

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

BÜYÜK BOYUTLU VERİ SINIFLANDIRMADA DOĞRUSAL BOYUT
İNDİRGEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Elektrik- Elektronik Müh. Eray YILDIZ

OCAK 2018
TRABZON



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce

Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : / /

Tezin Savunma Tarihi : / /

Tez Danışmanı :

Trabzon

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

başlıklı bu çalışma, Enstitü Yönetim Kurulunun / / gün ve sayılı
kararıyla oluşturulan jüri tarafından yapılan sınavda
YÜKSEK LİSANS TEZİ
olarak kabul edilmiştir.

Jüri Üyeleri

Başkan :

Üye :

Üye :

Prof. Dr. Sadettin KORKMAZ

Enstitü Müdürü

ÖNSÖZ

Bu tez, Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı, Elektronik Mühendisliği Yüksek Lisans Programı'nda yapılan bir çalışmadır.

Öncelikle yüksek lisans tezi danışmanlığımı üstlenerek, gerek konu seçiminde ve gerekse çalışmaların yürütülmesinde yardımlarını esirgemeyen sayın hocam Yrd. Doç. Dr. Yusuf SEVİM'e en içten teşekkürlerimi sunuyorum.

Ayrıca bu yaşa gelmemde büyük emekleri geçen sevgilerini ve güvenlerini hiç eksik etmeyen, her an yanımda olan aileme çok teşekkür ederim.

Eray YILDIZ

Trabzon 2018

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduđum “Büyük Boyutlu Veri Sınıflandırmada Doğrusal Boyut İndirgeme Yöntemlerinin Karşılaştırılması” başlıklı bu çalışmayı baştan sona kadar danışmanım Yrd. Doç. Dr. Yusuf SEVİM’in sorumluluğunda tamamladıđımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiđimi beyan ederim. 05/01/2018

Eray YILDIZ

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	VII
SUMMARY	VIII
ŞEKİLLER DİZİNİ	IX
TABLolar DİZİNİ.....	X
SEMBOLLER DİZİNİ	XII
1. GENEL BİLGİLER	1
1.1. Giriş.....	1
1.2. Boyut İndirgeme	2
1.3. Doğrusal Boyut İndirgeme Yöntemleri.....	3
1.3.1. Temel Bileşen Analizi.....	4
1.3.2. Doğrusal Diskriminant Analizi	9
1.3.3. Yerellik Koruyan İzdüşüm.....	13
1.3.4. Yerellik Duyarlı Ayırma Analizi	16
1.3.5. Komşuluk Koruyan Gömme	21
1.3.6. İzometrik İzdüşüm	24
1.4. Sınıflandırma Yöntemleri	29
1.4.1. k -En Yakın Komşuluk Yöntemi.....	30
1.4.2. Yapay Sinir Ağları	30
1.4.2.1. Çok Katmanlı İleri Beslemeli Yapay Sinir Ağları	31
1.4.2.2. Geri Yayılım Algoritması	32
1.4.3. Destek Vektör Makinaları.....	33
1.4.3.1. İki Sınıflı Doğrusal Ayrılabilir Durum	34
1.4.3.2. İki Sınıflı Doğrusal Ayrılmayan Durum	36
1.4.3.3. Çok Sınıflı DVM.....	37
1.5. Karşılaştırma Ölçütleri.....	37
1.5.1. Doğruluk Değerlendirmesi.....	37

1.5.2.	Sınıflandırma Süreleri	38
2.	YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME	39
2.1.	Giriş.....	39
2.2.	ORL Veri Kümesi	39
2.3.	USPS Veri Kümesi.....	39
2.3.1.	Öznelikler.....	42
2.3.1.1.	Gradyan Öznelikler.....	42
2.3.1.2.	Diyagonal Öznelikler	43
2.4.	ORL Veri Kümesine Ait Bulgular	44
2.4.1.	Orijinal Öznelikler Üzerinde Sınıflandırma	44
2.4.2.	Boyut İndirgeme	45
2.4.3.	Boyut İndirgeme İşlemi Ardından Sınıflandırma	46
2.5.	USPS Veri Kümesine Ait Bulgular.....	49
2.5.1.	Gradyan Özneliklere Ait Bulgular	49
2.5.1.1.	Orijinal Öznelikler Üzerinde Sınıflandırma	49
2.5.1.2.	Boyut İndirgeme	50
2.5.1.3.	Boyut İndirgeme İşlemi Ardından Sınıflandırma	51
2.5.2.	Diyagonal Özneliklere Ait Bulgular	54
2.5.2.1.	Orijinal Öznelikler Üzerinde Sınıflandırma	54
2.5.2.2.	Boyut İndirgeme	55
2.5.2.3.	Boyut İndirgeme İşlemi Ardından Sınıflandırma	55
2.5.3.	Gradyan ve Diyagonal Özneliklere Ait Bulgular	59
2.5.3.1.	Orijinal Öznelikler Üzerinde Sınıflandırma	59
2.5.3.2.	Boyut İndirgeme	60
2.5.3.3.	Boyut İndirgeme İşlemi Ardından Sınıflandırma	60
3.	SONUÇLAR VE ÖNERİLER.....	65
4.	KAYNAKLAR	67
ÖZGEÇMİŞ		

Yüksek Lisans Tezi

ÖZET

BÜYÜK BOYUTLU VERİ SINIFLANDIRMADA DOĞRUSAL BOYUT İNDİRGEME
YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Eray YILDIZ

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı
Danışman: Yrd. Doç. Dr. Yusuf SEVİM
2018, 70 Sayfa

Yüksek boyutlu verinin analiz edilmesi ile birçok alanda karşılaşılır. Yüksek boyutlu veri analizinde, işlem yükü ve zaman açısından verinin etkin ve hızlı bir şekilde işlenmesi zor olabilir. Bu zorluğun üstesinden gelebilmek için boyut indirgeme yöntemleri yaygın bir şekilde kullanılmaktadır. Doğrusal boyut indirgeme yöntemleri, işlem yükü ve zaman açısından doğrusal olmayan yöntemlere göre daha iyi sonuç verirler. Bu yüzden yüksek boyutu veri analizinde daha çok tercih edilmektedirler.

Bu tez çalışmasında yüz ve rakam veri kümeleri kullanılarak çeşitli sınıflandırma yöntemleri üzerinde popüler doğrusal boyut indirgeme yöntemleri ve performansları incelenmiştir. Bu yöntemler, Temel Bileşen Analizi (TBA), Doğrusal Ayırma Analizi (DAA), Yerellik Koruyan İzdüşüm (YKİ), Komşuluk Koruyan Gömme (KKG), Yerellik Duyarlı Ayırma Analizi (YDAA) ve İzometrik İzdüşümdür (İZİ).

Doğrusal boyut indirgeme yöntemleri kullanılarak bir veri kümesinin sınıflandırılmasında doğrulukta bir artma veya azalma olmasının veri kümesinin türüne bağlı olduğu görülmüştür. Bazı boyut indirgeme yöntemlerinin sınıflandırma doğrulukları açısından veri kümelerinin türüne göre öne çıktığı görülmüştür. Hatta veri kümelerinden elde edilen özniteliklere göre bazı yöntemlerin öne çıktığı da görülmüştür.

Anahtar Kelimeler: Doğrusal Boyut İndirgeme, Yüz Tanıma, Rakam Tanıma.

Master Thesis

SUMMARY

COMPARISON OF LINEAR DIMENSIONALITY REDUCTION METHODS ON
BIG DATA CLASSIFICATION

Eray YILDIZ

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Electrical and Electronics Engineering Graduate Program
Supervisor: Asst. Prof. Yusuf SEVİM
2018, 70 Pages

Analysis of high-dimensional data is encountered in a lot of areas. In analysis of high-dimensional data, processing data effectively and fast may be troublesome in terms of time and process load. Dimensionality reduction methods are used commonly to overcome this trouble. Linear dimensionality reduction methods give better results than nonlinear dimensionality reduction methods in terms of time and process load. Therefore, they are preferred more in analysis of high-dimensional data.

In this thesis, various popular linear dimensionality reduction methods and their performance were investigated on various classification methods by using face and digit datasets. These methods are Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Locality Preserving Projection (LPP), Neighborhood Preserving Embedding (NPE), Locality Sensitive Discriminant Analysis (LSDA) and Isometric Projections(IsoP).

In the classification of a dataset by using linear dimensionality reduction methods, it was seen that an increase or a decrease in accuracy depends on the kind of the dataset. And it was seen that some linear dimensionality reduction methods become prominent according to the kind of datasets in terms of classification accuracy. Also, it was seen that some linear dimensionality reduction methods become prominent according to features obtained from datasets.

Key Words: Linear Dimensionality Reduction, Face Recognition, Digit Recognition.

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1.1.	DAA yöntemi için bir örnek [14]..... 9
Şekil 1.2.	a) G_w ve G_b komşuluk grafları b) farklı sınıf örnekleri arasındaki marjin [29]..... 18
Şekil 1.3.	Jeodezik uzaklık örneği..... 25
Şekil 1.4.	Çok katmanlı ileri beslemeli bir yapay sinir ağı örneği..... 31
Şekil 1.5.	a) Dar marjin b) Geniş marjin [40]. 35
Şekil 2.1.	ORL veri kümesi için bazı örnekler [42]. 40
Şekil 2.2.	Yüz kısmı kesip çıkartılmış örnekler 40
Şekil 2.3.	USPS veri kümesi için bazı örnekler 41
Şekil 2.4.	a) x b) y eksenî yönünde gradyan vektörünün bileşenini yaklaşık hesaplayan filtre 42
Şekil 2.5.	Standart yönler ve bileşenlere ayırma..... 43
Şekil 2.6.	Diyagonal öznelik çıkarma 44

TABLolar DİZİNİ

	<u>Sayfa No</u>
Tablo 2.1. USPS veri kümesi örneklerinin sınıflara göre sayısı	41
Tablo 2.2. USPS veri kümesi örneklerinin sınıflara göre oranları	41
Tablo 2.3. Orijinal öznitelikler üzerinde sınıflandırma.....	45
Tablo 2.4. Boyut indirgeme işlemi için süreler	45
Tablo 2.5. <i>k</i> -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	46
Tablo 2.6. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	47
Tablo 2.7. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	47
Tablo 2.8. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(ms).....	48
Tablo 2.9. ORL veri kümesi için tüm sonuçlar	48
Tablo 2.10. Orijinal öznitelikler üzerinde sınıflandırma.....	50
Tablo 2.11. Boyut indirgeme işlemi için süreler	50
Tablo 2.12. <i>k</i> -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	51
Tablo 2.13. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	52
Tablo 2.14. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	52
Tablo 2.15. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(s).....	53
Tablo 2.16. USPS-Gradyan öznitelikler için tüm sonuçlar	53
Tablo 2.17. Orijinal öznitelikler üzerinde sınıflandırma.....	55
Tablo 2.18. Boyut indirgeme işlemi için süreler	55
Tablo 2.19. <i>k</i> -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	56
Tablo 2.20. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	57
Tablo 2.21. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	57
Tablo 2.22. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(s).....	58
Tablo 2.23. USPS-Diyagonal öznitelikler için tüm sonuçlar	58
Tablo 2.24. Orijinal öznitelikler üzerinde sınıflandırma.....	59

Tablo 2.25.	Boyut indirgeme işlemi için süreler.....	60
Tablo 2.26.	k -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	61
Tablo 2.27.	DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	62
Tablo 2.28.	YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%).....	62
Tablo 2.29.	Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(s).....	63
Tablo 2.30.	USPS-Gradyan ve Diyagonal öznitelikler için tüm sonuçlar	63



SEMBOLLER DİZİNİ

DAA	Doğrusal Ayırma Analizi
FDA	Fisher Doğrusal Ayırma
İZİ	İzometrik İzdüşüm
KKG	Komşuluk Koruyan Gömme
TBA	Temel Bileşenler Analizi
YDAA	Yerellik Duyarlı Ayırma Analizi
YKİ	Yerellik Koruyan İzdüşüm
C	Kovaryans Matrisi
d	X Giriş Verisi Boyutu
D_J	Jeodezik Uzaklık Matrisi
G	Komşuluk Grafi
G_B	Sınıflar Arası Komşuluk Grafi
G_W	Sınıf İçi Komşuluk Grafi
N	k -En Yakın Komşular Kümesi
P	Dönüşüm Matrisi
r	Y Çıkış Verisi Boyutu
S_B	Sınıflar Arası Saçılım Matrisi
S_W	Sınıf İçi Saçılım Matrisi
W	Ağırlık Matrisi
W_B	Sınıflar Arası Ağırlık Matrisi
W_W	Sınıf İçi Ağırlık Matrisi
X	Giriş Verisi
Y	Çıkış Verisi
Ψ	Lagrange Fonksiyonu

1. GENEL BİLGİLER

1.1. Giriş

Günümüzde gelişen teknoloji ile mühendislik, bilim, iş dünyası, tıp, toplum ve benzer alanlardan bilgisayar ağlarına veya çeşitli veri depolama alanlarına çok miktarda veri akmaktadır. İleri derecede sayısallaşan dünya ve gelişmiş veri depolama ile gün geçtikçe ortaya çıkan verinin miktarı artmaktadır. Bu tür yüksek boyutlu verilerin içinde önemli bilgiler yer almaktadır. Dolayısıyla bu önemli bilgileri elde edebilmek için yüksek boyutlu verileri verimli bir şekilde işlemek gerekmektedir. Yüksek boyutlu verilerin analiz edildiği birçok uygulama alanı vardır. Bu alanlar arasında ses işleme, bilgisayarla görme, uzaktan algılama, yüz tanıma ve veri madenciliği örnek gösterilebilir.

İstatistik, mühendislik veya birçok ilgili alanda yüksek boyutlu verilerin üzerinde daha kolay ve verimli çalışabilmek için farklı yöntemler bulunmuştur ve hala daha bu konu üzerinde çalışmalar yapılmaya devam edilmektedir. Her geçen gün artan veri miktarıyla, yüksek boyutlu verileri işleyerek analiz etmek için güçlü ve çok yönlü yöntemlere ihtiyacımız vardır.

Yüksek boyutlu verilerde, büyük veri boyutu işlem yükü ve zaman açısından verinin etkin ve hızlı bir şekilde işlenmesini zorlaştırmaktadır. Buna yüksek boyutluluğun laneti denilmektedir [1]. Bu zorluğun üstesinden gelebilmek için boyut indirgeme yöntemleri yaygın bir şekilde kullanılmaktadır.

Boyut indirgeme, orijinal verinin temsil edileceği daha düşük boyutlu bir uzay bulma işlemidir [2]. Boyut indirgeme işlemi ile ilgilenmemizin birkaç sebebi vardır:

- Veri sıkıştırma ve veriyi depolamak için gerekli alanın azalmasını sağlar.
- Verideki gereksiz veya tekrarlanan özelliklerin tespit edilmesini sağlar.
- Sınıflandırma, kümeleme, olasılık yoğunluk kestirimi gibi çoğu öğrenme ve istatistiksel yöntemlerde karmaşıklık verinin boyutuna bağlıdır. Azaltılmış işlem ve bellek için veri boyutu azaltılmalıdır.

- Veri daha düşük boyut ile ifade edildiğinde, verinin içinde yatan yapılar hakkında daha kolay fikir elde edilir ve bu da daha kolay bilgi çıkarımına olanak sağlar.
- Veri çok düşük boyutlarda (2 veya 3) kayıpsız bir şekilde temsil edildiğinde, saçılım grafikleri çizdirilerek verinin görsel olarak yapısı gözlemlenebilir ve aykırı değer analizi yapılabilir.

1.2. Boyut İndirgeme

Veri kümesinin bir veri örneği denklem (1.1)'de gösterildiği gibi ifade edilirse, d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ denklem (1.2)'deki gibi $d \times n$ boyutlu bir X matrisi ile temsil edilir. Denklem (1.2)'deki matrisin her bir sütunundaki \mathbf{x}_i 'ler verinin bir örneğini ve her bir satırı ise verinin bir değişkenini (X_i) veya boyutunu temsil etmektedir.

$$\mathbf{x}_i = \begin{bmatrix} x_{i,1} \\ x_{i,2} \\ \vdots \\ x_{i,d} \end{bmatrix} \quad (1.1)$$

$$X = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n] = \begin{bmatrix} x_{1,1} & \dots & x_{n,1} \\ \vdots & \ddots & \vdots \\ x_{1,d} & \dots & x_{n,d} \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix} \quad (1.2)$$

Üsteki denklemdeki veriye boyut indirgeme işlemi uygulandıktan sonra elde edilecek r boyutlu n tane veri örneği $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ denklem (1.3)'teki gibi $r \times n$ boyutlu bir matris Y ile ifade edilir.

$$Y = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \dots \quad \mathbf{y}_n] = \begin{bmatrix} y_{1,1} & \dots & y_{n,1} \\ \vdots & \ddots & \vdots \\ y_{1,r} & \dots & y_{n,r} \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_r \end{bmatrix} \quad (1.3)$$

Boyut indirgeme işlemi d boyutlu X veri kümesini r boyutlu Y veri kümesine dönüştüren bir fonksiyon olarak görülebilir ($r < d$). Burada X giriş verisi ve Y çıkış verisi olarak adlandırılabilir. Genel olarak veri boyutu indirgeme işlemi aşağıdaki denklemdeki gibi bir fonksiyon ile ifade edilebilir:

$$Y = f(X) \quad (1.4)$$

Burada kullanılan fonksiyon doğrusal veya doğrusal olmayan bir fonksiyon olabilir. Dolayısıyla boyut indirgeme yöntemleri doğrusal ve doğrusal olmayan olmak üzere ikiye ayrılır. Doğrusal boyut indirgeme yöntemleri, işlem yükü ve çalışma zamanı açısından doğrusal olmayan yöntemlere göre daha performanslıdır. Bu yüzden, doğrusal olmayan yöntemlere göre uygulanması daha kolay ve basit oldukları için yüksek boyutlu veri analizinde daha çok tercih edilmektedirler. Ayrıca yapılan çalışmalarda, doğrusal olmayan yöntemlerin sahip olduğu karmaşık ve doğrusal olmayan verilerin yapısını öğrenme yeteneğine rağmen, doğrusal olmayan yöntemlerin gerçek veri kümeleri üzerinde doğrusal yöntemlere göre başarılı sonuçlar ortaya koymadığı görülmektedir [3].

1.3. Doğrusal Boyut İndirgeme Yöntemleri

Doğrusal boyut indirgeme yöntemleri istatistik, makine öğrenmesi ve ilgili alanlarda geliştirilen ve yüksek boyutlu veri analizinde sıkça başvurulan yöntemlerdendir. Sahip olduğu özellikler olan geometrik olarak kolay yorumlanabilme ve doğrusal olmayan yöntemlere göre daha az hesaplama karmaşıklığına sahip olma, bu yöntemleri yüksek boyutlu veri analizinde öne çıkarmaktadır. Bu yöntemler verinin ilgilenilen varyans, dinamik yapısı, veri sınıfları arasındaki marjinler gibi birçok özelliğini ortaya çıkarabilmektedir [4].

Doğrusal boyut indirgeme yöntemlerinin çalışma prensibi, bir kriteri optimize ederek yüksek boyutlu verinin doğrusal olarak izdüşüm yapılacağı daha düşük boyutlu bir uzayda

ifade edilmesidir [5]. Bir başka bakış açısıyla, yeni veri değişkenleri orijinal veri değişkenlerinin doğrusal kombinasyonlarıyla elde edilmektedir. Bu yöntemler, yüksek d boyutlu x_i veri örneklerini giriş olarak alır ve bunu düşük r boyutlu y_i veri örneklerine dönüştüren bir $d \times r$ dönüşüm matrisi P bulurlar. Matematiksel olarak d boyutlu n örneğe sahip bir veri kümesinin örnekleri P dönüşüm matrisinin transpozu ile çarpılarak r boyutlu örnekler elde edilebilir.

$$y_i = P^T x_i \quad i = 1, \dots, n \quad (1.5)$$

Giriş ve çıkış verileri denklem (1.2) ve (1.3)'teki gibi $d \times n$ boyutlu bir matris X ve $r \times n$ boyutlu bir matris Y ile ifade edilirse denklem (1.6)'daki gibi P dönüşüm matrisi kullanılarak doğrusal boyut indirgeme işlemi tanımlanabilir.

$$Y = P^T X \quad (1.6)$$

Doğrusal boyut indirgeme yöntemlerindeki temel fark ise, farklı doğrusal boyut indirgeme yöntemlerinin farklı P dönüşüm matrisi oluşturmalarıdır. Yöntemlere göre, bu matris oluşturulurken verinin sahip olduğu özelliklerden hangisinin yansıtılmak ve korunmak istendiğine göre P matrisi şekillenmektedir. Verinin ilgilenilen özelliğinin korunması için uygun bir kriter bulunur ve optimize edilir. Bu optimizasyon işlemi ile P dönüşüm matrisi elde edilir.

1.3.1. Temel Bileşen Analizi

Temel Bileşen Analizi (TBA) yöntemi ilk olarak 1901 yılında Karl Pearson tarafından ortaya konulmuştur [6]. 1933 yılında Hotelling tarafından farklı bir bakış açısıyla yorumlanmış ve yine Hotelling tarafından hızlandırılmış bir versiyonu da geliştirilmiştir [7-9]. 1964 yılında Rao, TBA yönteminin kullanımı ile ilgili yeni fikirler ve yorumlamalar ortaya koymuştur [9, 10]. Bu gelişmelerden sonra birçok alanda TBA yöntemi kullanılmaya başlanmıştır.

