ÇANKIRI KARATEKİN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

BETON KARIŞIM VERİLERİNİN YAPAY ZEKÂ TEKNİKLERİ KULLANILARAK TAHMİN EDİLMESİ

Bahadır ERİN

İNŞAAT MÜHENDİSLİĞİ ANABİLİM DALI

ÇANKIRI
2022

Her hakkı saklıdır

TEZ ONAYI

Bahadır ERİN tarafından hazırlanan **"Beton Karışım Verilerinin Yapay Zekâ Teknikleri Kullanılarak Tahmin Edilmesi"** adlı tez çalışması 10/08/2022 tarihinde aşağıdaki jüri tarafından oy birliği/oy çokluğu ile Çankırı Karatekin Üniversitesi Fen Bilimleri Enstitüsü İnşaat Mühendisliği Anabilim Dalında **Yüksek Lisans Tezi** olarak kabul edilmiştir.

Danışman **:** Doç. Dr. Cihan DOĞRUÖZ

Jüri Üyeleri **:**

Başkan **:** Doç. Dr. Cihan DOĞRUÖZ
İnşaat Mühendisliği Anabilim Dalı
Çankırı Karatekin Üniversitesi

Üye **:** Prof. Dr. Ender SARIFAKIOĞLU
İnşaat Mühendisliği Anabilim Dalı
Çankırı Karatekin Üniversitesi

Üye **:** Doç. Dr. Murat YAYLACI
İnşaat Mühendisliği Anabilim Dalı
Recep Tayyip Erdoğan Üniversitesi

Yukarıdaki sonucu onaylarım

Prof. Dr. İbrahim ÇİFTÇİ

Enstitü Müdürü

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Çankırı Karatekin Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğine göre hazırlamış olduğum "**Beton Karışım Verilerinin Yapay Zekâ Teknikleri Kullanılarak Tahmin Edilmesi**" konulu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, tezin Çankırı Karatekin Üniversitesi Fen Bilimleri Enstitüsü'nden başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve bu çalışmanın Çankırı Karatekin Üniversitesi tarafından kullanılan "Bilimsel İntihal Tespit Programı"yla tarandığını, "intihal içermediğini" beyan ederim. Çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması halinde ortaya çıkacak tüm ahlaki ve hukuki sonuçlara razı olduğumu bildiririm. Çankırı Karatekin Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim (10/08/2022).

Bahadır ERİN

ÖZET

Yüksek Lisans Tezi

BETON KARIŞIM VERİLERİNİN YAPAY ZEKÂ TEKNİKLERİ KULLANILARAK TAHMİN EDİLMESİ

Bahadır ERİN

Çankırı Karatekin Üniversitesi
Fen Bilimleri Enstitüsü
İnşaat Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Cihan DOĞRUÖZ

Yapay zekâ bilimi ve bir alt dalı olan makine öğrenmesi konularında günden güne inanılmaz gelişmeler yaşanmaktadır. Asıl amacı problemi insan zekasından ilham alarak çözmek olan bu disiplin, gayesini çok sayıda bilim dalında başarıyla gerçekleştirmektedir. Yapılan çalışmada, inşaat mühendisliği alanında yorucu yöntemlerle belirlenebilen ve zaman alan beton basınç dayanımı tespiti için makine öğrenmesi yöntemleriyle sayısız dijital tahmin deneyleri yapılmıştır. Yapılan deneyler sırasında özellikle veri ön işleme aşaması için özgün yöntemler geliştirilmiş ve geliştirilen yöntemlerin tahminlere katkısı test edilmiştir. Birçok yaygın model ve yaygın modellerin birleştirildiği topluluk modelleriyle, geliştirilen ön işleme yöntemlerinin verimliliği kanıtlanmıştır. Ayrıca bu sayede tahmin algoritmalarının veri setine uyumluluğu gözlemlenmiş ve performans sıralaması yapılmıştır.

2022, 59 sayfa

ANAHTAR KELİMELER: Beton karışımı, Yapay zekâ, Makine öğrenmesi, Orange yazılımı, Veri ön işleme

ABSTRACT

Master of Science Thesis

PREDICTION OF CONCRETE MIXTURE DATA BY APPLYING ARTIFICIAL INTELLIGENCE TECHNIQUES

Bahadır ERİN

Çankırı Karatekin University
Graduate School of Natural and Applied Sciences
Department of Civil Engineering

Advisor: Assoc. Prof. Dr. Cihan DOĞRUÖZ

Artificial intelligence and machine learning science are progressing impressively. This discipline, whose main purpose is to solve problems is inspired by human intelligence and achieves its aim in many branches of science. Concrete compressive strength can be determined by difficult methods in civil engineering. In this study, a lot of digital prediction experiments were conducted using machine learning methods for concrete compressive strength prediction. Novel methods were developed for data preprocessing. And the contribution of the developed methods to the predictions were tested. The efficiency of preprocessing methods were proved with many prediction models. In addition, the compatibility of the estimation algorithms with the dataset was observed. According to these observations, the performance ranking of the models was made.

2022, 59 pages

Keywords: Concrete mixture, Artificial intelligence, Machine learning, Orange software, Data preprocessing

ÖNSÖZ VE TEŞEKKÜR

“Beton Karışım Verilerinin Yapay Zekâ Teknikleri Kullanarak Tahmin Edilmesi” adlı bu çalışma, Çankırı Karatekin Üniversitesi, Fen Bilimleri Enstitüsü, İnşaat Mühendisliği Anabilim Dalı’nda Yüksek Lisans Tezi olarak hazırlanmıştır. Çalışmada genel geçer yapı malzemesi olan betonda, basınç dayanımı parametresi tahmininde yapay zekâ yöntemlerinin kullanılabilirliği irdelenmiş ve çalışma kapsamında geliştirilen tahminin verimini artıracak özgün yöntemlerle bilim dünyasına katkı sağlamaya çalışılmıştır.

Çalışmanın başından sonunda kadar gösterdiği özverili, samimi ve yardım sever davranışlarıyla çalışma sürecini ilerletmemde büyük çaba sarf eden danışmanım Sayın Doç. Dr. Cihan DOĞRUÖZ' e teşekkürlerimi sunarım.

Desteğini ve sabrını eksik etmediği için sevgili eşim Gökçen Tuba Erin' e, varlığıyla beni mutlu eden oğlum Mert ERİN' e sonsuz teşekkürler.

Bahadır ERİN
Çankırı, Ağustos 2022

İÇİNDEKİLER

ÖZET i
ABSTRACT ii
ÖNSÖZ VE TEŞEKKÜR iii
İÇİNDEKİLER iv
SİMGELER DİZİNİ vi
KISALTMALAR DİZİNİ vii
ŞEKİLLER DİZİNİ ix
ÇİZELGELER DİZİNİ x
1. GİRİŞ 1
2. LİTERATÜR ÖZETİ 4
3. MATERYAL VE YÖNTEM 13
3.1 Çalışmada Kullanılan Veri Setleri ve İlgili Bazı Tanımlar 13
3.1.1 Veri ön işleme 14
3.1.2 Makine öğrenmesinde gürültü tanımı 14
3.1.3 Özellik seçimi yöntemi 15
3.1.4 Veri setinde bulunan bazı parametrelerin tanımları 15
3.2 Kullanılan Yazılımlar 16
3.3 Kullanılan Makine Öğrenmesi Yöntemleri ve Tayin Edilen Model Parametreleri 16
3.3.1 K en yakın komşu 16
3.3.2 Karar ağaçları 17
3.3.3 Destek vektör makineleri 18
3.3.4 Olasılıksal dereceli azalma 19
3.3.5 Rassal orman 20
3.3.6 Yapay sinir ağları 21
3.3.7 Doğrusal regresyon 22
3.3.8 Gradyan artırma 23
3.3.9 Adaptif yükseltme 24
3.4 Diğer Tanımlar 25
3.4.1 İstifleme yöntemi 25

3.4.2 K katmanlı çapraz doğrulama 25
3.4.3 Performans ölçüm yöntemleri 26
3.5 Yöntem 27
3.5.1 Aşama 1 27
3.5.2 Aşama 2 29
3.5.3 Aşama 3 29
3.5.4 Aşama 4 29
4. BULGULAR 32
4.1 Veri Setine Özellik Seçimi Yöntemi Uygulanması 32
4.2 Aşama 1'de Elde Edilen Sonuçlar 33
4.3 Aşama 2'de Elde Edilen Sonuçlar 35
4.4 Aşama 3'te Elde Edilen Sonuçlar 36
4.5 Aşama 4'te Elde Edilen Sonuçlar 36
5. TARTIŞMA VE SONUÇ 40
KAYNAKLAR 43
EKLER 48
ÖZGEÇMİŞ 59

SİMGELER DİZİNİ

E	Verimlilik katsayısı
R	Korelasyon katsayısı 2
R2	Korelasyon katsayısı
Kg	Kilogram
M3	Metreküp
Mpa	Megapascal

KISALTMALAR DİZİNİ

AB	Adaptive boosting
ABC	Artificial bee colony
AI	Artificial intelligence
ANFIS	Adaptive network based fuzzy inference system
ANN	Artificial neural network
BANN	Bagged artificial neural network
BPNN	Back propagation neural network
BRT	Bagging regression trees
CART	Classification and regression tree
DT	Decision trees
DVM	Destek vektör makineleri
ET	Extra trees
FIS	Fuzzy inference system
GAANN	Genetic algorithm artificial neural network
GANN	Gradient boosted artificial neural network
GB	Gradient boosting
GBM	Gradient boosting machine
GP	Genetic programming
GPR	Gaussian process regression
IGAER	Info gain attribute evaluator with ranker search
KNN	K-nearest neighbors
KRR	Kernel ridge regression
LaR	Lasso regression
LGBM	Light gradient boosting machine
LR	Linear regression
LSSVM	Least squares support vector machine
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MARS	Multivariate adaptive regression splines
MART	Multiple additive regression trees
MLP	Multilayer perceptron
MLR	Multiple linear regression
MR	Multiple regression
MSE	Mean squared error
RAE	Relative absolute error
RBFNN	Radial basis function neural networks
RF	Random forest
RMSE	Root mean square error
RO	Rassal orman

RR	Ridge regression
RRSE	Root relative squared error
SGD	Stochastic gradient descent
SMO	Sequential minimal optimization
SVM	Support vector machine
UCI	University of california irvine
WBANN	Wavelet bagged artificial neural network
WGBANN	Wavelet gradient boosted artificial neural network
XGB	Extreme gradient boosting
YSA	Yapay sinir ağları

ŞEKİLLER DİZİNİ

Şekil 3.1 K-En Yakın Komşu modeli için tayin edilen parametreler 17
Şekil 3.2 Karar ağaçları modelleri için Orange yazılımında tayin edilen parametreler 18
Şekil 3.3 Destek Vektör Makineleri modeli parametreleri 19
Şekil 3.4 Olasılıksal Dereceli Azalma modeli parametreleri 20
Şekil 3.5 Rassal Orman modeli parametreleri 21
Şekil 3.6 Yapay Sinir Ağları parametreleri 22
Şekil 3.7 Doğrusal Regresyon parametreleri 23
Şekil 3.8 Gradyan Artırma parametreleri 24
Şekil 3.9 Adaptif Yükseltme parametreleri 25
Şekil 3.10 Orange yazılımında oluşturulan tahmin modelleri şeması 28
Şekil 3.11 Orange yazılımında bütün veri seti kombinasyonlarını denemek için oluşturulan örnek bir şema 31
Şekil 4.1 Tek Değişkenli Regresyon uygulama sonuçları 32
Şekil 4.2 RReliefF metodu uygulama sonuçları 33
Şekil 4.3 Geliştirilen beton basınç dayanımı yazılımına girilen 1. örnek 39
Şekil 4.4 Geliştirilen beton basınç dayanımı yazılımına girilen 2. örnek 39

ÇİZELGELER DİZİNİ

Çizelge 3.1 Kullanılan temel veri seti ile ilgili birtakım istatistikler 14
Çizelge 3.2 1030 örnekli beton basınç dayanımı ham veri setinden örnek bir kısım ... 28
Çizelge 3.3 Ön işleme yöntemleri açıklamaları 29
Çizelge 3.4 "Numune Yaşı" parametresine uygulanan formül sonuçları 30
Çizelge 3.5 Üretilen örnek bir veri setinden kısımlar 31
Çizelge 4.1 Aşama 1 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları 34
Çizelge 4.2 Aşama 1 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları 34
Çizelge 4.3 Aşama 2 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları 35
Çizelge 4.4 Aşama 2 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları 35
Çizelge 4.5 Aşama 3 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları 36
Çizelge 4.6 Aşama 3 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları 36
Çizelge 4.7 Aşama 4 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları 37
Çizelge 4.8 Aşama 4 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları 37
Çizelge 4.9 Aşama 4 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tüm modellere göre tahmin sonuçları 38
Çizelge 4.10 Aşama 4 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tüm modellere göre tahmin sonuçları 38

1. GİRİŞ

Beton geniş bir uygulama alanına sahip olan temel bir yapı malzemesidir. Betonun su, agrega ve çimento temel bileşenleridir. Bunlara ek olarak dayanımını artırmak veya bazı fiziksel ve mekanik özelliklerini değiştirmek amacıyla bazı katkılar ilave edilmektedir. Bunlar mineral veya kimyasal katkılar olabilir. Göründüğünden daha karmaşık bir yapısı vardır (Özbakır 2015). Tüm bu bileşenlerin farklı oranlarıyla farklı beton basınç dayanımları elde edilir. Daha başka şekilde ifade edersek beton basınç dayanımı içerdiği materyallere ve döküldükten sonra geçen süreye bağlı olarak değişebilen bir nonlineer fonksiyondur. Ve basınç dayanımını belirlemek uzun süren laboratuvar çalışmaları ve yeterli sayıda numune gerektirir (Yeh 1998).

Yapay zekâ (AI), makinelerin ve özellikle de bilgisayarların insan zekasıyla benzer bir şekilde; düşünme, analiz etme, eyleme geçme gibi olguları yapmasını hedefleyen bir bilim dalıdır (Pirim 2006). Teknolojinin insan hayatındaki yeri yadsınamaz olduğundan yapay zekâ disiplini de günümüzde inanılmaz bir hızla gelişmektedir. Ve çok farklı alanlarda kendine bir yer edinmeye başlamıştır. Günümüz yapay zekâ teknolojisiyle görüntüler ve sesler makineler tarafından algılanabilmekte, matematiksel teoremler kanıtlanabilmekte ve gelecek adına çok tutarlı tahminler yapılabilmektedir.

Makine öğrenmesi (machine learning), yapay zekâ biliminin bir alt dalıdır. Makine öğrenmesi yöntemleri ile mevcut veri setleri algoritmalar tarafından incelenir ve algoritmaların aynı insan zekâsı gibi eğitilmeleri sağlanır. Veri setlerindeki kalıpları ve bağıntıları öğrenen algoritma ileriye dönük tutarlı tahminlerde de bulunmaktadır. Bunun için bolca matematiksel ve istatistiksel yöntemlerden faydalanır (Çağlayan 2018). Veri seti büyüklüğü arttıkça tahmin başarısı azalmak yerine daha da artar. Çünkü veri setini algoritmanın tüm detaylarıyla anlamasına olanak tanınmış olur. Veri setleri birçok bilim dalına ait olabileceğinden çoğu bilim dalında makine öğrenmesi yöntemleri verimli bir şekilde kullanılabilmektedir.

İnşaat mühendisliği alanının temel yapı taşlarından biri betondur. Yapılan çalışmada, beton basınç dayanımı verileri temin edilerek, makine öğrenmesi yöntemleriyle tahminler yapılmıştır. Temin edilen veri seti, 17 adet farklı laboratuvardaki deneyler sonucunda oluşturulan 1030 örnekli bir veridir. Veri setinde toplam 9 adet parametre (özellik) bulunur. Bunlardan, "Çimento", "Yüksek Fırın Cürufu", "Uçucu Kül", "Su", "Süper Akışkanlaştırıcı", "Kaba Agrega", "İnce Agrega" ve "Numune Yaşı" giriş parametresi ve "Beton Basınç Dayanımı" ise tahmin hedefi parametresi olarak kullanılmaktadır. Yapılan çalışmada, diğer çalışmalardan farklı olarak tahmin verimini artıran ön işleme (preprocessing) yöntemleri üzerinde durulmuştur. Kullanılan beton basınç dayanımı veri setine özgü ön işleme yöntemleri geliştirilmiş ve geliştirilen bu yöntemler farklı makine öğrenmesi tahmin modelleri ile süreçteki her aşamada test edilmiştir. Üstelik yapılan bu çalışmalar sayesinde kullanılan yazılım ve seçilen model parametrelerine göre tahmin modellerinden hangisinin daha verimli olduğu da incelenmiş olmaktadır.

Çalışmada veri madenciliği ve makine öğrenmesi yazılımı olan Orange yazılımı kullanılmıştır. Yazılımın olanak tanıdığı makine öğrenmesi modellerinden K-En Yakın Komşu (K-Nearest Neighbors), Karar Ağaçları (Decision Trees), Destek Vektör Makineleri (Support Vector Machine), Olasılıksal Dereceli Azalma (Stochastic Gradient Descent), Rassal Orman (Random Forest), Yapay Sinir Ağları (Artificial Neural Network), Doğrusal Regresyon (Linear regression), Gradyan Artırma (Gradient Boosting) ve Adaptif Yükseltme (Adaptive Boosting) modelleri ile tahminler yapılmıştır. Ayrıca bu modellerden en verim alınanlarıyla İstifleme (Stacking) yöntemi ile topluluk modelleri oluşturulmuştur.

Oluşturulan tüm tahmin modellerinde, performans ölçüm değerleri incelendiğinde, değerlerin uygulanan özgün ön işleme (preprocessing) yöntemleri ile her safhada dahada iyileştiği görülmektedir. Ayrıca nihai sonuçlara göre kalıplaşmış modellerden Gradyan Artırma (Gradient Boosting) modeli performansı en iyi modeldir. Bununla birlikte, İstifleme (Stacking) medotodu ile birleştirilen topluluk (ensemble) modellerinden, Gradyan Artırma (Gradient Boosting) ve Yapay Sinir Ağları (Artificial Neural Network) yöntemlerinin birleştiği model en verim alınan topluluk modelidir. Bu

çalışmada performans ölçüm değerleri Ortalama Karesel Hata (Mean Squared Error), Ortalama Hata Karesinin Kökü (Root Mean Square Error), Ortalama Mutlak Hata (Mean Absolute Error) ve Korelasyon Katsayısı (Correlation Coefficient(R2)) formülleri kullanılarak elde edilmektedir.

Son olarak, konu bütünlüğünün sağlanması amacıyla Orange yazılımı altyapısını kullanan ve çalışmada en iyi sonucu veren veri seti ile eğitilen modellerle çalışan basit bir beton basınç dayanımı tahmin yazılımı da geliştirilmiştir. Yazılım, Autocad yazılımının bir eklentisi olan Visual Basic for Applications dili ile geliştirilmiştir. Bu yazılım vasıtasıyla, denenen bireysel örneklerinin tutarlılığı gözlemlenmiştir.

