





Makine Öğrenme Teknikleri Kullanılarak Kükürt Giderme İşleminde Kullanılan Malzeme Miktarının Tahmini

Esra Özcan¹ , Emrullah Sonuç^{2*} ,

^{1,2} Karabük Üniversitesi, Bilgisayar Mühendisliği Bölümü, Karabük, Türkiye

e.esracetinkaya@gmail.com, esonuc@karabuk.edu.tr

Öz

Bazik oksijen fırınlı bir demir çelik fabrikasının yüksek fırın tesislerinde kok, sinter ve diğer demir cevheri malzemelerinin kullanılmasıyla sıvı ham demir üretilmektedir. Bu üretimden sonraki adım çelik üretim süreci olup, hemen öncesinde sıvı ham demir, içerisindeki kükürt oranının belirli bir miktar düşürülmesi amacıyla kükürt giderme tesisinde işlem görmektedir. Desülfürizasyon olarak adlandırılan bu işlemin amacı bazı kükürt giderici reaktifler ilave edilerek hedef kükürt değerini yakalamaktır. İlave edilecek malzeme miktarlarını belirlemek için farklı yöntemler kullanılmaktadır. Genel olarak temel ve veriye dayalı modellerin uygulandığı çalışmalar görülmektedir. Ancak yapay zekâ tekniklerinin bu alandaki kullanımı oldukça kısıtlıdır. Bu çalışmada kükürt giderme işlemindeki malzeme (magnezyum, kireç, florit) oranları makine öğrenme teknikleri ile tahmin edilmiştir. Problem bir regresyon problemi olup altı farklı yöntem (Lineer Regresyon, K-En Yakın Komşu, Karar Ağaçları, Rastgele Orman, XGBoost, Yapay Sinir Ağları) veri seti üzerinde test edilmiştir. Çalışmada kullanılan veriler 2020 yılına ait olup kükürt giderme tesisinden alınmıştır. Verilerin %80'i eğitim, %20'si test için kullanılacak şekilde ayrılmıştır. Değerlendirme ölçütü olarak Doğruluk ve Ortalama Mutlak Hata Yüzdesi (Mean Absolute Percentage Error, MAPE) kullanılmıştır. Elde edilen sonuçlara göre Yapay Sinir Ağı modeli magnezyum, kireç ve florit için sırasıyla %85, %95,4 ve %80,14 doğruluk değerlerine ulaşmıştır. MAPE değerleri ise sırasıyla 14,99, 4,59 ve 19,85 olup bu da modelin başarılı bir tahmin gerçekleştirdiğini ortaya koymaktadır.

Anahtar kelimeler: Makine Öğrenmesi, Yapay Sinir Ağları, Regresyon, Kükürt Giderme.

Prediction of Amount of Material used in Desulfurization Process using Machine Learning Techniques

Abstract

Liquid raw iron is produced by using coke, sinter and other iron ore materials in the blast furnace facilities of an iron and steel factory with a basic oxygen furnace. The next step after the production, is the steelmaking process, and just before that, the liquid raw iron is processed in the desulfurization plant in order to reduce the available sulfur content by a certain amount. The purpose of this process, called desulphurization, is to achieve the target sulfur value by adding some desulfurization reagents. Different methods are used to determine the amount of material to be added. There are many studies in which basic and data-based models are applied in this desulphurization process. However, the use of artificial intelligence techniques in this area is quite limited. In this study, the material (magnesium, lime, fluorite) ratios in the desulfurization process were predicted by machine learning techniques. This problem is a regression problem and six different methods (Linear Regression, K-Nearest Neighbor, Decision Trees, Random Forest, XGBoost, Artificial Neural Networks) are tested on the dataset. The data used in the study belonged to the year 2020 and were taken from the desulfurization plant. 80% of the data is used for training and 20% for testing. Accuracy and Mean Absolute Percentage Error (MAPE) were used as evaluation criteria. According to the results obtained, Artificial Neural Network model obtained 85%, 95.4% and 80.14% accuracy for magnesium, lime and fluorite, respectively. The MAPE values are 14.99, 4.59 and 19.85, respectively, which shows that the model makes a successful prediction.

Keywords: Machine Learning, Artificial Neural Networks, Regression, Desulfurization.

* Sorumlu yazar.
E-posta adresi: esonuc@karabuk.edu.tr

Alındı : 13 Eylül 2021
Revizyon : 4 Kasım 2021
Kabul : 6 Aralık 2021

1. Giriş (Introduction)

Çelik sektöründeki farklı üretim yöntemleri ülkelerin gelişmişlik seviyeleri ve ekonomik durumuna göre farklılık göstermektedir (Ersöz vd., 2016). Bunlar arasında en yaygın olan yöntem ise demir cevherinden sağlanan sıvı ham demir kullanımınıdır. Çelik üretimi iki farklı yolla yapılmaktadır. Bunlardan birincisi yüksek fırınlarda oksijen üfleme yöntemi, diğeri ise elektrik ark ocaklarında hurdanın eritilmesi yöntemidir (Yıldız, 2017).

