



Comparative Analysis of Machine Learning Algorithms for Irrigation Status Prediction

Betül Demir^{1,a,*}, Yeşim Dokuz^{2,b}, Burak Şen^{3,c}

¹Niğde Ömer Halisdemir Üniversitesi, Ayhan Şahenk Tarım Bilimleri ve Teknolojileri Fakültesi, Disiplinlerarası Dijital Tarım Anabilim Dalı, 51240, Niğde

²Niğde Ömer Halisdemir Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği, 51240, Niğde, Türkiye

³Niğde Ömer Halisdemir Üniversitesi, Ayhan Şahenk Tarım Bilimleri ve Teknolojileri Fakültesi, Biyosistem Mühendisliği, 51240, Niğde, Türkiye

*Corresponding author

ARTICLE INFO	ABSTRACT
<p><i>Research Article</i></p> <p>Received : 21.01.2025 Accepted : 27.02.2025</p> <p>Keywords: Irrigation status estimation Machine learning Agricultural irrigation Artificial intelligence. Spatial Irrigation Estimation</p>	<p>Compared to traditional methods, systems supported by machine learning have been found to make more accurate irrigation decisions. In this study, Kaggle, a data sharing platform frequently used in data science, was utilized. The “Weather Data” dataset was used for irrigation status prediction. Missing and outlier data were corrected and dependent (irrigation status) and independent (air temperature, humidity, soil moisture value, precipitation) parameters were obtained. Focusing on the provinces in the Southeastern Anatolia Region (Adıyaman, Batman, Diyarbakır, Gaziantep, Kilis, Mardin, Siirt, Şanlıurfa and Şırnak), the accuracy of the algorithms was tested under different conditions. In the analyses conducted separately for each province, irrigation status estimation was performed using machine learning algorithms such as Decision Trees, Support Vector Machines, Random Forest, Naive Bayes, Gradient Boosting, Logistic Regression, K-Nearest Neighbor and Artificial Neural Network models. As a result of the predictions made on the dataset, the algorithms were compared using the accuracy metric and the most effective algorithms were Random Forest (95%), Decision Tree (97%), Gradient Boosting (93%) and Artificial Neural Network (98%) models with over 90% accuracy in all cities. Other algorithms also showed remarkable performance with high accuracy rates (above 75%). In the analysis for each city, the algorithm performance ranking was similar. In conclusion, this study shows that machine learning algorithms can be used in agricultural irrigation with very high performance.</p>

Türk Tarım – Gıda Bilim ve Teknoloji Dergisi, 13(2): 497-503, 2025

Sulama Durumu Tahmini için Makine Öğrenimi Algoritmalarının Karşılaştırmalı Analizi

MAKALE BİLGİSİ	ÖZ
<p><i>Araştırma Makalesi</i></p> <p>Geliş : 21.01.2025 Kabul : 27.02.2025</p> <p>Anahtar Kelimeler: Sulama durumu tahmini Makine öğrenimi Tarımsal sulama Yapay zekâ Mekansal Sulama Tahmini</p>	<p>Geleneksel yöntemlere kıyasla, makine öğrenimi ile desteklenen sistemlerin, daha hassas sulama kararları verebildiği tespit edilmiştir. Bu çalışmada, veri bilimi alanında sıkça kullanılan veri paylaşım platformu olan Kaggle'dan faydalanılmıştır. Sulama durumu tahmini için “Weather Data” veri kümesi kullanılmıştır. Veri kümesinde eksik ve aykırı veriler düzeltilmiş, bağımlı (sulama durumu) ve bağımsız (hava sıcaklığı, nemi, toprak nem değeri, yağış durumu) parametreler elde edilmiştir. Güneydoğu Anadolu Bölgesi'ndeki illere (Adıyaman, Batman, Diyarbakır, Gaziantep, Kilis, Mardin, Siirt, Şanlıurfa ve Şırnak) odaklanılarak farklı koşullarda algoritmaların doğrulukları test edilmiştir. Her bir il için ayrı ayrı yapılan analizlerde, makine öğrenimi algoritmalarından Karar Ağaçları, Destek Vektör Makineleri, Rastgele Orman, Naive Bayes, Gradyan Artırma, Lojistik Regresyon, K-En Yakın Komşu ve Yapay Sinir Ağı modelleri kullanılarak sulama durumu tahmini gerçekleştirilmiştir. Veri kümesi üzerinde yapılan tahminler sonucunda algoritmalar, doğruluk (accuracy) metriği kullanılarak karşılaştırılmış ve en etkili algoritmaların Rastgele Orman (%95), Karar Ağacı (%97), Gradyan Artırma (%93) ve Yapay Sinir Ağı (%98) modeli ile tüm şehirlerde %90'ın üzerinde sulama durumu tahmini doğruluğu elde edilmiştir. Diğer algoritmalar da yüksek doğruluk oranları ile (%75 üzeri) dikkate değer performans sergilemişlerdir. Her bir il için yapılan analizlerde, algoritma performans sıralamasının benzer olduğu belirlenmiştir. Sonuç olarak, bu çalışma makine öğrenimi algoritmalarının tarımsal sulamada oldukça yüksek bir performansla kullanılabileceğini göstermektedir.</p>

^a betuldemir339@gmail.com

^b <https://orcid.org/0000-0002-4388-1122>

^c ysen@ohu.edu.tr

^d <https://orcid.org/0000-0001-7202-2899>

^e bsen@ohu.edu.tr

^f <https://orcid.org/0000-0001-8105-1106>



Giriş

Su, tarımsal üretimin temel yapı taşlarından biridir. Bitkiler, kökleri aracılığıyla topraktan suyu emerek ihtiyaç duydukları besinleri alır ve bu besinler, temel metabolik süreçlerin sağlıklı bir şekilde gerçekleşmesi için kullanılır (Sargıncı ve ark., 2022). Sulama suyunun doğru yönetimi, su güvenliğinin sağlanması ve tarımsal üretimin sürdürülebilirliğinin korunması açısından kritik bir öneme sahiptir.

