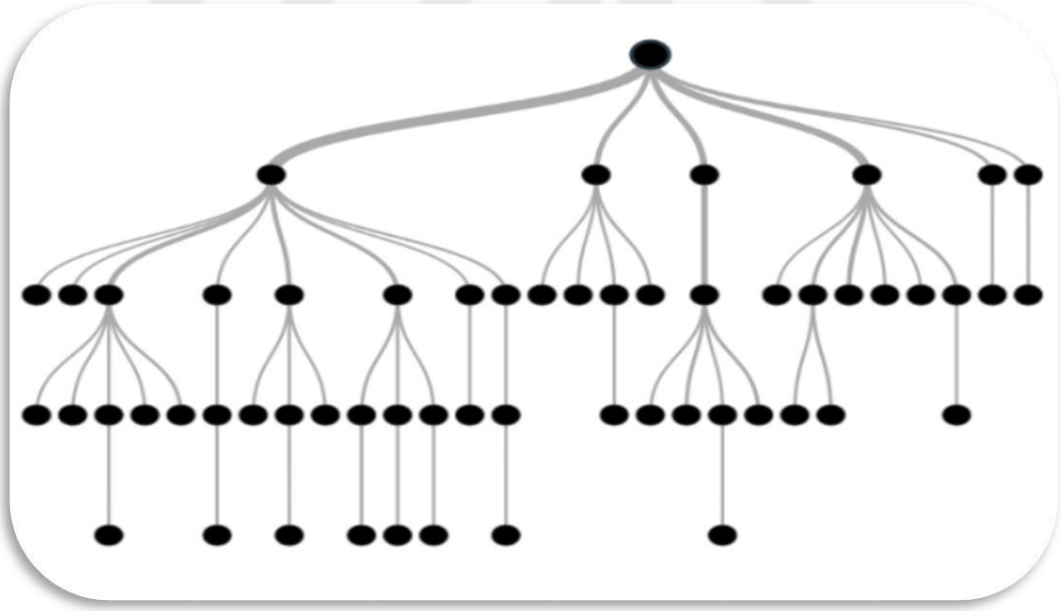
Denetimli öğrenme eğitim verisinin sonuç bilgisi bulunmaktadır. Sonucu belli olan veriler üzerinde çalışan metotlara denir (Kutlugün, 2017). Bu yöntem ile oluşturulan model sonuç bilgisi mevcut olan verilerin sonuçlarının tahmin edilmesi amaçlanmaktadır.

Denetimsiz öğrenme yöntemi ise sonuç bilgisi yoktur. Veriler arasında bir bağlantı kurulması gruplama yapılması amaçlanmaktadır (Kutlugün, 2017).

### 3.2.5.1. Karar ağaçları (KA)

Büyük miktarlardaki verileri basit adımlar uygulanarak, toplu halde olan veri kümesini parçalayarak çok küçük verilere bölen bir yapıdır. Bir karar ağacının yapısı incelendiğinde tüm düğümler bir özelliği test ederek dala geçer dallar ise ana düğümün test etmesi sonucuna karşılık gelir ve dallardan sonra tekrar test olan veri sınıf etiketi belli olur. Karar ağaçları kök, dal ve yapraklardan oluşur. Kökünde başlayan test işlemi dallara ve en son yaprakla sona erer. Kök en baştaki yapı verinin ham halidir, yaprak sınıf etiketinin belli olduğu kendisinden sonra herhangi bir yapı gelmeyen düğümdür. Düğümlerde veri kaybı olmadan veri dallara doğru ayrılır (Özekes, 2003)

Karar ağaçlarının çalışma prensibi öncelikle eğitimin gerçekleştiği küme sonra alt kümelere ayrılır. İşlem tahmin bitene kadar tekrarlanır (Aydın, 2018). Basit bir karar ağacı şeması Şekil 3.4.'teki gibidir.



Şekil 3.4. Karar Ağacı Yapısı (Ülgen, 2017)

Karar ağaçlarının çalışma prensibine bakıldığı zaman bir düğümü iki veya daha fazla alt düğüme bölmek için birden fazla algoritma kullanılır. Homojenliği arttırmak

için alt düğümler oluşturulur. Diğer bir deyişle, düğümün homojenliğinin hedef değişkenlere göre arttığı söylenebilir.

Karar ağaçlarının birçok avantajı olması ile popüler hale gelmiştir. Basit olduğu için insanlar tarafından kolay yorumlayabilmesi, hızlı olmaları, çok boyutlu verileri işleyebilmeleri, veri dağılımında birkaç özellik yapıldığı için çok çeşitli veri dağılımlarını modelleyebilmesi, doğruluğunun ve başarısının iyi olması, tahmini değerler yapraklardan köküne geriye doğru izlenebilmesi, yeni bir veri veya sorgunun sonucunu tahmin edilebilir olması, Karar ağacının dezavantajları da şöyledir Aşırı uyumsuzluk içerebilmesi, dikdörtgen bölme ile sınıflandırması, sayısal olmayan verileri kolaylıkla ele almaması, veri özelliklerine göre oluşturulan ağacın oldukça büyük ve karmaşık olması, ezbere öğrenme olayı gerçekleşebilir (Aksoy, 2017).

Seçeceğimiz algoritma değişkenin ve verinin tipine ve sayısına göre değişmektedir. Karar ağaçlarında en çok tercih edilen algoritmalara bakıldığı zaman en küçük kareler yöntemi, kategorik değişkenler için Entropi, Gini ve Sınıflandırma Hatası yöntemi biçimindedir. Entropi bir veri kümesinin sadece bir sonucu varsa düşük bir entropiye sahip olduğu söylenebilir. Dolayısıyla veriler, entropiyi en aza indirecek bir biçimde bölünmelidir. Bölünmeler iyi yapılması, tahminlerin de iyi olacağı anlamına gelmektedir.

$$H = -\sum p(x).\log p(x) \tag{3.3}$$

Eşitlik (3.3)'te p(x) bir sınıfın grup yüzdesini, H ise entropiyi göstermektedir.

İyi bir sınıf bölünmeyi belirlemek amacıyla bilgi kazancını kullanılmalıdır. Bilgi kazancının hesabı eşitlik (3.4)'te gösterilmiştir.

$$\text{Gain}(S, D) = H(S) - \sum_{E \epsilon D} \frac{|V|}{|S|} H(V) \tag{3.4}$$

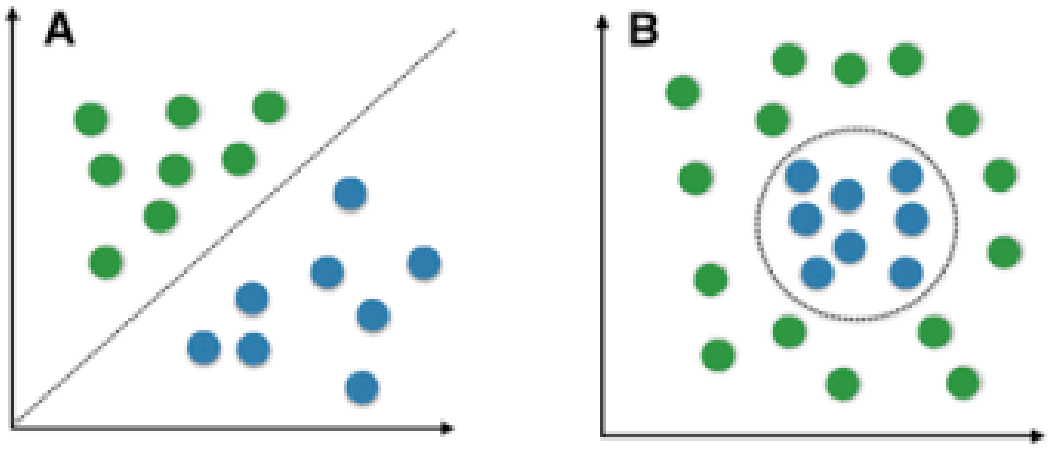
Eşitlik (3.4) de S ham veri kümesini, D ise kümenin bölünmüş bir parçasını ifade etmektedir.

V'nin tümü ayrıktır bir araya gelerek S'yi oluşturmaktadır. Bilgi kazancı, ham veri setinin entropisi ile her özniteliğin entropi değeri arasındaki farka eşittir (Doğan ve ark., 2019).

### 3.2.5.2. Destek vektör makineleri (DVM)

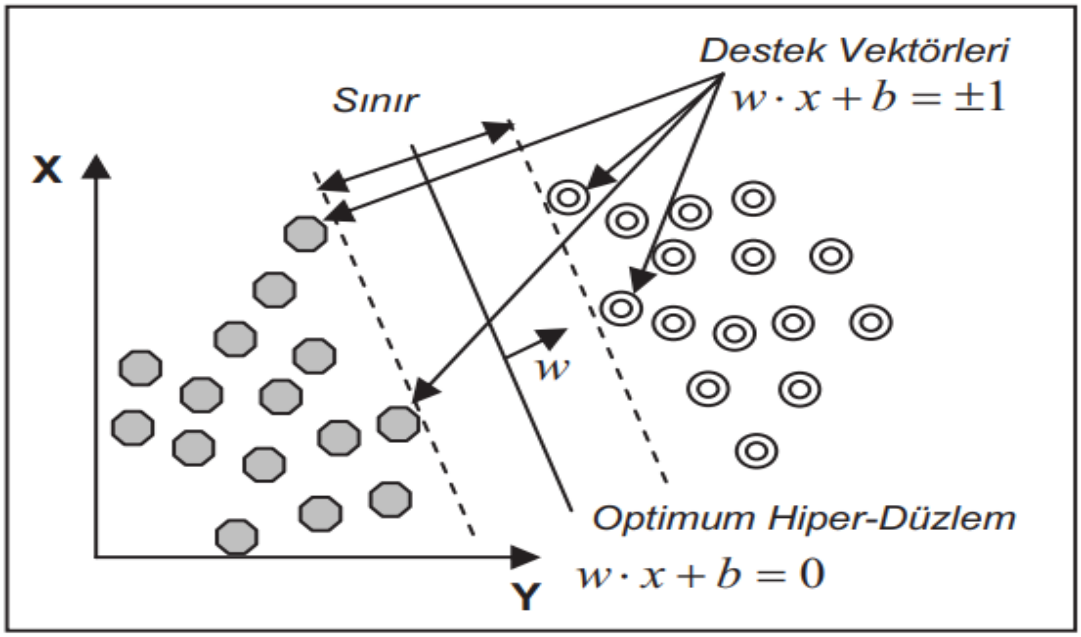
Genel anlamda istatiksel öğrenme teorisine dayalı olarak geliştirilmiş olan bir eğitimli sınıflandırma algoritmasıdır (Seçkin, 2011). Veri madenciliği algoritmaları arasında en dayanıklı ve en doğru metotlardandır (Cortes ve Vapnik, 1995). Destek vektör makineleri orijinal olarak Vapnik tarafından 1990 yıllarında geliştirilmiştir (Cortes ve Vapnik, 1995). Verileri sınıflandırma işlemi için oldukça basit ve kullanışlı bir yöntemdir. Sınıflandırma için düzlem üzerinde bulunan iki ayrı sınıfı belirleyip ayırmak için aralarında bir sınır çizgisi çekilmesi ile ayırmak mümkündür. Çizgi çekilirken dikkat edilmesi gereken konu iki sınıfın elemanlarına eşit uzaklıkta olmalıdır. Destek vektör makineleri eşit uzaklık çizginin nasıl çizileceğini belirler. DVM hem lineer hem de lineer olmayan veri kümelerini sınıflayabilir.

Sınıflar doğrusal olarak veya doğrusal olarak ayrılamayan olarak iki sınıfa ayrılır. Doğrusal olarak ayrılamayan veriler yüksek boyutlu uzaya taşınarak sınıflandırma yapılır.




Şekil 3.5. A) Doğrusal ayrılabilen veriler B) Doğrusal olarak ayrılmayan veriler (Akar, 2013)

Destek vektör makinelerinin avantajlarına bakıldığı zaman doğrusal veya doğrusal olmayan verilerde kullanılabilmesi, regresyon, aykırı değer tespiti, yüz tanıma sistemleri ve ses analizine kadar birçok sınıflandırma probleminde kullanılması, yüksek boyutlu kümelerde etkili olması, belleğin verimli şekilde kullanılması, çok sayıda değişkenle çalışmasına rağmen yüksek oranda doğruluk değerine sahip olması, karmaşık sınıfları modelleyebilmesi gibi özellikleri ile avantajlara sahipken büyük veri sayısına sahip kümelerde destek vektör makinelerinin eğitim işlemi yani öğrenmesi zaman alması dezavantajlarını göstermektedir (Doğan ve ark., 2019).