Bazik oksijen fırınlı entegre bir tesiste kömür ve cevher ham maddeleri ile üretim sürecinin ilk aşamasına başlanmaktadır. Kok üretimi için kok bataryalarında kömür işlem görmektedir. Cevher ise sinter tesisinde sinterleme için kullanılmaktadır. Elde edilen kok ve sinter ile diğeri demir cevherleri, yüksek fırınlarda bir araya gelerek, sıvı ham demir üretimi gerçekleştirilmektedir. Sıvı ham demirin sıvı çeliğe dönüşümü ise hurda, kireç ve oksijen kullanımıyla bazik oksijen fırınlarında gerçekleşmektedir. Bu aşamadan sonra ise sıvı çeliğin kalitesini artırmak amacıyla pota fırınlarında alaşım malzemeler eklenir. Döküme hazır hale gelerek döküm tesislerinde slab, kütük vb. formlarına getirilir (Türkoğlu ve Özyıldırım, 2017). Bu yarı mamuller haddeleme işlemiyle de son ürün haline gelmektedir.

Entegre tesislerde çelik üretiminin en önemli girdisi sıvı ham demirdir (Özcan ve Köprü, 2020). Yüksek fırınlarda üretimi yapılan sıvı ham demir, çelik üretimine geçmeden önce desülfürizasyon işlemine tabii tutulmaktadır (Çetin, 2016). Amaç; sıvı ham demir içerisindeki kükürt, fosfor ve silis gibi değerlerin istenilen seviyeye getirilmesidir.

Bu maddelerden kükürdün seviyesi oldukça önemlidir; çünkü kükürdün çeliğin mekanik özellikleri üzerinde olumsuz etkileri mevcuttur. Ayrıca darbe mukavemetini azaltan zararlı bir element olup malzemenin sertleşebilirliğini ve kaynaklanabilirliğini de etkileyebilmektedir. Bu nedenle çelik içinde miktarının sınırlı olması beklenmektedir (Yıldız, 2017).

Kükürt oranını azaltmak için kireç, magnezyum, florit ve karpit gibi kükürt giderici reaktif malzemeler kullanılmaktadır. Bu malzemelerin kullanımı ve miktarları istenilen kükürt değerine ulaşmak için bir modele ya da yönteme göre kurgulanmalıdır (Visuri vd., 2020).

Literatürde yer alan bir çalışmada torpedo arabasındaki sıcak metalde kükürt giderme işlemi için yapay sinir ağı kullanılmıştır (Deo vd., 1994). Modelde sıcak metal ağırlığı, ilk kükürt içeriği, işlem süresi reaktif enjeksiyon hızı gibi değişkenler kullanılarak bir tahmin işlemi gerçekleştirilmiştir. Başka bir çalışmada geri yayımlı bir yapay sinir ağı modeli geliştirilerek kükürt giderme işlem parametrelerini tahmin etme hedeflenmiştir (Rong vd., 2005). Diğeri ise geri yayımlı bir sinir ağı ile sıcak metal ağırlığı, sıcaklık değeri, sıcak metaldeki başlangıç ile son kükürt içeriği gibi kriterler kullanılarak magnezyum tozunu tahmin

etmek için bir model kurgulanmıştır (Zhan vd., 2010). Magnezyum bazlı ve kireç bazlı enjeksiyonların karşılaştırması için kükürt gidermede kinetik bir model önerilmiştir (Jin vd., 2006). Bir diğeri yapılan çalışmada ise proses sonuçları regresyon yaklaşımı kullanılarak tahmin edilmiştir (Vino vd., 2007). Tahmin için sıcaklık faktörü baz alınarak regresyon modelleri oluşturulmuştur. Başka bir çalışmada kullanılan geri yayımlı sinir ağına moment teriminin eklenmesi önerilmiş ve bu şekilde kükürt giderici ekleme miktarı tahmin modeli kurgulanmıştır (Liang vd., 2011).

İlgili literatür incelendiğinde kükürt giderici malzemelerin miktarına yönelik tahmin çalışmalarının kısıtlı olduğu görülmektedir. Son yıllarda makine öğrenme tekniklerinin gelişimi ve ülkemizdeki sanayi tesislerindeki üretim artışı bu çalışmanın motivasyon kaynağı olmuştur. Çalışmada, kükürt giderme işleminde kullanılacak malzemelerin miktarını tespit etmek için farklı regresyon modelleri test edilmiştir. Bu modellerden elde edilen sonuçlar kıyaslanarak literatüre bir katkı sağlanması amaçlanmıştır. Makalenin geriye kalan kısımları şu şekilde organize edilmiştir: İkinci kısımda kullanılan makine öğrenme tekniklerine kısaca değinilmiş ve veri seti tanımlanmıştır. Üçüncü bölümde kullanılan yöntemlerden elde edilen sonuçlar tablolarla sunulmuş ve yorumlanmıştır. Son bölümde ise çalışmadan elde edilen kazanımlara değinilmiştir.

