

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

FOURIER DÖNÜŞÜMLÜ KIZILÖTESİ SPEKTROSKOPİSİ VE
MAKİNE ÖĞRENMESİ KULLANILARAK POLİOLLERDE
BAZI KARAKTERİSTİK ÖZELLİKLERİN ÖNGÖRÜLMESİ

YUSUF SERDAR BEKDAŞ

YÜKSEK LİSANS TEZİ
KİMYA MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMAN: DR. ÖĞRETİM ÜYESİ ÖZGÜN YÜCEL

OCAK 2024

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

FOURIER DÖNÜŞÜMLÜ KIZILÖTESİ
SPEKTROSKOPİSİ VE MAKİNE ÖĞRENMESİ
KULLANILARAK POLİOLLERDE BAZI
KARAKTERİSTİK ÖZELLİKLERİN ÖNGÖRÜLMESİ

YUSUF SERDAR BEKDAŞ

YÜKSEK LİSANS TEZİ
KİMYA MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMAN: DR. ÖĞRETİM ÜYESİ ÖZGÜN YÜCEL

OCAK 2024

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL

**PREDICTION OF SOME CHARACTERISTIC
FEATURES IN POLYOLS USING FOURIER
TRANSFORMED INFRARED SPECTROSCOPY AND
MACHINE LEARNING**

YUSUF SERDAR BEKDAŞ

**A THESIS OF MASTER OF SCIENCE
DEPARTMENT OF CHEMICAL ENGINEERING**

ADVISOR: ASSISTANT PROFESSOR ÖZGÜN YÜCEL

JANUARY 2024

YÜKSEKLİSANS/DOKTORA JÜRİ ONAY FORMU

GTÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulunun 10/01/2024 tarih ve 2024/05 sayılı kararıyla oluşturulan jüri tarafından 29/01/2024 tarihinde tez savunma sınavı yapılan Yusuf Serdar BEKDAŞ'ın tez çalışması Kimya Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Dr. Öğr. Üyesi Özgün YÜCEL

ÜYE

: Dr. Öğr. Üyesi Emrah KIRTIL

ÜYE

: Prof.Dr.Hasan SADIKOĞLU

ONAY

Gebze Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Yönetim Kurulunun
.....tarih ve/.....sayılı kararı.

İMZA/MÜHÜR



Īthaf / Dedication

Aileme

ÖZET

Günlük hayatta poliüretanları birçok alanda kullanmaktayız. Bu poliüretanların ve diğer bir çok ürünün hammaddesi olan poliöl ürünlerinin bazı ana karakteristiklerini bulunmaktadır. Bu çalışma, bu karakteristikleri belirlemek amacıyla çeşitli makine öğrenme yöntemlerini incelemektedir.

Piöl ürünleri, endüstriyel süreçlerde yaygın olarak kullanılan bir kimyasal bileşenlerdir ve bu ürünlerin ana karakteristiklerinin doğru bir şekilde tahmin edilmesi, üretim süreçlerinin yönetimi, kalite kontrolü ve ürün geliřtirmenin temelini oluşturur.

Çalışma, piöl ürünlerinin ana karakteristiklerini etkileyebilecek faktörleri belirlemek amacıyla derinlemesine bir literatür taraması yapar. Daha sonra, çeşitli kaynaklardan elde edilen geniş veri setlerini kullanarak, dört farklı makine öğrenme yöntemi olan lineer regresyon, Gaussian Process Regresyon (GPR), Destek Vektör Makinesi (SVM) ve Yapay Sinir Ağları (ANN) ile piöl karakteristiklerini tahmin etmek için kapsamlı analizler gerçekleştirir.

Her bir modelin performansı, kök-kare ortalama hatası (RMSE), R-kare değeri ve diğer metrikler kullanılarak ayrıntılı bir şekilde değerlendirilir. Sonuçlar, hangi makine öğrenme yönteminin piöl ürünlerinin ana karakteristiklerini en iyi tahmin ettiğini belirlemeye yardımcı olacaktır.

Ayrıca, bu tez, endüstriyel süreçlerin verimliliğini artırmayı hedefleyen mühendisler ve endüstri profesyonelleri için değerli bir kaynak oluşturacaktır. Piöl ürünlerinin kalite kontrolü ve süreç iyileřtirmesi için gelişmiş veri analitiğı tekniklerinin nasıl kullanılabileceğini vurgulayarak, gelecekteki çalışmalar için temel oluşturacaktır. Sadece poliollerin değil tüm kimyasalların birçok karakteristik özelliğini (Derişim, pH, mol kütlesi, bağ yapısı, öz ısı, öz kütle, ışıık geçirgenlik vb.) tahmin etmek için de kaynak olacaktır.

Sonuç olarak, bu çalışma, makine öğrenme yöntemlerinin piöl ürünlerinin ana karakteristiklerini tahmin etmek için nasıl kullanılabileceğini ve endüstriyel süreçlerin geliřtirilmesine nasıl katkı sağlayabileceğini inceliyor. Bu çalışma, endüstriyel uygulamalar ve araştırma alanlarında değerli bir bilgi kaynağı olacaktır.

Anahtar Kelimeler: Makine öğrenmesi, FTIR, Piöl, OH sayısı

ABSTRACT

We use polyurethane in many areas in daily life. Polyol products, which are the raw materials of these polyurethanes and many other products, have some main characteristics. This study examines various machine learning methods to identify these characteristics. Polyol products are commonly used chemical components in industrial processes, and accurately predicting their key characteristics is crucial for process management, quality control, and product development.

The study commences with an in-depth literature review to identify the factors that may influence the key characteristics of polyol products. Subsequently, by utilizing extensive datasets from various sources, the research employs four different machine learning methods linear regression, Gaussian Process Regression (GPR), Support Vector Machines (SVM), and Artificial Neural Networks (ANN) - to predict the key characteristics of polyol products, conducting comprehensive analyses in the process.

Each model's performance is assessed in detail using metrics such as Root Mean Square Error (RMSE), R-squared value, and others. The results provide valuable insights into which machine learning method offers the most accurate predictions for the key characteristics of polyol products.

Furthermore, this thesis serves as a valuable resource for engineers and industry professionals aiming to enhance industrial processes. By highlighting how advanced data analytics techniques can be applied for quality control and process improvement of polyol products, it paves the way for future research. It will also be a resource for estimating many characteristic properties (concentration, pH, molar mass, bond structure, specific heat, specific mass, light transmittance, etc.) not only of polyols but of all chemicals.

In conclusion, this study delves into the application of machine learning methods to predict the key characteristics of polyol products and how they can contribute to the enhancement of industrial processes. It stands as an invaluable source of knowledge for industrial applications and research fields.

Keywords: Machine learning, FTIR, Polyol, OH number

TEŞEKKÜR

Başta, yüksek lisans eğitimimde ve akademik hayatımda desteğini ve yardımlarını hiçbir zaman esirgemeyip tecrübesi ve bilgisi ile bu çalışmanın oluşmasının yolunu açan danışmanım Dr. Öğretim Üyesi Özgün YÜCEL' e, bütün yaşamım boyunca yanımda olan, bilgi ve tecrübelerini benimle paylaşan canım aileme ve göstermiş olduğu desteklerinden dolayı çalıştığım kuruma en içten teşekkürlerimi sunarım.



İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	v
ABSTRACT	vi
TEŞEKKÜR	vii
İÇİNDEKİLER	viii
SİMGELER ve KISALTMALAR DİZİNİ	ix
ŞEKİLLER DİZİNİ	x
TABLOLAR/ÇİZELGELER DİZİNİ	xi
1. GİRİŞ	1
1.1. Tezin Amacı, Katkısı ve İçeriği	1
1.2. Literatür Tarama	2
2. MAKİNE ÖĞRENMESİ VE ALGORİTMALARI	6
2.1. Karar Ağacı (Decision Tree)	7
2.2. Rasgele Orman (Random Forest)	8
2.3. Gauss Süreci (Gaussian Processes)	10
2.4. Genelleştirilmiş Katkı Modeli (Generalized Additive Model)	10
2.5. Destek Vektör Makinesi (Support Vector Machine)	11
2.6. Yapay Sinir Ağları (Artificial Neural Networks)	12
3. MALZEME VE YÖNTEM	13
3.1. Ön modeller ve Nitelik çıkarımı	15
3.2. Model performansları	21
4. SONUÇLAR ve YORUMLAR	28
KAYNAKLAR	29
ÖZGEÇMİŞ	33
TEZ ÇALIŞMASI KAPSAMINDA YAPILAN YAYINLAR	34

SİMGELER ve KISALTMALAR DİZİNİ

ANN	: Artificial neural network
DVM	: Destek Vektör Makinesi (Support Vector Machine)
FTIR	: Fourier Transformed Infrared Spectroscopy
GAM	: Generalized Additive Model
GTÜ	: Gebze Teknik Üniversitesi
GPR	: Gaussian Processes for Regression
KA	: Karar ağacı (Decision Tree)
PU,PUR	: Poliüretan
RO,RF	: Rasgele Orman (Random Forest)
PCA	: Principal Component Analysis (Temel Bileşenler Analizi)
RMSE	: Root Mean Square Error (Ortalama Karekök Sapması)
SVM	: Support Vector Machine
MAE	: Mean absolute error (Ortalama mutlak hata)
MRMR	: Minimum redundancy (Minimum artıklık)
MSE	: Mean Squared Error (Ortalama Hata Karesi)
YSA	: Yapay Sinir Ağları(Convolutional Neural Network)

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 1.1: Yapay Zeka ile ilgili yayınların yüzdesi .	3
Şekil 1.2: FT-NIR spektrometresi ile alınan NIR spektrumu.	4
Şekil 1.3: Test seti için PLSR'nin çözümlemesi. Şekil PLSR için test seti örneğinin yanıtları.	5
Şekil 2.1: Öğrenme süreçleri.	6
Şekil 2.2: Makina Öğrenmesi Sınıflandırması.	6
Şekil 2.3: Karar ağacı.	8
Şekil 2.4: Rasgele orman şeması.	9
Şekil 2.5: GPR grafiği.	10
Şekil 2.6: Doğrusal bir model kullanmanın GAM'a karşı etkisi .	11
Şekil 2.7: DVM sınıflandırma şeması.	12
Şekil 2.8: YSA yapısı.	12
Şekil 3.1: Denemelerde kullanılan malzemelerin FTIR spektrumları.	14
Şekil 3.2: İkinci derece SVM grafiği.	15
Şekil 3.3: Üç katmanlı Sinir Ağı grafiği.	16
Şekil 3.4: Üstel GPR grafiği.	17
Şekil 3.5: Doğrusal (Lineer) grafiği.	18
Şekil 3.6: Denemelerde kullanılan FTIR pik bölgeleri.	19
Şekil 3.7: F Test algoritmasına göre özellik - önem puanları grafiği.	20
Şekil 3.8: RReliefF algoritmasına göre özellik - önem puanları grafiği.	20
Şekil 3.9: Veri Kümesi Eğitim ve değerlendirme.	21
Şekil 3.10: SVM tahminleri ve gerçek değerler.	22
Şekil 3.11: SVM parametreleri.	22
Şekil 3.12: GP tahminleri ve gerçek değerler.	23
Şekil 3.13: GP parametreleri ve aralıkları.	24
Şekil 3.14: ANN tahminleri ve gerçek değerler.	25
Şekil 3.15: ANN parametreleri ve aralıkları.	25
Şekil 3.16: DT tahminleri ve gerçek değerler.	26
Şekil 3.17: DT parametreleri.	26

TABLolar/ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Tablo 1.1: 2000'den 2020'ye kadar yapay zeka ile ilgili yayınların ülke/bölge ve şirkete göre dağılımı.	3
Tablo 1.2: Modelin değerlendirme endeksleri.	5
Tablo 3.1: Bazı FTIR piklerinin görülme bölgeleri.	13
Tablo 3.2: Çalışmada kullanılan örnekler ve OH aralıkları.	14
Tablo 3.3: MRMR testine göre en anlamlı sütunlar.	19
Tablo 3.4: F Test e göre sonuç için en anlamlı sütunlar.	20
Tablo 3.5: RReliefF e göre sonuç için en anlamlı sütunlar.	21
Tablo 3.6: Yapılan benzetimler sonucu elde edilen sonuçlar.	21
Tablo 3.7: SVM parametreleri.	23
Tablo 3.8: GP parametreleri ve aralıkları.	24
Tablo 3.9: ANN parametreleri ve aralıkları.	26
Tablo 3.10: DT parametreleri.	27
Tablo 3.11: Denemelere ait ortalama değerler.	27

1. GİRİŞ

Teknolojinin hayatımızın vazgeçilmez bir ögesi olduğu artık reddedilemez bir gerçek. Günlük hayatta, birçok alanda, birçok üründe, kullandığımız neredeyse her nesnede teknolojinin nimetlerinden faydalanılmıştır ve faydalanmaya devam edilmektedir. Bizler teknolojiyi ve gelişen endüstriyi sürekli takip etmekte ve üretimlerimize entegre etmek durumundayız.

Son yıllarda artan yapay zekâ teknolojisi de her alanda merak uyandırmaktadır. İngiliz bilimci Alan Turing'in sorduğu “Makineler düşünebilir mi?” sorusuyla [Turing, 1950] başlayan, günümüzde zirveye ulaşan yapay zekâ teknolojisi ile bilgisayarlar ve benzeri cihazlar insan gibi düşünüp karar verme yetisine sahip olmuştur [İnce vd., 2021]. Burada, var olan veriler bilgisayara işlenerek çeşitli uygulamalar yardımıyla yeni sorulara cevaplar almak amaçlanmaktadır. Ucu bucağı bilinmemekle beraber, yapay zekâ ile bugünkü bilgi ile telefonların, kameraların yüz tanıma sistemleri, bankacılık uygulamalarının asistanları, yol bulma uygulamaları, otomatik insansız araçlar, askeri alanda kullanılan birçok cihaz bunların sadece birkaçıdır [Gümüş vd., 2020; Çetiner vd., 2022; Königstorfer ve Thalmann, 2020].

