ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

YÜKSEK ÇÖZÜNÜRLÜKLÜ MENZİL PROFİLİ KULLANARAK DENİZ HEDEFLERİNİN SINIFLANDIRILMASI

Özlem ERGÜN

ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA
2018

Her hakkı saklıdır

TEZ ONAYI

Özlem ERGÜN tarafından hazırlanan **"Yüksek Çözünürlüklü Menzil Profili Kullanarak Deniz Hedeflerinin Sınıflandırılması"** adlı tez çalışması 29/08/2018 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü Elektrik Elektronik Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Dr. Öğr. Üyesi Gökhan SOYSAL
Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı

Jüri Üyeleri :

Başkan : Doç. Dr. Umut ORGUNER
Orta Doğu Teknik Üniversitesi, Elektrik-Elektronik Mühendisliği Anabilim Dalı

Üye : Prof. Dr. Murat EFE
Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı

Üye : Dr. Öğr. Üyesi Gökhan SOYSAL
Ankara Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı

Yukarıdaki sonucu onaylarım

Prof. Dr. Atila YETİŞEMİYEN
Enstitü Müdürü

ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davranıldığı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

29/08/2018



Özlem ERGÜN

ÖZET

Yüksek Lisans Tezi

YÜKSEK ÇÖZÜNÜRLÜKLÜ MENZİL PROFİLİ KULLANARAK DENİZ HEDEFLERİNİN SINIFLANDIRILMASI

Özlem ERGÜN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik Elektronik Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Gökhan SOYSAL

Deniz yolu, uzun zamanlardan beri, ticari, sivil ve askeri amaçlı olarak sıklıkla kullanılmakta olup, deniz gözetimi vasıtasıyla kaçakçılık faaliyetlerinin tespiti, sahil güvenlik operasyonlarının takibi, liman ve deniz şeridi gözetimin sağlanması yapılabilmektedir. Deniz gözetimi sonucunda tespit edilen araçların türünün belirlenmesi, bu araçların manevra, sürat, taşınan yük vb. kabiliyetleri, tehdit seviyeleri, kaçakçılık faaliyetlere uygunluk durumları hakkında ön bilgi sağlamaktadır. Bu bağlamda deniz araçlarının sınıflandırılması önemli bir problem olup, tez çalışmasında deniz gözetleme radarına yakın bölgelerdeki hedeflerin sınıflandırılması üzerine araştırma yapılmıştır. Deniz araçlarının türünün tespit edilmesi için yaygın olarak sentetik açıklıklı radarlar (SAR) kullanılmaktadır. Ancak bu radarlar kullanılarak yapılan sınıflandırma işlemleri yüksek maliyet ve işlem yüküne sahiptir. Bu tezde, deniz hedeflerinin sınıflandırılması için yüksek menzil çözünürlüklü (HRR) radar kullanılmıştır. HRR radarlar, SAR gibi hedefin silueti hakkında bilgi vermemektedir ve hedefin oryantasyonuna göre elde edilen veri değişmektedir. Ancak her açıdan ve kısa sürede veri toplanabilmesi, işlem yükünün az ve maliyeti düşük olmasından dolayı sınıflandırma çalışmalarında tercih edilmektedir. Tez çalışmasında, sınıflandırma kavramı, sınıflandırıcılar, öznitelik kavramı, yüksek çözünürlüklü menzil profili, ve bu profilden öznitelik çıkartılması ve öznitelik seçimleri incelenmiştir. HRR radar simülatörü çıktıları kullanılarak radar yakın bölgeler için balıkçı teknesi, sahil güvenlik gemisi, kargo gemisi ve orta boy tanker hedeflerinin farklı sınıflandırıcılar kullanılarak elde edilen sınıflandırma başarımları incelenmiş ve tartışılmıştır. Benzetim sonuçları, özellikle ortalama RCS, RCS standart sapması ve uzunluk özniteliklerinin sınıflandırma başarımında baskın rol oynadığını göstermiştir. Uzunluk özniteliği gemi oryantasyonundan çok fazla etkilendiğinde, hedef takibinden elde edilen kinematik bilgiler kullanılarak daha doğru hesaplanması için bir yaklaşım önerilmiş ve bu yaklaşım ile sınıflandırma başarımının arttığı gösterilmiştir.

Ağustos 2018, 97 sayfa

Anahtar Kelimeler: yüksek menzil çözünürlüklü radar, yüksek çözünürlüklü menzil profili, gemi sınıflandırılması, öznitelik çıkartılması, güdümlü sınıflandırma

ABSTRACT

Master Thesis

SHIP TARGET CLASSIFICATION USING HIGH RANGE RESOLUTION PROFILE

Özlem ERGÜN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Electrical and Electronics Engineering

Supervisor: Asst. Prof. Dr. Gökhan SOYSAL

In this thesis, the classification of ship targets was investigated. The sea route has long been used for commercial, civil and military purposes. The detection of smuggling activities, search and rescue operations, control of ports and coast line security can be made by sea surveillance. Determination of the type of sea vehicles (targets) via sea surveillance can provide preliminary information about vehicle kinematics (maneuver capability, speed, etc.) and vehicle attitude (load, thread level, potential of smuggling and piracy). In this context, classification of sea vehicles is an important problem and, this thesis study has been focused on classification of the ship targets located in the area close to the sea surveillance radar. Synthetic aperture radars (SAR) are widely used to determine the type of the sea vehicles. However, using SAR for classification is costly and its computation load is high as well. In this thesis, high range resolution radar data where range profile of the target changes with the orientation is used for classification instead of SAR where image of the target can be obtained. High range resolution radar is preferred in classification studies due to the fact that one can collect data in every aspect of the target and in a short period of time and, it has computation load. In this thesis, the concept of classification, classifiers, features, feature extraction and selection and, high range resolution profile has been investigated. Classification of fishing boat, coaster, bulk carrier and medium tanker has been studied by using simulated high resolution range profiles and, performance of both classifiers and the features obtained from data has been investigated and discussed. Simulation results was shown that the features mean RCS, standard deviation of the RCS and, length have major influence on classification performance. Since the length feature highly depends on orientation of the ship, a new approach based on ship kinematic information obtained from the target tracking has been proposed for calculating length feature more accurately and, it was shown that calculation of the length feature using kinematic information increases classification performance.

August 2018, 97 pages

Key Words: high range resolution radar, high resolution range profile; ship classification; feature extraction; supervised classification.

TEŞEKKÜR

Çalışmalarımı yönlendiren ve araştırmalarım sırasında bilgi, öneri ve yardımlarını sunarak gelişmeme katkıda bulunan danışman hocam sayın Dr. Öğretim Üyesi Gökhan SOYSAL'a, (Ankara Üniversitesi Elektrik ve Elektronik Mühendisliği Anabilim Dalı Öğretim Üyesi) çalışmalarıma olan katkılarından dolayı Tevfik Bahadır SARIKAYA'ya, tez çalışmalarım süresince desteklerini hep hissettiğim başta ablam Elif ERGÜN TUNCER ve abim İbrahim ERGÜN olmak üzere tüm aileme teşekkür ederim.

Özlem ERGÜN
Ankara, Ağustos 2018

İÇİNDEKİLER

TEZ ONAYI
ETİK i
ÖZET ii
ABSTRACT iii
TEŞEKKÜR iv
KISALTMALAR DİZİNİ vii
ŞEKİLLER DİZİNİ viii
ÇİZELGELER DİZİNİ ix
1. GİRİŞ….. 1
2. KURAMSAL TEMELLER 7
2.1 Radar ile Hedef Tanıma ve Yüksek Çözünürlüklü Menzil Profili 9
2.2 Sınıflandırma İşlemi Temelleri 13
2.3 Gemi Sınıflandırması için Öznitelik Seçimi ve Elde Edilmesi 17
2.3.1 Öznitelik seçilmesi 21
2.4 Gemi Sınıflandırması için Sınıflandırıcı Yaklaşımları 24
2.4.1 K-En yakın komşu kuralı 25
2.4.2 Bayes sınıflandırıcısı 29
2.4.3 Bayes ağı sınıflandırıcısı 33
2.4.4 Karar ağaçları 35
2.4.5 Çok katmanlı algılayıcı 39
2.4.6 Destek vektör makineleri 41
3. MATERYAL VE YÖNTEM 46
3.1 Radar Simülatör Özellikleri 46
3.2 WEKA Temel Özellikler 50
3.2.1 WEKA sınıflandırıcı başarı parametreleri 51
4. ARAŞTIRMA BULGULARI 54
4.1 Özniteliklerin Çıkartılması ve Sınıflandırma Sonuçları 54
4.2 Gürbüz Özniteliklerin Tespit Edilmesi ve Sınıflandırma Sonuçları 74
4.3 Uzunluk Özniteliğinin Gemilerin Baş Açısına Göre Hesaplaması ve Sınıflandırma Sonuçları 81

4.4 Radar Tespit Olasılığının Sınıflandırma Başarısına Etkisi..... 88
5. TARTIŞMA VE SONUÇ..... 90
KAYNAKLAR 94
ÖZGEÇMİŞ..... 97

KISALTMALAR DİZİNİ

FMT	Fourier-Mellin Dönüşümü
GSFS	Genelleştirilmiş İleriye Doğru Sıralı Seçim
HRRP	Yüksek Çözünürlüklü Menzil Profili
HRR	Yüksek Çözünürlüklü Menzil
KNN	K- En Yakın Komşu
LOS	Radar Görüş Hattı
MLP	Çok Katmanlı Algılayıcı
NB	Naive Bayes
RCS	Radar Kesit Alanı
SFS	İleriye Doğru Sıralı Seçim
SNR	Sinyal Gürültü Oranı
SVM	Destek Vektör Makineleri

ŞEKİLLER DİZİNİ

Şekil 2.1 Hedef kimliği belirleme blok şeması 8
Şekil 2.2 Menzil çözünürlüğünün yansıma sinyalinin elde edilmesine etkisi 10
Şekil 2.3 Menzil çözünürlüğünün hedefin siluet bilgisinin edinilmesine etkisi 10
Şekil 2.4 Temsili bir gemi için yapısal özelliklerin elde edilen profile etkisi 11
Şekil 2.5 Sınıflandırma işlemlerinde model eğitim şeması 16
Şekil 2.6 Sınıflandırma işleminde tahmin sürecinin örnek akış şeması 17
Şekil 2.7 Tez çalışmasında kullanılan sınıflandırıcılar 25
Şekil 2.8 Test verisine k=3 için bulunan en yakın komşular 26
Şekil 2.9 Bayes Ağı öznitelik örnek ilişki yapısı çizgesi 34
Şekil 2.10 Karar Ağacı öznitelik uzayının bölünmesi örneği 37
Şekil 2.11 Şekil 2.10'da verilen problem için oluşturulan karar ağacı 38
Şekil 2.12 Üç katmanlı örnek MLP gösterimi 40
Şekil 2.14 Destek vektörleri kullanılarak en iyi sınırın bulunması örneği 43
Şekil 3.1 Radar simülatöründen elde edilen örnek benzetim verileri 49
Şekil 3.2 Başarı oranlarının temsili gösterimi 52
Şekil 4.1 Radar-Gemi geometrisi temsili gösterim 55
Şekil 4.2 Kargo gemisi için farklı kerteriz açıları için elde edilen simülatör çıktıları 56
Şekil 4.3 Hizalama işlemi uygulaması sonrasında profilin değişimi 58
Şekil 4.5 Uzunluk özniteliğinin çıkarımı blok şeması 60
Şekil 4.6 Farklı konumlar ve baş açıları için uzunluk özniteliğinin değişimi 62
Şekil 4.7 Faklı konum ve baş açıları için dağıtıcı özniteliği değişimi 64
Şekil 4.8 Faklı konum ve baş açıları için simetri özniteliği değişimi 65
Şekil 4.9 Faklı konum ve baş açıları için ortalama RCS özniteliği değişimi 66
Şekil 4.10 Faklı konum ve baş açıları için ortalama SNR özniteliği değişimi 67
Şekil 4.11 Faklı konum ve baş açıları için standart sapma RCS özniteliği değişimi 69
Şekil 4.12 Faklı konum ve baş açıları için standart sapma SNR özniteliği değişimi 70
Şekil 4.7 Baş açısı 45°için geminin radar göre duruşu 83
Şekil 4.8 Baş açısı 75°için geminin radar göre duruşu 83
Şekil 4.9 Baş açısı 75° için örnek uzunluk hesaplama gösterimi 84
Şekil 4.10 Baş açısı 135°için geminin radar göre duruşu 85

ÇİZELGELER DİZİNİ

Çizelge 2.1 Literatürde uygulanan bazı ön işleme yöntemleri 12
Çizelge 2.2 İçecek sınıflandırma problemi için örnek veri sunumu 14
Çizelge 2.3 Bayes karar kuralları 30
Çizelge 3.1 Radar simülatörü ve veri özellikleri 47
Çizelge 3.2 Sınıflandırma için kullanılan hedef özellikleri 48
Çizelge 3.3 WEKA'da tanımlı parametrelerin açıklaması 53
Çizelge 3.4 WEKA kullanılan değişkenlerin matematiksel ifadesi 53
Çizelge 4.1 Tüm öznitelikler kullanıldığında elde edilen sınıflandırma başarıları 72
Çizelge 4.2 Tüm öznitelikler kullanıldığında sınıflandırıcıların doğruluk analizi 73
Çizelge 4.3 Çeşitli öznitelik kümelerinin sınıflandırma performansları 76
Çizelge 4.4 Özniteliklerin bireysel sınıflandırma performansları hata matrisleri 78
Çizelge 4.5 Gemilerin baş açılarına göre elde edilen profillerden çıkartılan uzunluk değerleri 78
Çizelge 4.6 Baş açısının uzunluk özniteliğinin sınıflandırma başarısına etkisi 80
Çizelge 4.7 Öznitelik alt kümesi ile elde edilen sınıflandırma başarıları 81
Çizelge 4.8 Kinematik bilgi varlığında gemilerin baş açılarına göre uzunluk hesaplama 85
Çizelge 4.9 Kinematik bilgilerden hesaplanan uzunluğun yer altığı öznitelik alt kümesinin sınıflandırma sonuçları 86
Çizelge 4.10 Simetri, Dağıtıcı ve Uzunluk öznitelik kümesinin farklı uzunluk hesaplama yöntemlerine göre sınıflandırma başarısının değişimi 87
Çizelge 4.11 Bayes sınıflandırıcısının radar tespit olasılığına göre sınıflandırma başarısı 89

1. GİRİŞ

Radar, nesnelerin tespit edilmesi ve konumlarının belirlenmesi amacıyla kullanılan radyo dalgaları yayan elektronik bir cihazdır. Modern radar sistemleri, elektromanyetik enerjiyi darbeler halinde yollamaktadır. Bu dalgalar, herhangi bir engel bulunmayan hava veya uzay içerisinde izlediği yolda bir nesne ile karşılaşıncaya dek hareketine devam eder. Elektromanyetik dalgalar, nesneye çarptıkları zaman, dalgaların içerdiği enerjinin bir kısmı nesne tarafından geri yansıtılırken, bir kısmı nesne tarafından sönümlenmektedir ve kalan kısmı radara geri iletilmektedir (Tait 2005). Yakın zamana kadar radar, hedefleri tespit ve takip eden bir sensör olarak kullanılmaktaydı. Yüksek menzil çözünürlüğüne sahip radarların geliştirilmesiyle ortamda bulunan nesnelerden yansıyan enerjiden hedefin türüne ait bilginin de çıkarılabilmesi mümkün olmuştur. Nesneden yansıyan bu sinyal yüksek çözünürlüklü menzil profili (High Resolution Range Profile - HRRP) olarak adlandırılmaktadır. Yüksek çözünürlüklü menzil profili, radarın görüş hattı (line of sight -LOS) üzerinde meydana gelen saçılma sinyallerinin toplanması ile elde edilen (Du vd. 2005) hedefin tek boyutlu görüntüsünü ifade etmektedir. HRRP radyal hat boyunca hedefin saçılım merkezinin dağılım bilgisini sağladığından (Liu vd. 2011a), hedefin iskeleti ile ilgili bilgi içermektedir. HRR profilinden elde edilen bu bilgi kullanılarak, öznitelik çıkarılması ve hedeflerin sınıflandırılmasına yönelik çalışmalar son dönemlerde yoğunlaşmıştır.