Kuraklık, dünya nüfusunun hızla artması, tatlı su kaynaklarının sınırlı olması tarımsal üretimi tehdit eden önemli etmenler arasındadır (Türkeş, 2012). Türkiye'nin yıllık su potansiyeli yaklaşık 112 milyar metreküptür. Bu suyun %74'ü tarımsal sulama için kullanılmaktadır. Tarımsal sulamada kullanılan suyun %50'si ise verimsiz kullanım ve geleneksel yöntemler nedeniyle israf edilmektedir (Çakmak & Gökalp, 2013). Bu nedenle Türkiye gibi birçok ülke, yenilikçi çözümler geliştirme ihtiyacı ile karşı karşıyadır. Akıllı sulama sistemleri, bu ihtiyaca yanıt olarak, suyun doğru zamanda, doğru miktarda ve doğru noktaya uygulanmasını sağlayarak su kullanımını optimize eder (Şahin, 2024). Sulama planlaması; toprağın su tutma kapasitesi, bitki türü, iklim koşulları ve sulama yöntemine göre değişkenlik gösterir. Bu nedenle, sulama kararları alınırken birçok faktörün göz önünde bulundurulması gerekmektedir.

Geleneksel sulama kararları genellikle çiftçilerin tecrübelerine ve gözlemlerine dayanır (Çakmak ve ark., 2008). Çiftçiler, bitkilerdeki yaprak solgunluğu, renk değişimi veya büyüme duraklaması gibi belirtileri inceleyerek su ihtiyacını tahmin eder. Toprağın yüzey rengine bakarak veya el ile sıkarak, toprağa bir çubuk batırarak, bir su kabının buharlaşma oranına bakarak, bölgenin yağış ve sıcaklık düzenlerini göz önünde bulundurarak sulama zamanlamasını belirler ve önceki sulamaların ürün verimi üzerindeki etkilerini inceleyerek kararlar alır. Geleneksel yöntemler, düşük maliyetli ve uygulanabilir olmaları nedeniyle hala yaygın olarak tercih edilmekle birlikte, modern tekniklerle karşılaştırıldığında hassasiyet ve esneklik açısından sınırlamalar göstermektedir (Şahin, 2024). Bu durum, tarımsal üretimde daha yenilikçi yaklaşımların benimsenmesine zemin hazırlamıştır. Sensörlerle yapılan ölçümler ve veri tabanlı sulama tekniklerine kısmen geçilmiş olsa da, bu teknolojilerin sulama kararlarında istenen doğruluk seviyesine ulaşmadığı ve tarımsal verimliliği tam anlamıyla optimize edemediği gözlemlenmektedir. Bunun temel nedenleri arasında, sensörlerin sınırlı kapsama alanı, toprak ve bitki türlerine göre değişkenlik gösteren nem ihtiyacının tam olarak belirlenememesi, altyapı ve veri analizinde yaşanan eksiklikler ile çiftçilerin bu teknolojilere erişim ve adaptasyon konusundaki zorlukları yer almaktadır.

Çeşitli koşullarda en iyi sulama kararlarını sağlayan yöntemler arasında; FAO Penman-Monteith yöntemi, Hargreaves-Samani denklemi gibi evapotranspirasyon modelleri, tansiyometre ve toprak nem sensörleri gibi doğrudan ölçüm teknikleri, su dengesi modelleri, iklim tahmin modelleri, zaman serisi analizleri, regresyon modelleri ve makine öğrenimi tabanlı karar destek sistemleri bulunmaktadır (Taştan, 2019). Bu yöntemler, sulama kararlarının daha hassas bir şekilde alınmasına

katkı sağlarken, uygulamada karşılaşılan veri eksikliği ve bölgesel değişkenlikler gibi faktörler nedeniyle beklenen doğruluk seviyesine tam olarak ulaşamamaktadır.

Allen ve ark. (1998) tarafından geliştirilen FAO Penman-Monteith yöntemi, evapotranspirasyonu hesaplamak için kullanılan ve dünya genelinde yaygın kabul gören bir modeldir. Bununla birlikte, tansiyometre ve toprak nem sensörleri ile toprak nemine dayalı sulama kararları alınabilir (Doğru ve ark., 2012). Çiftçiler, bu sensörlerden gelen verilerle bitkilerin su ihtiyacını tahmin eder ve sulama zamanını optimize eder. Bitkilerdeki yaprak solgunluğu, renk değişimleri veya büyüme duraklamaları gibi anormal durumlar gözlemlenerek sulama gereksinimleri belirlenebilir.

İstatistiksel yöntemler, geleneksel yaklaşımların önemli bir parçasıdır. Yılmaz ve ark. (2007), geçmiş sulama verilerinin analizine dayalı regresyon modellerinin sulama ihtiyaçlarını tahmin etmede etkili olduğunu göstermiştir. Benzer şekilde, zaman serisi analizleri, yağış ve sıcaklık gibi iklim parametrelerinin de sulama zamanı tahmini için kullanıldığı bilinmektedir. Ancak bu yöntemlerin doğruluğu, veri kalitesine ve ölçüm sıklığına bağlı olarak sınırlı kalabilmektedir.