Şekil 3.6. Doğrusal olarak ayrılabilen veri setlerinde hiper-düzlemin belirlenmesi (Kavzoğlu ve Çölkesen, 2010)

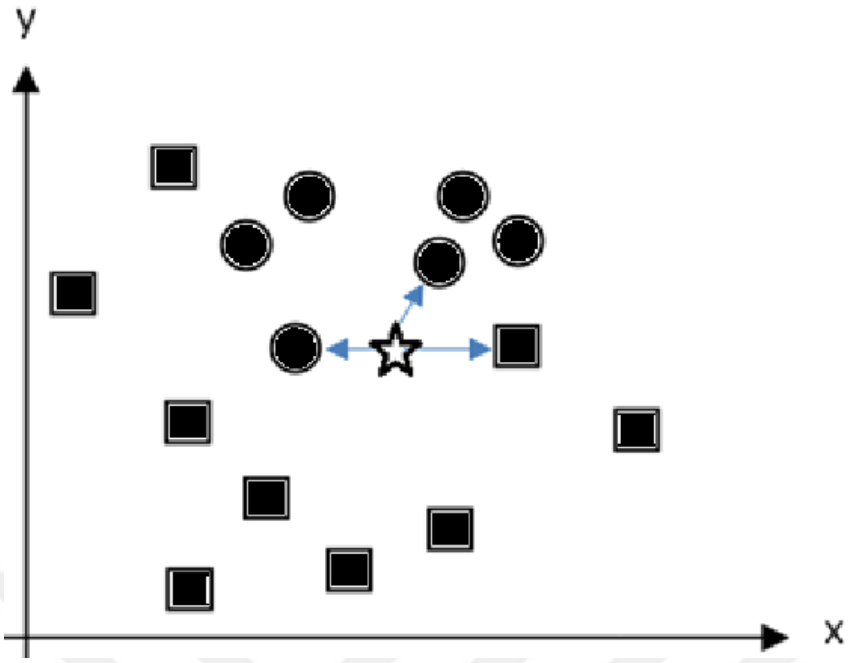
DVM en büyük değerli aralık sınıflandırıcılarındandır. $n$ boyutlu uzayda verileri sınıflandırabilmek için belirtilen bir ayırıcı fonksiyon eşitlik (3.5)'te verilmiştir.

$$F(x) = w.xi + b = 0 \quad (3.5)$$

(3.5) eşitliğinde verilen fonksiyonda, w ağırlık vektörünü olan değerin bulunması ile gerçekleşir. B ise eğilim değerini gösterir. N tane eğitim verisi olan bir kümede w ve b değerleri |w.xi+b|=1 ayırıcı düzleme yakın olan değerler seçilir. Veri setleri farklı sınıflardan oluştuğu için sonuç +1 veya -1 olabilmektedir. Ağırlık vektörü olarak bilinen w değerinin küçük olması istenir çünkü sınıfları ayıran düzlemlerin büyük olmasını ile sınıfların daha kolay ayrılmasını sağlar (Kavzoğlu ve Çölkesen, 2010).

### 3.2.5.3. K en yakın komşu (KNN)

K en yakın komşu tahmin yöntemi kullanımı kolay ve basit olması geniş alanda kullanımını sağlamıştır. Bir verinin kimliği kendisine en yakın komşu verinin kimlik sonucu ile aynı olur mantığıyla çalışır. k harfi kaç tane komşu sayısını olduğunu göstermektedir. Sınıflamada başarılı olmak için k sayısının değeri çok önemlidir. K sayısının değerini bulmak için çapraz doğrulama veya deneme yanılma yoluyla yapılmaktadır (Yiğiter ve ark., 2018). K en yakın komşu bir örneğin hangi sınıfa ait olduğunu bulurken seçilen örneğe en yakın sınıf sayıları k değerini seçer ve k'nın grup üyeliğine göre örneğin hangi sınıfa benzediği ve o sınıfa ait olduğuna karar verir. K en yakın komşu yönteminin avantajlarına bakıldığı zaman eski ve basit olan gürültülüye sahip verilere karşı dirençli olması, genetik ve biyo-enformatiğin birçok alanında, veri madenciliğinde, bilgi güvenliğinin sağlanmasında, saldırı tespit sistemleri gibi birçok sistemde kullanılmaktadır (Pilavcılar, 2007). Fakat bunun yanında uzaklık hesabında büyük veriler kullanıldığı ve çok sayıda bellek alanına gereksinim duyulduğu için dezavantajı da mevcuttur (Bütüner ve Calp, 2020).




Şekil 3.7. KNN Çalışma şeması

Sınıf grubu daire olanlar kare olanlardan fazla ise sınıfı dairedir. k değeri seçerken tek sayı seçilerek komşu örneklerin sayısı eşit çıkması önlenmiş olur. k değeri çift sayı seçilirse k tane örnek için her bir sınıfa ait örnekler kendi aralarında toplanarak ortalamaları bulunur. Ortalaması en küçük olan sınıf, test edilen örneğe daha yakın olduğu için test örneğinin sınıfı en küçük ortalamaya sahip olan sınıf olur (Karadağ ve Taşaltın, 2016).

Hangi veri setine katılacağı belli olmayan örneğin kümedeki verilere uzaklığı ilgili fonksiyonlar yardımıyla hesaplanır.

K en yakın komşu yönteminde yeni örneğin komşu sınıflara olan uzaklığı bulmak için en çok kullanılan uzaklıklar öklid, manhattan, ve Chebyschev uzaklıklarıdır. Öklid uzaklığı denklem (3.6)'da tanımlanmıştır (Kabak ve ark., 2017).

$$di,j = \sqrt{[\,(xi1 - xj1\,)\,2 + (xi2 - xj2\,)\,2 + \cdots + (xip - xjp)\,2]} \qquad (3.6)$$

Manhattan uzaklığı denklem (3.7)'de tanımlanmıştır (Kabak ve ark., 2017).

$$\mathrm{di,j} = (|\mathrm{xi1} - \mathrm{x1}| + |\mathrm{xi2} - \mathrm{xj2}| + \cdots + |\mathrm{xip} - \mathrm{xjp}|) \quad (3.7)$$

Chebyshev uzaklık ölçüsü (max. değer uzaklığı) iki nesnenin farkının mutlak değerine denir. Denklem (3.8)'de tanımlanmıştır (Kabak ve ark., 2017).

$$\mathrm{di,j} = \mathrm{maxm}|\mathrm{xi1} - \mathrm{xj1}| \quad (3.8)$$

### 3.2.6. Performans metrikleri

Bir sınıflandırma problemi için bir model oluşturulduğunda modelin başarısı bakıldığı zaman gelen doğru tahminlerin sayısı gibi düşünülür. Ancak bu bilgi sadece sınıflandırmanın doğruluğunu vermektedir. Bir modelin yeterince iyi olup olmadığına karar vermek için tek başına sınıflandırma doğruluğu genellikle yeterli bilgi değildir.

Karmaşıklık matrisi: Sınıflandırıcının tahmin sonuçlarını sunmanın açık ve net bir yolu, bir karmaşıklık matrisi kullanmaktır. 4 parametresi vardır.

Doğru Pozitif (DP): Doğru tahmin edilen pozitif değerlerdir.

Doğru Negatif (DN): Doğru tahmin edilen negatif değerlerdir. Doğru tahminler ile gerçek sınıfın değeri ve tahmini sınıfın aynı olduğunu gösterir.

Yanlış Pozitif (YP): Gerçek sınıfınız ve tahmin edilen sınıfla uyuşmadığı ortaya çıkar.

Yanlış Negatif (YN): Gerçek sınıfınız ve tahmin edilen sınıfla uyuşmadığı ortaya çıkar (Başer ve ark., 2021).

Doğru tahmin edilen pozitif ve negatif sayıların artması istenirken, yanlış tahin edilen pozitif ve negatif sayılarının azaltılması ile sınıflandırma performansının artması sağlanır.

### 3.2.7. Başarı ölçütü

Tez çalışmasında nohut türlerinin sınıflandırılmasında başarı ölçütü olarak doğruluk kullanılmıştır. Doğruluk; doğru tahmin edilen değerlerin toplam değerlere

oranıdır. Yüksek doğruluk varsa modelin en iyisi olduğunu düşünülebilir. Ancak yanlış tahminlerin farklı ve çok olduğu durumlarda modelin performansını değerlendirmek için diğer parametrelere bakılması gerekir.

$$Doğruluk = (DP + DN)/(DP + DN + YP + YN) \quad (3.9)$$

Tez çalışmasının ikinci aşamasında verim tahmininde başarı ölçütü olarak rmse yöntemi kullanılmıştır. RMSE, gerçek değerlerle tahmin edilen değerler arasındaki farkın standart sapmasıdır. Denklem (3.10)'da gösterildiği gibi hesaplanmaktadır. Tahmin edilen değerlerin veri noktalarından ne kadar uzakta olduğunun bir ölçüsüdür; RMSE değeri 0'dan sonsuza kadar değişebilir. RMSE değerinin sıfır olması modelin hiç hata yapmadığı, sayı değeri sıfıra ne kadar yakınsa o kadar iyi olduğu anlamına gelmektedir (Chai ve Draxler, 2014).

$$RMSE = \sqrt{[1/\,\mathrm{n}\ \sum_{i=1}^{n}(\mathrm{yi}\ -\ \mathrm{yi}')^{2}]} \quad (3.10)$$

yi:gerçek değer,

yi': tahmin edilen değer

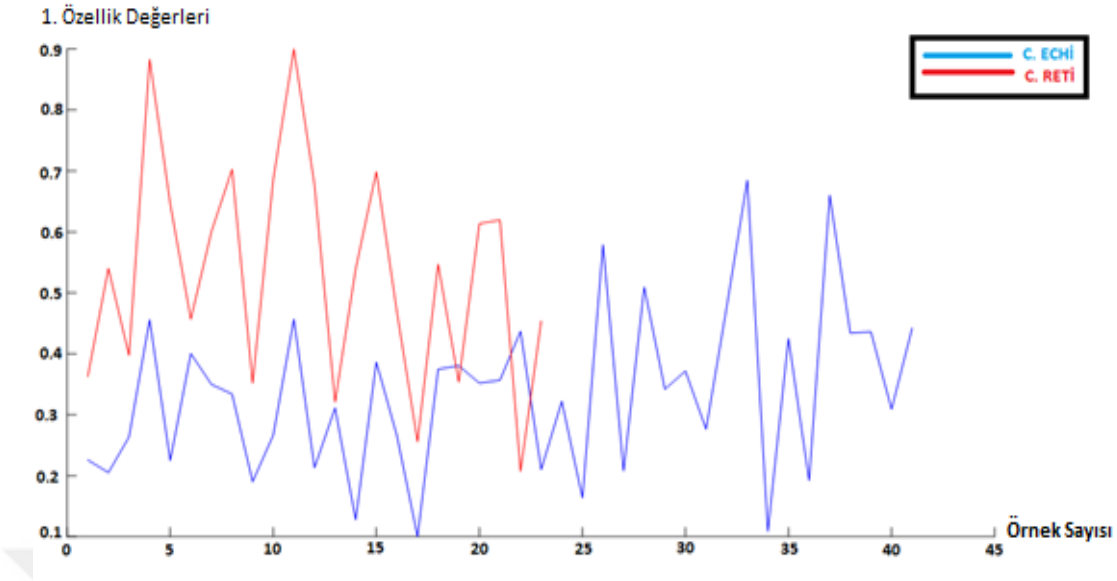
n: ise gözlem sayısı değerlerini göstermektedir.

# 4. ARAŞTIRMA BULGULARI ve TARTIŞMA

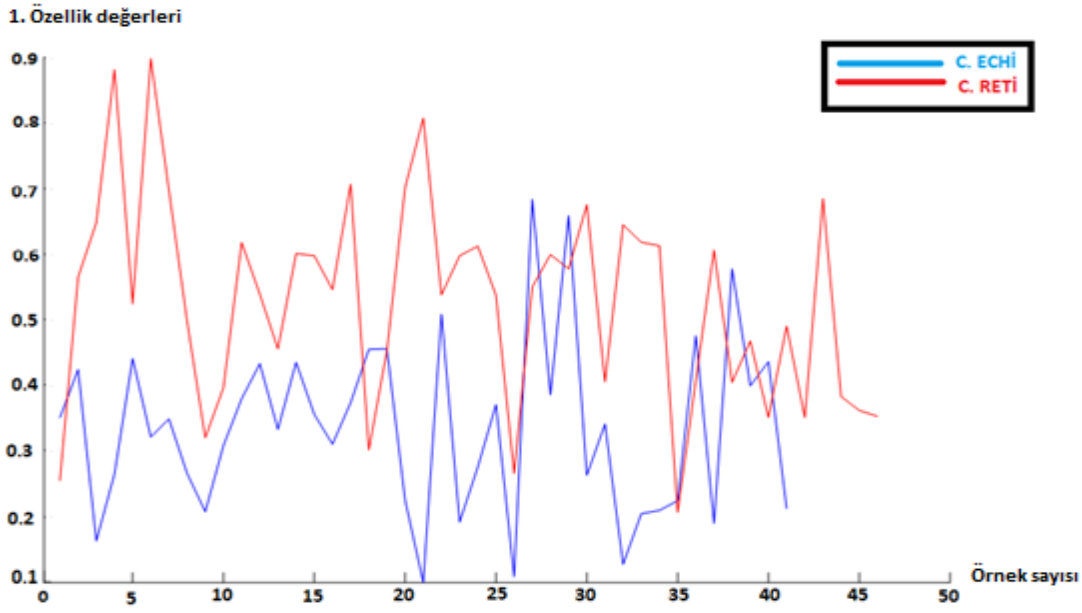
Tez çalışmamızda kullandığımız nohut özelliklerine ait değerler iki farklı uygulama olarak ele alınmıştır. Birinci uygulama için, Harran Üniversitesi Ziraat Fakültesi deneme arazileri kullanılarak ekilen C. ech. ve C. reti. nohut türlerinden elde edilen veri setleri kullanılarak Sınıflandırma işlemi yapılmış, ikinci uygulama için ise yine aynı veri setleri kullanılarak regresyon işlemi yapılmıştır. Veri setlerinin kullanımı üç şekilde gerçekleştirilmiştir. İlk olarak sadece 64 örnekten oluşan ham veriler kullanılmış, ikinci aşamada ise ham verilere ilave olarak diğer nohut türüne nispeten daha az örneğe sahip olan nohut türünden sentetik veriler üretilerek toplam 87 örnek kullanılmıştır. Üçüncü ve son kullanım şekli olarak her iki nohut türüne ait sadece sentetik verilerin yer aldığı 87 örnek kullanılmıştır. Veri setlerinin eğitim ve test için ayrıştırma işlemi 5 çapraz doğrulama (ÇD-5) ve veri setinin %50'si eğitim, %50'si test (%50-%50) işleminde kullanılacak şekilde yapılmıştır. Sınıflandırma yöntemleri olarak KA, DVM ve KNN kullanılmış, regresyon yöntemleri olarak ise KA ve DVM yöntemleri tercih edilmiştir. Performans ölçütleri olarak sınıflandırma için doğruluk, regresyon için rmse ölçütleri kullanılmıştır.