2. Materyal ve Yöntem (Material and Method)

Çalışmada makine öğrenme teknikleri kullanılmıştır. Bu teknikleri uygulamak için yararlanılan veri seti, giriş ve çıkış parametreleri belirlenerek kullanıma uygun hale getirilmiştir. Tahmin edilecek veri sayısal bir değerdir. Bu şekilde sayısal bir değeri tahmin etmeyi içeren problemler regresyon olarak tanımlanmaktadır. Regresyon, bağımsız değişkenlerin bağımlı değişkendeki değişimini Eşitlik (1)'deki gibi ifade etmektedir.

$$Y = a + bX_1 + cX_2 + dX_3 + \varepsilon \quad (1)$$

Tahmine dayalı modelleme, yani veriler üzerinde tahmin yapmak için geçmiş verilerden yararlanmayı gerektirir.

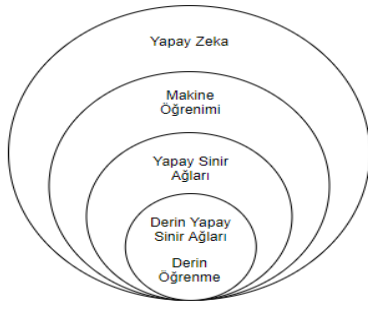
Regresyon problemleri, makine öğrenme mimarileri ile çözülebilmektedir. Bu kapsamda aşağıdaki yöntemler kullanılarak farklı modeller kurgulanmış ve tahmin işlemi gerçekleştirilmiştir:

- Lineer Regresyon,
- K-En Yakın Komşu,
- Karar Ağaçları,
- Rastgele Orman,
- XGBoost,
- Yapay Sinir Ağı (YSA).

Makine öğrenmesi, yapay zekanın bir alt dalı olup ilgili probleme ait verilerden faydalanarak bir modelin kurgulandığı teknikleri içermektedir. Şekil 1'de yapay zekâ ve alt dalları gösterilmektedir. Makine öğrenmesi

algoritmaları verileri kullanarak birtakım kurallar türetir ve bu sayede öğrenme işlemi gerçekleştirerek problemi çözme başarısının artırılmasını hedefler (Alan, 2020).

Makine öğrenmesinin temeli Alan Turing'in 1950'li yıllarda yaptığı çalışmalarla atılmıştır. Alan Turing'in bir makinenin zekaya sahip olabileceğini araştırması araştırmacıların ilgisi çekmiştir. Arthur Samuel'in hazırladığı bir dama programı ile 1959 yılında makine öğrenmesi kavramı olarak ele alınmıştır (Görgün, 2020).



Şekil 1. Makine Öğrenme Metodolojisi (Machine learning methodology)

Makine öğrenme teknikleri günümüzde birçok alanda yaygın olarak kullanılmaktadır. Bunlardan başlıcaları; savunma sanayii, yüz tanıma, nesne ve ses tanıma, otonom sistemler, medikal ve tıp alanında hastalık teşhisi, bankacılıkta sahtelik tespitidir (Elen ve Avuçlu, 2020, Aswad ve Sonuç, 2020, Dolapci ve Özcan, 2021). Makine öğreniminde kullanılacak modeller veriye göre değişiklik göstermektedir. Yine verinin yoğunluğunun yani veri setinin genişlik ve varyasyonlara sahip olma durumunun model başarısındaki rolü büyüktür (Baydilli, 2021).

2.1. Veri Seti (Data Set)

Çalışmada entegre bir demir çelik tesisinin kükürt giderme prosesinden elde edilen veriler kullanılmıştır.

Veri dosyası analiz edilerek tahmin işlemi için gerekli kriterler tespit edilmiştir. Bu kapsamda, 7 kriterin baz alındığı tamamı numerik değer içeren 4214 kayıttan oluşan veri seti elde edilmiştir. Tahmin edilecek değişkenler Tablo 1'deki gibi belirlenmiştir.

Tablo 1. Tahmin Edilecek Değişkenler (Predictable Variables)

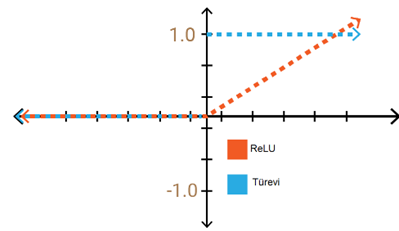
| Çıkış Değişkenleri |
|----------------------------|
| Magnezyum (Mg) |
| Kireç (CaO) |
| Florit (CaF ₂) |

Üç malzeme için ayrı ayrı tahmin modelleri oluşturularak; sıvı ham demir tonaj miktarı, hedeflenen kükürt değeri, mevcut kükürt değeri ve mevcut silisyum değerleri sabit girdi değişkenleri seçilmiştir. Magnezyum miktarı tahmininde bu kriterler kullanılmış olup Kireç miktarı tahmininde girdiye Magnezyum; Florit miktarı tahmininde ise girdiye Magnezyum ve Kireç değişkenleri de ilave edilmiştir.