Radarlar yetenekleri dâhilinde çok uzun menziller arasında sinyal iletimi ve alımı yapabilmektedir. Hem sivil hem de askeri alanda uzun menzilli iletişim ve hedef tespitinin önemi çoktur, bunun yanında tespit edilen hedefin sınıflandırılması da bir o kadar önem arz etmektedir. Hedefin daha önceden sınıflandırılması, ilgili hedefe nasıl tepki verileceği, askeri durumlar için hedefin çeşidine göre hangi donanımlarla müdahale edileceği gibi ön hazırlıkların yapılmasına olanak sağlamaktadır. Radarlar birçok koşulda çalışabilen cihazlar olduklarından dolayı radar kullanılarak kara araçları, hava araçları, deniz araçları sınıflandırılabilmektedir. Bu sebeple radar ile hedef sınıflandırma farklı çevre ortamları ve farklı çeşitlilikteki hedefler için yapılabilmektedir.

Tez çalışması kapsamında, deniz hedeflerinin sınıflandırılması incelenmiştir. Deniz yolu, uzun zamanlardan beri, ticari, sivil ve askeri amaçlı olarak sıklıkla kullanılmaktadır. Etrafı denizle veya okyanusla çevrili olan ülkelerin deniz trafiğini ve karasuları sınırlarını kontrol etmesi kaçınılmazdır. Deniz gözetimi ile deniz yolu kullanılarak yapılan kaçakçılık faaliyetlerinin tespiti, arama kurtarma ve sahil güvenlik operasyonlarının takibi, ticari faaliyetlerin olduğu limanların kontrolü, yurt güvenliğinin ve deniz şeridi gözetimin sağlanması yapılabilmektedir. Bu kontrollerde tespit edilen deniz aracının türü büyük önem arz etmektedir. Tespit edilen deniz aracının sınıflandırma süreci sonucunda elde edilecek bilgi; kimliği bilinmeyen araçların manevra, sürat, taşınan yük ve benzeri kabiliyetleri, tehdit seviyeleri, kaçakçılık ve korsanlık gibi faaliyetlere uygunluk gibi özelliklerin belirlenmesi için ön bilgi sunacağından büyük önem arz etmektedir.

Deniz gözetim radarının ortama yaptığı yayına karşılık ortamdaki hedeflerden yansıyan sinyallerin modellenmesiyle bu sinyallerden elde edilen menzil profillerinden öznitelik çıkartılması ile hedef sınıflandırılması yapılmaktadır. Farklı türlerdeki deniz araçları farklı yapı malzemelerine ve tasarımlara sahip olduklarından dolayı radar yayınına karşılık geri yansıttıkları sinyaller birbirinden farklı olmaktadır. Örneğin bir balıkçı teknesinin geri yansıttığı sinyal ile bir yük gemisinin geri yansıttığı sinyalden elde edilen profiller birbirinden farklılık göstermektedir. Benzer şekilde, aynı tasarıma sahip olan aynı tür gemiler, örneğin iki farklı tersanede üretilmiş yük gemisinin menzil profilleri benzer özellikler göstermektedir. Bu sebeple HRR profili gemilerin sınıflandırılması için temel veri kaynağı olarak kullanılmaktadır.

Tez kapsamında incelenen araştırma konuları aşağıda verilmektedir:

- HRR profilinden gürbüz özniteliklerin çıkartılması
- Sınıflandırma başarısına öznitelik sayısının ve türünün etkisinin araştırılması
- Yüksek performans gösteren sınıflandırıcıların tespit edilmesi
- Kinematik verilerin sınıflandırma başarısı üzerine etkilerinin araştırılması

HRR profili, çevre ve sistem gürültüsü, menzil değeri, radarın çözünürlüğü ve darbe şekli, deniz aracının radara (gözlemciye) göre olan konumu (Slomka vd. 1999), menzil göçü (range migration), benek dalgalanmaları (speckle fluctation) (Lu vd. 2011) ve benzeri faktörlerden etkilemektedir. Bu durum aynı türdeki aracın HRR profillerinin birbirlerinden çok farklı çıkmasına dolayısı ile aynı deniz aracının çok farklı şekilde sınıflandırılmasına sebep olabilmektedir. Sınıflandırma işleminin her koşulda maksimum doğrulukta yapılabilmesi için HRR profilinin iyileştirilerek öz bir temsilinin elde edilmesi gerekmektedir. Tez çalışması kapsamında, çevre ve sistem gürültüsü, menzil değeri ve deniz aracının radara (gözlemciye) göre olan konumuna ilişkin etkenlerin iyileştirilmesi adına çalışmalar yapılmıştır. HRR profilinin iyileştirmesi üzerine yapılan işlemler önişleme olarak adlandırılarak sunulmuştur. Literatürde HRR profilinin iyileştirilmesi için çok sayıda önişleme yöntemi yer almaktadır ve bu yöntemlerden öne çıkanlar Bölüm 2 Kuramsal Temeller başlığı altında verilmiştir.

Öznitelik çıkartma terimi, geniş bir veri kümesini tanımlamak için gerekli olan veri miktarını azaltmak olarak açıklanabilmektedir. Öznitelikler, üzerinde inceleme yapılan veriyi temsil eden özet bilgiler olmalıdır. Karmaşık ve büyük verilerin analizi yapılırken karşılaşılan başlıca sorunlardan biri, verinin işlenmesi sırasında büyük miktarda bellek ve hesaplama gücü ihtiyacının olmasıdır. Bunun yanı sıra sınıflandırma yöntemleri özelinde, veriyi temsil eden parametre sayısının büyük olması sınıflandırıcıların eğitim aşamasına aşırı yük oluşturmakta ve yeni veri örneklerinin genellemesinin zayıf bir şekilde yapılmasına neden olabilmektedir (Alpaydın 2010). Sınıflandırma algoritmalarının eğitim kümesi için yüksek performansta çalışması ve sınıflandırma başarımının arttırılması amacıyla elde edilen profil verisinden öznitelikler çıkartılmaktadır.

Sınıfları birbirinden ayıran özniteliklerin değerleri ayrıklaştıkça, sınıflandırma işleminin başarısı artmaktadır. Bu nedenle özniteliklerin, ilgilenilen veriyi doğru temsil edecek şekilde çıkartılması ile sınıflandırma başarısı arasında güçlü bir ilişki bulunmaktadır. HRR profili temeline dayalı öznitelik çıkarımında, hedefin radara göre duruş açısı, menzil profilinde büyük değişikliklere sebep olmaktadır ve aynı hedef için farklı açılarda çok farklı menzil profilleri elde edilebilmektedir. Duruş açısına bağlı olan

menzil profili değişkenliği ve dolayısı ile değişmez öznitelik tanımlama sorunu sınıflandırma başarımını doğrudan etkilemektedir. Tez çalışması kapsamında bu problem üzerine odaklanılmıştır ve problemin sınıflandırma başarımına olan etkileri ve başarımı arttırıcı yöntemler araştırılarak, özniteliklerin yüksek doğrulukta çıkartılabilmesi için uygulanacak yöntemler Bölüm 4 Araştırma Bulguları başlığı altında sunulmuştur.

Sınıflandırma işleminde öznitelik sayısının arttırılmasının başarımı arttırıcı yönde etkisinin olacağını söylemek her zaman doğru değildir. Aksine, bazı durumlarda öznitelik sayısının arttırılması sınıflandırma başarımını düşürmektedir. Bu sebeple, ölçüm maliyeti ve sınıflandırma doğruluğu için öznitelik sayısının olabildiğince az tutulması gerekmektedir. Sınıflandırma işleminde sınıflandırıcının seçilmesi kadar kullanılacak özniteliğin seçimi de önemlidir. Hangi özniteliğin daha iyi ayırıcı özellik göstereceği ya da kullanılacak öznitelik sayısı, araştırılması gereken konulardandır. Öznitelik seçilmesi, sınıflandırma algoritması için öznitelikler içerisinden en iyi alt kümenin seçilmesidir. Seçilen bu öznitelikler sınıflar arasında bulunan sınırın genişliğini arttırmakta ve böylece sınıflandırma başarısı da artmaktadır. Öznitelik seçilmesi üzerine ulaşılan literatürde yapılmış çalışmalar Bölüm 2 Kuramsal Temeller başlığı altında sunulmuştur.

Literatürde kullanılan birçok sınıflandırıcı algoritması bulunmaktadır. Her bir algoritma farklı ihtiyaçlar doğrultusunda geliştirilmiş, çeşitli iç dinamiklere sahiptir. Bazı algoritmalar yüksek başarımı vaat ederken, hafıza kapasitesini ve zamanı daha fazla kullanmaktadır. Bazı algoritmalar çok iyi tanımlanmış problemler için uygun iken bazı algoritmalar problem modelinin değişimine uyum sağlayabilmektedir. Bu sebeple sahip olunan probleme ve kısıtlara uygun sınıflandırıcı bulunması, sınıflandırma sürecindeki en son ve en önemli adımdır. Bu kapsamda, tez çalışmasında deniz araçlarının sınıflandırılması için K-En Yakın Komşu, Bayes, Bayes Ağı, Karar Ağaçları ve Çok Katmanlı Algılayıcı, Destek Vektör Makineleri sınıflandırıcıları kullanılmıştır ve sınıflandırıcıların performansları sınıflandırma başarılarına göre karşılaştırılmıştır. Sınıflandırıcılar hakkında kuramsal bilgiler Bölüm 2 Kuramsal Temeller başlığı altında, performans değerlendirmeleri Bölüm 4 Araştırma Bulguları başlığı altında verilmiştir.

Radar ile hedef sınıflandırılması uygulamalarında, kinematik verilerin katkısını kullanarak sınıflandırma yapılmasına dair ulaşılan literatürde bir çalışma bulunmamaktadır. Hedefe ait kinematik verilerin, özniteliklerin değerlendirilmesi aşamasında gürbüzlüğü arttıracak bir bilgi sağladığı düşünülmektedir. Bu sebeple tez çalışması kapsamında, deniz araçlarının gözlemciye göre olan baş açısı bilgisi kullanılarak deniz araçlarının uzunluk özniteliğinin hesaplanması üzerine çalışılmış olup, geliştirilen yöntem Bölüm 4 Araştırma Bulguları başlığı altında detaylı olarak açıklanmıştır.

Tez çalışması süresince yapılan araştırmalarda, uzunluk özniteliğinin kinematik bilgiler varlığında çok daha doğru hesaplandığı ve bunun sınıflandırma başarısına büyük oranda etki ettiği ortaya konmuştur. Sınıflandırma başarısına en çok etki eden özniteliğin uzunluk olduğu ve bunu radar kesit alanının (radar cross section - RCS) ve sinyal gürültü oranının (signal-to-noise ratio – SNR) ortalama değerlerini içeren özniteliklerin takip ettiği gözlenmiştir. Dağıtıcı ve simetri özniteliklerin tez çalışmasında kullanılan iki boyutlu gemi modellemesi için yeterli bilgi vermediği bu nedenle performanslarının düşük olduğu gözlenmiştir. Benzer şekilde, RCS ve SNR'ın standart sapmalarının hesaplanması sonucu elde edilen özniteliklerin, sınıflara göre değişimi incelendiğinde ve bireysel sınıflandırma performansına bakıldığında yeterli katkı sağlayamadıkları gözlenmiştir.

Bu tezin literatüre katkıları ve özgün değeri aşağıdaki gibi özetlenebilir:

- Ulusal fayda sağlayacak bir çalışma olması: Ülkemizin üç tarafının denizlerle çevrili olması sivil taşımacılık, ticari faaliyetler ve askeri durumlar için sahil şeridinin kontrol altında tutulması gerekmektedir. Tespit edilen hedeflerin sınıflandırılması bu bakımdan kıymet arz etmektedir.
- Gemilerin kinematik bilgileri kullanılarak başarımın arttırması: Ulaşılan literatürde hedefe ait kinematik bilgilerin sınıflandırma sürecine katkısı bulunamamıştır. Araştırma sonucunda, gemilerin baş açısı ve radardan uzaklıklarına ait kinematik bilgiler kullanılarak uzunluk özniteliğinin daha doğru hesaplandığı ortaya konmuştur. Bunun bir çıktısı olarak, sınıflandırma başarısını

arttırmıştır. Bu sonuç, tez çalışmasının literatüre katkısı olarak değerlendirilmektedir.

Tez, beş bölümden oluşmaktadır. Birinci bölümde, tez çalışmasının amacı ve kapsamı belirtilmiştir. İkinci bölümde, tez çalışmasında incelenen problem hakkında mevcut literatür incelenmiştir. Sürecin tüm parçaları tek tek incelenerek sisteme genel bir bakış elde edilmiştir. Profilin elde edilmesi, işlenmesi, öznitelik çıkartılması ve sınıflandırıcılar hakkında temel teori ve gerekli matematiksel altyapı verilmiştir.

Üçüncü bölümde, tezde kullanılan materyal ve yöntemler anlatılarak tez çalışmasında kullanılan ve veri edinimini sağlayan radar simülatörünün özellikleri hakkında bilgi verilmiştir. Sınıflandırma sonuçlarını sağlayan WEKA aracı tanıtılmış ve özellikleri sunulmuştur.

Bilgisayar ortamında gerçekleştirilen benzetim çalışmaları ve sonuçları dördüncü bölümde sunulmuş ve elde edilen sonuçlar yorumlanmıştır. Altı farklı sınıflandırıcının, sınıflandırma performansı karşılaştırılmış ve tartışılmıştır. Beşinci bölümde, genel bir değerlendirme yapılarak tez çalışmasında elde edilen sonuçlar özetlenmiş ve gelecekte yapılabilecek çalışmalara yönelik öngörüler sunulmuştur.

2. KURAMSAL TEMELLER

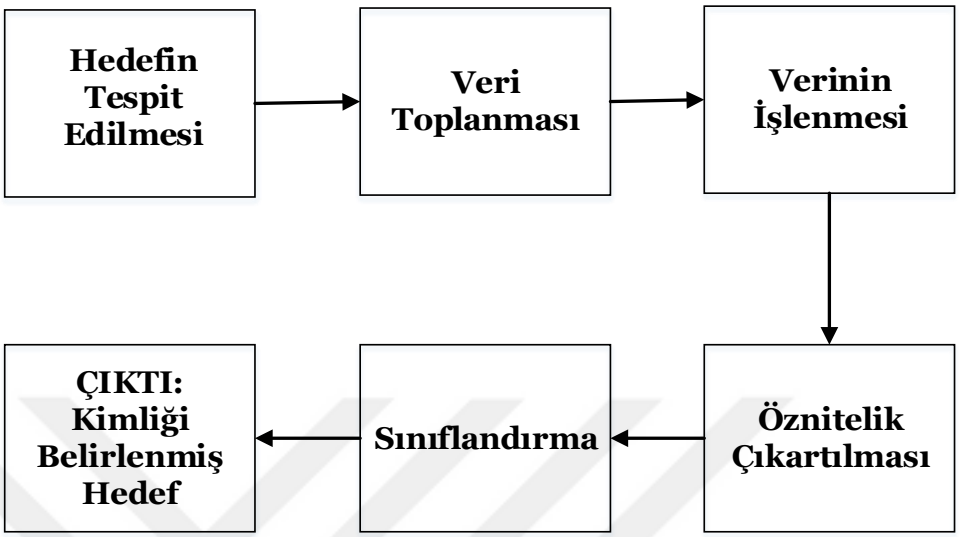
Bu bölümde tez kapsamında kullanılan yöntemlerin kavranması amacıyla kuramsal bilgilere yer verilmiştir. Öncelikle yüksek menzil çözünürlüklü radar ile hedef tanıma genel olarak özetlenerek temel kavramlar açıklanmıştır. Alt bölümlerde kavramlar daha detaylı olarak incelenmiştir.