Makine öğrenmesi, sulama kararlarının verilmesinde geleneksel yöntemlere kıyasla daha hassas ve veriye dayalı çözümler sunarak tarımsal süreçlerde önemli bir dönüşüm yaratmıştır. Destek Vektör Makineleri (SVM), farklı parametreler arasındaki ilişkileri modelleyerek sulama ihtiyaçlarının tahmini için sıkça kullanılan yöntemlerden biridir (Demir ve ark., 2024). Bu yöntem, özellikle karmaşık ve çok değişkenli verilerin analiz edilmesinde etkili bir araç olarak öne çıkar. Benzer şekilde, Karar Ağacı ve Rastgele Orman algoritmaları gibi yöntemler, toprak koşulları ve iklim verileri gibi çok boyutlu verilerle sulama planlamasının otomasyonu için başarıyla uygulanmıştır (Zhang ve ark., 2015). Hosseini ve ark. (2019), Yapay Sinir Ağları (ANN) kullanarak geliştirdikleri modelle sulama zamanlamasının daha yüksek doğrulukla tahmin edilebileceğini göstermiştir. Ayrıca, Konvolüsyonel Sinir Ağları (CNN) gibi yöntemler, görüntü verilerini analiz ederek bitki sağlığı ve toprak durumunun değerlendirilmesini sağlayarak sulama kararlarına destek olmaktadır. Optimizasyon tabanlı yöntemlerden Genetik Algoritmalar, sulama sistemlerinde su dağıtımını optimize etmek için kullanılırken, Karınca Kolonisi Optimizasyonu gibi yöntemler, sulama yollarının verimliliğini artırmak için uygulanmıştır. Bu yaklaşımlar, sulama sistemlerinin enerji ve su tasarrufunu maksimize etme potansiyeline sahiptir (Rahman ve ark., 2020). Makine öğrenimi yöntemleri, sulama kararlarının daha hassas ve etkili bir şekilde verilmesine olanak tanımaktadır. Geleneksel yöntemlerin sınırlarını aşan bu teknikler, büyük veri setlerinin analiz edilmesi ve kararların optimize edilmesi süreçlerinde önemli bir rol oynamaktadır.

Türkiye'de de makine öğrenimi tabanlı sulama sistemleri giderek daha fazla kullanılmakta, verimliliğin artırılması için önemli bir araç olarak değerlendirilmektedir (Bayrakçı ve ark., 2021). Ancak sulama planlamasında hangi makine öğrenimi algoritmasının daha iyi sonuçlar verdiği değişkenlik göstermektedir. Her algoritmanın belirli veri setlerine ve

koşullara göre performansı farklılık gösterebilir. Bu çalışmanın amacı, sulama planlamasında hangi algoritmanın daha uygun olduğunu tespit ederek, doğru planlama kararlarının alınmasına fayda sağlamaktır. Su kaynaklarının daha etkin yönetilmesi ve tarımsal verimliliğin en üst düzeye çıkarılması için büyük önem taşımaktadır.

Bu çalışmada Güneydoğu Anadolu bölgesi çalışma alanı olarak seçilmiştir. Bölgedeki toplam su potansiyeli, yaklaşık 53 milyar metreküp civarındadır ve bunun büyük bir kısmı tarımsal sulama için kullanılmaktadır (GAP BKİ, 2023). Sulamanın büyük ölçüde geleneksel yöntemlerle (salma sulama vb.) yapılması, su kayıplarına ve gereksiz su kullanımına neden olmaktadır. Tarımsal sulamanın modern yöntemler ve makine öğrenimi tabanlı teknolojilerle optimize edilmesi hem su kaynaklarının korunması hem de tarımsal üretimin sürdürülebilirliği açısından büyük bir fırsat sunmaktadır. Bu süreçte kullanılacak makine öğrenimi algoritmalarının doğru seçimi, suyun verimliliğinin artırılmasında kritik bir rol oynayacaktır.

Materyal ve Yöntem

Makine öğrenmesi, girdi olarak aldığı verileri kullanarak anlamlı bilgiler üreterek tahminler yapan algoritmalar geliştiren bir bilim dalıdır (Bayrakçı ve ark., 2021). Makine öğrenmesi, sulama yönetiminde, geçmiş verilere dayanarak su kullanımının tahmin edilmesi ve tarımsal verimliliğin artırılması için kullanılır.

Tarımsal sulama yönetiminde kullanılan makine öğrenmesi algoritmaları, veri setinin özelliklerine göre seçilir ve su yönetimini optimize etmek, çiftçilerin karar alma süreçlerini iyileştirmek için güçlü araçlar sunar.

Bu çalışmada kullanılan makine öğrenimi algoritmaları aşağıda sunulmuştur.

Lojistik Regresyon (Logistic Regression): Bağımsız değişkenlerin doğrusal kombinasyonunu lojistik (sigmoid) fonksiyon aracılığıyla bir olasılık değerine dönüştüren ve çıktı değişkeninin belirli bir sınıfa ait olma ihtimalini hesaplayan bir sınıflandırma yöntemidir. Genellikle iki sınıflı (binary) sınıflandırma problemlerinde kullanılır ve modelin çıktısı belirlenen bir eşik değerine göre sınıflandırılır. Lojistik regresyon, kolay yorumlanabilir olması ve hesaplama açısından verimli çalışması nedeniyle tercih edilen yöntemlerden biridir (Bircan, 2004).

Karar Ağaçları (Decision Tree): Veri setindeki özelliklere göre karar kuralları oluşturur ve bu kuralları kullanarak tahminler yapar. Basit ve sezgisel bir yapıya sahiptir. Kök düğüm olarak adlandırılan başlangıç noktasından başlayarak, her düğüm bir özelliği temsil eder ve dallara ayrılır. Bu dallar, karar ağacının oluşturduğu karar kurallarını temsil eder. Sonuç olarak, veri setinin her bir girdisi, ağacın dallarını takip ederek bir sonuca ulaşır ve böylece tahmin yapılır (Özcan & Özer, 2021). Karar Ağaçları algoritması, karmaşık algoritmalarla birlikte kullanılır veya birleştirilerek daha güçlü modeller oluşturulur.

Rastgele Orman (Random Forest): Birçok karar ağacının bir araya gelerek oluşturduğu bir modeldir. Her bir karar ağacı, veri setinin farklı alt kümeleri üzerinde eğitilir ve farklı kararlar alır. Daha sonra, bu karar

ağaçlarının tahminleri bir araya getirilerek, son tahmin yapılır. Bu yaklaşım, modelin aşırı öğrenme (overfitting) yapma olasılığını azaltır ve genellikle yüksek doğruluk oranları ile sonuçlanır (Sevgen & Aliefendioğlu, 2020). Rastgele Orman algoritması, bireysel karar ağaçlarının sınırlamalarını aşarak, büyük veri setlerinde bile etkili bir performans sergiler. Birden fazla ağacın kullanılmasıyla model, her bir ağacın yaptığı hataları minimize eder. Bu sayede, özellikle büyük ve karmaşık veri setlerinde yüksek tahmin doğruluğuna ulaşabilir.

