



**T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANABİLİM DALI**

YÜKSEK LİSANS TEZİ

**XRD'DE X-IŞINI KIRINIM VERİLERİNİN YAPAY ZEKA
YÖNTEMLERİ İLE ANALİZİ**

Özcan Ali KALKAN

**Eylül-2024
BATMAN**

T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANA BİLİM DALI

YÜKSEK LİSANS TEZİ

XRD'DE X-IŞINI KIRINIM VERİLERİNİN YAPAY ZEKA
YÖNTEMLERİ İLE ANALİZİ

Özcan Ali KALKAN

Danışman
Prof. Dr. Ömer Faruk ERTUĞRUL

Eylül-2024
BATMAN

TEZ KABUL VE ONAYI

Özcan Ali KALKAN tarafından hazırlanan “XRD’DE X-IŞINI KIRINIM VERİLERİNİN YAPAY ZEKA YÖNTEMLERİ İLE ANALİZİ” adlı tez çalışması 05/09/2024 tarihinde aşağıdaki jüri tarafından oy birliği ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgi Teknolojileri Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Doç. Dr. Cafer BUDAK

.....

Danışman

Prof. Dr. Ömer Faruk ERTUĞRUL

.....

Üye

Dr. Öğr. Üyesi Abdulkerim ÖZTEKİN

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Özcan Ali KALKAN

05/09/2024

ÖZET

YÜKSEK LİSANS TEZİ

XRD'DE X-IŞINI KIRINIM VERİLERİNİN YAPAY ZEKA YÖNTEMLERİ İLE ANALİZİ

Özcan Ali KALKAN

Danışman: Prof. Dr. Ömer Faruk ERTUĞRUL

2024, 43 Sayfa

Jüri

Danışmanın Prof. Dr. Ömer Faruk ERTUĞRUL

Doç. Dr. Cafer BUDAK

Dr. Öğr. Üyesi Abdülkerim ÖZTEKİN

Bu tezde, X-Işını Kırınımı (XRD) verilerinin yapay sinir ağları (ANN) ve k-En Yakın Komşu (kNN) algoritmaları kullanılarak analiz edilmesi incelenmiştir. XRD, kristal yapılarının belirlenmesinde yaygın olarak kullanılan güçlü bir teknik olmasına rağmen, karmaşık ve doğrusal olmayan verilerin analizi geleneksel yöntemlerle sınırlı kalmaktadır. Bu çalışmada, XRD verilerinin analizine yapay zeka temelli makine öğrenimi tekniklerinin entegre edilmesiyle, bu sınırlamaların aşılması hedeflenmiştir. Tez kapsamında, Erzurum ve çevresinden alınan volkanosedimanter kayaç örneklerinden elde edilen XRD verileri üzerinde ANN ve kNN algoritmaları uygulanmıştır. ANN, malzeme özelliklerini yüksek doğrulukla tahmin edebilmiş ve kristal yapılar arasındaki karmaşık ilişkileri etkili bir şekilde modellemiştir. KNN algoritması ise, XRD verilerinde faz ve kristal yapıların sınıflandırılmasında başarıyla kullanılmıştır. Farklı K değerleri ile yapılan deneyler, küçük K değerlerinin yerel yapılara daha duyarlı olduğunu, büyük K değerlerinin ise daha genelleştirilebilir tahminler sunduğunu ortaya koymuştur. Ayrıca, model performansını değerlendirmek için gelişmiş çapraz doğrulama teknikleri kullanılmıştır. Bu teknikler, aşırı öğrenme riskini azaltarak, modelin genelleme yeteneğini artırmıştır. Çalışmada kullanılan ANN ve kNN modelleri, faz değişimlerinin tespitinde ve kristal yapıların sınıflandırılmasında önemli başarılar elde etmiştir. Bu tez, XRD verilerinin analizinde yapay zeka temelli yöntemlerin uygulanabilirliğini göstermiş ve bu alandaki bilimsel araştırmaların daha derinlemesine ve doğru sonuçlara ulaşmasına katkı sağlamıştır.

Anahtar Kelimeler: Yapay Sinir Ağları, k-En Yakın Komşu, X-Işını Kırınımı, Makine Öğrenmesi, Çapraz Doğrulama, Kristal Yapı Analizi

ABSTRACT

MS THESIS

ANALYSIS OF X-RAY DIFFRACTION (XRD) DATA USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Özcan Ali KALKAN

Batman University Graduate Education Institute

Information Technologies Department of Science

Advisor: Prof. Dr. Ömer Faruk ERTUĞRUL

2024, 43 Pages

Jury

Advisor Assoc. Prof. Dr. Ömer Faruk ERTUĞRUL

Doç. Dr. Cafer BUDAK

Dr. Öğr. Üyesi Abdülkerim ÖZTEKİN

This thesis investigates the analysis of X-Ray Diffraction (XRD) data using artificial neural networks (ANN) and k-Nearest Neighbors (kNN) algorithms. While XRD is widely used to determine crystal structures, analyzing complex and nonlinear data remains a challenge with traditional methods. In this study, machine learning techniques based on artificial intelligence are integrated into the analysis of XRD data to overcome these limitations. XRD data obtained from volcanosedimentary rock samples collected from the Erzurum region were analyzed using ANN and kNN algorithms. ANN demonstrated high accuracy in predicting material properties and effectively modeled complex relationships between crystal structures. The kNN algorithm, on the other hand, was successfully employed in classifying phases and crystal structures within the XRD data. Experiments with different K values revealed that smaller K values were more sensitive to local structures, while larger K values provided more generalizable predictions. Advanced cross-validation techniques were used to evaluate model performance, reducing the risk of overfitting and improving the generalization capability of the models. The ANN and kNN models applied in this study achieved significant success in detecting phase transitions and classifying crystal structures. This thesis demonstrates the applicability of artificial intelligence-based methods in the analysis of XRD data and contributes to more in-depth and accurate outcomes in scientific research within this field.

Keywords: Artificial Neural Networks, k-Nearest Neighbors, X-Ray Diffraction, Machine Learning, Cross-Validation, Crystal Structure Analysis

ÖNSÖZ

Öncelikle, bu zorlu süreçte her zaman yanımda olan ve bana her türlü desteği sağlayan sevgili aileme, özellikle de canım anneme sonsuz teşekkürlerimi sunuyorum. Tezimin hazırlanmasında bana rehberlik eden çok değerli danışmanım Prof. Dr. Ömer Faruk Ertuğrul'a ve katkılarından dolayı tüm jüri üyelerine en içten teşekkürlerimi iletiyorum. Ayrıca, bu süreçte bana moral ve motivasyon sağlayan arkadaşlarıma da minnettarım.

Özcan Ali KALKAN
BATMAN-2024



İÇİNDEKİLER

ÖZET	v
ÖNSÖZ	vii
İÇİNDEKİLER	viii
SİMGELER VE KISALTMALAR	ix
ŞEKİLLER DİZİNİ	x
TABLolar DİZİNİ	xi
1. GİRİŞ	1
2. LİTERATÜR ARAŞTIRMASI	3
3. MATERYAL VE YÖNTEM	6
3.1. Veri Setinin Hazırlanması	6
3.1.1. Bölgenin jeolojisi	6
3.1.2. Laboratuvar çalışmaları	7
3.2. Yöntem ve Veri Setinin Hazırlanması	10
3.2.1. Bragg yasası	10
3.2.2. Rietveld yöntemi	12
3.3. Veri Setinin İşlenmesi	12
3.3.1. ANN generalized regression	12
3.3.2. K-Nearest neighbors (KNN)	14
3.3.3. Çapraz doğrulama yöntemleri	16
3.4. Kullanılan Kodlar ve Yöntemler	17
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA	21
4.1. KNN ve XRD	21
4.2. Çapraz Doğrulama, Transfer Fonksiyonları ve Nöron Sayılarının ANN Sonuçlarına Etkisi	29
4.3. Karşılaştırmalı Analizler	33
5. SONUÇLAR ve ÖNERİLER	36
KAYNAKLAR	38

SİMGELER VE KISALTMALAR

Simgeler

XRD: X-Işını Kırınımı

ANN: Yapay Sinir Ağı (Artificial Neural Networks)

KNN: K En Yakın Komşu (K-Nearest Neighbors)

RMSE: Kök Ortalama Kare Hatası (Root Mean Squared Error)

MAE: Ortalama Mutlak Hata (Mean Absolute Error)

KFold: K Katmanlı Çapraz Doğrulama Yöntemi



ŞEKİLLER DİZİNİ

Şekil 3.1 Pasinler Yöresi Jeolojik Haritası, Qe ve Qy örneklerin alındığı volkanoklastik bölgeleri göstermektedir.	6
Şekil 3.2 Kayaç örneklerinin XRD çekimi için öğütülmesi.	8
Şekil 3.3 XRD çekimi için hazırlanan, agat havanda öğütülmüş kayaç	8
Şekil 3.4 Xcalibur Oxford CCD Difraktometre cihazı (Perugia Üniversitesi).	9
Şekil 3.5 Tek kristal XRD kırınımı çalışma prensibi (Gentili, 2011'den uyarlanmıştır.	10
Şekil 3.6 Bragg yasası ile minerali oluşturan atomların oluşturduğu düzlemler arası mesafe ölçümü (Yeniyol, 2004'ten uyarlanmıştır)	11
Şekil 4.1 Makine öğrenmesi esnasında hata veren eğitim sonuçları.	24
Şekil 4.2 ANN basit şematik görünümü.	25
Şekil 4.3 12 parçalı çapraz doğrulamada en iyi performans.	26
Şekil 4.4 Çapraz doğrulama öncesi regresyon analizi.	27
Şekil 4.5 Bazı tuzların 1.0 atm basınçta sudaki çözünürlüklerinin sıcaklıkla değişimi..	28
Şekil 4.6 ANN ağının eğitim, doğrulama ve test aşamalarındaki hata çıktıları histogram eğrisi.....	29
Şekil 4.7 Çapraz doğrulama sonrası regresyon analiz grafikleri.	31

TABLÖLAR DİZİNİ

Tablo 4.1 Farklı Nöron Sayıları ve aktivasyon fonksiyonları ile MAE sonuçları	31
Tablo 4.2 Farklı Nöron Sayıları ve aktivasyon fonksiyonları ile MSE sonuçları.....	32
Tablo 4.3 Farklı Nöron Sayıları ve aktivasyon fonksiyonları ile RMSE sonuçları	32
Tablo 4.4 Modellerin 8 ve 12 Parça Temelli Performans Karşılaştırması.....	34
Tablo 4.5 Farklı Modellerin RMSE ve MAE Değerleri ile Karşılaştırılması.....	35



1. GİRİŞ

Bu tez çalışmasında, Erzurum Pasinler Platosu'ndan alınan volkanosedimanter kayaç örnekleri üzerinde gerçekleştirilen X-Işını Kırınımı (XRD) analizleri ve bu verilerin makine öğrenimi teknikleri ile işlenmesi ele alınmaktadır. XRD, kristal yapıların ve düzenlemelerinin incelenmesinde kritik bir teknoloji olarak öne çıkmaktadır. Ancak, XRD verilerinin karmaşık ve doğrusal olmayan yapısı, geleneksel analitik yöntemlerle tam anlamıyla çözümlenememektedir. Bu noktada makine öğrenimi, bu karmaşık veri setlerini işlemek ve anlamlandırmak için güçlü bir alternatif olarak karşımıza çıkmaktadır.

Bu çalışmanın temel amacı, XRD verilerinin makine öğrenimi ve yapay sinir ağları (ANN) gibi ileri düzey teknikler kullanılarak daha doğru ve güvenilir bir şekilde analiz edilebileceğini göstermektir. Özellikle, ANN Genelleştirilmiş Regresyon ve K-En Yakın Komşular (kNN) algoritmalarının, XRD verilerinin karmaşık yapısını modelleme ve tahmin etme potansiyelleri değerlendirilecektir. Bu tekniklerin kullanımı, kristalografi ve malzeme bilimi gibi alanlarda veri analizine yeni bir bakış açısı getirecek ve elde edilen sonuçların doğruluğunu önemli ölçüde artıracaktır.

XRD verilerinin analizinde karşılaşılan en önemli zorluklardan biri, verilerin karmaşık doğası ve düşük sinyal-gürültü oranlarıdır. Geleneksel yöntemler, bu tür karmaşık veri yapılarında yetersiz kalabilmekte ve çoğu zaman düşük doğrulukla sonuçlanmaktadır. Bu tezde önerilen yöntemler, özellikle makine öğrenimi tabanlı yaklaşımlar, XRD verilerindeki gizli ilişkileri ortaya çıkarmak ve bu doğrusal olmayan ilişkileri başarılı bir şekilde modellemek üzere tasarlanmıştır.

ANN, XRD verilerinin analizinde yüksek doğrulukla malzeme özelliklerini tahmin edebilme yeteneği sunarken, KNN algoritması, veri noktaları arasındaki sınıflandırma ve regresyon problemlerinde etkili sonuçlar sağlamaktadır. ANN, veri setlerinde doğrusal olmayan ilişkileri modelleme kabiliyeti ile öne çıkarken, KNN algoritması özellikle sınıflandırma problemlerinde başarılı sonuçlar vermektedir. Her iki algoritmanın da avantajları, XRD analizleri için kritik olan faz analizi, kristal yapı tanımlama ve malzeme özelliklerinin tahmini gibi işlemlerde önemli katkılar sağlayacaktır.

Çalışmamızın en önemli yeniliği, geleneksel yöntemlerin sınırlamalarını aşarak makine öğrenimi tekniklerinin XRD analizine uygulanabilirliğini ortaya koymaktır. Özellikle, volkanosedimanter kayalardaki eser miktarda bulunan minerallerin tespitini kolaylaştırmak amacıyla karbonat gürültüsünün eliminasyonu hedeflenmiştir. Bu sayede, daha doğru ve güvenilir mineral analizleri gerçekleştirilmesi amaçlanmıştır.

Bu tez çalışmasının sonuçları, XRD verilerinin analizinde makine öğrenimi tekniklerinin kullanılabilirliğini ve bu yöntemlerin malzeme bilimi araştırmalarına katkı sağlayabilecek potansiyelini ortaya koyarak, bilimsel literatüre önemli bir katkı sağlayacaktır.



2. LİTERATÜR ARAŞTIRMASI

Bu bölümde X ışınlarının keşfedilişi, çeşitli bilim dallarına katkılarına değinilmiş ve yerbilimlerinde uygulanmaya başladığı 1950'li yıllardan günümüze uygulanış biçimleri ile ilgili detaylı kaynaklar özetlenerek teknikleri üzerinde durulmuştur.

X ışınlarının keşfi, bilim tarihinde dönüm noktalarından biridir ve çeşitli bilim dallarına yaptığı katkılar son derece önemlidir. X ışınları, 1895 yılında Wilhelm Conrad Röntgen tarafından keşfedilmiştir. Röntgen, vakum tüpleriyle gerçekleştirdiği deneylerde, bazı ışınların maddelerden geçerek bir ekran üzerinde görüntü oluşturduğunu fark etmiş ve bu ışınları X ışınları olarak adlandırmıştır. Bu keşif, özellikle tıp alanında, iç organların zarar görmeden görüntülenebilmesine olanak tanıyan radyolojinin doğuşunu sağlamıştır (Cullity, 1956).

