



BİLİM-TEKNOLOJİ-YENİLİK EKOSİSTEMİ DERGİSİ

JOURNAL OF SCIENCE-TECHNOLOGY-INNOVATION ECOSYSTEM

E-ISSN : 2757-6140

Volume : 7

Issue : 2

Year : 2026



RESEARCH ARTICLE

Vis-NIR Spektral Yansıma ve Makine Öğrenmesi Kullanarak Domateste Sertlik Tahmini

Prediction of Tomato Firmness Using Vis-NIR Spectral Reflection and Machine Learning

Aysun İğdeci¹ , İsmail Kavdır²

¹ Çanakkale Onsekiz Mart Üniversitesi, Lisansüstü Eğitim Enstitüsü, Tarım Makinaları ve Teknolojileri Mühendisliği Anabilim Dalı, Çanakkale, Türkiye

² Çanakkale Onsekiz Mart Üniversitesi, Ziraat Fakültesi, Tarım Makinaları ve Teknolojileri Mühendisliği Bölümü, Çanakkale, Türkiye

MAKALE BİLGİSİ

Araştırma Makalesi

Geliş: 04.06.2026
Kabul: 24.06.2026
Yayın: 01.07.2026

Anahtar Kelimeler:

Makine öğrenmesi
Vis/NIR spektroskopisi
Domates sertlik

Mizanpaj Editörü:

Habibe Doğan

ÖZ

Bu çalışmada domateslerde sertlik özelliğinin hasarsız olarak tahmin edilmesi için iki farklı ışık kaynağı kullanarak iki farklı yansıma modunda alınan Vis/NIR spektral okumaları, standart yöntemler ile ölçülen sertlik özellikleri ve çalışma kapsamında geliştirilen doğrusal ve makine öğrenmesine dayalı doğrusal olmayan farklı tahmin modelleri kullanılmıştır. Çalışmada hedeflenen en önemli amaçlardan biri doğrusal tahmin modelleri ile makine öğrenmesine ve derin öğrenme tekniklerine dayalı olarak geliştirilen tahmin modellerinin domates sertlik özelliklerinin tahmin edilmesi sürecinde karşılaştırılması olmuştur. Bu kapsamda aynı veri grubunu kullanarak daha önceki bir çalışmada geliştirilip kullanılan PLS (kısmi en küçük kareler) regresyon modeli ile bu çalışma kapsamında geliştirilen ve kullanılan 1DCNN (bir boyutlu evreşimli sinir ağları) regresyon, LSTM (uzun-kısa süreli bellek) regresyon, SVM (destek vektör makinesi) regresyon ve PCR (temel bileşen regresyonu) regresyon modelleri ile elde edilen sertlik tahmin sonuçları karşılaştırılmıştır. Elde edilen sonuçlara bakıldığında DH-2000-BAL ışık kaynağı yardımı ile alınan yansıma modu spektrumlarının kullanılması ile daha iyi sertlik tahmin değerlerinin alındığı sonucuna varılmıştır. Bunun yanında, çalışmada ele alınan farklı sertlik tahmin modelleri kıyaslandığında, DH-2000-BAL ışık kaynağının kullanılması ile alınan spektrumların kullanıldığı makine öğrenmesine dayalı 1DCNN regresyon modeli ile elde edilen tahmin sonuçlarına ait kararlılık katsayısı (R^2) 0.83 olurken (RMSEP=263.49), çalışma kapsamında geliştirilen tüm tahmin modelleri (1DCNN, LSTM, SVM ve PCR) ile elde edilen sertlik tahmin değerleri önceki çalışmada PLS regresyon modeli ile elde edilen sertlik tahmin değerlerinden daha iyi olmuştur.

ARTICLE INFO

Research Article

Received: 04.06.2026
Accepted: 24.06.2026
Published: 01.07.2026

ABSTRACT

This study utilized Vis/NIR spectral readings obtained in two different reflection modes using two different light sources, firmness features measured using standard methods, and various linear and machine learning-based nonlinear prediction models developed within the scope of this study to predict tomato firmness characteristics non-destructively. One of the most important objectives of the study was to compare linear prediction models with the models developed based on machine learning and deep learning techniques in the process of estimating tomato firmness properties. In this context, using the same data set, the firmness prediction results obtained using the PLS (partial least squares) regression model developed and used in a previous study were

✉ Correspondence: ikavdir17@gmail.com (İsmail Kavdır)

Citation: İğdeci, A., & Kavdır, İ. (2026). Vis-NIR spektral yansıma ve makine öğrenmesi kullanarak domateste sertlik tahmini. *Bilim-Teknoloji-Yenilik Ekosistemi Dergisi*, 7(2), 115-130.

This work is licensed under a Creative Commons Attribution CC-BY 4.0 International License.

Keywords:

Machine learning
Vis/NIR spectroscopy
Tomato firmness

compared with the 1DCNN (one-dimensional convolutional neural networks) regression, LSTM (long-short-term memory) regression, SVM (support vector machine) regression, and PCR (principal component regression) regression models developed and used in this study. Based on the results obtained, it was concluded that better firmness prediction values were obtained by using reflection mode spectra acquired with the help of the DH-2000-BAL light source. Furthermore, when comparing the different firmness prediction models considered in this study, the coefficient of determination (R^2) of the prediction results obtained with the machine learning-based 1DCNN regression model using spectra acquired with the DH-2000-BAL light source was 0.83 (RMSEP=263.49), while the firmness prediction values obtained with all the prediction models developed in this study (1DCNN, LSTM, SVM, and PCR) were better than the firmness prediction values obtained with the PLS regression model in the previous study.

Giriş

Üretimi ve tüketimi mevsimsel bağımlılıktan kurtularak giderek artan oranlarda talep görmekte olan domatestte bir kalite kriteri olarak sertlik, hasat sonrası değerlendirme süreçlerinde ele alınan başlıca kalite göstergelerinden biri olmaktadır. Diğer taraftan hasat sonrası kalite değerlendirme süreçlerinin giderek otomasyona doğru yönelmesi ve bu alanda artan dijitalleşme ve sensör teknolojilerinin bir sonucu olarak tarımsal ürünlerin iç ve dış özelliklerinin hasarsız teknikler ile algılanması kaçınılmaz bir gelişme olarak karşımıza çıkmaktadır. Bu bağlamda, Vis/NIR (Visible/Near-Infrared) spektroskopi tarımsal ürünlerin kalite parametrelerinin algılanmasında önemli potansiyeller sunmaktadır. Bunun yanında, yapay zekâ ile yakın ilişki içinde olan makine öğrenmesi tekniklerinin de farklı alanlarda ve tarımsal süreçlerde giderek yaygın olarak kullanıldığı görülmektedir. Bu gelişmelerin sonucu olarak, Vis/NIR spektral teknikler ile makine öğrenmesi yaklaşımlarının birlikte kullanılması tarımsal ürünlerin hasarsız ve hızlı bir şekilde değerlendirilmesinde kabul gören ve gelişime açık alanlar olarak değerlendirilmektedir.

Vis/NIR spektroskopisi, özellikle yüksek nem ihtiva eden sağlam meyve ve sebzelerin iç ve yüzey özelliklerinin kalite kontrolünde, kalitatif ve kantitatif analizlerde güvenilir sonuçlar vermesi nedeniyle avantaj sağlamaktadır (Kavdır ve ark., 2009; Jamshidi ve ark., 2016). Doğrusal analizlerde başarılı sonuçlar veren kısmi en küçük kareler (PLS) regresyonu ile tarımsal ürünlerin sertlik özelliğinin tahmin edilmesine yönelik araştırmalara literatürde sıklıkla rastlanılmaktadır. Ecartot ve ark. (2013) domateslerde sertlik tahmini için Vis/NIR spektroskopi ve PLS regresyonu kullanmışlar ve sertlik tahminini $R^2=0.76$, $RPD=2.14$ korelasyon değerleri ile bulmuşlardır. Huang ve ark. (2018) Vis/NIR spektroskopi ve PLS regresyon tekniklerini kullanarak domatestte darbe sertliği, sıkıştırma alanı ve delinme eğimi özelliklerinin tahmini için sırasıyla 0.899, 0.917 ve 0.935 regresyon kararlılık katsayısı değerlerini elde ederken, NIR spektroskopi ölçümlerini kullanarak 0.846, 0.831 ve 0.853 kararlılık katsayısı değerlerini (R^2) elde etmişlerdir.

Ürünün kimyasal yapısı hakkında bilgiler içeren spektral verilerin anlamlı hale getirilmesinde, bilinen kemometrik hesaplama araçlarına alternatif olarak geliştirilen makine öğrenmesi algoritmaları, amaca yönelik güçlü model doğruluğu sunmaktadırlar (Kazemi ve ark., 2026). Makine öğrenmesi, tarımsal uygulamalarda yoğun veri üretilen süreçleri çözmek, ölçülebilir hale getirmek ve anlamak için yeni fırsatlar yaratmak üzere büyük veri teknolojileri ve yüksek performanslı hesaplama arayışı ile birlikte kullanılmaya başlanmıştır (Liakos ve ark., 2018). Geniş bir uygulama yelpazesine sahip olan makine öğrenmesi yöntemleri, görüntü tanıma, doğal dil işleme, lineer olmayan sistem davranışlarının modellenmesi, tahmin ve karar verme gibi alanlarda verimli bir şekilde kullanılmaktadır. Özellikle spektrum verilerinin girdi olarak kullanıldığı makine öğrenmesi (ML-machine learning) tabanlı tahmin modellerinde; temel bileşen regresyonu (PCR), destek vektör makinesi (SVM), evrişimli sinir ağları (CNN), yinelemeli sinir ağları/uzun-kısa süreli bellek (RNN/LSTM), rasgele orman (RF), k-en yakın komşu (KNN) ve gradyan arttırmalı karar ağaçları (XG Boost) gibi mimarilerin kullanımının yaygınlaştığı görülmektedir.