2. LİTERATÜR ÖZETİ

Çeşitli istatistik ve yapay zekâ yöntemleri ile tahmin yapmak için kullanılan alanlar her geçen gün artmaktadır. Bilgisayar teknolojisinin de gelişmesi ile tahmin sonuçlarında büyük oranda iyileşmeler gözlemlenmektedir. Birçok bilim ve mühendislik alanlarında farklı metotlar uygulanarak tatmin edici sonuçlar alınmaktadır. Bu tahminler ile gerek duyulan kapsamlı laboratuvar çalışmaları azalmaktadır. Bu sebeple gerekli iş gücü, zaman ve maliyetten oldukça tasarruf edilmektedir. Birçok alanda olduğu gibi inşaat mühendisliği alanında da anılan yöntemler sıklıkla kullanılmakta ve diğer çoğu branşta olduğu gibi inşaat mühendisliği alanında da oldukça verimli tahmin sonuçları elde edilmektedir. İnşaat mühendisliği bilimi ve teknolojisinde oldukça geniş bir alana sahip beton araştırmalarında bahse konu yöntemler hızla yaygınlaşmaktadır. Birçok akademik çalışmada oldukça tutarlı sonuçlar elde edilmiştir. Yapılan literatür taramasında, bu çalışma kapsamında da kullanılan Kaliforniya Üniversitesi (UCI) bünyesinde açık kaynak olarak dağıtılan 1030 örnekli beton basınç dayanımı ve yine aynı üniversitenin yayınladığı beton slumpı veri setlerini kullanan çalışmalara bolca rastlanmıştır. Bazı çalışmalarda hibrit yöntemler bazılarında ise hibrit olmayan yöntemler denenmiştir. Bahsi geçen çalışmaların verimleri, çalışmaları hazırlayanlar tarafından genel itibariyle Korelasyon Katsayısı (R2), Ortalama Mutlak Hata (MAE), Ortalama Karesel Hata (MSE), Ortalama Hata Karesinin Kökü (RMSE) vb. performans ölçüm yöntemleri ile sayısal olarak ölçülmüştür.

Chine *et al.* (2010) topladıkları verilerle beton slumpı tahmini yapmışlardır. “Çimento” (C), “Uçucu Kül” (FA), “Yüksek Fırın Cürufu” (SL), “Su” (W), “Süper Akışkanlaştırıcı” (SP), “İnce Agrega” (FN) ve “Kaba Agrega” (CA) olmak üzere 7 giriş verisine (W+SP) / (C), (W+SP) / (C+FL+SL), (W+SP) / (C+FL+SL+CA+FA) ve (CA+FA) / (C+FL+SL) oranlarını ekleyerek toplam 11 giriş bir çıkışlı (slump) bir model oluşturmuşlardır. Veri seti eğitim ve test verileri olarak ikiye ayrılmış ve BPNN ve MLR yöntemleri ile tahmin yapılmıştır. Çalışma sonucunda BPNN modeli daha iyi sonuç vermiştir. Bu modelde 2 adet gizli katman ve gizli katman başına 8 nöron bulunmaktadır. Modelde Dereceli Azalma algoritması da mevcuttur. Anılan modelin test verilerindeki performansı 0.5366 R2 ve 7.4636 RMSE değeri ile ölçülmüştür.

Chou *et al.* (2011) beton basınç dayanımı tahmininde farklı yöntemleri karşılaştırmışlardır. Çalışmada iki adet makine öğrenmesi yöntemi (Yapay Sinir Ağları, Destek Vektör Makineleri), bir adet istatistiksel yöntem (MR) ve iki adet topluluk öğrenmesi yöntemi (MART, BRT) kullanılmıştır. Kullanılan veri setine 10 Katmanlı Çapraz Doğrulma uygulanmıştır. Tüm yöntemlerin pozitif yönleri olsa da MART yöntemi diğerlerine göre ön plana çıkan yöntem olmuştur. Çalışma sonucunda bahse konu yöntemle ortalama R2 değeri 0.9108, ortalama MAPE değeri 13.8855, ortalama RMSE değeri 4.9489 olarak elde edilmiştir.

Erdal *et al.* (2013) bir takım birleşik modellerle beton basınç dayanımı tahminleri yapmışlardır. Kullanılan verile seti rastgele olarak eğitim ve test verilerine birkaç farklı şekilde ayrılmıştır. Ayrıca çalışmada gizli katman sayıları ve gizli katmanların nöron sayılarında bazı denemeler yapılmıştır. En uygun sonuçlar %90 eğitim ve %10 test verisi olarak bölünen veri setinde, her birinde 10 düğüm bulunan ve 10 gizli katman sayısı olan modelden alınmıştır. BANN, GBANN, WBANN ve WGBANN modelleri kullanılmıştır. Bu modellerden en iyi verimi WGANN modeli vermiştir. 0.9528 R2, 5.75 RMSE ve 4.83 MAE değeri elde edilmiştir.

Chou *et al.* (2014) 5 farklı beton basınç dayanımı veri seti ile tekil ve kombine modellerle çalışma yapmışlardır. Verilere 10 Katmanlı Çapraz Doğrulama uygulanmıştır. Kullanılan tekil yöntemler MLP, ANN, DVM, CART ve LR'dir. Bu yöntemlerin Oylama (Voting), Torbalama (Bagging) ve İstifleme (Stacking) yöntemleri ile birleştirilmiş haliyle de tahminler yapılmıştır. Genel itibari ile bir veri seti hariç, topluluk modelleri daha güzel sonuçlar vermiştir. Ayrıca İstifleme yöntemi diğer yöntemlere göre tahmin edilen verilere biraz daha uyum sağlamıştır. Verilerin hemen hemen hepsinden tatmin edici sonuçlar alınabilmiştir. Fakat kullanılan veri çeşitlerinin birinde İstifleme metodu ile birleştirilerek kullanılan CART, DVM ve LR modelleri ile çok daha iyi performans ölçüm değerleri elde edilmiştir. Bahsedilen modelde MAE, RMSE, MAPE hata değerleri sırasıyla 1.09, 1.51 ve 2.11 bulunmuştur.

Yuan *et al.* (2014) beton basınç dayanımı tahmininde iki adet hibrit model kullanmışlardır. Bunlar GAANN ve ANFIS modelleridir. Kullandığı veri setindeki

“Numune Yaşı” parametresinden bir takım 28 günlük örnekleri seçmiş ve 180 satırlı bir veri seti elde etmişlerdir. Bu verinin 150 örneğini eğitim, 30 örneğini test verisi olarak kullanmışlardır. Sonuç olarak 7 giriş ve 1 çıkışlı bir model oluşturulmuştur. Alınan sonuçlara göre ANFIS modeli verim olarak bir adım öndedir. R2 değeri 0.95, RMSE değeri 1.46 olarak ölçülmüştür.

Aydoğmuş *et al.* (2015) tekil yöntemler olan CART, DVM, MLP ve RBFNN ayrıca bu tekil yöntemlerin toplu formlarıyla beton basınç dayanımı tahminleri üretmişlerdir. Kullanılan veri setine 3 ve 5 Katmanlı Çapraz Doğrulama uygulanmıştır. Karşılaştırma sonucunda 5 Katmanlı Çapraz Doğrulama uygulanmış Torbalama (Bagging) MLP modelinden en yüksek verim alınmıştır. Ayrıca torbalama yöntemi uygulanmış modellerin tekil modellere göre daha verimli olduğu anlaşılmıştır. Torbalama yöntemi uygulanmış MLP modelinde R, RMSE ve MAE değerleri sırasıyla 0.7741, 11.24, 8.27 olarak elde edilmiştir.

Erdal (2015) çalışmasında Yapay Sinir Ağları (ANN) ve Destek Vektör Makineleri (SVM) yöntemlerini karşılaştırmıştır. Kullanılan 1030 örnekli veri seti Rassal Ayırma (Random Split) yöntemi ile eğitim ve test verilerine ayrılmıştır. Farklı oranlarda iki model denenmiştir (Model 1: %90 eğitim %10 test, Model 2: %80 eğitim, %10 test). DVM için birkaç çekirdek fonksiyonu denendikten sonra polynomial fonksiyonda karar kılınmıştır. Bu model ve YSA modeli ile ilgili en uygun parametreler seçildikten sonra Model 1’den daha iyi performans alınmış ve ayrıca DVM yöntemi, YSA modeline göre daha iyi sonuçlar vermiştir. Daha önce yapılan çalışmalar ile karşılaştırma da yapılmıştır. Önceki çalışmaların aksine DVM yönteminin üstün olması çekirdek fonksiyonunun kullanılmasına bağlanmıştır. Anılan yöntemde RMSE değeri 4.54, MAE değeri 3.30 ve R2 değeri 0.952 olarak elde edilmiştir.

Durap ve Doğan (2015) beton basınç dayanımı ile ilgili bir uzman sistem geliştirme çalışması yapmışlardır. Çalışmanın bir safhasında kullanılan veri setiyle Doğrusal Regresyon ve Yapay Sinir Ağları yöntemleriyle beton basınç dayanımı tahminleri yapmışlardır. Tahmin sonucunda YSA yönteminden daha fazla verim alınmıştır. Fakat bulunan ölçüm değerleri (MAE=7.13, RMSE=8.94, RAE=%52.88, RRSE=%53.49)

tatmin edici bulunmamıştır. Bu sebeple veriler normalizasyona tabi tutulduktan sonra K En Yakın Komşu algoritması kullanarak bir yazılım geliştirilmiştir. Yazılımda sadece beton basınç dayanımı değil karışımdaki tüm birimler için değerler elde edilebilmektedir. Eğitilmiş bir desenle bir fonksiyon elde etmeden, verilerin kendisi kullanılarak oluşturulmuş K En Yakın Komşu algoritması ile hata değerlinin tamamı sıfıra eşitlenmiştir.

Özkan and Altın (2016) çalışmalarında beton basınç dayananımı tahmini yapmışlardır. Kullanılan 1030 örnekli veri setini rastgele dağıtarak %80 eğitim ve %20 test verisi olarak kullanmışlardır. 8 nöronlu giriş katmanı, 1 nöronlu çıkış katmanı ve 1'den 50'ye kadar nöron sayıları denenerek oluşturulan 1 adet gizli katman içeren bir yapay sinir ağı modeli geliştirmişlerdir. Oluşturulan yapay sinir ağında farklı eğitim algoritmaları ve gizli katman transfer fonksiyonları denenmiştir. En iyi sonuç gizli katmanında 27 nöron bulunan, Levenberg–Marquardt Geri Yayılım eğitim algoritması ve Tansig gizli katman transfer fonksiyonu kullanılan modelde alınmıştır. Eğitim veri seti R2 değeri 0.99031 ve test veri seti R2 değeri 0.97106 olarak elde edilmiştir.

Siraj *et al.* (2016), 1030 örnekli veri setinde "Numune Yaşı" parametresinde değeri 28 olan örnekleri seçmiştir. Diğerlerini iptal ederek veri örnek sayısını 425'e düşürmüştür. Beton basınç dayanımı tahmini yapılmış (çıkış) ve anılan sebepten giriş sayısı 8'den 7'e düşürülmüştür. Veri seti %70 ve %30 oranında rastgele eğitim ve test verisi olarak ayrılmıştır. Yapay Sinir Ağları (ANN), FIS ve ANFIS yöntemleri kullanılmıştır. YSA modelinde 1 adet gizli katman bulunmakta olup, çalışmada gizli katman nöron sayıları 1'den 25'e kadar denenmiştir. Seçilen diğer parametrelere göre (Logsig, Purelin vs.) optimum sonuç 24 gizli katman nöron sayısı olan modelden alınmıştır. Bütün çalışmanın sonunda yapılan ölçümlere göre FIS en kötü sonuçları vermiştir. YSA ve ANFIS'ten neredeyse denk sonuçlar alınmıştır. ANFIS modelinde MAE değeri 4.19, MSE değeri 29.40, RMSE değeri 5.42 ve R2 değeri 0.86 olarak elde edilmiştir. ANN modelinde ise bu değerler sırasıyla 4.03, 30.41, 5.51 ve 0.86'dır.

Saad *et al.* (2016) çalışmalarında giriş verilerinden en etkili olanları bir Özellik Seçimi metodu olan IGAER metodu ile seçerek beton basınç dayanımı tahmin doğruluğunu

artırmışlardır. Özellik Seçimi sonucunda “Kaba Agrega” en az etkili parametre olarak tespit edilmiştir. Bu sebeple giriş verileri 8’den 7’ye düşürülerek, 7 giriş 1 çıkışlı bir yapay sinir ağı modeli tasarlanmıştır. Kullanılan veriler %70 eğitim, %15 test ve %15 doğrulama olmak üzere üç kısma ayrılmıştır. Levenberg-Marquardt yöntemi ile eğitilen İleri Beslemeli Yapay Sinir Ağı modeli kullanılmıştır. Modelin genel doğruluk oranı %97.949 olarak bulunmuştur.

Khashman and Akpınar (2017) beton basınç dayanımı tahmininde farklı bir Yapay Sinir Ağları modeli geliştirmişlerdir. Kullanılan orijinal veri setinde “Çimento”, “Su”, “Yüksek Fırın Cürufu”, “Uçucu Kül”, “Süper Akışkanlaştırıcı”, “Kaba Agrega”, “İnce Agrega”, “Numune Yaşı” Ve “Beton Basınç Dayanımı” parametreleri vardır. Tahmin edilen hedef parametre “Beton Basınç Dayanımı” dır. Diğerleri giriş katmanında kullanılmıştır. Fakat giriş katmanına 9. giriş değeri olarak orijinal verilerde net olarak bulunmayan “su/bağlayıcı oranı” eklenmiştir. Su orijinal veri setinde bulunmaktadır. Bağlayıcı ise yine orijinal veri setinde Bulunan “Yüksek Fırın Cürufu”, “Uçucu Kül” ve “Çimento” parametrelerinin toplamıdır. Çıkış katmanında bulunan beton basınç dayanımı ise düşük, orta ve yüksek dayanım olmak üzere 3 guruba ayrılmıştır. 20’den küçük “düşük”, 20 ve 50 arası “orta” ve 50’den büyük “yüksek dayanım” olarak tasarlanmıştır. Çıkış katmanının nöron sayısı 3 tanedir. Çünkü “düşük”, “orta” ve “yüksek dayanım” ikili sistem (binary) olarak işlenmiştir. Ağ modeli ayrıca 8 nöronlu 1 adet gizli katmana sahiptir. Eğitim ve test verisi olarak ta 40:60, 50:50, 60:40 gibi oranlar denenmiştir ve en iyi sonuçlar 50:50 eğitim ve test oranından alınmıştır. Sonuç olarak modelin genel doğruluk oranı 86,02 olarak bulunmuştur.

Santosa and Santosa (2017) beton basınç dayanımı tahmininde Yapay Sinir Ağları modelini Genetik Algoritma Tabanlı Yapay Sinir Ağları Modeli ile karşılaştırmışlardır. Yaptıkları tahminlerde veri seti için Çapraz Doğrulama yöntemi uygulamışlardır. Kullandıkları ağ modelleri 7 giriş, 1 çıkışlıdır ve ağlarda 1 adet gizli katman mevcuttur. Denenen parametrelerle beraber karşılaştırma sonunda Genetik Algoritma modelinin RMSE değeri 1.389871 bulunmuştur ve öne çıkan model olarak tespit edilmiştir.

Keleş *et al.* (2018) çalışmalarında ABC algoritması olarak isimlendirilen bir yöntemle kullanılan veri setine Özellik Seçimi (Feature Selection) metodu uygulamışlardır. Kullanılan veriler beton basınç dayanımının yanında slump değerleri de içermektedir. Beton basınç dayanımı tahmini yapılmasından ötürü verilerden "slump" ve "flow" parametreleri çıkarılarak ilk önce 7 giriş 1 çıkışlı bir model tasarlanmıştır. Ayrıca revize edilen veri setlerine 10 katmanlı Çapraz Doğrulama uygulanmıştır. Rassal Orman, SMO, K-En Yakın Komşu ve Karar Ağaçları gibi sınıflandırma algoritmaları uygulandıktan sonra doğruluk ve "F-Measure" değerleri elde edilmiştir. En uygun değerler Rassal Orman metodu ile elde edilmiştir. ABC yöntemi ile Özellik Seçimi uygulandıktan sonra Veri Seti 7 parametreden (özellik) 3 parametreye düşülmüştür ("Uçucu Kül", "Su", "Çimento"). 91.2621 doğruluk ve 0.913 "F-measure" değerleri ile Özellik Seçimi uygulanmadan önceki değerler birebir yakalanmıştır.

Sun *et al.* (2019) geliştirilmiş Destek Vektör Makinesi çalışması yapmışlardır. Farklı bir veri seti ve farklı bir konu ile ilgili çalışılsa da makalesinin bir bölümünde önerdiği yöntemi bu çalışmada kullanılan verilerle yani 1030 örnekli beton basınç dayanımı veri seti ile test etmişlerdir. R değerini 0.94 elde ederek olumlu sonuç almışlardır.

Mohan (2019) çalışmasında Destek Vektör Makineleri ve Yapay Sinir Ağları yöntemlerini karşılaştırmıştır. Kullanılan 1030 örnekli verilere Çapraz Doğrulama yöntemi uygulanmıştır. Geri Yayılım yöntemi kullanılan 8 Giriş 1 çıkışlı ve farklı gizli katman nöron sayıları olan Yapay Sinir Ağları modeli, DVM yöntemine göre daha yüksek doğruluk oranlarına ulaşmıştır. Doğruluk oranı değeri YSA modelinde 0.94, DVM modelinde ise 0.78 olarak bulunmuştur.

Lakshmanarao *et al.* (2019) iki aşamalı çalışmalarında beton basınç dayanımı tahmini yapmışlardır. İlk aşamada veri setindeki tüm giriş verileri kullanılarak tahmin yapılmıştır. İkinci aşamada ise aynı tahmin yöntemleri kullanılmış fakat MLR ile geriye doğru eleme yaparak Özellik Seçimi yöntemi uygulanmıştır. Bunun sonucunda beton basınç dayanımı tahmini için en etkili parametreler "Çimento", "Yüksek Fırın Cürufu", "Uçucu Kül", "Su" ve "Numune Yaşı" olarak bulunmuştur. Tahmin için 4 adet yöntem kullanılmıştır. Bunlar MLR, DVM, Karar Ağaçları ve Rassal Orman metotlarıdır. İki

aşamanın da sonuçları hemen hemen aynı olup en iyi sonuç Rassal Orman yöntemi ile alınmıştır. R2, MAE ve RMSE değerleri sırasıyla 0.91, 3.5 ve 5 şeklinde elde edilmiştir.

Yunani (2019) çalışmasının ilk safhasında Gradyan Artırma, ET, LGBM, XGB gibi ağaç tabanlı yöntemlerle ve ayrıca bu yöntemleri Oylama (Voting) yöntemiyle birleştirerek beton basınç dayanımı tahmini yapmıştır. Modellerde 10 Katmanlı Çapraz Doğrulama uygulanmıştır. Çalışmanın ikinci safhasında, kullanılan veri setine Özellik Seçimi (Feature Selection) uygulanarak tekrar aynı yöntemlerle modeller eğitilmiştir. Yapılan çalışmalarda "Uçucu Kül" en az etkiye sahip parametre olarak belirlendiğinden veri setinden çıkarılmıştır. Ana veri setine ve Özellik Seçimi yöntemi uygulanmış veri setine göre yapılan tahminler karşılaştırıldığında Özellik Seçimi uygulanmış verilerde tahmin başarısı olarak cüzi bir iyileşme görülmüştür. Anılan modelde R2 değeri 0.824 olarak bulunmuştur.

Silva *et al.* (2020a) Rassal Orman, DVM ve YSA yöntemleriyle beton basınç dayanımı tahmini yapmışlardır. Tahminde kullanılan veri setine önce Çapraz Doğrulama ve bir hiperparametre optimizasyonu yöntemi olan GridSearchCV yöntemleriyle ön değerlendirmeden geçirilmiş ve akabinde 8 giriş 1 çıkışlı tahmin modelleri oluşturulmuştur. Bulunan ölçüm değerleri karşılaştırıldığında RO yönteminin hafif bir farkla önde olduğu anlaşılmaktadır. RO yönteminde R2 değeri 0.902, RMSE değeri 5.614 olarak bulunmuştur.