Kimya teknolojisinde ise kimyasalların özellikleri başta olmak üzere, sentezlerin, deneylerin sonuçlarını tahmin etmek amacıyla kullanılmak üzere çalışmalar devam etmektedir [Kapdan, 2022].

1.1 Tezin Amacı, Katkısı ve İçeriği

Tezin amacını anlamak için öncelikleri polioller dolayısıyla poliüretanları tanımak gerekir. PU, PUR, bir poliöl ve bir izosiyanatın karışımı ile meydana gelen plastik malzemelerdir [Yılğör, 2020]. Oldukça geniş bir kullanım alanı vardır. Günlük hayatta en çok köpük olarak kozmetik, yatak, mobilya, izolasyon ve ayakkabılarda kullanılmaktadır [Baysal ve Kasapbaşı, 2017]. 1937'de Prof. Dr. Otto Bayer ve çalışma arkadaşları tarafından Alman I. G. Farben Industrie laboratuvarlarında keşfedildi [Bayer, 1947]. Poliüretanları birçok şekilde sınıflandırmak mümkün. Bu sınıflandırmayı yaparken nereye hitap edeceğini bilmek gerekir. Örneğin esnek

köpükleri sertlik yumuşaklığa göre, dolgulu dolgusuz yapısına göre, yanmazlık durumlarına göre, kimyasal yapısına, yoğunluğuna veya kullanılan sektöre göre bile sınıflandırma yapılabilir [Choi vd., 2022]. Bir diğer sınıflandırma yöntemi ise hammadde özelliğine göredir. Poliüretan oluşması için bir izosiyanat ve bir alkol yeterlidir. Reaksiyonu hızlandırmak için, sistemin ayakta durmasını sağlamak için, bazı ekstra özellikler eklemek için yardımcı kimyasallar kullanılabilir [Eroğlu, 2008]. Sistemde oran olarak en çok kullanılan polioller burada devreye girmektedir. Polioller; bir alkol ile fonksiyonalize edilmiş düşük mol kütleli (3000-5000 dalton arası) polimerler olarak adlandırılabilir [Çalıkoğlu, 2017]. Petrol esaslı malzemelerdir genellikle, fakat üretim maliyeti, enerji maliyeti ve tehlikeli bir proses olması sebebiyle daha çevreci olan, doğal yağlardan yapılan poliollerin kullanımı da son zamanlarda artış göstermiştir [Yargıç vd., 2020]. Polioller en önemli karakteristikleri OH sayısı su oranı, asit sayısı ve viskozitesidir [Yenen, 2020]. En çok kullanılan polioller poliester ve polieterdir.

Bu çalışmada glikollerin bazı özelliklerinin FTIR ile tespit edilmesi amaçlanmıştır. Bunun için yapay zekanın bazı algoritmaları kullanılmıştır. Ayrıca şu sorulara da cevaplar aranmıştır:

FTIR ile poliollerin karakteristik özellikleri tespit edilebilir mi?

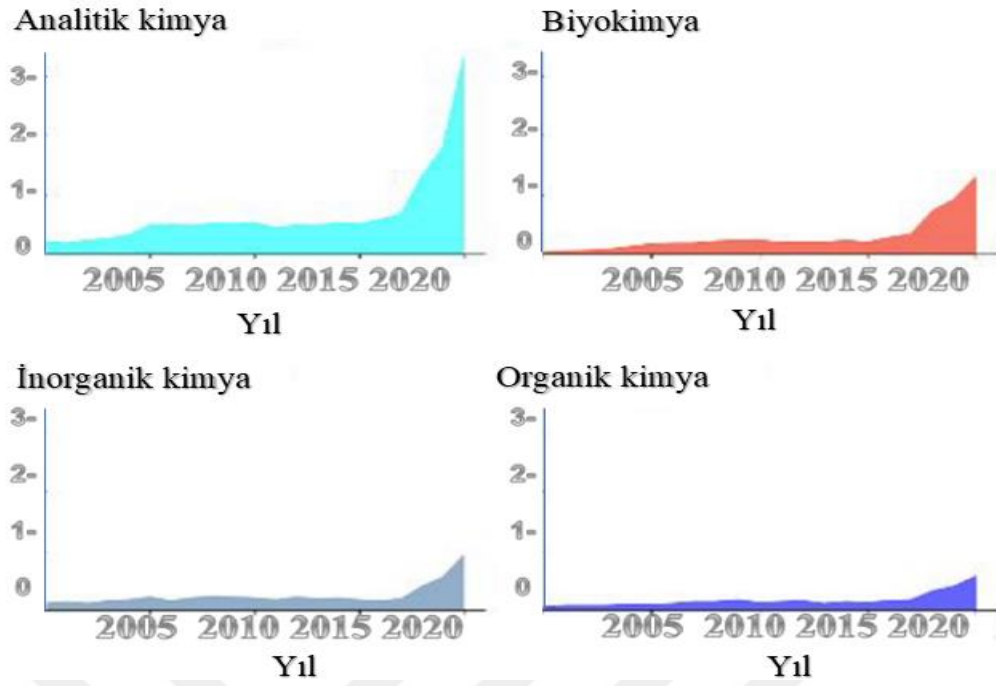
Yapay zekâ ile FTIR birlikte çalışabilir mi?

Hangi yapay zekâ algoritması bu tür araştırmalar için daha uygun?

1.2 Literatür Tarama

Kimya mühendisliğinde yapay zekanın kullanılması, endüstride ve akademide, bilim dünyasında son yıllarda hızla artmaktadır. Öncelikle yapay zekâ dünyasında kimya endüstrisi ile ilgili yapılan bazı istatistikî verilere bakalım;

2000 li yıllarda bilgisayar teknolojisinin iyiden iyiye yaygınlaşması ile yapay zeka ile ilgili yayınlar da yazılmaya başlanmıştır. Sektörde yapılan yayınların yüzdeleri **Şekil 1.1** de verilmiştir.



Şekil 1.1: Yapay Zeka ile ilgili yayınların yüzdesi [Zachary vd., 2021].

Bu yayınların bir çoğu Çin, Amerika ve Hindistan’da yayınlanırken Patentlerin büyük kısmı teknoloji şirketleri LG, IBM ve Fanuc tarafından yazılmıştır. Bununla ilgili diğer veriler **Tablo 1.1** de belirtilmiştir.

Tablo1.1: 2000'den 2020'ye kadar yapay zeka ile ilgili yayınların ülke/bölge ve şirkete göre dağılımı [Zachary vd., 2021].

Ülke	%	Şirket	Patent sayısı
Çin	26,5	LG	263
Amerika	17,2	IBM	222
Hindistan	5,8	Fanuc	151
İran	5,5	Siemens	115
İngiltere	3,9	Ping An Tec.	95
Almanya	3,8	Philips	94
Japonya	3,2	Samsung	86

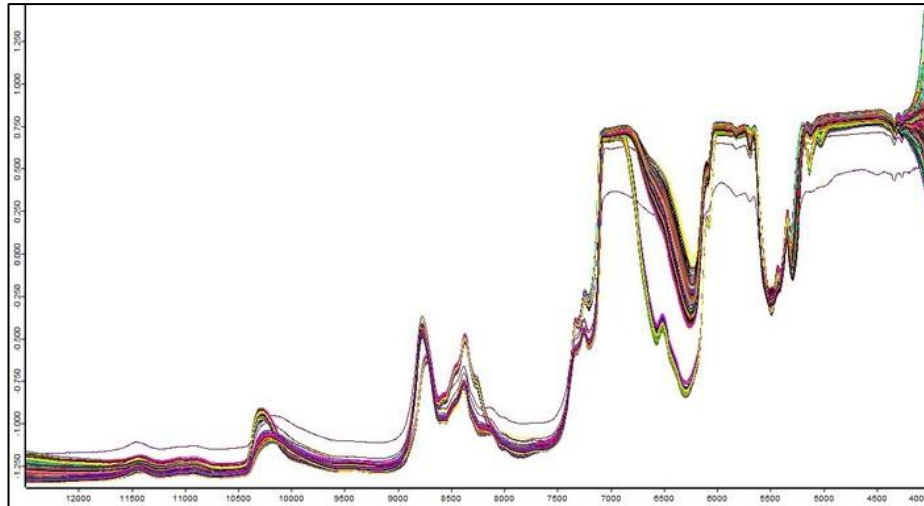
Literatürde konuyla alakalı bazı çalışmaları sıralayacak olursak;

Ian Robertson 2017 yılında “The Determination of Hydroxyl Number in Polyols Using the Spectrum Two N FT-NIR Spectrometer” adlı çalışmasında NIR tabanlı bir uygulama ile poliollerde OH sayısının analizi için titrasyonla yapılan referans yöntemine alternatif yöntem mümkün olduğunu, benzer doğruluk yakalayarak saptamıştır. NIR spektroskopisinin OH numarası tespitinde, mevcut yöntemlerden önemli ölçüde hızlı ve daha kolay olduğunu belirtmiştir.

2014 yılında Bonnie Tay Yen Ping ve arkadaşlarının “Measurement Of Hydroxyl Value Of Palm-Based Polyol By Attenuated Total Reflectance/Fourier Transform Infrared Spectroscopy” adlı çalışmasında doğal yağlardan yapılan poliollerin OH sayısını tahmin etmeye yönelik bazı algoritmalar kullanmıştır. O da Robertson gibi bu yöntemin daha hızlı olduğunu çünkü yaş analize gerek kalmadığını belirtmiştir. Buna rağmen kalibrasyon için yaş analizlere her zaman ihtiyaç vardır.

“Evaluation of Techniques for Determination of Hydroxyl Value in Materials for Different Industrial Applications” adlı çalışmada da Laís L. J. ve arkadaşları Yüzey aktif maddelerde (IOH aralığı 46–104 mg KOH/g) ve polyesterlerde (IOH aralığı 37–310 mg KOH/g) MIR ve NIR bölgelerinde türevlendirme olmaksızın IOH tayini için FT-IR metodolojileri geliştirmiştir. FT-IR metodolojisi, IOH değerleriyle iyi bir doğrusallık gösterdiğini ve ölçülen değerlerde ıslak kimya yöntemiyle gözlemlenenlerden daha iyi sonuçlar verdiğini belirtmiş, daha hızlı olduğu da kanıtlanmıştır.

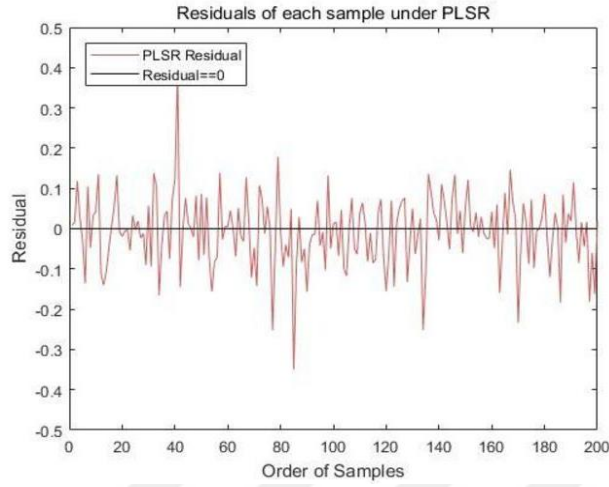
Makine öğrenmesi algoritmaları kullanılarak kimyasalların özellikleri tespit edilebildiği bir diğer çalışmada yakın kızılötesi spektroskopi ve kısmi en küçük kareler regresyonu kullanılarak polifenilen eter sentezi sırasında o-kresol derişimi tahmin etmek amaçlanmış [Chen vd., 2019]. Şekil 1.3 de, FT-NIR spektrometresi tarafından toplanan, önceden işlenmemiş spektral veri kümesini göstermektedir.



Şekil 1.2: FT-NIR spektrometresi ile alınan NIR spektrumu.

Spektrumlardan görüldüğü gibi büyük çoğunluğu neredeyse aynıdır. Uygulama bu durumlarda %95-99 aralığında benzerlik vermektedir. Bu da kimyasalların

karakteristiklerini belirlemeyi zorlaştırmaktadır. Burada makine öğrenmesi ile algoritmaları devreye girmektedir.



Şekil 1.3: Test seti için PLSR'nin çözümlemesi. Şekil PLSR için test seti örneğinin yanıtları.

Bu süreçteki kritik adımlar, orijinal NIR spektral verilerinin ön işlenmesi ve NIR spektral verilerinin değişken seçimidir.

Tablo 1.2: Modelin değerlendirme endeksleri. Tablo, 10 kat çapraz doğrulama model değerlendirme göstergeleri.

	Yorumlanabilirlik	Tahmin doğruluğu	
PLSR	R²	R_p	MRPECV
	0,98332	0,9916	0,01106

Sistematiik olmayan arka plan, NIR spektrumunun saçılma sapmasına veya taban çizgisi kaymasına yol açar. İlk kritik adım olarak uygun NIR spektral işleme yöntemi (türev, normalleştirme, çok değişkenli saçılma düzeltmesi vb.), arka planı kaldırır ve sinyal-gürültü oranını artırır.