## 4.1. Verilerin Grafiksel Gösterimi

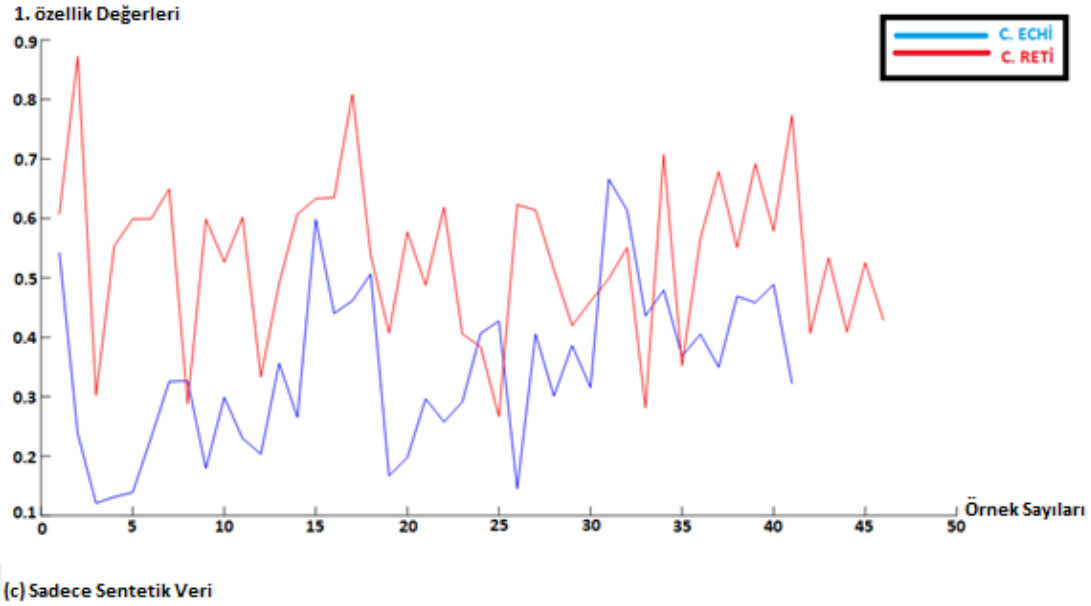
Veri seti sınıflandırmaya alınmadan önce grafiksel gösterimi incelenmiştir. Şekil 4.1. (a) Sadece ham verilerin olduğu 41 adet C. ech. ve 23 adet C. reti. nohut türü verisinden oluşmaktadır. C. ech. nohut türünün örnek sayısının C. reti.'nin örnek sayısından fazla olması grafikteki kırmızı rengin daha erken bitmesine sebep olmuştur. Şekil 4.1. (b) ham veriler ile birlikte üretilen 23 adet C. reti. sentetik verinin eklenmesi ile oluşan 41 adet C. ech. ve 46 adet C. reti. nohut türü verisinden oluşmaktadır. Şekil 4.1. (c)' de ise sadece sentetik verilerden oluşan 41 adet C. ech. ve 46 adet C. reti. nohut türü verisinden oluşmaktadır.




(a) Ham veri




(b) Ham + Sentetik Veri




Şekil 4.1. C. ech. ve C. reti. grafiksel gösterimi

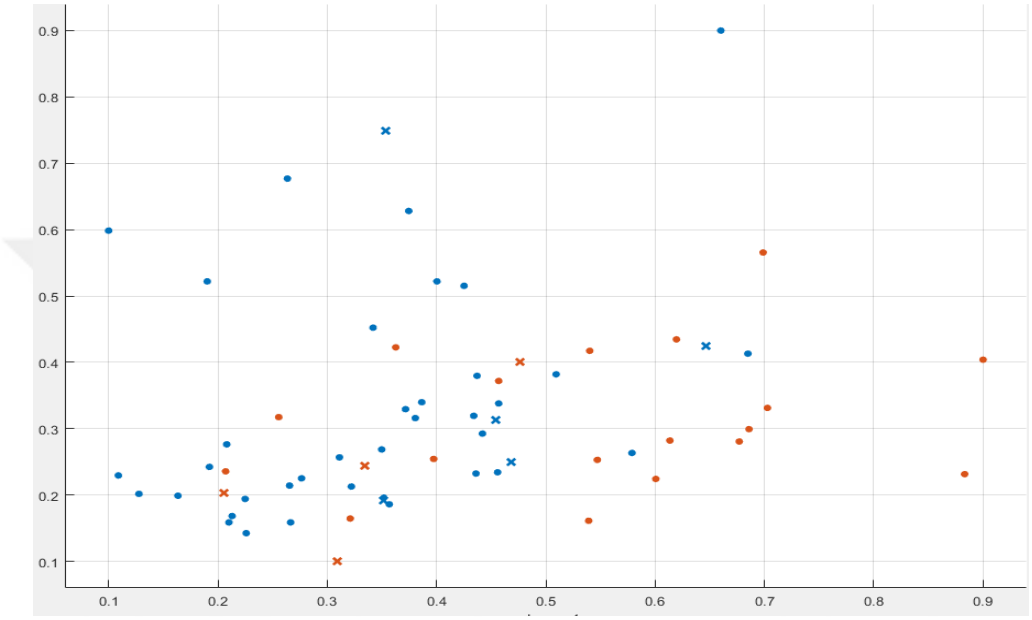
Şekil 4.1.'de verilen grafikler C. ech. ve C. reti nohut türlerine ait, örnek sayısı ve özellikleri göstermektedir. Mavi renk C. ech., kırmızı renk ise C. reti türünü göstermektedir. Yalnızca ham verilerin olduğu grafikler, ham verilerle birlikte sentetik verilerin olduğu grafikler ve yalnızca sentetik örneklerin olduğu grafikler arasında farklılıkların olduğunu gözlemlemek, iki tür arasında sınıflandırma veya regresyon işlemlerinin başarılı sonuçlar vereceğini göstermektedir.

## 4.2. Çalışmada Sınıflandırma Sonucu Elde Edilen Sonuçların Karşılaştırılması

Çalışmada kullanılan veri setlerinin örnek sayılarının farklı olması başarı oranlarını etkilemiştir. İlk olarak elimizdeki ham veriler kullanılmıştır. Örnek sayıların farklı olması başarı oranlarının düşük olmasına sebep olmuştur ikinci aşama olarak iki türün örnek sayılarını birbirine yakın olacak şekilde sentetik veri üretilmiş ve başarı oranlarının arttığı gözlemlenmiştir. Üçüncü aşama olarak elimizdeki verilerden tamamen farklı olan sentetik veriler üretilmiş ve başarı oranları gözlemlenmiştir.

İlk olarak sadece ham verilerden oluşan, verilerin eğitim ve test için ayrıştırılması işleminde ÇD-5 ve %50-%50 işlemleri kullanılarak KA, DVM ve KNN sınıflandırma yöntemleri ile elde edilen sınıflandırma sonuçları incelenmiştir.

Ham verilere ait KA sınıflandırma yöntemleri kullanılarak veri setinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımları Şekil 4.2.'de verilmiştir.



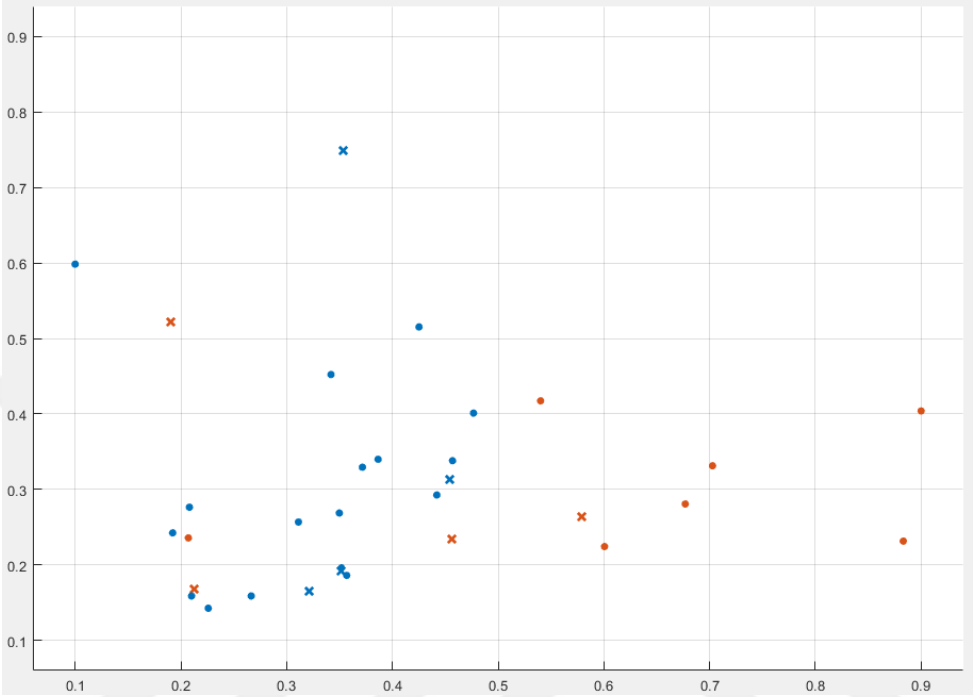

Şekil 4.2. Ham veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 41 adetten 37 tanesi doğru 4 tanesi yanlış yani çarpı (x) tahmin edilmiştir. C. reti. İse kırmızı renk olup 23 adetten 18 adet doğru tahmin edilip 5 tanesi yanlış yani çarpı (x) tahmin edilmiş sonuçlar Çizelge 4.1.'de verilmiştir.

Çizelge 4.1. Ham veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| KA | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 37 | 5 |
| C. reti. | 4 | 18 |

Ham verilere ait KA sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.3.'te verilmiştir.



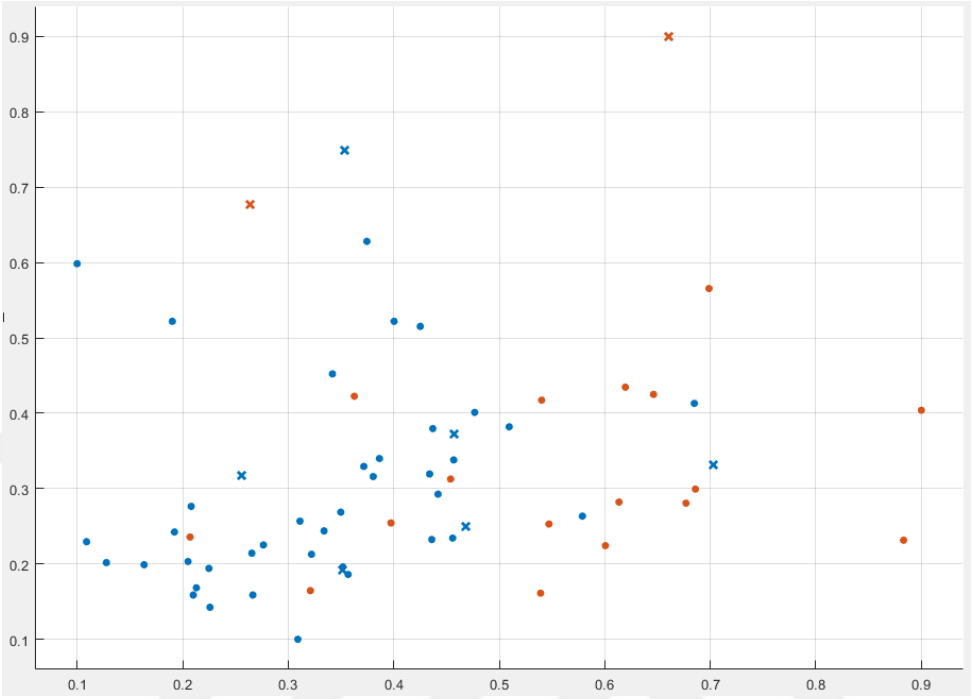

Şekil 4.3. Ham veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 17 tanesi doğru 4 tanesi yanlış tahmin edilmiştir. C. reti. İse kırmızı renk olup 7 adet doğru tahmin edilip 4 tanesi yanlış tahmin edilmiş sonuçlar Çizelge 4.2.'de verilmiştir.

Çizelge 4.2. Ham veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| KA | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 17 | 4 |
| C. reti. | 4 | 7 |

Ham verilere ait DVM sınıflandırma yöntemleri kullanılarak veri setinin ÇD-5 ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.4.'te verilmiştir.



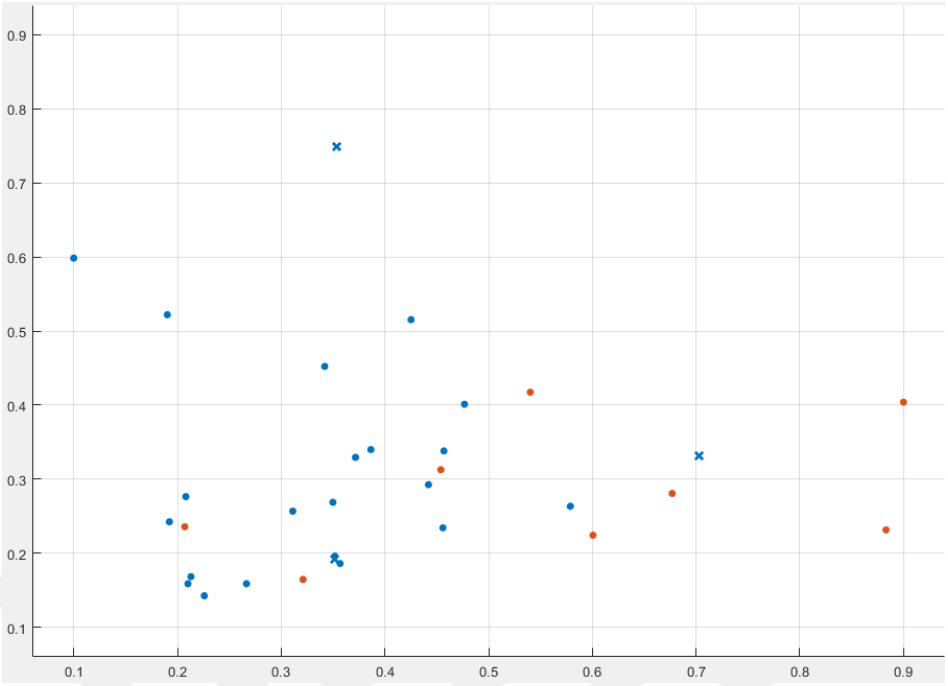

Şekil 4.4. Ham veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 39 tanesi doğru 2 tanesi yanlış edilmiştir. C. reti. İse kırmızı renk olup 17 adet doğru tahmin edilip 6 tanesi yanlış tahmin edilmiş sonuçlar Çizelge 4.3.'te verilmiştir.