TBA yöntemi, çok boyutlu veri analizinde sıkça kullanılan gözetimsiz bir doğrusal boyut indirgeme yöntemidir. TBA yönteminin uygulamalı doğrusal cebirden gelen en

değerli sonuçlardan biri olduğu söylenmektedir [11]. TBA yöntemi verinin sınıf bilgisini boyut indirgeme işleminde kullanmaz. Bu sebepten dolayı gözetimsiz öğrenme yöntemi olarak isimlendirilmektedir.

TBA, veri kümesinde bulunan muhtemel ilintili değişkenleri temel bileşenler olarak adlandırılan doğrusal ilintisiz değişkenlere dönüştürmek için ortogonal dönüşüm kullanan bir yöntemdir. TBA bu dönüştürme işlemini yaparken, boyutu azaltılmış Y çıkış veri kümesinin, X giriş veri kümesinde var olan varyansı mümkün olduğunca muhafaza etmesi sağlanır [5].

TBA yönteminde veriden elde edilen birinci ve ikinci dereceden istatistik değerlerini kullanılmaktadır. X giriş verisinin d boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ olduğu varsayılırsa, birincil istatistiksel değer veri örneklerinin ortalama değeridir. Bir veri kümesinin ortalama değeri, denklem (1.7)'de gösterildiği gibi elde edilir:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad (1.7)$$

İkinci dereceden istatistiksel değer veri değişkenlerinin varyansıdır. Bir rastgele değişkenin varyansı, o değişkenin ortalama değer etrafında yayılımını veya değişkenliğini gösterir. Verinin bir değişkeni olan X_i için varyans denklem (1.8) ile hesaplanır. Bu denklemde \bar{x}_i ve $\mathbf{x}_{k,i}$, X_i değişkenin ortalama değerini ve X_i değişkeninin k . örneğini ifade etmektedir.

$$\text{var}(X_i) = \sigma_{x_i}^2 = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{x}_{k,i} - \bar{x}_i)^2 \quad (1.8)$$

İki rastgele değişken arasındaki birlikte değişim kovaryans ile tespit edilir. Verinin iki değişkeni X_i ve X_j için kovaryans denklem (1.9) ile hesaplanır. Bu denklemde \bar{x}_i ve \bar{x}_j , X_i ve X_j değişkenlerinin ortalama değeridir. $\mathbf{x}_{k,i}$ ve $\mathbf{x}_{k,j}$, X_i ve X_j değişkenlerinin k . örneklerini ifade etmektedir.

$$\text{cov}(X_i, X_j) = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{x}_{k,i} - \bar{\mathbf{x}}_i)(x_{k,j} - \bar{x}_j) \quad (1.9)$$

Çok değişkenli verilerde verinin değişkenleri arasındaki kovaryansların tümünü görebilmek için bir kovaryans matrisi oluşturabilir. Denklem (1.10) ile ifade edilen kovaryans matrisi, $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i)$ olduğundan simetrik bir matristir.

$$C = \begin{bmatrix} \text{cov}(X_1, X_1) & \cdots & \text{cov}(X_1, X_d) \\ \vdots & \ddots & \vdots \\ \text{cov}(X_d, X_1) & \cdots & \text{cov}(X_d, X_d) \end{bmatrix} \quad (1.10)$$

X veri kümesinin kovaryans matrisi elde edilirken verinin ortalama değeri de kullanılır. Denklem (1.11)'de veri örnekleri üzerinden kovaryans matrisinin elde edilmesi gösterilmiştir.

$$C = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (1.11)$$

TBA yönteminde ilk önce verinin eğer ortalama değeri 0'dan farklı ise, ortalama olarak adlandırılan işlem yapılır. Ortalama işlemi için bütün veri örneklerinden denklem (1.12)'deki gibi ortalama değer çıkartılır. Daha sonra ortalanmış veri denklem (1.2)'deki gibi dxn boyutlu bir matris X_c ile ifade edilirse, C kovaryans matrisi denklem (1.13) kullanılarak da bulunabilir.

$$\mathbf{x}_{c,i} = (\mathbf{x}_i - \bar{\mathbf{x}}) \quad (1.12)$$

$$C = \frac{1}{n-1} X_c X_c^T \quad (1.13)$$

Bir d boyutlu veri örneğinin, $dx1$ boyutlu \mathbf{p} vektörü üzerine izdüşümü alınırken denklem (1.14)'teki gibi çarpılır ve y_i değerleri elde edilir. Eğer veri (1.1) denklemdeki gibi

$d \times n$ boyutlu bir matris X_c ile ifade edilirse, denklem (1.15) kullanılarak $1 \times n$ boyutlu Y çıkış verisi elde edilir.

$$y_i = \mathbf{p}^T \mathbf{x}_{c,i} \quad i = 1, \dots, n \quad (1.14)$$

$$Y = \mathbf{p}^T X_c \quad (1.15)$$

TBA yönteminde amaç verinin varyansını mümkün olduğunca korumak olduğu için, X_c verisi öyle bir vektör üzerine izdüşüm yapılmalı ki bu yön maksimum varyansa sahip olsun. Bunun için Y çıkış verisinin varyansı denklem (1.16)'da gösterilen şekilde ifade edilir. Y verisi elde edilirken X verisinin ortalananmış hali kullanıldığı için, Y verisinin ortalama değeri 0'dır. Dolayısıyla Y verisinin varyansı bulunurken ortalama değer kullanılmasına gerek yoktur.

$$\begin{aligned} \sigma_Y^2 &= \frac{1}{n-1} Y Y^T = \frac{1}{n-1} (\mathbf{p}^T X_c) (\mathbf{p}^T X_c)^T \\ &= \mathbf{p}^T \left(\frac{1}{n-1} X_c X_c^T \right) \mathbf{p} = \mathbf{p}^T C \mathbf{p} \end{aligned} \quad (1.16)$$

Denklem (1.16)'da görüldüğü gibi Y verisinin varyansı, X verisinin kovaryans matrisi üzerinden elde edilebilir. Y verisinin varyansını maksimum yapan \mathbf{p} vektörünün yönelimi önemlidir. Bu yüzden $|\mathbf{p}|=1$ alınabilir. $|\mathbf{p}|=1$ kısıtlaması ile maksimize edilmesi gereken J amaç fonksiyonu denklem (1.17)'de gösterilen şekilde ifade edilir:

$$J(\mathbf{p}) = \mathbf{p}^T C \mathbf{p} \quad (1.17)$$

Bu optimizasyon problemi Lagrange çarpanları tekniği ile (λ lagrange çarpanıdır) çözülmek istenirse, ilk önce denklem (1.18)'te yer alan ψ Lagrange fonksiyonu yazılır. Denklem (1.18)'in \mathbf{p} vektörüne göre türevi alınıp 0'a eşitlenirse, bu sonucu sağlayan \mathbf{p} vektörü denklem (1.17)'de yer alan J amaç fonksiyonunu maksimum yapar. Bu türevin sonucu denklem (1.19)'da gösterilmiştir.

$$\psi = \mathbf{p}^T C \mathbf{p} - \lambda (\mathbf{p}^T \mathbf{p} - 1) \quad (1.18)$$

$$\begin{aligned} \frac{\partial \psi}{\partial \mathbf{p}} &= 2C\mathbf{p} - 2\lambda\mathbf{p} = 0 \\ \Rightarrow C\mathbf{p} &= \lambda\mathbf{p} \end{aligned} \quad (1.19)$$

Denklem (1.19)'da sonucun 0 olması için, \mathbf{p} vektörü ve λ skalar değeri verinin C kovaryans matrisinin özvektörleri ve özdeğerleri olmalıdır. Denklem (1.20)'de ifade edildiği gibi, denklem (1.17)'deki J amaç fonksiyonunun maksimum olması için, verinin C kovaryans matrisinin değer olarak en büyük özdeğer λ_1 ve bu özdeğere ait özvektör \mathbf{u}_1 izdüşüm vektörü olarak kullanılmalıdır.

$$\begin{aligned} J(\mathbf{u}_1) &= \mathbf{u}_1^T C \mathbf{u}_1 = \mathbf{u}_1^T \lambda_1 \mathbf{u}_1 \\ &= \lambda_1 \mathbf{u}_1^T \mathbf{u}_1 = \lambda_1 \end{aligned} \quad (1.20)$$

Birden daha fazla boyut için Lagrange çarpanları yöntemini kullanarak benzer türetmeler ile diğer vektörlerin C kovaryans matrisinin değerce en büyük şekilde özdeğerlerine karşılık gelen özvektörler olduğu bulunur [12]. Dolayısıyla en büyük değere sahip r tane özdeğer $\lambda_1, \lambda_2, \dots, \lambda_r$ seçilir ve P dönüşüm matrisi bu öz değerlere karşılık gelen r tane özvektörler $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ kullanılarak oluşturulur. Denklem (1.5) ve (1.6) kullanılarak boyutu azaltılmış veriler elde edilir.

TBA yönteminin adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin \bar{x} ortalama değerinin bulunması.
2. X veri kümesinin ortalanması.
3. C kovaryans matrisinin bulunması.
4. C kovaryans matrisinin r tane en büyük özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ ve bu özdeğerlere ait özvektörlerinin $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ bulunması.
5. P dönüşüm matrisinin oluşturulması.

Çıkış: $d \times r$ boyutlu P dönüşüm matrisi.

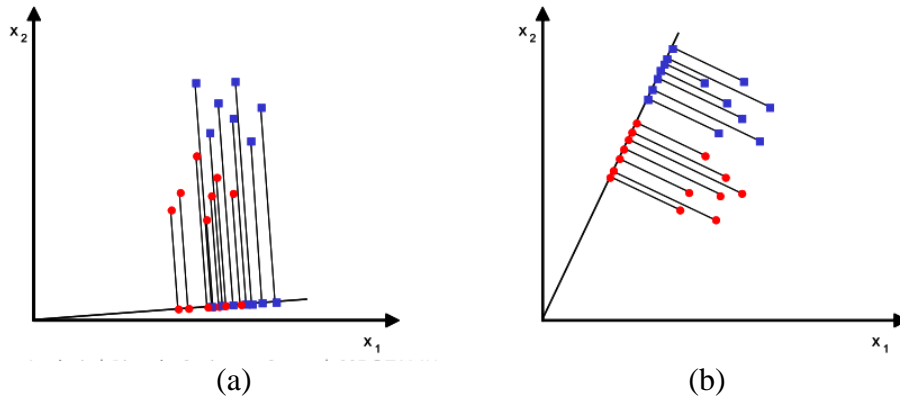
r değerinin matematiksel olarak belirlenmesi gerekirse, bu boyutu azaltılmış veride yansıtılacak varyans değeri ile bulunur. Bir C kovaryans matrisinin t tane sıfırdan farklı özdeğeri olursa, X verisinde var olan toplam varyansın $\%n$ 'lik kısmını, boyutu azaltılmış Y verisinde yansıtmak için denklem (1.21) kullanılarak uygun r değeri bulunur.

$$\frac{\sum_{i=1}^r \lambda_i}{\sum_{i=1}^t \lambda_i} \geq \frac{n}{100} \quad (1.21)$$

1.3.2. Doğrusal Diskriminant Analizi

Doğrusal Ayırma Analiz (DAA) yöntemi ilk olarak R.A. Fisher tarafından 1936 yılında ortaya konulmuştur [13]. Fisher Doğrusal Ayırma (FDA) ismiyle de bilinir. Eski bir yöntem olduğu için o tarihten sonra birçok alanda kullanılmıştır.

DAA yöntemi, yüksek boyutlu veri analizinde sıkça kullanılan bir boyut indirgeme yöntemidir. DAA yönteminde, verinin sınıf bilgisi de boyut indirgeme işleminde kullanılır. Bu sebeple TBA yönteminden farklı olarak gözetimli öğrenme yöntemlerine girmektedir. DAA yöntemi, veri sınıfları arasındaki ayrımı maksimum yapan P dönüşüm matrisinin bulunmasını amaçlayan bir yöntemdir. Şekil 1.1'de DAA yöntemi için bir örnek verilmiştir. Şekil incelendiğinde veri sınıfları arasındaki ayrımı, Şekil 1.1(b)'deki izdüşüm vektörünün Şekil 1.1(a)'daki izdüşüm vektörüne göre daha iyi yaptığı görülmektedir.



Şekil 1.1. DAA yöntemi için bir örnek [14].

DAA yönteminde sınıflar arası ayrımı yakalayabilmek için veriden elde edilen bazı istatistiksel değerler kullanılmaktadır. X giriş verisinin d boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ve her biri n_i tane örneğe sahip c tane sınıfa sahip olduğu varsayılırsa bu istatistiksel değerlerden ilki veri sınıflarının ortalama değerleridir. Veri sınıflarının ortalama değerleri $\bar{\mathbf{x}}_i$ ile ifade edilir ve denklem (1.22)'de gösterilen şekilde her bir sınıfa ait veri örneklerinin ortalaması ile elde edilirler.

$$\bar{\mathbf{x}}_i = \frac{1}{n_i} \sum_{i \in c} \mathbf{x}_i \quad (1.22)$$

DAA yönteminde kullanılan bir diğer istatistik ise sınıf içi saçılım matrisidir ve S_W ile gösterilir. Sınıf içi saçılım matrisi veri sınıflarının örneklerinin, sınıfların ortalama değerleri etrafındaki dağılımını veya yayılımını gösterir [15]. Sınıf içi saçılım matrisi aşağıdaki gibi hesaplanır:

$$S_W = \sum_{i=1}^c \sum_{j \in c} (\mathbf{x}_j - \bar{\mathbf{x}}_i)(\mathbf{x}_j - \bar{\mathbf{x}}_i)^T \quad (1.23)$$

DAA yönteminde kullanılan bir diğer istatistik, sınıflar arası saçılım matrisidir ve S_B ile gösterilir. Sınıflar arası saçılım matrisi veri sınıflarının ortalama değerlerinin, verinin ortalama değeri etrafındaki dağılımını veya yayılımını gösterir [15]. Sınıflar arası saçılım matrisi aşağıdaki gibi hesaplanır:

$$S_B = \sum_{i=1}^c n_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T \quad (1.24)$$

DAA yönteminde amaç veri sınıfları arasındaki ayrımı maksimum yapmak olduğu için, veri örnekleri öyle bir \mathbf{p} vektörü üzerine izdüşüm yapılmalı ki, bu yönde Y çıkış verisinin $S_{B,Y}$ sınıflar arası saçılımı maksimum ve $S_{W,Y}$ sınıf içi saçılımı minimum olsun. X verisinin \mathbf{p} vektörü üzerine izdüşümü alınır, $1 \times n$ boyutlu Y verisi elde edilir. Tek izdüşüm vektörü kullanıldığı için denklem (1.25)'te gösterilen $S_{B,Y}$ ve denklem (1.26)'da gösterilen

$S_{W,Y}$ skaler değer olurlar ve iki skaler değer kullanılarak denklem (1.27)'deki amaç fonksiyonu elde edilir.

$$S_{B,Y} = \mathbf{p}^T S_{B,X} \mathbf{p} \quad (1.25)$$

$$S_{W,Y} = \mathbf{p}^T S_{W,X} \mathbf{p} \quad (1.26)$$

$$J(\mathbf{p}) = \frac{S_{B,Y}}{S_{W,Y}} = \frac{\mathbf{p}^T S_{B,X} \mathbf{p}}{\mathbf{p}^T S_{W,X} \mathbf{p}} \quad (1.27)$$

Bu optimizasyon problemini çözmek için, J amaç fonksiyonunun \mathbf{p} vektörüne göre türevi alınıp 0'a eşitlenirse, J amaç fonksiyonunu maksimum yapan \mathbf{p} vektörü elde edilir. Bu türevin sonucu denklem (1.28)'de gösterilmiştir.

$$\frac{\partial J}{\partial \mathbf{p}} = \frac{2\left(\left(\mathbf{p}^T S_{W,X} \mathbf{p}\right) S_{B,X} \mathbf{p} - \left(\mathbf{p}^T S_{B,X} \mathbf{p}\right) S_{W,X} \mathbf{p}\right)}{\left(\mathbf{p}^T S_{W,X} \mathbf{p}\right)^2} = 0 \quad (1.28)$$

$$S_{B,X} \mathbf{p} = \left(\frac{\mathbf{p}^T S_{B,X} \mathbf{p}}{\mathbf{p}^T S_{W,X} \mathbf{p}} \right) S_{W,X} \mathbf{p} \quad (1.29)$$

Denklem (1.28)'in 0'a eşit olması için denklem (1.29)'daki eşitliğin sağlanması gerekmektedir. Bu denklemde parantez içindeki ifade skaler bir değerdir. Böylece denklem (1.29), denklem (1.30)'daki genelleştirilmiş özdeğer problemine dönüşür. Parantez içindeki ifade λ skalar değeri ile gösterilirse, λ skalar değeri S_W ve S_B matrislerinin genelleştirilmiş özdeğeri ve \mathbf{p} vektörü de bu matrislerin genelleştirilmiş özvektörü olmalıdır.

$$S_{B,X} \mathbf{p} = \lambda S_{W,X} \mathbf{p} \quad (1.30)$$

Denklem (1.27)'deki J amaç fonksiyonunun maksimum olması için, S_W ve S_B matrislerinin değer olarak en büyük genelleştirilmiş özdeğeri λ_1 ve bu genelleştirilmiş özdeğere ait genelleştirilmiş özvektör \mathbf{u}_1 izdüşüm vektörü olarak kullanılmalıdır.

Birden daha fazla boyut bulmak için denklem (1.31) veya (1.32)'deki amaç fonksiyonlarından herhangi biri kullanılabilir. Bu amaç fonksiyonlarında S_W ve S_B için, önceki amaç fonksiyonundaki gibi skaler değerler değil matrisler elde edilir. Dolayısıyla bu amaç fonksiyonlarında bu matrislerin trace veya determinantları kullanılmıştır.

$$J(P) = tr \left(\frac{P^T S_{B,X} P}{P^T S_{W,X} P} \right) \quad (1.31)$$

$$J(P) = \frac{|P^T S_{B,X} P|}{|P^T S_{W,X} P|} \quad (1.32)$$

Bu iki amaç fonksiyonunu maksimum yapan P dönüşüm matrisleri her iki amaç fonksiyonu için de aynı olmaktadır. Dolayısıyla S_W ve S_B matrislerinin en büyük değere sahip r tane genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ seçilir ve P dönüşüm matrisi bu öz değerlere karşılık gelen r tane genelleştirilmiş özvektörler $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ kullanılarak oluşturulur. Denklem (5) ve (6) kullanılarak boyutu azaltılmış veriler elde edilir.

DAA yönteminde c sınıflı bir veri için bir diğer önemli nokta ise, S_B matrisinin rank değerinin en fazla $(c-1)$ olmasıdır. Dolayısıyla elde edilen değeri sıfırdan farklı olan özdeğer sayısı $(c-1)$ olmaktadır. Diğer özdeğerler sıfır değerine sahip olmaktadır. DAA yönteminde c sınıflı bir veri için seçilen r boyut değeri en fazla $(c-1)$ olabilmektedir. Bu sebeple $r \leq (c-1)$ olmalıdır [5, 16].

DAA yönteminin adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin ortalama değerinin $\bar{\mathbf{x}}$ ve sınıfların ortalama değerlerinin $\bar{\mathbf{x}}_i$ bulunması.
2. S_B sınıflar arası saçılım matrisinin bulunması.
3. S_W sınıf içi saçılım matrisinin bulunması.

4. S_W ve S_B matrislerinin r tane en büyük genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ ve bu özdeğerlere ait özvektörlerinin $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ bulunması.

5. P dönüşüm matrisinin oluşturulması.

Çıktı: $d \times r$ boyutlu P dönüşüm matrisi.

1.3.3. Yerellik Koruyan İzdüşüm

Yerellik Koruyan İzdüşüm (YKİ) yöntemi yüksek boyutlu veri analizinde son zamanlarda karşılaşılan bir boyut indirgeme yöntemidir. İlk olarak 2003 yılında Xiaofei He ve arkadaşları tarafından geliştirilmiştir [17, 18]. YKİ yöntemi bundan sonra yüz tanıma [19, 20], konuşma tanıma [21], hiperspektral görüntü sınıflandırma [22, 23], avuç izi tanıma [24] gibi alanlarda kullanılmıştır.

YKİ yöntemi verinin sınıf bilgisini boyut indirgeme işleminde göz önünde bulundurmaz. Bu sebeple gözetimsiz öğrenme yöntemlerine girmektedir. YKİ yöntemi verinin komşuluk yapısını koruyan bir amaç fonksiyonunu optimize ederek P dönüşüm matrisini elde eder.