Yüksek boyutlu NIR spektral verilerinde mevcut olan çoklu doğrusal korelasyonlar, gereksiz hesaplamaya neden olur ve modelin çalışma verimliliğini azaltır. Ayrıca, polifenil eter ürünlerindeki o-kresol içeriğiyle zayıf veya hatta hiç ilişkisi olmayan gereksiz değişkenler vardır. İkinci kritik adım olarak PLS'ye dayalı değişken seçimi, yukarıdaki iki sorunu çözer; bu da modelin karmaşıklığını azaltır ve modelleme için OLSR'nin doğrudan kullanımına kıyasla daha gerçekçi kılar.

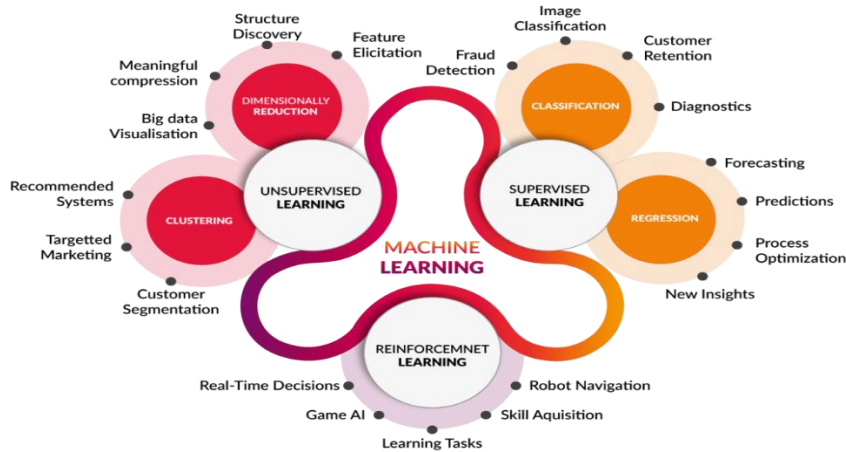
2. MAKİNA ÖĞRENMESİ VE ALGORİTMALARI

Son yılların en çok konuşulan konusu yapay zekâ ve onun bir dalı olan Makine Öğrenmesi, bilgisayarların birtakım verileri kullanarak öğrenme işlemini gerçekleştirmelerini amaçlayan algoritma ve tekniklerle ilgili bir çalışma alanıdır. Kökleri istatistiğe dayanan veri kümelerinden sonuç üretmemize olanak sağlayan tekniklere sahiptir. Bu bilgiler, verilerin Makine Öğrenmesi algoritmaları aracılığıyla öğrenilmesi neticesinde tahminlerde bulunmamızı sağlayan modellerden elde edilir. Öğrenmeyi gerçekleştirmek için birçok algoritma geliştirilmiştir. Özetleyecek olursak bilgi (veriler) sisteme tanıtılır. Uygun algoritma seçilir ve gerisini makine hesaplar.



Şekil 2.1: Öğrenme süreçleri.

Makina öğrenmesi karmaşık görülse de ilgi çekici ve uçsuz bucaksız olsa da çalışmada birkaç algoritmasını kullanacağız. Ana başlıklar olarak inceleyecek olursak Gözetimsiz öğrenme (Unsupervised learning), Gözetimli öğrenme (Supervised learning) ve Pekiştirmeli öğrenme (Reinforcement learning) olarak 3 başlıkta inceleyebiliriz.



Şekil 2.2: Makina Öğrenmesi Sınıflandırması [Ersoz ve Kasapşahin, 2021].

Gözetimli öğrenme (Supervised learning): Bir dizi X girdi değişkeni ile bir çıktı değişkeni Y arasında bir ilişki öğrenmeyi ve bu ilişkiyi bilinmeyen verilerin çıktılarını denemeye dayanır [Cunningham vd., 2008].

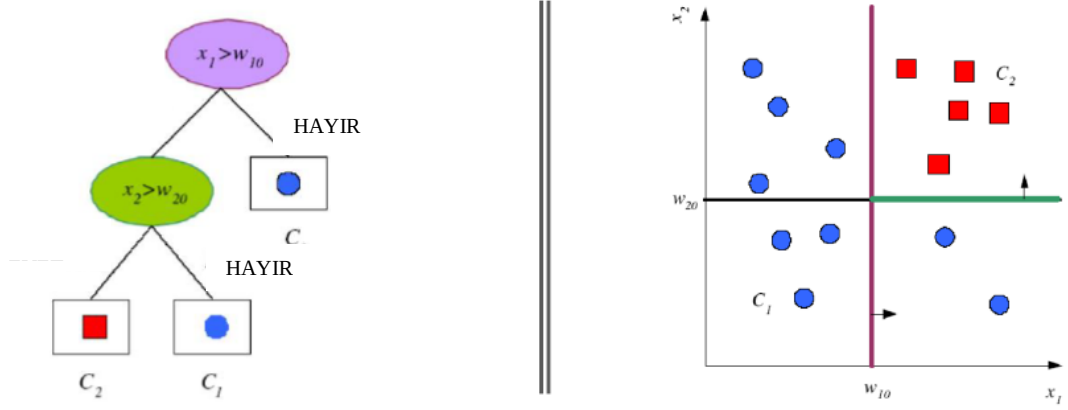
Gözetimsiz öğrenme (Unsupervised learning): Burada denetimli öğrenmeden farklı olarak "öğretmen olmadan öğrenme" olarak tanımlayabiliriz. Bu durumda, $Pr(X)$ ortak yoğunluğuna sahip rastgele bir p -vektörü X 'in bir dizi N gözlemi $(x_1, x_2, \dots, x_N)$ vardır. Burada, her gözlem için doğru cevaplar veya hata derecesi sağlayan bir gözetken veya öğretken olmadan direkt bu olasılık yoğunluğunun özelliklerini anlamaktır [Hastie vd., 2009].

Pekiştirmeli öğrenme (Reinforcement learning): Temel olarak ortamdan alınan geri bildirimlere bağlı olarak sonuç oluşturan bir fonksiyonun veri kullanılarak eğitilmesini amaçlar. Yapısal olarak incelendiğinde ajan, durum ve ortam gibi 3 temel bileşenden meydana gelir. Öğrenilen fonksiyonda gözlenen durum karşısında uzun vadede en başarılı sonucu getirecek olan aksiyonun seçilmesi sağlanır [Coşkun ve İtik, 2023].

Bu çalışmada kullanacağımız algoritmalar birer supervised learning olan Regresyon ve Sınıflandırmadır. Bunları da kısaca örnekleyecek olursak basit bir şekilde problemlerde tahmin edeceğimiz değer sayısal(maaş, boy vb) ise regresyon, hasta/sağlıklı, karlı/güneşli, başarılı/başarısız, pozitif/negatif gibi kategorik değerler ise sınıflandırma diyebiliriz [Kandar, 2021].

2.1. Karar Ağacı (Decision Tree)

Karar ağaçları(KA, DT), bir problemi çözmek için problem yapısına göre bir gövde, ağaç yapısı şeklinde bir sınıflandırma modeli oluşturur. Ağacın yapısının ve kuralların anlaşılabilir olması, algoritmanın uygulanmasını kolaylaştırır. Karar ağacı, birbiri ardına basit karar verme işlemlerinden oluşur.

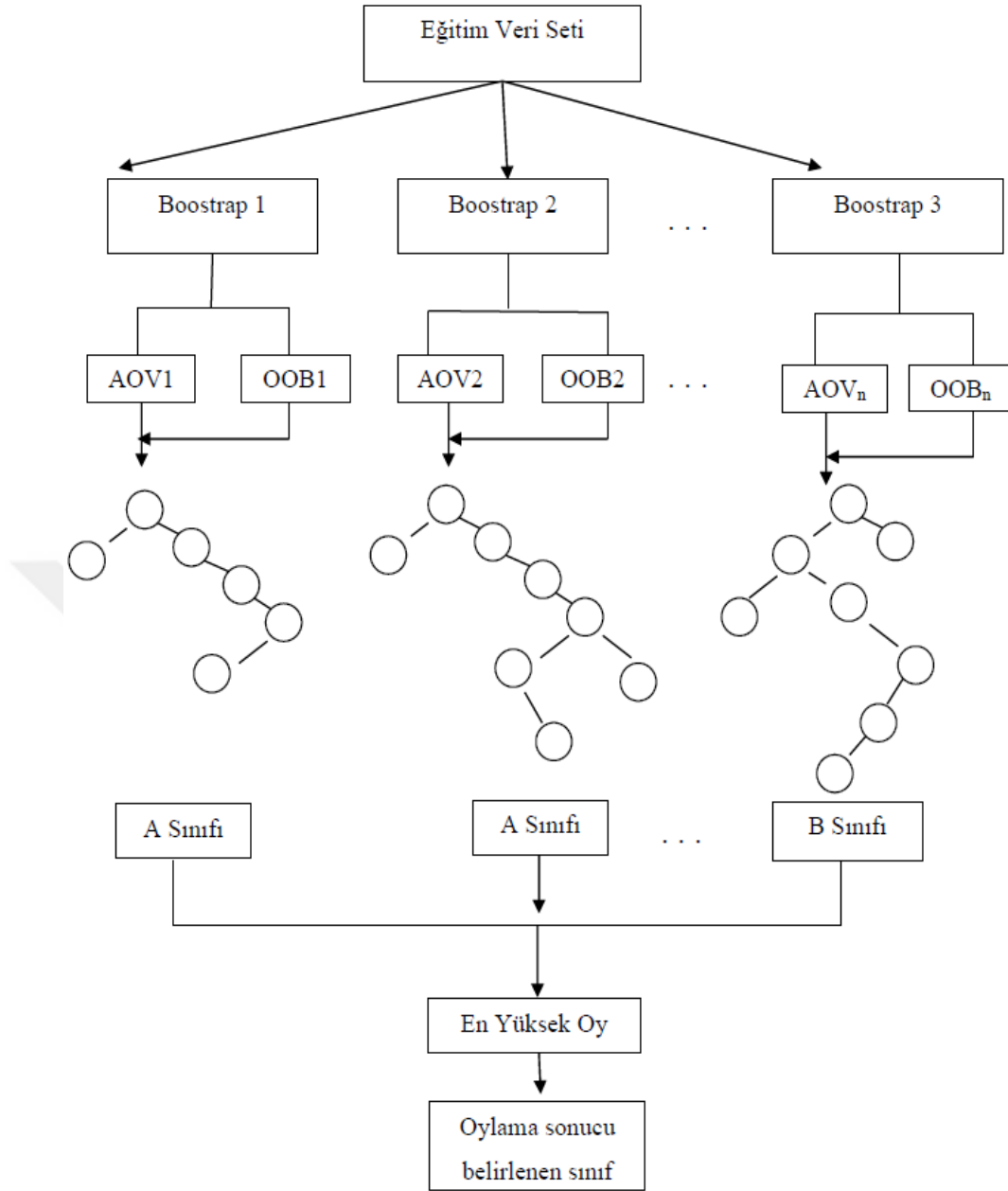


Şekil 2.3: Karar ağacı [Alpaydın, 2014].

Ağaç şeklinde gösterildiği gibi algoritmanın mantığında düğümler ve dallar bulunur. Algoritma çalışırken ilk düğümden başlayarak sorular sormaya başlanır ve son elemana ulaşana kadar dallanmaya devam eder [Şengür, 2013].

2.2. Rasgele Orman Algoritması (Random Forest)

Rasgele Orman (RO, RF) çok sayıda karar ağacının sonuçlarını çoğunluk oyuyla karar verip toplayan, topluluk sınıflandırma algoritmasıdır. Topluluk sınıflandırmada (ensemble learning) birden çok topluluğun ortaya koyduğu sonuçlar değerlendirilip bir araya getirilerek, tek bir karar verilmektedir.



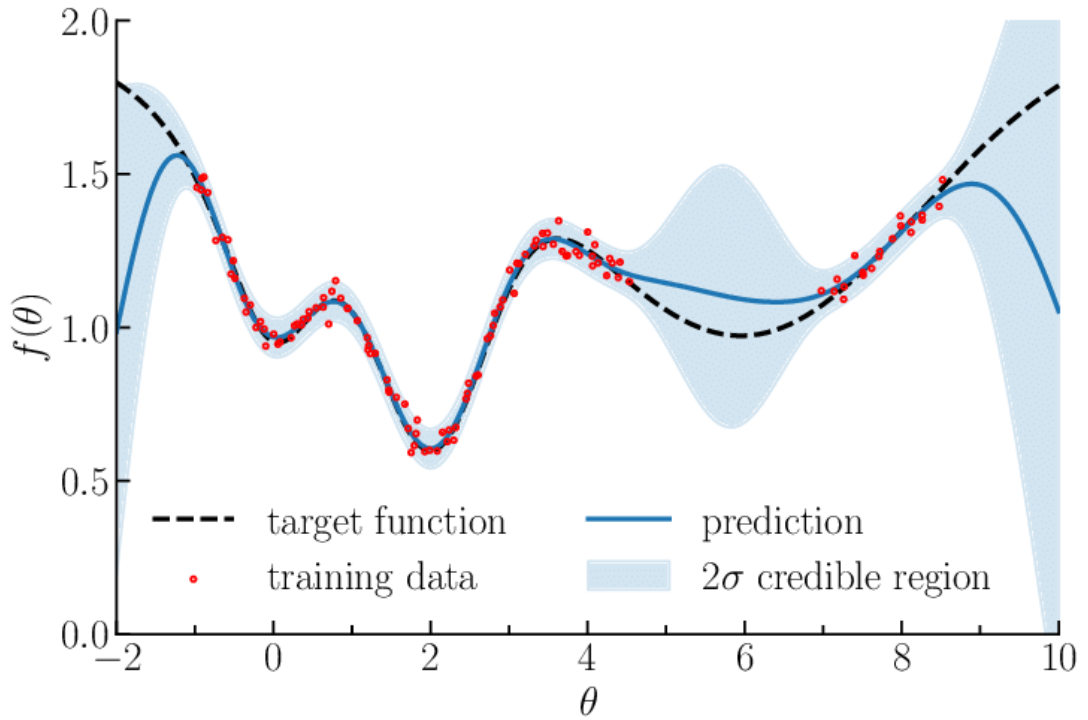
Şekil 2.4: Rasgele orman şeması [Yu vd., 2018].