Radar, cisimlerin varlıklarının ve konumlarının tespit edilmesinde kullanılan elektromanyetik sistemdir (Skolnik 1981). Radarlar temel olarak bir alıcı bir de verici antene sahiptir, verici anten vasıtasıyla ortama elektromanyetik sinyal yayını yapılmaktadır. Elektromanyetik sinyaller, elektriksel iletken bir yüzeye çarpmaları halinde, sinyalin enerjisinin bir kısmı cisim tarafından sönümlenirken, kalan kısmı yansımaktadır. Radarda bulunan alıcı anten yayın yapılan ortamdan yansıyan sinyalleri toplamaktadır. Toplanan sinyalin işlenmesi yoluyla cisimlerin, menzil, yön ve hız gibi durumları elde edilmektedir.

Geçmiş zamanlarda radarlar; ilgilenilen hedefin gözlemciye göre olan hızı, yönü ve menzilinin ölçülmesi amacıyla kullanılmaktaydı (Skolnik 1981). Bu bilgiler, hedef ile ilgili çok önemli olsa da hedefin kimliğini belirlemek için her zaman yeterli değildi. Hedefin kimliğinin belirlenmesi, bir tipteki hedefin diğer tipteki hedeften ayırt edilip tanınması olarak açıklanabilmektedir. Deniz hedefleri için tespit edilen geminin tanker mi veya yük gemisi mi olduğunun belirlenmesi buna örnek olarak verilebilir. Bu yetenek, hedef sınıflandırma veya hedef tanıma olarak ifade edilmektedir.

Hedefin kimliğinin belirlenmesi için birçok aşamadan geçilmektedir. Bu aşamalar Şekil 2.1'de verilmektedir. İlk iki aşama hedefin tespit edilmesi ve hedeften veri toplanmasıdır. Hedefin tespit edilmesi ve veri toplanmasında çok farklı sensörler kullanılabilmektedir. Örneğin hedef görülebilir menzilde ise kamera vb. optik sistemler kullanılarak hedef tespit edilip elde edilen görüntülerle hedef hakkında veri toplanabilmektedir. Genellikle hedef ile gözlemci arasındaki mesafe, optik sistemler kullanılamayacak kadar uzaktır. Bu nedenle uzak menzilde etkin olarak çalışabilecek

bir sensör seçilmesi gerekmektedir. Uzak mesafede bu amaçla kullanılabilecek uygun araç radardır.



Şekil 2.1 Hedef kimliği belirleme blok şeması

Sensör tarafından toplanan ve üzerinde hiçbir işlem yapılmamış veri, ham veri olarak adlandırılmaktadır. Ham verinin, gürültü sinyallerinden temizlenmesi, düzenlenmesi, girdi olarak verileceği sistemin kullanım biçimine dönüştürülmesi gibi işlemlerden biri veya daha fazlası uygulanarak işlenmiş veri elde edilmektedir. Tüm bunların yapıldığı aşama da veri işleme olarak adlandırılmaktadır.

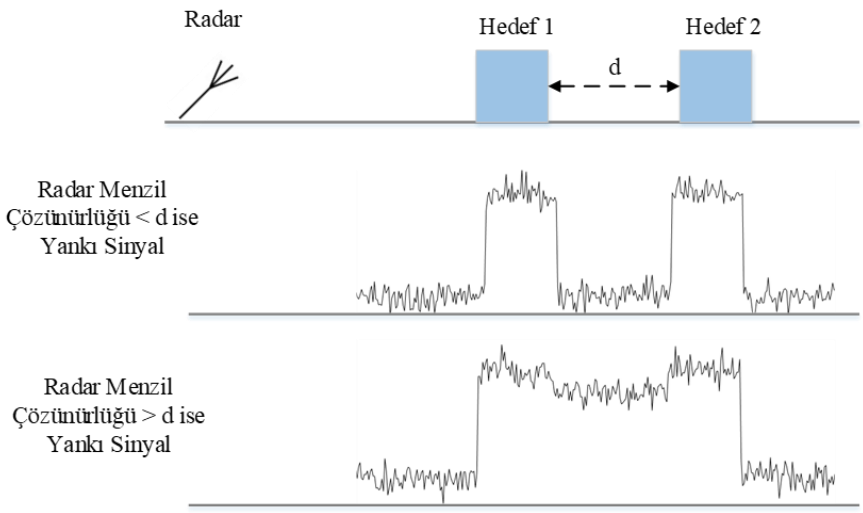
Veri işlendikten sonra, öznitelik çıkartılması aşamasına geçilmektedir. Bu çalışma için, radar dönüşünden elde edilen sinyalden karakteristik özellikler çıkartılmaya çalışılmaktadır. Hedefleri bir birinden ayırt edebilmek için bu karakteristik özellikler sınıflandırıcıya girdi olarak verilmektedir. Sınıflandırıcı bu bilgilere göre seçme koşullarını belirlemekte ve çıktı olarak kimliği belirlenmiş hedef bilgisini sunmaktadır. Alt bölümlerde hedef kimliği belirleme aşamaları detaylı olarak sunulmaktadır:

2.1 Radar ile Hedef Tanıma ve Yüksek Çözünürlüklü Menzil Profili

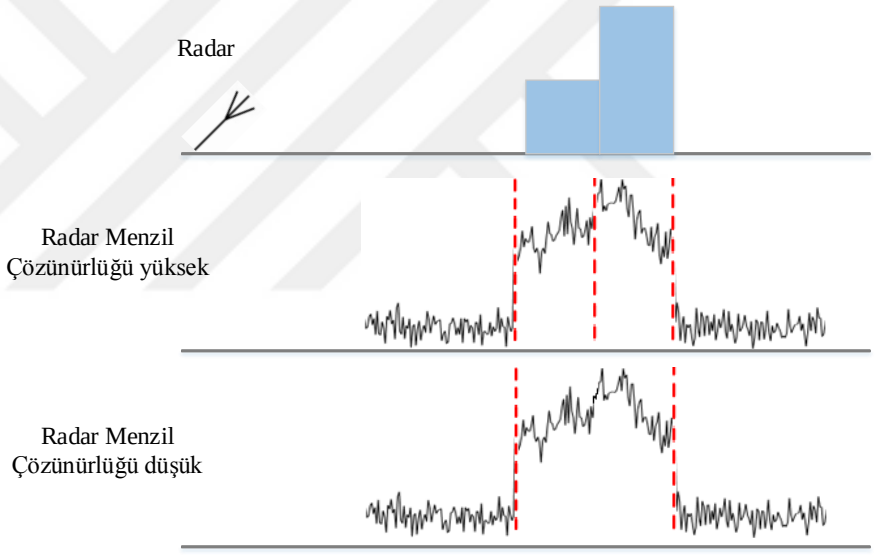
Genel olarak radar ile hedef tanıma, hedeften yansıyan sinyalin ayrıntılı olarak incelenmesi ile gerçekleştirilmektedir. Bunu gerçekleştirebilmek için hedef tespitinde kullanılan radarlara göre daha yüksek bir sinyal-gürültü oranı gerekmektedir (Skolnik 1981). Bunun sonucu olarak, sınıflandırma için gereken menzil çözünürlüğü, tespit için gerekenden daha fazla olmalıdır. Bu nedenle, çalışma kapsamında gemilerin sınıflandırılması problemi için yüksek menzil çözünürlüklü radar kullanılmaktadır.

Menzil çözünürlüğünün en genel tanımı, iki noktasal hedefin farklı iki hedef olarak algılanması için gereken, hedefler arası minimum uzaklıktır. Şekil 2.2'de aralarında d kadar uzaklık olan iki hedef, radar görüş hattı üzerinde bulunmaktadır. Eğer radarın, menzil çözünürlüğü d'den küçük ise şekilde görüleceği üzere yankı sinyalinden hedefler ayırt edilebilmektedir. Eğer radar çözünürlüğü d'den fazla ise bu durumda yankı sinyali tek bir hedeften dönüyormuş gibi elde edilmektedir bunun sonucu olarak tek hedef algılanmaktadır.

Radar menzil çözünürlüğü, hedeflerin ayırt edilmesinin yanında sınıflandırma için hedefin yapısal özelliklerine ait daha detaylı verinin elde edilmesine de olanak sağlamaktadır. Şekil 2.3'te radar görüş hattında yapısal özelliği şekil 2.2'de verilenlere göre farklı tek bir hedef bulunmaktadır. Şekilde yer alan kesikli çizgiler hedefin temsil edildiği hücre sayısını temsil etmektedir. Menzil çözünürlüğü yüksek olduğunda, hedef iki farklı hücre ile temsil edildiğinden daha detaylı analiz edilme ve buradan hedefin siluetine dair bilgi çıkartabilme şansı doğarken, çözünürlük düşük olduğunda hedef tek hücre ile temsil edilmektedir. Bu tek hücre analiz edilirken hedefin gerçek silueti çıkartılamamaktadır.

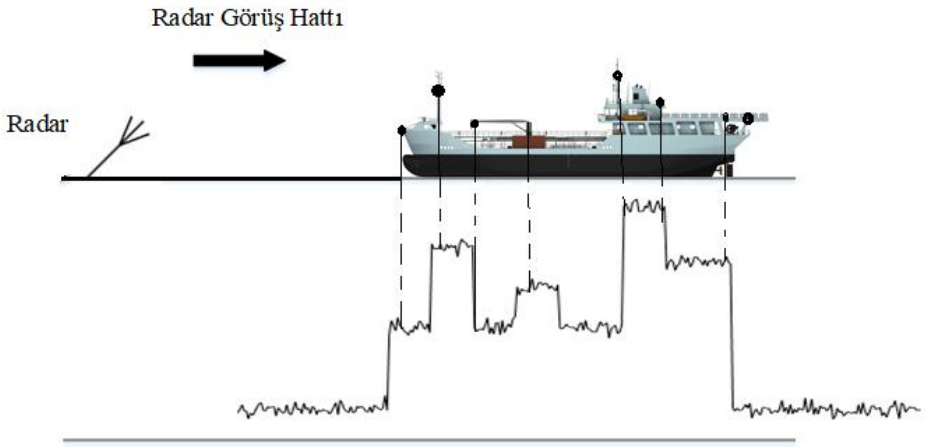


Şekil 2.2 Menzil çözünürlüğünün yansıma sinyalinin elde edilmesine etkisi



Şekil 2.3 Menzil çözünürlüğünün hedefin siluet bilgisinin edinilmesine etkisi

Yüksek menzil çözünürlüklü radarlar vasıtasıyla hedefin siluet bilgisi, yankı sinyalinden elde edilmektedir ve bu yankı sinyali literatürde yüksek çözünürlüklü menzil profili (high range resolution profile - HRRP) olarak adlandırılmaktadır. Şekil 2.4'te örnek bir gemi için yüksek menzil çözünürlüklü radar ile temsili olarak elde edilen menzil profili verilmektedir. Gemi üzerinde yer alan noktalar, radar tarafından iletilen sinyalin hedef üzerinde çarptığı noktaları göstermektedir. Her bir noktadan dönen sinyal ilgili menzile denk gelecek şekilde birleştirilmektedir.



Şekil 2.4 Temsili bir gemi için yapısal özelliklerin elde edilen profile etkisi

Şekil 2.4'te de görüldüğü üzere geminin direk, baca, radar anteni, kamara ve benzeri bölümlerinden geri saçılan sinyallerin genlikleri farklı olmaktadır. Gemiler farklı yapısal özelliklere göre inşa edildiklerinden dolayı farklı gemilerden elde edilen HRRP sinyalleri de farklı olmaktadır. Bunun sonucu olarak HRRP verisi işlenerek çıkartılan karakteristik özellikler karşılaştırılarak gemiler sınıflandırılabilmektedir.

Gemiden yansıyan HRRP sinyali birçok faktörden etkilenebilmektedir. Bu faktörler,

- Geminin radara (gözlemciye) göre olan oryantasyonu,
- Menzil,
- Radarın çözünürlüğü,
- Radarın darbe şekli,
- Çevresel ve sistem gürültüleri,

olarak sıralanabilmektedir. Gemi, radara göre dönüş hareketi yaparsa, elde edilen HRRP sinyalinin ölçeği değişmektedir (Slomka vd. 1999). Gemi ile gözlemci arasındaki menzil arttıkça yansıyan sinyalin gücü düşerken, gürültü seviyesi artmaktadır. Çevresel gürültüye örnek olarak hava koşulları, örneğin yağmur verilebilmektedir.

Yukarıda verilen HRRP'yi etkileyen faktörlerden en önemlisi, geminin yaptığı harekete veya radara göre olan duruşuna bağlı olarak geri dönüş sinyalinin değişime uğramasıdır. Geminin hareketi iki başlık altında incelenebilmektedir: öteleme hareketi ve dönme hareketi. Öteleme hareketinde bulunan geminin, radarın görüş hattına göre duruşu

değişmediği için elde edilen menzil profilleri de değişmemektedir. Ancak elde edilen sinyalin zarfı kaymaktadır (Mengdao ve Zheng 2001). Eğer gemi dönme hareketi yaparsa, bu durumda saçılımın radyal menzili değişmektedir. Saçılımın radyal menzilinin değişmesi iki farklı etkiye neden olabilmektedir; bunlar faz farkı ve sinyal zarfının yer değiştirmesidir (Mengdao ve Zheng 2001). Eğer sinyal zarfının yer değiştirmesi radarın menzil çözünürlüğünden fazla ise menzil göçü adı verilen problem ortaya çıkmaktadır. Menzil göçü, saçılım sinyalinin dağıldığı tüm menzil hücrelerinin değişmesi durumunu ifade etmektedir. Menzil göçü, doğrusal ölçeklendirmeye karşılık gelmektedir. Diğer bir ifadeyle; menzil profili görüş hattı, hedefin oryantasyonuna bağlı olarak sıkıştırılmış veya genişletilmiş biçimde olmaktadır (Inggs ve Robinson 1999).

Bu tip bozulmalardan dolayı HRRP; hedefin gözlemciye göre olan açısına, genlik ölçeğine ve zaman kaymasına karşı hassasiyete sahiptir. Hedefin farklı zaman, konum, menzil ve duruş pozisyonuna göre elde edilen HRRP'leri eşitlemek adına önişleme uygulanmaktadır. Bu kapsamda HRRP'nin daha gürbüz hale getirilmesi için ulaşılan literatürde ön işleme için çeşitli yaklaşımlar bulunmaktadır. Geleneksel olarak kullanılan yöntemler Fourier-Mellin dönüşümü, dalgacık dönüşümü ve istatiksel öznitelik çıkartma yöntemleridir (Junping ve Yi 2006). Literatürde geçen yöntemler çizelge 2.1'de özetlenmiştir.