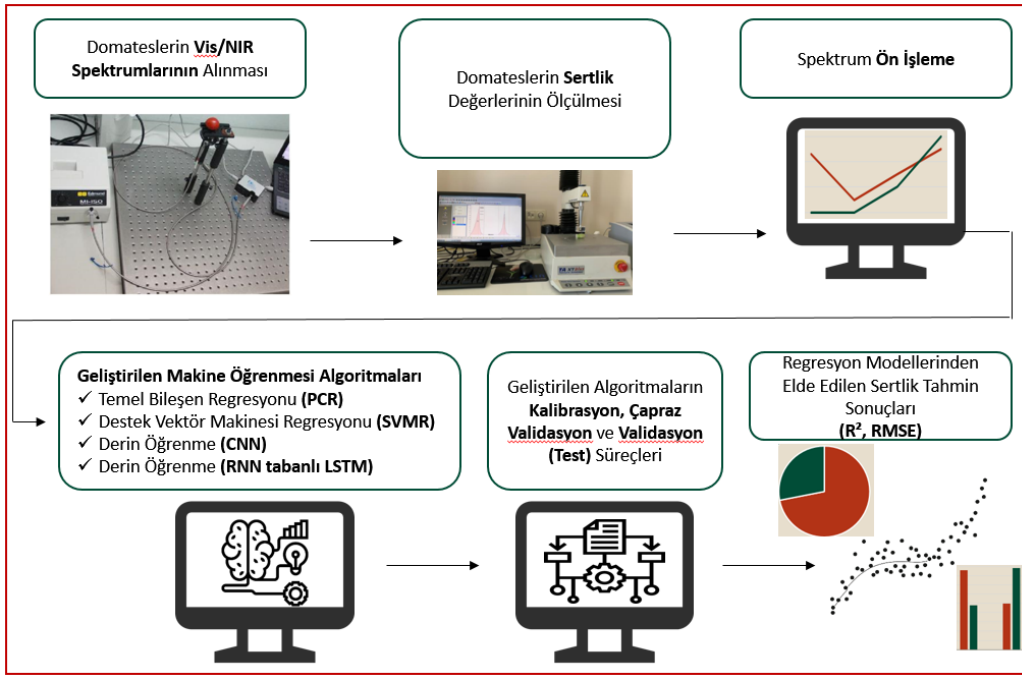
Yalçın ve ark. (2026) 118 oğlaktan elde ettikleri verileri kullanarak duyuşal et kalitesini tahmin etmek amacıyla sekiz farklı makine öğrenmesi algoritmasını (DT-karar ağacı, AB-AdaBoost, GB-gradient boost, RF, KNN, ANN- yapay sinir ağları ve SVM) karşılaştırmıştır. Et yumuşaklığı tahmininde PLS ($R^2=0.754$) ve SVM ($R^2=0.736$) algoritmalarının düşük hata payları ile öne çıkan modeller olduğu (sırasıyla %8.78 ve %8.38 MAPE değerleri ile) bildirilmiştir. Chakravartula ve ark. (2022), FT-NIR (Fourier transform yakın kızıl ötesi) spektroskopisi ile birlikte kullanılan CNN modelinin, spektral verilerden doğrudan özellik çıkarımı yaparak kahvedeki tağşiş içeriğini başarıyla tahmin edebildiğini ortaya koymuşlardır. Yang ve ark. (2020), toprak özelliklerini tahmin etmek amacıyla Vis-NIR spektrum verilerini CNN ve RNN mimarilerinin bir birleşimi olan hibrit bir model (CCNVR) kullanarak gerçekleştirmişlerdir. Söz konusu hibrit modelin performansı PLS, SVM, CNN, ANN ve RNN yöntemleri ile kıyaslanmış ve en başarılı sonuçların CCNVR modelinden elde edildiği ve bunu sırasıyla CNN, SVM ve RNN modellerinin izlediği kaydedilmiştir. Benzer şekilde Mou ve ark. (2017), hiperspektral görüntülerden elde edilen spektral verilerin ardışık (sıralı) yapıda olduğunu vurgulayarak, bu verilerin bir RNN türevi olan LSTM ağları için uygun bir girdi kaynağı oluşturduğunu belirtmişlerdir.

Ürün özelliklerinin tahminine yönelik önceki çalışmalara bakıldığında spektral verilerin yapısına ve ürün değişkenliğine bağlı olarak kullanılan parametrelerin bazı durumlarda farklı makine öğrenmesi kombinasyonlarını gerektirdiği görülebilmektedir. Kullanılan modellerin yer yer hibrit yapılara dönüştüğü ve bu dinamik süreçte, optimum sonuçların ancak ampirik yöntemlerle elde edilebildiği görülmektedir. Diğer taraftan, tarımsal ürünlerin endüstriyel ürünler gibi yekpare olmaması, iç ve dış kalite parametrelerinin hem birbirleri arasında hem de anlık değişiklik göstermesi gerçek zamanlı ve etkili ölçme ve veri değerlendirme teknikleri arayışlarını arttırmaktadır. Doğru enstrümanlarla güvenilirliği yüksek analizler yapabilmek, büyük veri setleriyle çalışmaya imkân tanıyarak bu tekniklerin sahada kullanım yollarını açmaktadır. Son zamanlarda yeni anılmaya başlanan Tarım 5.0'ın da gündeme aldığı en önemli konulardan biri olan çevresel, sosyal ve ekonomik sürdürülebilirliğin temelinde olan ve aynı zamanda, tarım teknolojilerinin de bir gerekliliği olarak veri toplama, verilerin doğru işlenmesi ve sonuç olarak güvenilir karar verme ve kontrol mekanizmalarının oluşturulması, giderek önem kazanan konular arasında görülmektedir (Bissadu ve ark., 2025).

Bu çalışmada, iki farklı ışık kaynağı ile alınan Vis/NIR spektral yansıma ölçümlerinin, standart sertlik ölçümlerinin ve PCR, SVM, CNN ve RNN tabanlı LSTM makine öğrenme algoritmalarının kullanılması ile domatestte sertlik tahmini için regresyon modellerinin geliştirilmesi ve birbirleri ile karşılaştırılması amaçlanmıştır. Çalışma kapsamında geliştirilen makine öğrenmesi algoritmaları ile elde edilen domates sertlik tahmin sonuçları kendi aralarında ve daha önce aynı veri grubu için elde edilen PLS regresyon analizi sonuçları (İğdeci, 2020) ile karşılaştırılmıştır.

Materyal ve Yöntem

Bu çalışmada kullanılan kiraz domates örneklerine ait spektral veriler ve standart sertlik ölçüm değerleri daha önce yürütülen ve tamamlanan bir çalışmaya ait verilerdir (İğdeci, 2020). İlgili çalışmada materyal olarak Yingzi F1 (*Lycopersicum esculentum* mill. Cv.) kiraz domates çeşidi ele alınmıştır. Kiraz domatesler yaklaşık 25 günlük bir periyotta haftada iki kez olacak şekilde 8 farklı günde ve her defasında yaklaşık 90 domates (toplam 720 örnek) olacak şekilde hasat edilmişlerdir. Domateslerin yaklaşık bir aylık bir süreç içerisinde aralıklı olarak hasat edilmeleri ile elde edilen örneklerin farklı renk ve olgunluk seviyelerinde olmaları sağlanmıştır (yeşil, yeşil-sarı, sarı-turuncu, turuncu, açık kırmızı ve kırmızı olacak şekilde 6 farklı olgunluk düzeyi). Spektral ölçümler hasattan bir gün sonra, sertlik ölçümleri de spektral ölçümlerden bir gün sonra gerçekleştirilmiştir. Ölçümler dışındaki sürelerde domates örnekleri 4 °C sıcaklıkta soğutucuda depolanmış ve ölçümlerden yaklaşık 3 saat önce soğutucudan çıkartılarak oda sıcaklığında (24°C) bekletilmişlerdir (İğdeci, 2020). Çalışma kapsamında gerçekleştirilen işlemler Şekil 1'de verilmiştir.

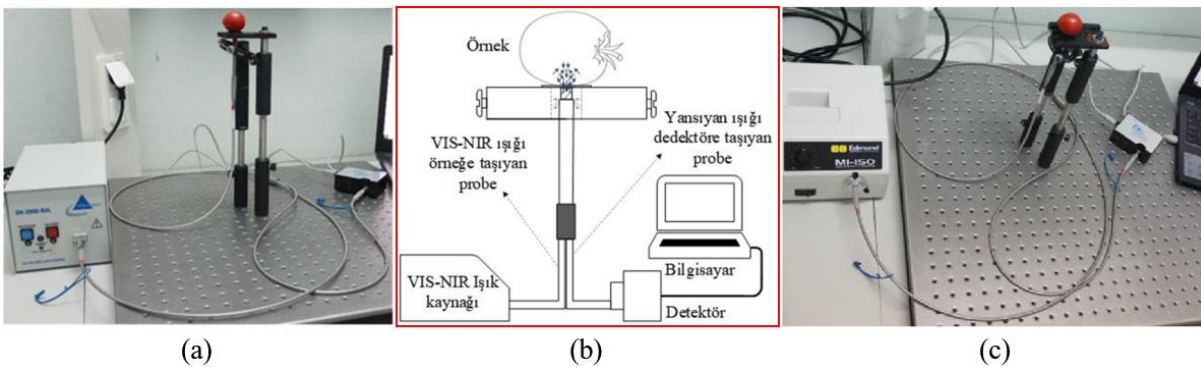


Şekil 1. Çalışmada yürütülen ölçüm ve model geliştirme süreçlerinin şeması

Figure 1. Diagram of the measuring and model development procedures carried out in the study

Vis/NIR Spektroskopi Düzenneği

Kiraz domateslerin spektral ölçümlerinin alımında spektrometre olarak 16 bit' lik CCD tabanlı USB4000-Vis/NIR spektrometre (Ocean Insight Inc., Duiven, The Netherlands) kullanılmıştır (İğdeci, 2020). Kullanılan spektrometrenin optik çözünürlüğü yaklaşık 1.5 nm ve dinamik aralığı 1300 (en iyi kapasitenin karanlık akım düzeyine oranı olarak tanımlanan) olmuştur. Bu spektrometre ile iki farklı yansıma modunda yapılan spektral ölçümler Vis/NIR bölgede 350-1100 nm dalga boyu aralığında gerçekleştirilmiştir. Örneklerin aydınlatılması için farklı etkili dalga boyu aralıklarına sahip iki farklı halojen ışık kaynağı kullanılmıştır. Birinci ışık kaynağı olarak 360-1500 nm arasında etkili dalga boyuna sahip halojen DH-2000-BAL (Ocean Insight Inc., Duiven, The Netherlands) ışık kaynağı kullanılırken (Şekil 2a-b), ikinci ışık kaynağı olarak 350-800 nm arasında etkili dalga boyu aralığına sahip halojen MI-150 ışık kaynağı (Dolan-Jenner Ind. 159 Swanson Road Boxborough, MA USA) kullanılmıştır (Şekil 2b-c). Spektrumların alımında OceanView (Version 1.6.7, Ocean Insight Inc., Duiven, The Netherlands) yazılımı kullanılmıştır. Bu spektrometre ile yapılan ölçümlerde örnekten gelen ışığı dedektöre taşıyan fiber optik problemlerin sabitlendiği ve spektrometrenin konumlandırıldığı titreşimlere karşı hassas, optik ölçümlerde kullanılan 596 x 596 x 46 mm ebatlarında paslanmaz çelik kaplama bir tezgahtan faydalanılmıştır (TMC, 15 Centennial Drive Peabody, MA 01960 USA) (İğdeci, 2020).