2.2. Yapay Sinir Ağları (Artificial Neural Network)

YSA makine öğrenimi yöntemlerinden biri olup insan beyninde bulunan nöronların görevinden esinlenerek geliştirilmiştir. Beyinde gerçekleşen öğrenme fonksiyonu simüle edilerek sınıflandırma, tahmin, kümeleme gibi problemleri çözmek için kullanılmaktadır (Karaatlı, 2012).

Sinir ağlarında doğrusal olmayan özellikleri tanıtmak amacıyla aktivasyon fonksiyonları kullanılmaktadır. Çalışmada, hız açısından avantajlı olması sebebiyle Şekil 2'de yer alan ReLU fonksiyonu kullanılmıştır. ReLU; 0 ile $+\infty$ arasında değer üreten bir aktivasyon fonksiyonudur.



Şekil 2. ReLU fonksiyonu (ReLU function)

2.3. Doğrusal Regresyon (Linear Regression)

Basit doğrusal regresyon, iki sürekli (nicel) değişken arasındaki ilişkileri özetlememize ve incelememize izin veren Eşitlik (2)'deki gibi istatistiksel bir yöntemdir. Değişkenlerden biri bağımsız değişken, diğeri bağımlı değişken olarak kabul edilir. Bağımlı değişkenin sürekli olması gerekirken, bağımsız değişkenler ya sürekli ya da kategorik özelliğe sahip değişkenler olabilir (Gök, 2017).

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (2)$$

2.4. Karar Ağacı Regresyonu (Decision Tree Regressor)

Karar ağacı, regresyon veya sınıflandırma modellerini kurgularken ağaç yapısından faydalanır. İşlenen veri kümesi alt kümelere parçalanırken buna paralel olarak Eşitlik (3)'te verilen karar ağacı modeli geliştirilir. Sonuçta düğüm ve yapraklardan oluşan bir model elde edilir (Başer vd., 2021).

$$S(T, X) = \sum_{c \in X} P(c) S(c) \quad (3)$$

2.5. K-En Yakın Komşu Regresyonu (K-neighbors Regressor)

Eşitlik (4)'te ifade edilen K-en yakın komşu yöntemi, eldeki tüm durumları hafızasında saklar ve mesafe fonksiyonu vb. benzerlik ölçüsü kullanılarak tahmin işlemini gerçekleştirir. Bu yöntem, ilk olarak 1970'lerin başında örüntü tanıma ve istatistiksel tahminlemede kullanılmıştır (Salman ve Sonuç, 2021).

$$\hat{y} = f(x) = \frac{1}{k} \sum_{j=1}^k y_{ij} \quad (4)$$

2.6. Rastgele Orman Regresyonu (Random Forest Regressor)

Karar ağaçlarının birden fazla kullanılmasıyla oluşmuş bir modeldir. Karar Ağacı Regresyonunun bazı durumlarda aşırı öğrenmeye neden olması beraberinde bir dezavantajı getirmektedir (Pekel, 2020). Bu sorunu ortadan kaldırmak için, Karar Ağacı Regresyonu yerine Rastgele Orman Regresyonu uygulanarak aşırı öğrenme sınırlandırılabilir (Görgün, 2020).

$$F(x) = \frac{1}{J} \sum_{j=1}^J c_{jfull} + \sum_{k=1}^K (\frac{1}{J} \sum_{j=1}^J contribution(x, k)) \quad (5)$$

2.7. Xgboost (Extreme Gradient Boosting)

Gradyan artırma algoritması; regresyon ve sınıflandırma problemlerinde kullanılan makine öğrenmesi yöntemidir. Veri setine farklı ağırlıklar verilmesi ile elde edilen ağaçlar topluluğundan tahminler yapılması ve zayıf tahminlerle modeli güçlendirmek temel amacıdır. Bu algoritmanın paralelleştirilmiş ve optimize edilmiş bir versiyonu ise Xgboost'dur. Veriler üzerinde mümkün olan en iyi modeli eğitmek yerine, eğitim veri kümesinin çeşitli alt kümelerindeki binlerce modeli eğitir ve ardından en iyi performans gösteren modelin çıktısı alınır (Dilwani, 2019). Verilerin parçalara ayrılarak analiz işlemine tabi tutulmasıyla gerçekleştirilmektedir. Buradaki amaç daha küçük parçalar kullanılarak daha iyi bir tahmin elde edilmesi ve başarının artırılmasıdır (Yangın, 2019).

$$\sum_{i=1}^n L(y_i, p_i) + \frac{1}{2} \alpha O^2_v \quad (6)$$

2.8. Model Performans Metrikleri (Model Performance Metrics)

Regresyon modellerinde, öğrenme işlemini başarılı kılmak için hata fonksiyonlarından çıkan değerler minimum olması hedeflenmektedir. Model için en uygun optimizasyon fonksiyonunu seçerek bu hedefi gerçekleştirmek mümkündür (Henrique vd., 2019).