Destek Vektör Makineleri (SVM): Veriler arasında bir ayrım çizgisi oluşturarak sınıflandırma yapan bir algoritmadır (Demirci, 2019). SVM, özellikle karmaşık verilerde yüksek doğruluğu ile geleneksel yöntemlere kıyasla daha etkili sonuçlar sunar.

K-En Yakın Komşu (KNN): Bir veri noktasını sınıflandırmak veya tahmin etmek için veri setindeki en yakın komşularını analiz eder. Algoritma, yeni bir veri noktasının etrafındaki K sayıda en yakın komşuya bakarak bu noktanın hangi sınıfa ait olduğunu veya hangi değeri alacağını belirler. Küçük veya orta ölçekli veri setlerinde oldukça hızlı analiz yapabilir. Daha büyük veri setlerinde veya çok fazla değişken içeren veri setlerinde, KNN'nin yavaş çalışması ve performans kaybı yaşaması muhtemeldir. Bu gibi durumlarda, veri boyutunu küçültmek için özellik seçimi veya boyut indirgeme teknikleri kullanılabilir, ya da daha uygun bir algoritmaya geçiş yapılabilir (Bütüner ve ark., 2023).

Gradyan Artırma (Gradient Boosting): Zayıf modelleri ardışık olarak eğitip birleştirerek güçlü ve yüksek performanslı tahmin modelleri oluşturur. Bu süreçte her bir yeni model, önceki modelin yaptığı hataları düzeltmeye odaklanarak nihai tahmin doğruluğunu artırır (Aydın ve ark., 2023).

Yapay Sinir Ağları (YSA): Yapay sinir ağları (YSA), biyolojik sinir sistemlerinden esinlenerek tasarlanmış yapay modellerdir (Öztürk & Şahin, 2018). Giriş, gizli ve çıkış katmanlarından oluşur. Giriş katmanı verileri alır, gizli katmanlar bu verilerdeki desenleri öğrenir ve çıkış katmanı tahminleri üretir. Model, ileri besleme (feedforward) ile veri akışını sağlar ve geri yayılım (backpropagation) yöntemiyle hatayı azaltarak öğrenme sürecini geliştirir. Geri yayılda, hata geriye doğru yayılıp ağırlıklar güncellenerek modelin doğruluğu artırılır (Akkaya, 2007).

Çalışma Alanı

Güneydoğu Anadolu Bölgesi, yarı kurak ve sıcak iklim özelliklerine sahiptir. Yazlar oldukça sıcak ve kurak geçerken, kışlar nispeten ılıman ve az yağışlıdır. Yıllık yağış miktarı ortalama 400-700 mm arasında değişmekte olup, bölgenin tarımsal üretimi büyük ölçüde sulamaya dayalıdır. Yaz aylarında sıcaklıklar 40°C'yi bulabilirken, kış aylarında sıcaklık nadiren 0°C'nin altına iner. Bu durum, bölgedeki tarımsal faaliyetlerde su yönetiminin kritik öneme sahip olmasına yol açmaktadır. Özellikle yaz aylarında aşırı sıcaklık ve yetersiz yağış, su kaynaklarının dikkatli ve verimli bir şekilde kullanılmasını zorunlu kılmaktadır.

Bu çalışmada, Güneydoğu Anadolu Bölgesi'ndeki tarımsal sulama yönetiminin optimizasyonu amaçlanmış, bölgedeki su kaynaklarının verimli kullanımı ve sulama süreçlerinin iyileştirilmesi için makine öğrenimi

algoritmaları kullanılmıştır. Bu algoritmalar; iklimsel veriler, toprak nem durumu gibi parametreleri analiz ederek en uygun sulama kararlarını verir. Projede, farklı makine öğrenimi algoritmalarının performansları karşılaştırılarak hangi algoritmanın daha iyi sonuçlar verdiği değerlendirilmiştir. Bu projede, veri toplama aşamasından itibaren kapsamlı bir yöntem uygulanmıştır.

Veri Toplama ve Veri Analizi

Sulama durumu tahmini için “Weather Data” veri kümesi kullanılmıştır (Meena, 2024). Hava durumu ile ilgili çeşitli özellikler içeren bu veri seti, sıcaklık (Temp_C), bağıl nem (Rel Hum %), toprak nem oranı (soilmoisture), hava durumu (Weather) ve pompa durumu (pump status) olmak üzere beş farklı özelliği kapsamaktadır. Toplamda 10.000 satırdan oluşan bu veri seti, tarımsal sulama sistemlerinin otomatik hale getirilmesi amacıyla sensörlerden alınan verilerin analiz edilmesi için kullanılmıştır. Bu çalışmada pompa durumu (pump status), tahmin edilmesi gereken bağımlı değişken olarak tanımlanırken; sıcaklık (Temp_C), bağıl nem (Rel Hum %), toprak nem oranı (soilmoisture) ve hava durumu (Weather) ise bağımsız değişkenler olarak kullanılmıştır.

Veri Tanımlama ve Düzenleme

Veriler öncelikle eksiklikler, aykırı değerler ve tutarsızlıklar açısından incelenmiştir. Hedef özellik olan “pompa durumu” ile bağımsız değişkenler arasındaki ilişkiler değerlendirilmiş, kategorik ve sayısal değişkenler ayrıştırılarak veri temizliği yapılmıştır. Eksik veri noktaları, veri setindeki sürekli değişkenler için önceki ve sonrakilerin ortalaması (interpolasyon) yöntemi kullanılarak tamamlanmıştır. Kategorik değişkenler için ise en sık görülen (mod) değerler doldurma yöntemi tercih edilmiştir. Veri setindeki aykırı değerler, IQR (Interquartile Range) yöntemiyle tespit edilmiş ve bu değerler eşik değerler doğrultusunda düzeltilmiştir. Tutarsız veriler, manuel inceleme ve veri doğrulama kuralları kullanılarak düzeltilmiştir. Kategorik Değişkenler: Pompa Durumu: “Açık (1)”, “Kapalı (0)”; Yağış Durumu: “Fog (0)”, “Cloudy (1)”, “Rain (2)”, “Clear (3)”, “Snow (4)”

Veri Dönüşümü

Bağımsız değişkenler arasında yer alan sıcaklık, bağıl nem, toprak nem oranı ve hava durumu gibi veriler, analiz öncesinde uygun formatlara dönüştürülmüştür. Özellikle kategorik değişken olan “Weather”, etiketleme (label encoding) yöntemiyle sayısal verilere dönüştürülmüştür. Ayrıca, modelin genelleme yeteneğinin test edilebilmesi için veri seti, eğitim (%80) ve test (%20) verisi olarak ikiye ayrılmıştır. Bu ayırım, modelin yeni verilere nasıl tepki verdiğini ölçmek ve aşırı öğrenmenin (overfitting) önüne geçmek amacıyla yapılmıştır.