Çizelge 2.1 Literatürde uygulanan bazı ön işleme yöntemleri

Ön İşleme Yöntemi	Kullanım Amacı
Fourier-Mellin Dönüşümü (FMT)	Fourier dönüşümünün zamanda değişmezlik özelliği kullanılarak geminin öteleme hareketi sonucunda meydana gelen sinyal zarfının kaymasının önüne geçilmektedir. Mellin dönüşümünün ölçeklendirme ile değişmezlik özelliği ile geminin dönüş hareketlerine bağlı olarak elde edilen daralan genişleyen profiller üzerinde iyileştirme yapılmaktadır
Dalgacık Dönüşümü	FMT uygulaması ile global spektrum elde edilmektedir. Sinyalin yerel özellikleri göz önüne alınarak gemi ile ilgili karakteristik özelliklerin çıkartılabilmesi dalgacık dönüşümü uygulanmaktadır
HRRP şablonu elde edilmesi	Bir zaman serisinden elde edilen profillerin istatiksel özelliklerinden faydalanılarak ilgili gemiye ait bir HRRP şablonu oluşturulmuştur. Profiller arası ilişkilerin tespit edilebilmesi için korelasyon ilişkisi araştırılmıştır.

FMT önişleme yönteminin detayları, Zwicke ve Kiss (1983), Inggs ve Robinson (1999) ve Liu vd. (2011b) çalışmalarından incelenebilmektedir. Dalgacık dönüşümün uygulanması Junping ve Yi (2006) çalışmasından ve HRRP şablonunun elde edilmesi Liu vd. (2011b)'den incelenebilmektedir.

Gemi ile radar arasındaki mesafe yeteri kadar uzak olduğunda tek bir radar huzmesinden geminin tamamına ait bir HRRP sinyali elde edilmektedir. Radarın tam bir tur taraması sonucunda gemi ile ilgili N tane HRRP sinyali elde edilmektedir ve çizelge 2.1'de verilen ön işleme yöntemleri aynı gemiye ait N tane sinyal üzerinde uygulanarak düzeltme işlemleri gerçekleştirilmektedir.

Çalışma kapsamında, yakın menzil gemilerin sınıflandırılması problemi incelendiği için literatürde geçen tek boyutlu HRRP sinyali, simülatör çıktısından elde edilememiştir. Bu nedenle radarın farklı merkez açılarına sahip huzmelerinden elde edilen veriler birleştirilerek bir kenarı kerteriz diğer kenarı menzil olacak şekilde iki boyutlu görüntü elde edilmiştir. Çizelge 2.1'de verilen ön işleme yöntemleri çalışma kapsamında uygulanmamıştır.

2.2 Sınıflandırma İşlemi Temelleri

İnsanoğlu veriyi tanıyabilmek için veriyi rengi, şekli, kokusu, ağırlığı, uzunluğu vb. özelliklerine göre analiz etmektedir ve verideki değişimleri takip ederek verileri ayırt edebilmektedir. Veri hacmi ve veriye ait özellikler arttıkça insanın bu verileri ayırt etmek için ortaya koyduğu yazılı kurallar ve analiz etme yeteneği kısıtlı kalmaktadır bu nedenle verideki değişimi analiz edebilecek otomatik sistemler geliştirilmiştir. Bu sistemlere en genel ifade ile makine öğrenmesi denilmektedir.

Makine öğrenmesi, insanoğlunun analiz etme, ayrıştırma ve anlama özelliklerinin matematiksel modeller kurularak ve yazılım dilleri kullanılarak bilgisayar ortamında taklit edilmesidir. Makine öğrenmesi günümüzde birçok alanda kullanılmaktadır, bunlara örnek olarak girilen anahtar kelimelere ilişik sonuçları sunan arama motorları,

elektronik postalarda istenmeyen postaların ayırt edilmesi, görüntü tanıma, öneri motorları, yazı ve konuşma anlama sistemleri verilebilmektedir.

Sınıflandırma, makine öğrenmesi yaklaşımlarından ortaya çıkan ve çıktı olarak etiket bilgisi sunan bir dalıdır. Örnek olarak, tez çalışmasında ilgilenilen problem için balıkçı teknesine ait veriler girildiğinde sonucun küçük boy gemi, tanker ile ilgili veriler girildiğinde sonucun çok büyük boy gemi çıkması verilebilmektedir. Özetlemek gerekirse, sınıflandırma en basit anlamda, verinin tanınması, ayrıştırılması, anlaşılması ve etiketlenmesi süreçleri olarak ifade edilebilmektedir.

Sınıflandırma süreci, bir bardak portakal suyu ile vişne suyunun ayırt edilmesi problemi üzerinden açıklanmıştır. İlk olarak bu iki içeceği birbirinden ayırt edecek özelliklerin tespit edilmesi gerekmektedir. Bu özellikler, öznitelik (feature) olarak adlandırılmaktadır ve bu problem için öznitelikler, içeceklerin rengi ve içerdikleri şeker oranı olarak seçilmiştir.

Sınıflandırma sürecinin ilk adımı veri toplanmasıdır bunun için yerel marketlerden çeşitli üreticilerin portakal ve vişne suları toplandığı düşünülsün. Ardından toplanan bu içeceklerden belirlenen öznitelikler olan renk ve şeker oranı bilgisi çıkartılmaktadır. Renk bilgisinin çıkartılması için spektrogram ve şeker oranı için hidrometre kullanıldığı düşünülerek çizelge 2.2'de verildiği gibi sonuçlar elde edilsin.

Çizelge 2.2 İçecek sınıflandırma problemi için örnek veri sunumu

Renk (nm)	Şeker Oranı (%)	Vişne veya Portakal Suyu?
610	7	Portakal suyu
690	10	Vişne Suyu
700	11	Vişne Suyu
590	6	Portakal suyu

Veriler toplandıktan ve öznitelikler çıkartıldıktan sonraki aşama, verilerin bir araya getirilmesidir, bu aşamada verilerin sayısı ve özellikleri modelin iyi kurulabilmesi için önemlidir.

Ölçümler alındıktan ve veriler bir araya getirildikten sonraki süreç, verinin hazırlanmasıdır. Veri hazırlanması, verilerin uygun bir yere yüklenmesi ve modelin eğitilmesi için hazır hale getirilmesi işlemlerini içermektedir. Genellikle veriler sıralandıktan sonra, sıraları rasgele bir şekilde dizilmektedir. Bunun yapılmasının nedeni, araştırılan problemin girdi olarak sunulan verinin hangi tür içecek olduğunun belirlenmesidir ve bu durum hangi içeceğin verisinin önce ya da sonra girdi olarak gelmesinden bağımsızdır. Bazı durumlarda verinin normalleştirilmesi, hataların düzeltilmesi vb. işlemlerin uygulanması gerekmektedir. Bu işlemler de veri hazırlama aşamasında yapılmaktadır.

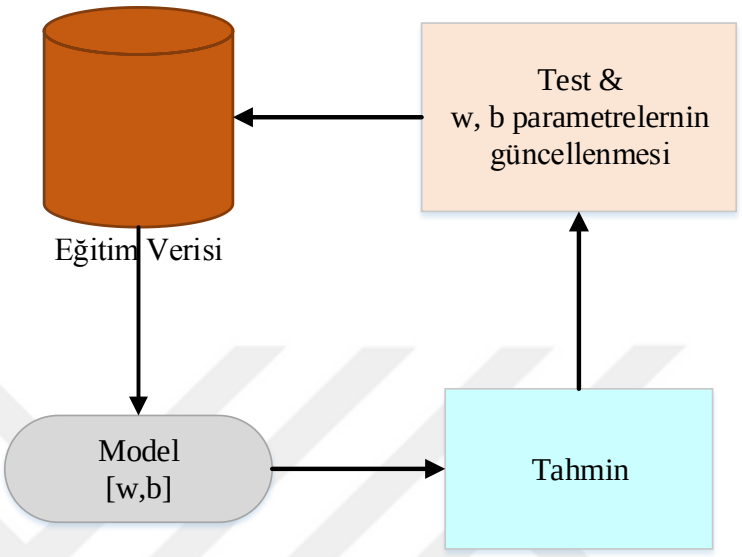
Bir sonraki aşama verinin analizini gerçekleştirecek modelin seçilmesidir. Yıllar boyunca araştırmacıların çalışmaları sonucunda birçok model ortaya konmuştur. Bu modellerden probleme uygun olacak bir tanesi seçilerek sürece devam edilmektedir.

Model belirlendikten sonraki aşama, modelin doğru tahmin yapabilmesini sağlamak için veri hazırlama aşamasında bir araya getirilen verilerle eğitilmesi işlemidir. Modelin eğitilmesi, tekrarlamalı bir işlemdir ve her defasında modelin daha iyi tahminde bulunması amaçlanmaktadır. Model daha önce de bahsedildiği üzere matematiksel bir alt yapıya sahiptir. İçecek türünün tespit edilmesi probleminde iki öznitelik olduğu için eşitlik (2.1)'de verilen doğrusal bir denklem model olarak seçilmiş olsun. Eşitlikte yer alan, y çıktı, x girdi, w ağırlık katsayısı ve b sapma değerini ifade etmektedir.

$$y = w * x + b \tag{2.1}$$

Bu modelde, eğitim sürecinde en uygun w ve b değerleri bulunmaya çalışılmaktadır. Buna göre herhangi bir içeceğe dair veri girişi yapıldığında denklemin çözülmesi ile elde edilecek değere göre içeceğin türü tahmin edilecektir. Eğitim aşamasında,

kullanılan verilerin çizelge 2.2'de verilen girdi değerleri sonucunda çıktının ne olacağı bilindiği için buna göre model parametreleri güncellenmektedir. Eğitim süreci kısaca şekil 2.5'te verildiği gibi işlemektedir.



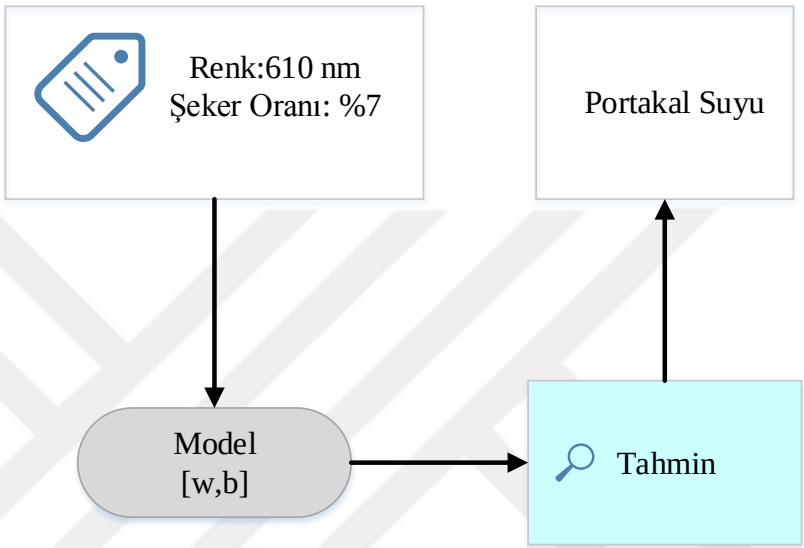
Şekil 2.5 Sınıflandırma işlemlerinde model eğitim şeması

Son aşama ise modelin değerlendirilmesi aşamasıdır. Bu aşamada, eğitim sonucunda evrimleşmiş model üzerinde eğitim amaçlı kullanılmayan harici bir veri girdi olarak verilerek model test edilmektedir. Böylece modelin daha önce karşılaşmadığı veriler karşısında nasıl bir performans gösterdiği değerlendirilmektedir. Bu aşama aslında modelin gerçek dünya şartlarında nasıl bir performans göstereceğine dair temsili bir sonuç vermektedir. Genellikle eğitim-test için veri oranları yüzde olarak 80-20 veya 70-30 olarak seçilmektedir.

Tüm bu sürecin son aşaması tahmin aşamasıdır, sürecin en başında amaçlanan bu içecek nedir sorusunun cevabının verilmesidir. Özetle, sınıflandırma süreci aşağıda verilen şekil 2.6'da olduğu gibi, girdi olarak verilen parametrelerin işlenmesi yoluyla cevabı aranan sorunun tahmin edilmesi ile sonuçlanmaktadır. Bunu gerçekleştirmek için aşağıda verilen adımlar sırasıyla uygulanmaktadır.

- Veriler toplanarak bir araya getirilmektedir,

- Veriler hazırlanmaktadır,
- Model seçilmektedir,
- Model eğitilmektedir,
- Test verisi kullanılarak model değerlendirilmektedir,
- Tahmin aşamasına geçilmektedir.



Şekil 2.6 Sınıflandırma işleminde tahmin sürecinin örnek akış şeması

Buradan sonraki bölümlerde, tez çalışması özelinde sınıflandırma süreci aşamaları verilmektedir. İlk olarak toplanan veriden gemilerin ayırt edilmesi için seçilen öznitelikler ve bunların veriden çıkartılma yöntemleri verilmektedir. Çalışma kapsamında kullanılan sınıflandırma modellerinin alt yapısı ve çalışma prensipleri verilerek bölüm tamamlanmıştır.

2.3 Gemi Sınıflandırması için Öznitelik Seçimi ve Elde Edilmesi

Öznitelik, gözlemlenen olayın veya cismin karakteristik veya tekil ölçülebilir özelliği olarak tanımlanmaktadır. Bölüm 2.2'de verilen portakal suyu ile vişne suyunun ayırt edilmesi probleminde olduğu gibi içeceklerin en büyük karakteristik ve ölçülebilir özelliklerinden biri renkleriydi. Bununla beraber bu özniteliğin sağlaması gereken bazı

kıstaslar bulunmaktadır. Bunlardan en önemlileri, bilgi sağlama seviyesi, ayırıcı özelliği ve elde edilme koşullarından bağımsız olmasıdır.

Örneğin sınıflandırma problemi nar suyu ile vişne suyunun ayırt edilmesi olsaydı renk özniteliği çok fazla bir bilgi sağlayamayacaktı çünkü iki içeceğin renklerinin dalga boyları birbirine yakın olmakla beraber yer yer üretime bağlı olarak da çakışabilmektedir. Bu nedenle probleme bağlı olarak bazı öznitelikler yüksek bilgi ve ayırma kapasitesine sahipken bazıları sahip değildir ve bu kıstaslara göre öznitelik seçimi yapmak önemlidir.

Toplanan veriden öznitelik kümesinin elde edilmesi öznitelik çıkartma (feature extraction) olarak adlandırılmaktadır ve tanımı Jain vd. (2000)'de orijinal öznitelik kümesinin birleşimlerinden veya dönüşümlerinden faydalanarak yeni özniteliklerin üretilmesi için uygulanan yöntemler olarak tanımlanmaktadır. Burada bahsedilen orijinal öznitelik kümesi, toplanan verinin kendisidir ve bu veri de aslında bir özniteliktedir. Tez çalışması kapsamında elde edilen HRRP'nin kendisi de bir özniteliktir.

Ulaşılan literatürde; HRRP'den çok sayıda öznitelik çıkartılabilmektedir ve bu konuda detaylı bilgi Guo ve Wenxian (1989), Slomka vd. (1999), Liu vd. (2011a), Chan ve Lee (2013) çalışmalarında yer almaktadır. Tez çalışması kapsamında, HRRP'den uzunluk, simetri, dağıtıcı (dispersant), radar kesit alanı (radar cross section - RCS) ve sinyal gürültü oranı (signal to noise ratio - SNR) öznitelikleri çıkartılmıştır. Bu öznitelikler, literatürde gemi sınıflandırması için en çok kullanılan öznitelikler olup, bu öznitelikler hakkında detaylı bilgi bu bölümün devamında verilmektedir.

Uzunluk Özniteliği:

Gemilerin boyları enlerine göre oldukça uzundur. Bu durum, gemilerin uzunluklarına göre kategorize edilebilmelerine olanak sağlamaktadır bu nedenle gemilerin uzunluk bilgisi sınıflandırma için önemli bir özniteliktir.