Modellerin performansı ise hata fonksiyonları ile ölçülmektedir. Regresyon problemleri için kullanılan hata fonksiyonları Ortalama mutlak hata (MAE), Hata kareler ortalaması (MSE) ve Ortalama mutlak hata yüzdesi (MAPE) şeklinde sıralanabilir. Tahmin sonucunun gerçek değerden ne kadar uzak olduğu bu metrikler ile analiz edilebilmektedir. Uygulanan metotlarda değerlendirme ölçütü olarak Eşitlik (7)'de ifade edilen MAPE tercih edilmiştir. Eşitlikte G gerçek değer, T ise tahmin değeri ifade etmektedir. MAPE değerinin %10'un altında çıkması yüksek doğruluğa sahip modeli, %10 ile %20 arasında olması ise doğru tahmin modelini temsil etmektedir (Karabıçak vd., 2018).

$$MAPE = \frac{\sum_{i=1}^n \frac{|G_i - T_i|}{|G_i|}}{n} \quad (7)$$

Modellerin başarısını ölçmek için kullanılan bir diğer metrik de doğruluktur. Eşitlik (8)'deki gibi modelde doğru tahmin edilen alanların toplam veri kümesine oranı ile doğruluğu hesaplamak mümkündür.

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Burada TP (True Positive, Doğru Pozitif), TN (True Negative, Doğru Negatif), FP (False Positive, Yanlış Pozitif) ve FN (False Negative, Yanlış Negatif)'i ifade etmektedir.

2.9. Uygulama (Application)

Çalışma için *python* programlama dili ve model kurgusu için desteklediği kütüphaneler ile *spyder* programından yararlanılmıştır.

2.9.1. Python (Python)

Python okunabilirliği kolay, modüler ve yorumlanabilir bir script dildir. Son yıllarda bilimsel hesaplama yöntemlerinin başında gelen yapay zekâ alanında yaygın şekilde kullanılmaktadır. Sonuçları kolaylıkla okuma, analiz etme ve görselleştirme imkânı sayesinde araştırmacılar tarafından tercih edilmektedir. Platform bağımsız olması en büyük avantajlarından bir tanesidir (Ürün, 2019). Çalışmada *python 3.7.4* sürümü kullanılmıştır.

2.9.2. Spyder (Spyder)

Python için pek çok editör bulunmakta olup çalışmada *Spyder* programı tercih edilmiştir. Ücretsiz ve açık kaynak bir yazılımdır. *Spyder* programının değişken görüntüleme ekranından değişkenlerin isim, tür, boyut ve aldıkları değerlere ulaşmak mümkündür (Alkan, 2019).

2.9.3. Keras (Keras)

Python pek çok kütüphane desteği vermektedir. Bunlardan, makine öğrenme modellerini tanımlamak ve eğitmek amacıyla *Keras* kütüphanesi çalışmada kullanılmıştır.

2.9.4. Tkinter (Tkinter)

Ayrıca ara yüz geliştirmeye yarayan, form uygulamaları için olanak sağlayan *Tkinter* kütüphanesi de uygulamaya dahil edilerek tekrar tekrar kullanılabilirliği imkân sağlanmıştır.

3. Deneysel Sonuçlar (Experimental Results)

Tesiste kullanılan kükürt giderici üç reaktif malzeme olan Magnezyum, Kireç ve Florit için modelin tahmin yetenekleri test edilmiştir. Bu değişkenlerin tahminine yönelik her bir değişken için modeller oluşturulmuştur. Dolayısıyla model çıkışı tek bir değişkenden oluşmaktadır.

Kükürt giderme işlemi için tesiste kullanılan en önemli malzeme Magnezyum'dur (Özmen vd., 2018). Bundan dolayı çalışmada ilk hedef Magnezyum miktarının tahmin edilmesidir. Magnezyum tahmin işlemi gerçekleştirildikten sonra Kireç tahmini gerçekleştirilmiş olup bu tahminde Magnezyum değerlerinden de yararlanılmıştır. Son olarak Florit tahmin işlemi gerçekleştirilmiş olup burada da Magnezyum ve Kireç değerleri kullanılarak tahmin işlemi gerçekleştirilmiştir.

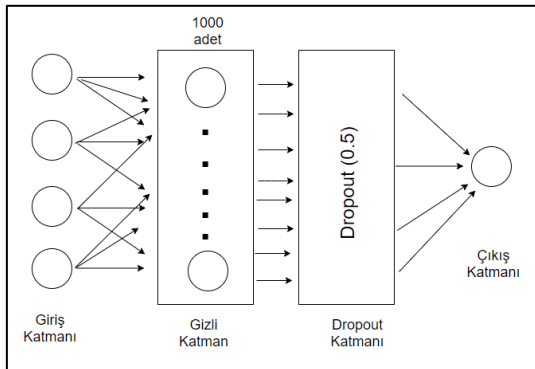
Bu kurgu dahilinde Linear Regression, KNeighborsRegressor, Decision Tree Regressor, Random Forest Regressor ve XGBoost metodları uygulanmıştır. Bu metodların algoritmaları *python* kütüphanelerinde tanımlı olup; LinearRegression, KNeighborsRegressor, DecisionTreeRegressor, RandomForestRegressor ve XGBRegressor fonksiyonları uygulamaya dahil edilmiştir. RandomForestRegressor metodunda *n_estimators* parametresi 100 olarak seçilmiştir. XGBRegressor için aşağıdaki şekilde parametreler belirlenmiştir. Diğer metodlar varsayılan hali ile kullanılmıştır.

```
XGBRegressor(base_score=0.5, booster='gbtree', colsample_bylevel=1,
             colsample_bynode=1, colsample_bytree=1, gamma=0,
             importance_type='gain', learning_rate=0.1, max_delta_step=0,
             max_depth=3, min_child_weight=1, n_estimators=100,
             n_jobs=1, nthread=None, random_state=0,
             reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None,
             silent=None, subsample=1, verbosity=1)
```