Algoritma şu şekilde çalışır;

1. Veri kümesinden rastgele örnekler seçilir,
2. Daha sonra, bu algoritma her durum için bir karar ağacı oluşturur, her karar ağacından bir tahmin alır,
3. Tahmin edilen sonuçlar oylanır,
4. Son olarak en çok oylanan tahmin sonucu seçilir.

2.3. Gauss Süreci Regresyonu(Gaussian Process Regression)

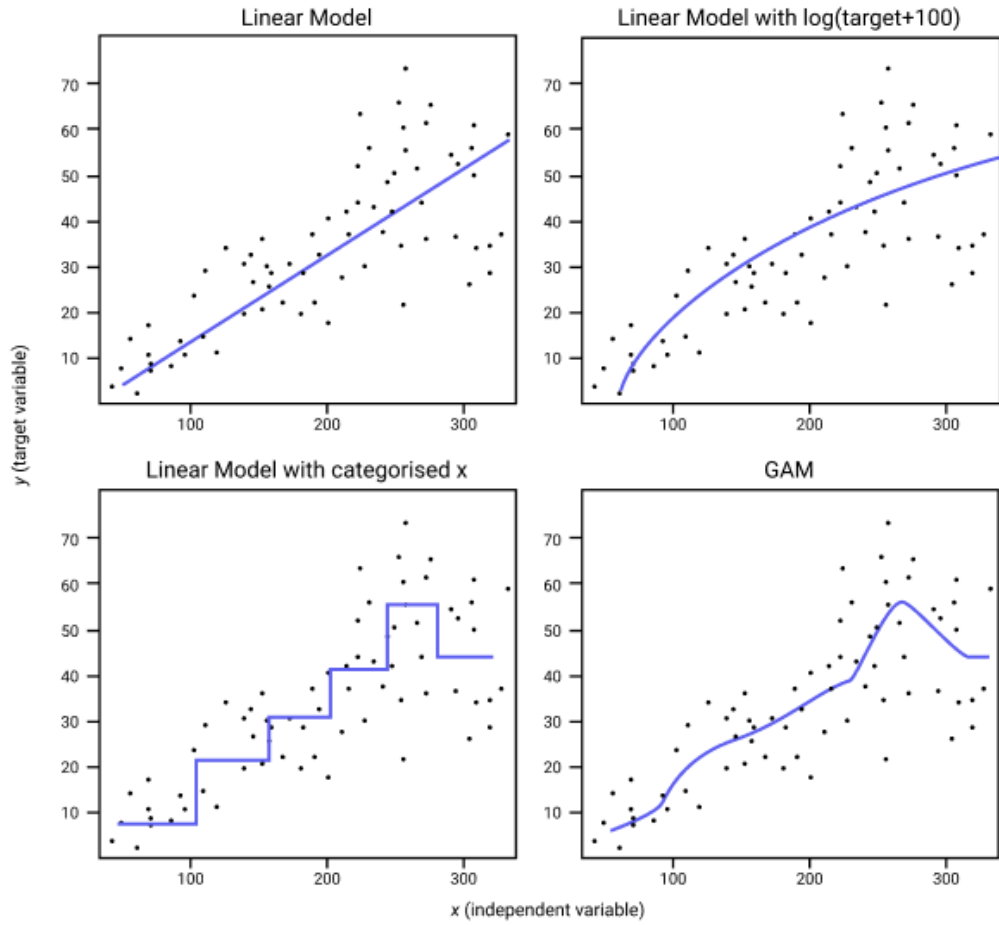
GPR, doğrusal olmayan, parametrik olmayan bir regresyon aracıdır ve yüksek boyutlu bir girdi uzayına dağılmış veri noktaları arasında enterpolasyon yapmak için kullanışlıdır. Bayes olasılık teorisine dayanır ve kernel ridge regresyon (KRR) ve radyal tabanlı fonksiyonlara sahip lineer regresyon gibi diğer regresyon teknikleriyle çok yakın bağlantıları vardır [Volker vd., 2021].



Şekil 2.5: GPR grafiği [Leclercq, 2018].

2.4. Genelleştirilmiş Katkı Modeli (Generalized Additive Model)

GAM, Doğrusal Regresyon gibi Genelleştirilmiş Doğrusal Modellerle karşılaştırıldığında daha farklı doğrusal bir modeldir. Burada metot, doğrusal olmayan özellikleri de öğrenmesi sağlanır.

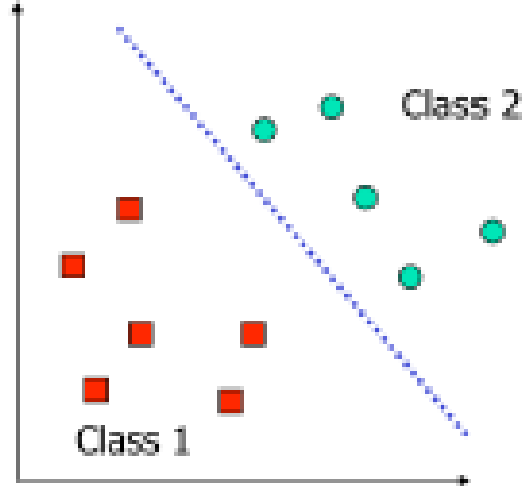


Şekil 2.6: Doğrusal bir model kullanmanın GAM'a karşı etkisi.

GAM'ler, ilişkinin basit bir ağırlıklı toplam olması gerektiği kısıtlamaz ve bunun yerine sonucun, her bir özelliğin rasgele fonksiyonlarının bir toplamı tarafından modeller [Shafi, 2021].

2.5. Destek Vektör Makinesi (Support Vector Machines)

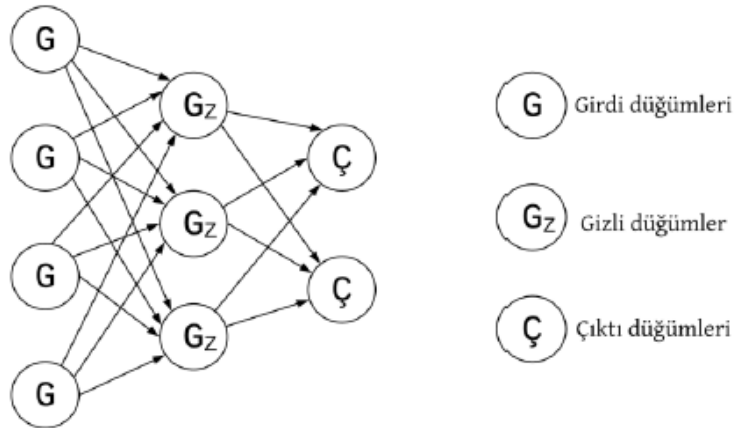
Destek Vektör Makinesi (DVM, SVM) ilk olarak Cortes ve arkadaşları tarafından 20.yy sonlarında geliştirilmiştir [Cortes ve Vapnik, 1995]. SVM de amaç, farklı örnek kategorileri arasında seçim yapmak için uzayda uygun bir sınıf sınırı (maksimum marjlı hiper düzlemler) meydana getirmektir. Boser çekirdek fonksiyonunu maksimum marjlı hiper düzlemlerde deneyerek, DVM en esnek ve etkili veri araştırma araçlarından biri haline gelmiştir [Boser vd., 1992]. Çekirdek işlevi ile, küçük bir alandan büyük alana veri eşleyebilir.



Şekil 2.7: DVM sınıflandırma şeması.

2.6. Yapay Sinir Ağları

Yapay Sinir Ağları (YSA, ANN), bir sinir hücresinin yapısından ve çalışma yönteminden esinlenilerek, karmaşık durumları tahminlemede kullanılan basit istatistiksel tekniklere alternatif bir yöntemdir. YSA’ da birbirlerine direkt bağlı nöronlar ağırlıkların belirlenmesi ve ateşleme fonksiyonu bulunmaktadır [Budak ve Erpolat, 2012]. Şekil 2.8’ de bir nöronun yapı bileşenleri gösterilmektedir.



Şekil 2.8: YSA yapısı.

3. MALZEME VE YÖNTEM

Çalışmada 58 poliol ve benzeri kimyasalın FTIR spektrumları çekilmiş, çeşitli makine öğrenmesi algoritması kullanılarak OH sayıları tahmin edilmeye çalışılmıştır.

Bazı bölgelerde farklılıklar alansal olarak bariz bir şekilde görülmektedir. **Tablo 3.1** de bazı FTIR piklerinin görülme bölgeleri verilmiştir. Bu bölgelere bakılarak piklerin nerelerde aranacağı belirlenir, bu da hesaplamalarda kolaylık sağlar.

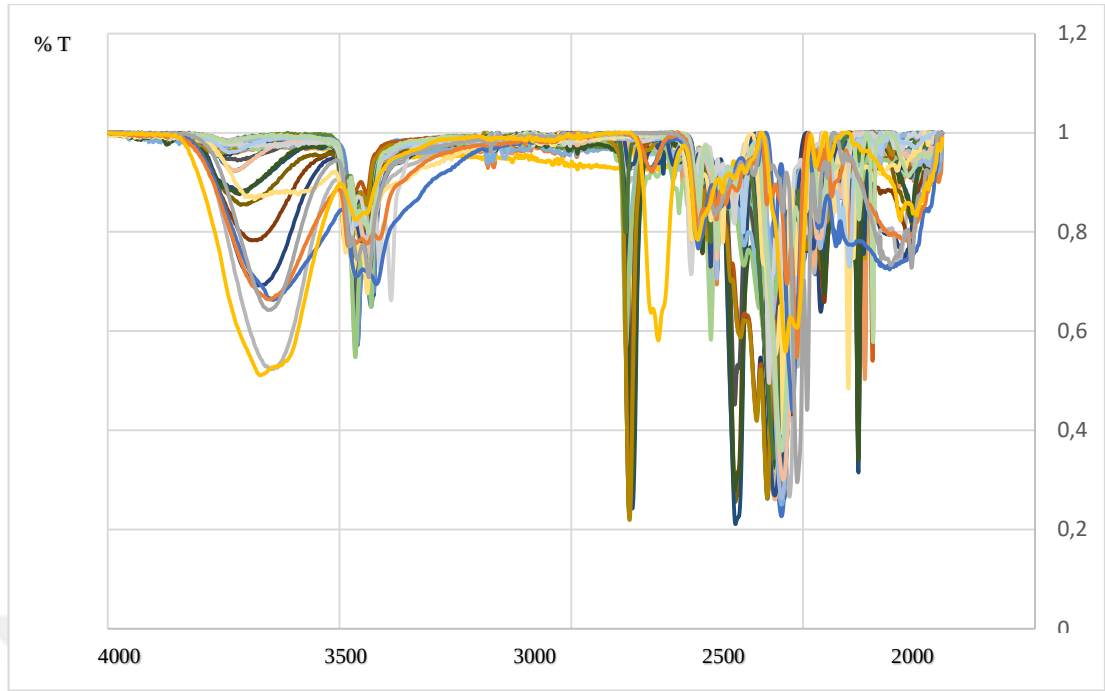
Tablo 3.1: Bazı FTIR piklerinin görülme bölgeleri [web1].

Bileşik Sınıfı	Grup	Absorpsiyon (cm ⁻¹)	Görünüş
Asit halojenür	C=O gerilme	1815-1785	Güçlü
Alkoller	O-H esneme	3700-3584	Orta, keskin
	O-H esneme	3550-3200	Güçlü, geniş
	O-H esneme	3200-2700	Zayıf, geniş
	O-H bükülme	1420-1330	Orta
Aldehit	C-H germe	2830-2695	Orta
	C=O germe	1740-1720	Güçlü
	C-H bükme	1390-1380	Orta
Karboksilik asit	O-H esneme	3300-2500	Güçlü, geniş
Ester	CO germe	1210-1163	Güçlü
İzosiyanat	N=C=O germe	2275-2250	Güçlü, geniş
Nitril	C≡N germe	2260-2222	Zayıf
Sülfat	S=O germe	1415-1380	Güçlü

FTIR yazılımı olarak OPUS, makine öğrenmesi için MATLAB kullanılmıştır. Sonuçlar, kimyasalların TDS leri ve ASTM D 1899 titrasyonuna göre doğrulanmıştır. Bu yöntem için 0,5g örnek 10ml Asetonitrilde çözülür. 10ml TSI çözeltisi eklenir. Manyetik karıştırıcıda 5dk karıştıktan sonra 0,5ml distile su ve 50ml Asetonitril eklenir ve 0,1 N TBA-OH ile titre edilir. Titrasyon sırasında 2 dönüm noktası görülecek. Aşağıdaki formül ile sonuç hesaplanır.

$$OHV [mg/g] = (EQ2 - EQ1) * 0,1 * 5,611 * F / W \quad (3.1)$$

(EQ: Dönüm noktası, F:Düzeltilme faktörü, W:Örnek ağırlığı)



Şekil 3.1: Denemelerde kullanılan FTIR spektrumları.