X ışınları, sadece tıp alanında değil, kristal yapıların analizinde de önemli bir rol oynamaktadır. 1914 yılında W. L. Bragg ve J. J. Thomson tarafından geliştirilen X-ışını kristalografisi, kristallerin atomik düzeyde incelenmesine olanak sağlamıştır. Bu teknik, kristal yapıları çözme ve analiz etme süreçlerinde yaygın olarak kullanılmaktadır (Bragg & Thomson, 1914). Bu yöntem, minerallerin kimyasal bileşimlerinin belirlenmesi, kristal yapıların anlaşılması ve çeşitli kristal defektlerinin tespit edilmesi gibi alanlarda büyük bir ilerleme sağlamıştır (Cullity, 1956; Nuffield, 1996).

1950'li yıllardan itibaren X ışınları, yer bilimlerinde kullanılmaya başlanmış ve bu dönemde jeolojik numunelerin mineral bileşimi ve yapısının incelenmesinde önemli bir araç haline gelmiştir. Mineralojide, kristal yapıların atomik düzenlemelerinin belirlenmesinde kullanılan X-ışını kırınımı (XRD) yöntemleri, jeoloji alanında kayaçların mineral içeriklerinin, dokularının ve jeolojik süreçlerin incelenmesinde geniş çaplı kullanılmıştır (Erdoğan & Soytürk, 1974; Hillier, 2000). Özellikle metamorfik ve tortul kayaçların oluşum süreçlerinin anlaşılmasında bu yöntem kritik bir öneme sahiptir. X ışınları ile yapılan bu incelemeler, jeolojik formasyonların evrimi ve jeolojik süreçlerin daha derinlemesine anlaşılmasına katkı sağlamaktadır.

X ışınları, yer bilimlerinin yanı sıra petrol ve madencilik endüstrisinde de önemli bir yer tutmaktadır. Bu alanlarda, X ışınları kayaçların iç yapısının incelenmesi ve hidrokarbon rezervlerinin tespit edilmesi için kullanılmıştır. Ayrıca, maden cevherlerinin mineral bileşimlerinin analiz edilmesi, metal miktarlarının belirlenmesi ve bu madenlerin ekonomik olarak değerlendirilmesi konularında da önemli bir araç olarak kullanılmıştır.

(Gentili, 2011; Hanawalt, 1976). Bu teknikler, petrol ve madencilik operasyonlarının verimliliğini artırarak bu sektörlere önemli katkılar sağlamıştır.

X-ışını kırınımı (XRD) yönteminin gelişmesiyle birlikte, bilim insanları daha karmaşık yapıların ve malzemelerin analizine yönelmiştir. Özellikle kristal yapılarının daha doğru analiz edilebilmesi için kullanılan Rietveld yöntemi ve bu yöntemin geliştirilmesi, malzeme bilimi alanında önemli bir adım olmuştur. Rietveld analizi, kristal yapılarının atomik düzeyde incelenmesini sağlar ve böylece malzemelerin fiziksel özelliklerini daha iyi anlamamıza yardımcı olur (Dollase, 1986). Bu yöntem, özellikle kil mineralleri gibi kompleks yapıları incelemek için kullanılmış ve bilim dünyasında geniş çapta kabul görmüştür (Hillier, 2000).

Nuffield (1996) tarafından yayımlanan X-ışını Kırınımı Yöntemleri kitabı, bu alandaki en önemli kaynaklardan biridir. Kitap, X ışınlarının kristal yapılar üzerindeki etkilerini, numune hazırlığı süreçlerini ve difraksiyon desenlerinin nasıl analiz edileceğini detaylandırarak bilim insanlarına rehberlik etmektedir. Nuffield'ın kitabı, özellikle laboratuvar ortamında X-ışını kırınımı deneylerini gerçekleştiren araştırmacılar için vazgeçilmez bir başvuru kaynağı olarak kabul edilir. Kitap, ayrıca X-ışını kırınım verilerinin işlenmesi ve modelleme süreçlerini de kapsamlı bir şekilde ele alır.

Diğer önemli kaynaklar arasında Cullity (1956) tarafından kaleme alınan Elements of X-Ray Diffraction kitabı yer almaktadır. Cullity, X-ışını kırınımı konusunu temel prensiplerden başlayarak uygulamalı tekniklerle açıklamış ve bu yöntemin malzeme bilimi ve jeoloji gibi farklı disiplinlerdeki uygulamalarını detaylandırmıştır. Bu kaynak, özellikle X ışını kırınımı konusunda derinlemesine bilgi edinmek isteyen öğrenciler ve araştırmacılar için önemli bir başvuru kaynağıdır.

Jeolojik araştırmalarda X ışınlarının kullanımıyla ilgili gelişmeler, mineral yapılarının detaylı bir şekilde incelenmesine olanak sağlamış, kayaçların oluşum süreçlerinin anlaşılmasını kolaylaştırmıştır. X-ışını difraksiyonu (XRD) sayesinde tortul kayaçların tabakalı yapıları, mineral içeriği ve oluşum süreçleri hakkında daha fazla bilgi edinilmiştir. Aynı zamanda, XRD ile elde edilen difraksiyon desenleri, malzemelerin kristal yapılarını anlamada kullanılan önemli bir araç olmuştur (Scarlett et al., 2001).

Petrol ve madencilik endüstrisinde, X ışınları ile yapılan analizler, kayaçların iç yapılarının incelenmesinde ve maden cevherlerinin metal içeriklerinin belirlenmesinde önemli bir rol oynamaktadır. X ışınlarının kullanımı, bu sektörlerde yapılan araştırmalarda daha derinlemesine ve güvenilir sonuçlar elde edilmesine olanak tanımaktadır. Özellikle hidrokarbon rezervlerinin belirlenmesi ve maden yataklarının

ekonomik deęerinin tespit edilmesinde bu teknikler geniř apta kullanılmaktadır (Hanawalt, 1976; Gentili, 2011).

XRD ile elde edilen bir difraksiyon deseni rneęi Őekil 1'de gsterilmiřtir. Difraksiyon deseninde grlen her bir pik, bir kristal dzleminden yansıyan X ıřınlarını temsil etmektedir. Bu desen, kristal yapının atomik dzenlemeleri hakkında bilgi verir ve kristalin yapısal zelliklerini anlamada nemli bir aratır.

Őekil 3.1'de XRD ile elde edilen difraksiyon deseni. Her bir pik, kristal yapısının belirli bir dzleminden yansıyan X ıřınlarını gstermektedir.

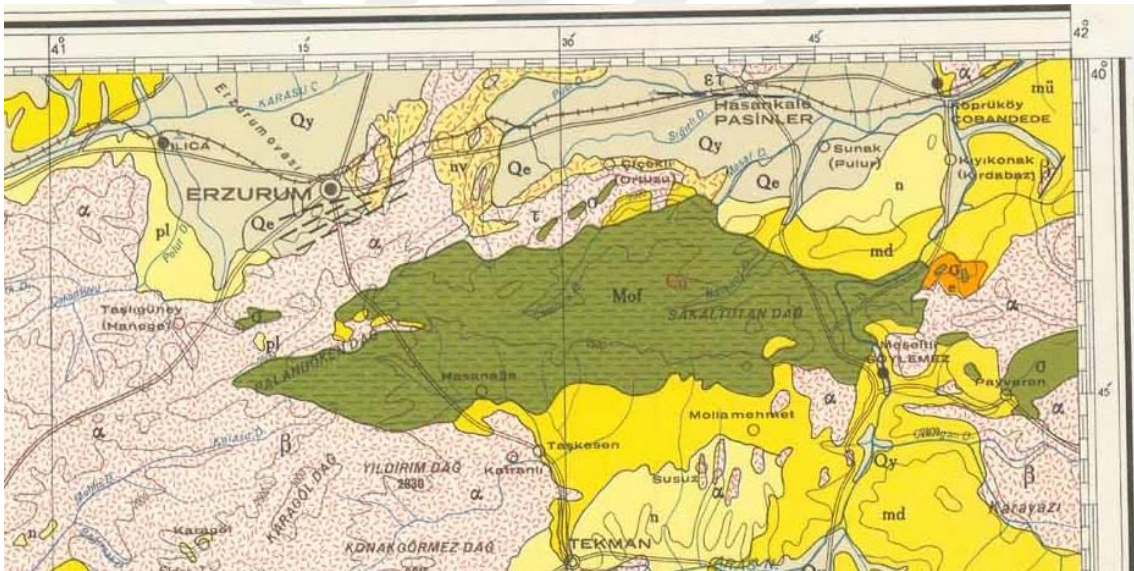
Sonuç olarak, X ıřınlarının keřfi ve kullanımı, bilim dnyasında devrim nitelięinde bir geliřme olarak kabul edilmiřtir. Mineralojiden jeolojiye, petrol ve madencilik endstrilerinden malzeme bilimine kadar geniř bir yelpazede uygulama bulan bu yntemler, bilimsel arařtırmaların daha doęru ve gvenilir olmasına katkı saęlamaktadır. Cullity (1956) ve Nuffield (1996) gibi nemli kaynaklar, X-ıřını kırınımının nasıl kullanıldığını ve bu tekniklerin geliřtirilmesini detaylı bir Őekilde ele alarak bilimsel arařtırmalara ıřık tutmaktadır. Bu yntemler sayesinde, Dnya'nın i yapısı, tarihi ve jeolojik sreleri daha iyi anlařılmakta ve bilimsel arařtırmaların doęruluęu artırılmaktadır.

3. MATERYAL VE YÖNTEM

3.1. Veri Setinin Hazırlanması

3.1.1. Bölgenin jeolojisi

Bu tez çalışmasında incelenen örnekler, Erzurum ili Pasinler yöresinde yer alan Neojen yaşlı volkanik-volkanosedimanter birimlere aittir. Örneklerin alındığı saha, Erzurum şehir merkezini kapsayan Erzurum İ-46 paftası, Erzurum İ-47 paftasının kuzeyi, Tortum H-47 ve Tortum H-48 paftalarının güneyini içine alan geniş bir bölgedir. Bölgenin jeolojik yapısı, Üst Miyosen yaşlı Varto Grubu'na ait Zirnik formasyonu, Varto formasyonu, Selçuk ignimbiriti, Varto lavları ve Alt Pliyosen yaşlı Hamurpet lavından oluşmaktadır.



Şekil 3.1 Pasinler Yöresi Jeolojik Haritası, Qe ve Qy örneklerin alındığı volkanoklastik bölgeleri göstermektedir.

Araştırma alanı, Türk-İran Yüksek Platosu'nun batı kısmında yer almakta olup, Torid, Anatolid ve Pontit'lerin birleştiği bir bölgede konumlanmaktadır. Önceki araştırmalara göre, Doğu Anadolu, kuzeyde Pontit ve güneyde Toros kütleleri tarafından sıkıştırılan merkezi bir plato olarak tanımlanmıştır. Bu çalışma alanı ise, Doğu Anadolu yığılma karmaşığına karşılık gelmektedir.

Hınis metaofiyoliti, Akdağ metamorfileri, ofiyolitli granitik kayalar ve ofiyolitli karışık granitik kayalar, Mesozoyik'in sonu ile Alt Miyosen dönemine ait birimler tarafından açıl uyumsuzlukla örtülmektedir. Bu birimler, Orta Miyosen yaşlı Bingöl Dağı grubu tarafından üstlenmiştir. Üst Miyosen yaşlı Varto Grubu ise Bingöl Dağı

grubunun üzerine gelerek gölssel çökeltiiler, volkanoklastikler ve lavlardan oluşmuştur. Bingöl Yanardağı'nın püskürükleri olan bu birimler, Alt Pliyosen yaşlı Hamurpet lavları ve Orta-Üst Pliyosen yaşlı gölssel çökeltiiler tarafından örtülmüştür.

Bölgeye ait ilk jeolojik harita, Pamir ve Baykal (1943) tarafından hazırlanmıştır. Bu çalışmada, bölgedeki volkanik kayaçların Miyosen döneminin son safhasına ait olduğu belirtilmiştir. Erentöz (1945) bölgedeki petrol emarelerini araştırmak amacıyla çalışmalar yürütmüştür. Gattinger (1954), bölgenin 1/100.000 ölçekli jeolojik etüdünü yaparak, önemli kayaç türleri ve stratigrafiyi belirlemiştir.

Altınlı (1963), Türkiye Jeoloji Haritası çalışmasında, bölgenin 1/500.000 ölçekli haritasını hazırlamış ve andezit, trakit ve bazalt kayaçlarının Üst Miyosen yaşlı olduğunu belirlemiştir. Akkus (1965), Erzurum çevresindeki volkanik kayaçları inceleyerek stratigrafik birimlerin detaylarını ortaya koymuştur. Bu çalışmalar, volkanik kayaçların yaşının Pre-Alt Miyosen olduğunu göstermiştir.

Erdogan ve Soyturk (1974), bölgedeki önceki çalışmaların bir derlemesini yaparak 1/25.000 ölçekli haritalar hazırlamıştır. Tokel (1979), Erzurum-Kars yöresindeki Neojen çökeltileriyle ilgili volkanizmayı incelemiş ve Üst Miyosen ile Üst Pliyosen volkanitleri arasındaki mineralojik ve kimyasal farklılıkları ortaya koymuştur. Innocenti ve diğerleri (1982), Erzurum-Kars bölgesindeki volkanizmanın, Arap ve Anadolu levhalarının çarpışması sonucunda geliştiğini ileri sürmüştür.

Yılmaz ve diğerleri (1986, 1989), Erzurum güneydoğusunda (Hınıs-Tekman-Karayazı ve Sakaltutan Dağı yöresi) yaptıkları çalışmalarda, ofiyolitli birimleri farklı ortamların ürünü olan üç yapısal birime ayırmışlardır. Şaroğlu ve Yılmaz (1986), Doğu Anadolu'da Neotektonik dönemde gerçekleşen jeolojik evrimi ve havza modellerini incelemiş ve kıvrımlar, bindirmeler, doğrultu atımlı faylar gibi yapısal unsurların bu bölgede geliştiğini belirtmişlerdir. Erzurum-Pasinler-Horasan havzasının doğrultu atımlı fayların etkisi altında gelişen bir dağ arası havza olduğu belirtilmiştir.

3.1.2. Laboratuvar çalışmaları

Örnekler, agat havanda toz haline getirilmiş (Şekil 3.2) ve İtalya'da Umbria Eyaleti Perugia Üniversitesi Jeoloji Bölümü laboratuvarlarında (Universita Studi Degli Perugia Schenze della Terra) XRD çekimleri yapılmıştır.



Şekil 3.2 Kayaç örneklerinin XRD çekimi için öğütülmesi.



Şekil 3.3 XRD çekimi için hazırlanan, agat havanda öğütülmüş kayaç

XRD çekimleri için Xcalibur Oxford CCD marka difraktometre kullanılmıştır (Şekil 3.3).



Şekil 3.4 Xcalibur Oxford CCD Difraktometre cihazı (Perugia Üniversitesi).