Şekil 3. XGBoost Modeli için Kullanılan Parametreler (Parameters Used for XGBoost Model)

Makine öğrenimi modellerinin başarısı, içerdiği parametrelere göre değişkenlik göstermektedir. Bu amaçla model farklı katman/nöron sayıları, maliyet fonksiyonları vb. parametreler kullanılarak test edilmiştir. Kullanılan farklı parametreler arasından en iyi performansa sahip mimari seçilerek ağı eğitimi tamamlanmıştır.

Geliştirilen YSA modeli Şekil 4'te gösterilmiştir (Özcan, 2021). Model sonuçlarının kıyaslanması için MAPE ve Doğruluk olmak üzere iki farklı metrik kullanılmıştır. Her bir çıkış değeri için sonuçlar Tablo 2, Tablo 3 ve Tablo 4'te gösterilmiştir.



Şekil 4. Önerilen YSA Modeli (Proposed ANN Model)

Tablo 2'ye göre MAPE ve doğruluk değerleri baz alındığında; Magnezyum için en iyi sonuçlar YSA modeli tarafından elde edilmiştir. Lineer Regresyon YSA'dan sonra başarılı olan ikinci yöntem olarak görülmektedir. Ayrıca K-En Yakın Komşu yöntemi de Lineer Regresyon'a yakın bir sonuç elde etmiştir. Bu üç yöntemin diğerlerine göre daha başarılı bir tahmin yaptığı görülmektedir.

Tablo 2. Mg Değişkeni için Metodların Sonuçları (Results of Methods for Mg Variable)

| Yöntem | MAPE | Doğruluk |
|------------------|-------|----------|
| Lineer Regresyon | 17,15 | 82,85 |
| K-En Yakın Komşu | 18,28 | 81,72 |
| Karar Ağaçları | 46,74 | 53,26 |
| Rastgele Orman | 44,70 | 55,29 |
| XGBoost | 44,37 | 55,63 |
| YSA | 14,99 | 85,00 |

Tablo 3'e göre Kireç için en iyi değerler Magnezyum da olduğu gibi yine YSA modeli tarafından elde edilmiştir. Yine Lineer Regresyon YSA'dan sonra ikinci sırada yer almakta ve onu K-En Yakın Komşu yöntemi izlemektedir. Diğer yöntemlerin elde ettikleri değerler ise kıyaslanacak seviyede değildir.

Tablo 3. CaO Değişkeni için Metodların Sonuçları (Results of Methods for CaO Variable)

| Yöntem | MAPE | Doğruluk |
|------------------|-------|----------|
| Lineer Regresyon | 6,03 | 93,96 |
| K-En Yakın Komşu | 7,81 | 92,19 |
| Karar Ağaçları | 48,53 | 51,47 |
| Rastgele Orman | 47,82 | 52,18 |
| XGBoost | 47,61 | 52,39 |
| YSA | 4,59 | 95,41 |

Tablo 4'e göre Florit için en iyi değerler diğer çıkış değişkenlerinde olduğu gibi YSA modeli tarafından elde edilmiştir. Lineer Regresyon ve K-En Yakın Komşu YSA'ya yakın değerleri elde eden iki yöntem olarak görülmektedir. Diğer yöntemlerin başarısı ise rekabet edecek seviyede değildir.

Tablo 4. CaF₂ Değişkeni için Metodların Sonuçları (Results of Methods for CaF₂ Variable)

| Yöntem | MAPE | Doğruluk |
|------------------|-------|----------|
| Lineer Regresyon | 25,05 | 74,94 |
| K-En Yakın Komşu | 25,48 | 74,51 |
| Karar Ağaçları | 65,86 | 34,14 |
| Rastgele Orman | 63,24 | 36,76 |
| XGBoost | 62,18 | 37,82 |
| YSA | 19,85 | 80,14 |

Tüm değerler ışığında tahmini gerçekleştirilen üç farklı malzeme için YSA modelinin diğer yöntemlere göre başarılı olduğu tespit edilmiştir. YSA Magnezyum için %85, Kireç için %95,4 ve Florit için %80,14 doğruluk değerlerini elde etmiştir. Yine bu değişkenler için YSA sırasıyla 14,99, 4,59 ve 19,85 MAPE değerlerine ulaşmıştır. Florit ve Kireç tahmininde

modele Magnezyum'un deęişken olarak eklenmesi, bu malzemelerin miktarının belirlenmesinde Magnezyum miktarının da etkili olmasından kaynaklanmaktadır.