Şekil 2. (a) DH-2000-BAL ışık kaynağı ile spektral ölçüm konfigürasyonu (b) yansıma modu ölçümlerinin şematik gösterimi (c) MI-150 ışık kaynağı ile spektral ölçüm konfigürasyonu

Figure 2. (a) Spectral measurement configuration with DH-2000-BAL light source (b) Schematic representation of reflection mode measurements (c) Spectral configuration with MI-150 light source

Spektral ölçümler Vis/NIR bölgesinde iki farklı ışık kaynağı ile iki farklı yansıma modunda Şekil 2’de gösterildiği şekilde gerçekleştirilmiştir. Hasarsız olarak alınan spektrum ölçümlerinin ardından her bir domates için sertlik özelliği değerleri standart ölçümler (sıkıştırma yöntemi) yardımıyla elde edilmiştir (İğdeci, 2020).

Sertlik Değeri Ölçümü

Kiraz domateslerinin sertlik ölçümleri işaretlenmiş spektral ölçüm bölgelerinden TextureExponent 32 yazılımına sahip 40 mm çaplı akrilik silindir uçlu TA-XTP Plustexture (Stable Micro Systems Ltd., UK) cihazının kullanılması ile, 1 mm s⁻¹ ilerleme hızında ve 4 mm ilerleme mesafesi parametreleri ile iki kez üst üste uygulanan sıkıştırma (5 g’lık sıkıştırma kuvveti ile) yöntemiyle ölçülmüştür (İğdeci, 2020). Bu çalışmada Kuvvet-1 (maksimum kuvvet: ürünün ilk sıkıştırılması sırasında elde edilen en yüksek pik değer) sertlik özelliği ele alınmıştır.

Spektrum Ön İşleme

Çalışma kapsamında etkili dalga boyu aralıkları ve bu aralıklar üzerinde uygulanan spektrum ön işleme teknikleri, önceki çalışma (İğdeci, 2020) kapsamında uygulanan PLS regresyon analizleri sonucunda her bir ışık kaynağı için belirlenmiştir. Buna göre DH-2000-BAL ışık kaynağının kullanıldığı spektral ölçümlerde gürültüyü azaltmak için filtreleme (smoothing) ve MSC (çarpımsal dağılım düzeltmesi-multiplicative scatter correction) spektrum ön işlem tekniklerinin ve etkili dalga boyu olarak ta 420-1037 nm arasının daha etkili olduğu belirlenmiştir (İğdeci, 2020). Benzer şekilde, MI-150 ışık kaynağının kullanıldığı spektral ölçümlerde gürültüyü azaltmak için filtreleme (smoothing) ve 1Dr (birinci türev-first derivative) spektrum ön işlem tekniklerinin ve etkili dalga boyu olarak ta 400-743 nm arasının daha etkili olduğu belirlenmiştir (İğdeci, 2020). Bununla birlikte ortalama merkezleme (mean centering) spektrum ön işlemesi her iki uygulamada da ortak olarak kullanılmıştır. Önceki çalışmada her iki ışık kaynağı ve iki farklı spektral yansıma ölçümü için belirlenen bu dalga boyu aralıkları (üzerlerinde belirtilen spektrum ön işleme işlemlerinin uygulandığı) bu çalışma kapsamında da aynen kullanılmıştır. Söz konusu dalga boyu aralıklarındaki spektral veriler çalışma kapsamında geliştirilen makine öğrenmesi algoritmalarına girdi olarak sunulmuşlardır. Aşağıda bu çalışmanın önemli bir süreci olan makine öğrenmesi algoritmalarının geliştirilmesi ve mimarileri hakkında bilgiler verilmiştir.

Makine Öğrenmesi Algoritmalarının Geliştirilmesi

Kiraz Domates örneklerine ait Vis/NIR spektral yansıma verilerinin ve standart sertlik değerlerinin kullanıldığı bu çalışmada, ikisi doğrusal ve üçü doğrusal olmayan makine öğrenmesine dayalı toplam beş regresyon modeli değerlendirilerek domates sertlik tahminindeki başarıları üzerinden birbirleri ile kıyaslanmışlardır. Çalışma kapsamında ele alınan makine öğrenmesi algoritmalarının geliştirilmesinde, 720 kiraz domates örneğine ait spektral veriden yaklaşık 490 adedi modellerin eğitiminde ve yaklaşık 225 adedi de test aşamalarında kullanılmıştır. Eğitim için ayrılan spektral ölçümlerin yaklaşık %20’si (yaklaşık 98 adet) eğitim içerisinde çapraz validasyon sürecinde kullanılmıştır.

Bu çalışmada kullanılan tüm algoritmalar (PCR, SVM, CNN ve LSTM) MATLAB yazılım ortamında (yazılım sürümü R2022b, The MathWorks, Inc., Natick, Massachusetts, USA) geliştirilmiş ve optimize edilmiştir.

Etkili dalga boyu aralıkları farklı iki halojen ışık kaynağı (DH-2000-BAL ve MI-150) ile iki farklı yansıma modunda alınan Vis/NIR spektral verileri MATLAB yazılımı ortamında geliştirilen makine öğrenmesi algoritmalarında giriş katmanları, DH-2000-BAL ışık kaynağı ile alınan verilerde her bir domates örneği için 3300 spektral okuma sunulacak şekilde oluşturulmuştur. M-150 ışık kaynağı ile alınan verilerde ise giriş katmanlarının yapısı her bir domates örneği için 1706 okuma değeri sunulacak şekilde oluşturulmuştur.

Geliştirilen tüm regresyon modellerinin tahmin başarısı, eğitim (calibration), çapraz validasyon (cross validation) ve validasyon (test-prediction) veri setleri için kararlılık katsayısı (R^2) ve kök ortalama kare hatası (RMSE) (kalibrasyon için RMSEC, çapraz validasyon için RMSECV ve validasyon-test için RMSEP) metriklerinin hesaplanması ile değerlendirilmiştir.

Temel Bileşen Regresyonu

Spektral verilerde yüksek korelasyonu azaltma, boyut indirgeme, gürültüyü filtreleme özelliklerinden faydalanmak ve diğer doğrusal (PLS) ve doğrusal olmayan (non-lineer) algoritmalarla doğrusal analize örnek

oluşturarak kıyaslama imkânı sağlaması açısından PCR yöntemi bu çalışmaya dahil edilmiştir. Temel bileşen regresyonu (PCR-Principle Component Regression), temel bileşen analizi (PCA-Principal Component Analysis) ve en küçük kareler regresyonu (LSR-Least Squares Regression) yöntemlerinin birlikte kullanımı ile gerçekleştirilmektedir (Çataltaş & Tütüncü, 2021). Bu yöntemde önce PCA yöntemi ile spektral verilerden temel bileşenler (PC-principal components) belirlenmektedir. Bu temel bileşenler içerisinde en fazla değişimi gösteren bileşenler seçilmekte ve bu temel bileşenlerin (seçilen) kullanılması ile regresyon modeli oluşturulmaktadır (Çataltaş & Tütüncü, 2021). PCR'nun altında yatan temel ilke, tipik en küçük kareler prosedürünü uygulayan bir doğrusal regresyon modelinde öngörücüler olarak seçilen temel bileşenlerin (PC'lerin) kullanılmasıdır (Tran ve ark., 2018). PCR çok sayıda ve birbiriyle ilişkili spektral veriyi kullanarak bir özellik vektörü (c) tahmin etmeye çalışmakta ve veriye ait temel bileşenleri (PC) bularak boyut indirgeme işlemi gerçekleştirmektedir (Mark & Workman, 2018).

Çalışma kapsamında kullanılan PCR modelinin yapısı iki aşamalı bir şekilde geliştirilmiştir. İlk aşamada, yanıt değişkeni tahmin matrisinin korunan temel bileşenlerine (PC) göre regresyona tabii tutularak çoklu doğrusallığın azaltılması işlemleri gerçekleştirilmiştir. Daha sonra, PCA eğitim verisi tahmin matrisi üzerine uygulanmış ve böylelikle orijinal ve potansiyel olarak ilişkili olan değişkenler, temel bileşenler (PC) olarak adlandırılan yeni ve doğrusal olarak ilişkisiz değişken kümesine dönüştürülmüştür. Daha sonraki ikinci aşamada ise aşağıdaki süreçler sırasıyla yerine getirilmiştir:

İlk temel bileşenler (PC) seçilerek spektral verilerdeki maksimum varyans yönleri korunmuş ve orijinal özellik uzayının dönüştürülmesi gerçekleştirilerek gürültü azaltılmıştır. Bir sonraki aşamada eğitim setindeki veriler azaltılmış boyutlu alana yansıtılmıştır. Daha sonra, yanıt değişkeninin temel bileşenlerin (PC) doğrusal bir kombinasyonu olarak modellendiği sıradan en küçük kareler (OLS) yöntemi kullanılarak çok değişkenli standart bir doğrusal regresyon modeli elde edilmiştir. Son olarak, geliştirilen model test seti verisi üzerinde kullanılarak geliştirilen sistemin performansı değerlendirilmiştir.