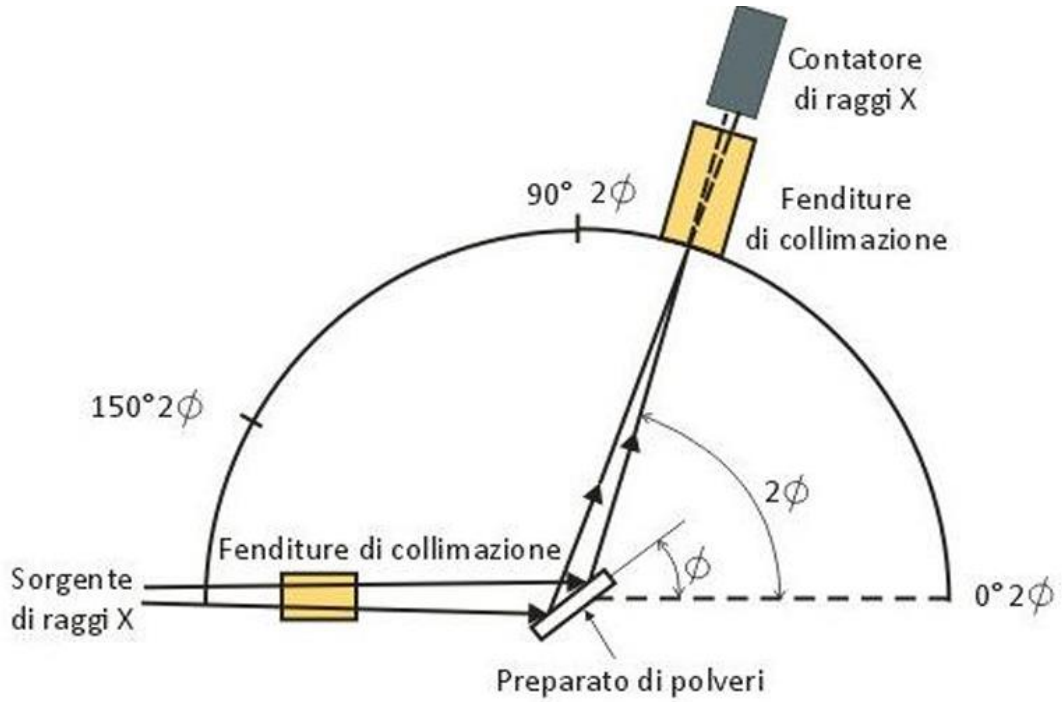
Tek Kristal Difraksiyonu:

Tek kristal difraksiyonu, tek bir mineral veya birkaç yüz mikrometre boyutlarındaki bir kristalin difraksiyon etkilerini ifade eder. X ışınları demeti, kristal yapısıyla etkileşime girerek, X ışınları dedektörü tarafından kaydedilen difraksiyon etkilerini üretir. Modern tekniklerle yapılan ölçümler, yüksek hassasiyetle birçok yansımanın ölçülmesini sağlayarak, manuel analizle yapılan eski deneylere göre doğruluğu önemli ölçüde artırmıştır. Bu çalışmada, Perugia Üniversitesi Yer Bilimleri Bölümü'nde kullanılan Xcalibur Oxford CCD difraktometresi (Şekil 5) ile difraksiyon verilerinin yoğunluk ölçümleri yapılarak yapısal rafinasyon için veri toplanmıştır.

Veri toplama işlemleri oda sıcaklığında, 50 kV ve 35 mA voltajında gerçekleştirilmiştir. Cihazın bileşenleri şunlardır:

X ışınları tüpü:

Tüp, tungsten bir filament içeren cam bir kılıf ve ısı etkili iyonik emisyon yoluyla elektron yayarak ışık üreten bir Mo (Molibden) hedef içeren anotla desteklenmiştir.



Şekil 3.5 Tek kristal XRD kırınımı çalışma prensibi (Gentili, 2011'den uyarlanmıştır).

Scintillation türü dedektör:

Bu dedektörde, gelen X ışınları bir sodyum iyodür (Na) kristaline çarparak fotoelektron üretir. Kristalin arka kısmı fotomultiplier ile desteklenir ve çıkan darbe birkaç voltur. Bu tip dedektörler, gelen radyasyonun yoğunluğuna orantılı olarak ışığın yoğunluğunu ölçer ve saniyede 105 darbeye kadar sayım yapabilir.

3.2. Yöntem ve Veri Setinin Hazırlanması

Çalışmanın bu kısmında, yerbilimciler tarafından 65 yıldır kullanılmakta olan temel kanunlar açıklanmış, Bragg Yasası ve Rietveld yöntemi ayrıntılı olarak ele alınmıştır.

X-ışını kırınımı (XRD), kristal yapıların belirlenmesinde kullanılan güçlü bir analitik tekniktir. XRD'nin temel prensibi Bragg Yasası'na dayanmaktadır. Bu bölümde, bu temel yasalar açıklanmış ve kabul gören Bragg Yasası ile Rietveld yöntemi tanımlanmıştır.

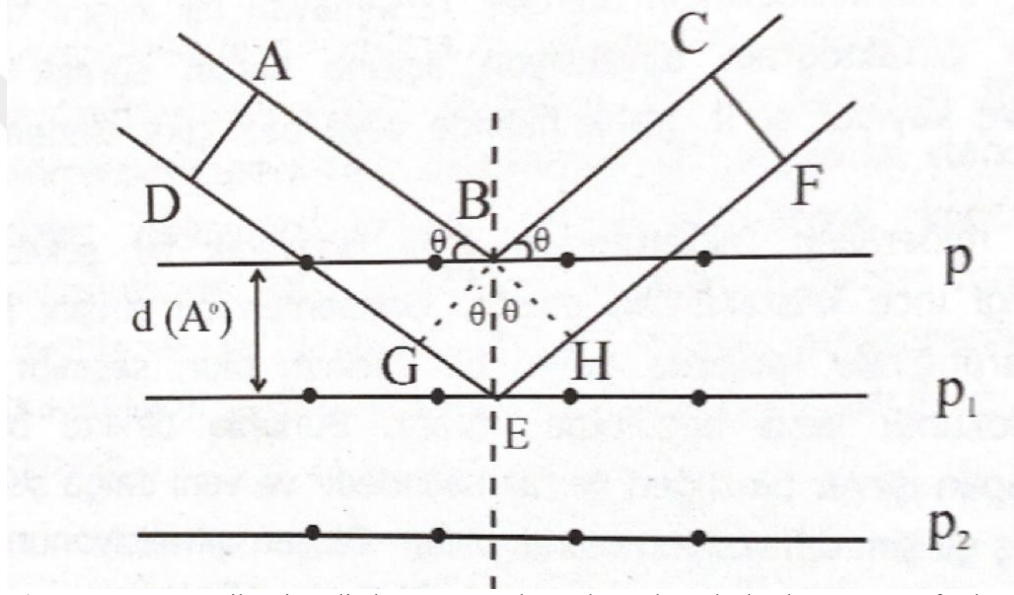
3.2.1. Bragg yasası

Bragg Yasası, kristal yapıdaki atomlar arasındaki düzenli uzaklıkların, X-ışınlarının kırınım desenlerine nasıl katkı sağladığını tanımlar. Kırınım desenleri ve yoğunlukları bu yasaya dayanarak hesaplanır. Bragg Yasası şu şekilde formüle edilir:

$$2d \sin(\theta) = n \lambda \quad (3.1)$$

Burada:

- d , kristal yapıdaki atomlar arası mesafeyi (d-spacing),
- θ , kırınım açısını,
- n , kırınımın sıra numarasını (1, 2, 3, ...),
- λ , X-ışınının dalga boyunu ifade eder.



Şekil 3.6 Bragg yasası ile minerali oluşturan atomların oluşturduğu düzlemler arası mesafe ölçümü (Yeniyol, 2004'ten uyarlanmıştır)

Bu denklem, kristal yapıdaki atomlar arasındaki mesafe (d) ile gelen X-ışınlarının açısı (θ) arasındaki ilişkiyi gösterir. X-ışınlarının belirli bir açıda kırınıma uğraması sonucu, karakteristik kırınım desenleri elde edilir. Bu desenler, kristal yapıların analiz edilmesini sağlar.

XRD deneylerinde, X-ışınları bir kristal yapıya çarptığında, bu yapıdaki atomlar X-ışınlarını farklı açılarda kırarak karakteristik kırınım desenlerini oluştururlar. Bu kırınım desenleri, Bragg Yasası'na göre belirli açılarda (kırınım açıları) belirli dalga boylarında görülür. Bu desenler, kristal yapıdaki atomlar arasındaki mesafelerin ve düzenlemelerin analiz edilmesine olanak tanır.

Bragg Yasası, X-ışını kırınımının temel matematiksel formülasyonunu sağlar ve XRD deneylerinde elde edilen verilerin analizinde önemli bir rol oynar. Bu prensip

sayesinde, XRD, kristal yapıların belirlenmesi, mineral tanımlama, kristal yapının incelenmesi ve malzemelerin karakterizasyonu gibi birçok alanda kullanılmaktadır. Bragg Yasası, X-ışını kırınımının temel prensibini tanımlayan bir fiziksel yasadır. Bu yasa, kristal yapıların X-ışınlarıyla etkileşimini açıklar. Aynı zamanda, Rietveld yöntemi, X-ışını kırınım verilerinin analizinde kullanılan bir yöntemdir.

3.2.2. Rietveld yöntemi

Rietveld yöntemi, X-ışını kırınım verilerinin analizinde kullanılan bir yöntemdir. Bu yöntem, deneysel X-ışını kırınım desenlerini kullanarak kristal yapıların detaylı bir şekilde karakterize edilmesini sağlar. Deneysel verilerle teorik modellerin karşılaştırılması yoluyla yapısal parametreler ayarlanır. Rietveld yöntemi, kristal yapıların mikroskobik düzeyde anlaşılmasında önemli bir araçtır ve yapı parametrelerinin doğru bir şekilde hesaplanmasını sağlar.

Sonuç olarak, Bragg Yasası X-ışını kırınımının temel prensiplerini tanımlarken, Rietveld yöntemi X-ışını kırınım verilerinin analizinde kullanılan matematiksel bir modelleme yöntemidir. Bu iki yöntem bir arada kullanılarak, kristal yapıların analizi yapılmaktadır.

3.3. Veri Setinin İşlenmesi

3.3.1. ANN generalized regression

Yapay Sinir Ağları (ANN), insan beyninin sinir hücreleri arasındaki bağlantıları taklit eden ve veri içindeki non-lineer (doğrusal olmayan) ilişkileri modelleme yeteneği nedeniyle makine öğrenmesi ve derin öğrenme alanında sıkça tercih edilen bir yöntemdir. ANN'ler, birçok bilimsel alanda karmaşık veri kümelerini analiz etme ve bu verilerden anlamlı sonuçlar çıkarma konusunda güçlü bir araç olarak kabul edilmektedir. Bu çalışmada, ANN'nin Genelleştirilmiş Regresyon modeli (ANN Generalized Regression) kullanılmıştır. Bu model, özellikle kristalografi ve malzeme bilimi alanındaki karmaşık verilerin işlenmesinde etkili bir çözüm sunmaktadır.

ANN Kullanımı:

ANN, çok katmanlı yapıları sayesinde öğrenme işlemi sırasında verilerdeki karmaşık non-lineer ilişkileri keşfetme ve tahmin etme yeteneğine sahiptir. Bu özellik,

özellikle geleneksel regresyon modellerinin yetersiz kaldığı durumlarda devreye girer. Geleneksel modeller, genellikle veri içindeki lineer ilişkiler üzerinde çalışırken, ANN'ler non-lineer ilişkileri de modelleyebilir ve daha karmaşık veri kümelerinde yüksek doğruluk oranlarına ulaşabilirler. X-ışını difraksiyon verileri gibi karmaşık ve yüksek boyutlu veri setlerinde, geleneksel yöntemlerle modellemenin zorluğu göz önüne alındığında, ANN'ler önemli bir çözüm sunmaktadır.

Bu çalışmada, XRD veri setleri üzerinde ANN kullanılarak kristal yapıların incelenmesi amaçlanmıştır. Kristalografik verilerde bulunan atomik düzenlemeler, kırınım açılarındaki varyasyonlar ve yapıların yoğunlukları gibi veriler, sinir ağı modelleri tarafından analiz edilerek daha doğru tahminler elde edilebilir. ANN, öğrenme algoritmaları aracılığıyla verilerdeki karmaşık ilişkileri keşfeder ve bu sayede geleneksel yöntemlerle çözilemeyen problemleri başarıyla çözebilir.

ANN Genelleştirilmiş Regresyonun XRD Verilerine Uygulanması:

ANN Genelleştirilmiş Regresyon modeli, X-ışını difraksiyon (XRD) deneylerinden elde edilen verilerin analizinde kullanılabilecek en etkili yaklaşımlardan biridir. Genelleştirilmiş Regresyon modeli, belirli bir malzeme örneğinin yapısal özelliklerini tahmin etmek için verilerdeki karmaşık ilişkileri modelleyerek daha doğru sonuçlar elde edilmesine olanak tanır. XRD verilerinin doğası gereği, atomlar arası mesafeler, kristal yapılar ve yoğunluklar gibi çok boyutlu ve karmaşık özellikler içerir. Bu karmaşıklık, geleneksel analiz yöntemlerinin yetersiz kaldığı durumlarda ANN modellerinin daha etkili olmasını sağlar.

ANN modelleri, verilerin yapısal özelliklerine göre optimize edilerek daha doğru tahminlerde bulunabilir. Özellikle XRD verileri, kristalografik yapılar ve atom düzenlemeleri gibi karmaşık ilişkileri içerdiğinden, ANN Genelleştirilmiş Regresyon modelinin uygulanması, bu verilerin daha doğru ve hızlı bir şekilde analiz edilmesine yardımcı olur. Geleneksel yöntemlerle elde edilemeyen doğruluk oranlarına ulaşmak amacıyla ANN modellerinin eğitilmesi ve optimize edilmesi gerekmektedir. Bu süreçte kullanılan geri yayılım algoritmaları, modelin her adımda hata payını düşürerek daha iyi sonuçlar elde edilmesini sağlar.

ANN Genelleştirilmiş Regresyon modelinin XRD veri setlerine uygulanması, kristalografi biliminde daha doğru modelleme ve tahmin yapabilme kapasitesini artırır. Bu sayede, X-ışını difraksiyon deneylerinin daha doğru analiz edilmesi ve yorumlanması mümkün hale gelir. ANN modelleri, verilerin yapısal karmaşıklığını anlamak ve tahminlerde bulunmak için özellikle güçlü bir yöntem olarak öne çıkar. Bu çalışmada

kullanılan ANN Genelleştirilmiş Regresyon yöntemi, XRD verilerinde bulunan atomik düzenlemeleri ve kırınım desenlerini başarılı bir şekilde modelleyerek, malzemelerin kristal yapıları hakkında daha kesin sonuçlar elde edilmesine olanak tanımıştır.