Magnezyumun modele eklenmeden ve eklendikten sonra elde edilen sonuçlar karşılaştırmalı olarak Tablo 5'te verilmiştir.

Tablo 5. Giriş Deęişkeni Olarak Mg'nin Model Sonuçlarına Etkisi (Effect of the Results of the Model for Mg as an Input Variable)

| Yöntem | Cao (Kireç) | | Mg'li Model | | CaF ₂ (Florit) | | Mg'li Model | |
|------------------|-------------|----------|-------------|----------|---------------------------|----------|-------------|----------|
| | MAPE | Doęruluk | MAPE | Doęruluk | MAPE | Doęruluk | MAPE | Doęruluk |
| Lineer Regresyon | 18,347 | 81,65 | 6,03 | 93,96 | 33,866 | 66,13 | 25,05 | 74,94 |
| K-En Yakın Komşu | 19,631 | 80,37 | 7,81 | 92,19 | 36,841 | 63,16 | 25,48 | 74,51 |
| Karar Ağaçları | 48,172 | 51,83 | 48,53 | 51,47 | 65,845 | 34,15 | 65,86 | 34,14 |
| Rastgele Orman | 45,543 | 54,46 | 47,82 | 52,18 | 59,809 | 40,19 | 63,24 | 36,76 |
| XGBoost | 45,075 | 54,92 | 47,61 | 52,39 | 59,492 | 40,51 | 62,18 | 37,82 |
| YSA | 16,507 | 83,49 | 4,59 | 95,41 | 36,005 | 63,99 | 19,85 | 80,14 |

4. Sonuçlar ve Tartışma (Conclusions and Discussions)

Çelik malzemelerin sanayinin gelişimiyle ülkemizde ve tüm dünyada birçok alanda kullanımı bu maddenin önemini artırmıştır. Sıvı ham demir kullanılarak elde edilen çeliğin kalitesi, bu aşamada gerçekleştirilen kükürt giderme işlemiyle yakından ilişkilidir. Kükürt giderme işleminde kullanılan başlıca malzemelerin en başında magnezyum, kireç ve florit gelmektedir.

Bu çalışmada, üretilen sıvı ham demirde kükürt giderme işlemi için yukarıda bahsedilen malzemelerin miktarının tahmini için farklı regresyon yöntemleri (Lineer Regresyon, K-En Yakın Komşu, Karar Ağaçları, Rastgele Orman, XGBoost, Yapay Sinir Ağı) test edilmiştir. Bu yöntemlerden elde edilen sonuçlar MAPE ve doğruluk metrikleri kullanılarak değerlendirilmiş ve YSA yöntemi diğer yöntemlere kıyasla başarılı olmuştur. MAPE değerleri Kireç için %5'in altında, Magnezyum ve Florit için ise %20'nin altında sonuçlanmıştır. Bu deęerin %20'den daha düşük bir deęerde olması, geliştirilen modeli doğru tahmin modeli sınıfına dahil etmektedir.

Veri setindeki verilerin sınırlı sayıda olması makine öğrenme modellerinin başarısını da sınırlı hale getirebilmektedir. Bu yüzden ileriki çalışmalarda veri artırma yöntemi kullanılarak makine öğrenme modelleri test edilerek başarı oranının daha da artırılabilmesi ve hata deęerlerinin minimum hale getirilmesi mümkün olabilir.

Kaynaklar (References)

Alan, A., 2020. Makine Öğrenmesi Sınıflandırma Yöntemlerinde Performans Metrikleri ile Test Tekniklerinin Farklı Veri Setleri Üzerinde Deęerlendirilmesi, Yüksek Lisans Tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Elazığ.

Aswad, S.A., Sonuç, E., 2020. Classification of VPN Network Traffic Flow Using Time Related Features on Apache Spark. In 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 1-8). IEEE.

Alkan, O., 2019. Parkinson Hastalığının Teşhisinde Derin Öğrenme Yöntemi ile Spect Görüntü Analizi, Yüksek Lisans Tezi, Ağrı İbrahim Çeçen Üniversitesi Fen Bilimleri Enstitüsü, Ağrı.

Başer, B.Ö., Yangın, M., Sarıdaş, E.S., 2021. Makine Öğrenmesi Teknikleriyle Diyabet Hastalığının Sınıflandırılması. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 25(1), 112-120.

Baydilli, Y.Y., 2021. Polen Taşıyan Bal Arılarının MobileNetV2 Mimarisi ile Sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, (21), pp.527-533.

Çetin, Z., 2016. Sürekli Döküm Prosesinde Pota Nozulu Tıkanma Probleminin Analizi ve Azaltılması, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.

Deo, B., Datta, A., Haresh, M., Kalra, P.K., Boom, R., 1994. Adaptive Neural Net (ANN) Models for Desulfurization of Hot Metal and Steel. Steel Research International, 65(11), 466-471.