Ele alınan çalışma kapsamında doğrusal regresyon modelleri olarak önceki çalışmada ele alınan kısmi en küçük kareler (PLS) regresyon modeli (İğdeci, 2020) ve yukarıda detayları verilen temel bileşen regresyonu (PCR) kullanılmıştır. Önceki çalışmada PLS regresyon modelinin kullanılması ile elde edilen tahmin sonuçları bu çalışmada sadece kıyaslama amacı ile kullanılmıştır.

Çalışmada ele alınan diğer regresyon modelleri doğrusal olmayan ve aşağıda verilen üç makine öğrenmesi algoritmalarından oluşmaktadır.

Destek Vektör Makinesi Regresyonu

SVM makine öğrenmesi uygulamalarında yaygın olarak kullanılan sınıflama ve regresyon modelidir (Chanda ve ark., 2019). Destek vektör makinalarının amacı, yüksek boyutlu bir özellik uzayında sınıflar arasında bir ayırım sağlayan ve bir ayırım oluşturan bir hiper-düzlem (hyper-plane) tanımlamaktır. Tipik olarak veri grubunu doğrusal olarak ayırabilen sonsuz sayıda hiper-düzlem mevcut olduğundan, optimum ya da en iyi uyum sağlayan bir hiper-düzlem, hiper-düzlem ve en yakın örnekler arasındaki mesafe ya da boşluklar maksimum olacak şekilde seçilmektedir. Ortaya çıkabilecek bazı sorunlar bu hiper-düzlemlerin gösterilmesi ya da tanımlanması için optimizasyon probleminin bir parçası olarak düzenleme (regularization) parametresi C (sınıflar arasında alansal çakışmalarda bir iyileşmeye izin veren) tarafından giderilebilmektedir. Burada C, çapraz validasyon hatasına bağlı olarak eldeki veri grubuna göre kararlaştırılmaktadır ve özellikle eğitim veri grubu tam anlamıyla doğrusal olarak ayıramaz durumda iken kullanıldığında fayda sağlamaktadır (Chanda ve ark., 2019). SVMR (destek vektör makinesi regresyonu) durumunda, aşağıdaki maliyet fonksiyonunu (Q) minimize etmek en iyi regresyon hattını tanımlamaktadır:

$$Q = \frac{1}{2} w^T w + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad \text{şu koşullara göre} \begin{cases} y - wx_i - b \leq \varepsilon + \xi_i \\ wx_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i \xi_i^* \geq 0 \end{cases} \quad (1)$$

Modelin tahmin ve genelleme kabiliyetlerini arttırmak için, SVMR'nun parametrelerinin kalibrasyon modelinin geliştirilmesinden önce belirlenmesi gerekmektedir. Düzenleme parametresi gama (gamma, δ) ve

radyal tabanlı fonksiyon (RBF; Radial Basis Function) çekirdek (kernel) parametresi sigma karenin (σ^2) optimize edilmesi gerekmektedir. Gama modelin performansını maksimize etme ve model karmaşıklığını minimize etme ile ilgilidir. δ 'nın büyük değerleri modelin az regüle edilebileceğini ve doğrusal olamama durumunun daha fazla olacağını ima etmektedir. Bunun yanında, sigma (σ^2) bant genişliği (bandwidth) parametresidir ve giriş düzleminden daha yüksek boyutsal özellik uzayına olan doğrusal olmayan eşlemeyi tanımlamaktadır. Sigmanın (σ) yüksek değerleri, daha doğrusal olmayan bir modele yönelen modeldeki daha fazla komşuyu ifade etmektedir (Chanda ve ark., 2019).

Maliyet fonksiyonu (1 nolu eşitlik) yardımı ile regresyon katsayıları üzerine iki koşullu bir ceza uygulanmaktadır. Bu koşullar bir hata terimi ve bir grup sabit ile çarpım halinde olan hata ağırlığını (C) içermektedir. Bu ceza ağırlığı, kullanıcı tarafından optimize edilmesi gereken ε 'nin değerine ek olarak C ile temsil edilmektedir. C 'nin çok yüksek değerlerde olması $\pm\varepsilon$ değerinden sapmalara neden olarak maliyet fonksiyonunun (Q) fazla artmasına neden olmaktadır, diğer taraftan C 'nin aşırı küçük seçilmesi durumunda, en iyi sonuç tamamen regresyon ağırlıkları tarafından belirlenmektedir.

Bu maliyet fonksiyonunun (Q) kullanılmasının amacı hem tahmin hatalarını hem de katsayıların büyüklüğünü eş zamanlı olarak en aza indirmektir. Hata terimindeki tahmin hataları, SVMR'de yaygın olarak izlenen bir prosedür olan $\pm\varepsilon$ sapmasının (Vapnik'in ε -duyarsız kayıp fonksiyonuna göre) ihmal edilmesiyle doğrusaldır. $+\varepsilon$ 'den fazla veya $-\varepsilon$ 'den az sapma gösteren tahminler, sözde gevşek değişkenler (ξ ve ξ^*) ile hesaplanmaktadır.

SVMR, sıradan en küçük kareler (OLS) ve PLS regresyon tekniğinin yaklaşımına benzer bir şekilde, açıklayıcı değişkenler ve bağımlı değişkenler arasında doğrusal bir ilişki bulmaktadır. OLS için amaç, aşağıdaki denkleme göre veriler üzerinden bir doğru uydurmaktır:

$$\hat{y}_i = wx_i + b \quad (2)$$

Burada, belirli bir maliyet fonksiyonunu minimize eden doğru, en iyi doğru olarak kabul edilmektedir. Yukarıdaki formülde, $\hat{y}_i$ belirli bir spektrum (x_i) için bağımlı değişkenin tahmini olurken, w ve b regresyon doğrusunun sırasıyla eğimi ve sapması olmaktadır.

SVM regresyon (SVMR), SVM ile aynı prensibe dayalı olarak çalışmaktadır. SVMR kernel fonksiyonlarına dayandığı için parametrik olmayan bir teknik olarak göz önüne alınmaktadır. SVMR'unda eğitim için olan veri grubu tahmin için olan değişkenleri ve gözlem değerlerini içermektedir. Burada amaç, eğitim için olan her bir x için $\hat{y}_i$ 'den (hedef) ε 'den daha büyük olmayan bir değerle sapma gösteren ve aynı zamanda mümkün olduğunca düz olan bir fonksiyon ($f(x)$) bulmaktır.

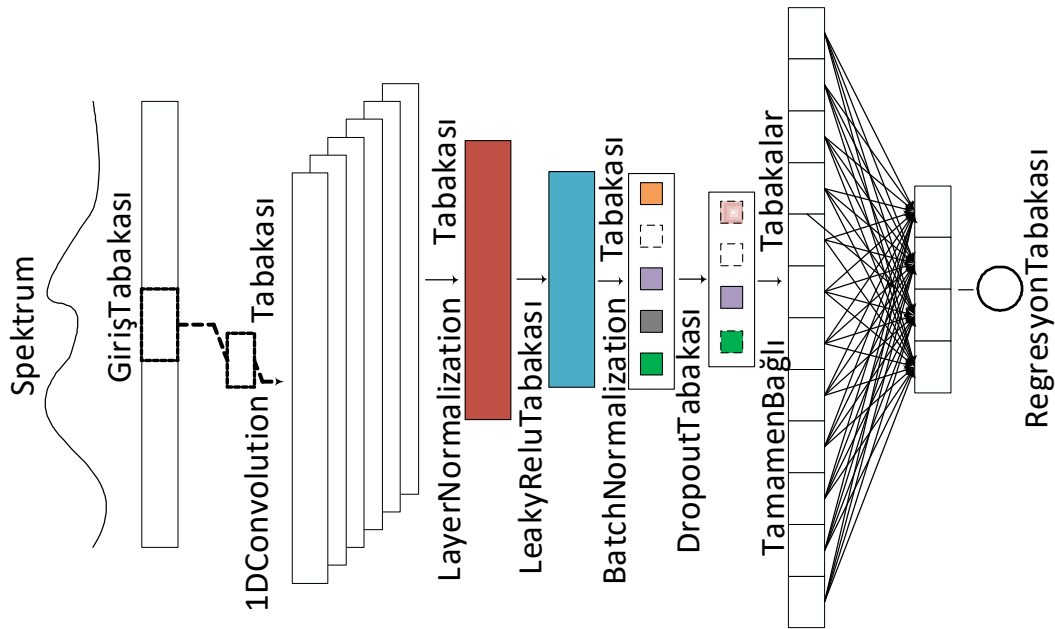
Bu çalışmada, Matlab (yazılım sürümü R2022b, The MathWorks, Inc., Natick, Massachusetts, USA) yazılımı ile spektral veri matrisinden hedef değerleri tahmin etmek için fitrsvm fonksiyonu kullanılarak bir destek vektör makinası regresyonu (SVMR) modeli geliştirilmiştir. Girdi değişkenleri arasında yüksek dereceli etkileşimlerin yakalanabilmesi için polinom çekirdek fonksiyonu (KernelFunction) kullanılmıştır. Spektumlarda bulunan doğrusal olmayan bağımlılıkları çözmek için, 2.2 dereceli genelleştirilmiş kesirli polinom çekirdek fonksiyonu uygulanmıştır. Spektral veriler analiz öncesinde kararlılığı sağlamak için standardize edilmiştir (ortalama sıfır, birim varyans). Model karmaşıklığı ve tahmin toleransı arasındaki dengeyi sağlamak için C düzenleme parametresi (BoxConstraint) kullanılmıştır. Çekirdek matris işlemleri için 98 MB'lık bir bellek tahsis edilerek programlama optimizasyonu sırasında hesaplama verimliliği artırılmıştır.