Avantajlar ve Sınırlamalar:

ANN Genelleştirilmiş Regresyon modelinin en büyük avantajlarından biri, büyük ve karmaşık veri setlerinde verimli bir şekilde çalışabilmesi ve geleneksel yöntemlere göre daha yüksek doğruluk oranları sağlamasıdır. Özellikle non-lineer ilişkilerin bulunduğu veri kümelerinde, ANN modelleri bu ilişkileri başarıyla yakalayabilir ve modelleyebilir. Bununla birlikte, ANN modellerinin eğitimi zaman alıcı olabilir ve yeterli sayıda veri ile beslenmesi gerekmektedir. Ayrıca, modelin aşırı öğrenme (overfitting) yapmaması için dikkatli bir şekilde optimize edilmesi gereklidir. Bu süreçte, çapraz doğrulama teknikleri ve modelin düzenli olarak test edilmesi, aşırı öğrenmenin önüne geçmek için kritik öneme sahiptir.

Genel olarak, ANN Genelleştirilmiş Regresyon modeli, X-ışını difraksiyon verileri gibi karmaşık veri setlerinin analizinde güçlü bir yöntem olarak öne çıkmaktadır. Bu modelin uygulanması, kristal yapıların daha doğru bir şekilde anlaşılmasına ve malzeme özelliklerinin tahmin edilmesine olanak tanımaktadır.

3.3.2. K-Nearest neighbors (KNN)

KNN algoritması, makine öğrenimi alanında sınıflandırma ve regresyon problemlerini çözmek için kullanılan basit ancak güçlü bir yöntemdir. kNN, verilerdeki noktaların bir uzayda yer aldığını varsayarak, bu noktalara en yakın komşuların sınıf veya değerlerine göre tahminlerde bulunur. Bu bölümde, kNN algoritmasının temel çalışma prensipleri ve performansını etkileyen faktörler detaylı bir şekilde açıklanacaktır.

KNN algoritması, eğitim verisindeki her veri noktasını uzayda temsil eder ve yeni bir veri noktası için tahmin yaparken, bu noktaya en yakın K komşunun sınıfı veya değeri dikkate alınarak tahmin gerçekleştirilir. Bu nedenle kNN, "örnek tabanlı öğrenme" veya "tembel öğrenme" (lazy learning) olarak da bilinir, çünkü model eğitimi sırasında herhangi bir genel fonksiyon oluşturulmaz. Bunun yerine, her tahmin sırasında eğitim verisi tekrar tekrar kullanılır.

KNN algoritmasının başarısını etkileyen en kritik parametrelerden biri, K değeridir, yani kaç komşunun dikkate alınacağını belirleyen parametre. Küçük K değerleri (örneğin $K = 1, 3$) modelin daha lokal bilgilerle tahmin yapmasına olanak

tanırken, büyük K değerleri (örneğin $K = 10, 20$) modelin daha genel eğilimleri dikkate almasını sağlar. Ancak, bu iki yaklaşımın farklı avantajları ve dezavantajları bulunmaktadır.

Küçük K Değerlerinin Avantaj ve Dezavantajları:

Avantajlar:

- Küçük K değerleri, modelin verideki küçük ayrıntılara ve yerel yapılara duyarlılığını artırır.
- Özellikle sınıflandırma problemlerinde $K = 1$ seçilirse, model doğrudan en yakın komşunun sınıfını tahmin eder ve basit, hızlı sonuçlar verir.

Dezavantajlar:

- Gürültüye karşı duyarlıdır. Eğitim veri kümesinde bulunan hatalı veya aykırı veriler, modelin genel performansını olumsuz etkileyebilir.
- Küçük K değerleri, aşırı öğrenme (overfitting) riski taşır. Model, eğitim verisine çok sıkı uyum sağladığından, test verisindeki performansı düşebilir.

Büyük K Değerlerinin Avantaj ve Dezavantajları

Avantajlar:

- Büyük K değerleri, modelin veri kümesindeki genel yapıları ve trendleri daha iyi yakalamasını sağlar.
- Gürültü ve aykırı değerlerin etkisi azalır, çünkü daha fazla komşunun ortalaması alınarak daha dengeli bir tahmin yapılır.

Dezavantajlar:

- Hesaplama maliyeti artar. Daha fazla komşunun hesaplanması gerektiğinden model daha fazla işlem gücü ve zaman gerektirir.
- Sınıflandırma problemlerinde, büyük K değerleri sınıf sınırlarını daha belirsiz hale getirebilir ve sınır bölgelerinde sınıf tahminleri kararsızlaşabilir.

XRD Verilerine Uygulama: KNN ve Aykırı Değerler:

X-ışını difraksiyon (XRD) verileri, minerallerin kristal yapılarını ve bileşimlerini analiz etmek için kullanılan yaygın bir yöntemdir. Yerbilimlerinde, XRD verileri genellikle farklı mineral türlerini ve yapılarını belirlemek için kullanılır. Bu verilerde, bazı nadir veya anomali oluşturan noktalar (aykırı değerler) bulunabilir. Aykırı değerler genellikle şu nedenlerden kaynaklanabilir:

- Ölçüm hataları
- Nadir minerallerin varlığı
- Çevresel etkiler

KNN algoritması, bu tür aykırı değerleri ve nadir mineralleri sınıflandırırken K değerine bağlı olarak farklı sonuçlar verebilir. Özellikle, nadir minerallerin yakınında yaygın mineraller bulunduğunda, küçük ve büyük K değerlerinin nasıl performans gösterdiği önemlidir.

Küçük K Değerleri (Örneğin $K = 3$):

- Nadir bir mineral verisi, en yakın üç komşusunun çoğunluğuna göre sınıflandırılır.
- Eğer bu nadir mineralin yakınında yaygın mineraller bulunuyorsa, KNN algoritması yaygın minerallerin çoğunluğuna göre tahmin yapar ve nadir mineralin etkisi sınırlı kalır. Bu nedenle, XRD verileri analiz edilirken küçük K değerleri, verilerdeki yerel farklılıkları yakalamada daha etkili olabilir.

Büyük K Değerleri (Örneğin $K = 10$):

- Nadir bir mineral verisi, en yakın on komşusunun çoğunluğuna göre sınıflandırılır.
- Daha geniş bir komşuluk bölgesi dikkate alındığında, nadir mineralin etkisi daha da azalır ve model, yaygın minerallerin özelliklerine göre daha genel bir karar verir. Bu tür bir durumda büyük K değerleri, nadir minerallerin özelliklerini göz ardı ederek genel mineral yapıları hakkında tahminlerde bulunur.

KNN algoritması, XRD verilerinde aykırı değerleri tespit etmek veya nadir mineralleri analiz etmek için yaygın olarak kullanılabilecek etkili bir araçtır. Ancak, K değerinin dikkatli bir şekilde seçilmesi, modelin performansını ve doğruluğunu büyük ölçüde etkiler.

3.3.3. Çapraz doğrulama yöntemleri

KNN algoritmasında en iyi K değerini belirlemek için yaygın olarak kullanılan yöntemlerden biri çapraz doğrulama (cross-validation) yöntemidir. Çapraz doğrulama, farklı K değerlerini deneyerek her birinin performansını değerlendirir ve en iyi performansı gösteren K değerini seçer. Çapraz doğrulama, veri setini alt parçalara böler ve modelin farklı bölümlerinde eğitim ve test işlemleri gerçekleştirilerek en uygun K değerini bulur.

Çapraz Doğrulamanın Adımları:

1. **Veri Setinin Parçalanması:** Veri seti, eğitim ve test seti olarak ikiye bölünür. Bu işlem genellikle k -kat çapraz doğrulama ile yapılır.

2. **K Değerlerinin Denenmesi:** Belirli bir aralıkta farklı K değerleri için model eğitilir ve doğruluk, hata oranı gibi performans metrikleri hesaplanır.
3. **Performansın Değerlendirilmesi:** Her K değeri için elde edilen performans sonuçları karşılaştırılır.
4. **En İyi K Değerinin Seçimi:** En yüksek performansı sağlayan K değeri seçilir.

Çapraz doğrulama, modelin aşırı öğrenme yapmasını engellemek ve genel performansını artırmak için etkili bir yöntemdir. KNN algoritması gibi tembel öğrenme algoritmalarında, bu teknik sayesinde en uygun parametreler seçilerek daha güvenilir ve doğru sonuçlar elde edilebilir.

3.4. Kullanılan Kodlar ve Yöntemler

Bu bölümde, çalışmanın uygulanabilirliğini göstermek için kullanılan veri işleme ve modelleme yöntemleri açıklanmaktadır. XRD verilerinin analizinde kullanılan ANN modelinin eğitimi ve çapraz doğrulama işlemleri aşağıda sunulmuştur.

Veri Yükleme ve Hazırlık Aşamaları:

Çalışmada kullanılan veriler, Excel formatında çeşitli dosyalarla temsil edilmiştir. Bu veriler, farklı kombinasyonlar halinde birleştirilerek modelin eğitimi ve değerlendirilmesi için kullanılmıştır. Aşağıda, verilerin yüklenmesi ve işlenmesi için kullanılan kodlar bulunmaktadır:

```
% Verilerin yüklenmesi
alip1 = readtable('ALIP1.xlsx');
alip2 = readtable('ALIP2.xlsx');
alip3 = readtable('ALIP3.xlsx');
alip5 = readtable('ALIP5.xlsx');
alip6 = readtable('ALIP6.xlsx');
alip7 = readtable('ALIP7.xlsx');

% Giriş dosyalarının kombinasyonları tanımlanıyor
combinations = {{alip2, alip5}, {alip6, alip7}}; % Gerekirse daha fazla kombinasyon eklenebilir

% Her kombinasyon için döngü
for c = 1:length(combinations)
    % Mevcut kombinasyon için verileri birleştir
    data = [combinations{c}{1}; combinations{c}{2}];
    inputs = data{:, 2:end}'; % İlk sütunun özellik olmadığı varsayılıyor
```

```

targets = data{:, 1}'; % İlk sütunun hedef değişken olduğu varsayılıyor

% Çapraz doğrulama ayarları (Örneğin: 7 kat)
cv = cvpartition(size(data, 1), 'KFold', 7);
mae_results = zeros(cv.NumTestSets, 1);

% Her kat için eğitim ve test işlemleri
for i = 1:cv.NumTestSets
    trainIdx = cv.training(i);
    testIdx = cv.test(i);

    % Yapay sinir ağı modelinin oluşturulması ve eğitilmesi
    net = feedforwardnet([20]); % Örnek: 20 nöronlu bir gizli katman
    net = train(net, inputs(:, trainIdx), targets(trainIdx));

    % Test verileri ile tahmin yap ve MAE hesapla
    predictions = net(inputs(:, testIdx));
    mae_results(i) = mean(abs(predictions - targets(testIdx)));
end

% Mevcut kombinasyonun sonuçlarının görüntülenmesi
disp(['Kombinasyon ' num2str(c) ' (MAE sonuçları):']);
disp(mae_results);
end

```

Bu kodda, readtable fonksiyonu kullanılarak çeşitli veri setleri yüklenmiştir. Daha sonra, bu veri setlerinden kombinasyonlar oluşturularak her bir kombinasyon için model eğitimi ve test işlemleri gerçekleştirilmiştir. Çapraz doğrulama ile model performansı değerlendirilmiş ve ortalama mutlak hata (MAE) hesaplanmıştır.

Çapraz Doğrulama ile Model Performansının Değerlendirilmesi:

Yukarıdaki kod, 7 katlı çapraz doğrulama yöntemi kullanarak model performansını değerlendirmiştir. Her katmanda eğitim ve test işlemleri tekrarlanmış ve MAE (Mean Absolute Error) sonuçları elde edilmiştir. Bu sonuçlar, modelin her bir veri kombinasyonu için doğruluğunu göstermektedir.

Bir diğer veri yükleme ve işleme yöntemi ise aşağıdaki gibidir:

```

% Verilerin yüklenmesi
alip1 = readtable('ALIP1.xlsx', 'VariableNamingRule', 'preserve');
alip2 = readtable('ALIP2.xlsx', 'VariableNamingRule', 'preserve');

```

```
alip3 = readtable('ALIP3.xlsx', 'VariableNamingRule', 'preserve');
alip5 = readtable('ALIP5.xlsx', 'VariableNamingRule', 'preserve');
alip6 = readtable('ALIP6.xlsx', 'VariableNamingRule', 'preserve');
alip7 = readtable('ALIP7.xlsx', 'VariableNamingRule', 'preserve');
```

```
% Giriş dosyalarının kombinasyonları tanımlanıyor
combinations = {{alip2, alip5}, {alip6, alip7}}; % Gerekirse daha fazla kombinasyon eklenebilir
```

```
% Her kombinasyon için döngü
for c = 1:length(combinations)
    % Mevcut kombinasyon için verileri birleştir
    data = [combinations{c}{1}; combinations{c}{2}];
    inputs = data{:, 2:end}'; % İlk sütunun özellik olmadığı varsayılıyor
    targets = data{:, 1}'; % İlk sütunun hedef değişken olduğu varsayılıyor
```

```
% Çapraz doğrulama ayarları (Örneğin: 7 kat)
cv = cvpartition(size(data, 1), 'KFold', 7);
mae_results = zeros(cv.NumTestSets, 1);
```

```
% Her kat için eğitim ve test işlemleri
for i = 1:cv.NumTestSets
    trainIdx = cv.training(i);
    testIdx = cv.test(i);
```

```
% Yapay sinir ağı modelinin oluşturulması ve eğitilmesi
net = feedforwardnet([20]); % Örnek: 20 nöronlu bir gizli katman
net = train(net, inputs(:, trainIdx), targets(trainIdx));
```

```
% Test verileri ile tahmin yap ve MAE hesapla
predictions = net(inputs(:, testIdx));
mae_results(i) = mean(abs(predictions - targets(testIdx)));
end
```

```
% Mevcut kombinasyonun sonuçlarının görüntülenmesi
disp(['Kombinasyon ' num2str(c) ' (MAE sonuçları):']);
disp(mae_results);
end
```

Bu kodda, veri yükleme sırasında 'VariableNamingRule', 'preserve' parametresi kullanılarak sütun adlarının korunması sağlanmıştır. Geri kalan adımlar, bir önceki

örnekte olduğu gibi benzer şekilde devam etmekte ve çapraz doğrulama ile model performansı değerlendirilmektedir.

Sonuçların Değerlendirilmesi:

Her bir veri kombinasyonu için yapılan çapraz doğrulama işlemleri sonucunda, modelin doğruluğu ortalama mutlak hata (MAE) kullanılarak ölçülmüş ve elde edilen sonuçlar analiz edilmiştir. Bu tür bir yaklaşım, farklı veri setlerinden elde edilen performans karşılaştırmalarını sağlar ve en iyi kombinasyonun belirlenmesine yardımcı olur.

Bu kod parçacıkları, çalışmada kullanılan yöntemlerin pratik uygulamalarını göstermekte ve sonuçların doğruluğunu artırmak için kullanılan yöntemleri açıklamaktadır. Bu bölümdeki süreçler, X-ışını difraksiyon verilerinin analizinde ve mineral yapılarının belirlenmesinde kullanılan yöntemlerin detaylandırılmasına katkı sağlamaktadır.