Biswas *et al.* (2020) yaptıkları beton basınç dayanımı tahminlerinde MARS, LSSVM ve GP gibi bu alanda yeni sayılabilecek yöntemler kullanmışlardır. Yapılan çalışmalar sonucunda tüm yöntemlerden tatmin edici sonuçlar alınsada MARS yönteminin küçük bir ölçüde önde olduğu kanısına varılmıştır. Bu yöntemde iki aşamada kullanılan veri setleri üzerinde işlem yapılmıştır. Birinci aşamada veri setleri %70 ve %30 eğitim ve test verisi olarak ayrılmıştır. İkinci aşamada veri setine 20 katmanlı Çapraz Doğrulama uygulanmıştır. Sırasıyla E, MAPE ve RMSE değerleri 0.8539, 18.6039 ve 6.6832 olarak bulunmuştur.

Jha *et al.* (2020) çalışmalarında Doğrusal Regresyon, RR, LaR, Karar Ağaçları, Rassal Orman, Adaptif Yükseltme ve son olarak Gradyan Artırma modelleriyle beton basınç dayanımı tahmini yapmışlardır. Tahmin için kullanılan veri setine 5 katmanlı Çapraz Doğrulama uygulanmıştır. İkinci aşamada da kullanılan verilerdeki parametrelerden sonuca en çok etki edeni saptanmıştır. Çalışma sonucunda Gradyan Artırma modeli en iyi sonucu vermiştir. Ayrıca çalışmaya göre veri setindeki "Çimento" ve "Numune Yaşı" parametrelerinin sonuca en fazla etki ettiği tespit edilmiştir. Hata değerlerinden en iyileri 0.91 R2 değeri ve 5.2 RMSE değeridir.

Tavares *et al.* (2020) beton basınç dayanımı tahmini için Levenberg-Marquardt Geri Yayılım algoritmasıyla beraber ileri beslemeli bir Yapay Sinir Ağı modeli geliştirmişlerdir. Kullanılan veri seti gereği 8 giriş verisi bir çıkış verisi mevcuttur. Çeşitli denemeler sonucunda iki gizli katman ve katman başına sırasıyla 15 ve 10 nöron kullanılmıştır. Veriler %70 eğitim, %15 test ve %15 doğrulama olarak bölümlere ayrılmıştır. Tüm bu seçimlerden sonra ortalama MSE değeri 27.4318, ortalama R değeri 0.96559 olarak bulunmuştur.

Silva *et al.* (2020b) beton basınç dayanımı tahmini için Karar Ağaçları ve Rassal Orman yöntemlerini kullanmışlardır. Kullanılan veri seti iki kısma ayrılmış ve %85 (875 örnek) eğitim verisi olarak, kalan %15 (155 örnek) test verisi olarak kullanılmıştır. Ayrıca kullanılan yöntemlerde parametre ayarları için bir parametre seçim yöntemi olan GridsearchCV isimli yöntem kullanılmış ve optimum parametre ayarları yapılmaya çalışılmıştır. Çalışmada Rassal Orman metodu, elde edilen 0.90 R2 değeri, 5.60 RMSE değeri ile düşük bir farkla daha iyi performans sergilemiştir.

Kaloop *et al.* (2020) beton basınç dayanımı tahminleri için hibrit yöntemler geliştirmişlerdir. İki aşamalı çalışmalarında ilk önce MARS yöntemini kullanarak veri setindeki giriş özelliklerinden en önemlilerini seçmişlerdir. Ham halinde 8 giriş değeri bir hedef değeri bulunan verilerin MARS yöntemi kullanılarak azaltılması sonucu "Uçucu Kül" ve "Kaba Agrega" çıkarılarak 6 giriş bir çıkışlı bir veri modeli oluşturulmuştur. Daha sonra en etkili özelliklerle oluşturulan bu yeni veri setiyle Kernel KRR, GBM ve GPR yöntemleri ile tahminler yapılmış ve sonuçlar karşılaştırılmıştır.

Bazı safhalarda 5 katmanlı Çapraz Doğrulama yöntemi de uygulanmıştır. En yüksek verim alınan model MARS tabanlı GBM modelidir. 0.965 R ve 0.037 MAE değerleri elde edilmiştir.

Muliauwan *et al.* (2020) çalışmalarında tekil ve hibrit yöntemler kullanarak beton basınç dayanımı tahminleri yapmışlardır. Veri seti %70'e %30 eğitim ve test verisi olarak bölünmüştür. Tekil yöntemler olarak YSA, DVM ve LR yöntemleri kullanılmıştır. Hibrit yöntemler ise YSA+DVM+LR, YSA+DVM, YSA+LR ve DVM+LR olmak üzere 4 çeşittir. Tüm bu çalışmalar sonucunda beklenilenin aksine tekil yaklaşımlardan olan ANN modelinden en yüksek verim alınmıştır. R, RMSE, MAPE, MAE değerleri sırasıyla 0.909, 7.176, 19.93 ve 5.550 olarak elde edilmiştir.

Yukarıdaki belirtilen birçok çalışmada, farklı yöntemlerle tatmin edici sonuçlar alınabilmiştir. Fakat veri setlerinde K Katmanlı Çapraz Doğrulama (K-Fold Cross Validation) yöntemi kullanılmayan çalışmalarda olası aşırı öğrenme (overfitting) sorunu sebebiyle belirlenen sonuçlar objektif bir şekilde elde edilememiş olabilir. Ayrıca bahsi geçen çalışmalar arasında veri ön işleme (preprocessing) yöntemlerinin önemini vurgulayan ve özgün ön işleme yöntemleri geliştirilen çalışma bulunmamaktadır.

3. MATERYAL VE YÖNTEM

Bu bölümde; yapılan çalışmada kullanılan veri seti detayları, kullanılan yöntemler, çalışma sürecine ait bilgiler ve bazı tanımlar açıklanacaktır.

3.1 Çalışmada Kullanılan Veri Setleri ve İlgili Bazı Tanımlar

Çalışmada kullandığımız temel veri seti; 17 Adet farkı laboratuvardaki deneyeler sonucunda oluşturulan, 1030 adet örnekli bir veri setidir (UCI IRVANE, https://archive.ics.uci.edu/ml/datasets/, 2007). Veri setindeki giriş parametreleri "Çimento", "Yüksek Fırın Cürufu", "Uçucu Kül", "Su", "Süper Akışkanlaştırıcı", "Kaba Agrega", "İnce Agrega" ve "Numune Yaşı" dır. Çıkış yani hedef parametresi ise yapılan çalışmanın genel olarak tahmin amacı olan "Beton Basınç Dayanımı" dır. Sonuç olarak 8 adet giriş parametresi ve 1 adet çıkış parametresinden oluşmaktadır. Veri seti ile ilgili detaylı bilgiler Çizelge 3.1'de gösterilmektedir.

Yapılan çalışmada, veri setine ilk uygulan algoritmalar Özellik Seçimi (Feature Selection) algoritmalarıdır. Bu metotla veri setini analiz etme ve anlama amacıyla birtakım tahliller yapılmıştır. Bu sayede veri seti hakkında bazı önemli çıkarımlar yapılmıştır. Fakat tahmin yaparken, veri seti parametrelerinin (özellik) sayısını azaltma işlemine gerek görülmemiştir.

Daha sonra ön işleme (preprocessing) yöntemleri veri setine uygulanmıştır. Bu safhada, bahsi geçen temel veri seti baz alınarak sayısız farklı veri seti üretilmiştir. Ama üretilen veri setlerinde verinin genel kalıbı korunmuştur. Bu işlemlerin amacı gürültüyü (noise) azaltmak ve veri seti kalıbını daha belirgin hale getirmektir. Ayrıca kullanılan tahmin modelinin veri setini daha rahat anlamasını sağlamaktır.

Çizelge 3.1 Kullanılan temel veri seti ile ilgili birtakım istatistikler

	Ortalama	Orta Değer	Minimum Değer	Maksimum Değer	Kayıp Oranı
Çimento (Kg/M3)	281.1656	272.9	102	540	0%
Yüksek Fırın Cürufu (Kg/M3)	73.8955	22	0	359.4	0%
Uçucu Kül (Kg/M3)	54.1871	0	0	200.1	0%
Su (Kg/M3)	181.5664	185	121.75	247	0%
Süper Akışkanlaştırıcı (Kg/M3)	6.20311	6.35	0	32.2	0%
Kaba Agrega (Kg/M3)	972.9186	968	801	1145	0%
İnce Agrega (Kg/M3)	773.5789	779.51	594	992.6	0%
Numune Yaşı (Gün)	45.66	28	1	365	0%
Beton Basınç Dayanımı (Mpa)	35.8178	34.4428	2.33181	82.5992	0%

3.1.1 Veri ön işleme

Veri ön işleme (preprocessing) safhası tahmin kalitesini artırmak için hayati önem taşımaktadır. Kullanılan makine öğrenmesi tahmin modelinin, veri setini daha iyi analiz edip anlaması için gerçekleştirilen bir takım sadeleştirme, düzeltme vb. işlemler bütünüdür. Örneğin veri içerisindeki kayıp değerleri temizleme, kalıbı bozmadan gereksiz ve anlamayı zorlaştıran örnekleri azaltma, verideki değerleri 0 ve 1 aralığına göre oranlayarak anlamayı kolaylaştırma, farklı verileri birleştirme ve daha birçok şey veri ön işleme olarak değerlendirilebilir (Acuña 2011).

3.1.2 Makine öğrenmesinde gürültü tanımı

Temin edilen veri setlerindeki değerler bozuk, anlamsız veya gereksiz olabilir. Sıklıkla karşılaşılan bu durumda tahmin kalitesi düşer ve sonuçlar zarar görür. Bu durum gürültü (noise) olarak tanımlanır (Zhu and Wu 2004). Bu sebeple bu değerler olabildiğince veri setinden temizlenmelidir. Bu sayede veri kalıbı dahada belirgin hale gelerek modelin veri setini öğrenmesi kolaylaşmaktadır.

3.1.3 Özellik seçimi yöntemi

Özellik Seçimi (Feature Selection) yöntemini kısa bir şekilde tanımlayacak olursak, veri setini en iyi temsil eden veri parametrelerinin bulunması olarak ifade edilebilir. Yapılan tahminlerde her veri seti parametresi hedef parametrenin tahmin edilmesinde etkili olmayabilir. Ve hatta anlamayı zorlaştırarak tahmin kalitesini de düşürebilir. Bu sebeple hedef parametre ile ilgisi olmayan parametreler çıkarılırsa daha iyi sonuçlar elde edilebilir (Budak 2018). Bazı durumlarda veri setini daha iyi anlamak adına parametre eksiltmeden de özellik seçimi yöntemine başvurulabilir. Özellik seçimi yöntemi için birtakım algoritmalar geliştirilmiştir. Veri seti parametrelerinin veri içindeki önem sırası bu algoritmalarla belirlenmektedir.

3.1.4 Veri setinde bulunan bazı parametrelerin tanımları

Yüksek fırın cürufu (Blast furnace slag): Yoğun demir madeni ihtiva eden toprak, demir çelik fabrikalarındaki büyük fırınlarda 1500 dereceye ulaşan sıcaklıklarda yakılır. İçeriğindeki demir alındıktan sonra geriye kalan maddeye yüksek fırın cürufu denir. Kızgın cürufun üzerine basınçlı su uygulanarak ani olarak soğutulur ve öğütülmesiyle kuma benzer bir fiziksel bir yapı alır. Yani demir çelik üretiminde oluşan bir yan üründür (Bilgen ve Kavak 2005). Beton karışımında bulunması halinde basınç dayanımı değerlerine olumlu bir etkisi bulunmaktadır.

Uçucu kül (Fly ash): Elektrik üreten termik santrallerde kömür yakılmasıyla oluşan külün büyük bir kısmı gazlarla beraber kazandan çıkar. Çıkan bu yan ürüne uçucu kül denilmektedir. Özellikleri kömürün cinsi, tesis tipi ve yanma süreci gibi birçok parametreye bağlıdır. Elde edilen bu yan ürün çimento üretiminde veya direk beton karışımına katılarak beton üretiminde kullanılabilir. Beton basınç dayanımına aynı yüksek fırın cürufu gibi önemli oranda katkı sağlamaktadır (Kaplan ve Gültekin 2010).

Süper akışkanlaştırıcı (Superplasticizer): Süper akışkanlaştırıcılar yüksek oranda su azaltıcı ikinci nesil bir beton katkısıdır. Betondaki su miktarını artırmadan betonun

işlenebilirliğini artırırlar. Hem beton karışımında dikkat edilmesi gereken bir etmen olan su ve çimento oranının artmasını engellerler, hem de düzgün işlenebilir bir beton elde edilmesine yardımcı olurlar. Farklı kimyasal yapıda birçok türü bulunmaktadır (Rasheed *et al*. 2018).

3.2 Kullanılan Yazılımlar

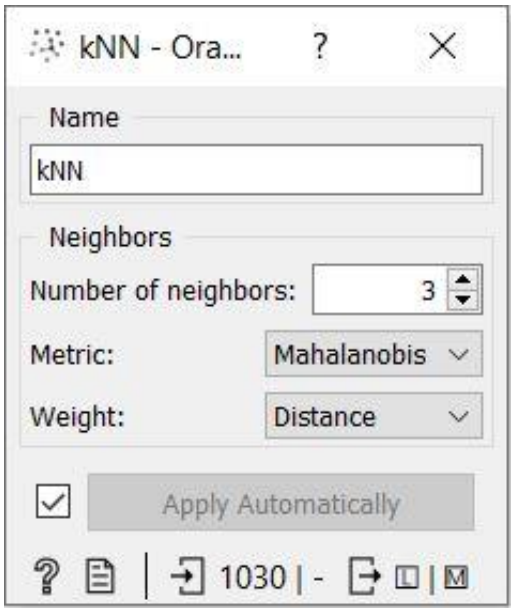
Bu çalışmada, bir makine öğrenmesi ve veri madenciliği yazılımı olan Orange kullanılmıştır. Orange birçok veri işleme, veri görselleştirme ve makine öğrenmesi modülü bulunan ve kolayca oluşturulabilen görsel şemalar sayesinde veri analizi ve tahminler yapılabilen açık kaynak kodlu bir yazılımdır. Bu çalışmada Orange yazılımının 3.32 versiyonu kullanılmıştır. Fakat, her ne kadar Orange yazılımının da veri işleme ile ilgili yetenekleri olsa da veri sıralama, veri bölme, tekrar veri üretme ve benzeri çalışmalarda Microsoft Excel yazılımından ve eklentisi olan Microsoft Visual Basic for Applications yazılımlarından yüksek oranda faydalanılmıştır. Ayrıca çalışma kapsamında geliştirilen beton basınç dayanımı yazılımı için Autocad Programının eklentisi olan Visual Basic for Applications yazılımı da kullanılmıştır.

3.3 Kullanılan Makine Öğrenmesi Yöntemleri ve Tayin Edilen Model Parametreleri

3.3.1 K en yakın komşu

K En Yakın Komşu (K-Nearest Neighbors), Orange yazılım programında modülü bulunan makine öğrenmesi metotlarından biridir. Temel ve yaygın makine öğrenmesi yöntemlerinden biri olan bu yöntem ayrıca basitliğiyle de ön plana çıkar. Veri sınıflandırma ve regresyon problemlerinde kullanılabilir. Tembel (lazy learner) bir öğrenme algoritmasıdır. Yani eğitim için kullanılacak veri örnekleri genelleştirilmeden sadece gerekli olduğunda işleme alınır. "K" burada komşuluk sayısı olarak tanımlanır. Çalışma mantığı şu şekildedir. Veri seti çok boyutlu bir uzay olarak düşünülürse eğer, noktanın çevresindeki en yakın K adet komşusunun ortalaması alınır ve bu şekilde

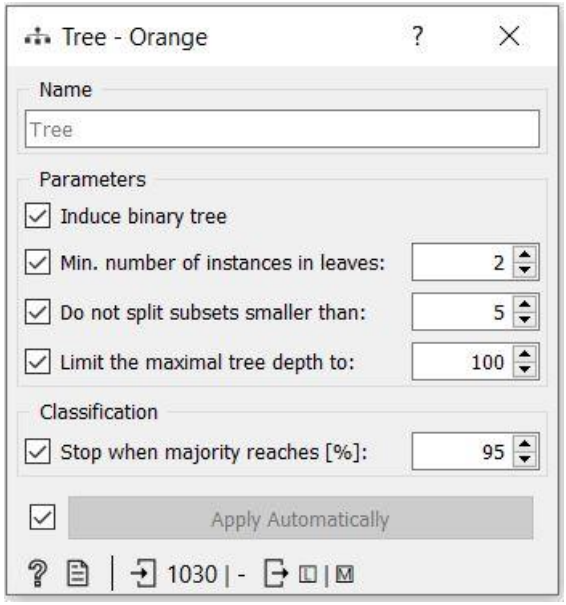
hangi sınıfta olacakları belirlenir. Ağırlık belirlenen türlerinde yakın komşu uzak komşuya göre daha fazla sonuca katkı sağlar (Imandoust and Bolandraftar 2013). Bu tahmin modeli ve bundan sonra açıklanacak olanlarda, Orange yazılımında model için neredeyse tüm parametreler denenmiş ve en verim alınan parametreler belirlenmiştir. Bu çalışmada, K-En Yakın Komşu yöntemi için "K" komşuluk sayısı 3 olarak tayin edilmiştir (Şekil 3.1).



Şekil 3.1 K-En Yakın Komşu modeli için tayin edilen parametreler

3.3.2 Karar ağaçları

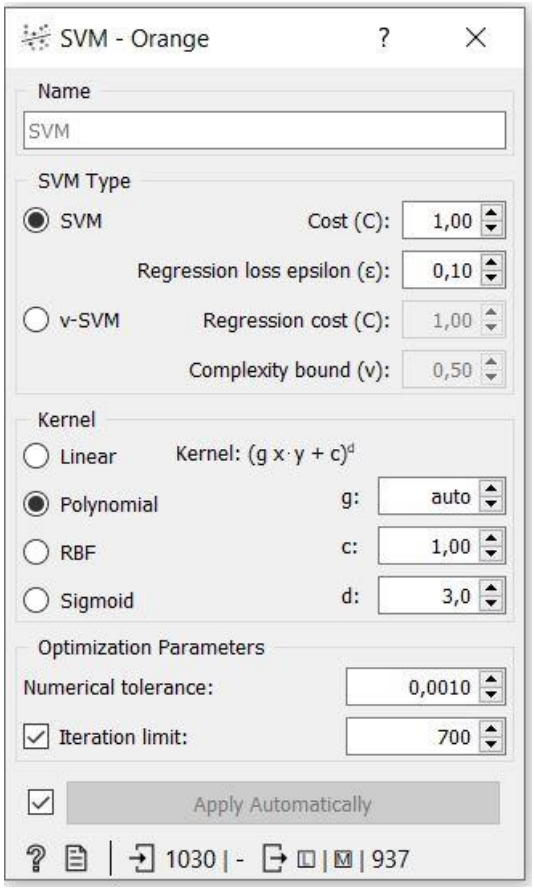
Karar Ağaçları (Decision Trees), veri sınıflandırma ve tahmin için yaygın olarak kullanılan ağaç tabanlı makine öğrenmesi metodudur. Hesap adımlarını takip etmenin imkânsız olduğu Yapay Sinir Ağları ve benzer yöntemlerin aksine kolayca anlaşılabilen ve yorumlanabilen bir metottur. İşlem maliyeti çok düşüktür. Verilerin belirli kurallara göre ağaç şeklinde dallara ayrılarak eğitilmesi ve yeni gelen verilerin bu kalıba göre değerlendirilmesi sistemine dayanır (Çalış vd. 2014). Yapılan çalışma kapsamında Orange yazılımında kullanılan karar ağacı modeli için belirlenen parametreler Şekil 3.2'de gösterilmektedir.