Uzunluk özniteliği, temel olarak yansıyan sinyalde menzil eksenine yayılan hücre sayısından bulunmaktadır. Ancak hedefin duruşuna bağlı olarak menzil eksenine yayılan hücreler, kerteriz ekseninde yayılım gösterebilmektedir. Bu durum deniz aracının uzunluğunun doğru tespit edilmesini zorlaştırmaktadır. Araştırma Bulguları bölümünde, uzunluk özniteliğinin daha doğru hesaplanması için gerçekleştirilmiş çalışmalar sunulmaktadır. Temel olarak uzunluk hesabı, radar yankısında menzil hücresi sayısının, radarın menzil çözünürlüğü ile çarpımına eşittir ve eşitlik (2.2)'de verilmektedir.

$$uzunluk = menzil.\text{çözünürlüğü} \times hücre.sayısı \tag{2.2}$$

Simetri Özniteliği:

Simetri bilgisi özellikle askeri gemiler ile sivil gemilerin ayırt edilmesinde kullanılan bir özniteliktir. Askeri gemiler genellikle simetrik bir biçimde inşa edilirler ve bunun sonucu olarak, ana iskeletten yansıyan radar sinyalinin enerji dağılımı incelendiğinde kesin bir simetriye sahip olduğu görülmektedir (Slomka vd. 1999).

Simetri özniteliği dolaylı olarak menzil hücrelerinin yayılımı ile ilişkilidir. Çünkü değerlerin hesaplanması, hedefin menzil hücrelerinde yayılımının tespit edilmesi ve buna göre orta noktasının bulunup bu noktaya göre olan enerji dağılımının bulunmasına dayanmaktadır. Simetri özniteliği, Liu vd. (2011b) referans alınarak eşitlik (2.3)'te verilmiştir. Eşitlikte yer alan $L_k(1)$ başlangıç, $L_k(Z(k))$ bitiş ve $n_c(k)$ merkez noktayı, $P(k,n)$ ise ilgili hücrenin genlik bilgisini temsil etmektedir.

$$S(k) = \left(\sum_{n=L_k(1)}^{n_c(k)-1} |P(k,n)|^2 \right) \Big/ \left(\sum_{n=n_c(k)+1}^{L_k(Z(k))} |P(k,n)|^2 \right) \tag{2.3}$$

Dağıtıcı Özniteliği:

Dağıtıcı özniteliği ile deniz aracının yansıma dağılımı incelenerek, geminin merkezinden uzaklaştıkça ölçülen enerjinin, toplam enerjiye oranından geminin yansıtma özelliği, Liu vd. (2011b) referans alınarak eşitlik (2.4)'te verildiği şekilde elde edilmektedir. Bu eşitlikteki parametrelerin temsil ettikleri anlamlar eşitlik (2.3) ile aynıdır, N, profildeki toplam menzil hücresi sayısını ifade etmektedir.

$$D = \left(\sum_{n=\text{hücre}_{min}}^{\text{hücre}_{max}} (n - n_c)^2 \, |P(k,n)|^2 \right) \Big/ (N^2) \sum_{n=\text{hücre}_{min}}^{\text{hücre}_{max}} |P(k,n)|^2 \qquad (2.4)$$

Radar Kesit Alanı (RCS) Özniteliği

Radar kesit alanı, nesneye iletilen ve geri yansıyan elektromanyetik dalganın bir ölçütüdür. Hedefe radar tarafından yollanan sinyal ulaştığında, sinyalin hedef üzerinde ışıma oluşturduğu alan, etkin alan olarak adlandırılmaktadır. Yansıma sinyalinin yoğunluğu RCS değeri ile tanımlanmaktadır. Çünkü yansıyan sinyalin karakteristiği; hedefin boyutlarına, geminin tasarım özelliklerine ve radar yayınına maruz kalan hedefin yüzeyinin yapısına bağlıdır (Chan ve Lee 2013).

Hedefin şekli, radar anteni tarafından yollanan sinyalin istenenden farklı yerlere saçılmasına sebebiyet verebilmektedir (Balanis 1982). Hedefin materyal özelliği ise radar anteninden gelen sinyalin yakalanmasını ve bunun bir kısmının ısı olarak dağıtılması ya da radara geri dönemeyecek yönlere saçılmasına etki etmektedir (Balanis 1982)RCS özniteliğinin Skolnik (1981)'den referans alınarak sunulan genel ifadesi eşitlik (2.5)'de verilmektedir. Eşitlikte yer alan $,R,$ hedef ile radar arasındaki mesafeyi, $P_r,$ radar yansıyan alan gücü ve P_i, gelen sinyalin alan gücünü ifade etmektedir.

$$\sigma = \frac{hedeften\ dönen\ güç/birim\ açı}{gelen\ gücün\ yoğunluğu/4\pi} = \lim_{R\to\infty} 4\pi R^2 \left|\frac{P_r}{P_i}\right|^2 \qquad (2.5)$$

Sinyal-Gürültü Oranı (SNR) Özniteliği

Sinyal-Gürültü oranı, analizi yapılmak istenen sinyalin gücünün ile gürültü sinyalinin gücünün oranlanması sonucu elde edilmektedir (Mahafza 2000) ve eşitlik (2.6-2.7)'de sunulmaktadır. Eşitlikte yer alan, P_r, istenen sinyalin gücünü ve N_0, gürültü sinyalinin gücünü temsil etmektedir. İstenen sinyalin gücü hesaplanırken, P_t iletilen gücü (Watt), G_t kazancı, A_{etk} etkin alanı ve R menzili (metre) ifade etmektedir.

$$SNR = \frac{P_r}{N_0} \tag{2.6}$$

$$P_r = \frac{P_t G_t}{4\pi R^2} \sigma \frac{A_{etk}}{4\pi R^2} \tag{2.7}$$

2.3.1 Öznitelik seçilmesi

Öznitelik seçilmesi terimi, algoritma için öznitelik giriş kümesinden en iyi alt kümenin seçilmesidir. Öznitelik seçme problemi şu şekilde tanımlanabilmektedir $m \leq d$ olmak üzere, d adet öznitelik içeren bir küme verildiğinde bu küme içerisinden en az sınıflandırma hatası verecek olan m elemanlı öznitelik alt kümesinin seçilmesidir.

Öznitelik boyutunun azaltılmasında temel motivasyon sınıflandırma başarısının arttırılmasıdır. Görece az sayıda veri ve buna oranla çıkartılan çok sayıda öznitelik varsa sınıflandırıcının zayıf genelleme yapması problemi ortaya çıkmaktadır. Bu durum boyut problemi (curse of dimesionality) olarak adlandırılmaktadır (Jain vd. 2000).

Öznitelik sayısının azaltılmasında bir diğer motivasyon ise ölçüm maliyetinin azaltılmasıdır. Burada dikkat edilmesi gereken nokta, öznitelik sayısının doğru şekilde azaltılmasıdır. Fazla azaltılması durumunda, sınıflandırıcının ayırt etme gücünün zayıflamasına yol açılabilmektedir.

Öznitelik vektörünün boyutunun azaltılmasındaki temel mesele, ölçüt fonksiyonunun seçilmesidir. Çoğunlukla kullanılan ölçüt fonksiyonu, öznitelik alt uzayının

sınıflandırma hatasıdır. Ancak özellikle örnek sayısının öznitelik sayısına oranının düşük olduğu durumlarda, sınıflandırma hatası kendi başına güvenilir bir kestirim yapmamaktadır. Ulaşılan literatürde yaygın olarak kullanılan öznitelik seçim yöntemleri aşağıda özetlenmiştir daha fazla bilgi için Jain vd. (2000) incelenebilir.

Kapsamlı Arama (Exhaustive Search)

Tüm özniteliklerin sayısı d ve seçilmek istenen öznitelik sayısı m olsun ve $m < d$ olmak üzere bu yöntemde olası tüm $\binom{d}{m}$ alt kümelerini değerlendirmektedir. Bu nedenle en uygun alt kümenin bulunmasını garanti eden bir yöntemdir.

Dal ve Sınır Arama (Branch-and-Bound Search)

Tüm olası öznitelik alt kümeleri, en iyi alt kümeyi bulabilmek için gözden geçirmektedir. En uygun alt kümenin bulunmasını, ölçüt fonksiyonunun monotonluk özelliğini sağlaması durumunda garanti etmektedir; en kötü durumda algoritmanın karmaşıklığı üsteldir.

En İyi Tekil Öznitelik (Best Individual Feature)

Tüm m özniteliği bireysel olarak değerlendirmektedir; en iyi m bireysel özniteliğini seçmektedir. İşlem yükü basittir ancak en uygun alt kümeyi bulamamaktadır.

İleriye Doğru Sıralı Seçim (Sequential Forward Selection -SFS)

İlk olarak öznitelik havuzundan en iyi özniteliği seçmektedir. Daha sonra bir başka öznitelik eklenmektedir. Arama seçilen ile en iyi olan özniteliklerin birleşimleri ölçüt fonksiyonunu maksimum değere ulaştırana kadar devam etmektedir. Bir öznitelik seçildiğinde, seçim kümesinin elemanı olarak kalmaktadır. Yöntemin işlem yükü olarak

caziptir çünkü boyutu 2 olan alt kümeleri seçmektedir, bu da yalnızca $(d-1)$ kadar olası alt kümenin incelenmesini gerektirmektedir.

Geriye Doğru Sıralı Seçim (Sequential Backward Selection- SBS)

Seçim kümesi ilk başta mevcut özniteliklerin hepsini içermektedir. Daha sonra gösterdikleri performans etkilerine göre düşük performansa sebebiyet veren öznitelikler seçim kümesinden silinmektedir. Bu süreçte bir öznitelik herhangi bir aşamada seçim kümesinden silinmiş ise bir daha seçim kümesine alınmamaktadır. Bu yöntem SFS yöntemine göre daha fazla işlem yüküne sahiptir.

"l'yi Ekle r'yi Çıkar" Seçimi ("Plus l -take away r" Selection)

İlk olarak öznitelik alt kümesi, ileriye dönük seçim uygulanarak boyutu l olacak şekilde genişletilmektedir ve daha sonra r kadar öznitelik geriye dönük seçim ile silinmektedir. SFS ve SBS yönteminde gözlenen öznitelik alt kümeleri ile çevrilme probleminden kaçınmaktadır; $l > r$ olacak şekilde l ve r değerlerinin seçilmesini gerektirmektedir

İleriye Doğru Sıralı Değişen Seçim (Sequential Forward Floating Selection- SFFS)

"l'yi Ekle r'yi Çıkar" Seçimi Yönteminin genelleştirilmiş halidir l ve r'nin değeri otomatik olarak belirlenmektedir ve dinamik olarak güncellenmektedir. Düşük maliyetli işlem yükünde en uygun sonuca yakın çözüm sağlamaktadır.

Tez çalışması kapsamında öznitelik seçimi için Genelleştirilmiş İleriye Doğru Sıralı Seçim (GSFS) yöntemi kullanılmıştır. Bu yöntem, öznitelik sayısının belirlenmesine ihtiyaç duyulmaması, karmaşık ve işlem yoğun bir algoritmaya sahip olmadığından dolayı seçilmiştir. Yöntemin kullanılması ile elde edilen sonuçlar Bölüm 4 Araştırma Bulguları başlığı altında verilmiştir.

2.4 Gemi Sınıflandırması için Sınıflandırıcı Yaklaşımları

Tez çalışması kapsamında, gemilerin sınıflandırılması probleminde üç farklı yaklaşıma ait sınıflandırıcılar kullanılmıştır ve performansları değerlendirilmiştir. Bu üç yaklaşım, benzerlik, olasılık ve geometrik tabanlıdır.

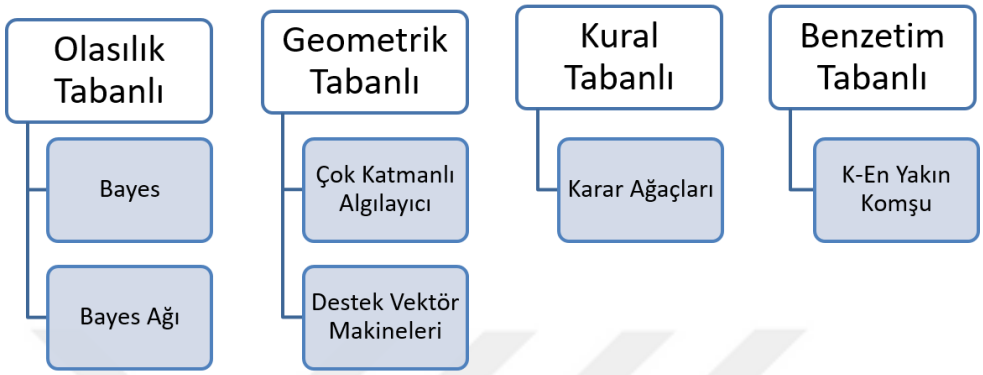
Sınıflandırıcı tasarımında, en basit ve en sezgisel yaklaşım, benzerlik tabanlı kavrama dayanmaktadır. Bu yaklaşımda örüntüler en benzer oldukları sınıflara atanmaktadır. Özellikle özniteliğin görüntü olması durumlarında şablon eşleştirme yapılarak tahminde bulunulmaktadır. Diğer tür veriler içinse, verinin belirlenmiş sınıflara olan uzaklığı ölçülmektedir. Bu ölçüm sonucuna en yakın çıktığı sınıfa atama yapılmaktadır.

Sınıflandırıcı tasarımındaki ikinci temel görüş, olasılık tabanlı yaklaşımdır. İdeal 0/1 zarar fonksiyonu ile birlikte oluşturulan Bayes karar kuralı, sonsal olasılığı maksimum olan sınıfa ilgili örüntüyü atamaktadır. Bu kural, farklı türlerdeki yanlış sınıflandırma ile ilişkili oluşan zarar göz önüne alınarak değiştirilmektedir. Sınıfların koşullu olasılık yoğunluk fonksiyonu bilindiği durumda, Bayes karar kuralı en uygun sınıflandırıcıyı oluşturmaktadır (Jain vd. 2000). Bu bakımdan önsel olasılık, zarar fonksiyonu ve sınıf koşullu olasılık yoğunluk bilgisi verildiği durum için diğer karar kurallarına göre düşük riske sahiptir.

Sınıflandırıcı tasarımında kullanılan yaklaşımlardan üçüncüsü, karar sınırlarını doğrudan hata ölçütünü en uygun hale getiren problemin çözümünden faydalanarak, karar sınırlarının kurulmasıdır. Bu yaklaşım geometrik yaklaşım olarak da ifade edilmektedir.

Bu üç yaklaşım dışında özel bir sınıflandırıcı türü olan karar ağaçları bulunmaktadır. Karar ağaçları, öncelik en belirgin öznitelik olacak şekilde yinelemeli olarak özniteliklerden bir tanesi seçilerek ağacın düğümü olarak yerleştirilmektedir. Bu şekilde her bir öznitelik eklenerek bir ağaç modeli ortaya çıkmaktadır. Yeni veri için ağaç üzerinde ilerleyerek hangi sınıfa atanması gerektiği belirlenmektedir.

Tez çalışmasında kullanılan sınıflandırıcılar sahip oldukları yaklaşımlara göre şekil 2.7'de verilen şemada sunulmaktadır. Sınıflandırıcılar hakkındaki detaylı bilgiye alt bölümlerde ulaşılabilmektedir.