Model Eğitimi

Toprak nemi, hava sıcaklığı, bağıl nem ve hava durumu gibi bağımsız değişkenler kullanılarak pompa durumunu tahmin edebilmek amacıyla Materyal ve Yöntem bölümünde sunulmuş olan makine öğrenmesi algoritmaları uygulanmıştır. “scikit-learn” kütüphanesi kullanılarak seçilen makine öğrenmesi algoritmalarıyla modeller oluşturulmuştur. Veri seti, eğitim (train) ve test setlerine

ayrılarak model eğitimleri gerçekleştirilmiştir. Eğitim sürecinde bağımsız değişkenler (sıcaklık, bağıl nem, toprak nemi, hava durumu) kullanılarak pompa durumu tahmin edilmiştir.

Model Doğrulama

Modellerin performanslarını değerlendirebilmek için doğruluk metriği (accuracy_score) kullanılmıştır. Doğruluk, modelin doğru sınıflandırdığı örneklerin tüm örneklere oranını ifade eder ve aşağıdaki formül ile hesaplanır:

$$\text{Doğruluk (Accuracy)} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP (True Positive): Doğru pozitif tahmin sayısı

TN (True Negative): Doğru negatif tahmin sayısı

FP (False Positive): Yanlış pozitif tahmin sayısı

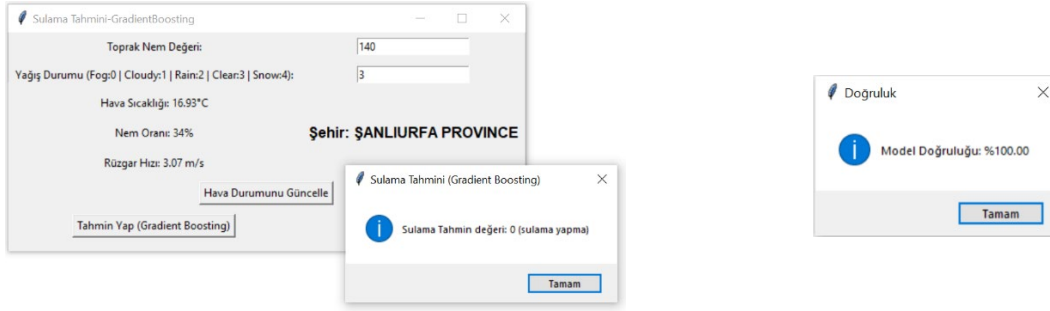
FN (False Negative): Yanlış negatif tahmin sayısı

Veri kümesi üzerinde yapılan tahminler sonucunda elde edilen doğruluk (accuracy) oranları karşılaştırılmış ve en etkili algoritmaların; Rastgele Orman (%95), Karar Ağacı (%97), Gradyan Artırma (%93) ve Yapay Sinir Ağı (%98) modeli ile tüm şehirlerde %90'ın üzerinde tahminleme doğruluğu elde edilmiştir. Diğer algoritmalar da yüksek doğruluk oranları ile (%75 üzeri) dikkate değer performans sergilemişlerdir.

Sonuçların Yorumlanması ve Raporlama

Yapılan analiz sonucunda, toprak nemi 300'ün üzerindeyse, hava sıcaklığı 30,3°C'yi aşıyor ve yağış yoksa sulama önerisi yapılmıştır. Bu koşullar altında “Sulama Tahmin Değeri: 1 (sulama yap)” şeklinde karar verilmiştir; aksi durumda “Sulama Tahmin Değeri: 0 (sulama yapma)” olarak belirlenmiştir. Bu karar destek sistemi, hem otomatik sulama sistemlerinde hem de manuel sulama süreçlerinde çiftçilere rehberlik edebilir. Otomatik sulama sistemleri, makine öğrenmesi modelleriyle entegre edilerek sulama işlemini tamamen otonom hale getirebilirken; manuel sulama yapan çiftçiler, önerilen tahminlerle sulama zamanlamalarını optimize edebilirler. Elde edilen sonuçlar raporlanmış ve hangi algoritmaların en etkili olduğu, hangi durumlarda kullanılması gerektiği belirlenmiştir. Sonuçlar, çiftçilerin mevcut meteorolojik parametreleri dikkate alarak tarımsal sulama işlemlerini optimize edebilmek için makine öğrenmesi modellerini hem manuel hem de otomatik sistemlerde kullanabileceklerini göstermektedir.

Ayrıca, Open Weather Map API kullanılarak anlık hava durumu verileri sistemde entegre edilmiştir. “update_weather” fonksiyonu ile şehirlerin hava durumu güncel verilere göre çekilmiş ve değerlendirilmiştir. Bu fonksiyon, sulama sisteminin hava durumunu gerçek zamanlı güncelleyebilmesine olanak tanımaktadır. Tahmin işlemi ise “predict” fonksiyonu ile yapılmış ve toprak nemi ve hava durumu bilgileri dayanarak sulama kararı alınmıştır. Bu süreç, Tkinter kullanılarak bir grafiksel kullanıcı arayüzü (GUI) üzerinden kullanıcıya sunulmuştur (Şekil 1). Arayüzde, hava durumu güncellemesi ve tahmin yapılması için butonlar eklenmiş, kullanıcılar kolayca güncel durumu görebilmiş ve tahmin sonuçlarına erişebilmiştir. Bu çalışmada kullanılan algoritmanın kaba kodu Çizelge 1’de sunulmuştur.