Şekil 3.2 Karar ağaçları modelleri için Orange yazılımında tayin edilen parametreler

3.3.3 Destek vektör makineleri

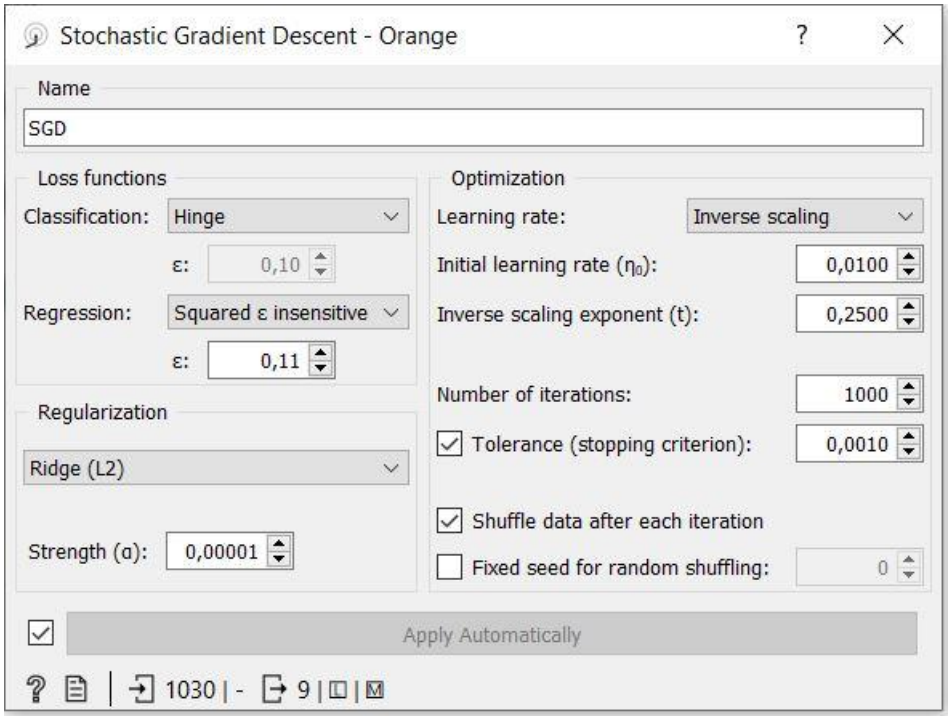
Destek Vektör Makineleri (Support Vector Machine), 90'lı yıllarda Bell Laboratuvarlarında geliştirilmiş ve hızla yaygın olarak kullanılan bir yöntem haline gelmiştir. İşleyiş şekli şu şekildedir. Örneklerin uzayda noktalar halinde gösterildiği düşünülür ve bir takım haritalama yaklaşımlarıyla bir düzlem oluşturulur. Yöntem noktaların birbirinden ayrılabildiği en iyi hiper düzlemi bulma amacını taşır. Hiper düzlem aslında doğrusal veya doğrusal olmayan bir fonksiyondur (Kartal 2020). Bu metot için seçilen parametreler Şekil 3.3'de belirtilmektedir.



Şekil 3.3 Destek Vektör Makineleri modeli parametreleri

3.3.4 Olasılıksal dereceli azalma

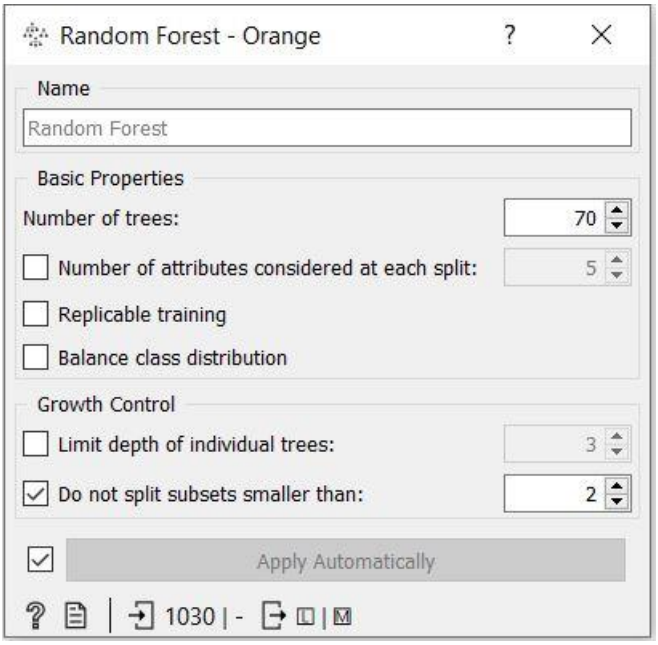
Orange yazılımında bir diğer makine öğrenmesi yöntemi olarak sunulan Olasılıksal Dereceli Azalma (Stochastic Gradient Descent), Derceli Azalma yönteminin farkı bir türüdür. Dereceli Azalma metodu sık kullanılan bir optimizasyon metodudur. Yani yinelemeli bir şekilde işlem yaparak fonksiyonun minimum değerine ulaşmayı hedefler (Deisenroth *et al*. 2020) Orange yazılımı için belirlenen Olasılıksal Dereceli Azalma parametreleri Şekil 3.4'te gösterilmektedir.



Şekil 3.4 Olasılıksal Dereceli Azalma modeli parametreleri

3.3.5 Rassal orman

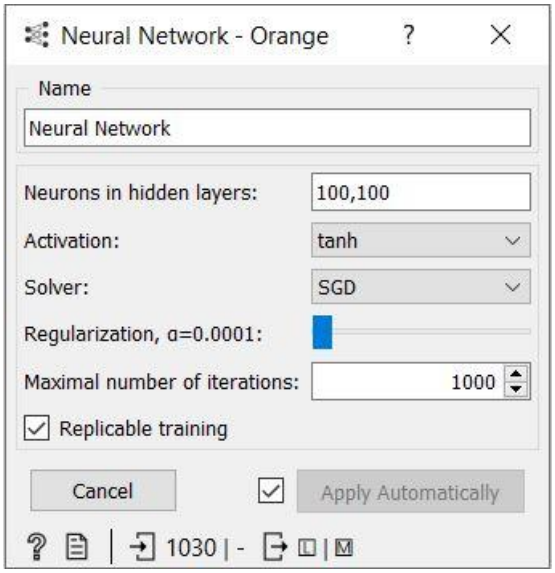
Rassal orman (Random Forest), birçok karar ağacının birleşmesi ile oluşan bir yöntemdir. Özünde bir topluluk öğrenmesi (ensemble learning) yöntemidir. Veri setiyle farklı işlem yapan birçok karar ağacı nihai sonuca ulaşmak için oylama yapar ve çoğunluk oyuyla karar verilir. Hem veri sınıflandırma ve hem de regresyon için başarılı sonuçlar verebilir (Kalaycı 2018). Bu yöntem için belirtilen parametreler Şekil 3.5'te gösterilmektedir.



Şekil 3.5 Rassal Orman modeli parametreleri

3.3.6 Yapay sinir ağları

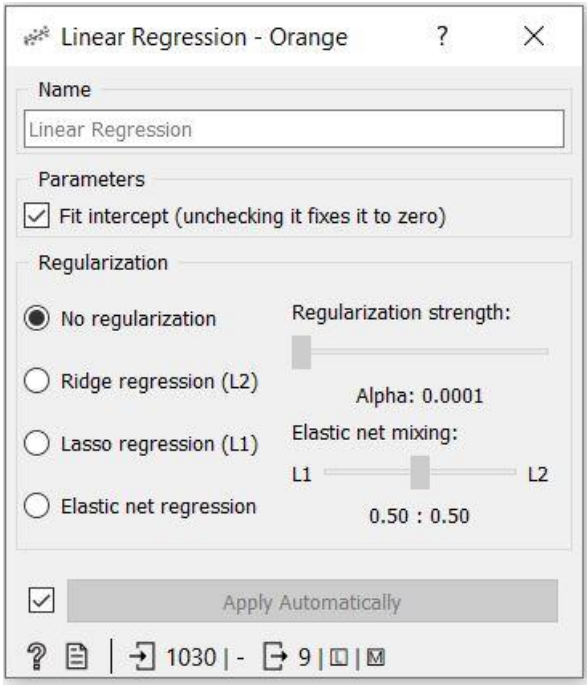
Bilgi toplama, öğrenme, anımsama, genelleme yapma ve tüm bunlardan sonra yeni tahminler üretme insan beyninin önemli niteliklerindendir. Yapay Sinir Ağları, bu özelliklerden esinlenerek oluşturulmuş bir makine öğrenmesi yöntemidir (Öztürk ve Şahin 2018). Yapay sinir ağları; giriş katmanı, gizli katmanlar ve çıkış katmanı olmak üzere katmanlara ayrılır. Bu çalışmada kullanılan veri setinde 8 adet giriş parametresi ve hedef parametrenin temsil ettiği 1 adet çıkış parametresi mevcuttur. Bu sebeple oluşturulan yapay sinir ağları tahmin modelinin giriş katmanında 8 adet, çıkış katmanında ise 1 adet nöron bulunmaktadır. Ayrıca modelde 2 adet gizli katman oluşturulmuştur. Bu gizli katmanlar, her birinde 100'er adet nöron olacak şekilde tasarlanmıştır. Kullanılan diğer parametreler Şekil 3.6 'da gösterilmektedir.



Şekil 3.6 Yapay Sinir Ağları parametreleri

3.3.7 Doğrusal regresyon

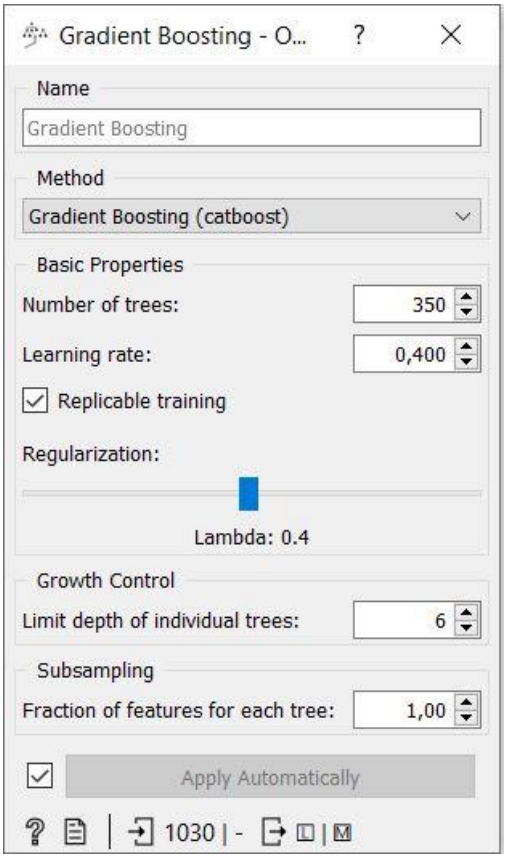
Doğrusal Regresyon (Linear Regression) yöntemi iki değişken arasındaki ilişkiyi belirlemek için yapılan işlemler bütünüdür. Verileri iki boyutlu bir uzayda nokta olarak düşünürsek, tahmin değerleri olarak tespit edilecek çizgi ile gerçek değerleri temsil eden noktalar arasında mesafe olarak minimum fark olması amaçlanır (Namlı vd. 2019). Orange yazılımında kullanmak için tespit ettiğimiz parametreler Şekil 3.7'de gösterilmektedir.



Şekil 3.7 Doğrusal Regresyon parametreleri

3.3.8 Gradyan artırma

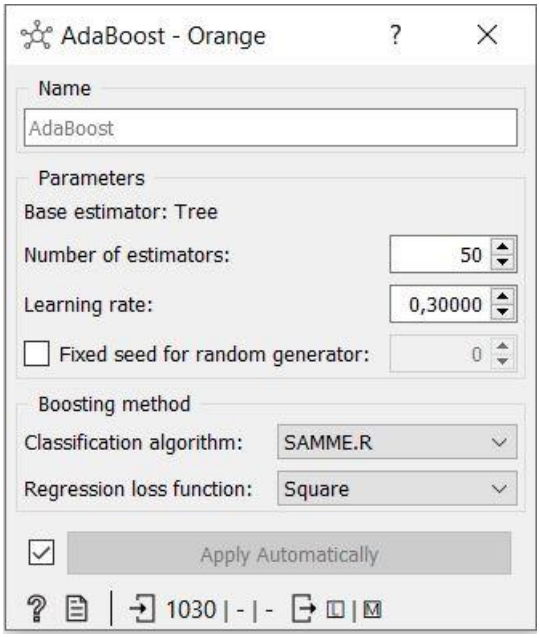
Genellikle karar ağaçlarından oluşan bir topluluk öğrenmesi (ensemble learning) metodudur. Karar ağaçları gibi yetersiz algoritmaların eksik yönlerini telafi ederek veri sınıflandırma ve regresyon problemleri için yüksek doğruluklu sonuçlar verir (Keleş vd. 2021). Bu çalışmada Orange yazılımında da sunulan Yandex tarafından geliştirilen açık kaynaklı bir Gradyan Artırma (Gradient Boosting) metodu kütüphanesi olan "CatBoost" kullanılmıştır (Şekil 3.8)



Şekil 3.8 Gradyan Artırma parametreleri

3.3.9 Adaptif yükseltme

Yapı olarak Gradyan Artırma yöntemi ile benzerlik gösterir. Adaptif Yükseltme (Adaptive Boosting) aynı Rassal Orman ve Gradyan Artırma yöntemleri gibi bir topluluk öğrenme (ensemble learning) yöntemidir. Uygulama adımları şu şekilde ilerler. Oluşturulan birçok zayıf öğreniciler veri setiyle eğitilir. Daha sonra her bir öğrenici diğerinin yanlışına odaklanır ve genel başarıya katkı sağlamaya çalışır. Yani yöntemde zayıf modelleri güçlendirme amaçlanmaktadır (Kalaycı 2018). Tayin edilen parametreler Şekil 3.9'de gösterilmektedir.



Şekil 3.9 Adaptif Yükseltme parametreleri

3.4 Diğer Tanımlar

3.4.1 İstifleme yöntemi

İstifleme (Stacking), makine öğrenmesi yaklaşımlarını birleştirilerek bir topluluk öğrenmesi (ensemble learning) geliştiren bir yöntemdir. Bu yöntemin diğer topluluk öğrenmesi yöntemlerinden en belirgin farkı, makine öğrenmesi modellerini birleştirirken sonuçları oylama, artırma gibi metotlar yerine başka bir makine öğrenmesi yöntemi kullanarak bulmasıdır (Gündüz 2020).

3.4.2 K katmanlı çapraz doğrulama

Veri seti üzerinden tahminler yapılırken, eğitim ve test verisinin dağılımı farklı sonuçlar alınmasına sebep olmaktadır. Üstelik aynı oranlarda eğitim ve test verisi olmasına rağmen yine farklı sonuçlar alınabilmektedir. Bu sorun K katmanlı Çapraz Doğrulama (K-fold Cross Validation) yöntemi ile çözülmektedir. Çalışma mantığına değinecek olursak, veri seti K sayısı kadar eşit parçalara ayrılır. Her K parçası sırasıyla test verisi geri kalan kısım ise eğitim verisi olarak kullanılır. Tüm bu kombinasyonlar için yapılan

tahminlerde çıkan sonuçların ortalaması alınarak optimum tahmin değerleri bulunmaya çalışılır. Görüldüğü üzere işlem maliyeti çok yüksektir. Fakat veri setini objektif olarak değerlendirme konusunda faydalı bir yöntemdir (Güvenoğlu ve Bağırgan 2019).

3.4.3 Performans ölçüm yöntemleri

Yapılan çalışmada, Orange yazılımından da faydalanarak bir takım performans ölçüm değerleri elde edilmiştir. Bu değerler tahmin kalitesi hakkında bilgi vermekte ve karşılaştırma imkânı sağlamaktadır. Literatürde de çok yaygın kullanılan bu ölçüm değerleri; Ortalama Karesel Hata (Mean Squared Error (MSE)), Ortalama Hata Karesinin Kökü (Root Mean Square Error (RMSE)), Ortalama Mutlak Hata (Mean Absolute Error (MAE)) ve Korelasyon Katsayısı (Correlation Coefficient (R2)) formülasyonu ile elde edilmektedir. Ortalama Karesel Hata, Ortalama Mutlak Hata ve Ortalama Hata Karesinin Kökü 0 değerine, Korelasyon Katsayısı 1 değerine ne kadar yaklaşırsa tahmin değeri gerçek değere o kadar yaklaşacaktır. Denklem (3.1), Denklem (3.2), Denklem (3.3) ve Denklem (3.4)'te performans ölçüm değerlerinin hesap adımları belirtilmektedir.

$$\text{Ortalama Karesel Hata(MSE)} = \frac{1}{N}\sum_{i=1}^{N}(y_{i-}\hat{y})^2 \tag{3.1}$$

$$\text{Ort. Hata Karesinin Kökü (RMSE} = \sqrt{\text{MSE}}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{i-}\hat{y})^2} \tag{3.2}$$

$$\text{Ortalama Mutlak Hata (MAE)} = \frac{1}{N}\sum_{i=1}^{N}|y_{i-}\hat{y}| \tag{3.3}$$

$$\text{Korelasyon Katsayısı } (R^2) = 1 - \frac{\sum(y_{i-}\hat{y})^2}{\sum(y_{i-}\bar{y})^2} \tag{3.4}$$

$\hat{y}$ =tahmin edilen değer

$\bar{y} = y$ nin ortalama değeri

3.5 Yöntem

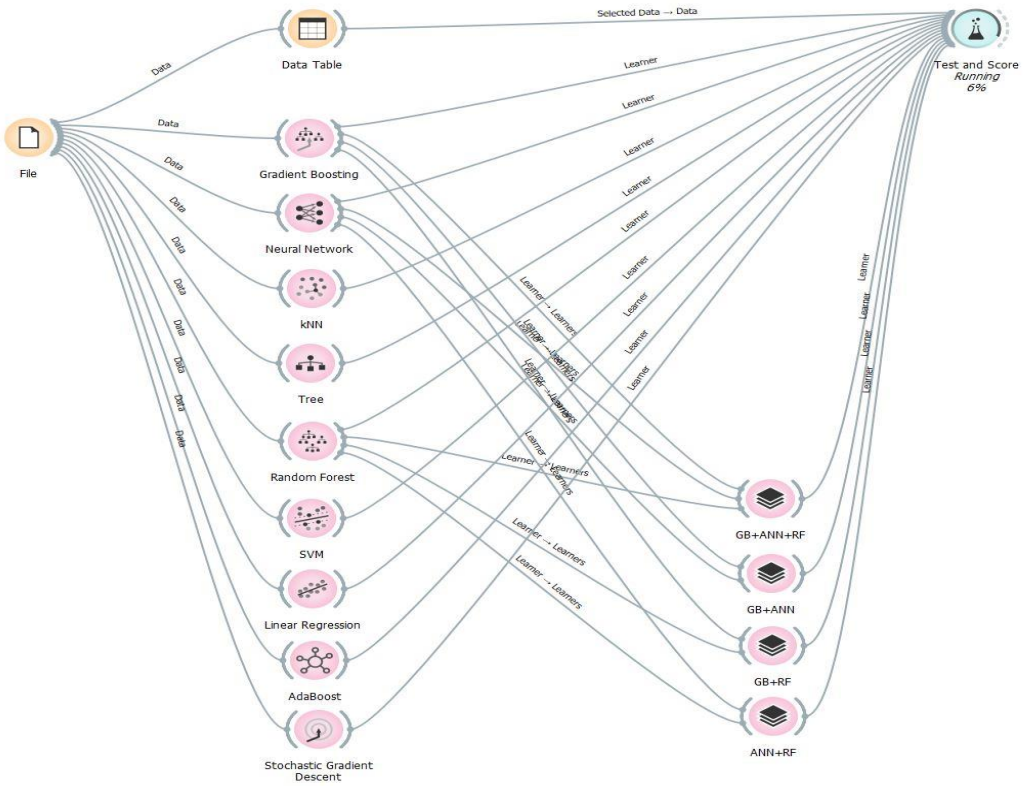
Genel itibari ile süreç şu şekilde ilerlemektedir. İlk önce ön işleme yöntemine tabi tutulmamış ham veri setine Özellik Seçimi (Feature Selection) tekniği uygulanarak verideki parametrelerinin sonuca hangi oranda etkili olduğu gözlemlenmeye çalışılmıştır. Özellik Seçimi tekniği veri setinden parametre çıkarma işlemi gibi aktif olarak kullanılmasa da uygulanacak ön işlemenin (preprocessing) hangi parametrelere uygulanacağı konusunda fikir vermiştir. Daha sonra beton basınç dayanımı tahmin sürecine geçilmiştir. Bu süreç, nihai sonuçları elde edene kadar birinin sonucunun bir sonrakine etki ettiği 4 adet aşamadan oluşmaktadır. Bu aşamaların her birinde ilk olarak veri setlerine 10 Katmanlı Çapraz Doğrulama (Cross Validation) tekniği uygulanmıştır. Daha sonra aynı aşama içerisinde veri setini %90 eğitim, %10 test kısımlarına rastgele ayırarak tahminler tekrarlanmış ve sonuçlar gözlemlenmiştir. Bu çalışmada dikkate aldığımız öncelikli sonuçlar 10 Katmanlı Çapraz Doğrulama (Cross Validation) tekniği yardımıyla elde edilen sonuçlardır. Çünkü olası aşırı öğrenme (overfitting) sorunları karşısında daha güvenli bir yoldur. %90 eğitim, %10 test kısımlarına rastgele ayrılmış veri setleriyle yapılan tahminler gözlem amacıyla yapılmıştır.