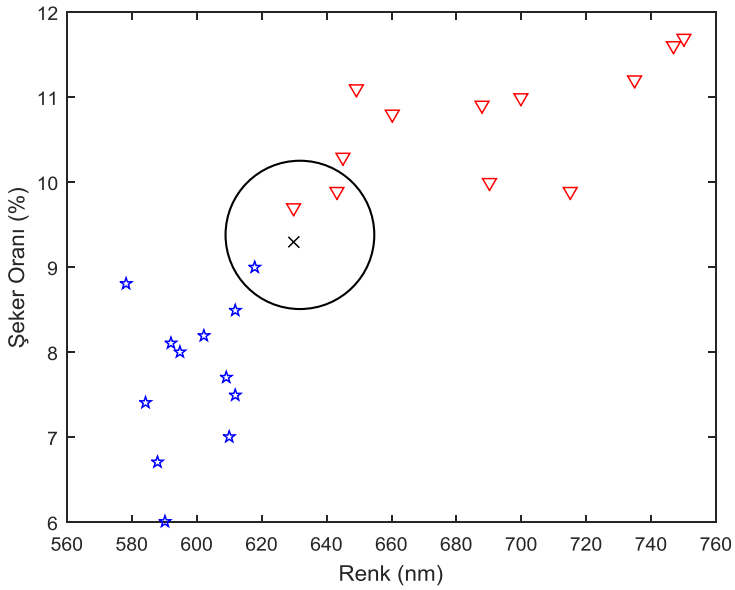
Şekil 2.7 Tez çalışmasında kullanılan sınıflandırıcılar

2.4.1 K-En yakın komşu kuralı

K-En Yakın Komşu kuralı (KNN), örüntü tanıma alanında sınıflandırma ve regresyon için kullanılan parametrik olmayan bir yöntemdir. Parametrik olmayan yöntemlerde, girdi olarak verilen uzay yerel alanlara bölünmektedir ve ilgilenilen probleme uygun bir uzaklık ölçütü kullanılarak her bir girdi için uzaklıklar hesaplanmaktadır (Alpaydın 2010).

Girdi olarak verilen eğitim örneklerinden bölünen bölgeler için yerel model hesaplanmaktadır. Test için farklı veriler kullanılarak modelin değerlendirilmesi yapılmaktadır. Bu durumu Bölüm 2.2'de verilen içeceklerin sınıflandırılması problemi üzerinden açıklamak gerekirse, toplanan içeceklerin renk bilgisine göre dağılımı şekil 2.8'de verildiği gibi olsun. Burada yıldız ile simgelenenler portakal suyu verisini, üçgen şeklinde simgelenenler de vişne suyu verisini temsil etmektedir. Model oluşturulduktan sonra, test verisi, şekilde 'x' ile temsil edilmektedir, KNN kullanılarak sınıflandırılmak istendiğinde ve en yakın komşu sayısı üç ($k = 3$) seçildiğinde, şekil 2.8'de verilen çember içerisinde test verisine en yakın olan üç komşu görülmektedir. Bu komşulardan

ikisi vişne suyu, biri portakal suyu olduğundan dolayı test verisi çoğunluk olan sınıfa yani vişne suyuna atanmaktadır.



Şekil 2.8 Test verisine k=3 için bulunan en yakın komşular

En yakın komşu sınıflandırıcısının yalancı kodu (pseudocode) aşağıda verilmektedir.

kNN (veri kümesi, test verisi)

1) veri kümesindeki her bir eleman ile test verisi arasındaki uzaklığı hesapla
2) test verisinin en yakın çıktığı k adet veri kümesindeki elemanlardan hangi sınıf daha çok çıkmışsa test verisini o sınıfa ata

En yakın komşu sınıflandırıcısında, test verisi ile komşular arasındaki uzaklığı belirlemek için birçok farklı metrik veya uzaklık ölçütü kullanılabilmektedir. Duda vd. (2001) yer alan KNN sınıflandırıcısı için en sık kullanılan uzaklık ölçütü ve metrikleri:

- Öklit uzaklığı
- Minkowski metriği
- Manhattan uzaklığı
- Tanimoto metriği

- Tanjant uzaklığı

olarak sıralanabilmektedir. Bu uzaklıklarının tanımını yapmadan önce metrik kavramını açıklamakta fayda vardır. Metrik, $D(\cdot,\cdot)$, iki değişken arasındaki mesafenin skaler değerini sunan genel bir işlevdir (Duda vd. 2001). Bir işlevin metrik sayılabilmesi için dört özelliğe sahip olması gerekmektedir. Bu özellikler, a, b, c problem örnek değişkenleri olmak üzere aşağıda sıralanmaktadır.

- Negatif olmama: $D(a,b) \geq 0$
- Yansıyabilir Olma (reflexivity): $D(a,b) = 0 \leftrightarrow a = b$
- Simetrik Olma: $D(a,b) = D(b,a)$
- Üçgen Eşitsizliğini Sağlama: $D(a,b) + D(b,c) \geq D(a,c)$

Öklit uzaklığı

Öklit uzaklığı, metrik özellikleri sağlayan ve d boyutlu durumlar için uzaklık hesaplanması için kolaylıkla kullanılabilen bir yöntemdir. Metrik özelliklerini uymasının yanı sıra değerlerin sabit bir katsayı ile çarpılması durumunda da metrik özelliklerine uymaktadır. Öklit uzaklığı eşitlik (2.8)'de verilmektedir.

$$D(a,b) = \left(\sum_{k=1}^{d} (a_k - b_k)^2 \right)^{1/2} \tag{2.8}$$

Minkowski Metriği

Minkowski metriği, d boyutlu örüntüler için kullanılan metriklerin genel bir sınıfını oluşturmaktadır ve L_k normu olarak da adlandırılmaktadır. Öklit uzaklığı L_2 normuna karşılık gelirken L_1 normu Manhattan uzaklığına karşılık gelmektedir. Minkowski metriği, eşitlik (2.9)'da verilmektedir.

$$L_k(a,b) = \left(\sum_{i=1}^{d} |a_i - b_i|^k\right)^{1/k} \tag{2.9}$$

Manhattan Uzaklığı

Manhattan uzaklığı veya şehir blok uzaklığı olarak da bilinmektedir, a, b değişkenleri arasındaki en kısa mesafenin koordinat eksenine paralel şekilde olacak yollar üzerinden bulunmasını sağlamaktadır. Adını, sokaklarının kuzey-güney ve doğu-batı yönünde inşa edilmiş olan Manhattan şehrinden almaktadır.

Tanimato Metriği

Bu metrik daha çok verilerin cinsine göre tasnif edilmesi (*taxonomy*) probleminde kullanılmaktadır (Duda vd. 2001). Tanimato metriğine göre iki sınıfın uzaklığı, S_1 ve S_2 olmak üzere iki küme ve her bir kümenin sahip olduğu eleman sayısı n_1 , n_2 ve her iki kümenin de sahip olduğu eleman sayısı n_{12} olmak üzere eşitlik (2.10) ile bulunmaktadır. Tanimato metriği yaygın olarak bir küme içerisindeki iki örüntünün ya da özniteliğin benzer veya farklı olduğunun belirlenmesi problemlerinde kullanılmaktadır (Duda vd. 2001).

$$D_{Tanimato}(S_1, S_2) = \frac{n_1 + n_2 - 2n_{12}}{n_1 + n_2 - n_{12}} \tag{2.10}$$

Sonuç olarak KNN sınıflandırıcısının kullanılmasında temel olarak seçilmesi gereken iki önemli parametre bulunmaktadır, bunlar kaç adet en yakın komşu seçileceği ve değişkenler arası uzaklığın hesaplama ölçütüdür. Belirlenen hesaplama ölçütün kullanılarak test verisine en yakın k adet komşu belirlenmektedir. Bu k adet komşudan en fazla hangi sınıfa ait eleman varsa test verisi o sınıfa atanmaktadır.

2.4.2 Bayes sınıflandırıcısı

Bayes sınıflandırıcısından bahsetmeden evvel Bayes öznel olasılık teorisinin genelleştirilmiş teorisi olan Dempster-Shafer teorisinden bahsetmekte fayda bulunmaktadır. Bu teorilerin temelinde kanaat fonksiyonları (belief functions) yer almaktadır ve ilgilenilen problemle ilişkin cevabı aranan sorunun olasılığının kanaat derecesini belirlemektedir. Kanaat dereceleri, olasılık teorisinin matematiksel kurallarına sahip olabilir veya olmayabilir, burada ayrımı oluşturan soruların ne kadar ilişkili olduğudur (Shafer 1992).

Kanaat fonksiyonları, problemle ile ilgili elde bulunan kanıtın doğruluğunun desteklenmesi için hesaplanan ve değeri (0-1) aralığında değişen çarpanları türetmektedir. Bu teori problem hakkında yeterli bilgi sağlanamadığı durumda kullanılmaktadır. Problem hakkında yeterli bilgi sağlanamadığı durumda hesaplanan olasılığın ağırlık değeri çarpılmaktadır ve bu değer kanaat fonksiyonlarının sonucundan elde edilmektedir. Bir diğer husus elde edilen ayrık olasılık bilgilerin problemin çözümü için birleştirilmesidir. Birleşim teorisi ve kanaat fonksiyonların işletilmesi ile ilgili detaylı bilgi Shafer (1976) kaynağından bulunabilir.

Bayes sınıflandırıcısı, her bir sınıfın sahip olduğu örnek veriler üzerinden dağılım fonksiyonunu çıkartarak sınıflandırma problemini ele almaktadır. Bayes sınıflandırıcısı yaklaşımında, sınıfların ve özniteliklerin ortak (joint) olasılıkları tahmin edilmektedir. Böylelikle, Bayes sınıflandırma öğrenmesinde ortak olasılık dağılımı kestirilmektedir. Bu kestirim yapıldıktan sonra, eğitim aşamasında kullanılmamış test verisinin öznitelik vektörüne bağlı olarak sınıflara ait olan koşullu olasılıkları bulunmakta ve test verisi en olası olan sınıfa atanmaktadır.

Bayes sınıflandırıcının standart gelişimi aşağıdaki eşitlikte açıklanmaktadır. Öncelikle temel tanımlamaları yapmakta fayda vardır, bu tanımlamalara göre Bayes kuralı eşitlik (2.11-2.12)'de verildiği gibidir.

- N, sınıf sayısı
- x, d boyutlu öznitelik vektörü
- w, durum (state of nature)
- w_j, ilgili j sınıfına ait olan obje
- $f(x)$, olay (event)
- $f(\bar{x}, w_j)$, pdf (likelihood)
- $P(w_j|x)$, sonsal (posterior) olasılık

$$P(w_j|x) = \frac{f(x, w_j)P(w_j)}{f(x)} \tag{2.11}$$

$$f(x) = \sum_{j=1}^{N} f(x, w_j)P(w_j) \tag{2.12}$$

Yukarıda verilen eşitlikler kullanılarak elde edilen olasılık bilgisi varlığında bir karar kuralı oluşturulup sınıflandırma işlemi gerçekleştirilebilmektedir, ikili sınıflandırma için örnek olarak oluşturulmuş kurallar, çizelge 2.3'te verilmektedir.

Çizelge 2.3 Bayes karar kuralları

Karar Kuralı	Matematiksel İfadesi
Sadece Önsel (*Prior*) Bilgi kullanılarak	eğer $P(w_1) > P(w_2)$ ise x'i w_1'e ata değilse x'i w_2'e ata
Sadece Sonsal (*Posterior*) Bilgi kullanılarak	eğer $f(x, w_1)P(w_1) > f(x, w_2)P(w_2)$ ise x'i w_1'e ata değilse x'i w_2'e ata
Bayes Risk kullanılarak	$R(\alpha_i\|x) = \sum_{j=1}^{2} \lambda(\alpha_i\|w_j)\, P(w_j\|x)$ eğer $R(\alpha_1\|x) < R(\alpha_2\|x)$ ise x'i w_1'e ata değilse x'i w_2'e ata

Çizelge 2.3'te görüleceği üzere Bayes Risk kullanılabilecek en iyi yöntemlerden biridir. Bayes Risk, temel olarak verilen kararın ne kadar riskli olduğunu ifade etmektedir. Bayes karar teorisindeki temel amaç, hatayı minimum yapmak ve en az riskli olan sınıfı seçmektir. Bu olgunun ortaya çıkmasının nedeni gerçek hayatta verilen her yanlış

kararın geri dönüşünün aynı olmamasından kaynaklanmaktadır. Örneğin, bir akciğer filmi üzerinde görüntü işleme sonucunda hastanın kanser belirtisi gösterip göstermediği bulunmak istensin. Bu akciğer filminden elde edilebilecek 4 farklı sonuç bulunmaktadır. Bunlar:

1) Yanlış + : Hasta gerçekte kanser değildir ancak kanser olarak sınıflandırılmıştır
2) Yanlış - : Hasta gerçekte kanserdir ancak kanser değildir olarak sınıflandırılmıştır.
3) Doğru + : Hasta gerçekte kanserdir ve kanser olarak sınıflandırılmıştır.
4) Doğru - : Hasta gerçekte kanser değildir ve kanser değildir olarak sınıflandırılmıştır.

Bu problemde hatalı karar verilen iki sonucun etkisi eş olarak alınamaz. Eğer alınırsa, ortaya çıkacak hatalı kararlardan hasta kanser iken, kanser değil teşhisi konulduğunda oluşacak zarar ve risk ile hasta kanser değilken, kanser olarak teşhis konmasında oluşacak zarar ve risk arasında büyük bir fark bulunmaktadır. İlk durumda sonuç hastanın hayatını kaybetmesi ile sonuçlanabilecekken, ikinci durumda ek tetkikler yapılarak kararın hatalı ortaya çıkabilmektedir. Bu sebeple çizelge 2.3'te verilen kurallardan Bayes Risk kuralının kullanılması iyi bir tercih olmaktadır.

Çizelge 2.3'te verilen Bayes Risk için oluşturulmuş karar verme modelinde $\lambda(\cdot)$ hasar (loss) fonksiyonunu, α_i ise eylem (action) temsil etmektedir. Hasar fonksiyonu verilecek hatalı kararın ağırlığını/maliyetini belirlemek için kullanılmaktadır örneğin kanser teşhisinin konulması probleminde eşitlik (2.13)'te verildiği gibi bir hasar fonksiyonu çıkabilmektedir.

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} D+ & Y+ \\ Y- & D- \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1000 & 0 \end{bmatrix} \tag{2.13}$$

Hasta gerçekte kanser ise ve kanser değil olarak sınıflandırılmışsa yani Yanlış (-) ise bu sınıflandırma hatasının riski ve vereceği hasar çok fazla olacağı çok yüksek bir katsayı ile modellenmesi gerekmektedir ve eşitlik (2.11)'de verilen modelde 1000 ağırlığı

verilmiştir. Diğer taraftan hasta sağlıklı iken kanser teşhisinin konulması yani Yanlış (+) hayati bir risk bulundurmadığı için çok daha düşük bir katsayı ile modellenmektedir.

Eylem, karar verildikten sonra ne yapılacağının belirlenmesidir örneğin kanser teşhisi probleminde hastaya kanser teşhisi yapılmışsa tedavinin gerçekleştirilmesi eylemi uygulanırken, hasta sağlıklı çıkmışsa tedaviye gerek yoktur eylemi uygulanmasıdır. Sınıflandırma problemi için eylemler sınıfların etiketleri olabilmektedir, içecek vişne suyu çıkmışsa bu içeceğin etiketini vişne suyu yap gibi bir eylem de tanımlanabilmektedir.

Bayes sınıflandırıcısında, tüm olasılık terimlerinin ve onlarla ilişkili olan olasılıkların bilindiği kabul edilmektedir. Bu nedenle problemin kolaylaştırılması açısından dağılımın Gauss olduğu kabul edilmesi avantaj sağlamaktadır. Gauss dağılımı doğada bulunan birçok olasılık dağılımını yüksek güvenirlikle temsil etmektedir bu nedenle de tercih edilmektedir (Duda vd. 2001). Bayes Risk kullanılarak sınıflandırma süreci yalancı kod (pseudocode) ile özetlenecek olursa sırası ile aşağıdaki adımlar uygulanmaktadır.