Çizelge 4.3. Ham veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| DVM | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 39 | 6 |
| C. reti. | 2 | 17 |

Ham verilere ait DVM sınıflandırma yöntemleri kullanılarak veri setinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.5.'te verilmiştir.



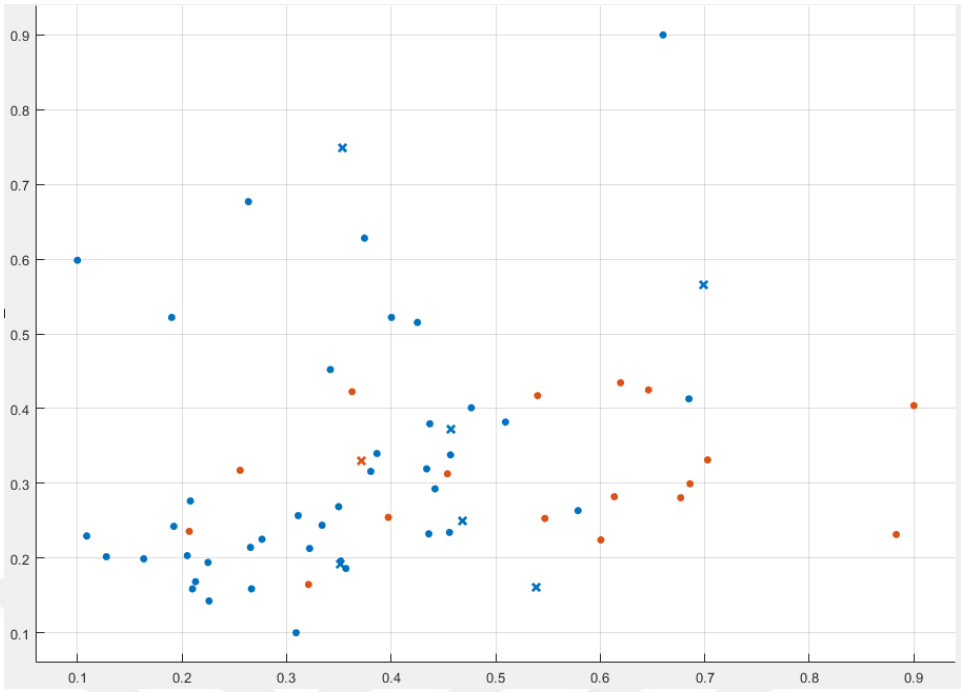

Şekil 4.5. Ham veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 21 tanesi doğru 0 tanesi yanlış tahmin edilmiştir. C. reti. İse kırmızı renk olup 8 adet doğru tahmin edilip 3 tanesi yanlış tahmin edilmiş sonuçlar Çizelge 4.4.'te verilmiştir.

Çizelge 4.4. Ham veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| DVM | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 21 | 3 |
| C. reti. | 0 | 8 |

Ham verilere ait KNN sınıflandırma yöntemleri kullanılarak veri setinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.6.'da verilmiştir.



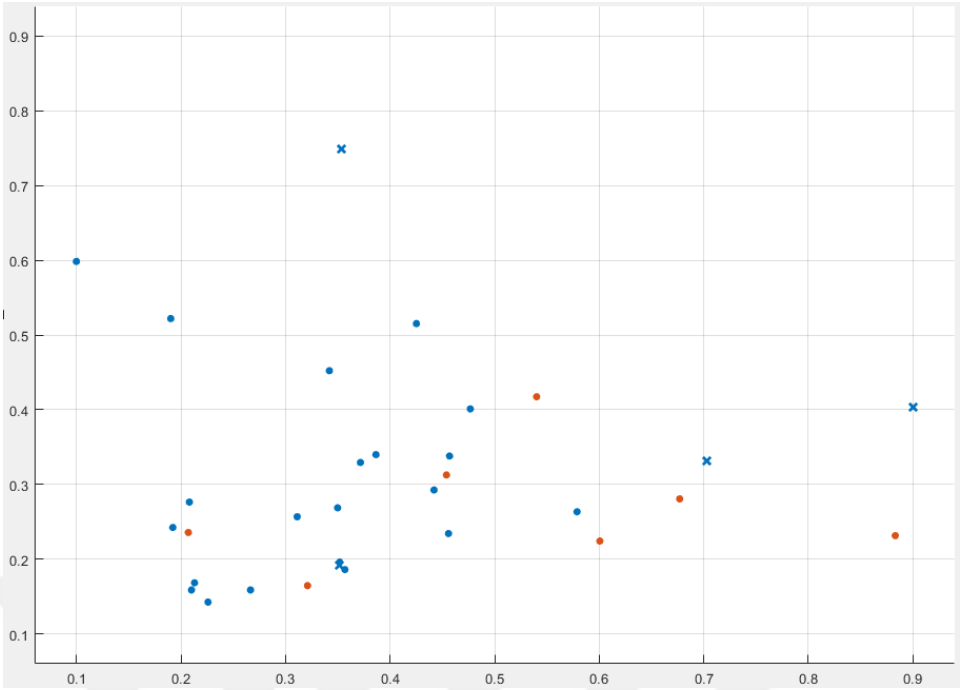

Şekil 4.6. Ham veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 40 tanesi doğru 1 tanesi yanlış tahmin edilmiştir. C. reti. İse kırmızı renk olup 17 adet doğru tahmin edilip 6 tanesi yanlış tahmin edilmiş sonuçlar Çizelge 4.5.'te verilmiştir.

Çizelge 4.5. Ham veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| KNN | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 40 | 6 |
| C. reti. | 1 | 17 |

Ham verilere ait KNN sınıflandırma yöntemleri kullanılarak veri setinin %50-%50 ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.7.'de verilmiştir.




Şekil 4.7. Ham veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 21 taneden hepsini doğru tahmin edilmiştir. C. reti. İse kırmızı renk olup 7 adet doğru tahmin edilip 4 tanesi yanlış tahmin edilmiş sonuçlar Çizelge 4.6.'da verilmiştir.

Çizelge 4.6. Ham veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| KNN | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 21 | 4 |
| C. reti. | 0 | 7 |

C. reti. ve C. ech. nohut türlerine ait sınıflandırma sonucu elde edilen grafikler, doğru ve yanlış sayıları doğrultusunda ham veri setine ait hesaplanan doğruluk oranları Çizelge 4.7.'de verilmiştir.

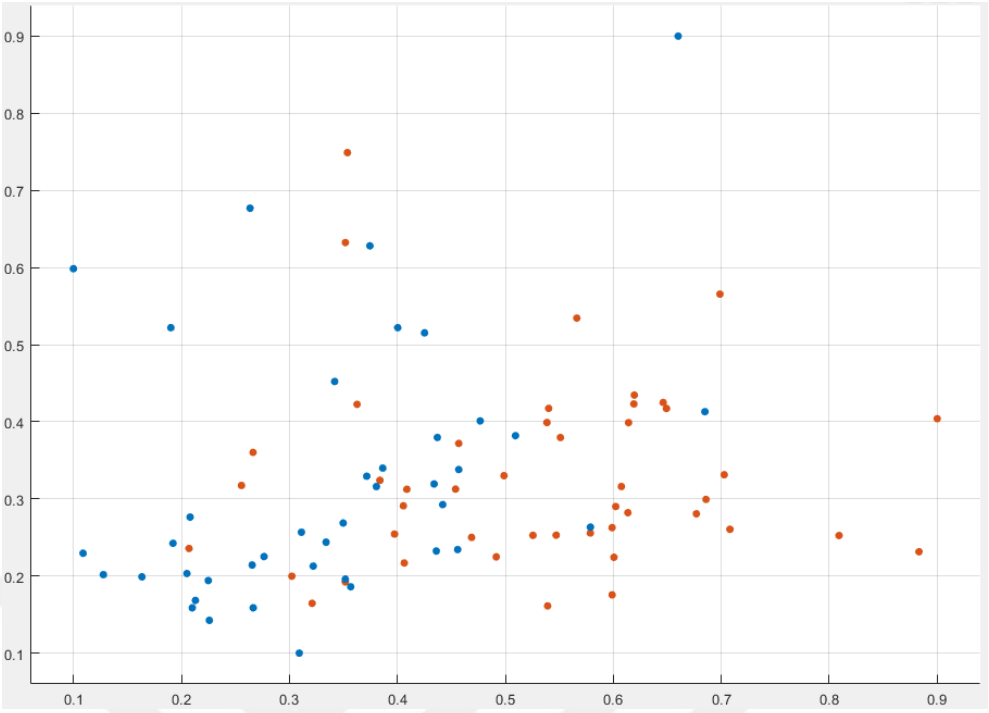
Çizelge 4.7. Ham veri (%) doğruluk oranları

| Kullanılan Yöntem | ÇD-5 | %50-%50 |
|---|---|---|
| Karar Ağaçları | 85.9 | 75 |
| Destek Vektör Makinaları | 87.5 | **90.6** |
| K En Yakın Komşuluk | 89.1 | 87.5 |

Elde edilen sonuçlar karşılaştırıldığında ham verilerden oluşan doğruluk değerleri incelendiğinde eğitim ve test verilerinin ayrıştırma işleminin %50-%50 şeklinde yapıldığı en yüksek başarı oranı %90.6 doğruluk yüzdesi ile DVM sınıflandırma yöntemi olmuştur.

İkinci aşama olarak ham verilere ilave olarak üretilen 21 adet C. reti. sentetik verisinin eklenmesi ile ham ve sentetik verilerinden oluşan, verilerin eğitim ve test için ayrıştırılması işleminde ÇD-5 ve %50-%50 işlemleri kullanılarak KA, DVM ve KNN sınıflandırma yöntemleri ile elde edilen sınıflandırma sonuçları incelenmiştir.

Ham + sentetik verilere ait KA sınıflandırma yöntemleri kullanılarak veri setinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.8.'de verilmiştir.



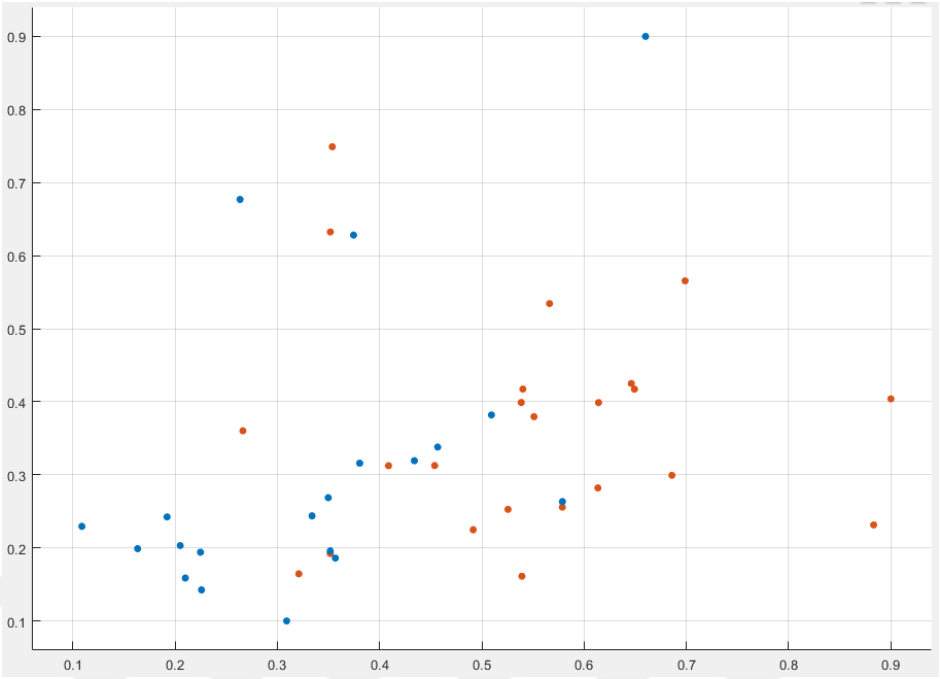

Şekil 4.8. Ham + sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 41 adetten hepsini doğru tahmin etmiştir. C. reti. İse kırmızı renk olup 46 adetten hepsini doğru tahmin etmiştir. Sonuçlar Çizelge 4.8.'de verilmiştir.

Çizelge 4.8. Ham + sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| KA | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 41 | 0 |
| C. reti. | 0 | 46 |

Ham + sentetik verilere ait KA sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.9.'da verilmiştir.



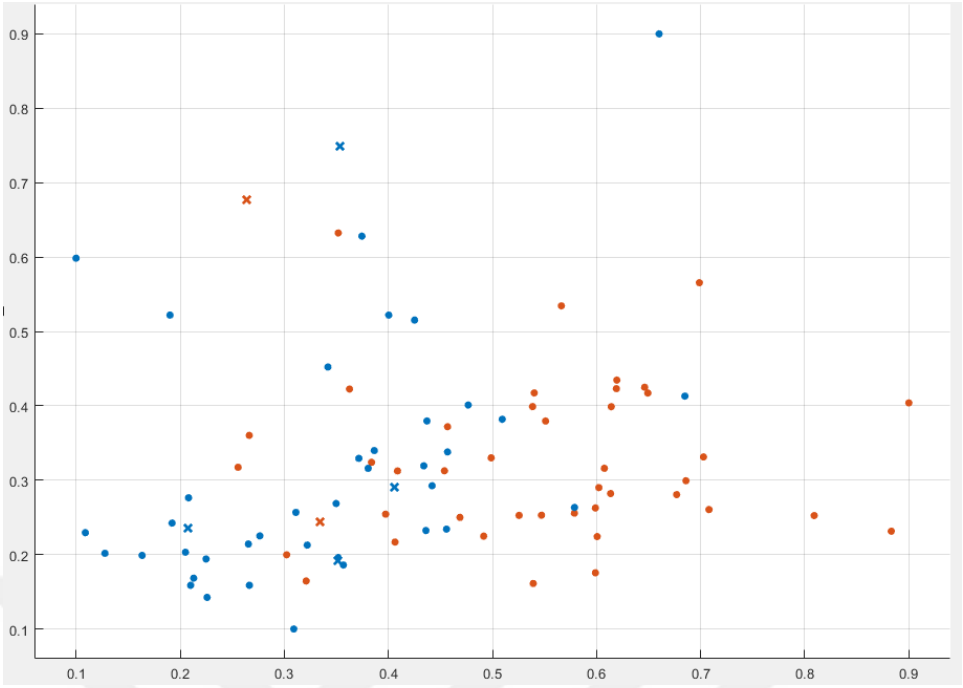

Şekil 4.9. Ham + sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 20 adetten hepsini doğru tahmin etmiştir. C. reti. İse kırmızı renk olup 23 adetten hepsini doğru tahmin etmiştir. Sonuçlar Çizelge 4.9.'da verilmiştir.