Yapılan denemelerde poliüretanlarda hammadde olarak kullanılan (farklı özelliklerde ve üretim sistemlerine göre ayrılan) birçok poliöl kullanılmıştır. Polioller ve kimyasallar grafikte görüldüğü gibi farklı spektrumlara sahiptir. Tabloda çalışmada kullanılan örnekler ve OH aralıkları verilmiştir.

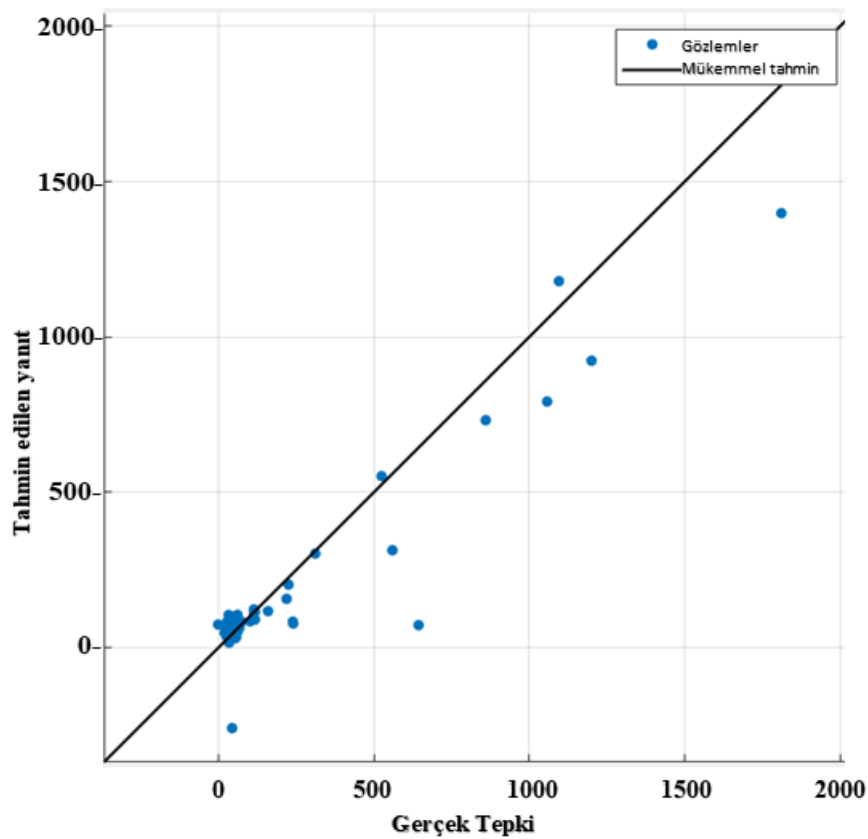
Tablo 3.2: Çalışmada kullanılan örnekler ve OH aralıkları.

Örnek Cinsi	OH Aralığı
HR Polyoller	26-29
SAN Polyoller	31-44
Hücre Açıcılar	32-34
Geri Dön. P.	35-49
Convansiyonel	47-49
P.Addition P.	42-75
GREEN P.	54-115
PET P.	160-500
VİSKO P.	250
AMİNLER	560-640
ORTEGOL204	860
DEG	1057
DIPA	1095
DP1022	1200
DEOA	1600
Gliserin	1810

Model eğitimi için, örneklerin FTIR spektrumları tablo haline getirilip tek tabloda birleştirildi. Matlab uygulamasının Regression Learner arayüzünde tüm algoritmalar çalıştırıldığında bütün tahminler ekrana düşmektedir.

3.1. Ön modeller ve Nitelik çıkarımı

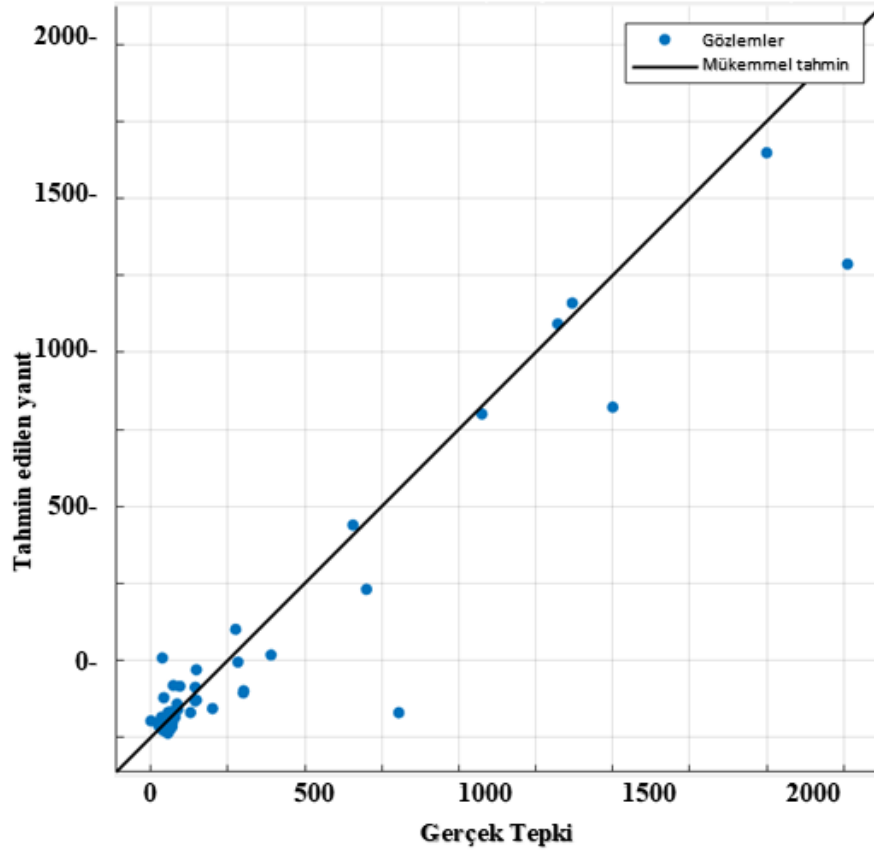
Matlab Regression learner arayüzü üzerinden pikler tanıtılmış ve en uygun metotlar seçilmiştir. Buna göre R^2 değeri en yüksek olan sonuçlar ve grafikler aşağıda listelenmiştir.



Şekil 3.2: İkinci derece SVM grafiği.

İlk deneme sonuçlarına göre RMSE: 132.52 , R^2 :0.89, MSE: 17562 , MAE: 68.729 olarak hesaplanmıştır. Burada kullanılan diğer parametreler; Tahmin hızı : ~85 obs/s, süresi ise 4.1422 saniye sürmüştür. Kernel fonksiyonu ikinci derece iken, Kernel ölçeği, kutu kısıtlaması ve epsilon otomatik olarak seçilmiştir. Son olarak verileri standartlaştırma aktif iken, optimize edici seçeneklerinde hiperparametre seçenekleri ve PCA öncesi modelde kullanılan tüm özellikler ve PCA devre dışı bırakılmıştır.

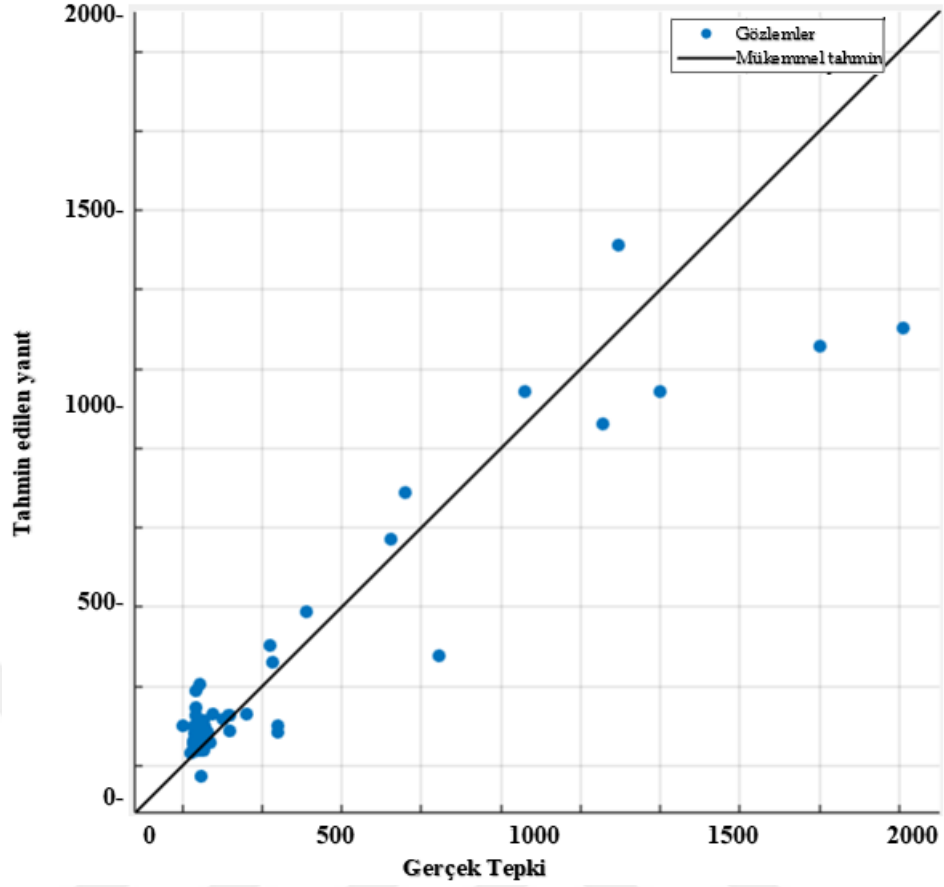
Çizgi dışında tahmin noktaları az olduğu için bu çalışmada en iyi seçeneklerden birisi ikinci derece SVM olmuştur.



Şekil 3.3: Üç katmanlı Sinir Ağı grafiği.

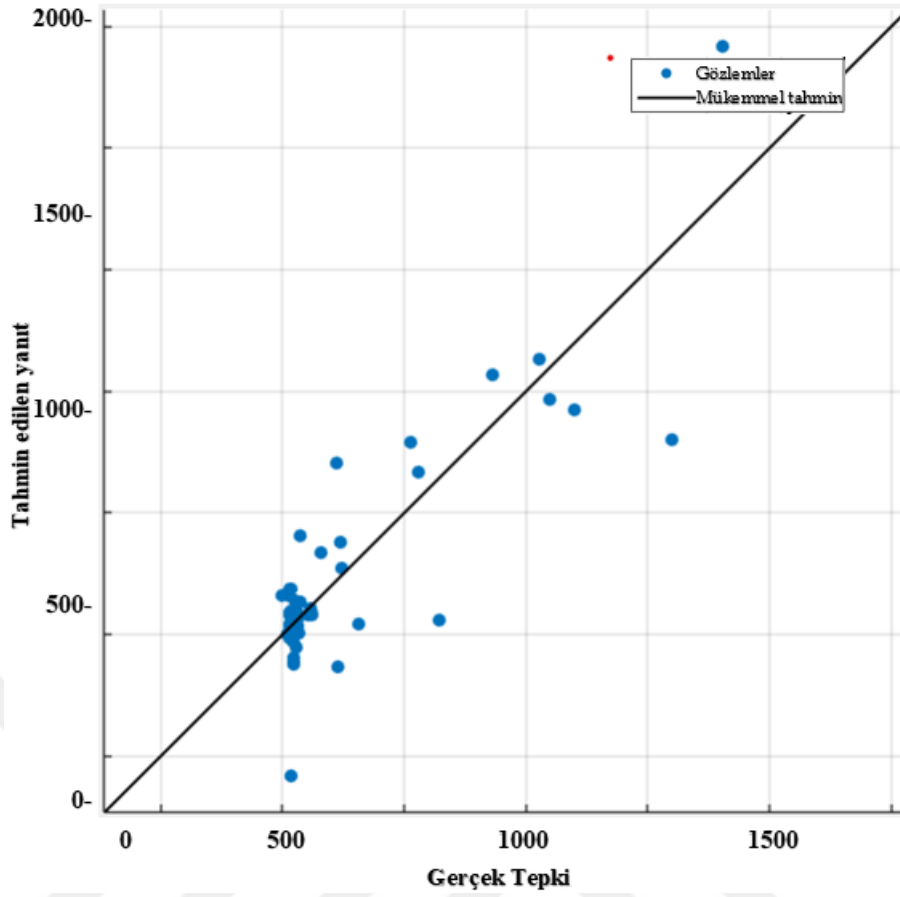
İkinci deneme olan Üç katmanlı Sinir Ağı sonuçlarında elde edilen veriler ; RMSE: 127.52 , R^2 : 0.90 , MSE: 16261 , MAE: 58.005 olarak tespit edilmiştir. Kullanılan parametrelerden tahmin hızı ~68 obs/s , eğitim süresi 19.86 saniye sürmüştür. Burada tamamen bağlı katmanların sayısı 3 iken, ilk, ikinci ve üçüncü katman boyutu 10 dur. Diğer parametreler de aktivasyon olarak ReLU, yineleme sınırı için 1000, Düzenleme gücü (Lambda) ise 0 dır. Data standartlaştırma kullanılırken, Optimize Edici Seçenekleri olarak hiperparametre seçenekleri ve öznitelik seçimi için PCA öncesi modelde kullanılan tüm özellikler ve PCA devre dışı olarak belirlenmiştir.

Çizgi dışında kalan tahmin noktaları az olduğu için bu algoritma da iyi sonuçlardan bir diğeridir.



Şekil 3.4: Üstel GPR grafiği.