3.5.1 Aşama 1

Henüz hiçbir işlem uygulanmamış ham veri seti kullanılarak (Çizelge 3.2) 9 adet tahmin modeli ile tahminler yapılmıştır. Daha sonra en iyi sonuçları veren modeller ile İstifleme (Stacking) yöntemi ile topluluk (ensemble) modelleri oluşturulmuştur (Şekil 3.10).

Çizelge 3.2 1030 örnekli beton basınç dayanımı ham veri setinden örnek bir kısım

Çimento (Kg/M3)	Yüksek Fırın Cürufu (Kg/M3)	Uçucu Kül (Kg/M3)	Su (Kg/M3)	Süper A. (Kg/M3)	Kaba Agrega (Kg/M3)	İnce Agrega (Kg/M3)	Numune Yaşı (Gün)	Beton Basınç Dayanımı (Mpa)
540.0	0.0	0.0	162.0	2.5	1040.0	676.0	28	79.99
540.0	0.0	0.0	162.0	2.5	1055.0	676.0	28	61.89
332.5	142.5	0.0	228.0	0.0	932.0	594.0	270	40.27
332.5	142.5	0.0	228.0	0.0	932.0	594.0	365	41.05
198.6	132.4	0.0	192.0	0.0	978.4	825.5	360	44.30
266.0	114.0	0.0	228.0	0.0	932.0	670.0	90	47.03
380.0	95.0	0.0	228.0	0.0	932.0	594.0	365	43.70
380.0	95.0	0.0	228.0	0.0	932.0	594.0	28	36.45
266.0	114.0	0.0	228.0	0.0	932.0	670.0	28	45.85
475.0	0.0	0.0	228.0	0.0	932.0	594.0	28	39.29
198.6	132.4	0.0	192.0	0.0	978.4	825.5	90	38.07
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.



Şekil 3.10 Orange yazılımında oluşturulan tahmin modelleri şeması

3.5.2 Aşama 2

“Süper Akışkanlaştırıcı” parametresi değerlerinin küçükten büyüğe sıralanması baz alınarak Aşama 1'de kullanılan ham veri seti tekrar sıralanmıştır. Daha sonra tahminler yapılarak performans ölçüm değerleri elde edilmiştir. İşlem yükünü ve süreyi azaltmak için bu ve bundan sonraki aşamalarda, Aşama 1'deki en performanslı iki model olan Gradyan Artırma ve Yapay Sinir Ağları modelleriyle tahminler yapılmıştır.

3.5.3 Aşama 3

Bu ve bir sonraki aşamada Çizelge 3.3‘te açıklanan birtakım yöntemler uygulamaya konulmuştur. Aşama 2'de kullanılan veri setindeki “Yüksek Fırın Cürufu”, “Uçucu Kül” ve “Süper Akışkanlaştırıcı” parametrelerine Yöntem 1, Yöntem 2 ve Yöntem 3 olarak adlandırılan özgün ön işleme (preprocessing) teknikleri tüm kombinasyonlarıyla uygulanarak 64 farklı veri elde edilmiştir (Ek 1). Bu verilerin tamamı ile yapılan tahminlerle en iyi sonuç veren veri seti ve makine öğrenmesi algoritması tespit edilmiştir.

Çizelge 3.3 Ön işleme yöntemleri açıklamaları

Yöntem 0: Parametre herhangi bir işleme tabi tutulmamaktadır
Yöntem 1: Parametre değeri sıfır ise “yok”, sıfırdan büyükse “var” değerini alır.
Yöntem 2: Parametre değeri sıfır ise “0”, sıfırdan büyükse eşit parçalara bölünmüş maks.-min. aralığındaki konumuna göre 1'den 3’e kadar bir değer alır.
Yöntem 3: Parametre değeri sıfır ise “0”, sıfırdan büyükse 1’den 5’e kadar maks. -min. aralığındaki konumuna göre bir değer alır.
Yöntem 4: Parametre herhangi bir işleme tabi tutulmamaktadır
Yöntem 5: Parametre değerine logaritmik bir formülle ön işleme yöntemi uygulanır.

3.5.4 Aşama 4

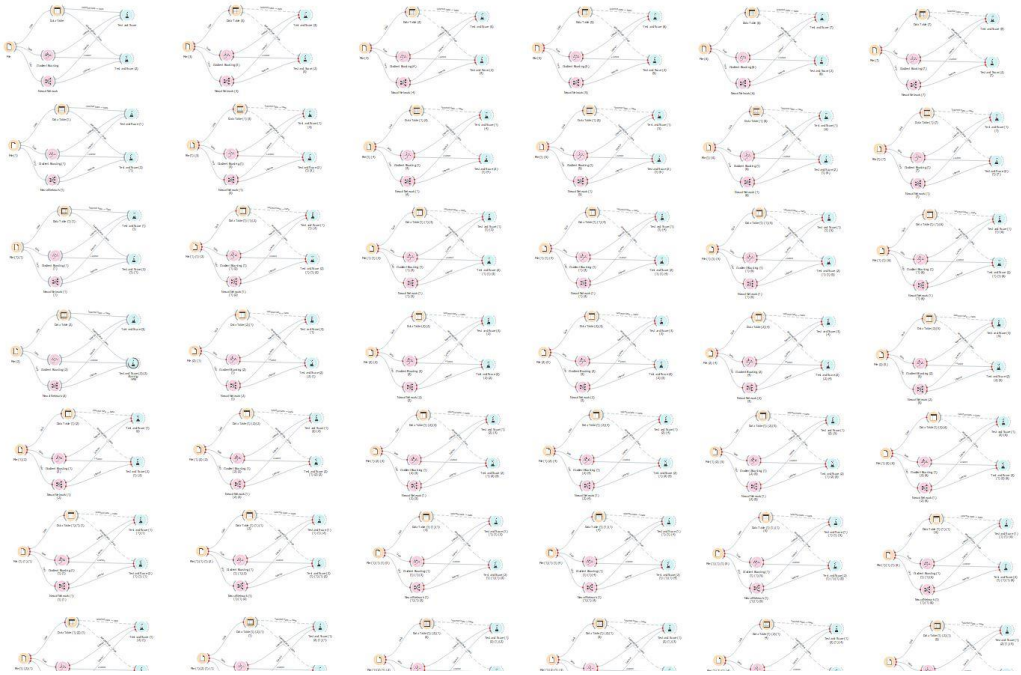
Aşama 3'ten farklı olarak “Numune Yaşı” parametresine logaritmik bir formülle ön işleme tekniği uygulanmıştır. Grafiksel benzerliği nedeniyle “Numune yaşı” parametresi değerlerinin 2 tabanında logaritması alınmış ve 1’den büyük değerler elde edebilmek

için sonuca 4 sayısı eklenmiştir. Son olarak sonuç 1 ondalık basamağa yuvarlanmıştır. Bu formül, 14 farklı "Numune Yaşı" parametresi değerine uygulanmıştır (Çizelge 3.4).

Çizelge 3.4 "Numune Yaşı" parametresine uygulanan formül sonuçları

Numune Yaşı	Yeni Değer
1	1.2
3	2.8
7	4
14	5
28	6
56	7
90	7.7
91	7.7
100	7.8
120	8.0
180	8.7
270	9.3
360	9.7
365	9.7

Yöntem 5 olarak adlandırılan bu yöntemle beraber Aşama 2'de kullanılan veri setine; Yöntem 1, Yöntem 2 ve Yöntem 3 de yine tüm kombinasyonlarıyla uygulanmış (Şekil 3.11) ve yine 64 farklı veri seti elde edilmiştir (Ek 1). Çizelge 3.5'te "Yüksek Fırın Cürufu" parametresine Yöntem 1, "Uçucu Kül" parametresine Yöntem 1, "Süper Akışkanlaştırıcı" parametresine Yöntem 3, "Numune Yaşı" parametresine Yöntem 5 uygulanarak üretilen "Veri_113_5" isimli örnek bir veri seti gösterilmektedir. Üretilen veri setlerinin tamamı ile ilk önce iki adet makine öğrenmesi algoritmasıyla (Gradyan Artırma ve Yapay Sinir Ağları) tahminler yapılmış ve en iyi sonucu veren veri seti tespit edilmiştir (Ek 2). Daha sonra en iyi sonucu veren veri seti kullanılarak aynı Aşama 1 de olduğu gibi 9 adet bilinen modellerle ve 3 adet İstifleme (Stacking) yöntemi ile oluşturulan topluluk modelleri ile nihai sonuca ulaşmak için tahminler yapılmıştır. Son olarak bu çalışma kapsamında Orange yazılımı altyapısını kullanarak çalışan basit bir beton basınç dayanımı yazılımı geliştirilmiştir. Geliştirilen bu yazılım aracılığıyla, üretilen veri setlerinden en performanslı olanıyla eğitilen modeller ile bir takım bireysel örneklerin tahmin sonuçları elde edilmiştir.



Şekil 3.11 Orange yazılımında bütün veri seti kombinasyonlarını denemek için oluşturulan örnek bir şema

Çizelge 3.5 Üretilen örnek bir veri setinden kısımlar

Çimento (Kg/M3)	Yüksek Fırın Cürufu	Uçucu Kül	Su (Kg/M3)	Süper A.	Kaba Agrega (Kg/M3)	İnce Agrega (Kg/M3)	Numune Yaşı	Beton Basınç Dayanımı (Mpa)
332.5	var	yok	228.0	0.0	932.0	594.0	9.30	40.27
332.5	var	yok	228.0	0.0	932.0	594.0	9.70	41.05
198.6	var	yok	192.0	0.0	978.4	825.5	9.70	44.30
266.0	var	yok	228.0	0.0	932.0	670.0	7.70	47.03
380.0	var	yok	228.0	0.0	932.0	594.0	9.70	43.70
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
425.0	var	yok	153.5	3.0	852.1	887.1	7.70	65.20
425.0	var	yok	153.5	3.0	852.1	887.1	7.70	65.20
333.0	var	var	167.0	3.0	996.0	652.0	6.00	47.28
143.6	yok	var	158.4	3.0	942.7	844.5	6.00	15.42
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

4. BULGULAR

Materyal ve yöntem bölümünde, beton basınç dayanımı tahmin modelleri ve mevcut veriye uygulanan ön işleme (preprocessing) yöntemlerinin uygulanmasındaki süreç detaylı olarak belirtilmiştir. Bu bölümde aynı sıralama ile uygulanan aşamaların bulgularına yer verilecektir.

4.1 Veri Setine Özellik Seçimi Yöntemi Uygulanması

Hiçbir işleme tabi tutulmamış ham veri setine iki farklı Özellik Seçimi (Feature Selection) yöntemi uygulanmıştır. Tek Değişkenli Regresyon (Univariate Regression) algoritması uygulandığı durumdaki elde edilen sonuçlarda (Şekil 4.1), veri setindeki en etkili parametrenin "çimento", en etkisiz parametrelerin ise sırasıyla "Uçucu Kül", "Yüksek Fırın Cürufu" ve "Kaba Agrega" olduğu görülmektedir.

		#	Uni...eq.
1	N Cement (component 1)(kg in a m^3 mixture)		338.726
2	N Superplasticizer (component 5)(kg in a m^3 mixture)		159.109
3	N Age (day)		124.673
4	N Water (component 4)(kg in a m^3 mixture)		94.119
5	N Fine Aggregate (component 7)(kg in a m^3 mixture)		29.583
6	N Coarse Aggregate (component 6)(kg in a m^3 mixture)		28.745
7	N Blast Furnace Slag (component 2)(kg in a m^3 mixture)		19.033
8	N Fly Ash (component 3)(kg in a m^3 mixture)		11.627

Şekil 4.1 Tek Değişkenli Regresyon uygulama sonuçları

RReliefF metodu kullanılan durumda ise (Şekil 4.2), "Numune Yaşı" veri setindeki en etkili parametredir. Etkisiz parametreler ise sırasıyla "Uçucu Kül", "Yüksek Fırın Cürufu" ve "Süper Akışkanlaştırıcı" olarak belirlenmektedir.

		#	RReliefF
1	N Age (day)		0.084
2	N Cement (component 1)(kg in a m^3 mixture)		0.081
3	N Water (component 4)(kg in a m^3 mixture)		0.072
4	N Coarse Aggregate (component 6)(kg in a m^3 mixture)		0.071
5	N Fine Aggregate (component 7)(kg in a m^3 mixture)		0.067
6	N Superplasticizer (component 5)(kg in a m^3 mixture)		0.054
7	N Blast Furnace Slag (component 2)(kg in a m^3 mixture)		0.052
8	N Fly Ash (component 3)(kg in a m^3 mixture)		0.026

Şekil 4.2 RReliefF metodu uygulama sonuçları

4.2 Aşama 1'de Elde Edilen Sonuçlar

Herhangi bir ön işleme (preprocessing) tabi tutulmamış ham veri setiyle 10 Katmalı Çapraz Doğrulama (Cross Validation) yöntemi uygulanarak yapılan tahminler için performans ölçüm değerleri Çizelge 4.1'de gösterilmektedir. Kullanılan 9 adet tahmin modelinden sırasıyla Gradyan Artırma ve Yapay Sinir Ağları modelleri en verimli sonuçları vermektedir. En verim alınan 3 modelin farklı kombinasyonlarda istifleme (Stacking) yöntemiyle birleştirilmesi sonucu oluşan modellerden en iyi performans gösteren Gradyan Artırma ve Yapay Sinir Ağları modellerinin birleştiği topluluk (ensemble) modelidir.

Aynı veri seti ve aynı modellerle yapılan, veri setinin %90 eğitim ve %10 test kısımlarına rastgele ayrılarak kullanıldığı tahminler için performans ölçüm değerleri Çizelge 4.2'de gösterilmektedir.

Çizelge 4.1 Aşama 1 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
K-En Yakın Komşu (kNN)	65.11	8.07	5.62	0.7665
Karar Ağaçları (DT)	42.30	6.50	4.34	0.8483
Destek Vektör Makineleri (SVM)	54.29	7.37	5.46	0.8053
Olasılıksal Dereceli Azalma (SGD)	110.46	10.51	8.36	0.6038
Rassal Orman (RF)	23.67	4.87	3.31	0.9151
Yapay Sinir Ağları (ANN)	**19.30**	**4.39**	**2.82**	**0.9308**
Doğrusal Regresyon (LR)	109.77	10.48	8.30	0.6063
Gradyan Artırma (GB)	**15.91**	**3.99**	**2.43**	**0.9429**
Adaptif Yükseltme (AB)	25.39	5.04	3.26	0.9089
GA+RO (GB+RF)	15.64	3.95	2.45	0.9439
GA + YSA (GB+ANN)	**15.03**	**3.88**	**2.38**	**0.9461**
YSA+RO (ANN+RF)	17.64	4.20	2.75	0.9367
GA + YSA + RO (GB+ANN+RF)	15.11	3.89	2.41	0.9458
En çok verim alınan iki model ve İstifleme (Stacking) metodu kullanılarak oluşturulan en verimli topluluk modeli kalın punto ile yazılmıştır.				

Çizelge 4.2 Aşama 1 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
K-En Yakın Komşu (kNN)	61.53	7.84	5.84	0.7463
Karar Ağaçları (DT)	43.78	6.62	4.75	0.8195
Destek Vektör Makineleri (SVM)	49.18	7.01	5.47	0.7972
Olasılıksal Dereceli Azalma (SGD)	106.19	10.31	8.18	0.5621
Rassal Orman (RF)	23.33	4.83	3.41	0.9038
Yapay Sinir Ağları (ANN)	**15.12**	**3.89**	**2.80**	**0.9376**
Doğrusal Regresyon (LR)	105.79	10.29	8.13	0.5638
Gradyan Artırma (GB)	**11.04**	**3.32**	**2.25**	**0.9545**
Adaptif Yükseltme (AB)	23.52	4.85	3.32	0.9030
GA+RO (GB+RF)	12.33	3.51	2.41	0.9492
GA + YSA (GB+ANN)	**10.21**	**3.19**	**2.19**	**0.9579**
YSA+RO (ANN+RF)	15.34	3.92	2.72	0.9368
GA + YSA + RO (GB+ANN+RF)	10.98	3.31	2.27	0.9547
En çok verim alınan iki model ve İstifleme (Stacking) metodu kullanılarak oluşturulan en verimli topluluk modeli kalın punto ile yazılmıştır.				

4.3 Aşama 2'de Elde Edilen Sonuçlar

Bu aşamada kullanılan veri seti, Aşama 1 de kullanılan ham veri setinin "Süper Akışkanlaştırıcı" parametresi değerlerinin küçükten büyüğe sıralanması baz alınarak tekrar sıralanmasıyla oluşturulmuştur. Aşama 1 de kullanılan kalıplaşmış modellerden en iyi performans gösterenler Gradyan Artırma ve Yapay Sinir Ağları modelleridir. Geçmiş konularda zaman ve işlemci tasarrufu işin bu ve bundan sonraki aşamalarda Aşama 1'in en iyi iki modeli ile devam edileceği belirtilmişti. Bu nedenle bu aşamada Gradyan Artırma ve Yapay Sinir Ağları modelleriyle tahminler yapılmıştır. 10 katmalı çapraz doğrulama (Cross Validation) yöntemi uygulanarak yapılan tahminler için performans ölçüm değerleri Çizelge 4.3'de gösterilmektedir.

Çizelge 4.3 Aşama 2 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
Yapay Sinir Ağları (ANN)	17.54	4.19	2.71	0.9371
Gradyan Artırma (GB)	14.63	3.83	2.32	0.9475

Aynı veri setinin %90 eğitim ve %10 test kısımlarına rastgele ayrılarak kullanıldığı tahminler için performans ölçüm değerleri Çizelge 4.4'de gösterilmektedir.

Çizelge 4.4 Aşama 2 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
Yapay Sinir Ağları (ANN)	10.85	3.29	2.35	0.9517
Gradyan Artırma (GB)	6.48	2.55	1.76	0.9712

4.4 Aşama 3'te Elde Edilen Sonuçlar

Geçmiş bölümlerde oluşturma süreci açıklanan, Aşama 2 de kullanılan veri seti baz alınarak üretilen 64 faklı veri seti ile tahminler yapılmıştır. "Veri_000_4" ten "Veri_333_4" e kadar (Ek 2) yapılan tüm tahminlerden en iyi performansı gösteren 2 numaralı, "Veri_001_4" olarak adlandırılan veri setidir. Çizelge 4.5'te "Veri_001_4" isimli veri seti ile 10 Katmalı Çapraz Doğrulama (Cross Validation) yöntemi uygulanarak yapılan tahminler için performans ölçüm değerleri gösterilmektedir.