Bir Boyutlu Evrişimli Sinir Ağları Regresyonu

CNN derin öğrenmede en çok kullanılan yapay sinir ağı modelidir ve genellikle görüntü verilerinin değerlendirilmesinde ve sınıflandırılmasında kullanılmaktadır. Açılar ve mesafeler gibi devam eden bir veri grubunun tahmini için ağın sonuna bir regresyon ağı dahil edilebilmektedir. CNN çıktısı belirlemek için verinin her bir tabakasında bulunan oldukça kompleks özelliklerin çıkartılması için dizayn edilmiş özel bir yapıya sahip çoklu-tabakalı bir yapay sinir ağı modelidir (Anonim-a, 2026). CNN en çok belli bir yapısı olmayan (örneğin görüntü gibi) veri grubu olduğu zaman kullanılmakta ve yapılarıdaki bilginin buradan

çıkartması beklenmektedir. Özellik öğrenme kabiliyetlerinden dolayı CNN herhangi bir özellik işleme yapmadan yüksek boyutlu işlenmemiş veri üzerinde sınıflama ve regresyon amacı ile kullanılabilir.

Çalışma kapsamında, spektral verilerin ardışık yapısını korumak ve dalga boyları arasındaki yerel komşuluk ilişkilerinden (spektral pik ve bant özellikleri) doğrudan özellik çıkarımı yapabilmek amacıyla tek boyutlu evrişimli sinir ağları (1DCNN) tabanlı bir regresyon modeli geliştirilmiştir (Şekil 3). Vektör formundaki spektral veriler geliştirilen bir boyutlu evrişimli sinir ağları (1DCNN) modeline girdi olarak sunularak kiraz domates örneklerinin sertlik tahmininde kullanılmıştır. Şekil 3'te genel yapısı verilen 1DCNN modelinde genel olarak şu tabakalar yer almıştır; evrişim (convolution) tabakası, tabaka normalizasyon (LayerNormalizationLayer) tabakası, sızıntı tabakası (leakyLayer), toplu normalizasyon tabakası (batchnormalizationLayer), dropout tabakası (dropoutLayer), tam bağlantılı tabaka (fully connectedLayer) ve regresyon tabakasıdır. Spektral verilere normalizasyon uygulanmıştır. Giriş tabakası değerleri yukarıda açıklandığı üzere ve kullanılan ışık kaynağına bağlı olarak farklı değerlerde (3300 ve 1706) seçilmiştir. Giriş ve regresyon tabakalarını da içine alacak şekilde 1DCNN içerisinde yaklaşık 8 tabaka yer almıştır. Veri matrisindeki genlik varyasyonlarını dengelemek için girdi aşamasında 0-1 aralığında min-max normalizasyonu (rescale-zero-one) uygulanmıştır. Benzer bir CNN mimarisi Yang ve ark. (2020)'nin yaptığı bir çalışmada da kullanılmıştır.



Şekil 3. Vis/NIR spektral verilerin değerlendirilmesinde kullanılan 1DCNN'nin genel yapısı

Figure 3. The general structure of 1DCNN used in the evaluation of Vis/NIR spectral data

Ağın öznelilik çıkarma bölümünde; 3 kernel boyutuna (filtre genişliği) sahip 128 adet filtreden oluşan tek boyutlu evrişim katmanı (1D Convolutional Layer) kullanılmıştır. Evrişim işlemi sırasında spektral sinyalin kenarlarındaki bilgi kaybını önlemek amacıyla "aynı boyutta dolgu" (same padding) stratejisi benimsenmiştir. Modelin eğitim kararlılığını artırmak ve gradyan problemlerini önlemek amacıyla evrişim katmanının ardından katman normalizasyonu (Layer Normalization) uygulanmıştır. Doğrusal olmayan ilişkilerin yakalanması ve nöron sönümlenmesi (dying ReLU) problemlerinin aşılması için aktivasyon fonksiyonu olarak sızıntı tabakası (Leaky ReLU) tercih edilmiştir.

Modelin veri setini ezberlemesini (overfitting) engellemek ve genelleme kabiliyetini optimize etmek amacıyla doğrusal katman öncesinde %50 oranında bir seyreltme (Dropout, 0.5) katmanı konumlandırılmıştır. Ağın son aşamasında, özetlenen spektral özellikleri hedef kalite parametresine eşlemek üzere sırasıyla 32, 12 ve 1 nöronlu üç adet tam bağlantılı katman (Fully Connected Layer) kullanılmıştır. 1DCNN regresyon

modelindeki son tahmin ise tek nöronlu bir doğrusal regresyon katmanı (Regression Output Layer) vasıtasıyla sürekli sayısal değer (regresyon çıktısı) olarak elde edilmiştir.

Uzun-Kısa Süreli Bellek Regresyonu

Derin öğrenmenin kullanımı farklı bilim alanlarında farklı problemlerin etkili analizlerinin yapılmasını sağlamaktadır. LSTM ağlarının zaman serileri, konuşma tahmini, sinyal analizi, stok tahmini gibi sırasal problemler için en uygun olduğu gözlenmektedir. LSTM ağları geleneksel ve tekrarlı sinir ağlarına (RNN) göre bazı avantajlara sahiptirler (Singh & Kasana, 2019). Uzun kısa-süreli bellek ağları (long short-term memory networks-LSTMN) tekrarlı sinir ağlarının (RNN) bir uzantısı olarak kabul edilmektedir (Hochreiter & Schmidhuber, 1997).

Geleneksel bir RNN, ağın uzun-dönemli bağımlılıkları öğrenmesini zorlaştıran ve zaman içerisinde geçilen saklı bir yapıya sahiptir. Diğer taraftan, LSTM ağları bilgiyi uzun zaman periyodu boyunca tutabilen ve bir konteyner görevi gören bir hafıza hücresi ilavesi ile RNN' nın sahip olduğu problemi çözebilmektedir. Hafıza hücresi üç kapı tarafından kontrol edilmektedir: Giriş kapısı, unutma kapısı ve çıkış kapısı. Bu kapılar, hafıza hücresine hangi bilginin ekleneceğine, hangi bilginin uzaklaştırılacağına ve hücreden hangi bilginin çıkacağına karar vermektedir. Bunun sayesinde, LSTM ağları, bilgi ağ içerisinde akış halinde iken söz konusu bilgiyi tutup tutmayacağına ya da atıp atmayacağına seçerek karar vermektedir ve bu da LTSM ağlarının uzun-dönem bağımlılıkları öğrenmesini sağlamaktadır (Anonim-b, 2026). Bu kapakçık mekanizmaları sayesinde bir LSTM modeli, örneğin spektrum üzerindeki işlevsel gruplara ait ana absorban bantları ile uzak bölgelerdeki üst ton (overtone) kombinasyonları arasındaki ilişkileri bütünsel olarak öğrenebilmektedir. LSTM ağları derin LSTM ağlarını oluşturmak adına yığın halinde birleştirilebilmektedirler ve bu sayede de sırasal veri içerisindeki daha kompleks yapıları öğrenebilmektedirler. Diğer taraftan, LSTM ağları görüntü analizi gibi görevler için CNN gibi diğer sinir ağları ile birlikte de kullanılabilirler.

Bu çalışma kapsamında, spektral veri matrisinin ardışık yapısını korumak ve dalga boyları arasındaki karmaşık ileri-geri yönlü kimyasal korelasyonları bütünsel olarak modellemek amacıyla çift yönlü uzun kısa vadeli bellek (BiLSTM) tabanlı bir derin öğrenme regresyon mimarisi geliştirilmiştir. Çalışmada kullanılan BiLSTM ağ yapısı, MATLAB kütüphanesi kullanılarak kodlanmıştır. Modelin giriş katmanı (Sequence Input Layer), BAL-2000-DN ışık kaynağı için 3300 (MI-150 ışık kaynağı için 1706) bağımsız dalga boyu değişkenine sahip spektral sinyalleri kabul edecek şekilde yapılandırılmıştır. Diğer taraftan, cihaz ve ortam kaynaklı taban çizgisi kaymalarını (baseline shifts) elimine etmek amacıyla girdi aşamasında Z-skoru standardizasyonu (z-score normalization) uygulanmıştır.

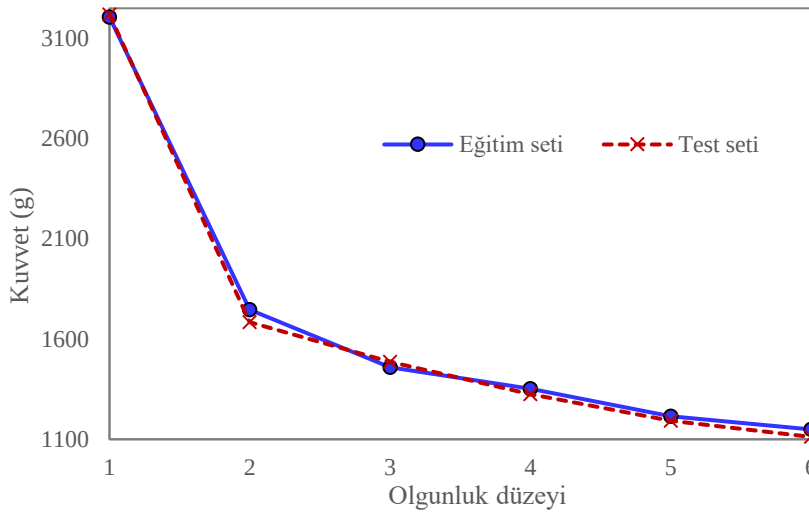
Ağın temelini oluşturan BiLSTM katmanı, 128 gizli birim (hidden units) içerecek şekilde tasarlanmıştır. Standart tek yönlü LSTM yapılarından farklı olarak BiLSTM katmanı, spektral sinyali hem düşük dalga boyundan yüksek dalga boyuna (ileri yönlü) hem de yüksek dalga boyundan düşük dalga boyuna (geri yönlü) olmak üzere çift taraflı olarak analiz etmektedir. Bu çift yönlü öğrenme mekanizması, spektrum üzerindeki spesifik absorban piklerinin hem öncesindeki hem de sonrasındaki spektral komşuluk ilişkilerini eş zamanlı yakalayarak modelin tahmin gücünü artırmaktadır.