YKİ yönteminin algoritmasına geçmeden önce veriden elde edilen G komşuluk grafi ve W ağırlık matrisinden bahsetmek gerekmektedir. X giriş verisinin d boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ olsun. G komşuluk grafi veri örneklerinin birbirleriyle olan ilişkisini ve komşuluklarını ortaya koyar. G komşuluk grafi verinin n tane örneğine karşılık gelen n tane düğümden oluşur. Eğer \mathbf{x}_i ve \mathbf{x}_j veri örnekleri birbirlerine yakınlarsa, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur. Örneklerin yakınlık ölçüsü için iki yöntem kullanılır [18]:

1. ε -komşuluğu: Eğer $\|\mathbf{x}_i - \mathbf{x}_j\| < \varepsilon$ ise, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur.

2. k -en yakın komşular: Eğer \mathbf{x}_i örneği \mathbf{x}_j örneğinin en yakın k komşusu arasında ise, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur.

G komşuluk grafında düğümleri birleştiren kenarların her biri için bir W_{ij} ağırlık değeri olmalıdır. Bu değer birbirleri arasında kenar olmayan düğümlerin kenarları için 0 alınır. Kenarı olan i ve j düğümleri için ise iki şekilde belirlenir [18]:

1. Eğer i ve j düğümleri birbirine bağlı ise, bu kenarın ağırlığı aşağıdaki denklem ile elde edilir ($t \in \mathcal{R}$).

$$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{t}} \quad (1.33)$$

2. Eğer i ve j düğümleri birbirine bağlı ise, bu kenarın ağırlığı $W_{ij} = 1$ olur.

Veri örneklerinin birbirleriyle olan ilişkisini gösteren W_{ij} ağırlık değerleri bir W ağırlık matrisi ile ifade edilebilir. Bu matrisin boyutu $n \times n$ dir ve simetriktir.

YKİ yönteminde veri örnekleri öyle bir \mathbf{p} vektör üzerine izdüşüm yapılmalı ki birbiriyle komşu olan örnekler bu vektör üzerinde birbirlerine yakın olsunlar. Bunun sağlanması için denklem (1.34)'teki amaç fonksiyonu seçilmiştir. Bu amaç fonksiyonunda $y_i = \mathbf{p}^T \mathbf{x}_i$ olmak üzere, W_{ij} ifadesi \mathbf{x}_i ve \mathbf{x}_j örneklerinin ağırlık değeridir. Eğer komşu örnekler izdüşüm yapıldığında birbirlerine uzak düşerlerse amaç fonksiyonu büyük bir değere sahip olur. Dolayısıyla komşu örnekleri birbirlerine yakın yerlere izdüşüm yapmak için J amaç fonksiyonu minimize edilmelidir.

$$\begin{aligned} J(\mathbf{p}) &= \frac{1}{2} \sum_{ij} (y_i - y_j)^2 W_{ij} \\ &= \frac{1}{2} \sum_{ij} (\mathbf{p}^T \mathbf{x}_i - \mathbf{p}^T \mathbf{x}_j)^2 W_{ij} \end{aligned} \quad (1.34)$$

Eğer amaç fonksiyonu düzenlenirse denklem (1.35) elde edilir.

$$\begin{aligned} J(\mathbf{p}) &= \sum_i \mathbf{p}^T \mathbf{x}_i D_{ii} \mathbf{x}_i^T \mathbf{p} - \sum_{ij} \mathbf{p}^T \mathbf{x}_i W_{ij} \mathbf{x}_j^T \mathbf{p} \\ &= \mathbf{p}^T X(D - W)X^T \mathbf{p} = \mathbf{p}^T XLX^T \mathbf{p} \end{aligned} \quad (1.35)$$

Bu denklemde D matrisi diyagonal bir matristir ve köşegende yer alan elemanlar denklem (1.36) ile elde edilir. L matrisi Laplacian matrisi olarak adlandırılır ve $(D - W)$ ile bulunur.

$$D_{ii} = \sum_j W_{ij} \quad (1.36)$$

D matrisi veri örnekleri üzerinde doğal bir ölçü verir. D_{ii} değeri ne kadar büyükse, y_i veri örneğinin önemi o kadar büyük olur. Dolayısıyla amaç fonksiyonu için denklem (1.37)'deki kısıtlama oluşturulmuştur [18].

$$Y^T D Y = 1 \Rightarrow \mathbf{p}^T X D X^T \mathbf{p} = 1 \quad (1.37)$$

Bu optimizasyon problemi Lagrange çarpanları tekniği ile çözülmek istenirse, ilk önce denklem (1.38)'te yer alan ψ Lagrange fonksiyonu yazılır. Denklem (1.38)'in \mathbf{p} vektörüne göre türevi alınıp 0'a eşitlenirse, bu sonucu sağlayan \mathbf{p} vektörü denklem (1.35)'te yer alan J amaç fonksiyonunu minimum yapar. Bu türevin sonucu denklem (1.39)'da gösterilmiştir.

$$\psi = \mathbf{p}^T X L X^T \mathbf{p} - \lambda (\mathbf{p}^T X D X^T \mathbf{p} - 1) \quad (1.38)$$

$$\begin{aligned} \frac{\partial \psi}{\partial \mathbf{p}} &= 2 X L X^T \mathbf{p} - 2 \lambda X D X^T \mathbf{p} = 0 \\ &\Rightarrow X L X^T \mathbf{p} = \lambda X D X^T \mathbf{p} \end{aligned} \quad (1.39)$$

Denklem (1.39)'da sonucun 0 olması için, λ skalar değeri $X L X^T$ ve $X D X^T$ matrislerinin genelleştirilmiş özdeğeri ve \mathbf{p} vektörü de bu matrislerin genelleştirilmiş özvektörü olmalıdır.

Denklem (1.35)'teki J amaç fonksiyonunun denklem (1.37)'deki kısıtlama altında minimum olması için, $X L X^T$ ve $X D X^T$ matrislerinin değer olarak en küçük genelleştirilmiş özdeğeri λ_1 ve bu genelleştirilmiş özdeğere ait genelleştirilmiş özvektör \mathbf{u}_1 izdüşüm vektörü olarak kullanılmalıdır.

Birden fazla boyut için $X L X^T$ ve $X D X^T$ matrislerinin en küçük değere sahip r tane genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ seçilir ve P dönüşüm matrisi bu özdeğerlere karşılık gelen r tane genelleştirilmiş özvektörler $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ kullanılarak oluşturulur. Denklem (1.5) ve (1.6) kullanılarak boyutu azaltılmış veriler elde edilir.

YKİ yönteminin adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin G komşuluk grafının ve W ağırlık matrisinin bulunması.
2. D diyagonal ve L Laplacian matrislerinin bulunması.
3. XLX^T ve XDX^T matrislerinin r tane en küçük genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ ve bu özdeğerlere ait özvektörlerinin $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ bulunması.
4. P dönüşüm matrisinin oluşturulması.

Çıkış: $d \times r$ boyutlu P dönüşüm matrisi.

1.3.4. Yerellik Duyarlı Ayırma Analizi

Yerellik Duyarlı Ayırma Analizi (YDAA) yöntemi yüksek boyutlu veri analizinde son zamanlarda karşılaşılan bir boyut indirgeme yöntemidir. YDAA yöntemi ilk olarak 2007 yılında Deng Cai ve arkadaşları tarafından çalışılmıştır [25]. YDAA yöntemi bundan sonra yüz tanıma [26], hiperspektral görüntü sınıflandırma [27,28], konuşmacı doğrulama [29] gibi alanlarda kullanılmıştır.

YDAA yönteminde verinin sınıf bilgisi boyut indirgeme işleminde kullanılır. Bu sebeple gözetimli öğrenme yöntemi olarak adlandırılmaktadır. Klasik DAA yöntemi verinin global geometrik yapısını göz önünde bulundurmaz, ama verinin yerel geometrik yapısını keşfetmekte başarısızdır. Bu yetersizliği gidermek için YDAA yöntemi geliştirilmiştir. YDAA yöntemi verinin yerel alanlarında farklı sınıflara ait veri örnekleri arasındaki marjini maksimum yapan P dönüşüm matrisini elde eder.

YDAA yönteminin temelini veriden elde edilen G_w sınıf içi komşuluk grafi, G_b sınıflar arası komşuluk grafi ve bu graflara ait W_w ve W_b ağırlık matrisleri oluşturmaktadır. X giriş verisinin d boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ olduğu varsayılır ise, her bir \mathbf{x}_i veri örneği için $N(\mathbf{x}_i) = \{\mathbf{x}_i^1, \mathbf{x}_i^2, \dots, \mathbf{x}_i^k\}$ kümesi, bu örneğin k -en yakın komşularından oluşturulur. Bu küme $N(\mathbf{x}_i)$, $N_w(\mathbf{x}_i)$ ve $N_b(\mathbf{x}_i)$ olarak gösterilen iki alt kümeye ayrılabilir. $N_b(\mathbf{x}_i)$ kümesi \mathbf{x}_i veri örneği ile aynı sınıf etiketini taşıyan komşulardan oluşur. $N_w(\mathbf{x}_i)$ kümesi ise \mathbf{x}_i veri örneği ile farklı sınıf etiketini taşıyan komşulardan oluşur. $l(\mathbf{x}_i)$, \mathbf{x}_i veri örneğinin sınıf

etiketini göstermek üzere, $N_w(\mathbf{x}_i)$ ve $N_b(\mathbf{x}_i)$ kümeleri denklem (1.40) ve (1.41)'de gösterildiği gibi tanımlanır.

$$N_w(\mathbf{x}_i) = \left\{ \mathbf{x}_i^j \mid l(\mathbf{x}_i^j) = l(\mathbf{x}_i), 1 \leq j \leq k \right\} \quad (1.40)$$

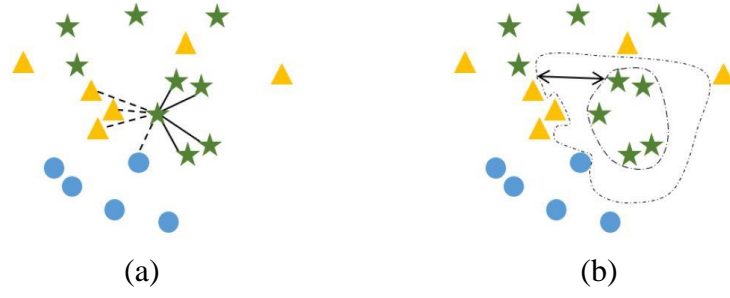
$$N_b(\mathbf{x}_i) = \left\{ \mathbf{x}_i^j \mid l(\mathbf{x}_i^j) \neq l(\mathbf{x}_i), 1 \leq j \leq k \right\} \quad (1.41)$$

G_w sınıf içi komşuluk grafi ve G_b sınıflar arası komşuluk grafi veri örneklerinin birbirleriyle ilişkisini ve komşuluklarını ortaya koyar. Bu komşuluk grafları verinin n tane örneğine karşılık gelen n tane düğümden oluşur. Eğer \mathbf{x}_i veri örneği için, \mathbf{x}_j veri örneği $N(\mathbf{x}_i)$ k -en yakın komşular kümesinde yer alıyorsa ve aynı sınıf etiketine sahip ise, komşuluk grafinda i ve j düğümleri arasına bir kenar koyulur. G_w sınıf içi komşuluk grafinda yer alan kenarların $W_{w,ij}$ ağırlıkları denklem (1.42)'de gösterilen şekilde bulunur. Eğer \mathbf{x}_i veri örneği için, \mathbf{x}_j veri örneği $N(\mathbf{x}_i)$ k -en yakın komşular kümesinde yer alıyorsa ve farklı sınıf etiketine sahip ise, komşuluk grafinda i ve j düğümleri arasına bir kenar koyulur. G_b sınıflar arası komşuluk grafinda yer alan kenarların $W_{b,ij}$ ağırlıkları denklem (1.43)'te gösterilen şekilde bulunur.

$$W_{w,ij} = \begin{cases} \mathbf{x}_i \in N_w(\mathbf{x}_j) \text{ yada } \mathbf{x}_j \in N_w(\mathbf{x}_i) \text{ ise, } 1 \\ \text{yoksa, } 0 \end{cases} \quad (1.42)$$

$$W_{b,ij} = \begin{cases} \mathbf{x}_i \in N_b(\mathbf{x}_j) \text{ yada } \mathbf{x}_j \in N_b(\mathbf{x}_i) \text{ ise, } 1 \\ \text{yoksa, } 0 \end{cases} \quad (1.43)$$

Veri örneklerinin birbirleriyle olan ilişkisini gösteren ağırlık değerleri boyutu $n \times n$ ve simetrik olan W_w ve W_b ağırlık matrisleri ile ifade edilebilir. Şekil 1.2(a)'da G_w ve G_b grafları için bir örnek ve Şekil 1.2(b)'de bu veri kümesi için farklı veri sınıfları arasındaki marjın gösterilmiştir.



Şekil 1.2. a) G_w ve G_b komşuluk grafları b) farklı sınıf örnekleri arasındaki marjin [29].

YDAA yönteminde veri örnekleri öyle bir \mathbf{p} vektör üzerine izdüşüm yapılmalı ki birbiriyle aynı sınıf etiketine sahip olan komşu örnekler bu vektör üzerinde birbirlerine yakın ve birbiriyle farklı sınıf etiketine sahip olan komşu örnekler ise bu vektör üzerinde birbirlerine uzak olsun. Bunun sağlanması için denklem (1.44)'ün minimize edilmesi ve denklem (1.45)'in maksimize edilmesi gereken iki amaç fonksiyonu oluşturulmuştur. Bu amaç fonksiyonunda $y_i = \mathbf{p}^T \mathbf{x}_i$ olur, $W_{w,ij}$ ve $W_{b,ij}$ ifadesi \mathbf{x}_i ve \mathbf{x}_j örneklerinin ağırlık değerleridir. Eğer aynı sınıfa ait komşu örnekler izdüşüm yapıldığında birbirlerine uzak düşerlerse J_1 amaç fonksiyonu büyük bir değere sahip olur. Dolayısıyla aynı sınıfa ait komşu örnekleri birbirlerine yakın yerlere izdüşüm yapmak için J_1 amaç fonksiyonu minimize edilmelidir. Eğer farklı sınıfa ait komşu örnekler izdüşüm yapıldığında birbirlerine yakın düşerlerse J_2 amaç fonksiyonu küçük bir değere sahip olur. Dolayısıyla farklı sınıfa ait komşu örnekleri birbirlerine uzak yerlere izdüşüm yapmak için J_2 amaç fonksiyonu maksimize edilmelidir.

$$J_1(\mathbf{p}) = \frac{1}{2} \sum_{ij} (y_i - y_j)^2 W_{w,ij} = \frac{1}{2} \sum_{ij} (\mathbf{p}^T \mathbf{x}_i - \mathbf{p}^T \mathbf{x}_j)^2 W_{w,ij} \quad (1.44)$$

$$J_2(\mathbf{p}) = \frac{1}{2} \sum_{ij} (y_i - y_j)^2 W_{b,ij} = \frac{1}{2} \sum_{ij} (\mathbf{p}^T \mathbf{x}_i - \mathbf{p}^T \mathbf{x}_j)^2 W_{b,ij} \quad (1.45)$$

Eğer amaç fonksiyonları düzenlenirse denklem (1.46) ve (1.47) elde edilir.

$$\begin{aligned}
J_1(\mathbf{p}) &= \sum_i \mathbf{p}^T \mathbf{x}_i D_{w,ii} \mathbf{x}_i^T \mathbf{p} - \sum_{ij} \mathbf{p}^T \mathbf{x}_i W_{w,ij} \mathbf{x}_j^T \mathbf{p} \\
&= \mathbf{p}^T X D_w X^T \mathbf{p} - \mathbf{p}^T X W_w X^T \mathbf{p}
\end{aligned} \tag{1.46}$$

$$\begin{aligned}
J_2(\mathbf{p}) &= \sum_i \mathbf{p}^T \mathbf{x}_i D_{b,ii} \mathbf{x}_i^T \mathbf{p} - \sum_{ij} \mathbf{p}^T \mathbf{x}_i W_{b,ij} \mathbf{x}_j^T \mathbf{p} \\
&= \mathbf{p}^T X (D_b - W_b) X^T \mathbf{p} = \mathbf{p}^T X L_b X^T \mathbf{p}
\end{aligned} \tag{1.47}$$

Denklem (1.46)'da D_w matrisi diyagonal bir matristir ve köşegende yer alan elemanları denklem (1.48) ile elde edilir. Denklem (1.47)'de D_b matrisi diyagonal bir matristir ve köşegende yer alan elemanları denklem (1.49) ile elde edilir. L_b matrisi Laplacian matrisi olarak adlandırılır ve $(D_b - W_b)$ ile bulunur.

$$D_{w,ii} = \sum_j W_{w,ij} \tag{1.48}$$

$$D_{b,ii} = \sum_j W_{b,ij} \tag{1.49}$$

D_w matrisi veri örnekleri \mathbf{x}_i üzerinde doğal bir ölçü verir. $D_{w,ii}$ değeri ne kadar büyükse, \mathbf{x}_i veri örneğinin önemi o kadar büyük olur. Dolayısıyla J_1 ve J_2 amaç fonksiyonları için denklem (1.50)'deki kısıtlama oluşturulmuştur [25].

$$Y^T D Y = 1 \Rightarrow \mathbf{p}^T X D X^T \mathbf{p} = 1 \tag{1.50}$$

Bu kısıtlama kullanılarak, denklem (1.51)'de gösterildiği gibi J_1 amaç fonksiyonu tekrar düzenlenir. Denklem (1.51)'deki J_1 amaç fonksiyonu minimize etmek yerine, denklem (1.52)'deki gibi düzenlenip bu amaç fonksiyonu maksimize edilirse aynı çözüm elde edilir.

$$J_1(\mathbf{p}) = 1 - \mathbf{p}^T X W_w X^T \mathbf{p} \tag{1.51}$$

$$J_1(\mathbf{p}) = \mathbf{p}^T X W_w X^T \mathbf{p} \tag{1.52}$$

Denklem (1.47) ve (1.52)'deki maksimize edilmesi gereken iki amaç fonksiyonu birleştirilirse, denklem (1.53)'te yer alan J amaç fonksiyonu elde edilir. Aşağıdaki denklemde α , 0 ile 1 arasında uygun bir sabit olarak seçilir. Bu amaç fonksiyonu denklem (1.50)'de yer alan kısıtlama altında çözümlenmelidir.

$$J(\mathbf{p}) = \mathbf{p}^T X (\alpha L_b + (1 - \alpha) W_w) X^T \mathbf{p} \quad (1.53)$$

Bu optimizasyon problemi Lagrange çarpanları tekniği ile çözülmek istenirse, ilk önce denklem (1.54)'te yer alan ψ Lagrange fonksiyonu yazılır. Denklem (1.54)'ün \mathbf{p} vektörüne göre türevi alınıp 0'a eşitlenirse, bu sonucu sağlayan \mathbf{p} vektörü denklem (1.53)'te yer alan J amaç fonksiyonunu maksimum yapar. Bu türevin sonucu denklem (1.55)'te gösterilmiştir.

$$\psi = \mathbf{p}^T X (\alpha L_b + (1 - \alpha) W_w) X^T \mathbf{p} - \lambda (\mathbf{p}^T X D_w X^T \mathbf{p} - 1) \quad (1.54)$$

$$\begin{aligned} \frac{\partial \psi}{\partial \mathbf{p}} &= 2X (\alpha L_b + (1 - \alpha) W_w) X^T \mathbf{p} - 2X D_w X^T \mathbf{p} = 0 \\ \Rightarrow X (\alpha L_b + (1 - \alpha) W_w) X^T \mathbf{p} &= \lambda X D_w X^T \mathbf{p} \end{aligned} \quad (1.55)$$

Denklem (1.55)'te sonucun 0 olması için, λ skalar değeri $X (\alpha L_b + (1 - \alpha) W_w) X^T$ ve $X D_w X^T$ matrislerinin genelleştirilmiş özdeğeri ve \mathbf{p} vektörü de bu matrislerin genelleştirilmiş özvektörü olmalıdır.

Denklem (1.53)'teki J amaç fonksiyonunun maksimum olması için, $X (\alpha L_b + (1 - \alpha) W_w) X^T$ ve $X D_w X^T$ matrislerinin değer olarak en büyük genelleştirilmiş özdeğeri λ_1 ve bu genelleştirilmiş özdeğere ait genelleştirilmiş özvektör \mathbf{u}_1 izdüşüm vektörü olarak kullanılmalıdır.

Birden fazla boyut için $X (\alpha L_b + (1 - \alpha) W_w) X^T$ ve $X D_w X^T$ matrislerinin en büyük değere sahip r tane genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ seçilir ve P dönüşüm matrisi bu özdeğerlere karşılık gelen r tane genelleştirilmiş özvektörler $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ kullanılarak oluşturulur. Denklem (1.5) ve (1.6) kullanılarak boyutu azaltılmış veriler elde edilir.

YDAA yönteminin adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin G_w ve G_b komşuluk graflarının ve bunlara ait W_w ve W_b ağırlık matrislerinin bulunması.
2. D_w diyagonal ve L_b Laplacian matrislerinin bulunması.
3. $X(\alpha L_b + (1-\alpha)W_w)X^T$ ve XD_wX^T matrislerinin r tane en büyük genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ ve bu özdeğerlere ait özvektörlerinin $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ bulunması.
4. P dönüşüm matrisinin oluşturulması.

Çıkış: $d \times r$ boyutlu P dönüşüm matrisi.