Şekil 1. Kullanıcı arayüzü
Figure 1. User Interface

Çizelge 1. Geliştirilen algoritmanın kaba kodu

Table 1. Pseudocode of the developed algorithms

Adım	Açıklama
Başla	Uygulamanın çalışmaya başlaması.
Veri Kümesini Yükle	Weather-Data1.csv dosyasından verileri yükle.
Modeli Oluştur ve Eğit	Karar Ağaçları modelini oluştur. Bağımsız değişkenleri (X) ve hedef değişkeni (y) belirle. Modeli eğit.
Pencereyi Başlat	Tkinter arayüzünü oluştur.
Etiketleri ve Giriş Kutularını Oluştur	Hava durumu bilgileri için etiketler, kullanıcı girişleri için kutular oluştur.
Butonları Ekle	“Hava Durumunu Güncelle” ve “Tahmin Yap” butonlarını ekle.
Fonksiyon: hava_durumu_al(api_anahatari, sehir)	API anahtarı ve şehir adı ile URL oluştur. API’den hava durumu verilerini al ve JSON formatında döndür.
Fonksiyon: hava_durumu_guncelle()	Hava durumu verilerini almak için hava_durumu_al fonksiyonunu çağır. Sıcaklık, nem, şehir, rüzgar hızı bilgilerini ekranda göster.
Fonksiyon: tahmin_yap()	Kullanıcıdan toprak nemi ve yağış durumu bilgilerini al. Model için giriş oluştur ve tahmin yap. Tahmine göre sulama mesajını göster. Modelin doğruluk oranını hesaplayıp ekranda göster.
Pencereyi Çalıştır	Tkinter ana döngüsünü başlatarak uygulamanın çalışmasını sağla.
Biti	Uygulama sonlanır.

Bu çalışmada kullanılan makine öğrenmesi algoritmaları, Türkiye’nin Güneydoğu Anadolu Bölgesi’ndeki Şanlıurfa, Mardin, Diyarbakır, Gaziantep ve Adıyaman illerine ait tarımsal sulama verileri üzerinde uygulanmış ve elde edilen sonuçlar karşılaştırmalı olarak analiz edilmiştir. Bu veriler, sulama yönetimini optimize etmek amacıyla çeşitli bağımsız değişkenlerden (örneğin: hava sıcaklığı, toprak nem oranı, bağıl nem, yağış durumu gibi meteorolojik veriler) ve bağımlı değişken olan “pompa durumu”ndan (sulama yapıp yapılmadığı bilgisi) oluşmaktadır.

Algoritmaların performansı, bu veriler kullanılarak değerlendirilmiş ve hangi algoritmaların sulama yönetiminde daha etkili olduğu incelenmiştir. Veriler, sulama yönetimini iyileştirmeyi amaçlayan tahminler yapmak için işlenmiş ve karşılaştırmalı analizlerde hangi modelin daha yüksek doğruluk sağladığı belirlenmiştir.

Bulgular

Bu çalışmada, Güneydoğu Anadolu Bölgesi’nin hava koşulları ve toprak yapısı değerlendirilerek tarımsal sulama tahmini için farklı makine öğrenmesi algoritmalarının performansları karşılaştırılmıştır. Bölgenin yarı kurak ve sıcak iklimi, su kaynaklarının sınırlı olması ve tarımsal üretimin büyük ölçüde sulamaya bağlı olması nedeniyle doğru sulama yönetimi hayati bir öneme sahiptir.

Çalışmada, Weather Data veri kümesi kullanılarak sulama durumu tahmini gerçekleştirilmiştir. Bu veri kümesi, hava sıcaklığı, nem, yağış miktarı ve toprak nemi gibi değişkenleri içermektedir. Bölgedeki iller üzerinde yapılan deneylerde, Karar Ağaçları (Decision Tree), Rastgele Orman (Random Forest), Gradient Boosting, Yapay Sinir Ağı (Artificial Neural Network - ANN), Destek Vektör Makineleri (Support Vector Machine - SVM), K-En Yakın Komşu (KNN), Naive Bayes ve Lojistik Regresyon algoritmaları değerlendirilmiştir.

Her bir algoritma, aynı eğitim ve test veri seti kullanılarak eğitilmiş ve sonuçları Çizelge 2’de sunulmuştur. Elde edilen bulgular, Yapay Sinir Ağı (ANN), Karar Ağaçları ve Gradient Boosting algoritmalarının en yüksek doğruluk oranlarına ulaştığını göstermektedir.

Yapay Sinir Ağı, %98 doğruluk oranıyla en yüksek performansı göstermiş, Karar Ağaçları algoritması ise %97 doğruluk oranıyla oldukça başarılı sonuçlar vermiştir. Şanlıurfa, Diyarbakır ve Mardin gibi tarımsal üretimin yoğun olduğu illerde bu iki algoritma özellikle öne çıkmıştır. Diğer algoritmalarla il bazında yapılan karşılaştırmalar da Çizelge 3’de gösterilmiştir. Buna göre, Rastgele Orman algoritması %95 doğruluk oranına ulaşmış, ancak K-En Yakın Komşu (KNN) ve Naive Bayes gibi daha basit algoritmaların doğruluk oranları %88 ve %84 civarında kalmıştır. Bu bulgular, daha karmaşık ve büyük veri setlerinde basit algoritmaların etkinliğinin sınırlı olduğunu göstermektedir.