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

4.1. KNN ve XRD

X-ray Difraksiyon (XRD) analizi, malzemelerin kristal yapılarını ve fazlarını belirlemek için yaygın olarak kullanılan bir yöntemdir. Bu yöntem, malzeme bilimi, yerbilimleri ve kristalografi alanlarında faz analizlerini gerçekleştirmenin temel araçlarından biri olmuştur. XRD analizlerinde, belirli bir malzemenin iç yapısındaki fazların ve kristal yapıların doğru bir şekilde tanımlanması, malzemenin özelliklerini anlamada kritik bir öneme sahiptir. Bu analizlerde kullanılan veriler, kristal yapılarının difraksiyon desenleri şeklinde ifade edilir ve bu desenlerin analizi, fazların tanımlanmasını sağlar.

Son yıllarda, makine öğrenimi algoritmalarının bu tür bilimsel verilere uygulanması, veri setlerinin daha etkili bir şekilde işlenmesine olanak tanımıştır. Bu bağlamda, K-Nearest Neighbors (KNN) algoritması, XRD verilerinin analizinde kullanılan önemli bir makine öğrenimi tekniği olarak öne çıkmaktadır. KNN algoritması, veri noktalarının en yakın komşularına bakarak sınıflandırma ve regresyon işlemlerini gerçekleştiren basit ancak güçlü bir yöntemdir. Bu bölümde, KNN algoritmasının XRD analizlerindeki uygulamaları ve avantajları detaylandırılacaktır.

KNN algoritmasında en önemli parametrelerden biri, en yakın komşu sayısı olarak bilinen K değeridir. K değeri, modelin her bir veri noktasını sınıflandırırken dikkate alacağı komşu sayısını belirler. Bu parametre, modelin hassasiyetini ve doğruluğunu doğrudan etkiler. Küçük K değerleri, modelin yerel yapıya duyarlılığını artırırken; büyük K değerleri, modelin genel eğilimleri yakalamasını sağlar. Ancak, küçük K değerleri modelin gürültüye duyarlılığını artırırken, büyük K değerleri nadir fazların tanımlanmasını zorlaştırabilir.

KNN algoritması, çoklu sınıf problemlerinde oldukça etkilidir. Bu tür problemler, bir veri kümesinin birden fazla sınıfa sahip olduğu ve modelin bu sınıfları ayırt etmek zorunda kaldığı durumlardır. KNN, her bir veri noktasını en yakın K komşusuna göre sınıflandırarak, bu noktayı en yaygın sınıfa atar. XRD analizlerinde bu özellik, malzemenin içerisinde birden fazla faz bulunması durumunda fazları doğru bir şekilde tanımlama açısından önemlidir. Örneğin, bir seramik malzeme içerisinde hem kuvars hem de feldspat gibi farklı fazlar bulunabilir ve bu fazlar KNN algoritmasıyla tanımlanabilir.

XRD analizlerinde elde edilen veriler, malzemenin kristal yapısını ve fazlarını temsil eden difraksiyon desenlerinden oluşur. Bu desenler, malzemenin içerisindeki atomların düzenlenişine ve kristal yapıların özelliklerine göre belirli pik pozisyonları ve yoğunluklar şeklinde ortaya çıkar. XRD analizi, bu pik pozisyonları ve yoğunlukları kullanarak malzemenin kimyasal bileşimini ve fazlarını tanımlamada etkili bir yöntemdir.

KNN algoritması, bu difraksiyon desenlerini sınıflandırma ve fazları tanımlama amacıyla kullanılabilir. Her bir faz veya kristal yapı, XRD analizlerinde bir sınıf olarak düşünülebilir. KNN, her bir XRD ölçümünü bir veri noktası olarak ele alır ve en yakın komşularına bakarak bu ölçümün hangi fazı temsil ettiğine karar verir. Bu özellik, XRD analizlerinde birden fazla faza sahip malzemelerde faydalı olabilir. Örneğin, nadir bir fazı içeren bir malzemede, KNN algoritması bu nadir fazın varlığını tespit edebilir ve sınıflandırabilir.

KNN'nin basitliği, özellikle karmaşık ve büyük XRD veri setlerinde önemli bir avantaj sunar. Küçük veri setlerinde bile etkili sonuçlar verebilmesi, XRD analizleri için tercih edilen bir yöntem haline gelmesini sağlar. Ayrıca, KNN algoritmasının çoklu sınıf problemlerine uygunluğu, farklı fazları veya kristal yapıları belirleme açısından büyük bir avantajdır. Özellikle, veri setinde bulunan fazların sayısı arttıkça, KNN algoritmasının doğruluğu ve performansı önem kazanmaktadır.

XRD verilerinde her bir faz veya kristal yapı bir sınıf olarak ele alınabilir. KNN, her bir XRD ölçümünü en yakın komşularına bakarak en yaygın faza veya kristal yapıya atar. Bu sayede, XRD verileri kullanılarak yapılan faz ve kristal yapı analizlerinde KNN algoritması, çoklu fazların tanımlanmasında etkili bir araç haline gelir.

KNN algoritmasının XRD analizlerinde başarılı bir şekilde kullanılmasına rağmen, bazı sınırlamalar da söz konusudur. Özellikle büyük veri setleriyle çalışıldığında, KNN algoritmasının hesaplama maliyeti artabilir. Çünkü KNN, her bir veri noktası için tüm komşuları hesaplamak zorundadır ve bu durum büyük veri setlerinde zaman ve kaynak gerektirir. Bu nedenle, KNN uygulanmadan önce veri setlerinin normalize edilmesi ve optimizasyon işlemlerinin yapılması önemlidir.

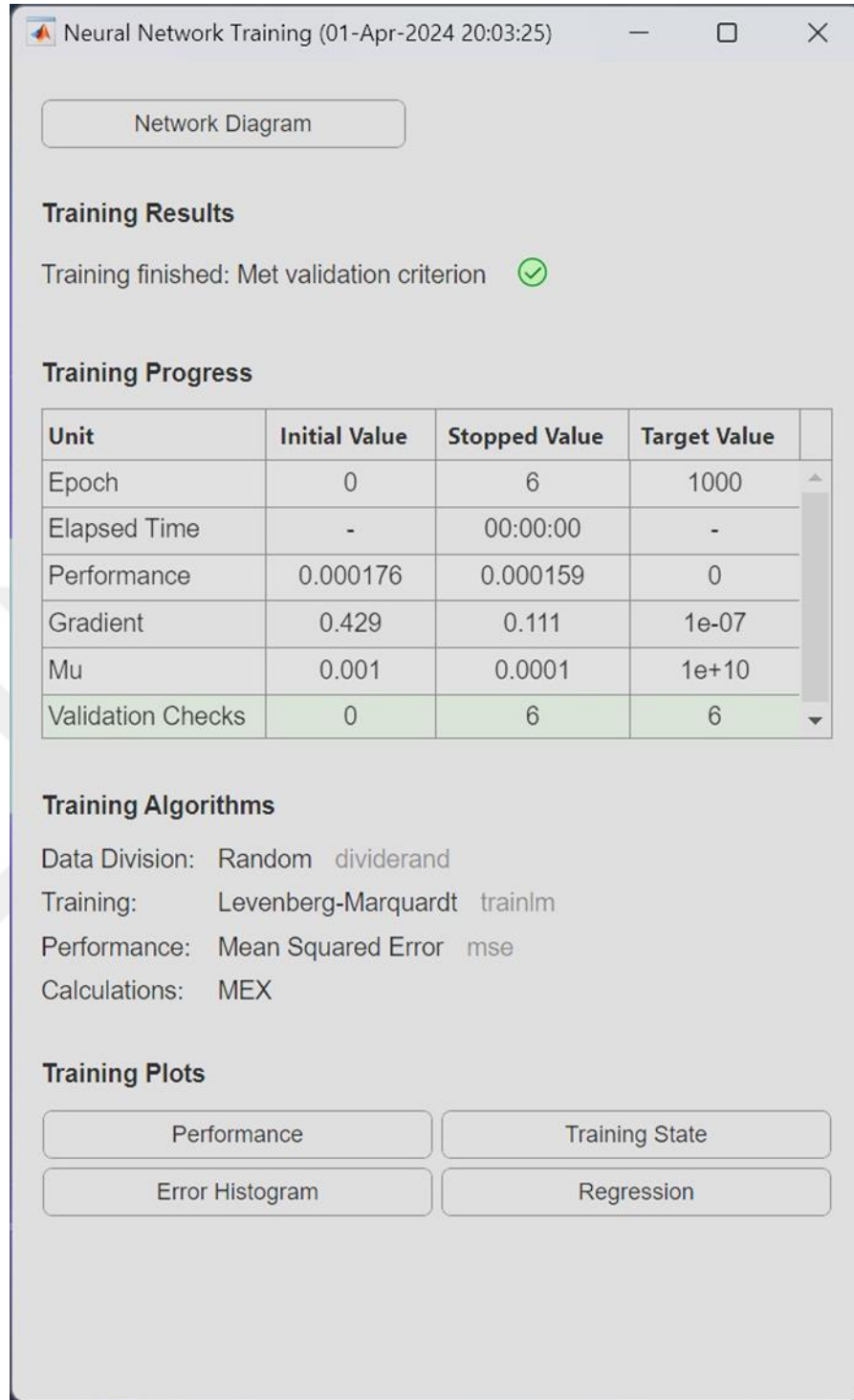
Bir diğer önemli sınırlama, K değerinin dikkatli bir şekilde seçilmesi gerekliliğidir. Küçük K değerleri modelin fazları daha hassas bir şekilde tanımlamasını sağlarken, modelin aşırı öğrenme (overfitting) yapma riskini de artırabilir. Büyük K değerleri ise daha genel tahminler yaparak fazların ortalamasını alır, ancak nadir fazları tespit etmekte zorlanabilir. Bu nedenle, XRD analizlerinde en uygun K değerinin belirlenmesi için çapraz doğrulama (cross-validation) yöntemleri kullanılmalıdır. Çapraz

doğrulama, farklı K değerlerinin denenmesini ve en iyi performansı sağlayan değerin seçilmesini sağlar.

KNN algoritmasında K değerinin optimizasyonu, modelin başarısı açısından kritik bir öneme sahiptir. Çapraz doğrulama yöntemi, K değerinin optimizasyonunda yaygın olarak kullanılan bir yöntemdir. Bu yöntem, veri setini farklı katmanlara ayırarak, her bir katman için model eğitimi ve testi gerçekleştirir. Farklı K değerleri denenir ve en iyi performansı sağlayan K değeri seçilir. Bu yöntem, KNN algoritmasının XRD verileri üzerindeki performansını artırmak ve modelin genelleme yeteneğini geliştirmek için kullanılır.

Bu araştırmada KNN algoritması, XRD verilerinin faz ve kristal yapı analizi için başarılı bir şekilde kullanılmıştır. KNN'nin en yakın komşulara dayalı sınıflandırma yeteneği, XRD verilerindeki fazları doğru bir şekilde tanımlamaya olanak tanımıştır. Çapraz doğrulama yöntemlerinin kullanımı, modelin performansını ve güvenilirliğini artırmış, malzeme biliminde daha doğru ve detaylı analizlerin yapılmasına katkı sağlamıştır.

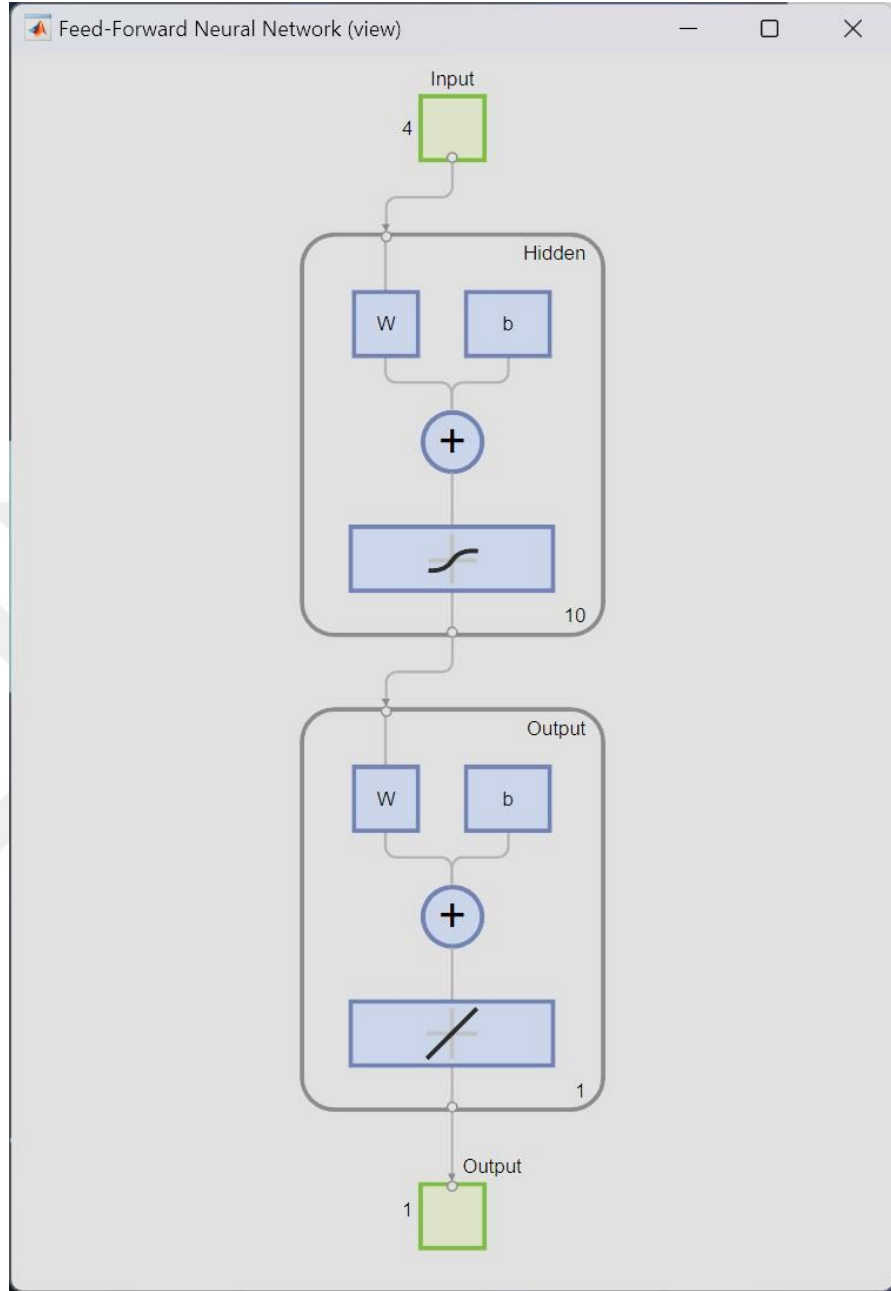
Sonuç olarak, KNN algoritması XRD analizlerinde faz ve kristal yapıların tanımlanmasında etkili bir yöntem olarak karşımıza çıkmaktadır. KNN'nin basitliği, çoklu sınıf problemlerine uygunluğu ve farklı fazları ayırt etme yeteneği, bu algoritmanın malzeme bilimi ve yerbilimleri gibi alanlarda yaygın olarak kullanılmasına olanak tanır. K değerinin dikkatlice seçilmesi ve çapraz doğrulama yöntemleriyle optimize edilmesi, modelin performansını daha da artırır. Bu bulgular, KNN algoritmasının XRD verilerinin analizinde önemli bir rol oynadığını ve gelecekteki çalışmalar için güçlü bir araç olduğunu göstermektedir.