Çizelge 4.5 Aşama 3 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
Yapay Sinir Ağları (ANN)	18.86	4.34	2.91	0.9324
Gradyan Artırma (GB)	13.83	3.72	2.31	0.9504

"Veri_001_4" isimli veri setinin %90 eğitim ve %10 test kısımlarına rastgele ayrılarak kullanıldığı tahminler için performans ölçüm değerleri Çizelge 4.6'da gösterilmektedir.

Çizelge 4.6 Aşama 3 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
Yapay Sinir Ağları (ANN)	11.95	3.46	2.48	0.9469
Gradyan Artırma (GB)	8.53	2.92	2.01	0.9621

4.5 Aşama 4'te Elde Edilen Sonuçlar

Bu aşamada da Aşama 3'te olduğu gibi üretilen 64 faklı veri seti ile tahminler yapılmıştır. Kullanılan veri setleri, Aşama 3'te kullanılan veri setleriyle bir noktada farklılık göstermektedir. Bu farklılık "Numune Yaşı" parametresine ön işleme amacıyla

uygulanan logaritmik bir formüldür. “Veri_000_5” ten “Veri_333_5” e kadar (Ek 2) yapılan tüm tahminlerden en iyi performansı gösteren 112 numaralı, “Veri_233_5” olarak adlandırılan veri setidir. Çizelge 4.7’de “Veri_233_5” isimli veri seti ile 10 Katmalı Çapraz Doğrulama (Cross Validation) yöntemi uygulanarak yapılan tahminler için performans ölçüm değerleri gösterilmektedir.

Çizelge 4.7 Aşama 4 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
Yapay Sinir Ağları (ANN)	17.59	4.19	2.58	0.9369
Gradyan Artırma (GB)	13.25	3.64	2.29	0.9525

“Veri_233_5” isimli veri setinin %90 eğitim ve %10 test kısımlarına rastgele ayrılarak kullanıldığı tahminler için performans ölçüm değerleri Çizelge 4.8’de gösterilmektedir.

Çizelge 4.8 Aşama 4 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
Yapay Sinir Ağları (ANN)	11.00	3.32	2.28	0.9511
Gradyan Artırma (GB)	6.70	2.59	1.78	0.9702

Tüm bu yapılan tahminlerden sonra nihai sonuca ulaşmak için 112 numaralı “Veri_233_5” isimli veri setiyle Aşama 1’de olduğu gibi 9 adet bilinen model ve bunlardan en iyi sonuç verenlerle oluşturulan topluluk modelleri ile tahminler yapılmıştır. Çizelge 4.9 ve Çizelge 4.10’da 10 Katmanlı Çapraz Doğrulama yöntemi için ve diğer ayırma yöntemi için (eğitim ve test) elde edilen tahmin sonuçları belirtilmektedir.

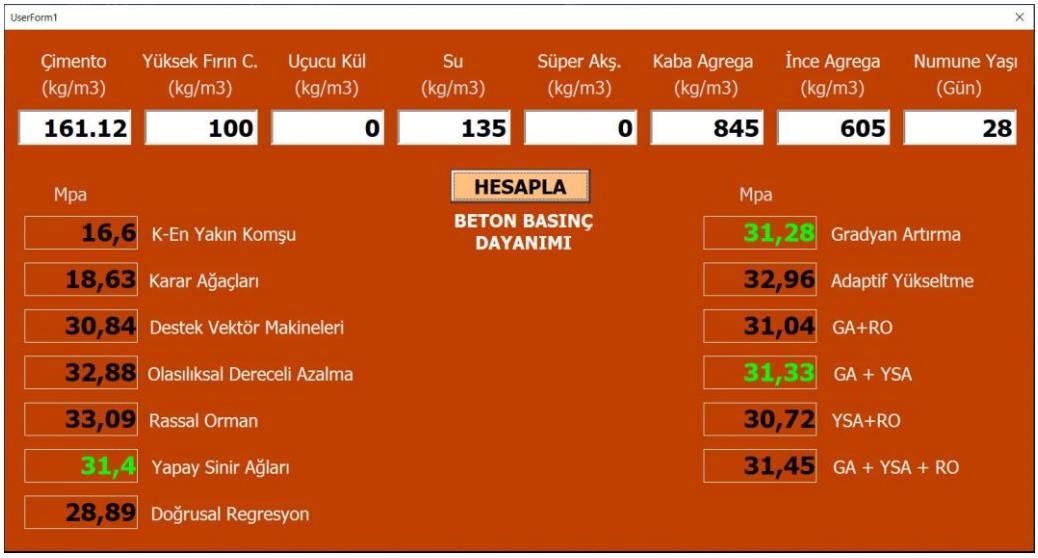
Çizelge 4.9 Aşama 4 de 10 Katmanlı Çapraz Doğrulama uygulanmış veri için tüm modellere göre tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
K-En Yakın Komşu (kNN)	26.88	5.18	3.29	0.9036
Karar Ağaçları (DT)	41.54	6.45	4.47	0.8510
Destek Vektör Makineleri (SVM)	34.44	5.87	4.28	0.8765
Olasılıksal Dereceli Azalma (SGD)	55.60	7.46	5.81	0.8006
Rassal Orman (RF)	23.31	4.83	3.30	0.9164
Yapay Sinir Ağları (ANN)	**17.59**	**4.19**	**2.58**	**0.9369**
Doğrusal Regresyon (LR)	55.63	7.46	5.80	0.8005
Gradyan Artırma (GB)	**13.25**	**3.64**	**2.29**	**0.9525**
Adaptif Yükseltme (AB)	22.18	4.71	3.02	0.9204
GA+RO (GB+RF)	13.21	3.63	2.32	0.9526
GA + YSA (GB+ANN)	**12.69**	**3.56**	**2.21**	**0.9545**
YSA+RO (ANN+RF)	16.45	4.06	2.62	0.9410
GA + YSA + RO (GB+ANN+RF)	12.78	3.58	2.24	0.9542
En çok verim alınan iki model ve İstifleme (Stacking) metodu kullanılarak oluşturulan en verimli topluluk modeli kalın punto ile yazılmıştır.				

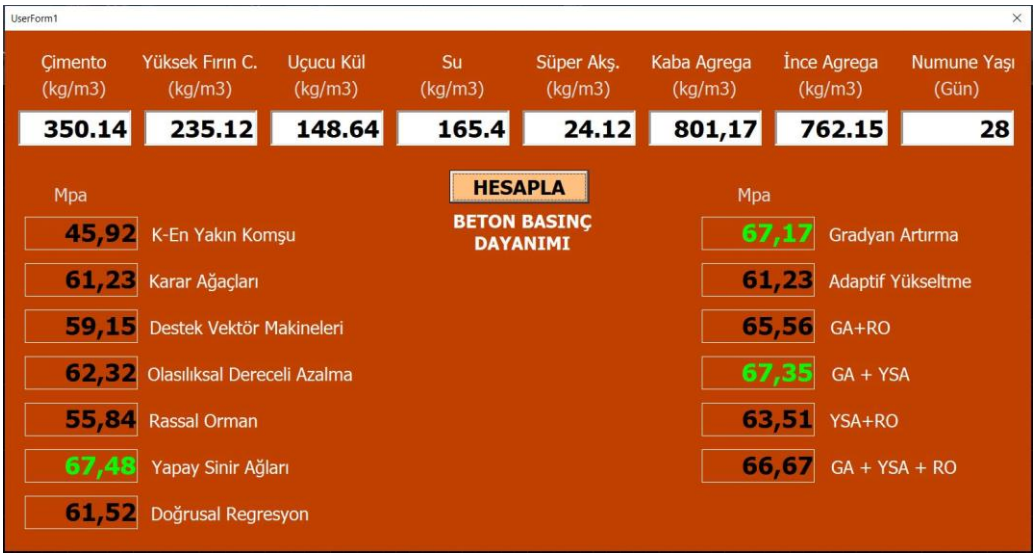
Çizelge 4.10 Aşama 4 de %90 eğitim, %10 test verisi olarak ayrılmış veri için tüm modellere göre tahmin sonuçları

Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
K-En Yakın Komşu (kNN)	16.52	4.06	2.90	0.9265
Karar Ağaçları (DT)	38.85	6.23	4.39	0.8272
Destek Vektör Makineleri (SVM)	26.08	5.11	3.92	0.8840
Olasılıksal Dereceli Azalma (SGD)	54.84	7.41	5.90	0.7560
Rassal Orman (RF)	15.78	3.97	2.84	0.9298
Yapay Sinir Ağları (ANN)	**11.00**	**3.32**	**2.28**	**0.9511**
Doğrusal Regresyon (LR)	54.83	7.40	5.92	0.7561
Gradyan Artırma (GB)	**6.70**	**2.59**	**1.78**	**0.9702**
Adaptif Yükseltme (AB)	16.19	4.02	2.72	0.9280
GA+RO (GB+RF)	6.45	2.54	1.77	0.9713
GA + YSA (GB+ANN)	6.33	2.52	1.70	0.9718
YSA+RO (ANN+RF)	9.44	3.07	2.15	0.9580
GA + YSA + RO (GB+ANN+RF)	**6.30**	**2.51**	**1.69**	**0.9720**
En çok verim alınan iki model ve İstifleme (Stacking) metodu kullanılarak oluşturulan en verimli topluluk modeli kalın punto ile yazılmıştır.				

Üretilen veri setlerinden en iyi sonuç alınan “Veri_233_5” isimli veri setidir. Bu veri seti ile eğitilen 13 adet tahmin modelini kullanan basit bir beton basınç dayanımı yazılım geliştirilmiştir. Bu yazılım ile denenen bir takım bireysel örneklerin sonuçları Şekil 4.3 ve Şekil 4.4’te gösterilmektedir. Denenen örnekler kullanılan veri setleri içinde bulunmayan ve tutarlılığı ölçmek adına rastgele girilmiş örneklerdir.



Şekil 4.3 Geliştirilen beton basınç dayanımı yazılımına girilen 1. örnek



Şekil 4.4 Geliştirilen beton basınç dayanımı yazılımına girilen 2. örnek

5. TARTIŞMA VE SONUÇ

Bu çalışmada, yapay zekanın bir alt dalı olan makine öğrenmesi teknikleriyle beton basınç dayanımı tahmini yapılması araştırılmış ve süreçteki iki önemli nokta üzerinde durulmuştur. Birincisinde aşamalar halinde özgün ön işleme (preprocessing) yöntemleri kullanarak bu yöntemlerin sürece etkisi araştırılmıştır. İkincisinde ise keşfedilen ön işleme yöntemlerini test etmek amaçlanmıştır. Bunun için 9 adet tahmin modeli ayrıca bu modellerden en fazla verim alınanlar ile 4 adet topluluk (ensemble) modeli oluşturulmuş ve hepsinin performansları karşılaştırılmıştır.

Başlangıçta uygulanan Özellik Seçimi (Feature Selection) yönteminde, denenen her iki algoritmada da "Uçucu Kül" ve "Yüksek Fırın Cürufu" parametrelerinin en etkisiz parametreler olduğu Şekil 4.1 ve Şekil 4.2'de görülmektedir. Keza RReliefF algoritması kullanılan durum baz alınırsa, "Süper Akışkanlaştırıcı" parametresinin de etkisiz bir parametre olduğu sonucuna varılmaktadır. Bu sebeple bahsi geçen parametreleri sadeleştirmek için uygulanacak bir ön işleme yönteminin gürültüyü (noise) azaltacağı ve tahminlerin verimini artıracağı değerlendirilmektedir. "Numune Yaşı" önemli bir parametredir ve diğerlerinden farklı bir ön işlemi yöntemi kullanarak tahmin veriminin artırılabileceği düşünülmüştür. Nitekim yapılan çalışmalar sonucu elde edilen bulgularda bu durumu destekler niteliktedir.

Aşama 1'den Aşama 4'e kadar, ön işleme (preprocessing) tabi tutulmamış ham veri setinden başlanılıp, özgün ön işleme yöntemleri kullanılarak sayısız veri üretilmiş ve bu verilerle beton basınç dayanımı tahminleri yapılmıştır. Her bir aşamada hem Çapraz Doğrulama (Cross Validation) uygulanmış hem de eğitim ve test verisi olarak ayrılmış veri setleriyle tahminler yapılmıştır. Bu çalışmada öncelikli olarak dikkate aldığımız sonuçlar Çapraz Doğrulama uygulanmış veri setleriyle yapılan tahmin sonuçlarıdır. Çünkü aşırı öğrenme (overfitting) riskine karşı çapraz doğrulama yöntemi tahmine güvenilirliği artırmaktadır. Modelin hiç görmediği bir veride daha iyi sonuçlar vereceği düşünülmektedir. Örneğin Aşama 2 de ham veri setinin sıralamasının değiştirilmesi aşırı öğrenmeye yatkınlığı artırabilir. Çizelge 4.3 ve Çizelge 4.4 teki alınan sonuçları incelersek eğer, Gradyan Artırma (Gradient Boosting) modelinin Aşama 1'e göre

tahmin sonuçlarının iyileştiği görülmektedir. Ama eğitim ve test verisi olarak ayrılan veri setinde (Çizelge 4.4) sonuçlar çok daha fazla iyileşmiştir. Bu durumun aşırı öğrenme (overfittingle) ile ilgili olduğu kanısına varılmıştır. Fakat ilerde bu işlem süreçleri referans alınarak oluşturulabilecek bir beton basınç dayanımı yazılımında, eğitim ve test verisi olarak ayrılmış veri seti sonuçlarının da yazılım tarafından gösterilmesinin sağlıklı olacağı düşünülmektedir. Bu kullanıcının tahmin sonucu hakkında daha iyi bir çıkarım sağlamasını kolaylaştırabilir.

Aşama 1'den sonraki süreçlerde genel itibari ile işlem yükünü ve süreyi azaltmak için Gradyan Artırma (Gradient Boosting) ve Yapay Sinir Ağları (Artificial Neural Network) modelleri ile tahminler yapılmıştır. Yapay Sinir Ağları modeline göre daha iyi sonuçlar veren Gradyan Artırma yöntemini baz alarak ifade edersek eğer, Aşama 1'den Aşama 4'e kadar her aşamada 10 katmanlı çapraz doğrulama kullanılarak yapılan tahminlerde, performans ölçüm değerlerinde sürekli iyileşmeler gözlenmektedir. Ayrıca Aşama 1'deki sonuçlar (Çizelge 4.1), Aşama 4 teki sonuçlarla (Çizelge 4.9) karşılaştırıldığında tüm modellerde performans ölçüm değerleri iyileşmiştir. Bununla birlikte; verim olarak diğer modellere göre düşük performans sergileseler de, ilk ve son değerler kıyaslandığında, K-En Yakın Komşu (K-Nearest Neighbors), Olasılıksal Dereceli Azalma (Stochastic Gradient Descent) ve Doğrusal Regresyon (Linear regression) modellerinde büyük çapta iyileşmeler görülmektedir. Sonuç güvenilirliği olarak ikinci önceliğimiz olsa da %90 eğitim, %10 test veri seti kullanılan tahminleri incelediğimizde aynı durum birkaç istisna hariç gözlenmektedir. Bu durum, araştırma kapsamında kullanılan özgün ön işleme (preprocessing) yöntemlerinin tahmin performansına olumlu katkısını kanıtlamaktadır.

Çizelge 4.9'daki nihai sonuçları incelediğimizde, bilinen ve kalıplaşmış modellerden Gradyan Artırma (Gradient Boosting) en başarılı performansı sergilemektedir. İstifleme yöntemi ile oluşturulan topluluk modellerinden ise Gradyan Artırma ve Yapay Sinir Ağları modellerinin birleştiği durumda en iyi performans ölçüm değerleri elde edilmiştir. Ve genel itibari ile topluluk modellerinin daha verimli olduğu gözlemlenmektedir. Kullandığımız makine öğrenmesi yazılımı çok detaylı ayarlar

içermese de geçmişte yapılan çalışmalarla karşılaştırıldığımızda hemen hemen denk performans ölçüm değerleri elde edilmiştir.

Çalışma kapsamındaki değerlendirmeler neticesinde şu sonuç çıkarılmıştır. Ön işleme yöntemlerinin, yapay zekanın bir alt dalı olan makine öğrenmesi konusunda çok etkili bir parametre olduğu söylenebilir. İleriki çalışmalarda bu yöntemler geliştirilerek beton basınç dayanımı tahmininde veya diğer iş kollarına ait veri setleriyle yapılan tahminlerde daha verimli sonuçlar elde edilmesi mümkündür.

KAYNAKLAR

Acuña, E. 2011. Preprocessing in data mining, In: International Encyclopedia of Statistical Science. M. Lovric (eds), Springer Berlin Heidelberg, pp. 1083-1085, Berlin.

Aydoğmus, H., Erdal, H., Karakurt, O., Namlı, E., Türkan, Y. and Erdal, H. 2015. A comparative assessment of bagging ensemble models for modeling concrete slump flow. Computers and Concrete, 16(5): 741-757.

Bilgen, G. ve Kavak, A. 2005. Yüksek fırın cürufu ve kirecin katkı olarak kil üzerindeki etkileri ve yol yapımında kullanımının araştırılması. 1. Geoteknik Sempozyumu, s. 393-408, Adana.

Biswas, R., Rai, B., Samui, P. and Roy, S. 2020. Estimating concrete compressive strength using MARS, LSSVM and GP. Engineering Journal, 24(2): 41-52.

Budak, H. 2018. Özellik seçim yöntemleri ve yeni bir yaklaşım. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 22: 21-31.

Chine, W., Chen, L., Hsu, H., Wang, T. and Chiu, C. 2010. Modeling slump of concrete using the artificial neural networks. International Conference on Artificial Intelligence and Computational Intelligence, pp. 236-239, China.

Chou, J., Chiu, C., Farfoura, M. and Altaharwa, I. 2011. Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data mining techniques. Journal of Computing in Civil Engineering, 25(3): 242-253.

Chou, J., Tsai, C., Pham, A. and Lu, Y. 2014. Machine learning in concrete strength simulations multination data analytics. Construction and Building Materials, 73: 771-780.

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

Çalış, A., Kayapınar, S. ve Çetinyokuş, T. 2014. Veri madenciliğinde karar ağacı algoritmaları ile bilgisayar ve internet güvenliği üzerine bir uygulama. Endüstri Mühendisliği Dergisi, 25 (3): 2-19.

Deisenroth, M. P., Faisal, A. A. and Ong, C. S. 2020. Mathematics for machine learning, Volume I. Cambridge University Press, 405 p., United Kingdom.

Durap, A. ve Doğan, Y. 2015. İnşaat mühendisliğinde bilişim kavramı ve veri madenciliği algoritmaları ile bir uzman sisteminin oluşturulması. XVII. Akademik Bilişim Konferansi, s.1164-1170, Eskişehir.

Erdal, H. 2015. Makine öğrenmesi yöntemlerinin inşaat sektörüne katkısı basınç dayanımı tahminlemesi. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 21(3): 109-114.

Erdal, H. I., Karakurt, O. and Namlı, E. 2013. High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform. Engineering Applications of Artificial Intelligence, 26(4): 1246-1254.

Gündüz, H. 2020. Filtre tabanlı nitelik seçimi ve topluluk öğrenme yaklaşımlarıyla borsa istanbul enerji endeksi yön tahmini. European Journal of Science and Technology, 1: 215-220.

Güvenoğlu, E. ve Bağırgan, M. 2019. Shearlet dönüşümü ve görüntü işleme teknikleri kullanılarak kot kumaşlar üzerinde gerçek zamanlı hata tespiti. El Cezeri Fen ve Mühendislik Dergisi, 6 (3): 491-502.

Imandoust, SB. and Bolandraftar, M. 2013. Application of k-nearest neighbor (KNN) approach for predicting economic events theoretical background. International Journal of Engineering Research and Applications, 3(5): 605-610.

Jha, A. K., Adhikari, S., Thapa, S., Kumar, A., Kumar, A. and Mishra, S. 2020. Evaluation of factors affecting compressive strength of concrete using machine learning. 2020 Advanced Computing and Communication Technologies for High Performance Applications Conference, pp. 70-74, India.