1.3.5. Komşuluk Koruyan Gömme

Komşuluk Koruyan Gömme (KKG) yöntemi yüksek boyutlu veri analizinde son zamanlarda karşılaşılan bir boyut indirgeme yöntemi olup, ilk olarak 2005 yılında Xiaofei He ve arkadaşları tarafından çalışılmıştır [30]. KKG yöntemi bundan sonra hiperspektral görüntüler [31], belge sınıflandırma ve erişim [32, 33], spam filtreleme [34] ve doku sınıflandırma [35] gibi alanlarda kullanılmıştır.

KKG yöntemi verinin sınıf bilgisini boyut indirgeme işleminde göz önünde bulundurmaz. Bu sebeple gözetimsiz öğrenme yöntemlerine girmektedir. KKG yöntemi verinin yerel komşuluk yapısını koruyan bir amaç fonksiyonunu optimize ederek P dönüşüm matrisini elde eder.

KKG yönteminin temelini veriden elde edilen G komşuluk grafi ve W ağırlık matrisleri oluşturmaktadır. X giriş verisinin d boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ olduğu varsayılır ise KKG yönteminde, YKİ yöntemine benzer olarak veri örneklerinin birbirleriyle ilişkisini ortaya koyan ve n tane düğümden oluşan G komşuluk grafi oluşturulmalıdır. Eğer \mathbf{x}_i ve \mathbf{x}_j veri örnekleri birbirlerine yakınlarsa, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur. Örneklerin yakınlık ölçüsü için iki yöntem kullanılır [30]:

1. ε -komşuluğu: Eğer $\|\mathbf{x}_i - \mathbf{x}_j\|^2 < \varepsilon$ ise, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur.

2. k -en yakın komşular: Eğer \mathbf{x}_i örneği \mathbf{x}_j örneğinin en yakın k komşusu arasında ise, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur.

G komşuluk grafında düğümleri birleştiren kenarların her biri için bir W_{ij} ağırlık değeri olmalıdır. Bu değer birbirleri arasında kenar olmayan düğümlerin kenarları için 0 alınır. KKG yönteminde her bir veri örneği \mathbf{x}_i , m tane komşu veri örneğinin doğrusal kombinasyonu ile elde edilmeye çalışılır. Burada önemli olan nokta m tane komşu veri örneği için bulunan katsayıların toplamı 1 olmalıdır. Bulunan bu katsayılar G komşuluk grafında kenarların ağırlık değerleri olur. Kenarı olan i ve j düğümleri için W_{ij} ağırlık değerleri, denklem (1.57)'de yer alan kısıtlama altında denklem (1.56)'daki amaç fonksiyonu minimize edilerek bulunur [30].

$$J_w(W) = \sum_i \left\| \mathbf{x}_i - \sum_j W_{ij} \mathbf{x}_j \right\|^2 \quad (1.56)$$

$$\sum_j W_{ij} = 1, \quad j = 1, \dots, m \quad (1.57)$$

Veri örneklerinin birbirleriyle olan ilişkisini gösteren W_{ij} ağırlık değerleri bir W ağırlık matrisi ile ifade edilebilir. Bu matrisin boyutu $n \times n$ olur.

KKG yönteminde veri örnekleri öyle bir \mathbf{p} vektör üzerine izdüşüm yapılmalı ki X giriş verisi bu vektör üzerine izdüşüm yapıldığında her bir veri örneği, denklem (1.56) ile bulunan ağırlık değerleri ve m tane komşu olan örnek ile ifade edilebilmelidir. Bunun sağlanması için denklem (1.58)'deki minimize edilmesi gereken amaç fonksiyonu oluşturulmuştur. Bu amaç fonksiyonunda $y_i = \mathbf{p}^T \mathbf{x}_i$ ve W_{ij} ifadesi \mathbf{x}_i ve \mathbf{x}_j örneklerinin ağırlık değeridir.

$$J(\mathbf{p}) = \sum_{ij} \left(y_i - \sum_j W_{ij} y_j \right)^2 \quad (1.58)$$

Eğer amaç fonksiyonundaki parantez içindeki ifade denklem (1.59)'da gösterilen şekilde tanımlanırsa ve $Y = \mathbf{p}^T X$ olduğundan, Z denklem (1.60)'ta gösterildiği gibi elde edilir.

$$z_i = y_i - \sum_j W_{ij} y_j \quad (1.59)$$

$$Z = Y - WY = (I - W)Y \quad (1.60)$$

Amaç fonksiyonu denklem (1.60) kullanılarak düzenlenirse denklem (1.61) elde edilir. Bu denklemde M matrisi $(I - W)^T (I - W)$ ifadesine eşittir. Bu matris simetrik olur.

$$\begin{aligned} J(\mathbf{p}) &= \sum_{ij} \left(y_i - \sum_j W_{ij} y_j \right)^2 \\ &= \sum_{ij} (z_i)^2 \\ &= \mathbf{Z}^T \mathbf{Z} \\ &= \mathbf{Y}^T (\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W}) \mathbf{Y} \\ &= \mathbf{p}^T \mathbf{X} (\mathbf{I} - \mathbf{W})^T (\mathbf{I} - \mathbf{W}) \mathbf{X}^T \mathbf{p} \\ &= \mathbf{p}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{p} \end{aligned} \quad (1.61)$$

Yapılan izdüşümde rastgele ölçekleme faktörünü kaldırmak için, denklem (1.62)'deki kısıtlama kullanılır [30].

$$\mathbf{Y}^T \mathbf{Y} = 1 \Rightarrow \mathbf{p}^T \mathbf{X} \mathbf{X}^T \mathbf{p} = 1 \quad (1.62)$$

Bu optimizasyon problemi Lagrange çarpanları tekniği ile çözülmek istenirse, ilk önce denklem (1.63)'te yer alan ψ Lagrange fonksiyonu yazılır. Denklem (1.63)'ün \mathbf{p} vektörüne göre türevi alınıp 0'a eşitlenirse, bu sonucu sağlayan \mathbf{p} vektörü denklem (1.61)'de yer alan J amaç fonksiyonunu minimum yapar. Bu türevin sonucu denklem (1.64)'te gösterilmiştir.

$$\psi = \mathbf{p}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{p} - \lambda (\mathbf{p}^T \mathbf{X} \mathbf{X}^T \mathbf{p} - 1) \quad (1.63)$$

$$\begin{aligned}\frac{\partial \psi}{\partial \mathbf{p}} &= 2\mathbf{XMX}^T \mathbf{p} - 2\lambda \mathbf{XX}^T \mathbf{p} = 0 \\ \Rightarrow \mathbf{XMX}^T \mathbf{p} &= \lambda \mathbf{XX}^T \mathbf{p}\end{aligned}\tag{1.64}$$

Denklem (1.64)'te sonucun 0 olması için, λ skalar değeri \mathbf{XMX}^T ve \mathbf{XX}^T matrislerinin genelleştirilmiş özdeğeri ve \mathbf{p} vektörü de bu matrislerin genelleştirilmiş özvektörü olmalıdır.

Denklem (1.61)'deki amaç fonksiyonunun $\mathbf{p}^T \mathbf{XX}^T \mathbf{p} = 1$ kısıtlaması altında minimum olması için, \mathbf{XMX}^T ve \mathbf{XX}^T matrislerinin değer olarak en küçük genelleştirilmiş özdeğeri λ_1 ve bu genelleştirilmiş özdeğere ait genelleştirilmiş özvektör \mathbf{u}_1 izdüşüm vektörü olarak kullanılmalıdır.

Birden fazla boyut için \mathbf{XMX}^T ve \mathbf{XX}^T matrislerinin en küçük değere sahip r tane genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ seçilir ve P dönüşüm matrisi bu özdeğerlere karşılık gelen r tane genelleştirilmiş özvektör $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ kullanılarak oluşturulur. Denklem (1.5) ve (1.6) kullanılarak düşük boyutlu veriler elde edilir.

KKG yönteminin adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin G komşuluk grafının ve W ağırlık matrisinin bulunması.
2. M matrisinin bulunması.
3. \mathbf{XMX}^T ve \mathbf{XX}^T matrislerinin r tane en küçük genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ ve bu özdeğerlere ait özvektörlerinin $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ bulunması.
4. P dönüşüm matrisinin oluşturulması.

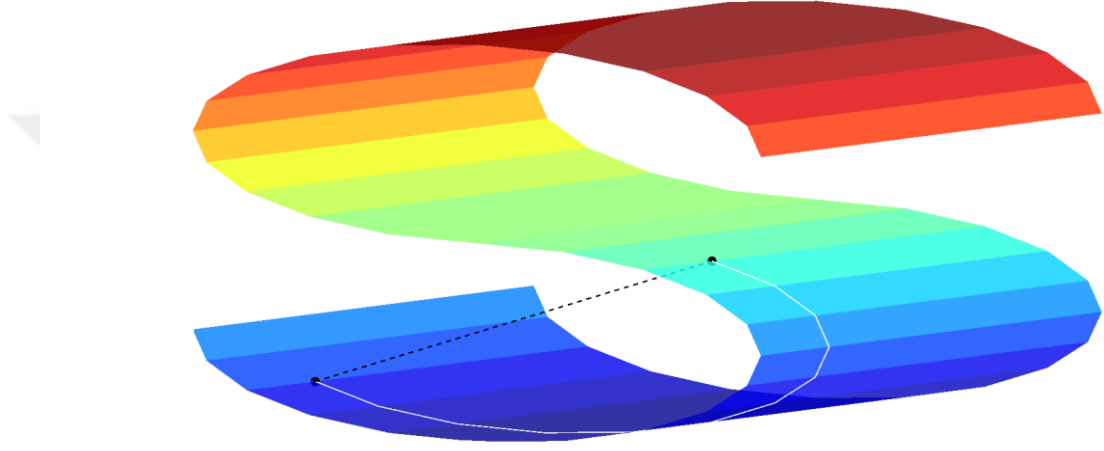
Çıkış: $d \times r$ boyutlu P dönüşüm matrisi.

1.3.6. İzometrik İzdüşüm

İzometrik İzdüşüm (İZİ) yöntemi ilk olarak 2007 yılında Deng Cai ve diğerleri tarafından ortaya konulmuştur [36]. İZİ yöntemi yüz tanıma [36], avuç içi izi tanıma [37] ve avuç içi damar tanıma [38] gibi alanlarda kullanılmıştır.

İZİ yöntemi yüksek boyutlu veri analizinde son zamanlarda karşılaşılan bir boyut indirgeme yöntemidir. İZİ yöntemi verinin sınıf bilgisini boyut indirgeme işleminde göz

önünde bulundurmaz. Bu sebeple gözetimsiz öğrenme yöntemlerine girmektedir. İZİ yöntemi verinin örnekleri arasındaki jeodezik uzaklıkları koruyan bir amaç fonksiyonunu optimize ederek P dönüşüm matrisini elde eder. Jeodezik uzaklık iki veri örneğini birbirine bağlayan en kısa eğrinin uzunluğu olarak tanımlanır. Şekil 1.3'te karmaşık bir verinin iki örneğinin arasındaki öklit uzaklığı siyah kesikli çizgi ile, jeodezik uzaklığı ise beyaz düz çizgi ile gösterilmiştir.



Şekil 1.3. Jeodezik uzaklık örneği

İZİ yöntemi de diğer yöntemlerin bazılarında olduğu gibi veriden elde edilen G komşuluk grafına ihtiyaç duymaktadır. X giriş verisinin d boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ olduğu varsayılırsa, İZİ yönteminde, YKİ yöntemine benzer olarak veri örneklerinin birbirleriyle ilişkisini ortaya koyan ve n tane düğümden oluşan G komşuluk grafi oluşturulmalıdır. Eğer \mathbf{x}_i ve \mathbf{x}_j veri örnekleri birbirlerine yakınlarsa, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur. Örneklerin yakınlık ölçüsü için iki yöntem kullanılır [36]:

1. ε -komşuluğu: Eğer $\|\mathbf{x}_i - \mathbf{x}_j\| < \varepsilon$ ise, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur.
2. k -en yakın komşular: Eğer \mathbf{x}_i örneği \mathbf{x}_j örneğinin en yakın k komşusu arasında ise, komşuluk grafında i ve j düğümleri arasına bir kenar koyulur.

İZİ yönteminde oluşturulan G komşuluk grafi kullanılarak her bir veri örnek çifti $(\mathbf{x}_i, \mathbf{x}_j)$ için jeodezik uzaklık $d_J(\mathbf{x}_i, \mathbf{x}_j)$ hesaplanmalıdır. Veri örneklerinin jeodezik uzaklıkları $d_J(\mathbf{x}_i, \mathbf{x}_j)$, bir D_J jeodezik uzaklık matrisi ile ifade edilebilir. Bu matrisin boyutu $n \times n$ olur. Bu matrisi bulmak için Floyd-Warshall algoritması kullanılır. Bu algoritmanın adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin G komşuluk grafinin bulunması.
2. Her bir veri çifti $(\mathbf{x}_i, \mathbf{x}_j)$ için, G grafında,

birbirine bağlı ise $d_J(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i - \mathbf{x}_j\|$

değil ise $d_J(\mathbf{x}_i, \mathbf{x}_j) = \infty$ olarak belirleme.

3. Bu adımın p değişkeni için 1'den n 'e kadar tekrarlanması.

Her bir veri çifti $(\mathbf{x}_i, \mathbf{x}_j)$ için,

$$d_J(\mathbf{x}_i, \mathbf{x}_j) = \min \{d_J(\mathbf{x}_i, \mathbf{x}_j), d_J(\mathbf{x}_i, \mathbf{x}_p) + d_J(\mathbf{x}_p, \mathbf{x}_j)\}$$

4. D_J jeodezik uzaklık matrisinin oluşturulması.

Çıkış: $n \times n$ boyutlu D_J jeodezik uzaklık matrisi.

İZİ yönteminde X veri örnekleri öyle bir \mathbf{p} vektör üzerine izdüşüm yapılmalı ki X veri örnekleri arasındaki jeodezik uzaklıklar bu vektör üzerinde Y veri örneklerinin öklit uzaklıkları ile korunmalıdırlar. Bunun sağlanması için $y_i = \mathbf{p}^T \mathbf{x}_i$ olmak üzere, denklem (1.65)'teki amaç fonksiyonu minimize edilmelidir.

$$\begin{aligned} J(\mathbf{p}) &= \sum_{ij} \left(d_J(\mathbf{x}_i, \mathbf{x}_j) - d(y_i, y_j) \right)^2 \\ &= \sum_{ij} \left(d_J(\mathbf{x}_i, \mathbf{x}_j) - d(\mathbf{p}^T \mathbf{x}_i, \mathbf{p}^T \mathbf{x}_j) \right)^2 \end{aligned} \quad (1.65)$$

Amaç fonksiyonunu düzenlemek için, elemanları $S_{ij} = (d_J(\mathbf{x}_i, \mathbf{x}_j))^2$ olan bir S matrisi ve merkezleme matrisi $H = I - \frac{1}{n} \mathbf{e} \mathbf{e}^T$ tanımlanır. I birim matris ve \mathbf{e} her elemanı 1 olan bir vektördür. S matrisi ve H merkezleme matrisi kullanılarak, bir D uzaklık matrisi için $\tau(D)$

iç çarpım matrisi denklem (1.66)'da gösterildiği gibi tanımlanır. Denklem (1.66) ile elde edilen iç çarpım matrisinin elemanlarının $\tau(D)_{ij}$ denklem (1.67)'deki gibi olduğu gösterilebilir [41].

$$\tau(D) = -\frac{1}{2}(HSH) \quad (1.66)$$

$$\tau(D)_{ij} = (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}}) \quad (1.67)$$

$\tau(D_J)$ ve $\tau(D_Y)$ matrisleri, X verisinin D_J jeodezik uzaklık matrisinin ve Y verisinin D_Y öklit uzaklık matrisinin iç çarpım matrisleri olmak üzere, amaç fonksiyonu denklem (1.68)'de gösterildiği gibi ifade edilebilir. $\|A\|$ ifadesi matris normu olarak adlandırılır ve

$\|A\| = \sqrt{\sum_{i,j} A_{i,j}^2}$ olarak bulunur.

$$J(\mathbf{p}) = \|\tau(D_J) - \tau(D_Y)\|^2 \quad (1.68)$$

$Y = [y_1, \dots, y_n] = \mathbf{p}^T X$ olduğundan, Y veri örneklerinin D_Y iç çarpım matrisi denklem (1.69)'da gösterildiği gibi ifade edilebilir. Dolayısıyla denklem (1.69)'u kullanarak amaç fonksiyonu denklem (1.70)'te gösterildiği gibi ifade edilir.

$$\tau(D_Y) = Y^T Y = X^T \mathbf{p} \mathbf{p}^T X \quad (1.69)$$

$$J(\mathbf{p}) = \|\tau(D_J) - X^T \mathbf{p} \mathbf{p}^T X\|^2 \quad (1.70)$$

Denklem (1.70)'teki amaç fonksiyonu $tr(A) = tr(A^T)$ eşitliği göz önüne alınarak düzenlenirse denklem (1.71) elde edilir.

$$\begin{aligned}
J(\mathbf{p}) &= \|\tau(D_j) - X^T \mathbf{p} \mathbf{p}^T X\|^2 \\
&= \text{tr} \left((\tau(D_j) - X^T \mathbf{p} \mathbf{p}^T X) (\tau(D_j) - X^T \mathbf{p} \mathbf{p}^T X)^T \right) \\
&= \text{tr} \begin{pmatrix} \tau(D_j) \tau(D_j)^T - X^T \mathbf{p} \mathbf{p}^T X \tau(D_j)^T \\ -\tau(D_j) X^T \mathbf{p} \mathbf{p}^T X + X^T \mathbf{p} \mathbf{p}^T X X^T \mathbf{p} \mathbf{p}^T X \end{pmatrix}
\end{aligned} \tag{1.71}$$

Yapılan izdüşümde rastgele ölçekleme faktörünü kaldırmak için, denklem (1.72)'deki kısıtlama kullanılır. Dolayısıyla denklem (1.73)'teki sonuç elde edilir.

$$Y^T Y = 1 \Rightarrow \mathbf{p}^T X X^T \mathbf{p} = 1 \tag{1.72}$$

$$\text{tr} \left(X^T \mathbf{p} \mathbf{p}^T X X^T \mathbf{p} \mathbf{p}^T X \right) = \text{tr} \left(\mathbf{p}^T X X^T \mathbf{p} \mathbf{p}^T X X^T \mathbf{p} \right) = 1 \tag{1.73}$$

Denklem (1.73)'teki sonuç ile denklem (1.71)'deki amaç fonksiyonu düzenlenirse denklem (1.74) elde edilir.

$$J(\mathbf{p}) = \text{tr} \left(\tau(D_j) \tau(D_j)^T \right) - 2 \text{tr} \left(\mathbf{p}^T X \tau(D_j) X^T \mathbf{p} \right) + 1 \tag{1.74}$$

Denklem (1.74)'teki amaç fonksiyonunun ilk terimi ve üçüncü terimi sabitlerdir ve optimizasyon işleminde etkileri yoktur. Dolayısıyla $\mathbf{p}^T X X^T \mathbf{p} = 1$ kısıtlamasıyla denklem (1.74)'teki amaç fonksiyonunu minimize etmek yerine $\mathbf{p}^T X X^T \mathbf{p} = 1$ kısıtlamasıyla denklem (1.75)'teki amaç fonksiyonunu maksimize etmek ile aynı çözüm elde edilir.

$$\begin{aligned}
J(\mathbf{p}) &= \text{tr} \left(\mathbf{p}^T X \tau(D_j) X^T \mathbf{p} \right) \\
&= \mathbf{p}^T X \tau(D_j) X^T \mathbf{p}
\end{aligned} \tag{1.75}$$

Bu optimizasyon problemi Lagrange çarpanları tekniği ile çözülmek istenirse, ilk önce denklem (1.76)'da yer alan ψ Lagrange fonksiyonu yazılır. Denklem (1.76)'nın \mathbf{p} vektörüne göre türevi alınıp 0'a eşitlenirse, bu sonucu sağlayan \mathbf{p} vektörü denklem (1.75)'te yer alan J amaç fonksiyonunu maksimum yapar. Bu türevin sonucu denklem (1.77)'de gösterilmiştir.

$$\psi = \mathbf{p}^T X \tau(D_J) X^T \mathbf{p} - \lambda (\mathbf{p}^T X X^T \mathbf{p} - 1) \quad (1.76)$$

$$\begin{aligned} \frac{\partial \psi}{\partial \mathbf{p}} &= 2X \tau(D_J) X^T \mathbf{p} - 2\lambda X X^T \mathbf{p} = 0 \\ \Rightarrow X \tau(D_J) X^T \mathbf{p} &= \lambda X X^T \mathbf{p} \end{aligned} \quad (1.77)$$

Denklem (1.77)'de sonucun 0 olması için, λ skalar değeri $X \tau(D_J) X^T$ ve $X X^T$ matrislerinin genelleştirilmiş özdeğeri ve \mathbf{p} vektörü de bu matrislerin genelleştirilmiş özvektörü olmalıdır. Denklem (1.75)'teki J amaç fonksiyonunun $\mathbf{p}^T X X^T \mathbf{p} = 1$ kısıtlaması altında maksimum olması için, $X \tau(D_J) X^T$ ve $X X^T$ matrislerinin değer olarak en büyük genelleştirilmiş özdeğeri λ_1 ve bu genelleştirilmiş özdeğere ait genelleştirilmiş özvektör \mathbf{u}_1 izdüşüm vektörü olarak kullanılmalıdır.