1. Her bir sınıf için eğitim verisi örneklerinden öznitelik vektörü için gerekli ölçümleri al.
2. Her bir sınıfın olabilirlik olasılık yoğunluk fonksiyonunu (likelihood pdf) ve önsel (prior) olasıklarını hesapla
3. Olabilirlik olasılıkları Gauss biçiminde olacak şekilde ifade et
4. Sonsal (posterior) olasılığı hesapla
5. Kayıp matrisini belirle ($\lambda(\cdot)$)
6. Bayes Riski hesapla ($R(\cdot)$)
7. Bayes Risk en düşük hangi sınıfın çıkıyorsa ona ata

Bu adımlar uygulanarak model geliştirilmektedir Bölüm 2.2'de verilen genel model kurulması süreçlerine göre, gerektiği takdirde Bayes modelinin geliştirilmesi için en uygun kayıp matrisinin belirlenmesi adına yinelemeli olarak model işletilebilmektedir. Bayes modeli, test için evrimleştiği zaman o test verisi için Bayes Risk hesaplanmaktadır ve hangi eylemin ($R(\alpha_i|x)$) olasılık değeri daha yüksek ise o eylem seçilmektedir.

2.4.3 Bayes ağı sınıflandırıcısı

Bayes ağı sınıflandırıcısı, grafiksel bir modele sahiptir. Grafiksel modeller, birçok avantaj sağlamaktadır. Öncelikle olasılık model yapısının, görselleştirilmesinin en basit yoludur. Koşullu bağımsızlık dahil olmak üzere, modelin özelliklerine ilişkin bilgiler grafiğin incelenmesi ile elde edilebilmektedir (Bishop 2006).

Bayes ağı sınıflandırıcısında var olan özniteliklere göre her birinin olasılık dağılımı hesaplanmaktadır. Bunu gerçekleştirmek için de eşitlik (2.14)'te verilen olasılık zincir kuralı kullanılmaktadır. Eşitlikte yer alan x_i var olan özniteliklerin her birine karşılık gelmektedir.

$$P(x_1, x_2, \cdots, x_l) = P(x_l|x_{l-1}, \cdots, x_1)P(x_{l-1}|x_{l-2}, \cdots, x_1) \cdots P(x_2|x_1)P(x_1) \tag{1.14}$$

Bu kural her zaman geçerlidir ve özniteliklerin sunulduğu sıraya bağlı değildir (Theodoridis ve Koutroumbas 2009). Bu kural, ortak olasılık yoğunluk fonksiyonunu; marjinal olasılık ile koşullu olasılıkların çarpımı olarak ifade edilmektedir. Burada her bir öznitelik için koşullu bağımlılık, çarpım içindeki her bir terimde bulunan özniteliklerin bir alt kümesiyle sınırlı olmaktadır (Theodoridis ve Koutroumbas 2009). Bu koşul altında eşitlik (2.15)'te aşağıda verildiği şekilde değişecektir.

$$P(x) = P(x_1)\prod_{i=2}^{l} P(x_i|A_i), \;\; A_i \subseteq \{x_{i-1}, x_{i-2}, \dots, x_1\} \tag{2.15}$$

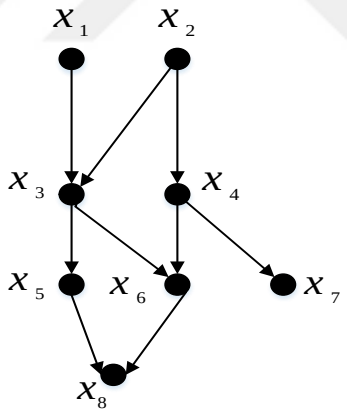
Şekil 2.9'da verilen örnek çizge için eşitlik (2.15) işletilirse aşağıda verilen denklem serisi elde edilmektedir.

$$\begin{aligned} &P(x_8|x_7, \cdots, x_1) = P(x_8|x_6,x_5) \\ &P(x_7|x_6, \cdots, x_1) = P(x_7|x_4) \\ &P(x_6|x_5, \cdots, x_1) = P(x_6|x_4,x_3) \\ &P(x_5|x_4, \cdots, x_1) = P(x_5|x_3) \\ &P(x_4|x_3, x_2, x_1) = P(x_4|x_2) \end{aligned} \tag{2.16}$$

$$P(x_3|x_2,x_1) = P(x_3|x_2,x_1)$$
$$P(x_2|x_1) = P(x_2)$$

$$A_8 = \{x_6, x_5\}, \quad A_7 = \{x_4\}, \quad A_6 = \{x_4, x_3\}, \quad A_5 = \{x_3\}$$
$$A_4 = \{x_2\}, \quad A_3 = \{x_2, x_1\}, \quad A_2 = \emptyset \tag{2.17}$$

Bu varsayım kapsamında ortak olasılığın kestirimi, daha basit ve az sayıda terim ile hesaplanabilir hale gelmiştir. Bu yaklaşımın getirisi olarak Bayes ağı düğümleri, rastgele özniteliklere karşılık gelen ve yönlendirilmiş döngü olmayan çizge (directed acyclic graph - DAG) yapısına sahiptir. Her bir düğüm, koşullu olasılık $P(x_i|A_i)$ ile ilişkilidir. Bu durum çizge üzerinde x_i özel bir düğüm ile ilişik olan bir değişken ve A_i ise çizgedeki ataların kümesini temsil etmektedir. Döngü olmayan kavramı ise çizge üzerinde herhangi bir geriye dönüş yayının bulunmadığını ifade etmektedir. Görsel olarak açıklamak gerekirse; şekil 2.9'da verilen çizgede x_8 olarak ifade edilen düğümden, x_1 olarak ifade edilen düğüme bir yay çizilmemesi durumu, çizgeyi döngü olmayan kılmaktadır.



Şekil 2.9 Bayes Ağı öznitelik örnek ilişki yapısı çizgesi

Bayes ağının bütününün tanımının yapılabilmesi için iki temel bilgi gerekmektedir. Bunlar, kök noktaların marjinal olasılıkları (ata bireyler olmadan) ve kök olamayan noktaların koşullu olasılıklarıdır. Koşullu olasılık, ilgili düğümün atalarının tüm olası birleşimleri için hesaplanmaktadır (Theodoridis ve Koutroumbas 2009). Bu durumda, değişkenlerin ortak olasılıkları, tüm koşullu olasılıkların, köklerin önsel olasılıkları ile birlikte olan çarpımından elde edilmektedir. Bunun gerçekleştirilebilmesi için gereken,

rastgele değişkenlerin topolojik olarak bölümlenmesi (sorting), diğer bir ifadeyle oluşturulan çizgede her değişkenin kendi soyundan (descendants) önce gelmesi için değişkenlerin sıralanmasıdır.

Oluşturulan bu ağdan olasılık çıkarımı (probability inference) yapılabilmesi için, bazı değişkenlerin değerlerinin bilinmesi gerekmektedir. Bu bilgi kanıt (evidence) olarak adlandırılmaktadır. Burada amaç, çizgede bulunan diğer değişkenlerin tamamının veya bir kısmının koşullu olasılıklarının verilen kanıt bilgisi ile hesaplanmasıdır. Bayes ağı, bir ağaç yapısına sahip olduğundan olasılık çıkarımı, ağaç boyunca aşağı ve yukarı yönde yayılan hesaplamaların birleşimlerinden elde edilmektedir.

Bayes ağının eğitilmesi iki bölümden oluşmaktadır. İlk bölüm ağın topolojisinin öğrenilmesidir. Topoloji, ya tüm bağımlılıkları bilen bir uzman tarafından belirlenir ya da eğitim kümesi üzerinde optimizasyon yöntemleri uygulanarak belirlenmektedir. Topoloji belirlendikten sonra, bilinmeyen parametreler (örneğin koşullu olasılıklar, marjinal olasılıklar) ulaşılan eğitim verilerinden kestirilmektedir. Özetlemek gerekirse, Bayes ağı kullanılarak sınıflandırma işlemi aşağıda verilen yalancı kod (pseudocode) adımları izlenerek uygulanmaktadır.

1. Ağın düğümleri (öznitelikleri) belirle
2. Ağın kenarları (öznitelikler arası ilişkiler/bağlar) belirle
3. Düğümlerin önsel olasılıkları hesapla
4. Düğümler arası bağlantıların koşullu olasılıkları hesapla
5. İlk dört işlemi eğitim verisi kullanılarak tekrarla ve modeli geliştir
6. Model evrimi tamamlandı ise test verisini ağa girdi olarak ver
7. Olasılık değerlerini hesapla, test verisini olasılık değeri yüksek olan sınıfa ata

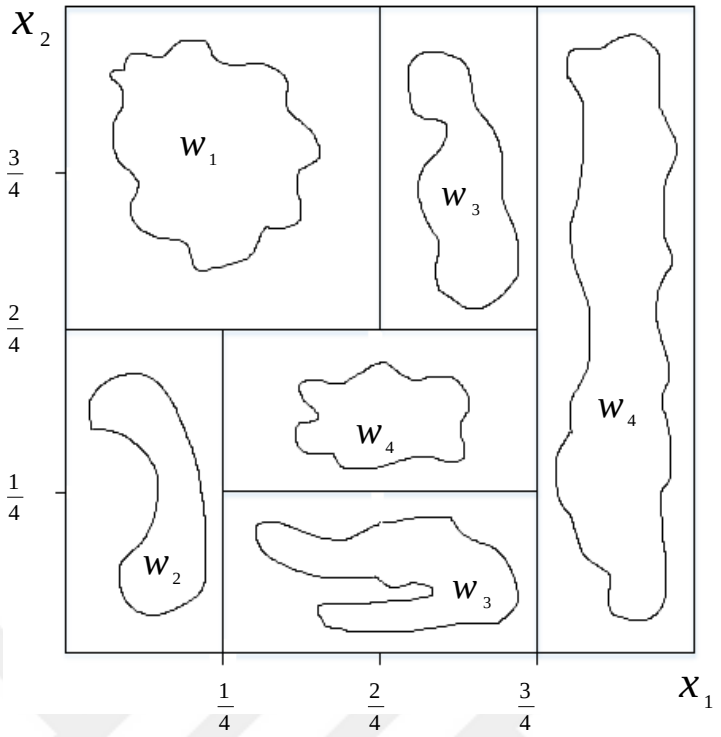
2.4.4 Karar ağaçları

Karar Ağaçları çok aşamalı (multistage) bir yöntemdir. Bu yöntemde koşulları sağlayan sınıf bulunana kadar uzaydaki tüm sınıflar sırayla değerlendirilmekte ve eşleşme

sağlanmayan sınıflar elenmektedir. Bu aşama bittikten sonra öznitelik uzayı, sınıflara karşılık gelen benzersiz bölgelere ayrılmaktadır ve bu işlem sıralı bir şekilde uygulanmaktadır (Theodoridis ve Koutroumbas 2009). Test girdisi olan öznitelik vektörünün atanacağı bölgenin aranması işlemi, uygun bir şekilde oluşturulmuş bir ağacın düğümleri boyunca, bir karar dizisi aracılığı ile uygulanmaktadır.

Bu tür ağaç yapıları, sınıf sayısı fazla olduğu zaman avantaj oluşturmaktadır (Theodoridis ve Koutroumbas 2009). En çok tercih edilen karar ağacı, uzayı eksenlere paralel olacak şekilde hiper dikdörtgenlere bölen ağaçtır. Bu tür ağaçlarda, bir öznitelik için uygun olduğu bölgenin aranması esnasında ağaç düğümlerinde ilerlenirken sorulan soru, "bu öznitelik $x_i \leq \theta$?" şeklinde olmaktadır. Bu soruda yer alan θ , eşik değerini belirtmektedir. Bu tür ağaçlar, sıradan ikili sınıflandırma ağacı (ordinary binary classification trees - OBCTs) olarak bilinmektedir, şekil 2.10'da örnek bir uzay bölünmesi verilmektedir. Şekilde yer alan w_i'ler sınıf etiketlerine karşılık gelmektedir.

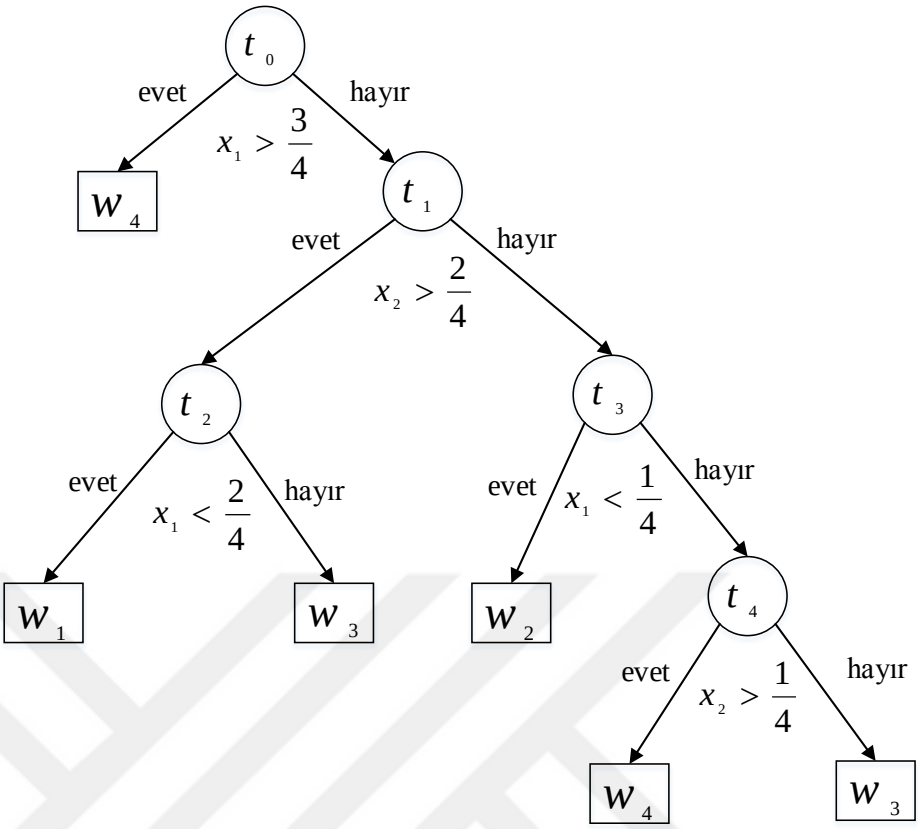
Şekil 2.10'da verilen örnek, basit iki boyutlu bir uzay için oluşturulmuş bir problemdir. Bu problem ve uzay bölünmesi için oluşturulan ağaç modeli şekil 2.11'de verilmektedir. Buna göre, ağacın her bir düğümünün eşik değerine göre verilen karar sonucunda, ikili olarak ayrıldığı görülmektedir. Bu problemde karar ağacı, temel geometri göz önüne alınarak çıkartılmıştır ve çok boyutlu olan problemlerde bunun gerçekleştirilmesi mümkün değildir. Şekil 2.11'de görüldüğü üzere ağaca x_1 özniteliğine karşılık olarak $\frac{3}{4}$ eşik değeri ile başlanılmıştır. Bu öznitelik yerine başka bir öznitelik veya başka bir eşik değeri seçimine karar verilen aşama, eğitim fazıdır. Eğitim fazında ağacın düğümlerdeki özniteliklerin seçimi ve eşik değerlerinin belirlenmesi gibi tasarım kararları verilmektedir.