Kalaycı, T. 2018. Kimlik hırsızı web sitelerinin sınıflandırılması için makine öğrenmesi yöntemlerinin karşılaştırılması. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 24(5): 870-878.

Kaloop, M. R., Kumar, D., Samui, P., Hu, J. W. and Kim, D. 2020. Compressive strength prediction of high-performance concrete using gradient tree boosting machine. Construction and Building Materials, 264: 1-11.

Kaplan, G. ve Gültekin, A. 2010. Yapı sektöründe uçucu kül kullanımının çevresel ve toplumsal etkiler açısından incelenmesi. Uluslararası Sürdürülebilir Yapılar Sempozyumu, s. 59-64, Ankara.

Kartal, C. 2020. Destek vektör makineleri ile borsa endekslerinin tahmini. İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 9(2): 1394-1418.

Keleş, M., Keleş, A. E. and Kılıç, Ü. 2018. Prediction of concrete strength with data mining methods using artificial bee colony as feature selector. 2018 International Conference on Artificial Intelligence and Data Processing, pp. 91-94, Malatya.

Keleş, M., Keleş, A. ve Keleş, A. 2021. Yapay zekâ teknolojisi ile uçuş fiyatı tahmin modeli geliştirme. Turkish Studies Information Technologies and Applied Sciences, 15(4): 511-520.

Khashman, A. and Akpınar, P. 2017. Nondestructive prediction of concrete compressive strength using neural networks. Procedia Computer Science, 108: 2358-2362.

Lakshmanarao, A., Swathi, Y. and Sharma, P. S. N. 2019. Estimation of compressive strength of concrete with machine learning algorithms. Journal of Information and Computational Science, 9(10): 180-187.

Mohan, P. 2019. Speculation of compressive strength of concrete in real-time. International Journal of Recent Technology and Engineering, 7(6): 988-992.

Muliauwan, H. N., Prayogo, D., Gaby, G. and Harsono, K. 2020. Prediction of concrete compressive strength using artificial intelligence methods. Journal of Physics, 1625(1): 1-10.

Namlı, E., Ünlü, R. ve GÜL, E. 2019. Fiyat tahminlemesinde makine öğrenmesi teknikleri ve doğrusal regresyon yöntemlerinin kıyaslanması; Türkiye'de satılan ikinci el araç fiyatlarının tahminlenmesine yönelik bir vaka çalışması. Konya Journal of Engineering Sciences, 7(4): 806-821.

Özbakir, O. 2015. Agregaların fiziksel özelliklerinden yola çıkılarak beton dayanımının yapay sinir ağlarıyla kestirilmesi. Doktora Tezi, İstanbul Teknik Üniversitesi, 135 sayfa, İstanbul.

Özkan, I. A. and Altın, M. 2016. Estimating of compressive strength of concrete with artificial neural network according to concrete mixture ratio and age. International Journal of Intelligent Systems and Applications in Engineering, 4(3): 76-79.

Öztürk, K. ve Şahin, M. E. 2018. Yapay sinir ağları ve yapay zekâ'ya genel bir bakış. Takvim-i Vekayi, 6 (2): 25-36.

Pirim, H. 2006. Yapay zekâ. Yaşar Üniversitesi E-Dergisi, 1(1): 81-93.

Rasheed, A., Usman, M., Farooq, S. and Hanif, A. 2018. Effect of super-plasticizer dosages on fresh state properties and early-age strength of concrete. Materials Science and Engineering Conference, pp. 1-7, Malaysia.

Saad, S., Ishtiyaque, M. and Malik, H. 2016. Selection of most relevant input parameters using WEKA for artificial neural network based concrete compressive strength prediction model. 2016 IEEE 7th Power India International Conference, pp. 1-6, India.

Santosa, S. and Santosa, Y. 2017. Evolutionary artificial neural networks for concrete mix design modelling. International Journal Of Computer Application, 7(5): 62-70.

Silva, P., Moita, G. and Arruda, V. 2020a. Machine learning techniques to predict the compressive strength of concrete. Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería, 36(4): 1-9.

Silva, P., Moita, G. and Arruda, V. 2020b. A computational method to predict the concrete compression strength using decision trees and random forest. XLI Ibero-Latin-American Congress on Computational Methods in Engineering Conference, pp. 1-7, Brazil.

Siraj, N., Fayek, A. and Tsehayae, A. 2016. Development and optimization of artificial intelligence based concrete compressive strength predictive models. International Journal of Structural and Civil Engineering, 5(3): 156-167.

Sun, J., Zhang, J., Gu, Y., Huang, Y., Sun, Y. and Ma, G. 2019. Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression. Construction and Building Materials, 207: 440-449.

Tavares, D., Ribeiro, D., Yanagi Junior, T., Lacerda, W., Tiradentes, E., Teixeira, R. and Garcia, H. 2020. Use of artificial neural networks to predict concrete compression strength. Brazilian Journal of Development, 6(7): 42815-42826.

UCI IRVANE, 2007. Web sitesi. https://archive.ics.uci.edu/ml/datasets/. Erişim Tarihi: 01.01.2022.

Yeh, I. 1998. Modeling of strength of high performance concrete using artificial neural networks. Cement and Concrete Research, 28(12): 1797-1808

Yuan, Z., Wang, L. and Ji, X. 2014. Prediction of concrete compressive strength research on hybrid models genetic based algorithms and ANFIS. Advances in Engineering Software, 67: 156-163.

Yunani, J. L. 2019. Concrete compressive strength analysis using tree based methods. 2019 International Multi-Conference on Industrial Engineering and Modern Technologies, pp. 1-4, Russia.

Zhu, X. and Wu, X. 2004. Class noise vs. attribute noise: a quantitative study. Artificial Intelligence Review, 22(3): 177–210.

EKLER

EK 1. Üretilecek veri setleri ön işleme kombinasyonları

EK 2. Üretilen veri setleri ile yapılan tüm tahmin sonuçları

EK 1. Üretilecek veri setleri ön işleme kombinasyonları

Yöntem 0: Parametre herhangi bir işleme tabi tutulmamaktadır.					
Yöntem 1: Parametre değeri sıfır ise "yok", sıfırdan büyükse "var" değerini alır.					
Yöntem 2: Parametre değeri sıfır ise "0", sıfırdan büyükse 1'den 3'e kadar aralıktaki konumuna göre bir değer alır.					
Yöntem 3: Parametre değeri sıfır ise "0", sıfırdan büyükse 1'den 5'e kadar aralıktaki konumuna göre bir değer alır.					
Yöntem 4: Parametre herhangi bir işleme tabi tutulmamaktadır.					
Yöntem 5: Parametre değerine logaritmik bir formülle ön işleme yöntemi uygulanır.					
		Yöntem Uygulanan Veri Parametreleri			
No	**Üretilen Verinin Adı**	**Yüksek Fırın Cürufu**	**Uçucu Kül**	**Süper Akışkanlaştırıcı**	**Numune Yaşı**
1	Veri_000_4	Yöntem 0	Yöntem 0	Yöntem 0	Yöntem 4
2	Veri_001_4	Yöntem 0	Yöntem 0	Yöntem 1	Yöntem 4
3	Veri_002_4	Yöntem 0	Yöntem 0	Yöntem 2	Yöntem 4
4	Veri_003_4	Yöntem 0	Yöntem 0	Yöntem 3	Yöntem 4
5	Veri_010_4	Yöntem 0	Yöntem 1	Yöntem 0	Yöntem 4
6	Veri_011_4	Yöntem 0	Yöntem 1	Yöntem 1	Yöntem 4
7	Veri_012_4	Yöntem 0	Yöntem 1	Yöntem 2	Yöntem 4
8	Veri_013_4	Yöntem 0	Yöntem 1	Yöntem 3	Yöntem 4
9	Veri_020_4	Yöntem 0	Yöntem 2	Yöntem 0	Yöntem 4
10	Veri_021_4	Yöntem 0	Yöntem 2	Yöntem 1	Yöntem 4
11	Veri_022_4	Yöntem 0	Yöntem 2	Yöntem 2	Yöntem 4
12	Veri_023_4	Yöntem 0	Yöntem 2	Yöntem 3	Yöntem 4
13	Veri_030_4	Yöntem 0	Yöntem 3	Yöntem 0	Yöntem 4
14	Veri_031_4	Yöntem 0	Yöntem 3	Yöntem 1	Yöntem 4
15	Veri_032_4	Yöntem 0	Yöntem 3	Yöntem 2	Yöntem 4
16	Veri_033_4	Yöntem 0	Yöntem 3	Yöntem 3	Yöntem 4
17	Veri_100_4	Yöntem 1	Yöntem 0	Yöntem 0	Yöntem 4
18	Veri_101_4	Yöntem 1	Yöntem 0	Yöntem 1	Yöntem 4
19	Veri_102_4	Yöntem 1	Yöntem 0	Yöntem 2	Yöntem 4
20	Veri_103_4	Yöntem 1	Yöntem 0	Yöntem 3	Yöntem 4
21	Veri_110_4	Yöntem 1	Yöntem 1	Yöntem 0	Yöntem 4
22	Veri_111_4	Yöntem 1	Yöntem 1	Yöntem 1	Yöntem 4
23	Veri_112_4	Yöntem 1	Yöntem 1	Yöntem 2	Yöntem 4
24	Veri_113_4	Yöntem 1	Yöntem 1	Yöntem 3	Yöntem 4
25	Veri_120_4	Yöntem 1	Yöntem 2	Yöntem 0	Yöntem 4
26	Veri_121_4	Yöntem 1	Yöntem 2	Yöntem 1	Yöntem 4
27	Veri_122_4	Yöntem 1	Yöntem 2	Yöntem 2	Yöntem 4
28	Veri_123_4	Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 4

29	Veri_130_4	Yöntem 1	Yöntem 3	Yöntem 0	Yöntem 4
30	Veri_131_4	Yöntem 1	Yöntem 3	Yöntem 1	Yöntem 4
31	Veri_132_4	Yöntem 1	Yöntem 3	Yöntem 2	Yöntem 4
32	Veri_133_4	Yöntem 1	Yöntem 3	Yöntem 3	Yöntem 4
33	Veri_200_4	Yöntem 2	Yöntem 0	Yöntem 0	Yöntem 4
34	Veri_201_4	Yöntem 2	Yöntem 0	Yöntem 1	Yöntem 4
35	Veri_202_4	Yöntem 2	Yöntem 0	Yöntem 2	Yöntem 4
36	Veri_203_4	Yöntem 2	Yöntem 0	Yöntem 3	Yöntem 4
37	Veri_210_4	Yöntem 2	Yöntem 1	Yöntem 0	Yöntem 4
38	Veri_211_4	Yöntem 2	Yöntem 1	Yöntem 1	Yöntem 4
39	Veri_212_4	Yöntem 2	Yöntem 1	Yöntem 2	Yöntem 4
40	Veri_213_4	Yöntem 2	Yöntem 1	Yöntem 3	Yöntem 4
41	Veri_220_4	Yöntem 2	Yöntem 2	Yöntem 0	Yöntem 4
42	Veri_221_4	Yöntem 2	Yöntem 2	Yöntem 1	Yöntem 4
43	Veri_222_4	Yöntem 2	Yöntem 2	Yöntem 2	Yöntem 4
44	Veri_223_4	Yöntem 2	Yöntem 2	Yöntem 3	Yöntem 4
45	Veri_230_4	Yöntem 2	Yöntem 3	Yöntem 0	Yöntem 4
46	Veri_231_4	Yöntem 2	Yöntem 3	Yöntem 1	Yöntem 4
47	Veri_232_4	Yöntem 2	Yöntem 3	Yöntem 2	Yöntem 4
48	Veri_233_4	Yöntem 2	Yöntem 3	Yöntem 3	Yöntem 4
49	Veri_300_4	Yöntem 3	Yöntem 0	Yöntem 0	Yöntem 4
50	Veri_301_4	Yöntem 3	Yöntem 0	Yöntem 1	Yöntem 4
51	Veri_302_4	Yöntem 3	Yöntem 0	Yöntem 2	Yöntem 4
52	Veri_303_4	Yöntem 3	Yöntem 0	Yöntem 3	Yöntem 4
53	Veri_310_4	Yöntem 3	Yöntem 1	Yöntem 0	Yöntem 4
54	Veri_311_4	Yöntem 3	Yöntem 1	Yöntem 1	Yöntem 4
55	Veri_312_4	Yöntem 3	Yöntem 1	Yöntem 2	Yöntem 4
56	Veri_313_4	Yöntem 3	Yöntem 1	Yöntem 3	Yöntem 4
57	Veri_320_4	Yöntem 3	Yöntem 2	Yöntem 0	Yöntem 4
58	Veri_321_4	Yöntem 3	Yöntem 2	Yöntem 1	Yöntem 4
59	Veri_322_4	Yöntem 3	Yöntem 2	Yöntem 2	Yöntem 4
60	Veri_323_4	Yöntem 3	Yöntem 2	Yöntem 3	Yöntem 4
61	Veri_330_4	Yöntem 3	Yöntem 3	Yöntem 0	Yöntem 4
62	Veri_331_4	Yöntem 3	Yöntem 3	Yöntem 1	Yöntem 4
63	Veri_332_4	Yöntem 3	Yöntem 3	Yöntem 2	Yöntem 4
64	Veri_333_4	Yöntem 3	Yöntem 3	Yöntem 3	Yöntem 4
65	Veri_000_5	Yöntem 0	Yöntem 0	Yöntem 0	Yöntem 5
66	Veri_001_5	Yöntem 0	Yöntem 0	Yöntem 1	Yöntem 5
67	Veri_002_5	Yöntem 0	Yöntem 0	Yöntem 2	Yöntem 5
68	Veri_003_5	Yöntem 0	Yöntem 0	Yöntem 3	Yöntem 5
69	Veri_010_5	Yöntem 0	Yöntem 1	Yöntem 0	Yöntem 5

70	Veri_011_5	Yöntem 0	Yöntem 1	Yöntem 1	Yöntem 5
71	Veri_012_5	Yöntem 0	Yöntem 1	Yöntem 2	Yöntem 5
72	Veri_013_5	Yöntem 0	Yöntem 1	Yöntem 3	Yöntem 5
73	Veri_020_5	Yöntem 0	Yöntem 2	Yöntem 0	Yöntem 5
74	Veri_021_5	Yöntem 0	Yöntem 2	Yöntem 1	Yöntem 5
75	Veri_022_5	Yöntem 0	Yöntem 2	Yöntem 2	Yöntem 5
76	Veri_023_5	Yöntem 0	Yöntem 2	Yöntem 3	Yöntem 5
77	Veri_030_5	Yöntem 0	Yöntem 3	Yöntem 0	Yöntem 5
78	Veri_031_5	Yöntem 0	Yöntem 3	Yöntem 1	Yöntem 5
79	Veri_032_5	Yöntem 0	Yöntem 3	Yöntem 2	Yöntem 5
80	Veri_033_5	Yöntem 0	Yöntem 3	Yöntem 3	Yöntem 5
81	Veri_100_5	Yöntem 1	Yöntem 0	Yöntem 0	Yöntem 5
82	Veri_101_5	Yöntem 1	Yöntem 0	Yöntem 1	Yöntem 5
83	Veri_102_5	Yöntem 1	Yöntem 0	Yöntem 2	Yöntem 5
84	Veri_103_5	Yöntem 1	Yöntem 0	Yöntem 3	Yöntem 5
85	Veri_110_5	Yöntem 1	Yöntem 1	Yöntem 0	Yöntem 5
86	Veri_111_5	Yöntem 1	Yöntem 1	Yöntem 1	Yöntem 5
87	Veri_112_5	Yöntem 1	Yöntem 1	Yöntem 2	Yöntem 5
88	Veri_113_5	Yöntem 1	Yöntem 1	Yöntem 3	Yöntem 5
89	Veri_120_5	Yöntem 1	Yöntem 2	Yöntem 0	Yöntem 5
90	Veri_121_5	Yöntem 1	Yöntem 2	Yöntem 1	Yöntem 5
91	Veri_122_5	Yöntem 1	Yöntem 2	Yöntem 2	Yöntem 5
92	Veri_123_5	Yöntem 1	Yöntem 2	Yöntem 3	Yöntem 5
93	Veri_130_5	Yöntem 1	Yöntem 3	Yöntem 0	Yöntem 5
94	Veri_131_5	Yöntem 1	Yöntem 3	Yöntem 1	Yöntem 5
95	Veri_132_5	Yöntem 1	Yöntem 3	Yöntem 2	Yöntem 5
96	Veri_133_5	Yöntem 1	Yöntem 3	Yöntem 3	Yöntem 5
97	Veri_200_5	Yöntem 2	Yöntem 0	Yöntem 0	Yöntem 5
98	Veri_201_5	Yöntem 2	Yöntem 0	Yöntem 1	Yöntem 5
99	Veri_202_5	Yöntem 2	Yöntem 0	Yöntem 2	Yöntem 5
100	Veri_203_5	Yöntem 2	Yöntem 0	Yöntem 3	Yöntem 5
101	Veri_210_5	Yöntem 2	Yöntem 1	Yöntem 0	Yöntem 5
102	Veri_211_5	Yöntem 2	Yöntem 1	Yöntem 1	Yöntem 5
103	Veri_212_5	Yöntem 2	Yöntem 1	Yöntem 2	Yöntem 5
104	Veri_213_5	Yöntem 2	Yöntem 1	Yöntem 3	Yöntem 5
105	Veri_220_5	Yöntem 2	Yöntem 2	Yöntem 0	Yöntem 5
106	Veri_221_5	Yöntem 2	Yöntem 2	Yöntem 1	Yöntem 5
107	Veri_222_5	Yöntem 2	Yöntem 2	Yöntem 2	Yöntem 5
108	Veri_223_5	Yöntem 2	Yöntem 2	Yöntem 3	Yöntem 5
109	Veri_230_5	Yöntem 2	Yöntem 3	Yöntem 0	Yöntem 5
110	Veri_231_5	Yöntem 2	Yöntem 3	Yöntem 1	Yöntem 5

111	Veri_232_5	Yöntem 2	Yöntem 3	Yöntem 2	Yöntem 5
112	Veri_233_5	Yöntem 2	Yöntem 3	Yöntem 3	Yöntem 5
113	Veri_300_5	Yöntem 3	Yöntem 0	Yöntem 0	Yöntem 5
114	Veri_301_5	Yöntem 3	Yöntem 0	Yöntem 1	Yöntem 5
115	Veri_302_5	Yöntem 3	Yöntem 0	Yöntem 2	Yöntem 5
116	Veri_303_5	Yöntem 3	Yöntem 0	Yöntem 3	Yöntem 5
117	Veri_310_5	Yöntem 3	Yöntem 1	Yöntem 0	Yöntem 5
118	Veri_311_5	Yöntem 3	Yöntem 1	Yöntem 1	Yöntem 5
119	Veri_312_5	Yöntem 3	Yöntem 1	Yöntem 2	Yöntem 5
120	Veri_313_5	Yöntem 3	Yöntem 1	Yöntem 3	Yöntem 5
121	Veri_320_5	Yöntem 3	Yöntem 2	Yöntem 0	Yöntem 5
122	Veri_321_5	Yöntem 3	Yöntem 2	Yöntem 1	Yöntem 5
123	Veri_322_5	Yöntem 3	Yöntem 2	Yöntem 2	Yöntem 5
124	Veri_323_5	Yöntem 3	Yöntem 2	Yöntem 3	Yöntem 5
125	Veri_330_5	Yöntem 3	Yöntem 3	Yöntem 0	Yöntem 5
126	Veri_331_5	Yöntem 3	Yöntem 3	Yöntem 1	Yöntem 5
127	Veri_332_5	Yöntem 3	Yöntem 3	Yöntem 2	Yöntem 5
128	Veri_333_5	Yöntem 3	Yöntem 3	Yöntem 3	Yöntem 5