Birden fazla boyut için $X \tau(D_J) X^T$ ve $X X^T$ matrislerinin en büyük değere sahip r tane genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ seçilir ve P dönüşüm matrisi bu özdeğerlere karşılık gelen r tane genelleştirilmiş özvektör $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ kullanılarak oluşturulur. Denklem (1.5) ve (1.6) kullanılarak düşük boyutlu veriler elde edilir.

İZİ yönteminin adımları aşağıda gösterildiği gibidir.

Giriş: d boyutlu n tane veri örneğine sahip bir veri kümesi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$.

1. X veri kümesinin G komşuluk grafının bulunması.
2. D_J jeodezik uzaklık matrisi ve $\tau(D_J)$ iç çarpım matrisinin bulunması.
3. $X \tau(D_J) X^T$ ve $X X^T$ matrislerinin r tane en büyük genelleştirilmiş özdeğerleri $\lambda_1, \lambda_2, \dots, \lambda_r$ ve bu özdeğerlere ait özvektörlerinin $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ bulunması.
4. P dönüşüm matrisinin oluşturulması.

Çıkış: $d \times r$ boyutlu P dönüşüm matrisi.

1.4. Sınıflandırma Yöntemleri

Sınıflandırma, veri sınıflarını tanımlayan modelleri çıkaran bir veri analiz tekniğidir. Bu modeller sınıflandırıcı olarak adlandırılırlar ve kategorik sınıf etiketlerini tahmin ederler.

Sınıflandırma yöntemleri eğitim ve sınıflandırma adımı olarak adlandırılan iki adımdan oluşur. Eğitim adımında, eğitim kümesi kullanılarak bir model oluşturulur. Sınıflandırma adımında ise eğitim kümesinden ayrı bir test kümesi kullanılarak bu modelin test kümesi üzerindeki başarımı hakkında bilgi elde edilir.

Makine öğrenmesi, örüntü tanıma ve istatistik gibi alanlarda çalışan araştırmacılar bir çok sınıflandırma yöntemi önermiştir. Bu yöntemlerden öne çıkanlar k -en yakın komşuluk, yapay sinir ağları ve destek vektör makineleri yöntemleridir. Bundan dolayı bu çalışmada bu üç yöntem sınıflandırıcı olarak kullanılmıştır.

1.4.1. k -En Yakın Komşuluk Yöntemi

k -En Yakın Komşuluk (k -EYK) [12, 40, 41] yöntemi benzerlik ile öğrenme mantığına dayanır. İlk olarak 1950'lerde tanımlanmıştır [41]. Basitliği ve başarımı sayesinde sıkça kullanılmaktadır.

k -EYK yöntemi, bir test örneğini kendisine benzer olan eğitim örneklerini inceleyerek bu test örneğini sınıflandırır. Bunun için d boyutlu bir test örneğine \mathbf{x}_i en yakın k tane komşu eğitim örneği bulunur. Yakınlık çeşitli uzaklık ölçütleri ile bulunabilir. En yaygın olarak kullanılanı denklem (1.78)'de gösterilen öklit uzaklığıdır. En yakın k komşunun bulunmasından sonra test örneğinin \mathbf{x}_i sınıfı, k en yakın komşulardaki en yaygın sınıf olarak belirlenir. $k=1$ için, bir test örneğinin sınıf etiketi değeri en yakın eğitim kümesi örneğinin sınıf etiketi değerini alır.

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{m=1}^d (x_{i,m} - x_{j,m})^2} \quad (1.78)$$

1.4.2. Yapay Sinir Ağları

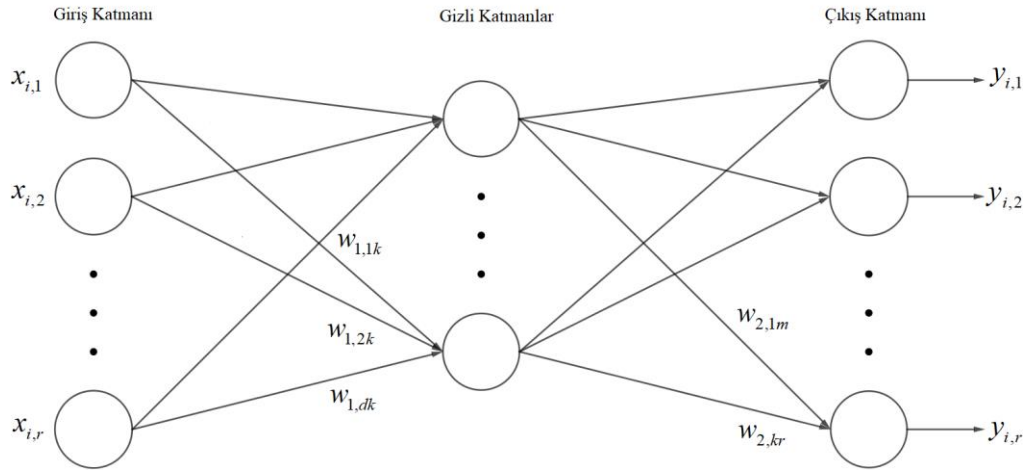
Bir Yapay Sinir Ağı (YSA) [12, 40, 41], yapay sinir hücresi olarak adlandırılan birimler topluluğudur. Bir yapay sinir ağında yer alan her bir yapay sinir hücresi bir giriş alır ve sahip oldukları bir aktivasyon fonksiyonu ve bu giriş ile bir çıkış oluştururlar. Yapay sinir ağları, yapay sinir hücrelerinin çıkışlarının diğer yapay sinir hücrelerine belirli bir

düzene göre bağlanmasıyla elde edilir. Bu bağlantıların her birinin bir ağırlık değeri olur. Bir yapay sinir ağı için eğitim adımı, yapay sinir ağı çıkışında test örneklerine ait doğru sınıf etiketlerini elde edebilmek için yapay sinir hücrelerini birbirine bağlayan bu ağırlık değerleri ayarlanarak gerçekleştirilir.

Literatürde bir çok farklı yapay sinir ağı ve yapay sinir ağı algoritması bulunmaktadır. Bunlardan en popülerleri çok katmanlı ileri beslemeli yapay sinir ağları üzerinde öğrenme gerçekleştiren geri yayılım algoritmasıdır.

1.4.2.1. Çok Katmanlı İleri Beslemeli Yapay Sinir Ağları

Çok Katmanlı İleri Beslemeli bir Yapay Sinir Ağı (ÇKİB-YSA) giriş katmanı, gizli katmanlar ve çıkış katmanı olarak adlandırılan katmanlarda yer alan yapay sinir birimlerinin sadece ileri yönlü olarak birbirine bağlanmasıyla oluşan yapay sinir ağıdır. ÇKİB-YSA için bir örnek Şekil 1.4'te gösterilmiştir. Gizli katman sayısı ve bu katmanlarda yer alan birim sayıları isteğe bağlıdır. Pratikte ise sadece bir gizli katman kullanılır.



Şekil 1.4. Çok katmanlı ileri beslemeli bir yapay sinir ağı örneği

Giriş katmanında yer alan giriş birimleri olarak adlandırılan birimler aldıkları girişleri üzerinde değişiklik yapmadan gizli katmanlara iletirler. Çıkış katmanında ve gizli katmanlarda yer alan birimler çıkış birimleri olarak adlandırılır. Bu çıkış birimleri kendinden bir önceki katmanda yer alan birimlerin çıkışlarını giriş olarak kullanır. Bir önceki katmanda

yer alan birimlerin çıkış değerleri O_i olmak üzere, denklem (1.79)'da gösterildiği gibi bir çıkış biriminde bunlar ilk olarak ağırlıklı olarak toplanır. Bu ağırlıklı toplama işlemi sonucunda net giriş I_j olarak adlandırılan değer elde edilir. w_0 değeri bias olarak adlandırılır.

$$I_j = \sum_k w_{1k} O_i + w_0 \quad (1.79)$$

Bir çıkış biriminde elde edilen net giriş değerine daha sonra aktivasyon fonksiyonu uygulanır. Literatürde farklı aktivasyon fonksiyonları olmasına rağmen genellikle denklem (1.80)'de gösterilen sigmoid fonksiyonunun kullanımı yaygındır [40].

$$O_j = \frac{1}{1 + I_j} \quad (1.80)$$

Gizli katmanlarda elde edilen bu çıkış değerleri bir sonraki katmanlara giriş değeri olarak iletilir.

1.4.2.2. Geri Yayılım Algoritması

ÇKİB-YSA üzerinde eğitim adımında geri yayılım algoritması kullanılır. Bu algoritmada, bir veri örneği için yapay sinir ağının tahmini sınıf değeri ile gerçek sınıf değerinin karşılaştırılması ve her bir veri örneğinin tekrarlamalı şekilde işlenmesi ile öğrenme gerçekleştirilir. Her bir eğitim örneği için yapay sinir ağının ağırlık değerleri tahmini değer ve gerçek sınıf değeri arasındaki hatayı azaltması için düzenlenir. Bu düzenleme çıkış katmanından geriye doğru gizli katmanlara doğru gerçekleşir.

Başlangıçta rastgele değerler olarak belirlenen ağırlıklar ile yapay sinir ağının bir eğitim örneği için denklem (1.79) ve (1.80) kullanılarak tahmini sınıf değeri elde edilir. Gerçek sınıf değeriyle olan hatayı azaltmak için çıkış katmanında bulunan sinir birimlerindeki hata denklem (1.81) ile bulunur. Bu denklemde T_j eğitim örneklerinin gerçek sınıf değeridir.

$$E_j = O_j(1 - O_j)(T_j - O_j) \quad (1.81)$$

Gizli katmanlarda bulunan sinir birimlerindeki hata denklem (1.82) ile bulunur. Bu denklemde w_{jk} , j . sinir birimi ile k . sinir birimi arasındaki bağlantının ağırlık değeridir. E_k ise k . sinir biriminde oluşan hata değeridir.

$$E_j = O_j(1 - O_j) \sum_k E_k w_{jk} \quad (1.82)$$

Bulunan hatalar ile ağırlık değerleri denklem (1.83)'te gösterildiği gibi düzenlenir. Denklem (1.83)'te yer alan α , öğrenme katsayısı olarak adlandırılır ve 0 ile 1 arasında uygun bir değer seçilir. Bias değeri için denklem (1.84) kullanılır.

$$w_{ij} = w_{ij} + \alpha E_j O_i \quad (1.83)$$

$$w_0 = w_0 + \alpha E_j \quad (1.84)$$

1.4.3. Destek Vektör Makinaları

Destek vektör makinaları (DVM) [12, 40, 41], 1992 yılında Vladimir Vapnik ve diğerleri tarafından ortaya konulmuştur. Ortaya çıkmasıyla beraber sınıflandırmadaki başarısı sayesinde bir çok alanda kullanılmaya başlamıştır.

DVM yöntemi hem doğrusal olarak ayrılabilen hem de doğrusal olarak ayrılamayan verilerde sınıflandırma gerçekleştirebilir. Doğrusal olarak ayrılabilen verilerde DVM yöntemi doğrusal en iyi ayırıcı hiperdüzlemi bulur. Doğrusal olarak ayrılamayan verilerde ise önce doğrusal olmayan bir fonksiyon ile veriyi daha yüksek boyutlu bir uzaya taşır ve burada doğrusal olarak ayrılabilen veriler gibi doğrusal en iyi ayırıcı hiperdüzlemi bulmaya çalışır.

Ayırıcı hiperdüzleme en yakın veri örneklerine destek vektörleri denir. DVM yöntemi bir verinin destek vektörleri arasındaki marjın olarak adlandırılan uzaklığı maksimize etmeyi

amaçlayan bir yöntemdir. Bir veri için elde edilen dar marjin ve geniş marjin Şekil 1.5'te gösterilmiştir.

DVM yöntemi iki sınıflı doğrusal olarak ayrılabilen, iki sınıflı doğrusal olarak ayrılamayan ve çok sınıflı verilerde incelenecektir.

1.4.3.1. İki Sınıflı Doğrusal Ayrılabilir Durum

X verisi iki sınıflı ve iki boyutlu n tane örneği $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ve sınıf etiketleri l_1, l_2, \dots, l_n olsun. Her bir etiket değeri sınıfa göre +1 veya -1 olabilir. Bu veri kümesi için ayırıcı hiperdüzlem denklem (1.85)'te gösterildiği gibi yazılabilir. Bu denklemde $\mathbf{w}^T = [w_1 \ w_2 \ \dots \ w_n]$ ağırlık vektörü ve b ise bias olarak adlandırılır.

$$\mathbf{w}^T \mathbf{x} + b = 0 \quad (1.85)$$

Bu hiperdüzlemin üzerinde kalan herhangi bir nokta denklem (1.86)'yı ve altında kalan herhangi bir nokta denklem (1.87)'yi sağlar.

$$\mathbf{w}^T \mathbf{x} + b > 0 \quad (1.86)$$

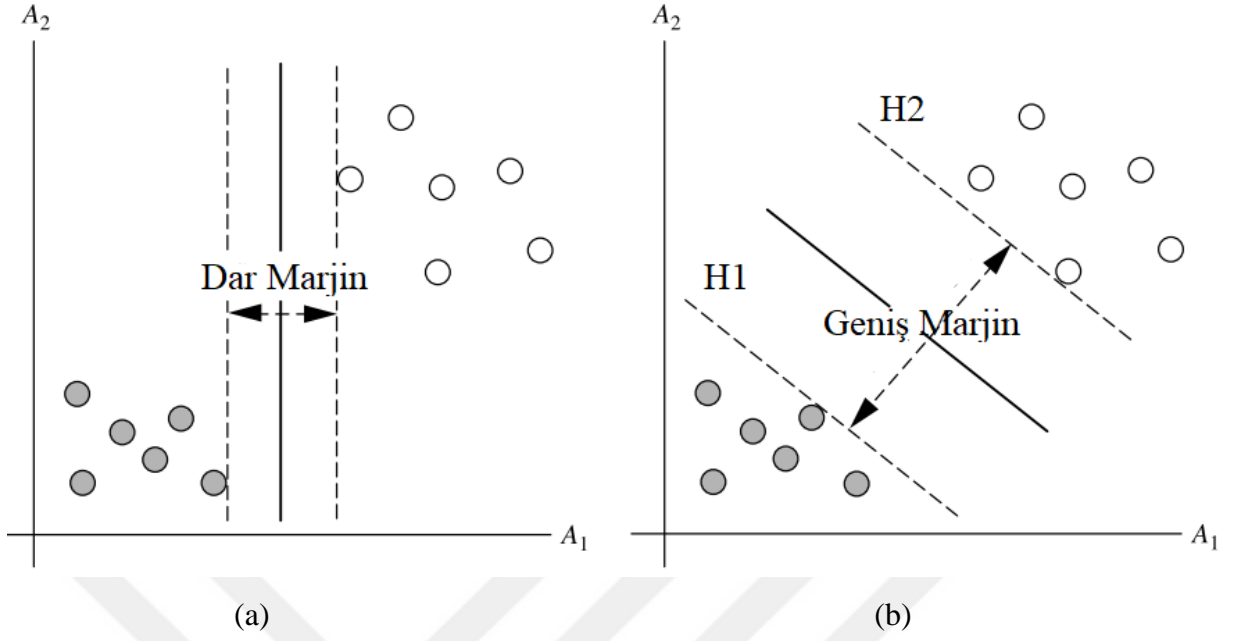
$$\mathbf{w}^T \mathbf{x} + b < 0 \quad (1.87)$$

Şekil 1.5'te görülen H1 ve H2 hiperdüzlemleri l_i sınıf etiketleri olmak üzere denklem (1.88) ve (1.89)'da gösterildiği gibi ifade edilebilir. İki eşitsizlik birleştirilirse denklem (1.90) elde edilir.

$$l_i = +1 \Rightarrow \text{H1: } \mathbf{w}^T \mathbf{x} + b \geq 1 \quad (1.88)$$

$$l_i = -1 \Rightarrow \text{H2: } \mathbf{w}^T \mathbf{x} + b \leq -1 \quad (1.89)$$

$$l_i (\mathbf{w}^T \mathbf{x} + b) \geq 1, \forall i \quad (1.90)$$



Şekil 1.5. a) Dar marjin b) Geniş marjin [40].

Ayırıcı düzlem üzerinde yer alan herhangi bir noktanın H1 ve H2 hiperdüzlemlerinden her ikisine uzaklıkları $1/\|\mathbf{w}\|$ olur. Dolayısıyla marjin uzaklığı $2/\|\mathbf{w}\|$ olur. Bu ifade amaç fonksiyonu olarak kullanılabilir. Denklem (1.90) kısıtı altında maksimize edilmesi gereken amaç fonksiyonu denklem (1.91)'de gösterildiği gibi olur.

$$J(\mathbf{w}) = \frac{2}{\|\mathbf{w}\|} \quad (1.91)$$

Lagrange çarpanları tekniği ile bu amaç fonksiyonu optimize edilirse, denklem (1.92)'de yer alan çözüm elde edilir. Bu denklemde m destek vektör sayısını, \mathbf{x}_i destek vektörlerini, l_i sınıf etiketleri, ve \mathbf{x} test veri örneğini göstermektedir. Lagrange çarpanları α_i ve b_0 parametreleri optimizasyon işlemi sonucunda bulunur.

$$d(\mathbf{x}) = \sum_{i=1}^m l_i \alpha_i \mathbf{x}_i \cdot \mathbf{x} + b_0 \quad (1.92)$$

Bir test verisi \mathbf{x} bu denklemde yerine konulduğunda çıkan sonucun pozitif veya negatif olmasına göre sınıf etiketi belirlenir.

1.4.3.2. İki Sınıflı Doğrusal Ayrılamayan Durum

İki sınıflı doğrusal ayrılamayan durumda ilk olarak veri doğrusal olmayan bir fonksiyon Φ ile daha yüksek boyutlu bir uzaya haritalanır. Bu yeni uzayda örnekler $\mathbf{z} = \Phi(\mathbf{x})$ ile bulunur. Daha sonra bu yeni yüksek boyutlu uzayda en iyi doğrusal ayırıcı hiperdüzlem aranır. Eğitim kısmında en iyi ayırıcı hiperdüzlem elde edilirken, bu yüksek boyutlu uzayda bir çok kez nokta çarpım $\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ hesabı yapmak gerekir. Nokta çarpım hesabının getirdiği yükten kurtulmak için çekirdek fonksiyonları kullanılır. Denklem (1.93)'te gösterildiği gibi çekirdek fonksiyonları nokta çarpım $\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ sonucunu orijinal veri değişkenlerini kullanarak hesaplar.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \quad (1.93)$$

Eğitim kısmında nokta çarpım $\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ yapılması gereken her yerde yüksek boyutlu uzay dönüşümü yapmadan sadece $K(\mathbf{x}_i, \mathbf{x}_j)$ ifadesini kullanarak işlem yapılır. En çok kullanılan çekirdek fonksiyonları ise denklem (1.94)'te gösterilen polinom çekirdek fonksiyonu, denklem (1.95)'te gösterilen gaussian radyal tabanlı çekirdek fonksiyonu ve denklem (1.96)'da gösterilen sigmoid çekirdek fonksiyonudur [40].

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^p \quad (1.94)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|} \quad (1.95)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\kappa \mathbf{x}_i \cdot \mathbf{x}_j - \delta) \quad (1.96)$$

1.4.3.3. Çok Sınıflı DVM

DVM yöntemi ile çok sınıflı veriler üzerinde sınıflandırma yapabilmek için çeşitli yöntemler kullanılır. Bunlardan en önemli iki tanesi bire-bir yaklaşımı ve bire-çok yaklaşımıdır.

Bire-bir yaklaşımda, bütün veri sınıflarının ikili grupları için DVM yönteminde ayrıcı hiperdüzlemler bulunur. Bu da m sınıflı bir veri kümesi için $m(m-1)/2$ tane ayrıcı hiperdüzlem bulunmasını gerektirir. Bir test veri örneği için bulunan tüm ayrıcı hiperdüzlemler ile sınıflandırma yapılır ve en yüksek oy oranına sahip sınıf etiketi o test verisi için seçilir.

Bire-çok yaklaşımında ise m sınıflı bir veri kümesi için m tane ayrıcı hiperdüzlem bulunur. j . hiperdüzlem j . sınıfın eğitim örneklerini pozitif sınıf olarak ve diğer sınıfların eğitim örneklerini negatif sınıf olarak kullanarak bulur. Dolayısıyla bu hiperdüzlem j . sınıfın test örnekleri için pozitif sonuç diğer sınıflar için negatif sonuç elde etmek için eğitilmiş olur. Bir test veri örneğini sınıflandırmak için j . hiperdüzlem pozitif sonuç verirse j . sınıfının oy değeri bir artırılır. Eğer negatif sonuç verirse diğer bütün sınıfların oy değerleri bir artırılır. En sonunda en yüksek oy sayısına sahip sınıf etiketi bu test verisi için seçilir.