Ağın eğitim performansını stabilize etmek ve derin katmanlardaki gradyan patlamalarını önlemek adına BiLSTM öncesinde katman normalizasyonu (Layer Normalization), sonrasında ise grup normalizasyonu (Batch Normalization) ardışık olarak konumlandırılmıştır. Modelin sınırlı veri setlerinde ezberleme (overfitting) eğilimini bastırmak ve genelleme kabiliyetini optimize etmek amacıyla tam bağlantılı katman öncesinde %50 oranında bir seyreltme (Dropout, 0.5) katmanı kullanılmıştır. Ağın çıkış kısmında, çıkarılan öznetelikleri hedef kalite parametresine eşlemek üzere 20 ve 1 nöronlu iki tam bağlantılı katman (Fully Connected Layer) konumlandırılmış ve nihai tahmin çıktısı ise tek nöronlu bir regresyon katmanı (Regression Output Layer) vasıtasıyla sürekli sayısal değer olarak üretilmiştir. Çalışmada kullanılan LSTM regresyon modelinin genel yapısı mimari açıdan Şekil 3'te verilen yapıya benzerlik göstermektedir.

Bulgular ve Tartışma

Çalışmaya dahil edilen 720 kiraz domates örneğine ait spektroskopi okumaları USB4000-Vis/NIR spektrometre ve iki farklı ışık kaynağı kullanarak iki farklı yansıma modunda gerçekleştirilmiştir. Kiraz domates örneklerinin sertlik özelliklerinin değişimi yaklaşık bir aylık hasat periyodu boyunca azalma eğiliminde olmuştur (Şekil 4). Kiraz domates örneklerinden Vis/NIR yansıma spektrumları ve standart ölçümlerle elde edilen sertlik özelliklerinin kullanılması ile geliştirilen doğrusal ve makine öğrenmesine dayalı regresyon modellerinin kullanılması elde edilen sertlik tahmin sonuçları aşağıdaki gibi değerlendirilmiştir.

Ele alınan iki adet doğrusal regresyon modeli (PLS ve PCR) ile üç adet doğrusal olmayan makine öğrenmesine dayalı regresyon modelinin (SVR, 1D-CNN ve LSTM) tahmin performansları kendi aralarında karşılaştırmalı olarak değerlendirilmiştir. PLS regresyon analizi ile aynı veri grubunu kullanarak elde edilen sertlik tahmin sonuçları daha önceki çalışmada (İğdeci, 2020) elde edilen sonuçlar olup kıyaslama amacı ile burada tekrar kullanılmışlardır. Çalışma kapsamında geliştirilen doğrusal ve doğrusal olmayan makine öğrenmesi regresyon modelleri ile kiraz domates örneklerinin sertlik değerleri için elde edilen tahmin değerleri aşağıda iki kısım altında değerlendirilmiştir.



Şekil 4. Eğitim ve test veri grubuna ait domates örneklerinin olgunluk düzeyine göre sertlik değerlerindeki değişim (İğdeci, 2020).

Figure 4. Changes in firmness values of tomato samples from the training and test data groups according to their maturity level (İğdeci, 2020).

Domates Sertlik Özelliğinin USB4000-Vis/NIR Spektrometre ve DH-2000-BAL Işık Kaynağı ile Alınan Spektrumların Kullanılması ile Tahmin Edilmesi

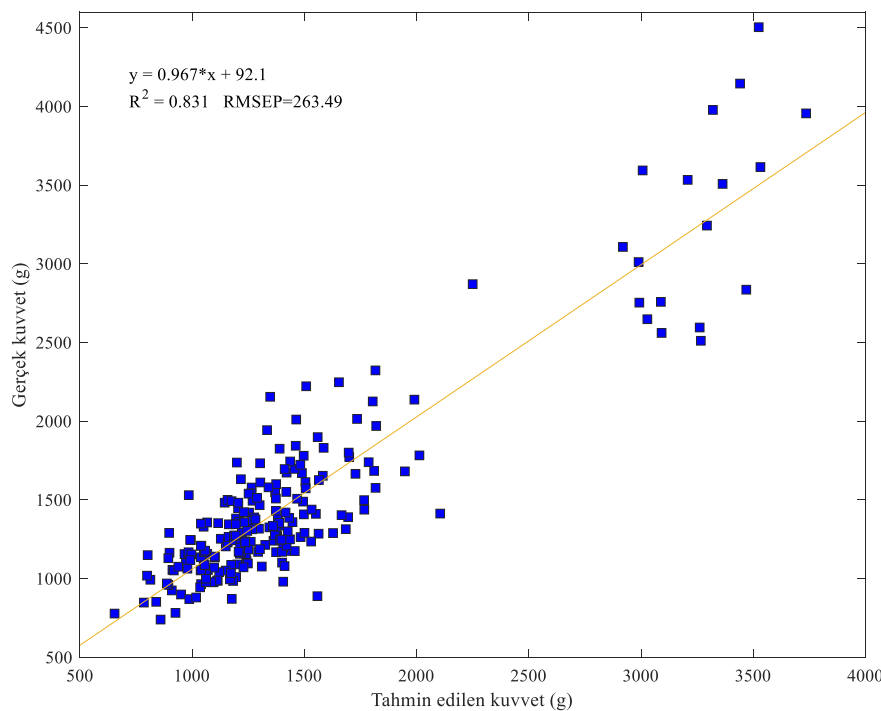
Kiraz domates sertlik özelliği USB4000-Vis/NIR spektrometre ve DH-2000-BAL ışık kaynağı ile yansıma modunda alınan spektrumların ve beş farklı regresyon modelinin (2 doğrusal ve 3 doğrusal olmayan-makine öğrenmesine dayalı) kullanılması ile başarılı olarak tahmin edilmiştir. Elde edilen tahmin sonuçları Çizelge 1’de verilmiştir. Çizelge 1 incelendiğinde, en başarılı sertlik tahmininin makine öğrenmesine dayalı olarak geliştirilen 1DCNN algoritmasının kullanılması ile elde edildiği görülmektedir. Bu algoritmanın kullanılması ile kiraz domates sertlik özelliği test (validasyon) veri grubu için 0.83 (R^2) kararlılık katsayısı (RMSEP=263.49) ile tahmin edilirken, bu uygulamayı PCR, LSTM ve SMV algoritmaları takip etmiştir. Bu üç algoritma ile elde edilen kararlılık katsayısı 0.82 (R^2) olurken, uygulamalara ait tahmindeki hata oranları (RMSEP) sırasıyla 266.68, 267.61 ve 269.08 olmuştur. Çalışma kapsamında geliştirilen ve ele alınan dört farklı regresyon modeli (PCR, SVR, 1DCNN ve LSTM) ile elde edilen sertlik tahmin değerleri önceki çalışmada PLS regresyon modeli ile elde edilen sertlik tahmin değerlerine ($R^2=0.80$, RMSEP=285.79) göre daha iyi olmuştur. Diğer taraftan çapraz validasyon sürecinde en iyi tahmin sonucu PCR algoritması ($R^2=0.81$, RMSEP=252.42) ile elde edilirken, kalibrasyon sürecinde en iyi tahmin değerleri LSTM regresyon modeli ($R^2=0.91$, RMSEP=197.02) ile elde edilmiştir.

Çizelge 1. DH-2000-BAL ışık kaynağı ile yansıma modundaki domates spektrumlarını kullanarak elde edilen sertlik tahminleri**Table 1.** Firmness estimations obtained using tomato spectra in reflection mode using DH-2000-BAL light source

Model	Kalibrasyon Seti		Çapraz Validasyon Seti		Validasyon Seti	
	R ²	RMSEC	R ²	RMSECV	R ²	RMSEP
PLS*	0.78	302.75	0.74	331.28	0.80	285.79
1D-CNN	0.87	231.55	0.79	325.77	0.83	263.49
LSTM	0.91	197.02	0.78	310.05	0.82	267.61
SVM	0.81	284.03	0.71	359.27	0.82	269.08
PCR	0.77	313.87	0.81	252.42	0.82	266.68

*Önceki çalışmada (İğdeci, 2020) elde edilen domates sertlik tahmin değerleri

DH-2000-BAL ışık kaynağı ile alınan kiraz domates örneklerine ait Vis/NIR spektral verilerin kullanılması ile test veri grubu (validasyon) için en iyi sonucun alındığı sertlik tahmin modeli (1DCNN) Şekil 5'te verilmiştir.

**Şekil 5.** DH-2000-BAL ışık kaynağı ile yansıma modunda alınan spektral verilerin kullanılması ile domatesler için elde edilen sertlik tahmin modeli (1DCNN regresyon)**Figure 5.** Firmness prediction model for tomatoes obtained using spectral data acquired in reflection mode with DH-2000-BAL light source (1DCNN regression)

Kiraz Domates Sertlik Özelliğinin USB4000-Vis/NIR Spektrometre ve MI-150 Işık Kaynağı ile Alınan Spektrumların Kullanılması ile Tahmin Edilmesi

Kiraz domates sertlik özelliği USB4000-Vis/NIR spektrometre ve MI-150 ışık kaynağı ile yansıma modunda alınan spektrumların ve beş farklı regresyon modelinin kullanılması ile başarılı olarak tahmin edilmiştir. Elde edilen tahmin sonuçları Çizelge 2'de verilmiştir. Çizelge 2 incelendiğinde, en başarılı sertlik tahmininin PCR algoritmasının kullanılması ile elde edildiği görülmektedir. Bu algoritmanın kullanılması ile kiraz domates sertlik özelliği test (validasyon) veri grubu için 0.82 (R²) kararlılık katsayısı (RMSEP=277.16) ile tahmin edilirken, bu uygulamayı sırasıyla 1D-CNN, SVM ve LSTM algoritmaları takip etmiştir. Bu üç algoritmadan 1D-CNN algoritması için elde edilen kararlılık katsayısı 0.82 (R²) olurken, tahmindeki hata oranı (RMSEP) 290.19 olmuştur. SVM ve LSTM algoritmaları için elde edilen kararlılık katsayıları (R²) sırasıyla 0.81 ve 0.80 olurken, tahmindeki hata oranları (RMSEP) 289.51 ve 300.48 olmuştur. Çalışma kapsamında

geliştirilen ve ele alınan dört farklı regresyon modelinin (PCR, SVR, 1DCNN ve LSTM) kullanılması ile sertlik tahmininde elde edilen kararlılık katsayı değerleri, önceki çalışmada PLS regresyon modeli ile elde edilen kararlılık katsayısı değerine ($R^2=0.79$) göre daha iyi olmuştur.