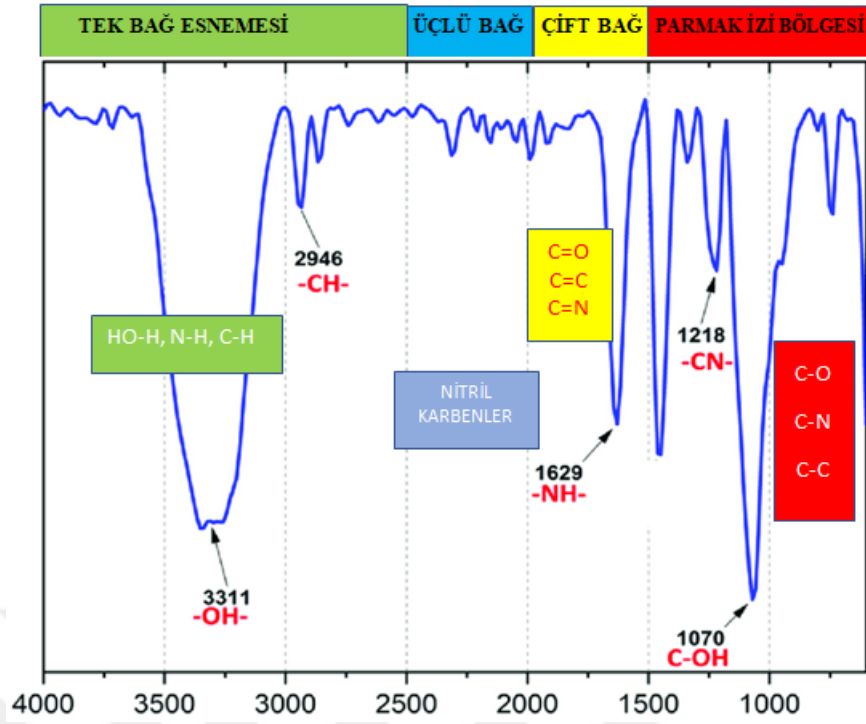
Diğer deneme Üstel GPR sonuçları ; RMSE :148.68 , R^2 : 0.86 , MSE : 22106, MAE: 78.874 Tahmin hızı: ~35 obs/s , eğitim süresi:50.167 saniye sürmüştür. Diğer parametreler ise, temel fonksiyon sürekli, kernel fonksiyonu üstel ve izotropik çekirdek kullanılırken, çekirdek ölçeği otomatik olarak belirlenmiştir. Sinyal standart sapması otomatik, sigma otomatik ve standartlaştırma kullanılmış olup, sayısal parametreler optimize edilmiştir. Son olarak Optimize Edici kullanılmzken PCA de kullanılmamıştır.



Şekil 3.5: Doğrusal (Lineer) grafiği.

Son deneme doğrusal (lineer) regresyon olarak belirlenmiştir. Sonuçlar ise RMSE: 219.54 , R^2 : 0.70 , MSE: 48199, MAE: 133.52 , Tahmin hızı: ~20 obs/s ve eğitim süresi 25.577 saniyedir. Doğrusal regresyonda başka parametre ve robust kullanılmamış ve PCA da devre dışı bırakılmıştır.

Denemelerde tahminleri oluştururken doğru sonuca en çok yaklaşan tahminleri seçmek gerekmektedir. Bunun için FTIR piklerini yorumlama ve Matlab ile çözümlemek gerekmektedir. **Şekil 3.6** da FTIR üzerinde görülen piklerin hangi bağa karşılık geldiği görülmektedir. Çalışmada, sistemi doğru sonuca götüren en uygun pikleri seçmek konusunda makine öğrenmesini kullanmıştır.



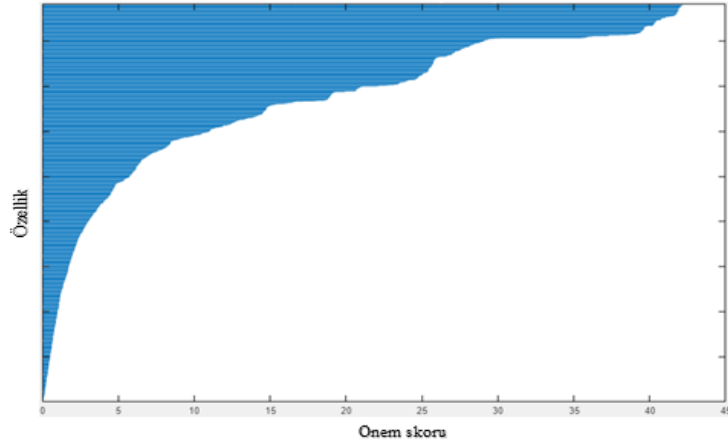
Şekil 3.6: Denemelerde kullanılan FTIR pik bölgeleri [Valencia vd.,2020].

Test sonuçlarına göre R^2 değeri en yüksek olan Trilayered Neural Network en uygun algoritma olduğu görülmüştür. Bunu da sırasıyla Rational Quadratic GPR, Quadratic SVM ve Lineer Regresyon takip etmektedir.

Denemeler esnasında Matlab programının Ftest ve MRMR özelliğinden de yararlanılmıştır. Bu özellikler en uygun pikleri seçip sisteme yorumlamaya yaramaktadır.

Tablo 3.3: MRMR testine göre en anlamlı sütunlar.

Veri sütunu	MRMR
323	0,4944
238	0,4267
1536	0,2106
1352	0,207
965	0,1643



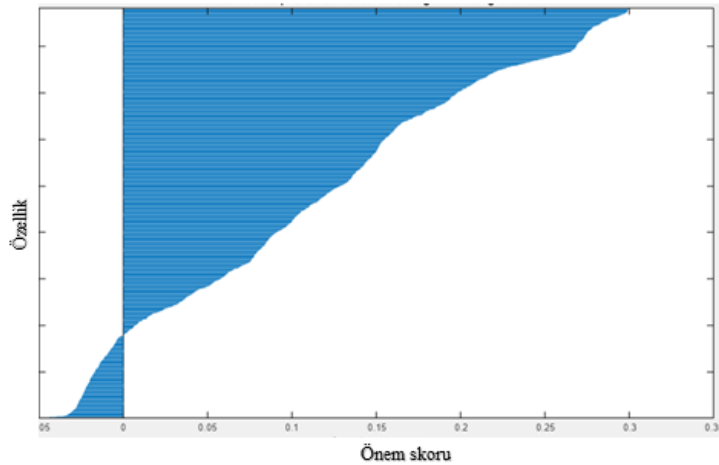
Şekil 3.7 : F Test algoritmasına göre özellik - önem puanları grafiği.

FTIR spektrumunda yaklaşık 2000 sütun vardır. Bu sütunların bazıları sonuç için anlamlıdır. F Test, sonuç için en anlamlı girdileri göstermektedir.

Tablo 3.4: F Test e göre sonuç için en anlamlı sütunlar.

Özellik	F Test
294	42,6944
295	42,693
292	42,3255
293	42,32
287	42,1791

ReliefF veya RReliefF algoritmasını kullanarak tahmin edicileri k en yakın komşuyla sıralar. Giriş matrisi X, öngörücü değişkenleri içerir ve y vektörü, bir yanıt vektörünü içerir. İşlev, en önemli tahmin edicilerin indekslerini içeren ve tahmin edicilerin ağırlıklarını içeren ağırlıkları döndürür.



Şekil 3.8: RReliefF algoritmasına göre özellik - önem puanları grafiği.

Y sayısalsa, kabartma varsayılan olarak regresyon için RReliefF analizini gerçekleştirir. Aksi takdirde, kabartma, sınıf başına k en yakın komşuyu kullanarak sınıflandırma için ReliefF analizini gerçekleştirir.

Tablo 3.5: RReliefF e göre sonuç için en anlamlı sütunlar.

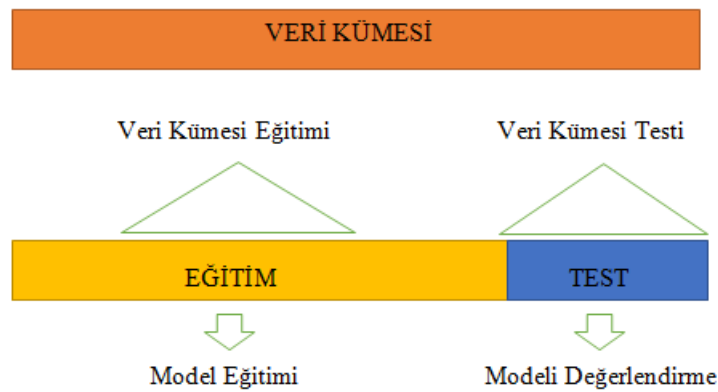
Özellik	RReliefF
349	0,3001
350	0,3
348	0,2995
351	0,2992
358	0,2991

Tablo 3.6:Yapılan benzetimler sonucu elde edilen sonuçlar.

	RMSE	MSE	R ²	MAE
Linear	219.54	48199	0.70	133.52
Exponential GPR	148.68	22106	0.86	78.874
Quadratic SVM	132.52	17562	0.89	68.729
Trilayered Neural Network	127.52	16261	0.90	58.005

3.2. Model performansları

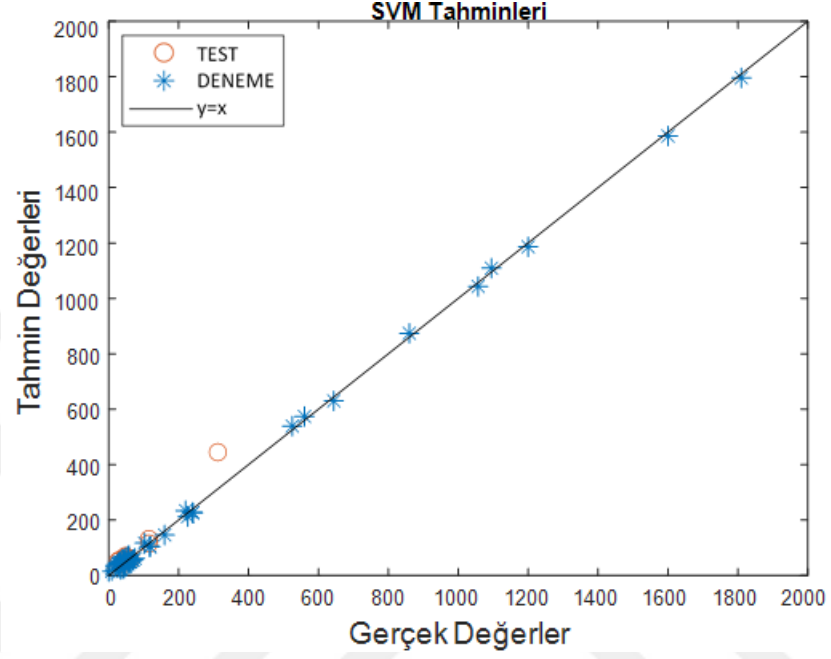
Modeller değerlendirilirken sonuçların uygunluğu tekrar bir deneme ile doğrulanmalıdır. Şekil 3.9 da değerlendirmelerin önem ölçü oranları verilmiştir.



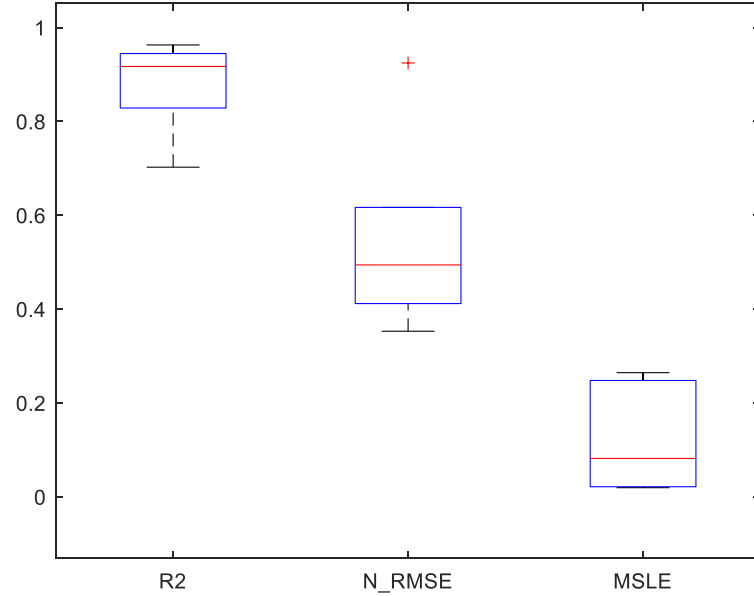
Şekil 3.9: Veri Kümesi Eğitim ve değerlendirme.

Deneme sonuçlarını, verilere göre test ettiğimizde aldığımız sonuçlar ise şu şekildedir:

SVM'e göre RMSE: 120.21 , R^2 : 0.91 , MSE: 14450 ve MAE: 63.952 olduğu görülmüştür. Bu denemenin hızı 120 osb/saniye iken deneme 3.0584 saniye sürmüştür. Denemede kullanılan parametreler; çekirdek fonksiyonu doğrusal, çekirdek ölçeği, kutu kısıtlaması ve epsilon otomatik olarak belirlenmiş, data standartlaştırma kullanılmıştır.



Şekil 3.10: SVM tahminleri ve gerçek değerler.