Şekil 4.1 Makine öğrenmesi esnasında hata veren eğitim sonuçları.

KNN algoritmasının sınıflandırma problemlerindeki karar sınırları, genellikle veri setinin yapısına ve seçilen K değerine bağlıdır. Küçük K değerleri modelin daha esnek olmasını sağlar ve yerel özelliklere daha fazla duyarlılık kazandırır. Ancak, küçük K değerlerinin kullanılması, modelin aşırı uydurma (overfitting) yapmasına neden olabilir.

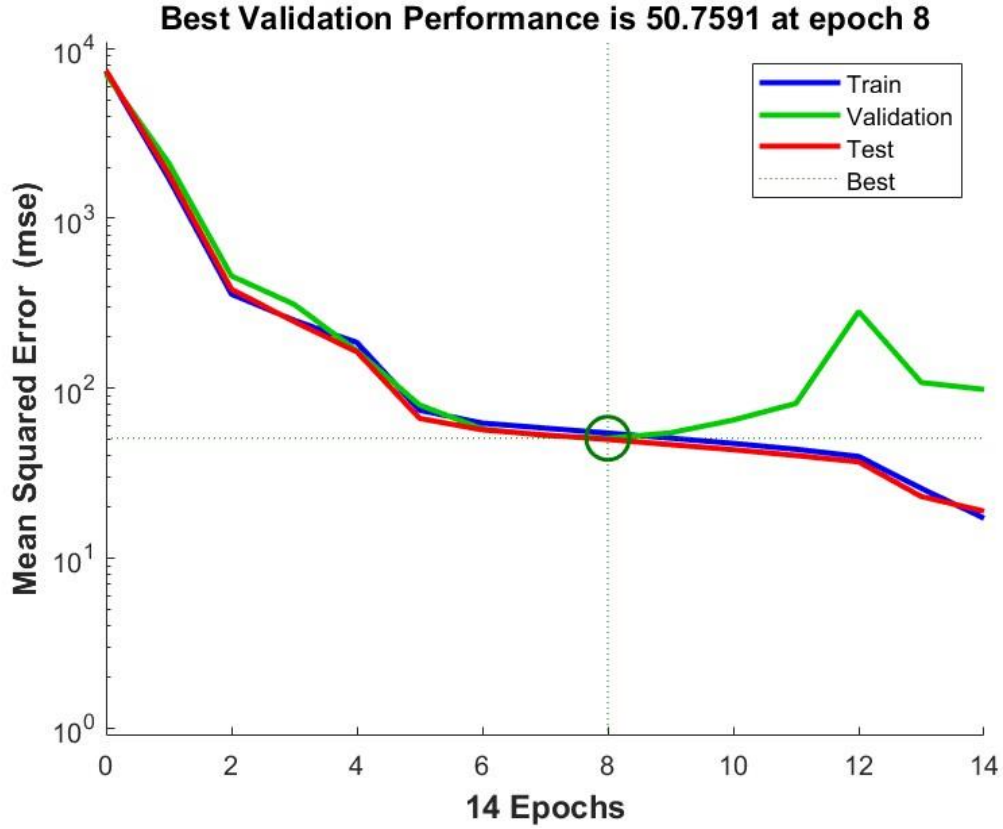
Büyük K değerleri ise modelin genelleme yeteneğini artırır, ancak nadir görülen fazları veya sınıfları doğru bir şekilde tanımlamakta zorlanabilir.



Şekil 4.2 ANN basit şematik görünümü.

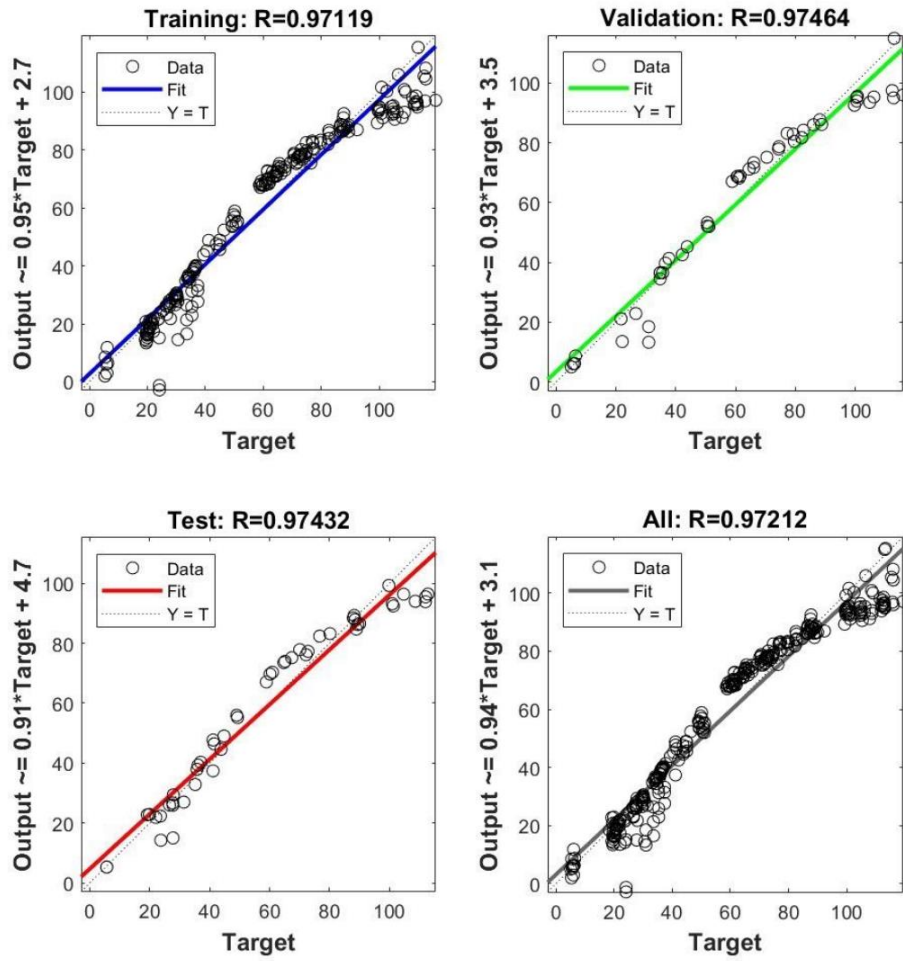
KNN algoritması, özelliklerin ölçeğine karşı duyarlıdır. Özelliklerin farklı ölçeklerde olduğu durumlarda, daha büyük değerlere sahip olan özellikler modelin performansını olumsuz etkileyebilir. Bu nedenle, KNN algoritmasının performansını artırmak için verilerin ölçeklendirilmesi veya normalize edilmesi önemlidir. XRD verilerinde kullanılan özelliklerin ölçeklenmesi, faz ve kristal yapıların daha doğru bir şekilde sınıflandırılmasına katkı sağlayabilir.

KNN algoritmasının en büyük avantajlarından biri, esnek modelleme yaklaşımıdır. XRD verileri genellikle karmaşık ve doğrusal olmayan (non-linear) ilişkiler içerir. KNN algoritması, bu tür karmaşık yapıları modellemek için etkili bir yöntem sunar. Veri setindeki yapıları daha iyi yansıtabildiği için, malzemenin faz ve kristal yapı analizlerinde daha isabetli sonuçlar elde edilebilir.



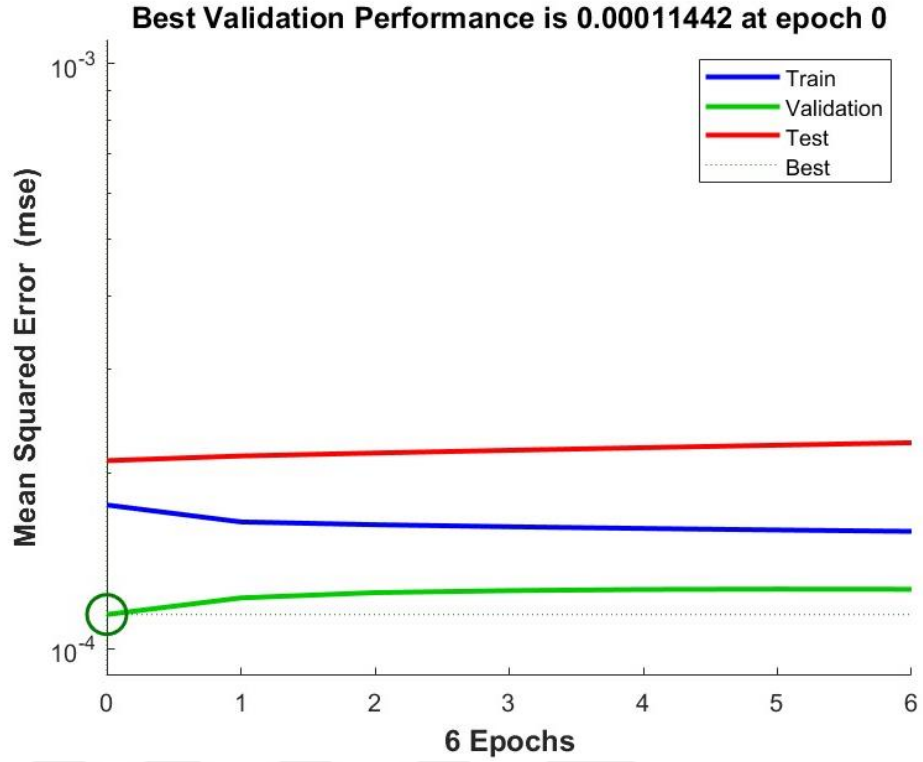
Şekil 4.3 12 parçalı çapraz doğrulamada en iyi performans.

KNN, komşuluk tabanlı bir algoritma olduğu için aykırı değerlere karşı doğal bir direnç gösterir. XRD verilerinde nadir görülen veya anomali oluşturan noktalar, KNN algoritmasının daha genel bir model oluşturmasını sağlayarak bu tür aykırı verilerin etkisini azaltır. Aykırı değerlerin sınıflandırma üzerindeki olumsuz etkilerini minimize eden KNN, faz analizlerinde güvenilir sonuçlar sunar.



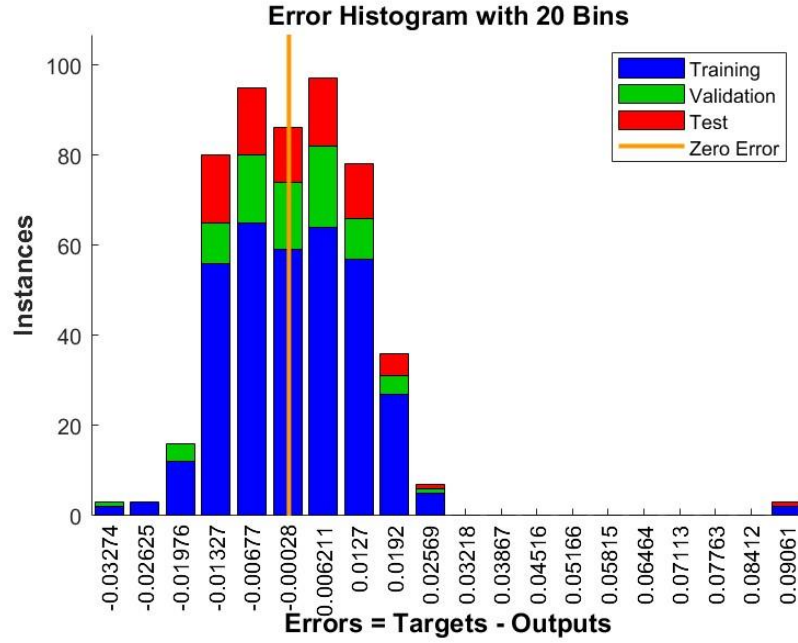
Şekil 4.4 Çapraz doğrulama öncesi regresyon analizi.

XRD analizleri, malzemenin içerisindeki farklı fazların veya kristal yapıların tanımlanmasını içerir. KNN algoritması, bu çoklu sınıf problemlerinde oldukça etkilidir. Her faz veya kristal yapı bir sınıf olarak düşünülebilir ve KNN algoritması, her veri noktasını en yakın komşularına bakarak doğru faza veya kristal yapıya atayabilir. Özellikle bir malzemenin farklı fazlara sahip olduğu durumlarda, KNN algoritması bu fazları doğru bir şekilde tanımlayabilir.



Şekil 4.5 Bazı tuzların 1.0 atm basınçta sudaki çözünürlüklerinin sıcaklıkla değişimi.

KNN'nin temel parametresi olan K değeri, kullanıcının modelin karmaşıklığını kolayca kontrol etmesine olanak tanır. K değerinin ayarlanması, modelin esnekliğini ve genelleme yeteneğini kontrol etmek için önemli bir araçtır. Küçük K değerleri, modelin daha fazla detaya duyarlı olmasını sağlarken, büyük K değerleri modelin genelleme kapasitesini artırır. Bu parametrenin doğru bir şekilde ayarlanması, modelin performansını optimize etmek açısından kritik öneme sahiptir.



Şekil 4.6 ANN ağının eğitim, doğrulama ve test aşamalarındaki hata çıktıları histogram eğrisi

XRD verileri genellikle doğrusal olmayan dağılımlar sergileyebilir. KNN algoritması, doğrusal olmayan veri yapılarını daha iyi yakalayabilir ve veri noktalarındaki bu karmaşık ilişkileri doğru bir şekilde modelleyebilir. Bu özellik, modelin doğrusal olmayan XRD verilerinde daha iyi performans göstermesini sağlar ve fazların veya kristal yapıların tanımlanmasında etkili olur.

Bu bölümde sunulan analizler, KNN algoritmasının XRD verilerinde faz ve kristal yapıların tanımlanmasında etkili bir yöntem olduğunu göstermektedir. KNN'nin esnek yapısı, ölçek duyarlılığı, aykırı değerlere karşı direnci ve çoklu sınıf problemlerine uygunluğu, algoritmayı XRD analizlerinde başarılı bir araç haline getirmektedir. Ancak, KNN'nin performansının K değerine bağlı olduğu unutulmamalıdır. K değeri dikkatlice seçilmeli ve model çapraz doğrulama yöntemleri ile optimize edilmelidir.