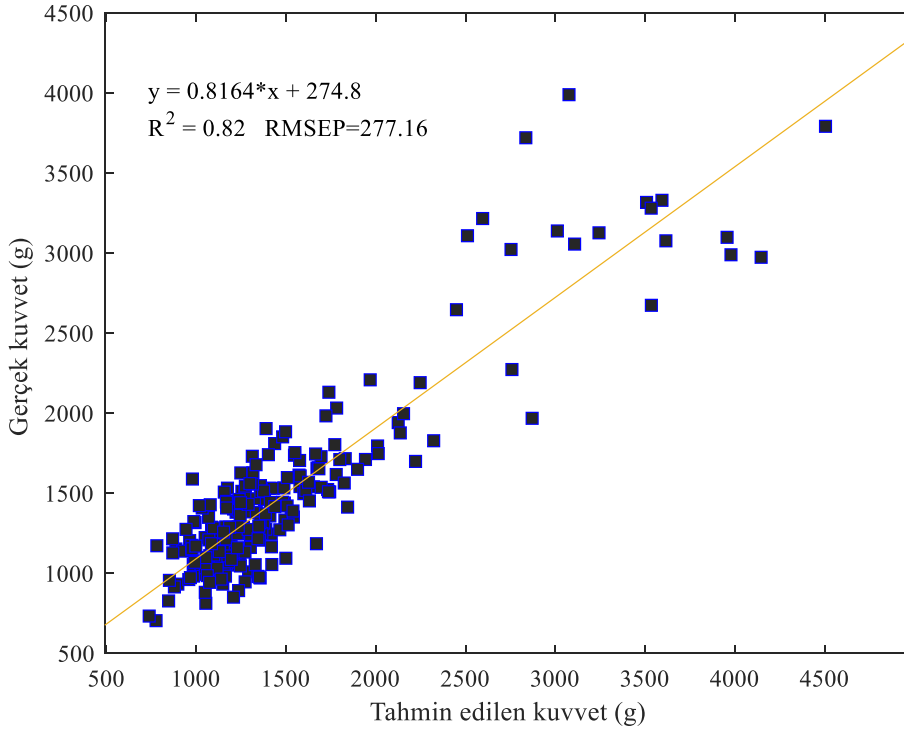
Bununla birlikte, çapraz validasyon sürecinde en iyi tahmin sonucu diğer ışık kaynağı (DH-2000-BAL) ile olan yansımaya uygulamasında olduğu gibi yine PCR algoritması ($R^2=0.80$, RMSEP=310.28) ile elde edilirken, kalibrasyon sürecinde en iyi tahmin değerleri LSTM regresyon modeli ($R^2=0.92$, RMSEP=172.35) ile elde edilmiştir (Çizelge 2).

Çizelge 2. MI-150 ışık kaynağı ile yansımaya modunda alınan domates spektrumlarını kullanarak elde edilen sertlik tahminleri
Table 2. Firmness estimations obtained using tomato spectra in reflection mode using MI-150 light source

Model	Kalibrasyon Seti		Çapraz Validasyon Seti		Validasyon Seti	
	R^2	RMSEC	R^2	RMSECV	R^2	RMSEP
PLS*	0.74	327.92	0.70	353.02	0.79	294.74
1D-CNN	0.89	221.74	0.75	369.10	0.82	290.19
LSTM	0.92	172.35	0.80	325.47	0.80	300.48
SVM	0.82	282.67	0.66	335.69	0.81	289.51
PCR	0.75	317.30	0.80	310.28	0.82	277.16

*Önceki çalışmada (İğdeci, 2020) elde edilen domates sertlik tahmin değerleri

MI-150 ışık kaynağı ile alınan domates örneklerine ait Vis/NIR spektral verilerin kullanılması ile test veri grubu (validasyon) için en iyi sonucun alındığı sertlik tahmin modeli (PCR regresyon) Şekil 6'da verilmiştir.



Şekil 6. MI-150 ışık kaynağı ile yansımaya modunda alınan spektral verilerin kullanılması ile domatesler için elde edilen sertlik tahmin modeli (PCR regresyon)

Figure 6. Firmness prediction model for tomatoes obtained using spectral data acquired in reflection mode with MI-150 light source (PCR regression)

Her iki ışık kaynağı ile alınan Vis/NIR spektral verilerin ve farklı regresyon modellerinin kullanılması ile kiraz domatestte sertliğin iyi bir şekilde tahmin edildiği görülmekle birlikte, DH-2000-BAL ışık kaynağı ile

elde edilen yansıma spektrumlarının kullanılması ile elde edilen tahmin sonuçlarının MI-150 ışık kaynağı ile alınan spektrumların kullanılmasına göre biraz daha iyi olmuştur (Çizelgeler 1 ve 2). DH-2000-BAL ışık kaynağı ile biraz daha iyi sonuçların alınmasında daha geniş bir dalga boyu aralığının ve dolayısıyla daha fazla verinin ele alınmış olmasının etkili olduğu değerlendirilmektedir.

Spektrumlarda gürültüyü azaltmak için uygulanan ön işleme tekniklerinden her iki ışık kaynağı ile alınan spektrumlara uygulanan ortalama merkezleme, DH-2000-BAL ışık kaynağının kullanıldığı spektral ölçümlere uygulanan filtreleme ve MSC ve MI-150 ışık kaynağının kullanıldığı spektral ölçümlere uygulanan filtreleme ve 1Dr uygulamalarının setlik tahmininde daha başarılı sonuçlar alınmasında etkili oldukları belirlenmiştir (İğdeci, 2020).

Bunun yanında, araştırma bulgularından çıkarılabilecek bir diğer sonuç ta doğrusal bir model olarak kabul edilen PCR regresyon modeli ile özellikle MI-150 ışık kaynağı ile alınan spektrumların kullanılmasıyla derin öğrenmeye dayalı modeller (1DCNN) kadar ve hatta daha yüksek tahmin başarılarının elde edildiğidir. PCR modeli ile BAL-2000-DH ışık kaynağı ile alınan spektrumların kullanılmasıyla da başarılı sonuçlar alınmıştır. Çalışma kapsamında ele alınan spektrumların çok karmaşık bir yapı sergilememesinden dolayı doğrusal modeller ile de başarılı tahmin sonuçlarının alındığı değerlendirilmektedir.

Yürütülen bu çalışma kapsamında kiraz domatesler için elde edilen sertlik tahmin değerlerinin literatürde yine domatesler için benzer yaklaşımlar ile yürütülen çalışmalarda rapor edilen sertlik tahmin sonuçlarına yakın olduğu değerlendirilmiştir.

Feng ve ark. (2019) tarafından yürütülen bir çalışmada kiraz domateslerde standart sertlik değerleri, NIR spektral verileri (950-1650 nm) ve kemometrik tekniklerin (PLS, SVM ve ELM-Extreme Learning Machine) kullanılması ile sertlik tahmin modelleri geliştirilmiştir. Sonuçta kiraz domateslerde sertlik değeri, herhangi bir ön işleme yapılmamış spektrumların kullanıldığı PLS, SVM ve ELM modelleri ile iyi sonuçlar elde edilerek tahmin edilmiştir (sırasıyla, $R^2=0.92$ ve $RMSEP=0.53$, $R^2=0.96$ ve $RMSEP=0.40$ ve $R^2=0.97$ ve $RMSEP=0.31$).

Sirisomboon ve ark. (2012) yürüttükleri bir çalışmada domateslerde standart sertlik ölçümü için sıkıştırma testi uygulamışlar ve daha sonra domateslerdeki sertlik özelliklerini tahmin etmek için, elde ettikleri sertlik değerlerini ve NIR spektral okumalarını (1100-1500 nm) kullanarak PLS regresyon modelleri geliştirmişlerdir. Sonuçta, domateslere ait test grubu verileri için sertliği 0.85 (r) korelasyon katsayısı ve 1.05 (SEP) tahmindeki hata oranı ile tahmin etmişlerdir.

Van Dijk ve ark. (2006) depolanan domateslerde sertlikteki azalma miktarını tahmin etmek için yürüttükleri çalışmalarında öncelikle domates standart sertlik değerlerini hasarsız sıkıştırma testleri ile ölçmüşlerdir. Daha sonra bu ölçümlerden elde ettikleri eğim (Nm^{-1}) ve mesafe (m) ölçümlerini (sertlik için) ve ayrıca 1100-2500 nm arasında elde ettikleri NIR spektral okumalarını kullanarak sertlik tahmini için PLS regresyon modelleri geliştirmişlerdir. Sonuçta, test grubu domatesleri için elde ettikleri sertlik tahminlerinde eğim (sertlik) için elde ettikleri kararlılık katsayısı (R^2) 0.83 olurken ($RMSEP= 0.51$), mesafe (sertlik) için elde ettikleri kararlılık katsayısı (R^2) 0.84 ($RMSEP= 0.11$) olmuştur.

Huang ve ark. (2018) domates sertliğini tahmin etmek için farklı dalda boyu aralıklarını içeren etkileşim (interaction) modunda alınmış spektral verileri (Vis/SWNIR ve NIR), farklı sertlik ölçüm tekniklerini (darbe, sıkıştırma, akustik ve hasarlı), farklı spektrum işleme tekniklerini ve PLS regresyon tahmin modellerini kullanmışlardır. Sonuçta sıkıştırma yöntemi ile elde edilen standart sertlik değerlerini ve Vis/SWNIR spektral verileri kullanarak sertlik tahmini için elde ettikleri korelasyon katsayısı (r_p) test grubu verileri için 0.92 ($RMSEP=0.90$) olmuştur.