1.5. Karşılaştırma Ölçütleri

1.5.1. Doğruluk Değerlendirmesi

Doğrusal boyut indirgeme yöntemlerinin sınıflandırma üzerindeki performansları incelenirken öncelikle sınıflandırma doğrulukları üzerindeki etkileri karşılaştırılabilir. Bir sınıflandırıcının doğruluğu, test verisi için bu sınıflandırıcının bulduğu tahmini sınıf etiketleri ile gerçek sınıf etiketleri arasındaki uygunluğu gösterir. Doğruluk bulunurken hata veya karışıklık matrisi olarak adlandırılan H matrisi kullanılır. Bu matris c sınıflı bir veri için $c \times c$ boyutunda olur.

$$H = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1c} \\ n_{21} & n_{22} & \cdots & n_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ n_{c1} & n_{c2} & \cdots & n_{cc} \end{bmatrix} \quad (1.96)$$

Hata matrisinde n_{ij} elemanı, gerçek sınıfı j olan ve sınıflandırıcı tarafından sınıf i olarak bulunan test verisi örneklerinin sayısını göstermektedir. Dolayısıyla köşegen üzerinde bulunan elemanlar n_{ii} doğru sınıflandırılan test veri örneklerinin sayısını gösterir. Köşegen dışındaki elemanlar yanlış sınıflandırılan test verisi örneklerinin sayısını gösterir.

H matrisi ile bir sınıflandırıcının doğruluğu n_t test veri örneği sayısını göstermek üzere denklem (1.97)'de gösterildiği gibi elde edilir.

$$d = \frac{\sum_{i=1}^c n_{ii}}{n_t} \quad (1.97)$$

1.5.2. Sınıflandırma Süreleri

Doğrusal boyut indirgeme yöntemleri bir veri üzerinde uygulandığında bu verinin sınıflandırılma işleminin süresine etkisi karşılaştırılabilir.

2. YAPILAN ÇALIŞMALAR, BULGULAR VE İRDELEME

2.1. Giriş

Bu bölümde doğrusal boyut indirgeme yöntemlerinin yüz tanıma ve rakam tanıma üzerindeki sınıflandırma performansları incelenmiştir. Yüz tanıma için Olivetti Research Laboratory (ORL) [42] olarak adlandırılan ve rakam tanıma için U.S. Postal Service (USPS) [41, 43] olarak adlandırılan veri kümeleri kullanılmıştır.

2.2. ORL Veri Kümesi

ORL veri kümesi 1992 ve 1994 yılları arasında Cambridge Üniversitesi'nde oluşturulmuştur. Veri kümesi 40 kişiden alınmış olup her bir kişi için 10 tane görüntü vardır. Bazı kişiler için görüntüler farklı zamanlarda, farklı ışıklandırma, farklı yüz ifadeleri (açık/kapalı göz, gülen/gülmeyen) ve farklı yüz detayları (gözlüklü/gözlüksüz) ile elde edilmiştir. Şekil 2.1'de 8 kişi için görüntü örnekleri verilmiştir.

Bu veri kümesinde her bir görüntü 92x112 piksele sahiptir. Bu görüntüler, kişilerin iki gözü aynı pozisyonda olacak şekilde normalize edilip, daha sonra yüz kısmı kesip çıkartılıp 32x32 piksele sahip görüntüler elde edilmiştir [30]. Bu şekilde elde edilen görüntüler Şekil 2.2'de verilmiştir.

ORL veri kümesinden elde edilen bu görüntülerde yer alan her bir piksel sınıflandırma için bir öznitelik olarak kullanılmıştır. Görüntüler 32x32 piksele sahip olduğu için bu veri kümesi için 1024 tane öznitelik vardır.

2.3. USPS Veri Kümesi

USPS veri kümesi, U.S. posta servisine ait zarflar üzerindeki elle yazılmış posta kodlarından elde edilmiş bir çok farklı insan tarafından farklı stillerle yazılmış rakamlardan oluşturulmuş bir veri kümesidir. Her bir görüntü 16x16 pikselden oluşmaktadır. Eğitim kümesi ve test kümesiyle sırayla 7291 ve 2007 örneğe sahiptir. Veri örneklerinin sınıflara

göre dağılımı ve oranları Tablo 2.1 ve Tablo 2.2’de verilmiştir. Şekil 2.3 de ise veri kümesinin bazı örnekleri verilmektedir.



Şekil 2.1. ORL veri kümesi için bazı örnekler [42].



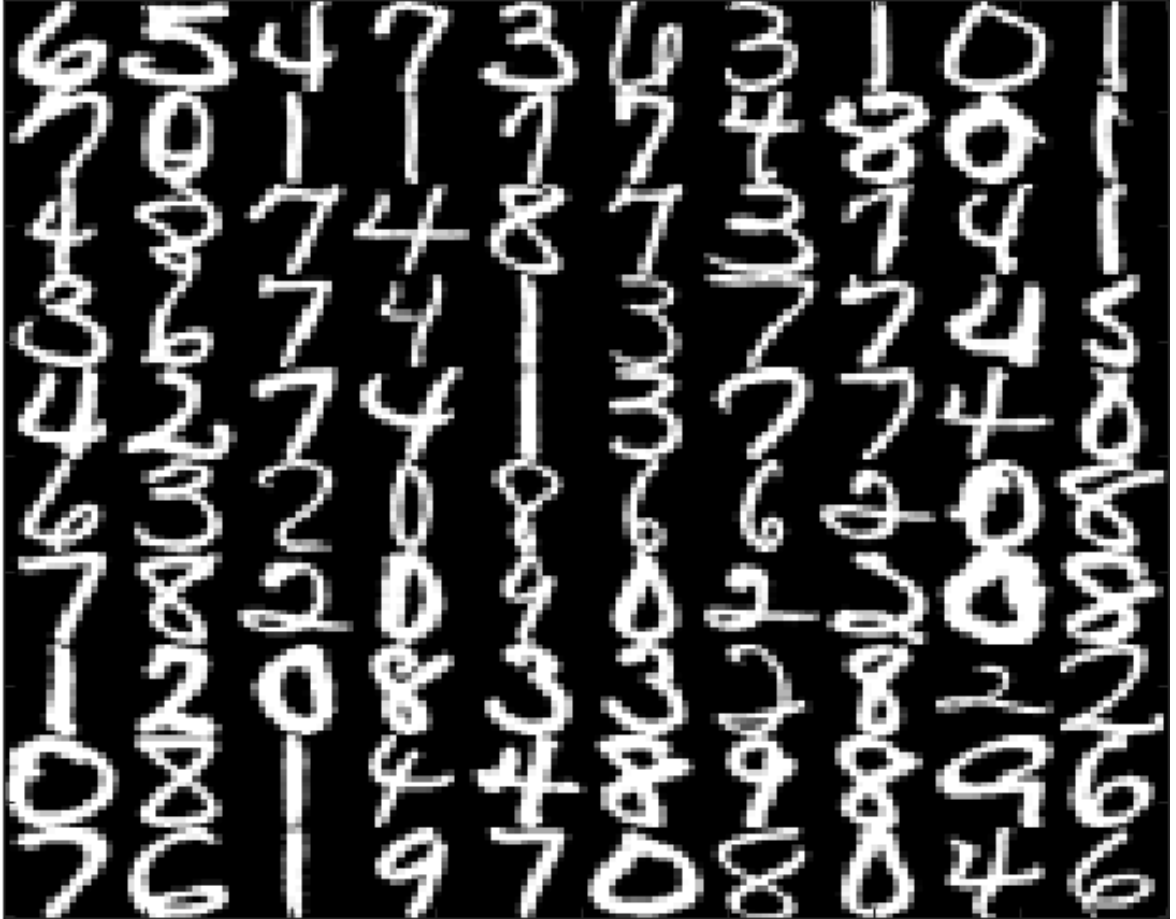
Şekil 2.2. Yüz kısmı kesip çıkartılmış örnekler

Tablo 2.1. USPS veri kümesi örneklerinin sınıflara göre sayısı

	0	1	2	3	4	5	6	7	8	9	Toplam
Eğitim	1194	1005	731	658	652	556	664	645	542	644	7291
Test	359	264	198	166	200	160	170	147	166	177	2007

Tablo 2.2. USPS veri kümesi örneklerinin sınıflara göre oranları

	0	1	2	3	4	5	6	7	8	9	Toplam
Eğitim	0.16	0.14	0.10	0.09	0.09	0.08	0.09	0.09	0.09	0.07	1
Test	0.18	0.13	0.10	0.08	0.10	0.08	0.08	0.07	0.08	0.08	1



Şekil 2.3. USPS veri kümesi için bazı örnekler

2.3.1. Öznitelikler

USPS veri kümesinde sınıflandırma işlemi için görüntü örneklerinden bazı öznitelikler elde edilmiştir. Bu öznitelikler gradyan öznitelikler ve diyagonal özniteliklerdir. Bu özniteliklerin nasıl oluşturulduğu aşağıda açıklanmıştır.

2.3.1.1. Gradyan Öznitelikler

USPS veri kümesi örneklerinden öznitelik çıkarmak için ilk olarak rakam veya karakter tanımadaki kullanılan gradyan öznitelikler tercih edilmiştir [44]. Gradyan öznitelikler Sobel operatörü kullanılarak bulunur. Sobel operatörü her piksel için gradyan vektörünü yaklaşık olarak bir çift filtreyi görüntü ile katlayarak hesaplar. Sobel operatörünün çift filtresi Şekil 2.4'te gösterilmiştir. Şekil 2.4.(a)'daki filtre x ekseninde gradyan vektörünün bileşenini yaklaşık olarak hesaplar. Şekil 2.4.(b)'deki filtre ise y ekseninde gradyan vektörünün bileşenini yaklaşık olarak hesaplar.

-1	0	1
-2	0	2
-1	0	1

(a)

-1	-2	-1
0	0	0
1	2	1

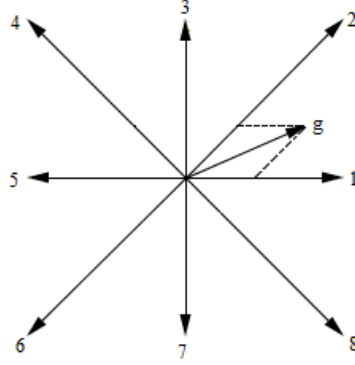
(b)

Şekil 2.4. a) x b) y ekseninde gradyan vektörünün bileşenini yaklaşık hesaplayan filtre

Elde edilen her bir gradyan vektörü standart sekiz yönden kendisine yakın olan iki yönde bileşenlere ayrılır. Şekil 2.5'te g gradyan vektörünün bileşenlerine ayrılması için bir örnek görülmektedir.

Daha sonra her bir görüntü örneği eşit piksel sayısına sahip $m \times n$ bölgeye ayrılır. Her bölgedeki her bir standart yön için, toplam bileşen vektörler hesaplanır. Bütün bölgelerdeki standart yönlerin şiddetleri, görüntü örnekleri için öznitelik vektörünü

oluşturur. Dolayısıyla her bir bölge için 8 tane öznitelik elde edilir. Böylece $m \times n \times 8$ tane öznitelik elde edilir.



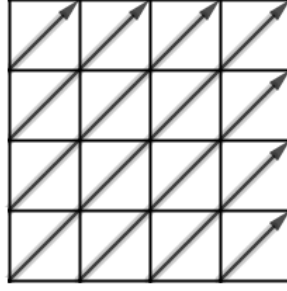
Şekil 2.5. Standart yönler ve bileşenlere ayırma

USPS veri kümesinde yer alan 16×16 piksele sahip görüntüler 4×4 bölgeye ayrılır. Her bölge için her 8 standart yön şiddeti bir öznitelik olduğu için, $4 \times 4 \times 8 = 128$ boyutlu öznitelik vektörleri oluşturulmuştur.

2.3.1.2. Diyagonal Öznitelikler

USPS veri kümesi örneklerinden öznitelik çıkarmak için ikinci olarak rakam veya karakter tanımada kullanılan diyagonal öznitelikler tercih edilmiştir [45]. Diyagonal öznitelikler bulunurken önce her bir görüntü örneği eşit piksel sayısına sahip $m \times n$ bölgeye ayrılır. Daha sonra bir bölgede Şekil 2.6'da gösterildiği gibi diyagonal olarak piksel değerleri toplanır. Bu değerler alt-öznitelikler olarak adlandırılır. Alt-özniteliklerin ortalamasının bulunması ile bu bölge için bir tane öznitelik elde edilir. Bu işlemler her bir bölge için tekrarlanır. Her bir görüntü örneği için $m \times n$ tane bölge olduğundan dolayı $m \times n$ tane öznitelik elde edilmiş olur.

USPS veri kümesinde yer alan 16×16 piksele sahip görüntüler 8×8 bölgeye ayrılır. Her bölgeden bir tane öznitelik elde edildiği için, $8 \times 8 = 64$ boyutlu öznitelik vektörleri oluşturulmuştur.



Şekil 2.6. Diyagonal öznitelik çıkarma

2.4. ORL Veri Kümesine Ait Bulgular

ORL veri kümesi üzerinde boyut indirgeme ve sınıflandırma yapılırken veride yer alan örneklerin yarısı eğitim yarısı test için ayrılmıştır. Bu ayırma işlemi örnekler rastgele seçilerek 5 kere tekrarlanmıştır. Sınıflandırma işlemlerinde bulunan doğruluklar bu 5 tane rastgele olarak oluşturulan eğitim ve test verileri üzerinden elde edilen değerlerin ortalaması alınarak bulunmuştur. Ayrıca boyut indirgeme yöntemlerinde ve sınıflandırma yöntemlerinde yer alan parametreler (k değeri, sınır birimi sayısı vb.) için sınıflandırma doğruluklarının en iyi çıktığı değerler kullanılmıştır.

USPS veri kümesinden elde edilen öznitelikler sırasıyla 11, 25 ve 39 boyuta indirgenmiştir. Veri boyutu için 39'dan sonraki değerlerin kullanılmamasının nedeni, 40 tane sınıfa sahip bu veri kümesi için DAA yönteminde en fazla 39 boyut elde edilebilmesidir. Veri boyutu için 11'den önceki değerlerin kullanılmamasının nedeni ise, 11'den önceki boyutlarda sınıflandırma doğruluklarının düşük elde edilmesidir.

2.4.1. Orijinal Öznitelikler Üzerinde Sınıflandırma

ORL veri kümesi için 32x32 piksele sahip görüntülerden 1024 öznitelik elde edilmiştir. Bu öznitelikler ile k -EYK, YSA ve DVM yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar, doğruluk ve süre açısından Tablo 2.3'te verilmiştir. k -EYK yönteminde k değeri 1 olarak kullanılmıştır. DVM yönteminde gaussian radyal tabanlı çekirdek fonksiyonu kullanılmıştır. YSA yönteminde ise 100 sınır birimine sahip tek gizli katmandan oluşan bir sınır ağı kullanılmıştır.

Tablo 2.3 incelendiğinde süre açısından en başarılı sınıflandırma yönteminin *k*-EYK olduğu görülmektedir. Süre açısından en yavaş yöntemin ise YSA olduğu açıktır. Doğruluk açısından ise YSA yöntemi diğer iki yönteme göre çok daha başarılı olmuştur. *k*-EYK süre konusundaki başarısının yanında doğruluk açısından ise en başarısız yöntem olmuştur.

Tablo 2.3. Orijinal öznitelikler üzerinde sınıflandırma

	Doğruluk(%)	Süre(ms)
<i>k</i> -EYK	86.80	102
DVM	88.70	394
YSA	96.20	206x10 ³

2.4.2. Boyut İndirgeme

ORL veri kümesi üzerinde boyut indirgeme işlemi için TBA, DAA, YKİ, YDAA, KKG ve İZİ yöntemleri uygulanmıştır. ORL veri kümesinin sahip olduğu 1024 boyut, sırasıyla 11, 25 ve 39'a indirgenmiştir. YKİ, YDAA, KKG ve İZİ yöntemlerinde *G* komşuluk grafları bulunurken *k*-en yakın komşular yaklaşımı kullanılmıştır. Sınıflandırma doğruluklarının en iyi çıktığı *k* değeri 4 olarak saptanmıştır. Bu değer bütün komşuluk grafları bulunurken kullanılmıştır.

Doğrusal boyut indirgeme yöntemlerine ait işlem süreleri Tablo 2.4'te ortalama olarak gösterilmiştir. Bir yöntem ile veri boyutu $r = (11, 25, 39)$ 'a indirgenğinde bu değerler için süreler birbirinden çok farklı olmamaktadır. Dolayısıyla ayrı ayrı olarak gösterilmemiştir, ortalama olarak tabloda yer alan değerler elde edilmiştir.

Tablo 2.4. Boyut indirgeme işlemi için süreler

	TBA	DAA	YKİ	YDAA	KKG	İZİ
Süre(ms)	42.061	64.259	72.707	84.730	101.931	301.288

Süre açısından yöntemler kıyaslandığında, TBA yöntemi diğerlerine göre en hızlı yöntem olmuştur. En yavaş yöntemin ise İZİ yöntemi olduğu açıktır.

2.4.3. Boyut İndirgeme İşlemi Ardından Sınıflandırma

ORL veri kümesinin özneliliklerine $r=(11, 25, 39)$ değerleri için boyut indirgeme işlemi ardından k -EYK, YSA ve DVM yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar doğruluk açısından Tablo 2.5, Tablo 2.6 ve Tablo 2.7’de verilmiştir. Bu sınıflandırma işlemlerinde yine k -EYK yönteminde k değeri için 1, DVM yönteminde gaussian radyal tabanlı çekirdek fonksiyonu ve YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

k -EYK yöntemine ait sınıflandırma doğrulukları için Tablo 2.5 incelendiğinde her bir r değeri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin İZİ olduğu açıktır. Yine bu tabloda $r=11$ değeri için YKİ, YDAA ve KKG yöntemleri diğer yöntemlere göre aşırı derecede başarısızlık göstermişlerdir. Ayrıca $r=39$ değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde k -EYK yöntemi ile yapılan sınıflandırma doğrulukları orijinal verinin k -EYK yöntemine ile sınıflandırma doğruluğuna %86.80 göre çok daha başarılı sonuçlar vermişlerdir.

Tablo 2.5. k -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=11$	$r=25$	$r=39$
TBA	84.50	90.00	88.60
DAA	83.40	91.00	93.30
YKİ	70.50	86.30	93.50
YDAA	71.00	88.20	93.80
KKG	71.30	87.50	93.90
İZİ	94.50	92.80	93.90

DVM yöntemine ait sınıflandırma doğrulukları için Tablo 2.6 incelendiğinde ve diğer yöntemlerle kıyaslandığında $r=(11, 25)$ değerleri için en başarılı yöntemin İZİ ve $r=39$ için DAA olduğu açıktır. Yine bu tabloda $r=11$ değeri için YKİ, YDAA ve KKG yöntemleri diğer yöntemlere göre aşırı derecede başarısızlık göstermişlerdir. Ayrıca $r=39$ değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde DVM yöntemi ile yapılan

sınıflandırma doğrulukları orijinal verinin DVM yöntemi ile sınıflandırma doğruluğu %88.70 değerine göre daha başarılı sonuçlar vermişlerdir.

Tablo 2.6. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=11$	$r=25$	$r=39$
TBA	85.10	90.80	93.10
DAA	83.30	91.80	94.60
YKİ	70.60	86.50	93.30
YDAA	71.00	88.20	93.80
KKG	71.30	87.40	93.90
İZİ	94.50	92.80	93.90

YSA yöntemine ait sınıflandırma doğrulukları için Tablo 2.7 incelendiğinde ve diğer yöntemlerle kıyaslandığında $r=(11, 25)$ değerleri için en başarılı yöntemin İZİ ve $r=39$ için TBA olduğu açıktır. Yine bu tabloda $r=11$ değeri için YKİ, YDAA ve KKG yöntemleri diğer yöntemlere göre aşırı derecede başarısızlık göstermişlerdir. Ayrıca her bir r değeri için her bir boyut indirgeme yöntemi ile bulunan veriler üzerinde YSA yöntemi ile yapılan sınıflandırma doğruluğu, orijinal verinin YSA yöntemi ile elde edilen sınıflandırma doğruluğuna %96.20 göre daha başarılı sonuç elde edilememiştir.

Tablo 2.7. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=11$	$r=25$	$r=39$
TBA	78.70	92.80	94.90
DAA	82.60	92.40	93.70
YKİ	70.60	86.00	90.50
YDAA	69.90	86.60	91.90
KKG	68.90	87.90	91.90
İZİ	91.50	93.30	94.50

Farklı r değerleri için sınıflandırma yöntemlerinin ortalama süreleri Tablo 2.8'de verilmiştir.

Tablo 2.8. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(ms)

	$r=11$	$r=25$	$r=39$
k-EYK	3.2	3.8	4.7
DVM	8.5	12.0	21.8
YSA	9.4×10^3	11.8×10^3	14.1×10^3

Bütün bulunan sonuçlar Tablo 2.9'da yer almaktadır. Bu tablo sınıflandırma doğruluğu açısından incelendiğinde en başarılı boyut indirgeme yönteminin, her sınıflandırma yöntemi için %90 ve üzeri başarılar sağlayan İZİ yöntemi olduğu açıkça görülmektedir. Ayrıca her bir boyut indirgeme yöntemi ile her bir sınıflandırma yönteminin işlem süresi kısalmıştır.