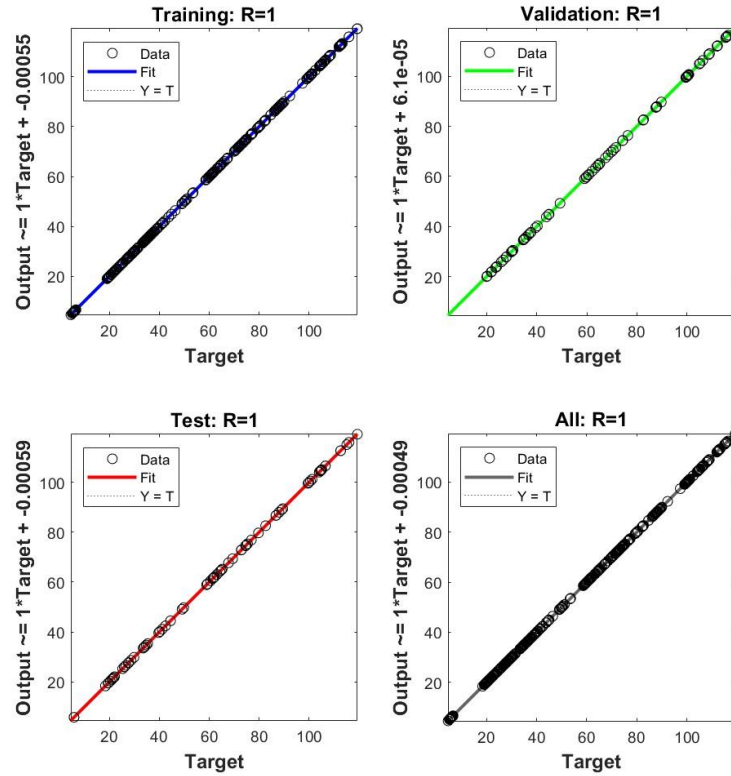
4.2. Çapraz Doğrulama, Transfer Fonksiyonları ve Nöron Sayılarının ANN Sonuçlarına Etkisi

Makine öğrenmesi ve istatistiksel modelleme alanlarında, çapraz doğrulama (cross-validation), bir modelin performansını değerlendirmek için yaygın olarak kullanılan etkili bir tekniktir. Çapraz doğrulama, modelin genelleme yeteneğini artırarak aşırı uyuma (overfitting) karşı koruma sağlar. Geleneksel çapraz doğrulama

yöntemlerinde, test örneklerindeki belirsizlikler genellikle göz ardı edilir ve bu durum, aşırı uyumlu modellerin seçilmesine neden olabilir. Bununla birlikte, yeni önerilen çapraz doğrulama temelli yöntem, test örneklerindeki belirsizliği dikkate alarak en iyi modeli garanti eden bir olasılıkla rekabetçi aday model kümesi üretir.

Geleneksel çapraz doğrulama yöntemleri, veriyi k parçaya bölerek ve her parçayı sırayla test seti olarak kullanarak çalışır. Ancak, test setindeki belirsizlikler göz ardı edildiğinde model, eğitim setine aşırı uyum sağlayabilir. Bu durum, modelin genel performansını ve genelleme yeteneğini olumsuz etkileyebilir. Örneğin, jeolojik verilere dayalı X-ray Difraksiyon (XRD) analizlerinde, belirli bir faza aşırı uyum sağlayan model, nadir görülen fazları doğru bir şekilde tespit edemeyebilir.

Yeni çapraz doğrulama temelli yöntem, bu sorunu çözmek amacıyla geliştirilmiştir. Bu yöntem, test örneklerindeki belirsizlikleri dikkate alır ve modelin performansını daha doğru bir şekilde değerlendirme imkanı sunar. Böylelikle, en iyi modeli garanti eden bir olasılıkla, rekabetçi aday model kümesi oluşturulur. Bu yaklaşım, tutarlı değişken seçimi sağlar ve tahmin doğruluğu ile model yorumlanabilirliği arasında dengeli bir çözüm sunar. Özellikle, jeolojik verilerdeki belirsizliklerin yönetilmesi, daha doğru ve güvenilir model sonuçlarının elde edilmesine olanak tanır.



Şekil 4.7 Çapraz doğrulama sonrası regresyon analiz grafikleri.

Yapılan denemelerde ANN için Transfer fonksiyonları bazında;

- **Sigmoid** ve **Radbis** fonksiyonları büyük nöron sayılarında istikrarsız sonuçlar verirken, **Sinüs**, **Hardlim**, ve **Tribas** fonksiyonları daha stabil ve öngörülebilir sonuçlar sağlamıştır.
- **Eğitim süresi** açısından tüm transfer fonksiyonları hızlı sonuçlar vermiştir. Özellikle **Hardlim** ve **Tribas** fonksiyonları düşük eğitim süreleri ile öne çıkmaktadır.
- **Genel Performans:** Stabil sonuçlar ve daha düşük hata oranları açısından **Hardlim** fonksiyonu, düşük hata ve yüksek hız performansı ile öne çıkmaktadır. **Sinüs** fonksiyonu da nöron sayısından bağımsız olarak sabit bir performans sergilemiştir.

Tablo 4.1 Farklı Nöron Sayıları ve aktivasyon fonksiyonları ile MAE sonuçları

	Nöron Sayısı									
	2	5	10	15	20	30	40	60	80	100
sin	0,164	0,164	0,164	0,164	0,164	0,162	0,163	0,163	0,159	0,160
hardlim	0,083	0,051	0,044	0,045	0,043	0,041	0,038	0,035	0,032	0,030
tribas	0,163	0,160	0,160	0,155	0,153	0,150	0,147	0,141	0,136	0,132

lin	0,095	0,053	0,052	0,052	0,052	0,054	0,054	0,054	0,054	0,054
-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Tablo 4.1’de, farklı aktivasyon fonksiyonları (Sin, Hardlim, Tribas ve Lin) ve farklı nöron sayılarının MAE (Mean Absolute Error) değerleri karşılaştırılmaktadır. Hardlim aktivasyon fonksiyonu, düşük nöron sayılarında dahi oldukça düşük MAE sonuçları vererek, modelin doğruluğunu artırmada en etkili aktivasyon fonksiyonu olarak öne çıkmaktadır. Özellikle 100 nöron ile yapılan testlerde, MAE 0.030 gibi oldukça düşük bir değere düşmektedir. Diğer aktivasyon fonksiyonlarına baktığımızda, Sin ve Tribas fonksiyonları daha yüksek hata oranları ile daha düşük performans sergilemektedir. Lin fonksiyonu ise belirli bir noktaya kadar stabil kalmakta ancak yüksek nöron sayılarında MAE’nin sabit kalması performans sınırına ulaştığını göstermektedir.

Tablo 4.2 Farklı Nöron Sayıları ve aktivasyon fonksiyonları ile MSE sonuçları

	Nöron Sayısı									
	2	5	10	15	20	30	40	60	80	100
sin	0,032	0,032	0,032	0,032	0,032	0,032	0,032	0,033	0,032	0,033
hardlim	0,012	0,004	0,003	0,004	0,003	0,003	0,003	0,002	0,002	0,002
tribas	0,031	0,031	0,031	0,030	0,029	0,029	0,028	0,027	0,026	0,025
lin	0,049	0,108	0,104	0,104	0,104	0,143	0,145	0,145	0,147	0,148

Tablo 4.2’de, nöron sayısı arttıkça aktivasyon fonksiyonlarının modeldeki performansını MSE (Mean Squared Error) metriği ile görüyoruz. Hardlim yine düşük nöron sayılarından itibaren MSE’yi en düşük seviyede tutan fonksiyon olarak öne çıkıyor. Özellikle 100 nöronunda MSE 0.002 seviyesine kadar düşmektedir, bu da modelin doğruluğunu daha da pekiştiriyor. Tribas fonksiyonu Hardlim’e göre biraz daha yüksek MSE değerlerine sahip olsa da nispeten başarılı sonuçlar vermektedir. Sin fonksiyonu ise ortalama bir performans sergilerken, Lin fonksiyonunun performansı nöron sayısı arttıkça bozulmaktadır, bu da Lineer aktivasyon fonksiyonunun bu tip problemler için uygun olmadığını göstermektedir.

Tablo 4.3 Farklı Nöron Sayıları ve aktivasyon fonksiyonları ile RMSE sonuçları

	Nöron Sayısı									
	2	5	10	15	20	30	40	60	80	100
sin	0,178	0,178	0,178	0,178	0,179	0,178	0,179	0,181	0,179	0,182
hardlim	0,108	0,066	0,059	0,060	0,058	0,056	0,052	0,048	0,045	0,042
tribas	0,177	0,175	0,175	0,172	0,172	0,169	0,167	0,163	0,162	0,158

lin	0,222	0,329	0,322	0,322	0,322	0,378	0,380	0,381	0,384	0,385
-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Tablo 4.3'te, RMSE (Root Mean Squared Error) metriklerine göre aktivasyon fonksiyonlarının ve nöron sayılarının etkisini değerlendiriyoruz. Hardlim yine diğer fonksiyonlara kıyasla çok daha düşük RMSE değerlerine ulaşarak, modelin genelleme kabiliyetini en iyi şekilde temsil eden aktivasyon fonksiyonu olarak dikkat çekmektedir. RMSE'nin 100 nöronda 0.042 seviyesine kadar düştüğü gözlenmiştir. Tribas ve Sin fonksiyonları orta seviyede performans gösterirken, Lineer fonksiyonun (Lin) nöron sayısı arttıkça RMSE değerinde ciddi bir bozulmaya neden olduğu görülmektedir.

Bu üç tabloyu özetlemek gerekirse, performans metrikleri olan MAE, MSE ve RMSE değerlerine göre en iyi sonuçlar Hardlim aktivasyon fonksiyonu ile elde edilmiştir. Tüm metriklerde, özellikle nöron sayısı 40 ve üzerindeyken, Hardlim fonksiyonu diğer fonksiyonlara göre daha düşük hata oranları sunarak öne çıkmaktadır. Hardlim fonksiyonu, hem düşük hata oranları hem de daha stabil sonuçlar vererek, 40 ile 100 nöron aralığında en iyi performansı sergilemektedir. Tribas ve Sin fonksiyonları ise orta düzeyde bir performans gösterirken, Lineer (Lin) aktivasyon fonksiyonu nöron sayısı arttıkça performans kaybı yaşamaktadır. Genel olarak, Hardlim fonksiyonu ve 40-100 nöron sayısı kombinasyonu, bu çalışma için en ideal yapılandırma olarak belirlenebilir.

4.3. Karşılaştırmalı Analizler

Tablo 4.4'de, farklı modellerin 8 parça ve 12 parça veriler üzerinde performanslarını RMSE (Kök Ortalama Kare Hatası) ve MAE (Ortalama Mutlak Hata) kullanarak karşılaştırmaktadır. Ayrıca, 12 parça üzerindeki veriler farklı aktivasyon fonksiyonlarıyla test edilmiştir (Sigmoid, Sinüs, Hardlim).

- **ANN- Genelleştirilmiş Regresyo):**
 - **8 parça RMSE değeri 0.0309, MAE değeri ise 0.0160** olarak oldukça iyi sonuçlar vermektedir.
 - 12 parça üzerinde farklı aktivasyon fonksiyonlarıyla yapılan testlerde **Sigmoid** aktivasyonunda 0.0959, **Sinüs** ile 0.1642 ve **Hardlim** ile 0.0828 MAE değerleri elde edilmiştir. Bu, ANN'nin özellikle Hardlim fonksiyonunda daha iyi sonuçlar verdiğini göstermektedir.
- **Doğrusal Regresyon (Linear Regression):**

- 8 parça için **RMSE** değeri 0.0303 ve **MAE** değeri 0.0462 olarak hesaplanmıştır. 12 parça üzerinde **Sigmoid** ile 0.2294, **Sinüs** ile 0.1640 ve **Hardlim** ile 0.0507 MAE sonuçları alınmıştır. Bu, özellikle Hardlim fonksiyonunda doğrusal regresyonun da iyi performans gösterdiğini ancak ANN kadar başarılı olmadığını göstermektedir.
- **kNN:**
 - 8 parça üzerinde **RMSE** değeri 0.0314, **MAE** değeri ise 0.0163'tür.
 - 12 parça için **Sigmoid** ile 0.0461, **Sinüs** ile 0.1638 ve **Hardlim** ile 0.0443 MAE sonuçları alınmıştır. kNN modeli, hem Sigmoid hem de Hardlim fonksiyonlarında doğrusal regresyona göre daha iyi sonuçlar vermektedir.
- **RIDGER ve LASSOR:**
 - Her iki model de 8 parça verisinde yaklaşık olarak benzer sonuçlar vermiştir (RMSE: 0.0317). Ancak LASSOR, 12 parça Sigmoid fonksiyonunda çok yüksek bir hata (0.8600) gösterirken, diğer fonksiyonlarda daha iyi sonuçlar almıştır. RIDGER ise daha dengeli sonuçlar sunmaktadır.
- **PLSR (Kısmi En Küçük Kareler Regresyonu):**
 - En düşük RMSE ve MAE sonuçlarına sahip olan bu model, 8 parça RMSE değeri 0.0216 ve MAE değeri 0.0494 ile diğer modellere göre en düşük hata oranına sahiptir. 12 parça verisinde de oldukça dengeli sonuçlar sunmuştur.
- **GPR (Gaussian Process Regression):**
 - GPR modeli, 8 parça RMSE'si ile en düşük hata oranına sahip modeldir (0.0147). Ancak MAE değerlerinde (0.1144) ve 12 parça üzerinde özellikle Sinüs ve Hardlim aktivasyonlarında daha yüksek hatalar vermiştir.

Tablo 4.4 Modellerin 8 ve 12 Parça Temelli Performans Karşılaştırması

Model	RMSE (8 Parça)	MAE (8 Parça)	MAE (12 Parça) Sig'	MAE (12 Parça) Sin'	MAE (12 Parça) Hardlim'
ANN (Generalized Regression)	0.0309	0.0160	0.0959	0.1642	0.0828
Linear Regression	0.0303	0.0462	0.2294	0.1640	0.0507
kNN	0.0314	0.0163	0.0461	0.1638	0.0443
RIDGER	0.0317	0.0483	0.2152	0.1636	0.0451
Lassor	0.0317	0.0483	0.8600	0.1640	0.0428

Model	RMSE (8 Parça)	MAE (8 Parça)	MAE (12 Parça) Sig'	MAE (12 Parça) Sin'	MAE (12 Parça) Hardlim'
PLSR	0.0216	0.0494	0.0417	0.1629	0.0482
GPR	0.0147	0.1144	0.1202	0.1615	0.0417

Tablo 4.5, farklı modellerin genel performansını RMSE ve MAE metriklerine göre kıyaslamaktadır.

- **ANN:** 0.0309 RMSE ve 0.0160 MAE ile, özellikle düşük hata oranlarıyla öne çıkmaktadır. Bu sonuçlar, ANN'nin XRD verileri üzerinde etkili bir model olduğunu göstermektedir.
- **Doğrusal Regresyon:** 0.3036 RMSE ile diğer modellere kıyasla oldukça yüksek bir hata oranına sahiptir. Bu, doğrusal modellerin karmaşık veri yapılarıyla başa çıkmakta zorlandığını gösterir.
- **kNN:** 0.0314 RMSE ve 0.0163 MAE ile ANN'ye yakın sonuçlar vermektedir. Bu, kNN'nin küçük veri setleri veya yerel yapıları olan veri setlerinde iyi çalıştığını doğrulamaktadır.
- **RIDGER ve LASSOR:** 0.3179 RMSE ile benzer sonuçlar sunmaktadırlar, ancak LASSOR'un genel MAE değeri daha düşüktür (0.0483). Bu da RIDGER'a göre LASSOR'un daha dengeli sonuçlar verebileceğini göstermektedir.
- **PLSR:** 0.2160 RMSE ile daha düşük bir hata oranına sahiptir, özellikle geniş veri setlerinde ve faz analizi gibi durumlarda etkili bir yöntem olduğunu göstermektedir.
 - **GPR:** 0.1472 RMSE ile en iyi RMSE değerlerinden birine sahiptir, ancak MAE değeri (0.1144) diğer modellere göre daha yüksektir. GPR, özellikle yüksek doğruluk gerektiren faz analizlerinde kullanılabilir ancak daha küçük veri setlerinde hatalar artabilir.