Benzer bir çalışmada Ecartot ve ark. (2013) Vis/NIR spektroskopisi ve PLS regresyonu kullanarak test grubu verileri için domateste sertlik tahmininde $R^2=0.76$ ve $RPD=2.14$ korelasyon sonuçlarını elde etmişlerdir.

Sonuç

İki farklı ışık kaynağı ile elde edilen Vis/NIR spektral yansıma ölçümlerinin, standart sertlik ölçüm değerlerinin ve geliştirilen doğrusal ve doğrusal olmayan (makine öğrenmesine dayalı) regresyon modellerinin kullanılması ile kiraz domatesler için elde edilen hasarsız sertlik tahmin değerleri aşağıda verilmiştir.

DH-2000-BAL ışık kaynağı ile elde edilen Vis/NIR spektral verilerin kullanılması ile test (validasyon) veri grubu için elde edilen en başarılı sertlik tahmini makine öğrenmesine dayalı 1DCNN regresyon modeli ile elde edilmiştir ($R^2=0.83$, RMSEP=263.49). Diğer taraftan, MI-150 ışık kaynağı ile elde edilen Vis/NIR spektral verilerin kullanılması ile test (validasyon) veri grubu için elde edilen en başarılı sertlik tahmini PCR regresyon modeli ile elde edilmiştir ($R^2=0.82$, RMSEP=277.16).

Genel olarak iki farklı ışık kaynağı ile alınan Vis/NIR spektral verilerin kullanılması ile elde edilen kiraz domates sertlik tahmin sonuçları birbirine yakın olmasına rağmen, DH-2000-BAL ışık kaynağı ile elde edilen yansıma spektrumlarının kullanılması ile elde edilen tahmin sonuçlarının biraz daha iyi olduğu görülmüştür. Bunun sebebi olarak söz konusu ışık kaynağı (DH-2000-BAL) ile daha geniş bir dalga boyu aralığının kaplanmış olması ve dolayısıyla modellerde daha fazla veri kullanılmış olması gösterilebilir.

Spektrumlarda gürültüyü azaltmak için uygulanan ön işleme tekniklerinden ortalama merkezleme tekniği her iki ışık kaynağı ile elde edilen spektrumlara uygulanırken, filtreleme-MSD DH-2000-BAL ışık kaynağının kullanıldığı spektral ölçümlerde ve filtreleme-1Dr MI-150 ışık kaynağının kullanıldığı spektral ölçümlerde kullanılmış ve etkili oldukları belirlenmiştir.

Bu çalışmada sertlik iki farklı ışık kaynağı ile yansıma modunda alınan Vis/NIR spektral verilerinin kullanılması ile başarılı bir şekilde tahmin edilirken, DH-2000-BAL ışık kaynağı ile alınan sonuçlar biraz daha iyi olmuştur. Bunun yanında, geliştirilen doğrusal ve doğrusal olmayan (makine öğrenmesine dayalı) regresyon modellerinden 1DCNN, PCR, LTSM ve SVM modellerinin daha önce geliştirilmiş olan PLS regresyon modeline göre daha iyi sertlik tahmin sonuçları verdiği belirlenmiştir.

Diğer taraftan bu çalışmada PCR regresyon modeli ile tecrübe edildiği gibi, spektral verilerin yapısının bazı durumlarda doğrusal olmayan regresyon modelleri ile de iyi uyum gösterebileceği ve bu regresyon modelleri ile de iyi tahmin sonuçlarının alınabileceği görülmüştür. Spektral veriler için uygun regresyon modeli geliştirme ve araştırma çalışmalarına doğrusal-doğrusal olmayan modeller çizgisinde gelecek çalışmalarla devam edilmesinin faydalı olacağı değerlendirilmiştir.

Ek Beyanlar ve Bilgiler

Teşekkür: Bu araştırma makalesi, Aysun İğdeci' nin "Makinalı öğrenme ve VIS/NIR spektroskopisi kullanarak domateslerin içsel kalite kriterlerinin tahmin edilmesi" başlıklı doktora tezinin bir parçasıdır. Çalışmada kullanılan veriler, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) 1180679 numaralı hızlı destek projesi ile desteklenmiş önceki bir çalışmada elde edilmiştir. Bunun için TÜBİTAK'a teşekkürlerimizi sunarız.

Birinci yazar, doktora öğrenimi süresince 100/2000 YÖK Doktora Bursu'ndan (2020–2024, Tarım ve Hayvancılıkta Dijital Teknolojiler alanında) ve TÜBİTAK-BİDEB 2211-C Yurt İçi Öncelikli Alanlar Doktora Burs Programı'ndan (2022–2025, Makine Öğrenmesi ve Yapay Zekâ öncelikli alanında) yararlanmıştır. Sağladıkları desteklerden dolayı her iki kuruma da teşekkür ederiz.

Araştırmacıların Katkı Oranı: Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan eder.

Çıkar Çatışması Beyanı: Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Kaynakça

- Anonim-a. (2026). Deep Learning Tutorial for Beginners: Neural Network Basics. <https://www.guru99.com/deep-learning-tutorial.html>. Son erişim tarihi: 23 Mayıs 2026.
- Anonim-b. (2026). Introduction to Long Short Term Memory. Geeks for Geeks. <https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/>. Son erişim tarihi: 23 Mayıs 2026.
- Bissadu, K. D., Sonko, S. ve Hossain, G. (2025). Society 5.0 enabled agriculture: Drivers, enabling technologies, architectures, opportunities, and challenges. *Information Processing in Agriculture*, 2025: 112-124.
- Chakravartula, S.S.N., Moschetti, R., Bedini, G., Nardella, M., & Massantini, R. (2022). Use of convolutional neural network (CNN) combined with FT-NIR spectroscopy to predict food adulteration: A case study on coffee. *Food Control*, 2022: 108816.

- Chanda, S., Hazarika, A.K., Choudhury, N., Islam, S.A., Manna, R., Sabhapondit, S., Tudu, B., & Bandyopadhyay, R. (2019). Support vector machine regression on selected wavelength regions for quantitative analysis of caffeine in tea leaves by near infrared spectroscopy. *Journal of Chemometrics*, (2019) 1-15.
- Çataltaş, O., & Tütüncü, K. (2021). A review of data analysis techniques used in near-infrared spectroscopy. *Avrupa Bilim ve Teknoloji Dergisi*, 25: 474-484.
- Ecarnot, M., Bączyk, P., Tessarotto, L., & Chervin, C. (2013). Rapid phenotyping of the tomato fruit model, Micro-Tom, with a portable VIS–NIR spectrometer. *Plant Physiology and Biochemistry*, 2013: 159-163.
- Feng, L., Zang, M., Adhikari, B., & Guo, Z. (2019). Nondestructive detection of postharvest quality of cherry tomatoes using a portable NIR spectrometer and chemometric algorithms. *Food Analytical Methods*, (2019) 12: 914-925.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9(8), 1735-1780.
- Huang, Y., Lu, R., & Chen, K. (2018) Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. *Journal of Food Engineering*, 2018: 185-198.
- İğdeci, A. (2020). Kiraz domateste bazı kalite kriterlerinin Vis/NIR spektroskopi ile yansıma ve etkileşim yöntemlerini kullanarak tahmin edilmesi (Yüksek Lisans Tezi). Çanakkale Onsekiz Mart Üniversitesi, Lisansüstü Eğitim Enstitüsü.
- Jamshidi, B., Mohajerani, E., & Jamshidi, J. (2016). Developing a Vis/NIR spectroscopic system for fast and non-destructive pesticide residue monitoring in agricultural product. *Measurement*, 2016: 1-6.
- Kavdır, İ., Büyükcın, M. B., Lu, R. ve Şeker, M., Kocabıyık, H. (2009). Prediction of olive quality using FT-NIR spectroscopy in reflectance and transmittance modes. *Biosystems Engineering*, 2009: 304-312.
- Kazemi, F., Shirzadifar, A., Afsharifar, A., Nasiri, S: M., & Mostowfizadeh-Ghalefarsa, R. (2026) Spectral Zucchini Index (SZI): A novel Vis/NIR-based machine learning approach for early and asymptomatic detection of zucchini yellow mosaic virus. *Smart Agricultural Technology* 14 (2026): 102066.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, P., & Bochtis, D. (2018). Machine Learning in Agriculture: A Review. *Sensors*, 2018: 2674.
- Mark, H., & Workman, J. J. (2018). Calculating the Solution for Regression Techniques: Part 2—Principal Component(s) Regression Made Simple. *Chemometrics in Spectroscopy* içinde, 121-123. Academic Press, 2018.
- Mou, L., Ghamisi, P. & Zhu, X. X. (2017). Deep recurrent neural networks for hyperspectral image classification *IEEE Transactions on Geoscience and Remote Sensing*, 2017: 3639 - 3655.
- Singh, S., & Kasana, S.S. (2019). Estimation of soil properties from the EU spectral library using long short-term memory networks. *Geoderma Regional*, 18 (2019) e00233.
- Sirisomboon, P., Tanaka, M., Kojima, T., & Williams, P. (2012). Nondestructive estimation of maturity and textural properties on tomato ‘Momotaro’ by near infrared spectroscopy. *Journal of Food Engineering*, 112 (2012): 218-228.
- Tran, H., Kim, J., Kim, D., Choi, M., & Choi M. (2018). Impact of air pollution on cause-specific mortality in Korea: Results from Bayesian Model Averaging and Principle Component Regression approaches. *Science of The Total Environment*, (2018):1020-1031.
- Van Dijk, C., Boeriu, C., Peter, F., Stolle-Smits, T., & Tijskens, L.M.M. (2006). The firmness of stored tomatoes (cv. *Tradiro*). I. Kinetics and near infrared models to describe firmness and moisture loss. *Journal of Food Engineering*, 77 (2006): 575-584.
- Yalçın, H., Keçici, P.D., & Ekiz, B. (2026). Prediction of meat sensory traits in goat kids using rearing factors, carcass, and instrumental meat quality characteristics with machine learning and regression methods. *Livestock Science*, 308: 105955.
- Yang, J., Wang, X., Ruihua, W., & Wang, H. (2020). Combination of convolutional neural networks and recurrent neural networks for predicting soil properties using Vis–NIR spectroscopy. *Geoderma*, 2020: 114616.