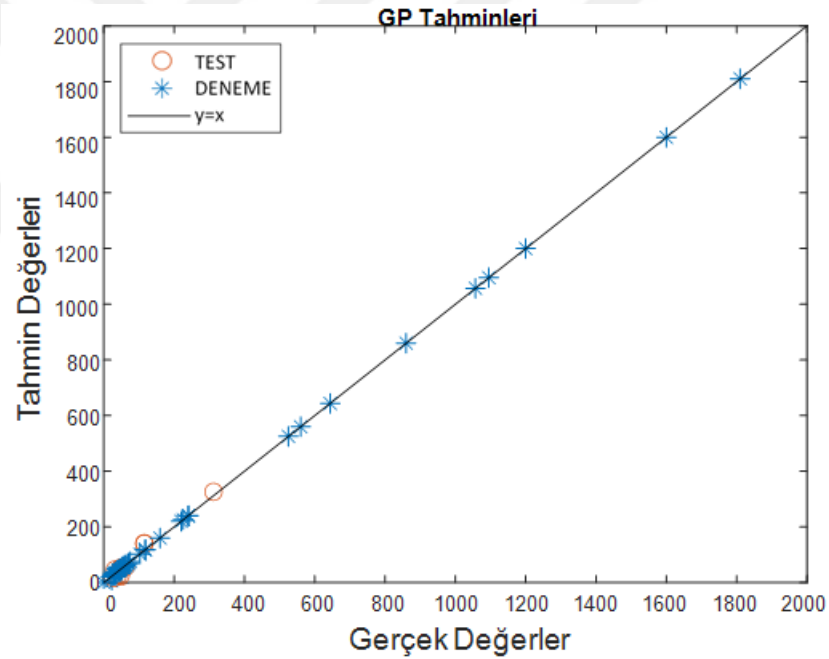
Şekil 3.11: SVM parametreleri

Şekil 3.11 de SVM algoritması parametreleri verilmiştir. Tablo 3.6 da ise bu grafiğe ait değerler verilmiştir.

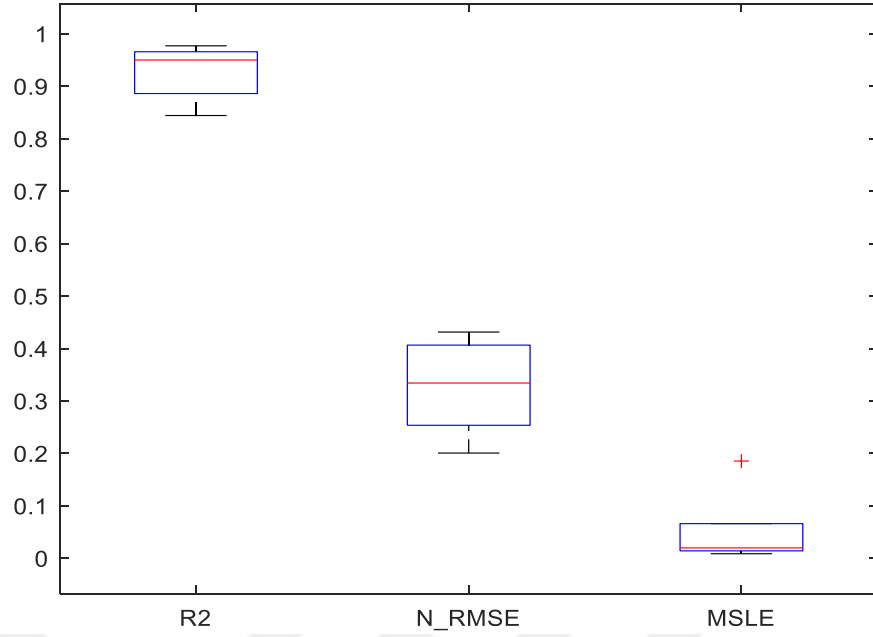
Tablo 3.7: SVM parametreleri.

	Üst Sınır	Alt Sınır	Ortalama
R^2	0,963	0,702	0,917
RMSE	0,616	0,352	0,494
MSLE	0,264	0,019	0,081

Diğer deneme olan GPR sonucuna göre RMSE: 120.97 , R^2 : 0.91 , MSE: 14634 ve MAE: 60.751 olduğu görülmüştür. Bu denemenin hızı 120 osb/saniye iken deneme 3.2723 saniye sürmüştür. Denemede kullanılan parametreler; ikinci derece GPR, temel fonksiyon devamlıdır. Çekirdek fonksiyonu rasyonel ikinci derece ve izotropik çekirdek kullanılmıştır. Çekirdek ölçeği, sinyal standart sapması ve sigma otomatik olarak belirlenmiş, data standartlaştırma kullanılmış ve sayısal parametreyi optimize edilmiştir. **Şekil 3.12** de Tahmin-Gerçek değerler grafiği verilmiştir.



Şekil 3.12: GP tahminleri ve gerçek değerler.



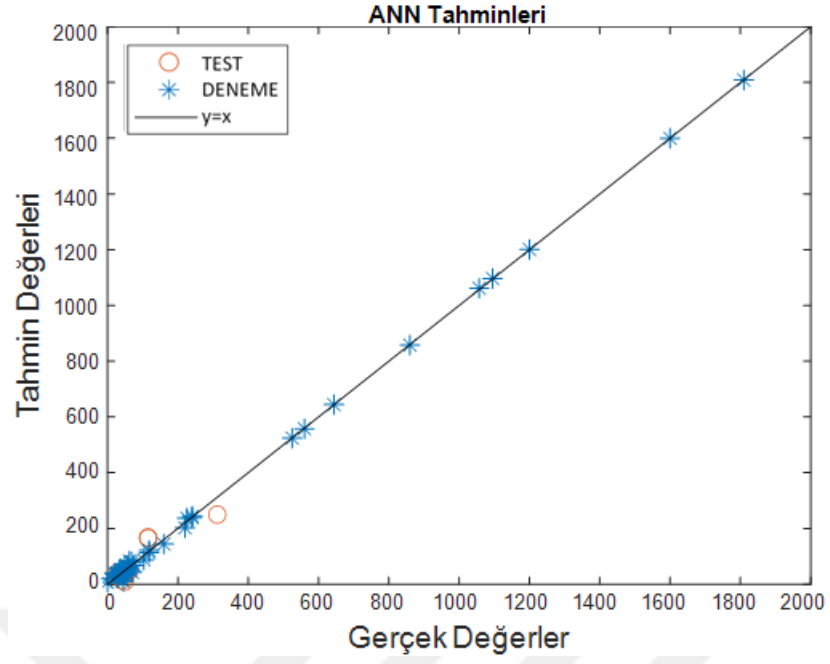
Şekil 3.13: GP parametreleri ve aralıkları.

Gauss Süreç Regresyonunda da tahmin performansı oldukça iyi olduğu görülmüştür. Bu denemeye ait grafik **Şekil 3.13** de sayısal sonuçlar **Tablo 3.7** de verilmiştir.

Tablo 3.8: GP parametreleri ve aralıkları.

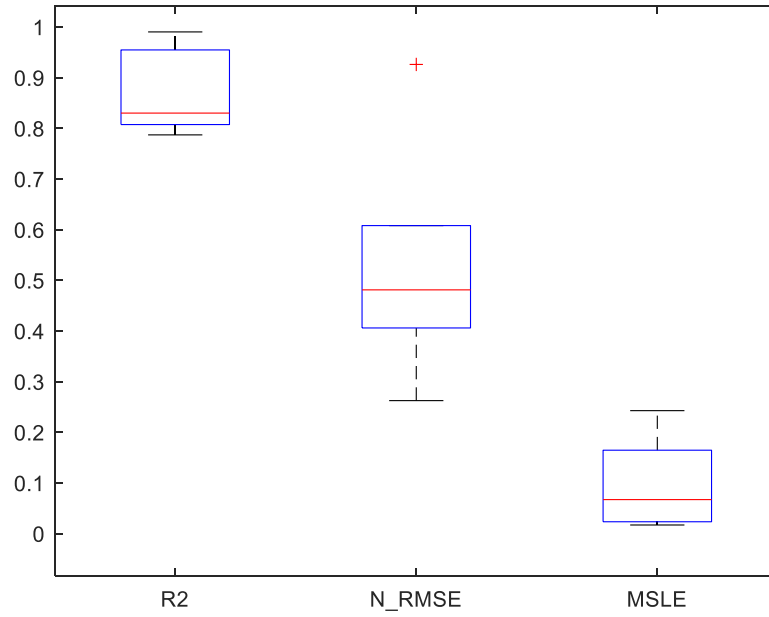
	Üst Sınır	Alt Sınır	Ortalama
R ²	0.977	0.844	0.950
RMSE	0.431	0.200	0.334
MSLE	0.185	0.008	0.019

Başka bir deneme olan ANN sonucuna göre RMSE: 121.17 , R²: 0.91 , MSE: 14683 ve MAE: 58.208 olduğu görülmüştür. Bu denemenin hızı 130 osb/saniye iken deneme 4.9467 saniye sürmüştür. Denemede kullanılan parametreler; dar sinir ağı, 1 tamamen bağlı katman kullanılmış ve ilk katman boyutu 10 dur. ReLU etkinleştirmesi seçilmiş ve iterasyon limiti 1000 olarak belirlenmiştir. Düzenleme gücü (Lambda) 0 ve data standartlaştırma kullanılmıştır.



Şekil 3.14: ANN tahminleri ve gerçek değerler.

Şekil 3.15 de ANN e göre doğruluk parametreleri , ANN Parametreleri sayısal olarak da **Tablo 3.8** de verilmiştir.

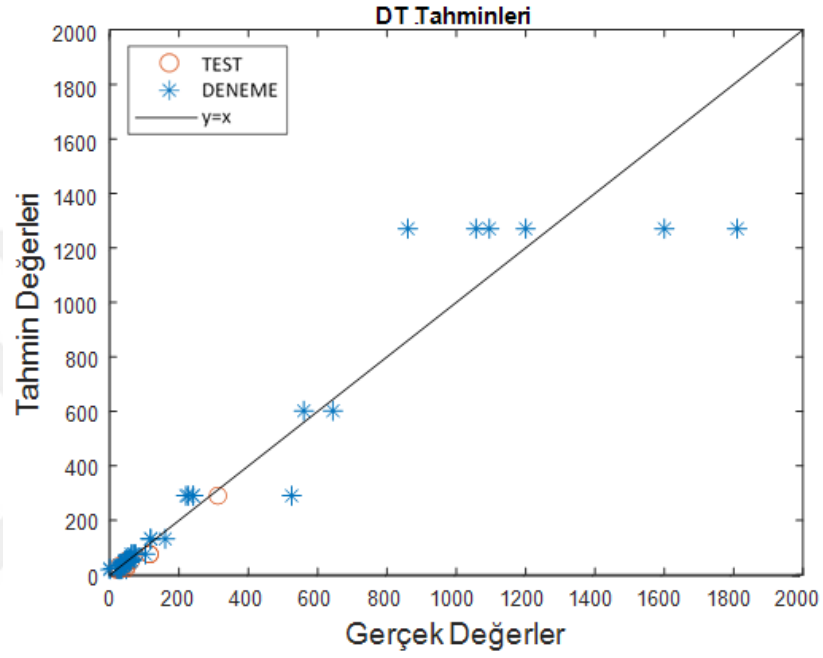


Şekil 3.15: ANN parametreleri ve aralıkları.

Tablo 3.9: ANN parametreleri ve aralıkları.

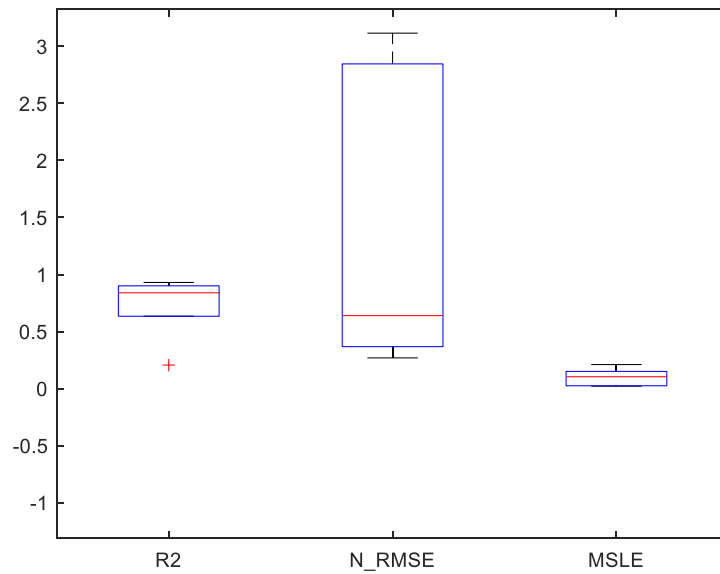
	Üst Sınır	Alt Sınır	Ortalama
R^2	0,99	0,787	0,83
RMSE	0,926	0,262	0,481
MSLE	0,242	0,017	0,067

Son deneme Karar ağacı algoritması. Bu algoritmaya ait tahmin grafiği **Şekil 3.16** da verilmiştir.



Şekil 3.16: DT tahminleri ve gerçek değerler.

Şekil 3.17 de bulunan sonuçlar **Tablo 3.10** da sayısal olarak verilmiştir.



Şekil 3.17: DT parametreleri.

Tablo 3.10: DT parametreleri.

	Üst Sınır	Alt Sınır	Ortalama
R^2	0,93	0,205	0,839
RMSE	3,116	0,269	0,639
MSLE	0,21	0,02	0,103

DT algoritması diğer algoritmalara göre yüksek bir başarı sağlayamamıştır. Tüm deneme sonuçları **Tablo 3.11** de verilmiştir.

Tablo 3.11: Denemelere ait ortalama değerler.