Dilwani, A.A.R., 2019. Makine Öğrenmesi ile Acil Triyaj ve Hastane Yatış Tahmini, Yüksek Lisans Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.

Dolapci, B., Özcan, C., 2021. Automatic Ship Detection and Classification using Machine Learning from Remote Sensing Images on Apache Spark. Journal of Intelligent Systems: Theory and Applications, 4(2), pp.94-102.

Elen, A., Avuclu, E., 2020. A Comparison of Classification Methods for Diagnosis of Parkinson's. International Journal of Intelligent Systems and Applications in Engineering, 8(4), 164-170.

Ersöz, F., Ersöz, T., Erkmen, İ.N., 2016. Dünyada ve Türkiye'de Ham Çelik Üretimine Bakış. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, 32(2), 1-12.

Gök, B., 2017. Makine Öğrenmesi Yöntemleri ile Akademik Başarının Tahmin Edilmesi. Gazi Üniversitesi Fen Bilimleri Dergisi, 5(3), 139-148.

Görgün, M., 2020. Makine Öğrenmesi Yöntemleri ile Kalp Hastalığının Teşhis Edilmesi, Yüksek Lisans Tezi, İstanbul Aydın Üniversitesi, Lisansüstü Eğitim Enstitüsü, İstanbul.

Henrique, B.M., Sobreiro, V.A., Kimura, H., 2018. Stock price prediction using support vector regression on daily and up to the minute prices. The Journal of finance and data science, 4(3), pp.183-201.

- Jin, Y., bi, X.G., Yu, S.R., 2006. Kinetic Model for Powder Injection Desulfurization. *Acta Metallurgica Sinica*, 19(4), 258-264.
- Karaatlı, M., 2012. Yapay Sinir Ağları ile Otomobil Satış Tahmini. *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 8(17), 87-100.
- Karabıçak, Ç., Avcı, S., Akman, G., Aladağ, Z., 2018. Determination of Demand Estimation Methods by Values and Variability Measures for Stock Items in a Cleaning Paper Company. *Journal of Current Researches on Engineering Science and Technology*, 4(1), 47-68.
- Liang, X.P, Ban, S.X., Wang, Y., Ge, W.S., Huang, Z.H., 2011. Research and Application of Desulfurizer Addition Amount Prediction Model Based on Neural Network. *Metallurgical Industry*.
- Özcan C., Köprü, E.Y., 2020. Yapay Sinir Ağları ile Sıvı Ham Demir Tahmini ve 5. Yüksek Fırın Uygulaması. *Avrupa Bilim ve Teknoloji Dergisi, Özel Sayı*, 155-162.
- Özcan, E., 2021. Kükürt Giderme İşlemi İçin Kullanılan Malzeme Miktarının Makine Öğrenme Yöntemleri İle Tahmini, Yüksek Lisans Tezi, Karabük Üniversitesi Lisansüstü Eğitim Enstitüsü, Karabük.
- Özmen, K., Eskiuyurt, T.G., Şahin, H., Erkal, H., Kocabaş, T., Çakır, M., Soysal Atan, B., 2018. Sıvı Ham Demir Kükürt Giderme Prosesinde Kullanılan Granüle Magnezyum Tüketiminin Seviye-2 Yazılımı ile Optimizasyonu. 19. Metalurji ve Malzeme Kongresi (IMMC 2018).
- Pekel, E., 2020. Estimation of soil moisture using decision tree regression. *Theoretical and Applied Climatology*, 139(3), pp.1111-1119.
- Rong, Z., Dan, B., Yi, J., 2005. A BP Neural Network Predictor Model for Desulfurization Molten Iron. *International Conference on Advanced Data Mining and Applications*, 22-24 July 2005, pp. 728-735.
- Salman K., Sonuç, E., 2021. Thyroid Disease Classification Using Machine Learning Algorithms. In *Journal of Physics: Conference Series* (Vol. 1963, No. 1, p. 012140). IOP Publishing.
- Türkoğlu, S., Özyıldırım, B. M., 2017. Developing Oxygen Amount Prediction Model of Basic Oxygen Furnace Steelmaking Process with Machine Learning Algorithms. *Çukurova Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 39(12), 22-31.
- Ürün, S., 2019. Python ile Programlamanın Temelleri.
- Vinoo, D.S., Mazumdar, D., Gupta, S.S., 2007. Optimization and Prediction Model of Hot Metal Desulphurisation. *Ironmaking & Steelmaking*, 34(6), 471-476.
- Visuri, V.V., Vuolio, T., Haas, T., Fabritius, T., 2020. A Review of Modeling Hot Metal Desulfurization. *Steel Research International*, 91(4), 1900454.
- Yangın, G., 2019. Xgboost ve Karar Ağacı Tabanlı Algoritmaların Diyabet Veri Setleri Üzerine Uygulaması, Yüksek Lisans Tezi, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Yıldız, K., 2017. Demir Çelik Metalurjisi. Sakarya.
- Zhan, D.P., Zhang, H, Jiang, Z., 2010. Prediction Model of Magnesium Powder Consumption During Hot Metal Pre-Desulfurization. *China Metallurgy*.