Çizelge 4.9. Ham + sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| KA | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 20 | 0 |
| C. reti. | 0 | 23 |

Ham + sentetik verilere ait DVM sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.10.'da verilmiştir.



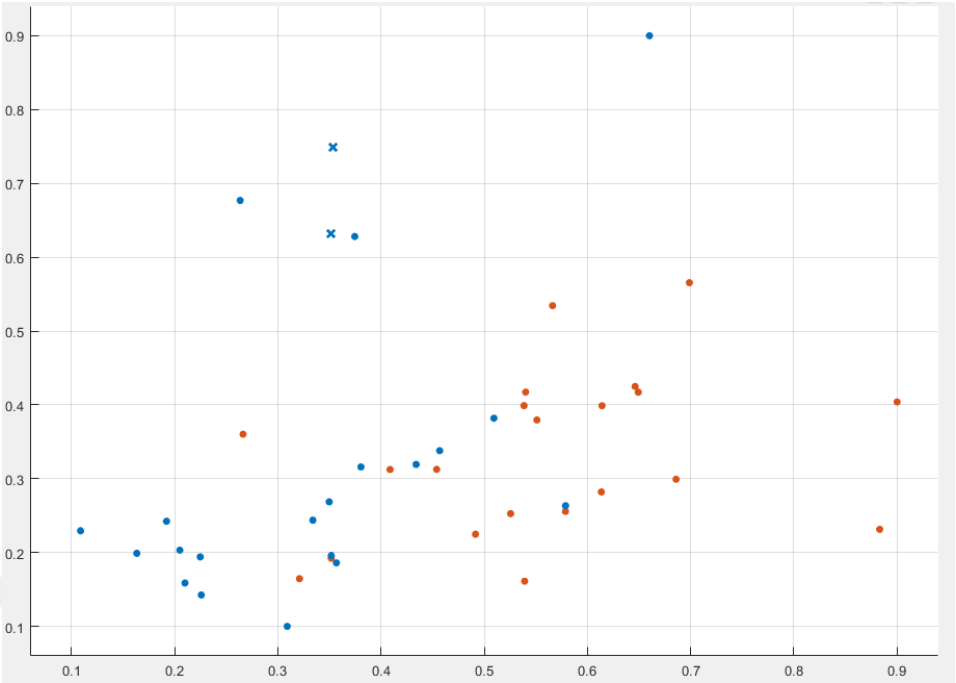

Şekil 4.10. Ham + sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 41 adetten 39 doğru 2 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 44 adetten 42 doğru 4 yanlış tahmin etmiş sonuçlar Çizelge 4.10.'da verilmiştir.

Çizelge 4.10. Ham + sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| DVM | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 39 | 4 |
| C. reti. | 2 | 42 |

Ham + sentetik verilere ait DVM sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.11.'de verilmiştir.



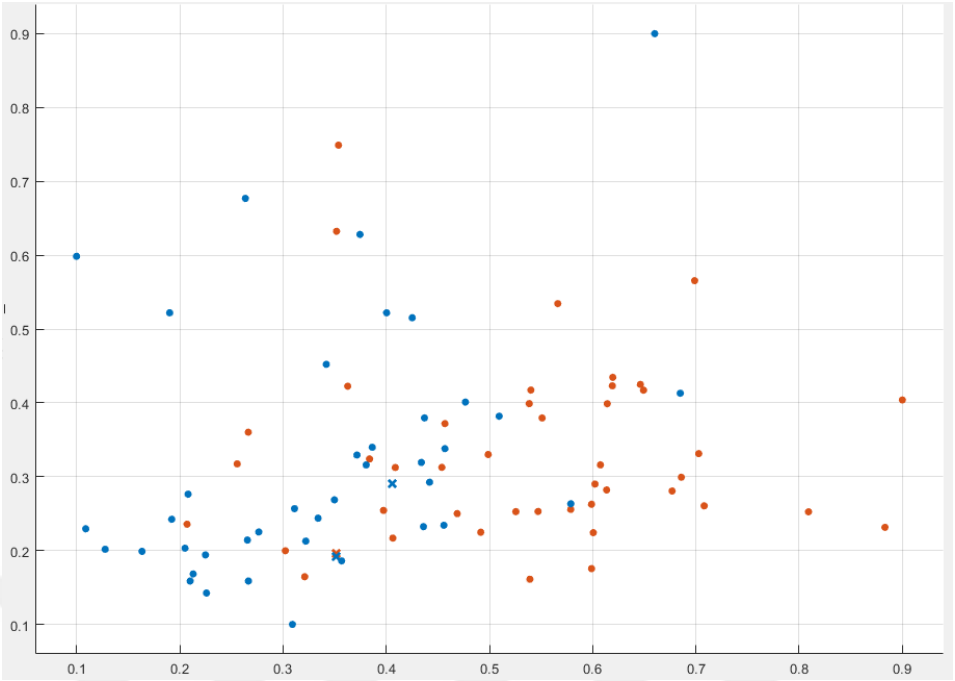

Şekil 4.11. Ham + sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 20 adetten hepsini doğru C. reti. İse kırmızı renk olup 21 doğru 2 yanlış tahmin etmiş sonuçlar Çizelge 4.11.'de verilmiştir.

Çizelge 4.11. Ham + sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| DVM | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 20 | 2 |
| C. reti. | 0 | 21 |

Ham + sentetik verilere ait KNN sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.12.'de verilmiştir.



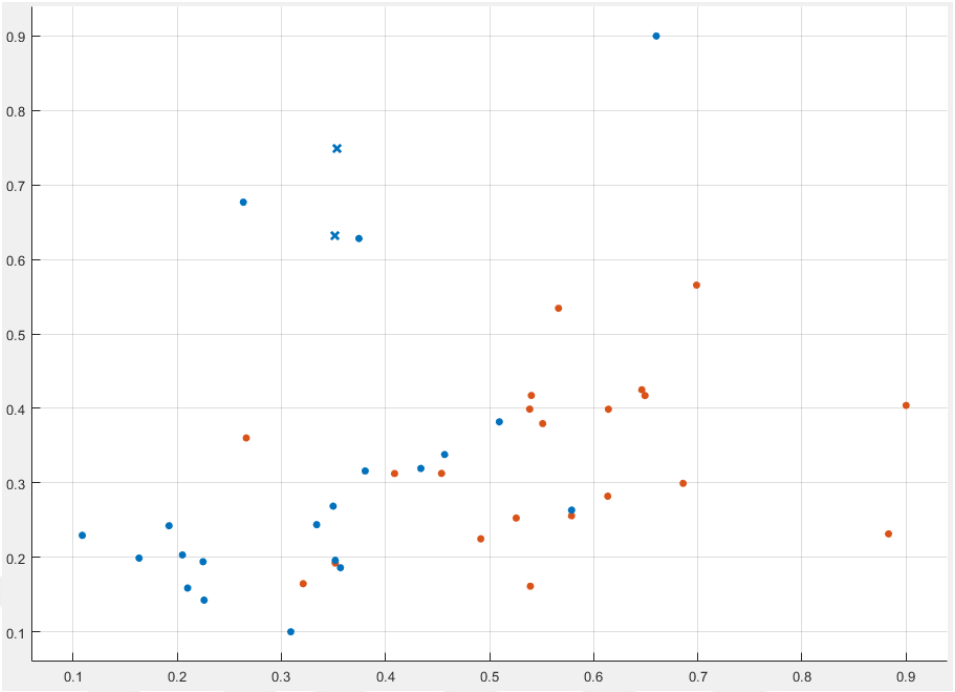

Şekil 4.12. Ham + sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 40 doğru 2 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 44 doğru 2 yanlış tahmin etmiş sonuçlar Çizelge 4.12.'de verilmiştir.

Çizelge 4.12. Ham + sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| KNN | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 40 | 2 |
| C. reti. | 2 | 44 |

Ham + sentetik verilere ait KNN sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.13.'te verilmiştir.




Şekil 4.13. Ham + sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 20 adetten hepsini doğru tahmin etmiştir. C. reti. İse kırmızı renk olup 21 doğru 2 yanlış tahmin etmiş sonuçlar Çizelge 4.13.'te verilmiştir.

Çizelge 4.13. Ham + sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| KNN | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 20 | 2 |
| C. reti. | 0 | 21 |

C. reti. ve C. ech. nohut türlerine ait sınıflandırma sonucu elde edilen grafikler, doğru ve yanlış sayıları doğrultusunda ham ve oluşturulan sentetik verilerin örneklerinden oluşan veri setine ait hesaplanan doğruluk oranları Çizelge 4.14.'te verilmiştir.

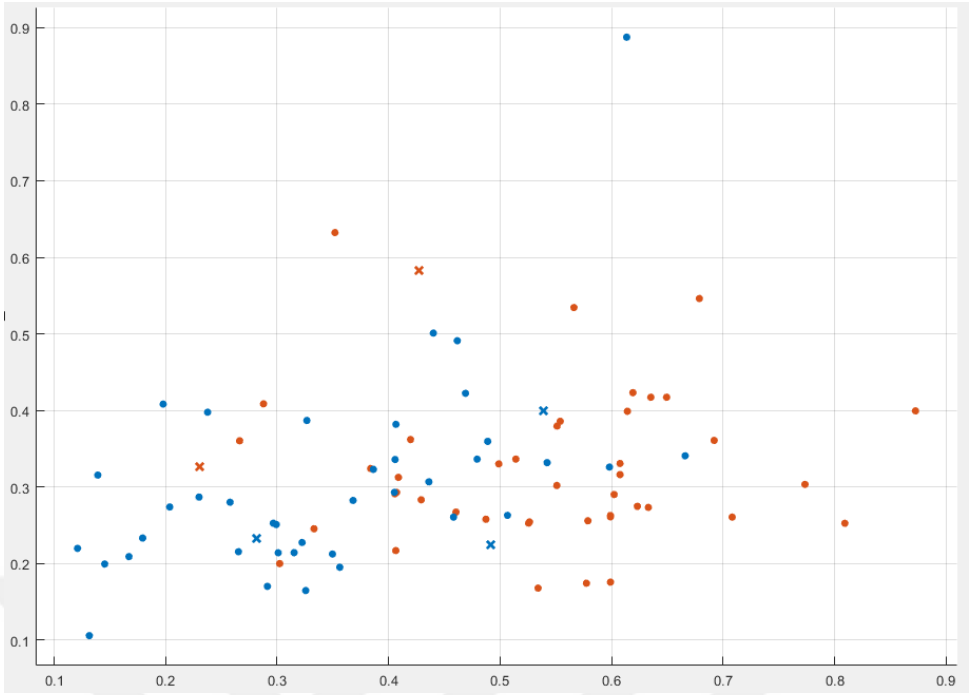
Çizelge 4.14. Ham + sentetik veri (%) doğruluk oranları

| Kullanılan Yöntem | ÇD-5 | %50-%50 |
|---|---|---|
| Karar Ağaçları | **100** | **100** |
| Destek Vektör Makinaları | 93.1 | 95.3 |
| K En Yakın Komşuluk | 96.6 | 95.3 |

Elde edilen sonuçlar karşılaştırıldığında ham ve sentetik verilerden oluşan doğruluk değerleri eğitim ve test verilerinin ayrıştırma işleminin %50-%50 ve ÇD-5 eşit doğruluk oranlarında yapıldığı en yüksek başarı oranı %100 doğruluk yüzdesi ile KA sınıflandırma yöntemi olmuştur.

Son olarak yalnızca sentetik verilerinden oluşan, verilerin eğitim ve test için ayrıştırılması işleminde ÇD-5 ve %50-%50 işlemleri kullanılarak KA, DVM ve KNN sınıflandırma yöntemleri ile elde edilen sınıflandırma sonuçları incelenmiştir.

Sadece sentetik verilere ait KA sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.14.'te verilmiştir.



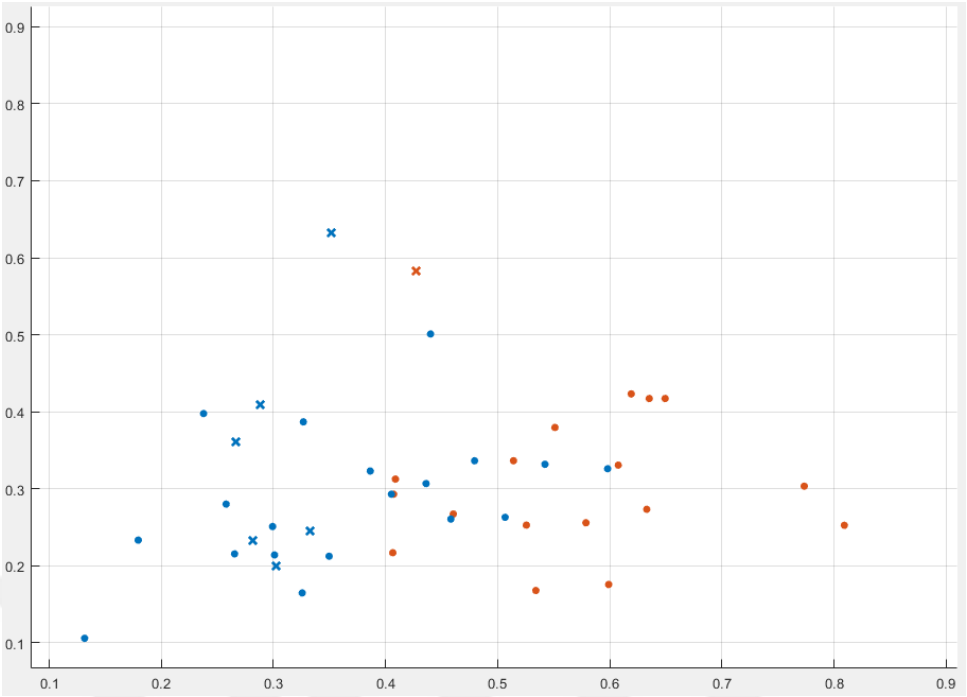

Şekil 4.14. Sadece sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 39 doğru 2 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 43 doğru 3 yanlış tahmin etmiş sonuçlar Çizelge 4.15.'te verilmiştir.