Çizelge 2. Makine öğrenmesi algoritmalarının performans analizi
Table 2. Performance analysis of machine learning algorithms

Algoritma	Doğruluk (%)
Karar Ağaçları	97
Rastgele Orman	95
Destek Vektör Makineleri	88
Gradient Boosting	93
Yapay Sinir Ağı (ANN)	98
K-En Yakın Komşu (KNN)	83
Lojistik Regresyon	75

Çizelge 3. İl bazında makine öğrenmesi algoritmalarının doğruluk metriği performansı (%)
Table 3. Accuracy metric performance of machine learning algorithms for cities

İl	Karar Ağaçları (%)	Rastgele Orman (%)	SVM (%)	Gradient Boosting (%)	Yapay Sinir Ağı (%)	KNN (%)	Lojistik Regresyon (%)
Adıyaman	96	94	87	92	97	88	85
Batman	95	93	86	91	96	87	84
Diyarbakır	98	96	89	94	99	90	86
Gaziantep	96	94	89	92	97	88	85
Kilis	94	92	89	90	95	86	83
Mardin	97	95	89	93	98	89	84
Siirt	95	93	89	91	96	87	83
Şanlıurfa	97	95	89	93	98	89	86
Şırnak	94	92	89	90	95	86	83

Çizelge 3'deki sonuçlar incelendiğinde, ANN ve Karar Ağaçları'nın, Şanlıurfa ve Diyarbakır gibi geniş tarım arazilerine sahip illerde en iyi sonuçları verdiği görülmektedir. Bu illerde tarımın büyük ölçüde sulamaya bağımlı olması ve su kaynaklarının dikkatli yönetilmesi gerekliliği, bu algoritmaların önemini bir kez daha ortaya koymaktadır.

Bu algoritmalar, sulama planlamasında verimliliği artırarak su israfını minimuma indirmiştir. Özellikle toprak nemi, hava sıcaklığı, bitki türü ve yağış miktarı gibi çevresel faktörler analiz edilerek, her tarım alanı için en uygun sulama zamanı ve miktarı belirlenmiştir. Yapay Sinir Ağı, sulama yönetimi için büyük veri setleriyle çalışabilmesi ve değişkenler arasındaki karmaşık ilişkileri başarılı bir şekilde modelleyebilmesi sayesinde en iyi sonuçları sağlamıştır. Karar Ağaçları ise, daha anlaşılır ve takip edilebilir karar süreçleri sunması nedeniyle tarım üreticileri için pratik bir karar destek aracı olarak kullanılabilir.

Tartışma

Güneydoğu Anadolu Bölgesi'nde sulama yönetimi, makine öğrenmesi algoritmaları sayesinde daha etkin hale getirilebilmektedir. İklimsel koşulların kurak ve düzensiz olması, tarımda su yönetiminin önemini artırmakta, ANN ve Karar Ağaçları algoritmaları bu ihtiyaçlara en uygun çözümleri sunmaktadır. ANN, karmaşık ve büyük veri setlerini başarıyla işleyerek yüksek doğruluk oranlarına ulaşmıştır. Karar Ağaçları ise, açıklayıcı yapısı ile tarım yöneticilerinin sulama kararlarını daha net bir şekilde almasına olanak sağlamaktadır.

Makine öğrenmesi algoritmalarıyla elde edilen bu bulgular, tarımsal üretimde sürdürülebilirlik ve su tasarrufunun sağlanması için önemli bir fırsat sunmaktadır.

Gelecekte IoT tabanlı veri toplama sistemlerinin yaygınlaşması ve bu verilerin makine öğrenmesi algoritmalarıyla işlenmesi, sulama yönetimini daha dinamik ve veriye dayalı bir yapıya kavuşturabilir. Böylece, su kaynakları daha verimli kullanılacak ve tarımsal üretim verimliliği artırılacaktır. Bu çalışmalar, özellikle Güneydoğu Anadolu gibi su kaynakları sınırlı olan bölgelerde büyük bir potansiyel taşımaktadır. Makine öğrenmesi algoritmaları ile entegre edilen bu sistemler, suyun ekonomik ve sürdürülebilir bir şekilde yönetilmesine katkıda bulunacaktır.

Sonuç ve Öneri

Bu çalışma, Güneydoğu Anadolu Bölgesi'nde tarımsal sulama yönetimi için farklı makine öğrenmesi algoritmalarının performanslarını değerlendirmiştir. Elde edilen bulgular, Yapay Sinir Ağı (ANN) ve Karar Ağaçları algoritmalarının sulama tahminlerinde en yüksek doğruluk oranlarına ulaştığını göstermektedir. Şanlıurfa, Diyarbakır ve Mardin gibi tarımsal üretimin yoğun olduğu illerde ANN algoritması %98, Karar Ağaçları algoritması ise %97 doğruluk oranı sağlamıştır. Bu sonuçlar, gelişmiş algoritmaların tarımsal sulama yönetiminde etkin bir karar destek sistemi olarak kullanılabilirliğini ortaya koymaktadır.

Makine öğrenmesi algoritmalarının tarımsal sulama sistemlerine entegrasyonu, su kaynaklarının daha verimli yönetilmesini ve israfın önlenmesini sağlayabilir. Özellikle hava durumu, toprak nemi, bitki türü ve yağış miktarı gibi değişkenler analiz edilerek, her bölge için optimum sulama stratejileri geliştirilebilir. ANN gibi karmaşık modeller, büyük veri setlerinden anlamlı sonuçlar çıkararak iklim değişikliği gibi çevresel faktörlere uyumlu dinamik sulama yönetimi sunabilir.

Gelecekteki çalışmalar için şu öneriler getirilebilir:

- *IoT ve sensör teknolojileri ile entegrasyon:* Toprak nem sensörleri ve hava durumu istasyonlarından gerçek zamanlı veri toplanarak modellerin doğruluk oranları artırılabilir.
- *Bölgesel ve ürün bazlı optimizasyon:* Farklı bitki türleri ve toprak yapıları göz önünde bulundurularak bölgeye özel sulama planları geliştirilebilir.
- *Derin öğrenme modellerinin karşılaştırılması:* ANN'nin yanı sıra derin öğrenme tabanlı algoritmalar (CNN, LSTM, Transformer vb.) test edilerek, uzun vadeli tahminlerdeki başarısı değerlendirilebilir.
- *Karar destek sistemlerinin geliştirilmesi:* Çiftçilerin ve tarım yöneticilerinin kolayca kullanabileceği, mobil veya web tabanlı arayüzler oluşturularak, sulama tahminlerinin sahada uygulanması sağlanabilir.