	DT	ANN	GP	SVM
R^2	0,839	0,83	0,95	0,917
RMSE	0,639	0,481	0,334	0,494
MSLE	0,103	0,067	0,019	0,081

4. SONUÇLAR VE YORUMLAR

Matlab ile yapılan regresyon sonuçları grafikleri ve sayısal sonuçlar neticesinde yukarıdaki sonuçlar elde edilmiştir. Denemeler gösteriyor ki, FTIR spektrumları ile Matlab birlikte kullanılarak kimyasalların bazı özellikleri tespit edilebilir. Denemeler sırasında en iyi modeller kernel kullanan SVM ve GP çıkmıştır. Buradan OH sayısı ve pikler arasında ilişkinin nonliner olduğu görülmektedir. Gauss süreci birçok kolay hesaplanabilir fonksiyonun birleşiminden oluşmaktadır. Böylece çıktılar arasında yararlı bilgileri yakalayabildiğinden, çoklu korelasyonlu çıktılarla başa çıkmayı sağlamaktadır [Lin vd., 2021].

Yapay zekâ algoritmalarının birçoğu bu tür sınıflandırma problemlerini çözebilmektedir. Fakat işlemler matematiksel denklemlere benzetildiğinde, denklemin kaçınıcı derece olduğu (zorluk derecesi veya çoklu değişkenler) ile bağlantılı olarak, sistem için en uygun tahmini oluşturan algoritmayı seçmek gerekmektedir.

Literatürdeki benzer çalışmalar incelendiğinde, benzer bir çalışmada yedi polipropilen glikol standardı FT-NIR spektrometresine kaydedilerek ve PLS algoritmasını kullanarak $R^2=0.99$ bulunmuştur [Ian R., 2017]. Bonnie'nin çalışmasında HATR/FTIR ve OH sayıları 59-229 arası 9 bitkisel poliöl ile dataseti oluşturmuş ve PLS algoritmasıyla $R^2=0.9925$ olarak bulunmuştur [Bonnie, 2017]. Diğer bir çalışmada ise 29-330 OH aralığında 7 poliesterin FT-IR/MIR spektrumlarını kullanarak regresyon yapılmış ve $R^2=0.977$ olarak bulunmuştur [Jesus vd., 2019]. Söz konusu çalışmaların hespsinde veri, eğitim ve test olarak ayrılmamıştır, bu nedenle herhangi bir validasyon söz konusu değildir. Bu çalışma kapsamında test verilerinin uyumları paylaşılmaktadır, kullanılan makine öğrenmelerinin eğitim verilerine uyumu yüksektir ($R^2>0.99$).

Bu çalışmada OH sayısı hesaplanırken, sonraki çalışmalarda, bir çoğu analiz sırasında uzun vakit harcamaya sebep olan asit tayini, ph tayini, izosiyanatlarda NCO % si, yağlarda iyot sayısı gibi yaş analizlerin hesaplanmasında kullanılabilir. Ayrıca bir çok özellik de 1 dakika gibi kısa bir sürede FTIR spektrumu çekilip, Matlaba tanıtılıp sayısal sonuçlar için formüller oluşturulabilir ve nihai tahmin tespit edilebilir.

KAYNAKLAR

Alpaydin E., (2014), “Introduction to machine learning”, MIT press.

Bayer O., (1947), “Das Di-Isocyanat-Polyadditions verfahren (Polyurethane)”, Angewandte Chemie, 59(9), 257-288.

Baysal G., Kasapbası E., (2017), “Polyurethanes and Usage Areas”, Global Journal of Science Frontier Research: B Chemistry, 17(1), 5-7.

Bonnie T. Y. P., Zulina A. M., Haliza A. A., (2017), “Measurement Of Hydroxyl Value Of Palm-Based Polyol By Attenuated Total Reflectance/Fourier Transform Infrared Spectroscopy”, Journal of Oil Palm Research, 27 (2), 190-193.

Boser B. E., Isabelle M. G., Vladimir N. V., (1992), “A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learning theory”, Association for Computing Machinery, New York, NY, USA, 144–152.

Breiman L., (2001), “Random forests”, Machine Learning , 45, 5–32.

Budak H., Erpolat S., (2012), “Kredi riski tahmininde yapay sinir ağıları ve lojistik regresyon analizi karşılaştırılması”, Academic Journal of Information Technology, 3(9), 23-30.

Chen Z., Zheng N., Luan X., Liu F., (2019), “O-cresol Concentration Online Measurement Based On Near Infrared Spectroscopy Via Partial Least Square Regression”, J. Vis. Exp, 8(153).

Choi J. H., Chae S. U., Hwang E. H., Choi D. M., (2022), “Fire Propagation Characteristics and Fire Risks of Polyurethanes: Effects of Material Type (Foam & Board) and Added Flame Retardant”, Fire, 5(4), 105.

Coşkun M. Y., İtik M., (2023), “Elektrohidrolik bir sistemin pekiştirmeli öğrenme tabanlı denetleyici ile konum denetiminin gerçekleştirilmesi”, NÖHÜ Müh. Bilim. Dergisi, 12(1), 280-288.

Cortes C., Vapnik V., (1995), “Support-vector networks”, Machine Learning, 20, 273–297.

Cunningham P., Cord M., Delany S.J., (2008), “Supervised Learning. In: Cord, M., Cunningham, P. (eds) Machine Learning Techniques for Multimedia”, Cognitive Technologies, Springer, Berlin, Heidelberg, 21-22.

Çetiner K. A., Zubari Ü., Berberoğlu G., Yavuz E., Yılmaz Ş., Mert A., (2022), “Yapay Zekâ Uygulamaları”, Aselsan, 114, 152-163.

Çalikoğlu Y., (2017), “Epoksitlendirilmiş Soya Yağından Tioglikolik Asit ve Tioglikolik Asit Metil Esteri ile Halka Açarak Oluşan Poliollerden Yeni Poliüretan Sentezi ve Karakterizasyonu”, Yüksek Lisans Tezi, Sakarya Üniversitesi.

Eroğlu A. R., (2008), “Yüzey Koruyucu Kaplamalarda Akrilik-Poliüretan Reçnelerin Kullanılması”, Yüksek Lisans Tezi, İstanbul Üniversitesi.

Ersoz F., Kapanşahin G., (2021), “Estimating the Cost of Crude Steel Production: A Machine Learning Approach”, 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Türkiye, 1-8.

Gümüş E., Medetoğlu B., Tutar S., (2020), “Finans ve Bankacılık Sisteminde Yapay Zekâ Kullanımı: Kullanıcılar Üzerine Bir Uygulama”, Bucak İşletme Fakültesi Dergisi, 3(1), 28-53.

Hastie T., Tibshirani R., Friedman J., (2009), "Unsupervised Learning", The Elements of Statistical Learning, Springer, 485–586.

Ian R., (2017), “The Determination of Hydroxyl Number in Polyols Using the Spectrum Two N FT-NIR Spectrometer”, PerkinElmer Inc., Seer Green, UK.

İnce H., İmamoğlu S. E., İmamoğlu, S. Z., (2021), “Yapay zeka uygulamalarının karar verme üzerine etkileri: Kavramsal bir çalışma”, International Review of Economics and Management, 9(1), 50-63.

Jesus L. L., Murakami L. M. S., Mello T. D. S. D., Diniz M. F., Silva L. M., Mattos E. D. C., Dutra R. D. C. L., (2019), “Evaluation of techniques for determination of hydroxyl value in materials for different industrial applications”, Journal of Aerospace Technology and Management, 11(1).

Kandar R., (2021), “Regresyon (Regression) - Sınıflandırma(Classification) Nedir? ” <https://rafetcankandar.medium.com/regresyon-regression-s%C4%B1n%C4%B1flandirma-classification-nedir-c2757850ae76>. (Erişim Tarihi:31.07.2023)

Kapdan S., (2022), “Kimyasal Proseslerde Yapay Zeka Kullanımı”, <https://inovatifkimyadergisi.com/kimyasal-proseslerde-yapay-zeka-kullanimi>. (Erişim Tarihi:30.01.2024)

Königstorfer F., Thalmann S., (2020), “Applications of Artificial Intelligence in commercial banks–A research agenda for behavioral finance”, Journal of Behavioral and Experimental Finance, 27.

Kulkarni V. Y., Sinha P. K., Petare M. C., (2016), “Weighted hybrid decision tree model for random forest classifier”, J. Inst. Eng., 97, 209–217.

Laís L. J., Lídia M. S. M., Talita de S. D. M., Milton F. D., Leandro M. S., Elizabeth da C. M., Rita de C. L. D., (2019), “Evaluation of Techniques for Determination of Hydroxyl Value in Materials for Different Industrial Applications”, J. Aerosp. Technol. Manag., São José dos Campos, 11.

Leclercq F., (2018), “Bayesian optimization for likelihood-free cosmological inference”, Physical Review, 98(6).

Lin Q., Hu, J., Zhou Q., Cheng Y., Hu Z., Couckuyt I., Dhaene T., (2021), “Multi-output Gaussian process prediction for computationally expensive problems with multiple levels of fidelity”, Knowledge-Based Systems, 227.

Mitchell T., (1997), “Machine learning”, McGraw Hill, New York.

Sargent D. J., (2001), “Comparison of artificial neural networks with other statistical approaches results from medical data sets”, Cancer, 91(8), 1636-1642.

Sarıman G., (2011), “Veri madenciliğinde kümeleme teknikleri üzerine bir çalışma: K-Means ve K-Medoids kümeleme algoritmalarının karşılaştırılması”, Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 15(3), 192-202.

Shafi A., (2021), “What is a Generalised Additive Model?”, Towards Data Science, <https://towardsdatascience.com/generalised-additive-models-6dfbedf1350a> (Erişim Tarihi: 31/07/2023).

Soucy P., Mineau G.W., (2001), “A simple KNN algorithm for text categorization, in: Data Mining”, ICDM, 647-648.

Şengür D., (2013), “Öğrencilerin akademik başarılarının veri madenciliği metotları ile tahmini”, Fırat Üniversitesi, Eğitim Bilimleri Enstitüsü, Bilgisayar ve Öğretim Teknolojileri Eğitimi, Yüksek Lisans Tezi.

Turing A. M., (1950), “Computing machinery and intelligence” Mind, New Series, 59(236), 433-460.

Valencia C., Valencia Y., Grande Tovar C.D., (2020), “Synthesis and Application of a Cationic Polyamine as Yankee Dryer Coating Agent for the Tissue Paper-Making Process”, Polymers, 12(1), 173.

Volker L. D., Albert P. B., Noam B., David M. W., Michele C., Gábor C., (2021), “Gaussian Process Regression for Materials and Molecules”, Chemical Reviews, 121 (16), 10073-10141.

Web1, (2023), <https://www.sigmaaldrich.com/TR/en/technical-documents/technical-article/analytical-chemistry/photometry-and-reflectometry/ir-spectrum-table> (Erişim Tarihi: 09/10/2023).

Yargıç A. Ş., Yarbay Ş. R. Z., Özbay N., (2020), “Biyo-poliol-esaslı karbon köpüğün yapısal özellikleri üzerinde çözücü türü etkisinin incelenmesi”, Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 36(1), 133-146.

Yenen E., (2020), “Sert Poliüretan Köpük Uygulamalarında Kullanılmak Üzere Asit İndisi Düşük Poliester poliollerin Geliştirilmesi”, Yüksek Lisans Tezi, İstanbul Üniversitesi.

Yılğör İ., Yılğör E., (2020), “Polyurethanes: Design, Synthesis and Structure-Property Behavior of Versatile Materials”, Hacettepe Journal of Biology and Chemistry, 48 (5), 425-445.

Yu F. W., Ho W. T., Ho C. K. T., Sit C. R. K. Y., (2018), “Logistic regression-based optimal control for aircooled Chiller”, *International journal of refrigeration*, 85, 200–212.

Zachary J. Baum, Xiang Yu, Philippe Y. Ayala, Yanan Zhao, Steven P. Watkins, and Qiongqiong Zhou, (2021), “Artificial Intelligence in Chemistry: Current Trends and Future Directions”, *Journal of Chemical Information and Modeling*, 61 (7), 3197–3212.

Zhang S., (2019), “Cost-Sensitive KNN Classification”, *Neurocomputing*, 391, 234–242.



ÖZGEÇMİŞ

Yusuf Serdar BEKDAŞ, 2004 yılında Sarıyer Vehbi Koç Vakfı (Y.D.A) Lisesinden mezun oldu. 2008 yılında Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik ve Mimarlık Fakültesinde başladığı Tekstil Mühendisliğini 2010 yılında sektörel krizden dolayı terk edip aynı yıl başladığı Konya Selçuk Üniversitesi Mühendislik ve Mimarlık Fakültesi Kimya Mühendisliğini 3 yılda tamamlayarak 2013 yılında lisans öğrenimini tamamlamıştır. 1 yıl boya ve tekstil, 8 yıl poliüretan sektöründe mesleki tecrübe edinmiştir. 2021 yılında başladığı Gebze Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Kimya Mühendisliği Anabilim dalından 2024 yılında mezun olarak yüksek lisans öğrenimini tamamlamıştır.



TEZ ÇALIŞMASI KAPSAMINDA YAPILAN YAYINLAR

Bekdaş Y. S., (2023), “Fourier Dönüşümlü Kızılötesi Spektroskopisi ve Makine Öğrenmesi Kullanılarak Poliollerde Bazı Karakteristik Özelliklerin Öngörülmesi”, Avrasya 9. Uluslararası Uygulamalı Bilimler Kongresi, 342-343, Tiflis, Gürcistan, 24 – 26 Kasım.