Çizelge 4.15. Sadece sentetik veri KA yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| KA | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 39 | 3 |
| C. reti. | 2 | 43 |

Sadece sentetik verilere ait KA sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.15.'te verilmiştir.



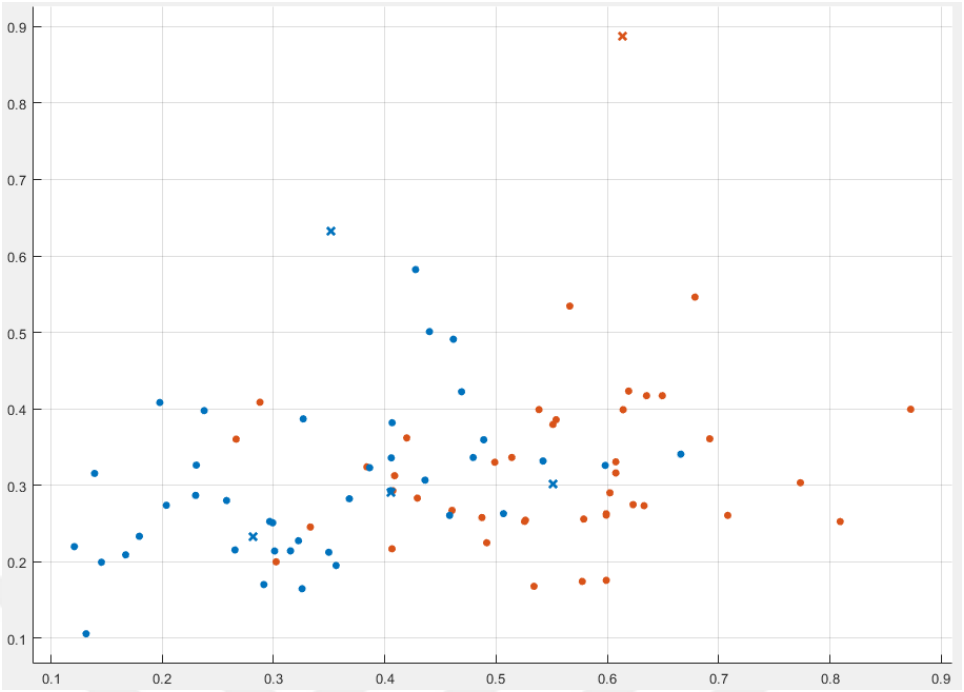

Şekil 4.15. Sadece sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 19 doğru 1 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 17 doğru 6 yanlış tahmin etmiş sonuçlar Çizelge 4.16.'da verilmiştir.

Çizelge 4.16. Sadece sentetik veri KA yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| KA | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 19 | 6 |
| C. reti. | 1 | 17 |

Sadece sentetik verilere ait DVM sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.16.'da verilmiştir.



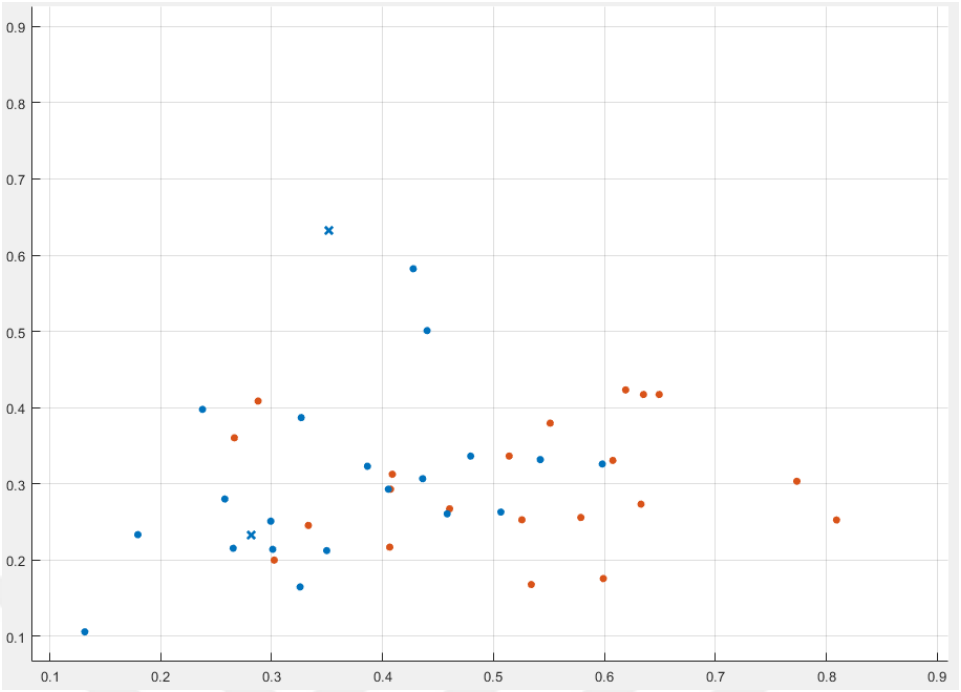

Şekil 4.16. Sadece sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 40 doğru 1 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 42 doğru 4 yanlış tahmin etmiş sonuçlar Çizelge 4.17.'de verilmiştir.

Çizelge 4.17. Sadece sentetik veri DVM yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| DVM | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 40 | 4 |
| C. reti. | 1 | 42 |

Sadece sentetik verilere ait DVM sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.17.'de verilmiştir.



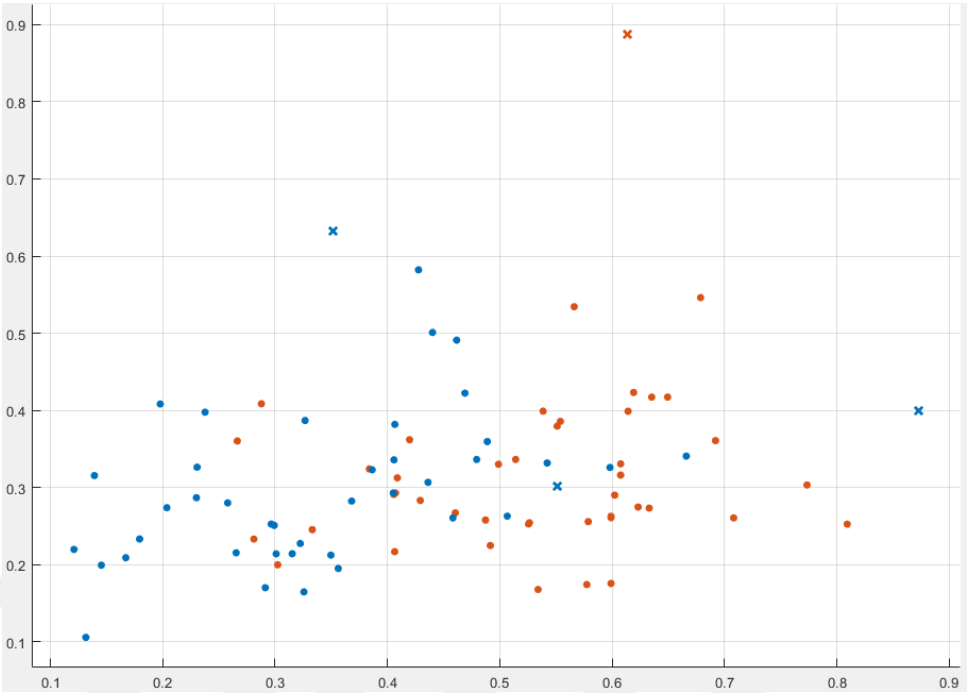

Şekil 4.17. Sadece sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 20 adetten hepsini doğru tahmin etmiştir. C. reti. İse kırmızı renk olup 21 doğru 2 yanlış tahmin etmiş sonuçlar Çizelge 4.18.'de verilmiştir.

Çizelge 4.18. Sadece sentetik veri DVM yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| DVM | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 20 | 2 |
| C. reti. | 0 | 21 |

Sadece sentetik verilere ait KNN sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin ÇD-5 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.18.'de verilmiştir.



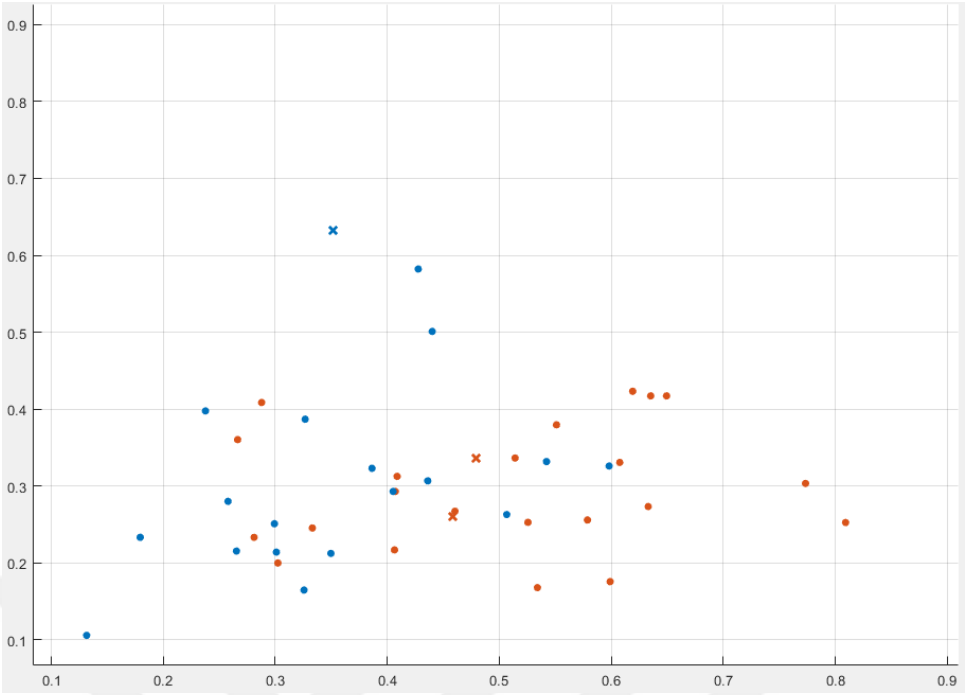

Şekil 4.18. Sadece sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 40 doğru 1 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 43 doğru 3 yanlış tahmin etmiş sonuçlar Çizelge 4.19.'da verilmiştir.

Çizelge 4.19. Sadece sentetik veri KNN yöntemi ÇD-5 ayrıştırma ile doğru ve yanlış tahminler

| KNN | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 40 | 3 |
| C. reti. | 1 | 43 |

Sadece sentetik verilere ait KNN sınıflandırma yöntemleri kullanılarak veri setinin eğitim ve test verilerinin %50-%50 ile ayrıştırılarak elde edilen doğru ve yanlış dağılımlar Şekil 4.19.'da verilmiştir.




Şekil 4.19. Sadece sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

C. ech. mavi renk olup 18 doğru 2 yanlış tahmin etmiştir. C. reti. İse kırmızı renk olup 22 doğru 1 yanlış tahmin etmiş sonuçlar Çizelge 4.20.'de verilmiştir.

Çizelge 4.20. Sadece sentetik veri KNN yöntemi %50-%50 ayrıştırma ile doğru ve yanlış tahminler

| KNN | C. ech. | C. reti. |
|---|---|---|
| C. ech. | 18 | 1 |
| C. reti. | 2 | 22 |

C. reti. ve C. ech. nohut türlerine ait sınıflandırma sonucu elde edilen grafikler, doğru ve yanlış sayıları doğrultusunda sadece sentetik verilerin örneklerinden oluşan veri setine ait hesaplanan doğruluk oranları Çizelge 4.21.'de verilmiştir

Çizelge 4.21. Sadece sentetik veri (%) doğruluk oranları

| Kullanılan Yöntem | ÇD-5 | %50-%50 |
|---|---|---|
| Karar Ağaçları | 94.3 | 83.7 |
| Destek Vektör Makinaları | 94.3 | 95.3 |
| K En Yakın Komşuluk | **95.4** | 93 |

Elde edilen sonuçlar karşılaştırıldığında sadece sentetik verilerden oluşan doğruluk değerleri eğitim ve test verilerinin ayrıştırma işleminin ÇD-5 şeklinde yapıldığı en yüksek başarı oranı %95.4 doğruluk yüzdesi ile KNN sınıflandırma yöntemi olmuştur.

Tüm uygulamalar ve veri seti incelendiğinde elde edilen başarı oranları karşılaştırılıp bakıldığı zaman kullanılan tüm yöntemlerin başarılı sonuçlar verdiği görülmektedir. Nohut tohum türleri arasında dengesiz örnek sayı dağılımları olduğu için sentetik veriler oluşturularak tekrardan makine öğrenme yöntemleri kullanılmıştır. Elde edilen sonuçlar karşılaştırıldığında, Sentetik veriler dahil edilerek yapılan sınıflandırma işleminde başarı oranlarının artığı gözlemlenmiştir. Bunun sebebi örnek sayısı artıkça özellikler arası değerlendirme hassasiyetinin de artmasından kaynaklanmaktadır. Bu şekilde elde edilen en yüksek başarı oranı eğitim ve test verilerinin ÇD-5 ile ayrılıp KA yöntemi kullanılarak yapılan sınıflandırmada olduğu gözlemlenmiştir.