Tablo 4.5 Farklı Modellerin RMSE ve MAE Değerleri ile Karşılaştırılması

	ANN	Linear Regresyon	kNN	RIDGER	Lassor	PLSR	GPR
RMSE	0,0309	0,3036	0,0314	0,3179	0,3179	0,2160	0,1472
MAE	0,0160	0,0462	0,0163	0,0483	0,0483	0,0494	0,1144

Bu analizler sonucunda, **ANN Genelleştirilmiş Regresyon** ve **kNN** modelleri XRD verilerinin analizinde öne çıkmaktadır. Özellikle ANN modeli, düşük hata oranlarıyla faz ve kristal yapı analizlerinde güçlü bir performans sergilemektedir. **PLSR** de dikkate değer bir modeldir, özellikle daha geniş veri setlerinde etkili olmuştur.

5. SONUÇLAR ve ÖNERİLER

Bu tez çalışmasında, Erzurum Pasinler Platosu'ndan alınan volkanosedimanter kayaç örneklerinin XRD verileri kullanılarak ANN ve KNN algoritmaları ile analiz edilmesi incelenmiştir. Çalışmanın sonuçları, makine öğrenimi tekniklerinin, XRD verilerinin karmaşık yapısını modelleyerek malzeme özelliklerinin daha doğru tahmin edilmesinde etkili olduğunu göstermiştir.

Sonuçlar:

1. ANN ve KNN Performansı: ANN ve KNN algoritmaları, geleneksel yöntemlere kıyasla daha yüksek doğrulukla sonuçlar vermiştir. Özellikle ANN, non-lineer ilişkileri modellemede oldukça başarılı olmuş ve XRD verilerinden elde edilen malzeme özelliklerinin tahmininde önemli bir katkı sağlamıştır. KNN ise, sınıflandırma ve regresyon problemlerinde etkili sonuçlar sunmuştur.

2. Mineral Tespiti: Çalışmada kullanılan makine öğrenimi yöntemleri, volkanosedimanter kayaçlarda eser miktarda bulunan minerallerin tespitinde başarılı olmuştur. Özellikle, XRD verilerinde gürültülerin eliminasyonu ve karmaşık yapıların tanımlanması sağlanmıştır. Karbonat gürültüsünün eliminasyonu, özellikle eser miktarda bulunan minerallerin doğru bir şekilde tespit edilmesine katkıda bulunmuştur.

3. Çapraz Doğrulama: Modelin performansını değerlendirmek için kullanılan çapraz doğrulama yöntemleri, modelin genelleme yeteneğini geliştirmiştir. Geleneksel çapraz doğrulama yöntemlerinden farklı olarak bu çalışmada kullanılan gelişmiş yöntem, modelin aşırı öğrenme riskini azaltmış ve test verileri üzerindeki doğruluğu artırmıştır.

4. Malzeme Bilimi Alanında Katkı: Bu tez, makine öğrenimi ve yapay zeka yöntemlerinin malzeme bilimi ve kristalografi alanında kullanılabileceğini göstermiştir. Özellikle XRD verilerinin işlenmesi ve analizinde, bu tekniklerin sağladığı yenilikçi çözümler, bilimsel araştırmaların doğruluğunu ve güvenilirliğini artırma potansiyeline sahiptir.

Öneriler:

1. Makine Öğrenimi Yöntemlerinin Geliştirilmesi: Gelecekte yapılacak çalışmalarda, XRD verilerinin daha geniş veri setleriyle analiz edilmesi ve daha karmaşık makine öğrenimi modellerinin kullanılması önerilmektedir. Özellikle, derin öğrenme yöntemlerinin bu veriler üzerinde uygulanarak doğruluk oranlarının daha da artırılabilmesi düşünülmektedir.

2. Farklı Veri Setleri Üzerinde Denemeler: Bu çalışmada kullanılan makine öğrenimi tekniklerinin farklı kayaç türleri ve jeolojik oluşumlar üzerinde test edilmesi, bu yöntemlerin genelleme yeteneğini değerlendirmek açısından faydalı olacaktır. Bu tür çalışmalar, makine öğrenimi yöntemlerinin farklı XRD veri setleri üzerinde ne kadar etkili olduğunu anlamamıza olanak sağlayacaktır.

3. Yapay Zeka ile Entegre Yeni Yöntemler: Makine öğrenimi yöntemlerinin yanı sıra, derin öğrenme ve sinir ağları gibi daha gelişmiş yapay zeka tekniklerinin de XRD verilerinin analizine entegrasyonu önerilmektedir. Özellikle, daha büyük veri setlerinde bu tür tekniklerin performansı test edilerek, daha güçlü modeller geliştirilebilir.

4. Endüstriyel Uygulamalar: Bu tezde kullanılan makine öğrenimi algoritmalarının endüstriyel uygulamalarda kullanılabilirliği araştırılmalıdır. Özellikle, malzeme bilimi ve üretim süreçlerinde, XRD analizleri yapay zeka destekli sistemler aracılığıyla hızlandırılabilir ve doğruluğu artırılabilir.

5. Gelişmiş Çapraz Doğrulama Yöntemleri: Modelin doğruluk ve genelleme kabiliyetini artırmak için çapraz doğrulama yöntemlerinin daha da geliştirilmesi ve farklı makine öğrenimi algoritmaları ile test edilmesi önerilmektedir.

Sonuç olarak, bu tez çalışması, XRD verilerinin analizinde makine öğrenimi tekniklerinin kullanılabilirliğini ve etkinliğini ortaya koymuştur. Gelecekte bu alanda yapılacak çalışmalar, daha büyük veri setleri ve gelişmiş modeller kullanılarak daha derin sonuçlar elde edebilir. Makine öğrenimi teknikleri, malzeme bilimi ve kristalografi gibi alanlarda önemli bir potansiyele sahip olup, bu disiplinlerdeki bilimsel araştırmalara önemli katkılar sağlayabilir.

KAYNAKLAR

- Akkus, I. (1965). Erzurum çevresindeki volkanik kayaçların jeolojisi. *Jeoloji Mühendisliği*, 15(18), 57-76.
- Altınli, İ. E. (1963). Türkiye jeoloji haritası: Erzurum bölgesi. Maden Tetkik ve Arama Genel Müdürlüğü Yayınları.
- Biran, A., & Breiner, M. (1996). *Matlab for engineers*. Addison & Wesley Publishing Company.
- Bragg, W. L., & Thomson, J. J. (1914). The diffraction of short electromagnetic waves by a crystal. *Proceedings of the Cambridge Philosophical Society, Mathematical and Physical Sciences*, 17, 43-57.
- Chaudhari, A., & S, K. (2022). Analysis of snow characteristics using satellite images. *Computers & Geosciences*, 163, 105315. <https://doi.org/10.1016/j.cageo.2022.105315>
- Colliot, O. (Ed.). (2023). *Machine learning for brain disorders; Neuromethods (Vol. 197)*. Springer.
- Cullity, B. D. (1956). *Elements of X-ray diffraction*. Addison-Wesley Publishing Company.
- David, K., & Cheney, W. (1996). *Numerical analysis*. Brooks /Cole Publishing Company.
- Dollase, W. A. (1986). Correction of intensities of preferred orientation in powder diffractometry: Application of the March model. *Journal of Applied Crystallography*, 19, 267-272.
- Epp, J. (2016). X-ray diffraction (XRD) techniques for materials characterization. In S. A. Pour (Ed.), *Materials characterization using nondestructive evaluation (NDE) methods* (pp. 81-124). Woodhead Publishing.
- Erdoğan, B., & Soytürk, M. (1974). Erzurum ve çevresinin jeolojik haritalanması üzerine bir çalışma. *Jeoloji Mühendisliği Dergisi*, 18(2), 83-97.
- Erentöz, C. (1945). Erzurum yakınlarındaki petrol imkanları hakkında bir ön rapor. *Türkiye Petrol Araştırmaları Dergisi*, 1(1), 34-45.
- Gadre, C. A., Yan, X., Song, Q., Li, J., Gu, L., Huyan, H., Aoki, T., Lee, S. W., Chen, G., Wu, R., et al. (2022). Nanoscale imaging of phonon dynamics by electron microscopy. *Nature*, 606, 292-297.
- Gentili, S. (2011). *Studio multi-metodico della mineralogia dell'Ignimbrite di Orvieto-Bagnoregio (O.B.I.): Un contributo alla comprensione dei processi sin e post-deposizionali*. Perugia, Umbria, İtalya.

- Greasley, J., & Hosein, P. (2023). Exploring supervised machine learning for multi-phase identification and quantification from powder X-ray diffraction spectra. *Journal of Materials Science*, 58, 5334-5348.
- Hanawalt, J. D. (1976). Phase identification by X-ray powder diffraction evaluation of various techniques. *Advances in X-ray Analysis*, 20, 63-73.
- Hillier, S. (2000). Accurate quantitative analysis of clay and other minerals in sandstones by XRD: Comparison of a Rietveld and a reference intensity ratio (RIR) method and the importance of sample preparation. *Clay Minerals*, 35, 291-302.
- Ige, A. O., & Mohd Noor, M. H. (2022). A survey on unsupervised learning for wearable sensor-based activity recognition. *Applied Soft Computing*, 127, 109363.
- Innocenti, F., Manetti, P., Peccerillo, A., & Poli, G. (1982). Anatolia ve çevresinde Neojen volkanizmasının kökeni. *Jeoloji ve Madencilik Dergisi*, 25(3), 131-144.
- Ishitsuka, K., Ojima, H., Mogi, T., Kajiwara, T., Sugimoto, T., & Asanuma, H. (2022). Characterization of hydrothermal alteration along geothermal wells using unsupervised machine-learning analysis of X-ray powder diffraction data. *Earth Science Informatics*, 15, 73-87.
- Jabnouni, H., Arfaoui, I., Cherni, M. A., Bouchouicha, M., & Sayadi, M. (2022). Machine learning based classification for fire and smoke images recognition. In *Proceedings of the 2022 8th International Conference on Control, Decision and Information Technologies (CODIT'22)* (pp. 425-430). Istanbul, Turkey.
- Jenkins, R., & Snyder, R. L. (1996). *Introduction to X-ray powder diffractometry*. Wiley.
- Li, Y., Beck, R., Huang, T., Choi, M. C., & Divinagracia, M. (2008). Scatterless hybrid metal-single-crystal slit for small-angle X-ray scattering and high-resolution X-ray diffraction. *Journal of Applied Crystallography*, 41, 1134-1139.
- Li, Z., Zhang, Y., & Wu, H. (2023). Advances in machine learning-based geological prediction. *Computers & Geosciences*, 172, 105312. <https://doi.org/10.1016/j.cageo.2023.105312>
- Mahadevkar, S. V., Khemani, B., Patil, S., Kotecha, K., Vora, D. R., Abraham, A., & Gabralla, L. A. (2022). A review on machine learning styles in computer vision-techniques and future directions. *IEEE Access*, 10, 107293-107329.
- McBirney, A. (2007). *Igneous petrology*. Johns & Bartlett Publishers.
- Meunier, A. (2005). *Clays*. Springer.
- Nuffield, E. W. (1966). *X-ray diffraction methods*. John Wiley & Sons.
- Oviedo, F., Ren, Z., Sun, S., Settens, C., Liu, Z., Hartono, N. T. P., Ramasamy, S., DeCost, B. L., Tian, S. I. P., Romano, G., et al. (2019). Fast and interpretable

- classification of small X-ray diffraction datasets using data augmentation and deep neural networks. *NPJ Computational Materials*, 5, 60.
- Pane, S. A., & Sihombing, F. M. H. (2021). Classification of rock mineral in field X based on spectral data (SWIR & TIR) using supervised machine learning methods. *IOP Conference Series: Earth and Environmental Science*, 830, 012042.
- Ramakrishnan, R., Dral, P. O., Rupp, M., & Von Lilienfeld, O. A. (2014). Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1, 140022.
- Ramesh, K., & Gupta, A. (2022). Fault detection in hydrocarbon pipelines using supervised learning. *Journal of Applied Geophysics*, 192, 105463. <https://doi.org/10.1016/j.cageo.2023.105463>
- Reventos, M. M., & Descarrega, J. M. A. (2012). Mineralogy and geology: The role of crystallography since the discovery of X-ray diffraction in 1912. *Revista de la Sociedad Geológica de España*, 25, 133-143.
- Scarlett, N. V. Y., Madsen, I. C., Manias, C., & Retallack, D. (2001). On-line X-ray diffraction for quantitative phase analysis: Application in the Portland cement industry. *Powder Diffraction*, 16, 71-80.
- Scherrer, P. (1918). Estimation of the size and internal structure of colloidal particles by means of Röntgen. *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 2, 96-100.
- Shcherbatov, I., Lisin, E., Rogalev, A., Tsurikov, G., Dvorak, M., & Strielkowski, W. (2021). Power equipment defects prediction based on the joint solution of classification and regression problems using machine learning methods. *Electronics*, 10(3145).
- Sun, H., Wang, Y., & Zhao, L. (2023). Improved X-ray diffraction data analysis for mineral identification. *Journal of Mineralogy*, 197, 103267. <https://www.sciencedirect.com/science/article/pii/S009830042300170X>
- Surdu, V., & György, R. (2023). X-ray diffraction data analysis by machine learning methods. Bucharest/Romania.
- Tokel, S. (1979). Erzurum-Kars bölgesinde Neojen çöküntülerle ilişkili volkanizmanın incelenmesi. *Yerbilimleri*, 4(2), 112-129.
- Uvarov, V., & Popov, I. (2007). Metrological characterization of X-ray diffraction methods for determination of crystallite size in nano-scale materials. *Materials Characterization*, 58, 883-891.
- Vecsei, P. M., Choo, K., Chang, J., & Neupert, T. (2019). Neural network based classification of crystal symmetries from X-ray diffraction patterns. *Physical Review B*, 99(245120).

- Xie, T., & Grossman, J. C. (2018). Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(145301).
- Xu, D., Chen, L. Q., Yu, C., Zhang, S., Zhao, X., & Lai, X. (2023). Failure analysis and control of natural gas pipelines under excavation. *Journal of Mineralogy*, 197, 103267.
- Xu, R., & Kam, H. L. (1992). Interaction of defects and strain in semiconductor materials. *Journal of Crystal Growth*, 123(1-2), 243-249. <https://www.sciencedirect.com/science/article/pii/002190459290081X>