Tablo 2.9. ORL veri kümesi için tüm sonuçlar

		Orijinal Öznitelikler		$r=11$	$r=25$	$r=39$
k-EYK	Doğruluk(%)	86.80	TBA	84.50	90.00	88.60
			DAA	83.40	91.00	93.30
			YKİ	70.50	86.30	93.50
			YDAA	71.00	88.20	93.80
			KKG	71.30	87.50	93.90
			İZİ	94.50	92.80	93.90
	Süre(ms)	102		3.2	3.8	4.7
DVM	Doğruluk(%)	88.70	TBA	85.10	90.80	93.10
			DAA	83.30	91.80	94.60
			YKİ	70.60	86.50	93.30
			YDAA	71.00	88.20	93.80
			KKG	71.30	87.40	93.90
			İZİ	94.50	92.80	93.90
	Süre(ms)	394		8.5	12.0	21.8

Tablo 2.9'un devamı

YSA	Doğruluk(%)	96.20	TBA	78.70	92.80	94.90
			DAA	82.60	92.40	93.70
			YKİ	70.60	86.00	90.50
			YDAA	69.90	86.60	91.90
			KKG	68.90	87.90	91.90
			İZİ	91.50	93.30	94.50
	Süre(ms)	206x10 ³		9.4x10 ³	11.8x10 ³	14.1x10 ³

2.5. USPS Veri Kümesine Ait Bulgular

USPS veri kümesinde eğitim ve test kümeleri ayrı ayrı verildiği için bir ayırma işlemi yapılmamıştır. Ayrıca boyut indirgeme yöntemlerinde ve sınıflandırma yöntemlerinde yer alan parametreler (k değeri, siniri birimi sayısı vb.) için sınıflandırma doğruluklarının en iyi çıktığı değerler kullanılmıştır.

USPS veri kümesinden elde edilen öznitelikler sırasıyla 5, 6, 7, 8 ve 9 boyuta indirgenmiştir. Veri boyutu için 9'dan sonraki değerlerin kullanılmamasının nedeni, 10 tane sınıfa sahip bu veri kümesi için DAA yönteminde en fazla 9 boyut elde edilebilmesidir. Veri boyutu için 5'ten önceki değerlerin kullanılmamasının nedeni ise, 5'ten önceki boyutlarda sınıflandırma doğruluklarının çok düşük elde edilmesidir.

2.5.1. Gradyan Özniteliklere Ait Bulgular

USPS veri kümesi üzerinde her bir veri örneği için ilk olarak gradyan öznitelikler ile 128 öznitelik elde edilmiştir.

2.5.1.1. Orijinal Öznitelikler Üzerinde Sınıflandırma

USPS veri kümesinden elde edilen gradyan öznitelikleri kullanarak k -EYK, YSA ve DVM yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar, doğruluk ve süre açısından Tablo 2.10'da verilmiştir. k -EYK yönteminde k değeri 5 olarak kullanılmıştır. DVM yönteminde ise gaussian radyal tabanlı çekirdek fonksiyonu

kullanılmıştır. YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

Tablo 2.10 incelendiğinde süre açısından en hızlı sınıflandırma yönteminin k -EYK olduğu görülmektedir. Süre açısından en yavaş yöntemin ise YSA olduğu açıktır. Doğruluk açısından ise DVM yöntemi diğer iki yönteme göre daha başarılı olmuştur.

Tablo 2.10. Orijinal öznitelikler üzerinde sınıflandırma

	Doğruluk(%)	Süre(s)
k -EYK	96.51	5.00
DVM	97.11	6.70
YSA	96.51	646

2.5.1.2. Boyut İndirgeme

USPS veri kümesinin gradyan öznitelikleri üzerinde boyut indirgeme işlemi için TBA, DAA, YKİ, YDAA, KKG ve İZİ yöntemleri uygulanmıştır. USPS veri kümesinin sahip olduğu 128 boyutu sırasıyla 5, 6, 7, 8 ve 9'a indirgenmiştir. YKİ, YDAA, KKG ve İZİ yöntemlerinde G komşuluk grafları bulunurken k -en yakın komşular yaklaşımı kullanılmıştır. Sınıflandırma doğruluklarının en iyi çıktığı k değeri, 5 olarak saptanmıştır. Bu değer bütün komşuluk grafları bulunurken kullanılmıştır.

Bu yöntemlere ait işlem süreleri Tablo 2.11'de ortalama olarak gösterilmiştir. Bir yöntem için veri boyutu $r = (5, 6, 7, 8, 9)$ 'a indirgendiğinde bu değerler için süreler birbirinden çok farklı olmamaktadır. Dolayısıyla ayrı ayrı olarak gösterilmemiştir, ortalama olarak tabloda yer alan değerler elde edilmiştir.

Süre açısından yöntemler kıyaslandığında, TBA yöntemi diğerlerine göre en hızlı yöntem olmuştur. En yavaş yöntemin ise İZİ yöntemi olduğu açıktır.

Tablo 2.11. Boyut indirgeme işlemi için süreler

	TBA	DAA	YKİ	YDAA	KKG	İZİ
Süre(s)	0.052	0.060	0.750	3.590	4.764	35.400

2.5.1.3. Boyut İndirgeme İşlemi Ardından Sınıflandırma

USPS veri kümesinin gradyan özneliklerine $r=(5, 6, 7, 8, 9)$ değerleri için boyut indirgeme işlemi ardından, k -EYK, DVM ve YSA yöntemleri ile sınıflandırma işlemi gerçekleştirilmiştir. Elde edilen sonuçlar doğruluk açısından Tablo 2.12, Tablo 2.13 ve Tablo 2.14'te verilmiştir. Bu sınıflandırma işlemlerinde yine k -EYK yönteminde k değeri için 5, DVM yönteminde gaussian radyal tabanlı çekirdek fonksiyonu ve YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

k -EYK yöntemine ait sınıflandırma doğrulukları için Tablo 2.12 incelendiğinde ve her bir r değeri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin DAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Yine bu tabloda $r=5$ değeri için TBA ve KKG yöntemleri diğer yöntemlere göre daha düşük performans göstermişlerdir. Ayrıca her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde k -EYK yöntemi ile yapılan sınıflandırma doğrulukları orijinal özneliklerin k -EYK yöntemi ile sınıflandırma doğruluğu %96.51 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.12. k -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	86.35	89.59	91.58	91.93	93.17
DAA	93.02	94.17	95.17	95.27	95.62
YKİ	92.63	93.02	93.87	94.67	94.82
YDAA	93.02	93.52	94.12	94.87	95.07
KKG	89.54	92.63	93.57	93.77	94.42
İZİ	92.48	93.67	94.57	94.72	95.17

DVM yöntemine ait sınıflandırma doğrulukları için Tablo 2.13 incelendiğinde ve $r=(6, 7, 8, 9)$ değerleri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin DAA olduğu açıktır. $r=5$ değeri için ise İZİ yöntemi başarılı olmuştur. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Yine bu tabloda $r=5$ değeri için TBA ve KKG yöntemleri diğer yöntemlere göre daha düşük performans

göstermişlerdir. Ayrıca her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde DVM yöntemi ile yapılan sınıflandırma doğrulukları orijinal özneliklerin DVM yöntemi ile sınıflandırma doğruluğu %97.11 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.13. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	87.05	90.58	92.43	93.12	93.82
DAA	92.83	93.87	95.27	95.57	95.96
YKİ	92.33	93.22	93.97	94.82	94.97
YDAA	92.53	93.42	93.82	95.07	95.32
KKG	89.84	92.53	94.07	94.72	95.22
İZİ	93.32	93.82	94.97	95.17	95.32

YSA yöntemine ait sınıflandırma doğrulukları için Tablo 2.14 incelendiğinde ve her bir r değeri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin DAA olduğu açıktır. $r=5$ değeri için ise İZİ yöntemi başarılı olmuştur. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Yine bu tabloda $r=5$ değeri için TBA ve KKG yöntemleri diğer yöntemlere göre daha düşük performans göstermişlerdir. Ayrıca her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde DVM yöntemi ile yapılan sınıflandırma doğrulukları orijinal özneliklerin YSA yöntemi ile sınıflandırma doğruluğu %96.51 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.14. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	87.05	90.98	92.38	92.40	93.17
DAA	92.87	94.02	95.12	96.01	95.71
YKİ	92.68	92.97	94.07	94.22	94.77

Tablo 2.14'ün devamı

YDAA	92.67	93.32	94.47	94.37	95.12
KKG	90.48	92.92	93.42	94.22	94.72
İZİ	92.53	92.78	94.62	94.62	95.12

Farklı r boyut değerleri için sınıflandırma yöntemlerinin ortalama süreleri Tablo 2.15'te verilmiştir.

Tablo 2.15. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(s)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
k -EYK	0.36	0.38	0.40	0.42	0.44
DVM	0.70	0.72	0.72	0.76	0.76
YSA	126	131	135	139	145

Bütün bulunan sonuçlar Tablo 2.16'da yer almaktadır. Bu tablo sınıflandırma doğruluğu açısından incelendiğinde en başarılı boyut indirgeme yönteminin, her bir sınıflandırma yöntemi için diğer boyut indirgeme yöntemlerinden daha fazla doğruluk sağlayan DAA yöntemi olduğu açıkça görülmektedir. Ayrıca her bir boyut indirgeme yöntemi ile yapılan her bir sınıflandırma yönteminin işlem süresi kısalmıştır.

Tablo 2.16. USPS-Gradyan öznitelikler için tüm sonuçlar

		Orijinal Öznitelikler		$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
k -EYK	Doğruluk(%)	96.51	TBA	86.35	89.59	91.58	91.93	93.17
			DAA	93.02	94.17	95.17	95.27	95.62
			YKİ	92.63	93.02	93.87	94.67	94.82
			YDAA	93.02	93.52	94.12	94.87	95.07
			KKG	89.54	92.63	93.57	93.77	94.42
			İZİ	92.48	93.67	94.57	94.72	95.17
	Süre(s)	5.00		0.36	0.38	0.40	0.42	0.44

Tablo 2.16'nın devamı

DVM	Doğruluk(%)	97.11	TBA	87.05	90.58	92.43	93.12	93.82
			DAA	92.83	93.87	95.27	95.57	95.96
			YKİ	92.33	93.22	93.97	94.82	94.97
			YDAA	92.53	93.42	93.82	95.07	95.32
			KKG	89.84	92.53	94.07	94.72	95.22
			İZİ	93.32	93.82	94.97	95.17	95.32
	Süre(s)	6.70		0.70	0.72	0.72	0.76	0.76
YSA	Doğruluk(%)	96.51	TBA	87.05	90.98	92.38	92.40	93.17
			DAA	92.87	94.02	95.12	96.01	95.71
			YKİ	92.68	92.97	94.07	94.22	94.77
			YDAA	92.67	93.32	94.47	94.37	95.12
			KKG	90.48	92.92	93.42	94.22	94.72
			İZİ	92.53	92.78	94.62	94.62	95.12
	Süre(s)	646		126	131	135	139	145

2.5.2. Diyagonal Özniteliklere Ait Bulgular

USPS veri kümesi üzerinde her bir veri örneği için ilk olarak diyagonal öznitelikler ile 64 öznitelik elde edilmiştir.

2.5.2.1. Orijinal Öznitelikler Üzerinde Sınıflandırma

USPS veri kümesinden elde edilen diyagonal öznitelikleri kullanarak k -EYK, YSA ve DVM yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar doğruluk ve süre açısından Tablo 2.17'de verilmiştir. k -EYK yönteminde k değeri 5 olarak kullanılmıştır. DVM yönteminde ise gaussian radyal tabanlı çekirdek fonksiyonu kullanılmıştır. YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

Tablo 2.17 incelendiğinde süre açısından en hızlı sınıflandırma yönteminin k -EYK olduğu görülmektedir. Süre açısından en yavaş yöntemin ise YSA olduğu açıktır. Doğruluk açısından ise DVM yöntemi diğer iki yönteme göre daha başarılı olmuştur.

Tablo 2.17. Orijinal öznitelikler üzerinde sınıflandırma

	Doğruluk(%)	Süre(s)
<i>k</i> -EYK	94.72	1.63
DVM	95.52	6.57
YSA	93.82	382

2.5.2.2. Boyut İndirgeme

USPS veri kümesinin diyagonal öznitelikleri üzerinde boyut indirgeme işlemi için TBA, DAA, YKİ, YDAA, KKG ve İZİ yöntemleri uygulanmıştır. USPS veri kümesinin sahip olduğu 64 boyutu sırasıyla 5, 6, 7, 8 ve 9'a indirgenmiştir. YKİ, YDAA, KKG ve İZİ yöntemlerinde G komşuluk grafları bulunurken k -en yakın komşular yaklaşımı kullanılmıştır. Sınıflandırma doğruluklarının en iyi çıktığı k değeri, 5 olarak saptanmıştır. Bu değer bütün komşuluk grafları bulunurken kullanılmıştır.

Bu yöntemlere ait işlem süreleri Tablo 2.18'de ortalama olarak gösterilmiştir. Bir yöntem için veri boyutu $r = (5, 6, 7, 8, 9)$ 'a indirgendiğinde bu değerler için süreler birbirinden çok farklı olmamaktadır. Dolayısıyla ayrı ayrı olarak gösterilmemiştir, ortalama olarak tabloda yer alan değerler elde edilmiştir.

Süre açısından yöntemler kıyaslandığında, TBA yöntemi diğerlerine göre en hızlı yöntem olmuştur. En yavaş yöntemin ise İZİ yöntemi olduğu açıktır.

Tablo 2.18. Boyut indirgeme işlemi için süreler

	TBA	DAA	YKİ	YDAA	KKG	İZİ
Süre(s)	0.041	0.044	0.660	3.501	4.371	35.533

2.5.2.3. Boyut İndirgeme İşlemi Ardından Sınıflandırma

USPS veri kümesinin diyagonal özniteliklerine $r = (5, 6, 7, 8, 9)$ değerleri için boyut indirgeme işlemi ardından k -EYK, DVM ve YSA yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar doğruluk açısından Tablo 2.19, Tablo 2.20 ve Tablo 2.21'de verilmiştir. Bu sınıflandırma işlemlerinde yine k -EYK yönteminde k değeri

için 5, DVM yönteminde gaussian radyal tabanlı çekirdek fonksiyonu ve YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

k -EYK yöntemine ait sınıflandırma doğrulukları için Tablo 2.19 incelendiğinde ve diğer yöntemlerle kıyaslandığında $r=(5, 6, 7, 8)$ değerleri için en başarılı yöntemin DAA ve $r=9$ değeri için en başarılı yöntemin YDAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Ama her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde k -EYK yöntemi ile yapılan sınıflandırma doğrulukları orijinal özniteliklerin k -EYK yöntemi ile sınıflandırma doğruluğu %94.72 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.19. k -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	82.26	84.26	87.14	88.59	89.94
DAA	85.35	88.74	89.89	91.03	90.93
YKİ	84.85	86.45	88.54	90.33	89.84
YDAA	85.05	87.34	89.54	90.73	90.98
KKG	83.36	86.20	88.14	89.59	90.38
İZİ	85.60	88.04	89.44	90.13	90.88

DVM yöntemine ait sınıflandırma doğrulukları için Tablo 2.20 incelendiğinde ve diğer yöntemlerle kıyaslandığında $r=5$ değeri için en başarılı yöntemin DAA, $r=(6, 8)$ değerleri için en başarılı yöntemin İZİ, $r=7$ değeri için en başarılı yöntemin YKİ ve $r=9$ değeri için en başarılı yöntemin YDAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Ama her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde DVM yöntemi ile yapılan sınıflandırma doğrulukları orijinal özniteliklerin DVM yöntemi ile sınıflandırma doğruluğu %95.52 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır

Tablo 2.20. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	84.45	86.25	88.39	90.18	91.53
DAA	86.65	88.99	90.08	90.83	91.53
YKİ	85.55	87.44	90.33	90.33	91.43
YDAA	85.95	87.79	90.23	90.88	91.68
KKG	84.21	87.24	88.84	90.23	91.18
İZİ	86.25	89.44	90.23	91.18	92.68

YSA yöntemine ait sınıflandırma doğrulukları için Tablo 2.21 incelendiğinde ve diğer yöntemlerle kıyaslandığında $r=(5, 8)$ değerleri için en başarılı yöntemin DAA, $r=(6, 7)$ değerleri için en başarılı yöntemin İZİ ve $r=9$ değeri için en başarılı yöntemin YDAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Ama her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde YSA yöntemi ile yapılan sınıflandırma doğrulukları orijinal özneliklerin YSA yöntemi ile sınıflandırma doğruluğu %93.82 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.21. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	83.71	85.70	88.19	89.49	89.94
DAA	85.95	88.19	89.09	90.38	91.28
YKİ	84.45	86.75	89.19	89.74	90.23
YDAA	85.80	87.39	89.34	89.99	91.43
KKG	84.60	86.90	88.19	89.19	90.23
İZİ	85.70	88.39	89.49	90.08	91.08

Farklı r boyut değerleri için sınıflandırma yöntemlerinin süreleri Tablo 2.22'de verilmiştir.

Tablo 2.22. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(s)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
k -EYK	0.36	0.38	0.40	0.42	0.44
DVM	1.23	1.25	1.36	1.40	1.61
YSA	126	131	135	139	145

Bütün bulunan sonuçlar Tablo 2.23'te yer almaktadır. Bu tablo sınıflandırma doğruluğu açısından incelediğinde boyut indirgeme yöntemleri arasında her sınıflandırma yöntemi için sürekli bir başarı gösteren bir yöntemden bahsetmek zordur. DAA, YKİ, YDAA ve İZİ yöntemleri bazı sınıflandırma yöntemleri ile yerel başarılar sağlamışlardır. Ayrıca her bir boyut indirgeme yöntemi ile yapılan her bir sınıflandırma yönteminin işlem süresi kısalmıştır.

Tablo 2.23. USPS-Diyagonal öznitelikler için tüm sonuçlar

		Orijinal Öznitelikler		$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
k -EYK	Doğruluk(%)	94.72	TBA	82.26	84.26	87.14	88.59	89.94
			DAA	85.35	88.74	89.89	91.03	90.93
			YKİ	84.85	86.45	88.54	90.33	89.84
			YDAA	85.05	87.34	89.54	90.73	90.98
			KKG	83.36	86.20	88.14	89.59	90.38
			İZİ	85.60	88.04	89.44	90.13	90.88
	Süre(s)	1.63		0.36	0.38	0.40	0.42	0.44
DVM	Doğruluk(%)	95.52	TBA	84.45	86.25	88.39	90.18	91.53
			DAA	86.65	88.99	90.08	90.83	91.53
			YKİ	85.55	87.44	90.33	90.33	91.43
			YDAA	85.95	87.79	90.23	90.88	91.68
			KKG	84.21	87.24	88.84	90.23	91.18
			İZİ	86.25	89.44	90.23	91.18	92.68
	Süre(s)	6.57		1.23	1.25	1.36	1.40	1.61

Tablo 2.23'ün devamı

YSA	Doğruluk(%)	93.82	TBA	83.71	85.70	88.19	89.49	89.94
			DAA	85.95	88.19	89.09	90.38	91.28
			YKİ	84.45	86.75	89.19	89.74	90.23
			YDAA	85.80	87.39	89.34	89.99	91.43
			KKG	84.60	86.90	88.19	89.19	90.23
			İZİ	85.70	88.39	89.49	90.08	91.08
	Süre(s)	382		126	131	135	139	145

2.5.3. Gradyan ve Diyagonal Özniteliklere Ait Bulgular

USPS veri kümesi üzerinde her bir veri örneği için gradyan öznitelikler için 128 öznitelik ve diyagonal öznitelikler için 64 öznitelik elde edilmiştir. Her bir veri örneği için bu iki öznitelik birlikte kullanılabilir. Dolayısıyla her bir veri örneği 192 öznitelik ile ifade edilmiş olur.

2.5.3.1. Orijinal Öznitelikler Üzerinde Sınıflandırma

USPS veri kümesinden elde edilen gradyan-diyagonal öznitelikleri kullanarak k -EYK, YSA ve DVM yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar doğruluk ve süre açısından Tablo 2.24'te verilmiştir. Orijinal öznitelikler ile sınıflandırma yaparken k -EYK yönteminde k değeri 5 olarak kullanılmıştır. DVM yönteminde ise gaussian radyal tabanlı çekirdek fonksiyonu kullanılmıştır. YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

Tablo 2.24. Orijinal öznitelikler üzerinde sınıflandırma

	Doğruluk(%)	Süre(s)
k -EYK	96.06	5.24
DVM	96.86	8.51
YSA	96.26	987

Tablo 2.24 incelendiğinde, süre açısından en hızlı sınıflandırma yönteminin k -EYK olduğu görülmektedir. Süre açısından en yavaş yöntemin ise YSA olduğu açıktır. Doğruluk açısından ise DVM yöntemi diğer iki yönetime göre daha başarılı olmuştur.