### 4.3. Çalışmada Regresyon Sonucu Elde Edilen Sonuçların Karşılaştırılması

Regresyon işlemi için yine sadece ham verilerden oluşan örnekler ele alınmış ve yöntem olarak, KA ve DVM yöntemleri kullanılmıştır. Verilerin ayrıştırılma işlemi eğitim ve test için ÇD-5 ve %50-%50 olarak belirlenmiştir. Başarı ölçütü

olarak rmse tercih edilmiştir. Regresyon ile nohut türlerine ait özelliklerden yararlanarak dekar başı elde edilen ürün miktarları tahmin edilmeye çalışılmıştır. Elde edilen başarı sonuçları Çizelge 4.22.'de verilmiştir.

Çizelge 4.22. Ham veri rmse başarı oranları

| Kullanılan Yöntem | ÇD-5 | %50-%50 |
|---|---|---|
| Karar Ağaçları | 0.17 | 0.17 |
| Destek Vektör Makinaları | 0.18 | 0.18 |

Başarı sonuçları karşılaştırıldığında rmse için 0'a yakın değerler daha başarılı sayılmaktadır. Kullanılan iki yöntemden KA ile elde edilen sonuçlar DVM yöntemi ile elde edilen sonuçlardan daha başarılı olmuştur. Verilerin ayrıştırılması ise sonuçları bu uygulama için çok etkilememiştir.

Aynı şekilde regresyon işlemi için ham verilere ilaveten sentetik veriler eklenerek oluşturulan veri seti kullanılmıştır. Regresyon yöntemleri olarak, KA ve DVM kullanılmıştır. Verilerin ayrıştırılma işlemi eğitim ve test için ÇD-5 ve %50-%50 olarak belirlenmiştir. Başarı ölçütü olarak rmse tercih edilmiştir. Elde edilen başarı sonuçları Çizelge 4.23.'te verilmiştir.

Çizelge 4.23. Ham veri+ sentetik veri rmse başarı oranları

| Kullanılan Yöntem | ÇD-5 | %50-%50 |
|---|---|---|
| Karar Ağaçları | 0.16 | 0.18 |
| Destek Vektör Makinaları | 0.16 | 0.19 |

Başarı sonuçları karşılaştırıldığında ÇD-5 ile yapılan veri setlerinin ayrıştırılmasında kullanılan her iki yönteminde başarı oranı 0.16 ile aynı olmuştur. Eğitim ve test verilerinin eşit ayrıştırıldığı uygulamada ise KA yöntemi ile elde edilen başarı oranı, DVM yöntemi ile elde edilen sonuçlardan daha başarılı olmuştur.

Yalnızca sentetik örneklerden oluşan veri setine, regresyon işlemi için makine öğrenme yöntemlerinden KA ve DVM uygulanmıştır. Verilerin ayrıştırılma işlemi daha önceki çalışmalarda olduğu gibi eğitim ve test için ÇD-5 ve %50-%50 olarak belirlenmiştir. Başarı ölçütü olarak rmse tercih edilmiştir. Elde edilen başarı sonuçları Çizelge 4.24.'te verilmiştir.

Çizelge 4.24. Yalnızca sentetik veri rmse başarı oranları

| Kullanılan Yöntem | ÇD-5 | %50-%50 |
|---|---|---|
| Karar Ağaçları | 0.16 | 0.19 |
| Destek Vektör Makinaları | 0.17 | 0.19 |

Başarı sonuçları karşılaştırıldığında eğitim ve test verilerinin eşit ayrıştırıldığı uygulamada her iki yönteminde başarı oranı 0.19 ile aynı olmuştur. Eğitim ve test verilerinin ÇD-5 ayrıştırılmasında elde edilen başarı sonuçları ise KA yöntemi ile elde edilen başarı oranı, DVM yöntemi ile elde edilen sonuçlardan daha başarılı olmuştur.

Tüm uygulamalar ve veri setleri incelendiğinde elde edilen başarı oranları karşılaştırılıp bakıldığı zaman kullanılan her iki yöntemin başarılı sonuçlar verdiği görülmektedir. Nohut tohum türleri arasında dengesiz örnek sayı dağılımları olduğu için sentetik veriler oluşturularak tekrardan makine öğrenme yöntemleri kullanılmıştır. Elde edilen sonuçlar karşılaştırıldığında, Sentetik veriler dahil edilerek yapılan regresyon işleminde başarı oranlarının artığı gözlemlenmiştir. Bunun sebebi örnek sayısı artıkça özellikler arası değerlendirme hassasiyetinin de artmasından kaynaklanmaktadır. Bu şekilde elde edilen en yüksek başarı oranı eğitim ve test

verilerinin ÇD-5 ile ayrılıp KA yöntemi kullanılarak yapılan regresyon uygulamalarında olduğu gözlemlenmiştir.

# 5. SONUÇLAR ve ÖNERİLER

Yapılan bu çalışmada, Harran Üniversitesi Ziraat Fakültesi deneme arazilerinden elde edilen veri seti makine öğrenme yöntemlerinden, KA, DVM ve KNN yöntemleri kullanılarak, nohut verim ve tür tahmini yapılmıştır. Elde edilen veriler üzerinde bazı ön işlemler gerçekleştirilmiştir. Ayırt etmede etkili olan bazı özellikler eklenip geri kalan özellikler yok sayılmıştır.

Değerlendirmenin daha hassas ve kabul edilebilir olması için D min. max. normalizasyon işlemi kullanılmış verilerin tamamı 0,1 ile 0,9 arasına alınmıştır. Veriler bu işlem ile daha düşük boyut haline getirilip işlenmesi kolaylaştırılmıştır.

Daha sonra SMOTE sentetik veri üretme işlemi uygulanmıştır. Dengesiz veri problemleri için yeniden örnekleme yöntemleri kullanılır. Bu yaklaşım ile az örnek sayısına sahip veriler kullanılarak az olan örneklerden yeni ve dengeli sınıflar oluşturularak iki sınıfın örnek sayıları birbirine yakın olmuştur. Türlerin yakın sayıda örnek sayılarına sahip olup oluşturulan sentetik verilerin dahil edildiği veri seti kullanıldığında başarı oranın arttığını gözlemlenmiştir.

Performans ölçütü olarak ise sınıflandırma için doğruluk regresyon için rmse kullanılmıştır. Elde edilen sonuçların sınıflandırma başarı oranları karşılaştırıldığında en iyi sonucu veren uygulama ham veriler ile birlikte oluşturulan sentetik verilerin kullanıldığı veri setinin eğitim ve test verilerinin ÇD-5 ile ayrıştırıldığı sınıflandırma başarı oranı en yüksek %100 doğruluk oranı ile KA sınıflandırma yöntemi olmuştur. Regresyon için ise yine en yüksek başarı oranı KA yöntemi kullanılarak 0.16 olmuştur. Kullanılan yöntemlerin başarı performansları değerlendirildiğinde genel olarak başarılarının yüksek oluşu kullanılan makine öğrenme yöntemlerinin bu alanda kullanılabileceğini göstermiştir.

## KAYNAKLAR


AÇIKGÖZ, E., TURAN, Z. M., and ÇELİK, N. (1987). Performance of some forage rape cultiv ars in Bursa province of Turkey.

ADAK, M. S., KAYAN, N., ve BENLİOĞLU, B. (2015). yemeklik tane baklagiller üretiminde değişimler ve yeni arayışlar.

AKAY, E. Ç. (2018). Ekonometride yeni bir ufuk: Büyük veri ve makine öğrenmesi. Sosyal Bilimler Araştırma Dergisi, 7(2), 41-53.

AKÇİN, A. (1988). Yemeklik dane baklagiller. Selçuk Üniversitesi Yayınları, 43, 307-367.

AKSOY, T. (2017). Müşteri memnuniyeti araştırmalarında cinsiyet faktörünün etkisi (Master's thesis, Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü).

ALPAY, Ö., ve ERDEM, E. (2018). Seralarda Isıtma Miktarının Yumuşak Hesaplama Teknikleri Kullanılarak Belirlenmesi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 30(2), 203-213.

ALTUN, M., ve TÜRKER, M. (2021). Çoklu Zamanlı Sentinel-2 Görüntülerinden Tarımsal Ürün Tespiti: Mardin–Kızıltepe Örneği. Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 21(4), 881-899.

ARI, B., ARI, A., ŞENGÜ, A., ve ARSLAN TUNCER, S. (2019). Derin öznitelikler kullanılarak aşırı öğrenme makineleri ile kayısı yapraklarının sınıflandırılması.

AYDIN, O. (2012). AB'de kırsal turizmde ilk 5 ülke ve Türkiye'de kırsal turizm. Karamanoğlu Mehmetbey Üniversitesi Sosyal ve Ekonomik Araştırmalar Dergisi, 2012(2), 39-46.

AYDIN, C. (2018). Makine öğrenmesi algoritmaları kullanılarak itfaiye istasyonu ihtiyacının sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, (14), 169-175.

AYDOĞAN, A., GÜRBÜZ, A., EVLİCE, AK ve KARACA, K. (2011). Nohut (Cicer arietinum L.) Çeşitlerinde Yaprak ve Un Rengi ile. Tarla Bitkileri Merkez Araştırma Enstitüsü Dergisi , 20 (2), 17-23.

AYELE, N. A., and TAMİRU, H. K. (2020). Developing Classification Model for Chickpea Types using Machine Learning Algorithms. International Journal of Innovative Technology and Exploring Engineering, 10(1), 5-11.

AZİMİ, S., and GANDHİ, T. K. (2020, December). Water Stress Identification in Chickpea Images using Machine Learning. In 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC) (pp. 1-6). IEEE.

BAKLAGİL WİKİPEDİA. (2022). https://tr.wikipedia.org/wiki/Baklagiller adresinden alındı. erişim tarihi 22.12.2022.

BAŞER, B. Ö., YANGIN, M., ve SARIDAŞ, E. S. (2021). Makine öğrenmesi teknikleriyle diyabet hastalığının sınıflandırılması. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 25(1), 112-120.

BAYAR, R. (2004). cumhuriyet döneminde türkiye'nin arazi bölünüşü ve tarım alanlarındaki değişmeler. coğrafi bilimler dergisi, 2(1), 41-55.

BAYAR, G. (2021). Increasing measurement accuracy of a chickpea pile weight estimation tool using Moore-neighbor tracing algorithm in sphericity calculation. Journal of Food Measurement and Characterization, 15(1), 296-308.

BAYRAKÇI, H. C., ÇİÇEKDEMİR, R. S., ve ÖZKAHRAMAN, M. Tarım Arazilerinde Harcanan Su Miktarını Yapay Zekâ Teknikleri Kullanarak Belirlenmesi1. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 9(6), 237-250.

BAYRAMOĞLU, Z. (2010). Tarımsal verimlilik ve önemi. Selcuk Journal of Agriculture and Food Sciences, 24(3), 52-61.

BELAY, A. J., SALAU, A. O., ASHAGRİE, M., ve HAİLE, M. B. (2022). Development of a chickpea disease detection and classification model using deep learning. Informatics in Medicine Unlocked, 100970.

BURUCU, D. (2017). Ürün Raporu:NOHUT. Tarımsal Ekonomi ve Politika Geliştirme Enstitüsü. ISBN: 978-605-2207-14-7.

BÜTÜNER, R., and CALP, M. H. (2020). Covid-19 detection from lung tomography images using deep learning and machine learning methods (No. 4097). EasyChair.

CHAİ, T., and DRAXLER, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250.

CORTES, C. and VAPNİK, V. (1995). Support vector networks. Machine learning, 20 (3), 273-297.

ÇAKMAK, A. (2019). Adıyaman, Diyarbakır ve Şanlıurfa illerinde toplanan yabani nohut devletinin karakterizasyonu/Adıyaman, Diyarbakır ve Şanlıurfa'dan toplanan yabani cicer türlerinin karakterizasyonu (Doktora tezi).

ÇETİN, N., OZAKTAN, H., UZUN, S., UZUN, O., and CİFTCİ, C. Y. (2023). Machine learning based mass prediction and discrimination of chickpea (Cicer arietinum L.) cultivars. *Euphytica*, *219*(1), 1-18.

ÇETİN, S., HAKAN, N. A. R., and KIZIL, Ü. Counting and Classification of Seed Using Machine Learning Methods. ÇOMÜ Ziraat Fakültesi Dergisi, 10(1), 55-62.

DEMİR, H., ERDOĞMUŞ, P., ve KEKEÇOĞLU, M. (2018). Destek Vektör Makineleri, YSA, K-Means ve KNN Kullanarak Arı Türlerinin Sınıflandırılması. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 6(1), 47-67.

DEMİRCİ, Ö., ve BİLDİRİCİ, N. (2020). Şanlıurfa ekolojik koşullarında yetiştirilen bazı nohut (Cicer arietinum L.) çeşitlerinin verim ve verim unsurlarının belirlenmesi. Avrupa Bilim ve Teknoloji Dergisi, (20), 656-662.

DOĞAN, R., GÜVEN, E. Y., ve AYDIN, M. A. (2019). Makine Öğrenmesi ile İşletim Sistemi Tarama Saldırısı Tespiti Operating System Scan Attack Detection with Machine Learning. In Proceedings of the International Conference on All Aspects of Cyber Security (Vo.

E.KAAN ÜLGEN. (2017). https://medium.com/@k.ulgen90/makine-%C3%B6%C4%9Frenimi-b%C3%B6l%C3%BCm-5-karar-a%C4%9Fa%C3%A7lar%C4%B1-c90bd7593010. adresinden alındı. erişim tarihi 22.12.2022

ERMAN, M., ÇİFTÇİ, V., ve GEÇİT, H. H., 1997. Nohut (Cicer arietinum L.)'ta Özellikler Arası İlişkiler ve Path Katsayısı Analizi Üzerine Bir Araştırma. Ankara Üniversitesi Ziraat Fakültesi Tarım Bilimleri Dergisi, 3 (3): 43-46.