Bu araştırmanın sonuçları, su kıtlığıyla mücadelede ve tarımsal verimliliğin artırılmasında makine öğreniminin büyük bir potansiyel taşıdığını göstermektedir. Gelecekte daha geniş kapsamlı çalışmalar ve saha uygulamaları ile bu sistemlerin tarım sektörüne entegrasyonu hızlandırılabilir.

Beyanlar

Yazar katkısı

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan eder

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan eder.

Kaynaklar

- Akkaya, G. (2007). Yapay sinir ağları ve tarım alanındaki uygulamaları. Atatürk Üniversitesi Ziraat Fakültesi Dergisi, 38(2), 195-202.
- Başakın, E. E., Ekmekçioğlu, Ö., & Özger, M. (2019). Makine öğrenmesi yöntemleri ile kuraklık analizi. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 25(8), 985-991.
- Bayrakçı, H. C., Çiçekdemir, R. S., & Özkahraman, M. (2021). Tarım Arazilerinde Harcanan Su Miktarını Yapay Zekâ Teknikleri Kullanarak Belirlenmesi. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 9(6), 237-250.
- Bircan, H. (2004). Lojistik regresyon analizi: Tıp verileri üzerine bir uygulama. Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2, 185-208.
- Bütüner, R., Şimşek, M., Aksöz, G., & Sadi, Ş. A. N. (2023). Yapay Sinir Ağları ile Mısır Yapraklarında Hastalık Tespiti. Journal of Information Systems and Management Research, 5(2), 21-31.
- Çakmak, B., & Gökalp, Z. (2013). Kuraklık ve tarımsal su yönetimi. Gaziosmanpaşa Bilimsel Araştırma Dergisi, (4), 1-11.
- Çakmak, B., Yıldırım, M., & Aküzüm, T. (2008). Türkiye'de tarımsal sulama yönetimi, sorunlar ve çözüm önerileri. TMMOB, 2, 215-224.

- Demir, E. (2003). Güneydoğu Anadolu Projesinin Ülke Ekonomisine Katkısı ve Bölge Yerleşim Alanlarına Etkisi. Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi, 23(3).
- Demirci, M. (2019). Destek vektör makineleri ve M5 karar ağacı yöntemleri kullanılarak yağış akış ilişkisinin tahmini. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 10(3), 1113-1124.
- Ergül Aydın, Z., İçmen Erdem, B., & Erzurum Çiçek, Z. İ. (2023). Prediction bike-sharing demand with gradient boosting methods. Pamukkale University Journal of Engineering Sciences, 29(8), 824-832.
- Güman, Z., & Gunay, F. B. Nesnelerin İnterneti Yardımıyla Akıllı Tarımda Yapay Zekâ Tabanlı Gübre ve Mahsul Tahmini. Journal of the Institute of Science and Technology, 14(3), 958-973.
- Güneydoğu Anadolu Projesi (GAP) Bölge Kalkınma İdaresi Başkanlığı. (2023). GAP Son Durum Raporu 2023. GAP Bölge Kalkınma İdaresi Başkanlığı. http://www.gap.gov.tr/upload/dosyalar/pdfler/2023_GAP_SON_DURUM_RAPORU.pdf
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.
- Kiliç, S. (2013). Linear regression analysis. Psychiatry and Behavioral Sciences, 3(2), 90.
- Kendirli, B. (2001). Harran ovası sulama birliklerinde antepfıstığının sulama planlaması. Journal of Agricultural Sciences, 7(04), 114-120.
- Meena, S. (2024). Weather Data [Data set]. Kaggle. Retrieved from <https://www.kaggle.com/code/sachinmeena04/weather-data-analysis/input?select=Weather+Data.csv>
- Nusrat, F., Uzbaş, B., & Baykan, Ö. K. (2020). Prediction of diabetes mellitus by using gradient boosting classification. Avrupa Bilim ve Teknoloji Dergisi, 268-272.
- Özcan, B., & Özer, E. H. (2021). Karar Teorisinin Tarım Sektöründe Uygulanması. Uluslararası Doğu Anadolu Fen Mühendislik ve Tasarım Dergisi, 3(1), 257-281.
- Öztürk, K., & Şahin, M. E. (2018). Yapay sinir ağları ve yapay zekâ'ya genel bir bakış. Takvim-i Vekayi, 6(2), 25-36.
- Sargıncı, M., Dönmez, A. H., & Yıldız, O. (2022). Biyosferde Su Döngüsü ve Bitkiler Tarafından Suyun Alınma Mekanizmaları. Düzce Üniversitesi Süs ve Tıbbi Bitkiler Botanik Bahçesi Dergisi, 1(1), 8-20.
- Sevgen, S. C., & Aliefendioğlu, Y. (2020). Mass appraisal with a machine learning algorithm: random forest regression. Bilişim Teknolojileri Dergisi, 13(3), 301-311.
- Şahin, H. (2024). Tarımsal Akıllı Sulama Sistemlerinde Yapay Zekâ, Derin Öğrenme ve Nesnelerin İnterneti Uygulamaları. Tarım Makinaları Bilimi Dergisi, 20(1), 41-60.
- Taştan, M. (2019). Nesnelerin İnterneti Tabanlı Akıllı Sulama ve Uzaktan İzleme Sistemi. Avrupa Bilim ve Teknoloji Dergisi, (15), 229-236.
- Terzi, İ., Özgüven, M. M., Altaş, Z., & Uygun, T. TARIMDA YAPAY ZEKA KULLANIMI.
- Türkes, M. (2012). Türkiye'de gözlenen ve öngörülen iklim değişikliği, kuraklık ve çölleşme. Ankara Üniversitesi Çevre Bilimleri Dergisi, 4(2), 1-32.
- Uzun, E. (2007). İnternet tabanlı bilgi erişimi destekli bir otomatik öğrenme sistemi.
- Yıldız, O., Dönmez, A. H., & Sargıncı, M. (2022). Su ve bitki sağlığına etkisi: Bitkilerde suyun alınma mekanizmaları. Düzce Üniversitesi Süs ve Tıbbi Bitkiler Botanik Bahçesi Dergisi, 1(1), 8-20.