2.5.3.2. Boyut İndirgeme

USPS veri kümesinin gradyan ve diyagonal öznitelikleri üzerinde boyut indirgeme işlemi için TBA, DAA, YKİ, YDAA, KKG ve İZİ yöntemleri uygulanmıştır. USPS veri kümesinin sahip olduğu 192 boyutu sırasıyla 5, 6, 7, 8 ve 9'a indirgenmiştir. YKİ, YDAA, KKG ve İZİ yöntemlerinde G komşuluk grafları bulunurken k -en yakın komşular yaklaşımı kullanılmıştır. Sınıflandırma doğruluklarının en iyi çıktığı k değeri, 5 olarak saptanmıştır. Bu değer bütün komşuluk grafları bulunurken kullanılmıştır.

Bu yöntemlere ait işlem süreleri Tablo 2.25'te ortalama olarak gösterilmiştir. Bir yöntem için veri boyutu $r = (5, 6, 7, 8, 9)$ 'a indirgenildiğinde bu değerler için süreler birbirinden çok farklı olmamaktadır. Dolayısıyla ayrı ayrı olarak gösterilmemiştir, ortalama olarak tabloda yer alan değerler elde edilmiştir.

Tablo 2.25. Boyut indirgeme işlemi için süreler

	TBA	DAA	YKİ	YDAA	KKG	İZİ
Süre(s)	0.070	0.076	0.852	3.605	4.568	36.351

Süre açısından yöntemler kıyaslandığında, TBA yöntemi diğerlerine göre en hızlı yöntem olmuştur. En yavaş yöntemin ise İZİ yöntemi olduğu açıktır.

2.5.3.3. Boyut İndirgeme İşlemi Ardından Sınıflandırma

USPS veri kümesinin gradyan ve diyagonal özniteliklerine $r=(5, 6, 7, 8, 9)$ değerleri için boyut indirgeme işlemi ardından k -EYK, DVM ve YSA yöntemleri ile sınıflandırma işlemi gerçekleştirildiğinde elde edilen sonuçlar doğruluk açısından Tablo 2.26, Tablo 2.27 ve Tablo 2.28'de verilmiştir. Bu sınıflandırma işlemlerinde yine k -EYK yönteminde k değeri

için 5, DVM yönteminde gaussian radyal tabanlı çekirdek fonksiyonu ve YSA yönteminde 100 sinir birimine sahip tek gizli katmandan oluşan bir sinir ağı kullanılmıştır.

k -EYK yöntemine ait sınıflandırma doğrulukları için Tablo 2.26 incelendiğinde ve her bir r değeri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin DAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Yine bu tabloda $r=5$ değeri için TBA, YDAA ve KKG yöntemleri diğer yöntemlere göre daha düşük performans göstermişlerdir. Ayrıca DAA yöntemi ve $r=9$ değeri durumu hariç, bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde k -EYK yöntemi ile yapılan sınıflandırma doğrulukları orijinal özniteliklerin k -EYK yöntemi ile sınıflandırma doğruluğu %96.06 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır. DAA yöntemi ve $r=9$ değeri için %96.26 ile daha fazla başarı elde edilmiştir.

Tablo 2.26. k -EYK yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	85.45	88.74	88.84	90.93	91.58
DAA	93.82	94.32	95.86	95.86	96.26
YKİ	91.23	92.23	93.42	93.72	94.42
YDAA	86.25	90.63	92.63	93.32	93.62
KKG	88.79	90.93	92.53	93.12	93.52
İZİ	92.38	92.87	93.97	94.57	94.72

DVM yöntemine ait sınıflandırma doğrulukları için Tablo 2.27 incelendiğinde her bir r değeri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin DAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Yine bu tabloda $r=5$ değeri için TBA, YDAA ve KKG yöntemleri diğer yöntemlere göre daha düşük performans göstermişlerdir. Ayrıca her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde DVM yöntemi ile yapılan sınıflandırma doğrulukları orijinal özniteliklerin DVM yöntemine ile sınıflandırma doğruluğu %96.86 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.27. DVM yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	86.00	89.09	90.1	91.58	92.38
DAA	93.97	94.47	95.96	95.86	96.31
YKİ	91.28	92.38	93.47	94.32	94.57
YDAA	85.95	90.48	92.48	93.42	94.37
KKG	89.09	91.78	92.83	93.77	94.12
İZİ	92.68	92.83	93.62	93.87	94.47

YSA yöntemine ait sınıflandırma doğrulukları için Tablo 2.28 incelendiğinde ve her bir r değeri için diğer yöntemlerle kıyaslandığında en başarılı yöntemin DAA olduğu açıktır. Her bir boyut indirgeme yöntemi ile elde edilen en iyi sınıflandırma doğrulukları $r=9$ değerinde bulunmuştur. Yine bu tabloda $r=5$ değeri için TBA, YDAA ve KKG yöntemleri diğer yöntemlere göre daha düşük performans göstermişlerdir. Ayrıca her bir r değeri için bütün boyut indirgeme yöntemleri ile bulunan veriler üzerinde YSA yöntemi ile yapılan sınıflandırma doğrulukları orijinal özneliklerin YSA yöntemine ile sınıflandırma doğruluğu %96.26 değerine göre daha başarılı sonuçlar ortaya koyamamışlardır.

Tablo 2.28. YSA yöntemi ile farklı boyut değerleri için sınıflandırma doğrulukları(%)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
TBA	85.95	89.14	89.49	90.78	92.03
DAA	93.87	94.42	95.76	95.96	96.11
YKİ	91.33	92.33	93.82	93.72	94.07
YDAA	87.84	90.68	92.53	93.47	93.82
KKG	89.09	91.63	93.07	93.02	93.77
İZİ	92.48	92.68	93.37	94.12	94.47

Farklı r değerleri için sınıflandırma yöntemlerinin ortalama süreleri Tablo 2.29'da verilmiştir.

Tablo 2.29. Farklı boyut değerleri için sınıflandırma yöntemlerinin süreleri(s)

	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
k -EYK	0.36	0.38	0.40	0.42	0.44
DVM	0.90	0.92	0.94	0.96	0.98
YSA	126	131	135	139	145

Bütün bulunan sonuçlar Tablo 2.30'da yer almaktadır. Bu tablo sınıflandırma doğruluğu açısından incelediğinde en başarılı boyut indirgeme yönteminin her sınıflandırma yöntemi için diğer boyut indirgeme yöntemlerinden daha fazla doğruluk sağlayan DAA yöntemi olduğu açıkça görülmektedir. Ayrıca her bir boyut indirgeme yöntemi ile yapılan her bir sınıflandırma yönteminin işlem süresi kısalmıştır.

Tablo 2.30. USPS-Gradyan ve Diyagonal öznitelikler için tüm sonuçlar

		Orijinal Öznitelikler		$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
k -EYK	Doğruluk(%)	96.06	TBA	85.45	88.74	88.84	90.93	91.58
			DAA	93.82	94.32	95.86	95.86	96.26
			YKİ	91.23	92.23	93.42	93.72	94.42
			YDAA	86.25	90.63	92.63	93.32	93.62
			KKG	88.79	90.93	92.53	93.12	93.52
			İZİ	92.38	92.87	93.97	94.57	94.72
	Süre(s)	5.24		0.36	0.38	0.40	0.42	0.44
DVM	Doğruluk(%)	96.86	TBA	86.00	89.09	90.1	91.58	92.38
			DAA	93.97	94.47	95.96	95.86	96.31
			YKİ	91.28	92.38	93.47	94.32	94.57
			YDAA	85.95	90.48	92.48	93.42	94.37
			KKG	89.09	91.78	92.83	93.77	94.12
			İZİ	92.68	92.83	93.62	93.87	94.47
	Süre(s)	8.51		0.90	0.92	0.94	0.96	0.98

Tablo 2.30'un devamı

YSA	Doğruluk(%)	96.26	TBA	85.95	89.14	89.49	90.78	92.03
			DAA	93.87	94.42	95.76	95.96	96.11
			YKİ	91.33	92.33	93.82	93.72	94.07
			YDAA	87.84	90.68	92.53	93.47	93.82
			KKG	89.09	91.63	93.07	93.02	93.77
			İZİ	92.48	92.68	93.37	94.12	94.47
	Süre(s)	987		126	131	135	139	145

3. SONUÇLAR VE ÖNERİLER

Bu tez çalışması kapsamında TBA, DAA, YKİ, YDAA, KKG ve İZİ olarak adlandırılan doğrusal boyut indirgeme yöntemlerinin yüz ve rakam sınıflandırma üzerindeki etkileri ve performansları incelenmiştir. Çalışmada yüz tanıma için ORL veri kümesi ve rakam tanıma için USPS veri kümesi kullanılmıştır.

ORL veri kümesi üzerinde orijinal öznitelikler olarak görüntülerin sahip olduğu 1024 piksel kullanılmıştır. Bu veri kümesi üzerinde her bir doğrusal boyut indirgeme yöntemi ile elde edilen özniteliklerle yapılan bazı sınıflandırma işlemlerinin, verinin orijinal öznitelikleri ile elde edilen sınıflandırma işlemlerinin doğruluklarına göre daha başarılı sonuçlar verdiği görülmüştür. Bunun yanında boyut indirgeme yöntemleri ile yapılan sınıflandırma işlemlerinin doğruluğu konusunda en başarılı yöntem olarak İZİ yöntemi öne çıkmıştır. Her bir boyut indirgeme işlemi ile sınıflandırma sürelerinde kayda değer azalmalar görülmüştür. Boyut indirgeme işlemi süresinde ise en başarılı yöntem olarak TBA yönteminin ve en başarısız yöntem olarak İZİ yönteminin olduğu görülmüştür.

USPS veri kümesi üzerinde orijinal öznitelikler olarak sadece gradyan öznitelikler, sadece diyagonal öznitelikler ve gradyan-diyagonal öznitelikler birlikte kullanılmıştır. Bu öznitelikler arasında bir karşılaştırma yapılırsa, orijinal öznitelikler ile yapılan sınıflandırmada yada boyut indirgeme yöntemleri ile yapılan sınıflandırmada en başarılı sonuçları genelde gradyan özniteliklerin verdiği görülmüştür.

Bu veri kümesi üzerinde her bir doğrusal boyut indirgeme yöntemi ile elde edilen özniteliklerle yapılan sınıflandırma işlemlerinin verinin orijinal öznitelikleri ile elde edilen sınıflandırma işlemlerinin doğruluklarına göre daha başarılı sonuçlar vermedikleri görülmüştür. Bunun yanında gradyan öznitelikler ve gradyan-diyagonal öznitelikler üzerinde boyut indirgeme yöntemleri ile yapılan sınıflandırma işlemlerinin doğruluğu konusunda en başarılı yöntem olarak DAA yöntemi öne çıkmıştır. Sadece diyagonal öznitelikler üzerinden boyut indirgeme yöntemleri ile yapılan sınıflandırma işlemlerinin doğruluklarına göre herhangi bir yöntem kesin başarılı olarak öne çıkmamıştır. Bu özniteliklerde DAA, YKİ, YDAA ve İZİ yöntemleri bazı sınıflandırma yöntemleri ile yerel başarılar sağlamışlardır. Her bir boyut indirgeme işlemi ile sınıflandırma sürelerinde kayda değer azalmalar görülmüştür. Boyut indirgeme süresinde ise en başarılı yöntem olarak TBA yönteminin ve en başarısız yöntem olarak İZİ yönteminin olduğu görülmüştür.

Elde edilen bu sonuçlar, doğrusal boyut indirgeme yöntemleri ile bir veri kümesinin sınıflandırılmasında doğruluğun artmasının veya azalmasının veri türüne bağlı olduğunu göstermiştir. Veriden elde edilen özneliklere göre bazı doğrusal boyut indirgeme yöntemlerinin öne çıktığı da belirlenmiştir.

Elde edilen sonuçlarda YKİ, YDAA, KKG ve İZİ yöntemlerinin boyut indirgeme işlem sürelerinin TBA ve DAA yöntemlerine göre uzun oldukları görülmüştür. Bu yöntemler için algoritma adımlarında boyut indirgeme işlem sürelerinde azalmalar sağlayan bazı değişiklikler ile işlem sürelerinin azaltılmaları sağlanabilir.

Rakam verisinin orijinal öznelikler ile yapılan sınıflandırma doğruluklarına göre boyut indirgeme ile elde edilen sınıflandırma doğrulukları daha düşük gözlemlendiği için, boyut indirgeme algoritmalarının adımlarında sınıflandırma doğruluklarında artışlar sağlayan bazı değişiklikler ile sınıflandırma doğrulukları artırılabilir.

Ayrıca bu çalışmada başarılı olarak öne çıkan boyut indirgeme yöntemleri de göz önünde bulundurularak, yeni boyut indirgeme yöntemleri geliştirilebilir.

4. KAYNAKLAR

1. Lee, J.A., Verleysen, M., Nonlinear Dimensionality Reduction, First Edition, Springer, 2007.
2. Martinez W.L., Martinez A.R., Solka J.L., Exploratory Data Analysis With MATLAB, Second Edition, Chapman & Hall /CRC, 2011.
3. L.J.P. van der Maaten, E.O. Postma, and H.J. van den Herik. Dimensionality Reduction: A Comparative Review, Tilburg University Technical Report, TiCC-TR 2009-005, 2009.
4. Cunningham, J.P. ve Ghahramani, Z., Linear Dimensionality Reduction: Survey, Insights and Generalizations, Journal of Machine Learning Research, 16 (2015) 2859–2900.
5. Lu, H., Plataniotis K.N., Venetsanopoulos, A.N., Multilinear Subspace Learning, First Edition, Chapman & Hall /CRC, 2013.
6. Pearson, K., On Lines and Planes of Closest Fit to Systems of Points in Space, Philosophical Magazine, 2, 11 (1901) 559–572.
7. Hotelling, H., Analysis of a Complex of Statistical Variables into Principal Components, Journal of Educational Psychology, 24,6 (1933) 417–441.
8. Hotelling, H., Simplified Calculation of Principal Components, Psychometrika, 1, 1 (1936) 27–35.
9. Jolliffe, I.T., Principal Component Analysis, Second Edition, Springer, 2002.
10. Rao, C.R., The Use and Interpretation of Principal Component Analysis in Applied Research, The Indian Journal of Statistics, Series A ,26, 4 (1964) 329–358.
11. Shlens, J., A Tutorial on Principal Component Analysis: Derivation, Discussion and Singular Value Decomposition, https://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf, 21.08.2017.
12. Alpaydın, E., Introduction To Machine Learning, Third Edition, MIT Press, 2014.
13. Fisher, R.A, The Use Of Multiple Measurements in Taxonomic Problems, Annals of Eugenics, 7, 2 (1936) 179-188.
14. Osuna, R.G., Linear Discriminant Analysis, http://courses.cs.tamu.edu/rgutier/csce666_f13/110.pdf, 16.09.2017.

15. Duda, R.O., Hart, P.E. ve Stok, D.G., Pattern Classification, Second Edition, Wiley, 2002.
16. Fukunaga, K., Introduction To Statistical Pattern Recognition, Second Edition, Academic Press, 1990.
17. He, X., Locality Preserving Projections, Doktora Tezi, University of Chicago, 2005.
18. He, X. ve Niyogi, P., Locality Preserving Projections, Advances in Neural Information Processing Systems, 2003.
19. He, X., Yan, S., Hu, Y., Niyogi, P. ve Zhang, H.J., Face Recognition Using Laplacianfaces, Ieee Transactions on Pattern Analysis and Machine Intellegence, 27, 3 (2005) 328-340.
20. Shermina, J., Application of Locality Preserving Projections in Face Recognition, International Journal of Advanced Computer Science and Applications, 1, 3 (2010) 82-85.
21. Tang, Y., ve Rose, R., A Study of Using Locality Preserving Projections for Feature Extraction in Speech Recognition, IEEE International Conference on Acoustics, Speech and Signal Processing, Mayıs 2008, Las Vegas, 1569-1572.
22. Zhang, M., Jia, P., Shen, Y., ve F. Shen, Hyperspectral Image Classification Method Based on Orthogonal NMF and LPP, 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Temmuz 2016, Taipei, 1-5.
23. Wang, Z. ve He, B., Locality Preserving Projections Algorithm for Hyperspectral Image Dimensionality Reduction, 2011 19th International Conference on Geoinformatics, Haziran 2011, Shanghai, 1-4.
24. Lu, J., Zhao, Y., Xue, Y. ve Hu, J., Palmprint Recognition via Locality Preserving Projections and Extreme Learning Machine Neural Network, 2008 9th International Conference on Signal Processing, Aralık 2008, Beijing, 2096-2099.
25. Cai, D., He, X., Zhou, K., Han, J. ve Bao, H., Locality Sensitive Discriminant Analysis, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Ocak 2007, Haydarabad, 708-713.
26. Liu, X., Wang, Z., Liu, J. Ve Feng, Z., Face Recognition with Locality Sensitive Discriminant Analysis Based on Matrix Representation, IEEE International Joint Conference on Neural Networks, Eylül 2008, Hong Kong, 4052-4058.
27. Yu, H., Gao, L., Li, W., Du, Q. ve Zhang, B., Locality Sensitive Discriminant Analysis for Group Sparse Representation-Based Hyperspectral Imagery Classification, IEEE Geoscience and Remote Sensing Letters, 14, 8 (2017) 1358-1362.

28. Huang, Y., Tang, K. ve Sun, Z., Locality Sensitive Discriminant Analysis for Classification of Hyperspectral Data, 7th International Congress on Image and Signal Processing, Ekim 2014, Dalian, 292-296.
29. Cai, D., Cai, W., Ni, Z. ve Li, M., Locality Sensitive Discriminant Analysis for Speaker Verification, Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Aralık 2016, Jeju, 1-5.
30. He, X., Cai, D., Hu, Y., Yan, S. ve Zhang, H.J., Neighbourhood Preserving Embedding, IEEE International Conference on Computer Vision (ICCV'05), Ekim 2005, Beijing, 1208-1213.
31. Khodr, J. ve Younes, R., Dimensionality Reduction on Hyperspectral Images: A Comparative Review Based on Artificial Datas, 4th *International Congress on Image and Signal Processing*, Ekim 2011, Shanghai, 1875-1883.
32. Wang, Z. ve Sun, X., Document Classification Algorithm Based on NPE and PSO, International Conference on E-Business and Information System Security, Mayıs 2009, Wuhan, 1-4.
33. Ziqiang, W. ve Xia, S., Web Document Retrieval Using Manifold Learning and ACO Algorithm, IEEE International Conference on Broadband Network & Multimedia Technology, Ekim 2009, Beijing, 152-155.
34. Wang, Z. ve Sun, X., An Efficient Spam Filtering Algorithm Based on NPE, IEEE International Symposium on Knowledge Acquisition and Modeling Workshop, Aralık 2008, Wuhan, 1102-1104.
35. Sun, X. ve Wang, Z.Q., Tissue Classification Algorithm Based on NPE and LS-SVM, IEEE International Symposium on IT in Medicine & Education, *Ağustos* 2009, Jinan, 1146-1149.
36. Cai, D., He, X. ve Han, J., Isometric Projections, The AAI Conference On Artificial Intellegence, 1 (2007) 528-533.
37. Younesi, A. ve Amirani, M. C., Palmprint Identification Using Isometric Projection and Linear Discriminant Analysis, International Conference on Communications, Signal Processing, and their Applications (ICCSPA), Şubat 2013, Sharjah, 1-5.
38. Al-juboori, A.M., Bu, W., Wu, X. ve Zhao, Q., Palm Vein Verification Using Multiple Features and Isometric Projections, International Journal Of Signal Processing and Pattern Recognition, 7, 1 (2014) 33-34.
39. Cai, D., He, X. ve Han, J., Isometric Projection Technical Report, Computer Science Department, UIUC, 2006, NO. UIUCDCS-R-2006-2747.

40. Han, J., Kamber, M. ve Pei, J., Data Mining: Concepts and Techniques, Third Edition, Morgan Kaufmann Publishers, MA, 2012.
41. Hastie, T., Tibshirani, R., Friedman, J., The Elements of Statistical Learning, Second Edition, Springer, 2009.
42. The ORL Database of Faces,
www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html, 16.09.2017.
43. The USPS Database of Digits,
<https://web.stanford.edu/~hastie/ElemStatLearn/data.html>, 16.09.2017.
44. Liu, C.L., Nkashima, K., Sako, H. ve Fujisawa, H., Handwritten Digit Recognition: Investigation of Normalization and Feature Extraction Techniques, Pattern Recognition, 37, 2 (2004) 265-279.
45. Pradeep, J., Srinivasan, E., ve Himavathi, S., Diagonal Based Feature Extraction for Handwritten Character Recognition System using Neural Network, 3rd International Conference on Electronics Computer Technology, Nisan 2011, Kanyakumari, 364-368.

ÖZGEÇMİŞ

Eray YILDIZ, 09.05.1990 tarihinde Bartın'da doğdu. İlköğretimini ve ortaöğretimini Bartın'da 2008 yılında tamamladı. 2014 yılında Hacettepe Üniversitesi Elektrik-Elektronik Mühendisliği Bölümünden mezun oldu. 2015 yılında Karadeniz Teknik Üniversitesi Elektrik-Elektronik Mühendisliği Devreler ve Sistemler Anabilim Dalı'na ÖYP araştırma görevlisi olarak atandı. Bu tarihten itibaren Karadeniz Teknik Üniversitesi Elektrik-Elektronik Mühendisliği Devreler ve Sistemler Anabilim Dalı'nda araştırma görevlisi olarak lisansüstü çalışmalarını sürdürmektedir. Yabancı dil olarak İngilizce bilmektedir.