ERMEYDAN, K. (2019). Pirinç çeltiğinin çekilmesi sonucu elde edilen pirinç tanelerinin sınıflandırılması işleminin makine öğrenmesi ile gerçekleştirilmesi (Master's thesis).

EROL, C. (2019). Mardin ili ve civarında nohut (Cicer reticulatum) gen kaynaklarının belirlenmesi toplanması ve karakterizasyonu (Master's thesis, Fen Bilimleri Enstitüsü).

EYİDOGAN, F. ve ÖZ, M.T. (2007). Effect of salinity on antioxidant responses of chickpea seedlings. Acta Physiol Plant, 29: 485-493.

FİELD, CB, BEHRENFELD, MJ, RANDERSON, JT, and FALKOWSKİ, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. science, 281 (5374), 237-240.

HEKİMOĞLU, B., ve ALTINDEĞER, M. (2006). Organik Tarım ve Bitki Koruma Açısından Organik Tarımda Kullanılacak Yöntemler. Samsun Valiliği Gıda Tarım Ve Hayvancılık İl Müdürlüğü.

IUCN. (2010). IUCN (International Union for Conservation of Nature). (2010). Numbers of threatened species by major groups of organisms (1996–2009).

JAHANBAKHSHİ, A., ABBASPOUR-GİLANDEH, Y., HEİDARBEİGİ, K., and MOMENY, M. (2021). Detection of fraud in ginger powder using an automatic sorting system based on image processing technique and deep learning. Computers in Biology and Medicine, 136, 104764.

JAYALAKSHMİ, T., and SANTHAKUMARAN, A. (2011). Statistical normalization and back propagation for classification. International Journal of Computer Theory and Engineering, 3(1), 1793-8201.

KABAK, M., SAGLAM, F., and AKTAŞ, A. (2017). Usability analysis of different distance measures on TOPSIS. Journal of the Faculty of Engineering and Architecture of Gazi University, 32(1).

KAR, S., PURBEY, V. K., SURADHANİWAR, S., KORBU, L. B., KHOLOVÁ, J., DURBHA, S. S., ... and VADEZ, V. (2021). An ensemble machine learning approach for determination of the optimum sampling time for evapotranspiration assessment from high-throughput phenoty.

KARADAG, K., TAŞALTİN, R., TENEKECİ, M. E., and GÜMÜŞÇÜ, A. (2018, May). Determination of water stress for pepper from spectral reflections through artificial learning methods. In 2018 26th Signal Processing and Communications Applications Conference (SIU).

KARADAĞ, K., ve TAŞALTIN, R. (2016). Biber bitkisinden alınan spektral yansımaların yapay sinir ağları kullanarak hastalık tespiti. Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi, 3(5), 50-59.

KARADAĞ, K., ve TAŞALTIN, R. (2016). Biber bitkisinden alınan spektral yansımaların yapay sinir ağları kullanarak hastalık tespiti. Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi, 3(5), 50-59.

KARAKULLUKÇU, E. (2008). Bazı nohut (Cicer arietinum L.) çeşitlerinin tuza toleranslarının belirlenmesi. Journal of Agricultural Sciences, 14(04).

KARAMAN, M., BARAN, Ö. G. D. N., ANDIRMAN, Ö. G. D. M., AKINCI, C., BİÇER, B. T., BAŞDEMİR, Ö. Ü. F., ... ve TURAN, Z. M. L. (2022). Tarla bitkilerinde ekonomik öneme sahip stratejik ürünlerin araştırılması.

KARTAL, S., CHOUDHARY, S., MASNER, J., KHOLOVÁ, J., STOČES, M., GATTU, P., ... and KİSSEL, E. (2021). Machine Learning-Based Plant

Detection Algorithms to Automate Counting Tasks Using 3D Canopy Scans. Sensors, 21(23), 8022.
KAVZOĞLU, T., ve ÇÖLKESEN, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. Harita Dergisi, 144(7), 73-82.
KHATRİ, A., AGRAWAL, S., and CHATTERJEE, J. M. (2022). Wheat Seed Classification: Utilizing Ensemble Machine Learning Approach. Scientific Programming, 2022.
KİNİ, A. S., PREMA, K. V., and PAİ, S. N. (2021, May). Intelligent classification model for Indian chickpea. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1847-1852). IEEE.
KİRTİS, A., AASİM, M., and KATIRCI, R. (2022). Application of artificial neural network and machine learning algorithms for modeling the in vitro regeneration of chickpea (Cicer arietinum L.). Plant Cell, Tissue and Organ Culture (PCTOC), 1-12.
KOSMOWSKİ, F., and WORKU, T. (2018). Evaluation of a miniaturized NIR spectrometer for cultivar identification: The case of barley, chickpea and sorghum in Ethiopia. PloS one, 13(3), e0193620.
KOYUNCU, B., ve MURAT, G. Ö. K. Bitki ıslahında genotip verim değerinin regresyon yöntemleri ile tahmini. Bahçe, 48(2), 79-85.
KUTLUGÜN, M. A. (2017). Gözetimli makine öğrenmesi yoluyla türe göre metinden ses sentezleme (Master's thesis, İstanbul Sabahattin Zaim Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı).
MAHMOOD, A., TİWARİ, AK, SİNGH, SK, and UDMALE, SS (2022). Contemporary machine learning applications in agriculture: Quo Vadis?. Concurrency and Computing: Practice and Experience, e6940.
MCALLİSTER, P., ZHENG, H., BOND, R., and MOORHEAD, A. (2018). Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets. Computers in biology and medicine, 95, 217-233.
MUSA, A. T. A. Ş. Fıstık sınıflandırma sistemi için Siirt fıstığı imgelerinden gürbüz özniteliklerin çıkarılması. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 7(1), 93-102.
NOHUT ENTEGRE MÜCADELE TEKNİK TALİMATI. (2017).
NOHUT WİKİPEDİA. (2022). https://tr.wikipedia.org/wiki/Nohut_(t%C3%BCr) adresinden alındı erişim tarihi 22.12.2022.
OLGUNLAR TOHUM NOHUT . (2022). https://olgunlar.com.tr/tohum/goksu/. adresinden alındı tarihi 22.12.2022.
ÖZÇAĞLAR, A. (2000). Coğrafyaya giriş. Ankara: Hilmi Usta Matbaacılık.
ÖZEKES, S. (2003). Veri madenciliği modelleri ve uygulama alanları.
ÖZİÇ, M. Ü., ÇANKAYA, N., ÖZCAN, M., ve GÖKÇE, B. (2020). Görüntü işlemede nesne koordinat özelliklerini kullanarak bakliyat sayma işlemine bir yaklaşım. konya mühendislik bilimleri dergisi, 8, 28-37..
ÖZTÜRK, M. (2011). Türkiye Cicer L.(nohut) cinsinin morfolojik, palinolojik, sitotaksonomik, moleküler filogenetik kapsamda revizyonu ile tohum proteini ve element analizleri yönünden incelenmesi. Selçuk Üniversitesi Fen Bilimleri Enstitüsü.

ÖZTÜRK, M. (2011). Türkiye Cicer L.(nohut) Cinsinin Morfolojik, Palinolojik,.
PİLAVCILAR, İ. F. (2007). Metin madenciliği ile metin sınıflandırma.
PODOLNY, BS, GURSKY, VV ve SAMSONOVA, MG (2020). A machine-learning analysis of flowering gene expression in the CDC frontier chickpea cultivar. Biophysics, 65(2), 225-236.
POURDARBANİ, R., SABZİ, S., KALANTARİ, D., HERNÁNDEZ-HERNÁNDEZ, J. L., and ARRİBAS, J. I. (2020). A computer vision system based on majority-voting ensemble neural network for the automatic classification of three chickpea varieties. Foods, 9(2), 113.
POURDARBANİ, R., SABZİ, S., GARCÍA-AMİCİS, V. M., GARCÍA-MATEOS, G., MOLİNA-MARTÍNEZ, J. M., and RUİZ-CANALES, A. (2019). Automatic classification of chickpea varieties using computer vision techniques. Agronomy, 9(11), 672.
SABANCI, K., AYDIN, C., ve ÜNLERŞEN, M. F. (2012). Görüntü İşleme ve Yapay Sinir Ağları Yardımıyla Patates Sınıflandırma Parametrelerinin Belirlenmesi. Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 2(2 Sp: A), 59-62.
SAĞLAM, F., and CENGİZ, M. A. (2019, SEPTEMBER). Noise detection in imbalanced classes using adaptive boosting. In 2019 4th International Conference on Computer Science and Engineering (UBMK) (pp. 449-452). IEEE.
SAHA, D. (2022). Evaluation of Machine Learning Techniques for Image Based Quality Assessment of Chickpea (Doctoral dissertation, University of Guelph).
SEÇKİN, K. (2011). Metin madenciliğinde kullanılan yöntemlerin karşılaştırılması: Siyasi parti liderlerinin grup genel toplantı konuşmaları ile bir uygulama (Master's thesis, Sakarya Üniversitesi).
SEPETOĞLU, H. (1996). Food grain legumes. Ege University Faculty of Agriculture Publications Lecture Notes, 24(3). Sepetoğlu, H. (1996). Food grain legumes. Ege University Faculty of Agriculture Publications Lecture Notes, 24(3). adresinden alındı
SEVLİ, Ö. Ü.O. Elma bitkisi hastalıklarının derin öğrenme ile tespiti.
SİNHA, R., IRULAPPAN, V., PATİL, B. S., REDDY, P. C. O., RAMEGOWDA, V., MOHAN-RAJU, B., ... and SENTHİL-KUMAR, M. (2021). Low soil moisture predisposes field-grown chickpea plants to dry root rot disease: evidence from simulation modeling and correlation an.
SİNGH, R. N., KRİSHNAN, P., BHARADWAJ, C., and DAS, B. (2023). Improving prediction of chickpea wilt severity using machine learning coupled with model combination techniques under field conditions. *Ecological Informatics*, *73*, 101933.
ŞENOL, E. M. İ. R., ve TİMOR, M. (2013). Yapay sinir ağları ve destek vektör makineleri yöntemlerinin sınıflandırma performanslarının karşılaştırılması: borsa endeks yönünün tahmini üzerine bir uygulama.
T.C TARIM ve ORMAN BAKANLIĞI ÜRÜN MASALARI DERGİSİ AĞUSTOS AYI NOHUT (2021). https://www.tarimorman.gov.tr/BUGEM/Belgeler/B%C3%BCltenler/A%C4%9EUSTOS%202021/Nohut%20A%C4%9Fustos%20B%C3%BClteni.pdf. adresinden alındı erişim tarihi 22.12.2022

TAHERİ-GARAVAND, A., NASİRİ, A., FANOURAKİS, D., FATAHİ, S., OMİD, M., and NİKOLOUDAKİS, N. (2021). Automated in situ seed variety identification via deep learning: a case study in chickpea. Plants, 10(7), 1406.

TEKE, M., MUTLU, N., TÜRKER, U., ÇULLU, M. A., KÜPCÜ, R., BOZGEYİK, F., ... ve HAKKI, İ. Gap bölgesinde hassas tarım uygulamalarının yaygınlaştırılması projesi-hassas wıdespread applıcatıon of sustaınable precısıon agrıculture practıces ın southeastern ana.

TİMUR, M., AYDIN, F., ve AKINCI, T. Ç. (2011). İstanbul Göztepe Bölgesinin Makine Öğrenmesi Yöntemi ile Rüzgâr Hızının Tahmin Edilmesi. Makine Teknolojileri Elektronik Dergisi, 8(4), 75-80.

TUNCA, E., ve KÖKSAL, E. (2021). Sentinel 2 Uydu Görüntülerinden Bitki Türlerinin Makine Öğrenmesi ile Belirlenmesi. ÇOMÜ Ziraat Fakültesi Dergisi, 9(1), 189-200.

TÜRKOĞLU, M., HANBAY, K., SİVRİKAYA, I. S., ve HANBAY, D. (2021). Derin Evrişimsel Sinir Ağı Kullanılarak Kayısı Hastalıklarının Sınıflandırılması. Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 9(1), 334-345.

ÜNSAL, A. (2019). Beslenmenin Önemi ve Temel Besin Öğeleri. Kırşehir Ahi Evran Üniversitesi Sağlık Bilimleri Dergisi, 2(3), 1-10.

ÜSTÜNER, M., ve ŞANLI, F. B. (2019). Çok zamanlı polarimetrik SAR verileri ile tarımsal ürünlerin sınıflandırılması. Jeodezi Ve Jeoinformasyon Dergisi, 7(1), 1-10.

WOOD, J. A., and GRUSAK, M. A. (2007). Nutritional value of chickpea. Chickpea breeding and management, 101-142.

YAMAN, O., ve TUNCER, T. Bitkilerdeki Yaprak Hastalığı Tespiti için Derin Özellik Çıkarma ve Makine Öğrenmesi Yöntemi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 34(1), 123-132.

YAVAŞ, M., GÜRAN, A., ve UYSAL, M. (2020). Covid-19 Veri kümesinin Smote tabanlı örnekleme yöntemi uygulanarak sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, 258-264.

YAVUZ, S., ve DEVECİ, M. (2012). İstatiksel Normalizasyon Tekniklerinin Yapay Sinir Ağin Performansina Etkisi. Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, (40), 167-187.

YİĞİTER, Ş. Y., SARI, S. S., KARABULUT, T., ve BAŞAKIN, E. E. (2018). Kira sertifikası fiyat değerlerinin makine öğrenmesi metodu ile tahmini. International Journal of Islamic Economics and Finance Studies, 4(3), 74-82.

YILDIZ, A. K., POLATCI, H., ve HARUN, U. Ç. U. N. (2015). Farklı Kurutma Şartlarında Muz (Musa cavendishii) Meyvesinin Kurutulması ve Kurutma Kinetiğinin Yapay Sinir Ağları ile Modellenmesi. Tarım Makinaları Bilimi Dergisi, 11(2), 173-178.