EK 2. Üretilen veri setleri ile yapılan tüm tahmin sonuçları

No	Tahmin Verisi	Model	Ortalama Karesel Hata (MSE)	Ortalama Hata Karesinin Kökü (RMSE)	Ortalama Mutlak Hata (MAE)	Korelasyon Katsayısı (R2)
1	Veri_000_4	YSA (ANN)	17.54	4.19	2.71	0.9371
		GA (GB)	14.63	3.83	2.32	0.9475
2	Veri_001_4	YSA (ANN)	**18.86**	**4.34**	**2.91**	**0.9324**
		GA (GB)	**13.83**	**3.72**	**2.31**	**0.9504**
3	Veri_002_4	YSA (ANN)	18.41	4.29	2.79	0.9340
		GA (GB)	15.05	3.88	2.35	0.9460
4	Veri_003_4	YSA (ANN)	17.43	4.17	2.73	0.9375
		GA (GB)	14.36	3.79	2.31	0.9485
5	Veri_010_4	YSA (ANN)	17.77	4.22	2.73	0.9363
		GA (GB)	14.75	3.84	2.35	0.9471
6	Veri_011_4	YSA (ANN)	18.51	4.30	2.83	0.9336
		GA (GB)	14.78	3.84	2.36	0.9470
7	Veri_012_4	YSA (ANN)	18.51	4.30	2.85	0.9336
		GA (GB)	14.93	3.86	2.34	0.9465
8	Veri_013_4	YSA (ANN)	18.03	4.25	2.73	0.9353
		GA (GB)	15.00	3.87	2.41	0.9462
9	Veri_020_4	YSA (ANN)	16.97	4.12	2.70	0.9391
		GA (GB)	14.75	3.84	2.35	0.9471
10	Veri_021_4	YSA (ANN)	18.72	4.33	2.86	0.9328
		GA (GB)	14.55	3.81	2.34	0.9478
11	Veri_022_4	YSA (ANN)	18.00	4.24	2.73	0.9354
		GA (GB)	15.31	3.91	2.37	0.9451
12	Veri_023_4	YSA (ANN)	17.96	4.24	2.80	0.9356
		GA (GB)	15.21	3.90	2.38	0.9454
13	Veri_030_4	YSA (ANN)	16.88	4.11	2.69	0.9395
		GA (GB)	14.24	3.77	2.35	0.9489
14	Veri_031_4	YSA (ANN)	17.80	4.22	2.80	0.9361
		GA (GB)	14.28	3.78	2.31	0.9488
15	Veri_032_4	YSA (ANN)	18.22	4.27	2.76	0.9347
		GA (GB)	15.22	3.90	2.36	0.9454
16	Veri_033_4	YSA (ANN)	16.73	4.09	2.72	0.9400
		GA (GB)	15.31	3.91	2.38	0.9451
17	Veri_100_4	YSA (ANN)	20.70	4.55	2.92	0.9258
		GA (GB)	14.18	3.77	2.32	0.9491
18	Veri_101_4	YSA (ANN)	18.83	4.34	2.85	0.9325
		GA (GB)	14.34	3.79	2.35	0.9486
19	Veri_102_4	YSA (ANN)	20.69	4.55	2.89	0.9258
		GA (GB)	15.89	3.99	2.40	0.9430
20	Veri_103_4	YSA (ANN)	20.80	4.56	2.98	0.9254
		GA (GB)	14.65	3.83	2.40	0.9475

21	Veri_110_4	YSA (ANN)	21.89	4.68	2.98	0.9215
		GA (GB)	14.69	3.83	2.35	0.9473
22	Veri_111_4	YSA (ANN)	18.32	4.28	2.84	0.9343
		GA (GB)	14.49	3.81	2.37	0.9480
23	Veri_112_4	YSA (ANN)	21.22	4.61	2.88	0.9239
		GA (GB)	15.50	3.94	2.36	0.9444
24	Veri_113_4	YSA (ANN)	19.99	4.47	2.94	0.9283
		GA (GB)	14.69	3.83	2.36	0.9473
25	Veri_120_4	YSA (ANN)	21.02	4.58	2.90	0.9246
		GA (GB)	15.95	3.99	2.41	0.9428
26	Veri_121_4	YSA (ANN)	18.20	4.27	2.81	0.9347
		GA (GB)	15.07	3.88	2.35	0.9459
27	Veri_122_4	YSA (ANN)	21.71	4.66	2.92	0.9221
		GA (GB)	16.48	4.06	2.40	0.9409
28	Veri_123_4	YSA (ANN)	19.53	4.42	2.83	0.9299
		GA (GB)	16.10	4.01	2.42	0.9422
29	Veri_130_4	YSA (ANN)	20.69	4.55	2.91	0.9258
		GA (GB)	15.86	3.98	2.37	0.9431
30	Veri_131_4	YSA (ANN)	19.16	4.38	2.88	0.9313
		GA (GB)	14.64	3.83	2.32	0.9475
31	Veri_132_4	YSA (ANN)	21.29	4.61	2.95	0.9236
		GA (GB)	16.78	4.10	2.45	0.9398
32	Veri_133_4	YSA (ANN)	19.94	4.47	2.89	0.9285
		GA (GB)	15.27	3.91	2.36	0.9452
33	Veri_200_4	YSA (ANN)	19.58	4.42	2.81	0.9298
		GA (GB)	14.68	3.83	2.35	0.9473
34	Veri_201_4	YSA (ANN)	18.77	4.33	2.91	0.9327
		GA (GB)	15.26	3.91	2.40	0.9453
35	Veri_202_4	YSA (ANN)	19.43	4.41	2.85	0.9303
		GA (GB)	15.43	3.93	2.36	0.9446
36	Veri_203_4	YSA (ANN)	18.88	4.35	2.78	0.9323
		GA (GB)	14.34	3.79	2.34	0.9486
37	Veri_210_4	YSA (ANN)	18.40	4.29	2.81	0.9340
		GA (GB)	15.60	3.95	2.38	0.9440
38	Veri_211_4	YSA (ANN)	17.96	4.24	2.80	0.9356
		GA (GB)	14.72	3.84	2.39	0.9472
39	Veri_212_4	YSA (ANN)	20.29	4.50	2.92	0.9272
		GA (GB)	15.39	3.92	2.38	0.9448
40	Veri_213_4	YSA (ANN)	19.40	4.40	2.83	0.9304
		GA (GB)	15.90	3.99	2.41	0.9430
41	Veri_220_4	YSA (ANN)	19.35	4.40	2.85	0.9306
		GA (GB)	15.25	3.90	2.39	0.9453
42	Veri_221_4	YSA (ANN)	17.62	4.20	2.82	0.9368
		GA (GB)	16.68	4.08	2.51	0.9402
43	Veri_222_4	YSA (ANN)	19.57	4.42	2.87	0.9298

		GA (GB)	16.30	4.04	2.47	0.9415
44	Veri_223_4	YSA (ANN)	19.35	4.40	2.88	0.9306
		GA (GB)	14.73	3.84	2.37	0.9472
45	Veri_230_4	YSA (ANN)	19.13	4.37	2.80	0.9314
		GA (GB)	14.93	3.86	2.38	0.9465
46	Veri_231_4	YSA (ANN)	18.85	4.34	2.94	0.9324
		GA (GB)	15.34	3.92	2.41	0.9450
47	Veri_232_4	YSA (ANN)	20.76	4.56	2.89	0.9255
		GA (GB)	15.33	3.91	2.40	0.9450
48	Veri_233_4	YSA (ANN)	18.96	4.35	2.85	0.9320
		GA (GB)	13.98	3.74	2.34	0.9498
49	Veri_300_4	YSA (ANN)	18.75	4.33	2.78	0.9327
		GA (GB)	15.09	3.88	2.36	0.9459
50	Veri_301_4	YSA (ANN)	18.59	4.31	2.86	0.9333
		GA (GB)	14.75	3.84	2.36	0.9471
51	Veri_302_4	YSA (ANN)	19.37	4.40	2.81	0.9305
		GA (GB)	14.69	3.83	2.33	0.9473
52	Veri_303_4	YSA (ANN)	19.35	4.40	2.84	0.9306
		GA (GB)	15.52	3.94	2.40	0.9444
53	Veri_310_4	YSA (ANN)	19.73	4.44	2.81	0.9292
		GA (GB)	14.94	3.86	2.33	0.9464
54	Veri_311_4	YSA (ANN)	17.73	4.21	2.79	0.9364
		GA (GB)	15.26	3.91	2.41	0.9453
55	Veri_312_4	YSA (ANN)	18.74	4.33	2.82	0.9328
		GA (GB)	16.03	4.00	2.35	0.9425
56	Veri_313_4	YSA (ANN)	20.75	4.56	2.91	0.9256
		GA (GB)	15.84	3.98	2.42	0.9432
57	Veri_320_4	YSA (ANN)	20.01	4.47	2.80	0.9282
		GA (GB)	14.96	3.87	2.36	0.9463
58	Veri_321_4	YSA (ANN)	17.29	4.16	2.78	0.9380
		GA (GB)	15.83	3.98	2.45	0.9432
59	Veri_322_4	YSA (ANN)	18.87	4.34	2.77	0.9323
		GA (GB)	15.20	3.90	2.38	0.9455
60	Veri_323_4	YSA (ANN)	19.04	4.36	2.82	0.9317
		GA (GB)	15.08	3.88	2.39	0.9459
61	Veri_330_4	YSA (ANN)	19.29	4.39	2.77	0.9308
		GA (GB)	15.14	3.89	2.33	0.9457
62	Veri_331_4	YSA (ANN)	18.31	4.28	2.83	0.9343
		GA (GB)	15.02	3.88	2.41	0.9461
63	Veri_332_4	YSA (ANN)	19.48	4.41	2.80	0.9301
		GA (GB)	15.04	3.88	2.38	0.9461
64	Veri_333_4	YSA (ANN)	19.05	4.36	2.82	0.9317
		GA (GB)	14.36	3.79	2.33	0.9485
65	Veri_000_5	YSA (ANN)	16.28	4.04	2.56	0.9416
		GA (GB)	14.84	3.85	2.36	0.9468

66	Veri_001_5	YSA (ANN)	16.41	4.05	2.54	0.9411
		GA (GB)	13.98	3.74	2.34	0.9499
67	Veri_002_5	YSA (ANN)	16.00	4.00	2.50	0.9426
		GA (GB)	15.61	3.95	2.40	0.9440
68	Veri_003_5	YSA (ANN)	15.45	3.93	2.52	0.9446
		GA (GB)	14.13	3.76	2.32	0.9493
69	Veri_010_5	YSA (ANN)	16.75	4.09	2.54	0.9399
		GA (GB)	15.14	3.89	2.37	0.9457
70	Veri_011_5	YSA (ANN)	16.99	4.12	2.59	0.9390
		GA (GB)	15.28	3.91	2.32	0.9452
71	Veri_012_5	YSA (ANN)	17.83	4.22	2.59	0.9361
		GA (GB)	16.14	4.02	2.42	0.9421
72	Veri_013_5	YSA (ANN)	16.43	4.05	2.51	0.9411
		GA (GB)	14.77	3.84	2.35	0.9470
73	Veri_020_5	YSA (ANN)	16.02	4.00	2.52	0.9425
		GA (GB)	14.93	3.86	2.33	0.9464
74	Veri_021_5	YSA (ANN)	16.33	4.04	2.56	0.9414
		GA (GB)	15.13	3.89	2.37	0.9457
75	Veri_022_5	YSA (ANN)	17.40	4.17	2.50	0.9376
		GA (GB)	14.99	3.87	2.39	0.9463
76	Veri_023_5	YSA (ANN)	15.52	3.94	2.48	0.9443
		GA (GB)	15.21	3.90	2.42	0.9454
77	Veri_030_5	YSA (ANN)	16.65	4.08	2.57	0.9403
		GA (GB)	13.84	3.72	2.28	0.9504
78	Veri_031_5	YSA (ANN)	15.65	3.96	2.46	0.9439
		GA (GB)	14.43	3.80	2.34	0.9482
79	Veri_032_5	YSA (ANN)	16.77	4.09	2.54	0.9399
		GA (GB)	15.38	3.92	2.39	0.9449
80	Veri_033_5	YSA (ANN)	15.74	3.97	2.51	0.9435
		GA (GB)	15.17	3.90	2.41	0.9456
81	Veri_100_5	YSA (ANN)	18.84	4.34	2.66	0.9324
		GA (GB)	15.81	3.98	2.42	0.9433
82	Veri_101_5	YSA (ANN)	17.83	4.22	2.58	0.9360
		GA (GB)	14.69	3.83	2.34	0.9473
83	Veri_102_5	YSA (ANN)	19.44	4.41	2.54	0.9303
		GA (GB)	15.98	4.00	2.48	0.9427
84	Veri_103_5	YSA (ANN)	20.33	4.51	2.72	0.9271
		GA (GB)	15.20	3.90	2.38	0.9455
85	Veri_110_5	YSA (ANN)	18.12	4.26	2.53	0.9350
		GA (GB)	15.31	3.91	2.34	0.9451
86	Veri_111_5	YSA (ANN)	18.30	4.28	2.63	0.9344
		GA (GB)	14.32	3.78	2.39	0.9486
87	Veri_112_5	YSA (ANN)	18.27	4.27	2.52	0.9345
		GA (GB)	16.06	4.01	2.43	0.9424
88	Veri_113_5	YSA (ANN)	19.01	4.36	2.59	0.9318

		GA (GB)	15.09	3.88	2.41	0.9459
89	Veri_120_5	YSA (ANN)	19.31	4.39	2.65	0.9307
		GA (GB)	14.92	3.86	2.37	0.9465
90	Veri_121_5	YSA (ANN)	20.49	4.53	2.68	0.9265
		GA (GB)	14.59	3.82	2.33	0.9477
91	Veri_122_5	YSA (ANN)	19.53	4.42	2.58	0.9300
		GA (GB)	15.96	3.99	2.42	0.9428
92	Veri_123_5	YSA (ANN)	19.55	4.42	2.68	0.9299
		GA (GB)	16.05	4.01	2.43	0.9424
93	Veri_130_5	YSA (ANN)	18.57	4.31	2.63	0.9334
		GA (GB)	15.70	3.96	2.41	0.9437
94	Veri_131_5	YSA (ANN)	18.10	4.25	2.60	0.9351
		GA (GB)	15.12	3.89	2.36	0.9458
95	Veri_132_5	YSA (ANN)	19.23	4.39	2.60	0.9310
		GA (GB)	15.71	3.96	2.39	0.9436
96	Veri_133_5	YSA (ANN)	20.16	4.49	2.73	0.9277
		GA (GB)	15.21	3.90	2.39	0.9454
97	Veri_200_5	YSA (ANN)	17.51	4.18	2.60	0.9372
		GA (GB)	14.63	3.82	2.32	0.9475
98	Veri_201_5	YSA (ANN)	17.30	4.16	2.60	0.9380
		GA (GB)	14.91	3.86	2.40	0.9465
99	Veri_202_5	YSA (ANN)	18.19	4.27	2.61	0.9347
		GA (GB)	16.05	4.01	2.42	0.9424
100	Veri_203_5	YSA (ANN)	16.97	4.12	2.59	0.9391
		GA (GB)	14.98	3.87	2.44	0.9463
101	Veri_210_5	YSA (ANN)	19.12	4.37	2.63	0.9314
		GA (GB)	15.25	3.91	2.40	0.9453
102	Veri_211_5	YSA (ANN)	18.24	4.27	2.59	0.9346
		GA (GB)	14.58	3.82	2.32	0.9477
103	Veri_212_5	YSA (ANN)	17.09	4.13	2.53	0.9387
		GA (GB)	16.87	4.11	2.45	0.9395
104	Veri_213_5	YSA (ANN)	19.48	4.41	2.65	0.9301
		GA (GB)	14.79	3.85	2.35	0.9470
105	Veri_220_5	YSA (ANN)	17.59	4.19	2.58	0.9369
		GA (GB)	15.16	3.89	2.35	0.9456
106	Veri_221_5	YSA (ANN)	16.16	4.02	2.55	0.9420
		GA (GB)	15.68	3.96	2.43	0.9438
107	Veri_222_5	YSA (ANN)	19.99	4.47	2.71	0.9283
		GA (GB)	15.74	3.97	2.44	0.9435
108	Veri_223_5	YSA (ANN)	16.40	4.05	2.57	0.9412
		GA (GB)	14.95	3.87	2.35	0.9464
109	Veri_230_5	YSA (ANN)	18.11	4.26	2.61	0.9351
		GA (GB)	14.20	3.77	2.33	0.9491
110	Veri_231_5	YSA (ANN)	18.51	4.30	2.64	0.9336
		GA (GB)	15.06	3.88	2.35	0.9460

111	Veri_232_5	YSA (ANN)	19.29	4.39	2.63	0.9308
		GA (GB)	15.08	3.88	2.38	0.9459
112	Veri_233_5	YSA (ANN)	**17.59**	**4.19**	**2.58**	**0.9369**
		GA (GB)	**13.25**	**3.64**	**2.29**	**0.9525**
113	Veri_300_5	YSA (ANN)	16.92	4.11	2.55	0.9393
		GA (GB)	15.07	3.88	2.36	0.9460
114	Veri_301_5	YSA (ANN)	17.12	4.14	2.57	0.9386
		GA (GB)	14.21	3.77	2.36	0.9490
115	Veri_302_5	YSA (ANN)	18.55	4.31	2.55	0.9335
		GA (GB)	14.67	3.83	2.37	0.9474
116	Veri_303_5	YSA (ANN)	15.83	3.98	2.50	0.9432
		GA (GB)	14.96	3.87	2.36	0.9464
117	Veri_310_5	YSA (ANN)	17.12	4.14	2.51	0.9386
		GA (GB)	14.39	3.79	2.33	0.9484
118	Veri_311_5	YSA (ANN)	17.84	4.22	2.61	0.9360
		GA (GB)	14.84	3.85	2.33	0.9468
119	Veri_312_5	YSA (ANN)	17.11	4.14	2.54	0.9386
		GA (GB)	15.59	3.95	2.35	0.9441
120	Veri_313_5	YSA (ANN)	18.72	4.33	2.60	0.9329
		GA (GB)	15.98	4.00	2.39	0.9427
121	Veri_320_5	YSA (ANN)	15.59	3.95	2.48	0.9441
		GA (GB)	15.12	3.89	2.36	0.9458
122	Veri_321_5	YSA (ANN)	16.53	4.07	2.62	0.9407
		GA (GB)	14.95	3.87	2.36	0.9464
123	Veri_322_5	YSA (ANN)	17.02	4.13	2.51	0.9390
		GA (GB)	14.92	3.86	2.38	0.9465
124	Veri_323_5	YSA (ANN)	16.08	4.01	2.54	0.9423
		GA (GB)	15.20	3.90	2.43	0.9455
125	Veri_330_5	YSA (ANN)	16.66	4.08	2.53	0.9403
		GA (GB)	14.70	3.83	2.38	0.9473
126	Veri_331_5	YSA (ANN)	16.54	4.07	2.56	0.9407
		GA (GB)	15.20	3.90	2.40	0.9455
127	Veri_332_5	YSA (ANN)	17.44	4.18	2.56	0.9375
		GA (GB)	15.00	3.87	2.37	0.9462
128	Veri_333_5	YSA (ANN)	16.01	4.00	2.53	0.9426
		GA (GB)	13.87	3.72	2.32	0.9502

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı ve Soyadı : Bahadır ERİN

Eğitim

Yüksek Lisans	Çankırı Karatekin Üniversitesi Fen Bilimleri Enstitüsü İnşaat Mühendisliği Anabilim Dalı	2019-Halen
Lisans	Erciyes Üniversitesi Yozgat Mühendislik Mimarlık Fakültesi İnşaat Mühendisliği Bölümü	2002-2008

İş Deneyimi

Yıl	**Kurum**	**Görev**
2010-Halen	ÇAKÜ, Yapı İşleri ve Teknik Daire başkanlığı	Mühendis