Şekil 2.10 Karar Ağacı öznitelik uzayının bölünmesi örneği

Eğitim fazında ağacın tasarımı dört aşamadan oluşmaktadır, bunlar:

- soru kümesinin oluşturulması,
- bölme ölçütünün belirlenmesi,
- bölme işlemini durdurma kuralı,
- sınıfa atama kuralı, olarak sıralanabilmektedir.



Şekil 2.11 Şekil 2.10'da verilen problem için oluşturulan karar ağacı

Soru kümesinin oluşturulması aşamasında, her bir düğüm için aday sorular bir küme oluşturacak şekilde toplanmaktadır. Bu sorularda aranan temel kıstas, soruya verilen cevap sonrasında ikili bir ayrım gerçekleşmesidir. Her bir düğüm, eğitim kümesinin bir alt kümesidir. Bu sebeple düğümlerin bölünmesi eğitim kümesinin bölünmesi anlamına gelmektedir.

Bölme ölçütü ise eğitim kümesini en iyi şekilde bölen soru adayının seçilmesi aşamasıdır. Bölme işlemini durdurma kuralı, ağacın büyümesini kontrol altında tutmak için gereklidir. Bu kural, bölme işlemi sonrasında elde edilen alt kümelerin boyutuna ait bir eşik değeri konularak karar verilmesini içermektedir. Ağaçta bir düğüm, bitirme düğümü olarak nitelendirilmişse o düğüme artık yaprak denmektedir. Ağaç üzerinde sınıfa atama kuralı yapraklar üzerinde yapılmaktadır. Yaprak üzerinde sınıf etiketleri belirlenmekte ve genellikle çoğunluk kuralına göre sınıf ataması yapılmaktadır.

Şekil 2.11'de verildiği gibi bir ağaç yapısı kurgulandıktan sonra, sıra ağaç yapısının test edilmesi sürecine gelmektedir. Bunun için bir test verisi ağaca girdi olarak verilmektedir. Test verisinin öznitelik vektörünün (x_1, x_2) değerleri ağaç yapısındaki eşik değerlerine bağlı olarak yer alan sorularla karşılaştırılarak ağaç üzerinde ilerlenip yapraklara gelinmektedir. Hangi yaprağa gelindiyse o yaprağı temsil eden sınıfa test verisi atanmaktadır.

2.4.5 Çok katmanlı algılayıcı

Çok katmanlı algılayıcı (Multi Layer Perceptron - MLP); yapay sinir ağı yapısına sahiptir, sınıflandırma ve regresyon için kullanılan parametrik olmayan bir yaklaşıma sahiptir ve yapıdaki algılayıcı sistemin en basit işlemci elemanıdır. Algılayıcıya bir girdi yapılır ve bu girdi ya çevresel elemanlardan oluşarak gelen ya da başka bir algılayıcının çıkışından elde edilen bir parametredir. Algılayıcıya girdi olarak gelen parametre belirli bir ağırlık ile çarpılmaktadır ve bu bağlantı ağırlığı (connection weight) olarak adlandırılmaktadır (Alpaydın 2010). Algılayıcıya girdi olarak verilen değişkenin işlenip çıktı oluşturma sürecinin genel biçimi eşitlik (2.18)'de verilmektedir. Bu eşitlikte, girdi $x_j \in \mathbb{R}, j = 1, \cdots, d$, bağlantı ağırlığı $w_j \in \mathbb{R}$ ve çıktı y ile simgelenmiştir. Modeli daha genel hale getirmek için ek olarak sapma birimi olarak tanımlanan w_0 ile simgelenmiş parametre eklenmiştir.

$$y = \sum_{j=1}^{d} w_j x_j + w_0 \tag{2.18}$$

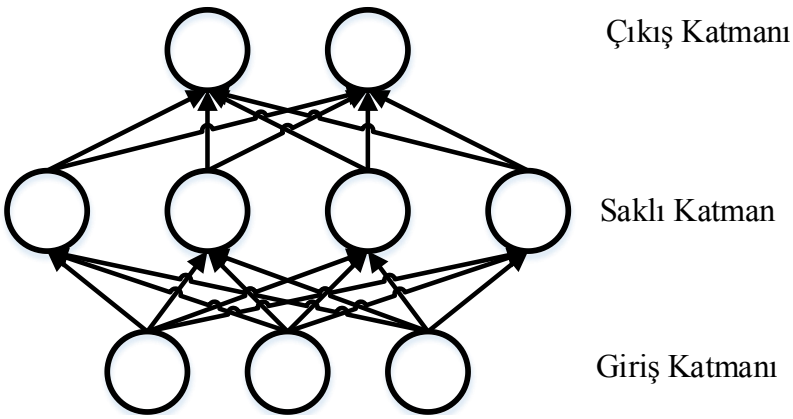
Sinir ağında, her bir algılayıcının değeri, girdi için hesaplanan yerel fonksiyon ve bağlantı ağırlığından hesaplanmaktadır. Ancak sınıflandırma işlemi için sonsal olasılık değerine ihtiyaç duyulmaktadır. Sinir ağına bunu uygulayabilmek için iki aşamalı bir süreç gereksinimi ortaya çıkmaktadır. İlk aşama, ağırlıklı toplamların hesaplanması; ikinci aşama ise değerlerin [0,1] aralığında kalmasını garanti etmek için normalleştirme işlemi (softmax scaling) işlemi uygulanmaktadır. Bu yöntem, üstel fonksiyon kullanılarak normalleştirme yapmaktadır, böylece doğrusal normalleştirme ile

değerlerin [-1,1] aralığında değişirken, bu işlem ile değerlerin [0,1] arasında değişmesi sağlanmıştır.

Sinir ağlarının eğitilmesi işlemi genellikle çevirim içi öğrenme şeklinde olmakta ve eğitim örneklerinin tamamı ilk etapta girdi olarak verilmemektedir. Örnekler birer birer ağa girdi olarak verilerek, ağın her girdi için parametrelerini güncellemesi ve kendisini zaman içinde yavaşça uygun hale getirmesi sağlanmaktadır. Bu yaklaşımının uygulanmasının üç ana motivasyonu vardır.

1) Eğitim örneklerini bellek etkin bir biçimde saklanmaktadır.
2) Problemin veya eğitim kümesinin zaman içindeki olabilecek değişimlerine karşı uyum sağlayabilmektedir.
3) Donanımsal problemler için fiziksel ortamın değişimine uyum sağlayabilmektir.

Tek katmanlı algılayıcı sadece girdinin doğrusal fonksiyonuna yaklaşabilmektedir. Ancak ayrımın doğrusal olmadığı problemlerde uygulanamamaktadır, problemi çözmemektedir. Bu nedenle, ileri doğru beslemeli ağlar için giriş ile çıkış arasında gizli katman olması gerekmektedir. Bu gereklilik çok katmanlı algılayıcıları oluşturmuştur, bu algılayıcı doğrusal olmayan ayırıcı fonksiyon (discriminant) uygulamaktadır. MLP, üç veya daha çok katmandan (bir adet giriş, bir adet çıkış, bir veya daha fazla saklı katman) oluşmaktadır. Her bir katman bir diğerine tamamen bağlantılıdır, diğer bir ifadeyle bir katmandaki her bir düğüm belirli bir ağırlık değeri ile bir sonraki katmanın her bir düğümü ile bağlantılıdır. Bu durum, şekil 2.12'de gösterilmektedir.



Şekil 2.12 Üç katmanlı örnek MLP gösterimi

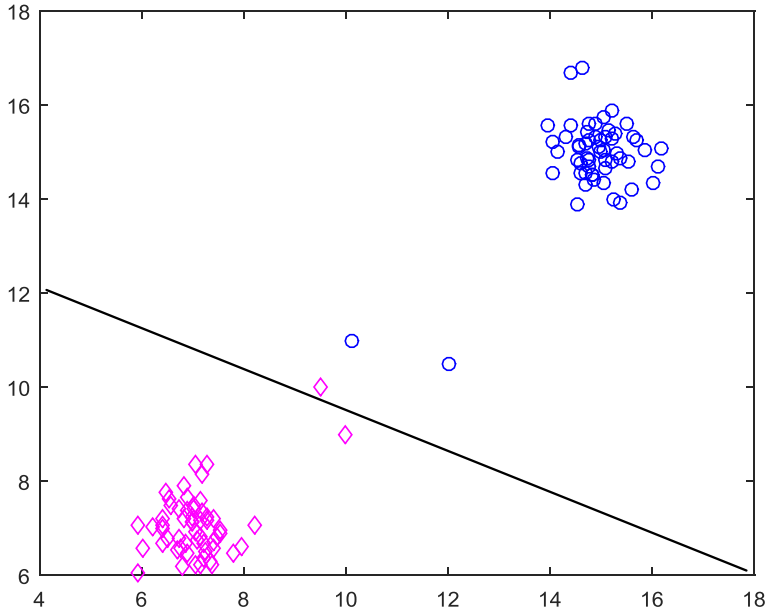
Ağın giriş katmanına uygulanan bir x girdisinin oluşturduğu aktivasyon ileri yönde yayılmaktadır ve saklı katmanın değerleri de eşitlik (2.19)'da verildiği üzere, doğrusal olmayan sigmoid fonksiyonu ile ağırlıklı olarak toplanmaktadır. Eşitlikte yer alan z_h ilgili saklı katman için hesaplanan çıktıyı, H toplam saklı katman sayısını ve alt indis h saklı katmana ait olan parametreleri ifade etmektedir.

$$z_h = sigmoid(w_h^T x) = \frac{1}{1 + \exp[-(\sum_{j=1}^{d} w_{hj} x_j + w_{h0})]}, h = 1, \cdots, H \tag{2.19}$$

Eğitim verileri ile ağ güncellenerek en iyi sonucu verecek bağlantı ağırlıkları belirlenmektedir. Bu değerler belirlendikten sonra test aşamasına geçilmektedir ve verinin öznitelik vektörü ağa girdi olarak verilmektedir. Ağ üzerinde yer alan bağlantı katsayıları, saklı katmanda yer alan fonksiyon çıkışları her biri tek tek hesaplanarak ağın sonuna varılmaktadır. En sona gelindiğinde hesaplanmış değer hangi sınıfı temsil ediyorsa test verisi o sınıfa atanmaktadır.

2.4.6 Destek vektör makineleri

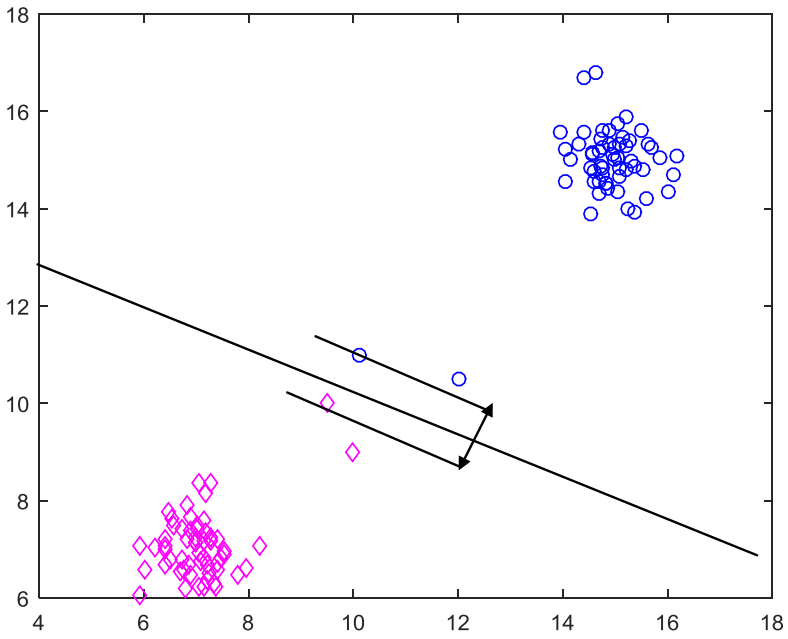
Destek Vektör Makineleri (Support Vector Machines-SVM) yöntemi sınıfların ortalama değerlerine göre doğrusal olarak sınırların belirlendiği durumda bazı noktaların yanlış sınıflandırılması problemini çözmek amacıyla geliştirilmiştir. Şekil 2.13'te doğrusal ortalama bir sınıflandırıcının çizdiği sınır sonucunda yanlış sınıflandırılan veriler görülmektedir.



Şekil 2.13 Doğrusal ortalama sınıflandırıcısının yanlış sınıflandırdığı veriler

SVM, her bir sınıftan olabildiğince uzak olacak şekilde sınırı belirlemeye çalışmaktadır. Bu sınır, maksimum aralıklı hiper düzlem (maximum-margin hyperplane) olarak adlandırılmakta (Segaran 2007) ve şekil 2.14'te verilmektedir. Sınırın yanında bulunan noktalar destek vektörleri (support vectors) olarak adlandırılmaktadır. Algoritma, ilk olarak destek vektörlerini bulmakta ve daha sonra bu vektörleri kullanarak sınıfları bölen sınırı bulmaktadır.

SVM kullanılmasının tercih edilmesinin bir diğer nedeni de sınır hattı üzerinde belirlenen kenar aralığının (*margin*) ayarlanmasıdır. Karar sınırın yakınlarında bulunan noktalar, sınıflandırma kararı verilmesinde oluşan kesinliği düşürmekte ve sınırdan ötede belirlenen kenar boşlukları geniş seçilirse kesinlik artmaktadır. Çünkü geniş kenar boşluğu olduğu durumda, ölçümde meydana gelen ufak hatalardan dolayı oluşan yanlış sınıflandırma engellenmektedir. Geniş ayırıcı aynı zamanda daha az karar verme demektir. Bu da modelin hafıza kapasitesinin düşmesine karşılık test aşamasında doğru sonucun elde edilmesini arttırmaktadır (Manning vd. 2009).



Şekil 2.14 Destek vektörleri kullanılarak en iyi sınırın bulunması örneği

SVM, yüksek boyutlu veri kümeleriyle birlikte iyi çalıştığı için bu yöntem veri-yoğun (data-intensive) bilimsel problemlerde ve oldukça karmaşık veri kümelerine sahip olan diğer problemlerle ilgilenmek amacıyla kullanılmaktadır. Yüz ifadelerinin sınıflandırılması, askeri veri kümesinden düşman/yabancı tespit edilmesi, el yazısı tanımlama vb. (Segaran 2007) bunlara örnek olarak verilebilmektedir.

SVM, ikili sınıflandırıcı (binary classifier) olarak tasarlanmıştır. Bununla birlikte, SVM sınıflandırıcısının çok sınıf için kullanılması üzerinde çok fazla çalışma bulunmaktadır. Genel olarak, K adet SVM kullanılarak K adet sınıfa sahip olan sınıflandırma probleminin çözülebileceğinden bahsedilmektedir (Mayoraz ve Alpaydın 1999, Liu vd. 2011a).

SVM sınıflandırıcısının çok sınıf sınıflandırmaya evirilmesi için kullanılan yöntemler; bir hepsine karşı (one-against-all), bir birine karşı (one-against-one) ve ikili hiyerarşidir (binary hierarchical).

Hepsine Karşı (one-against-all) Yöntemi

Hsu ve Lin (2002) çalışmasında yer alan bir hepsine karşı yaklaşımında; m'nci SVM için m'nci sınıfa ait tüm elemanlar pozitif olarak etiketlenmekte, geriye kalanlar ise negatif etiketlenmektedir. Buna göre çekirdek fonksiyonu kullanılarak m'nci sınıf ile diğerleri arasında sınır belirlenmektedir. Bu yaklaşım için kullanılan çekirdek fonksiyonu eşitlik (2.20)'de ve sınıf atama kuralı eşitlik (2.21)'de verilmiştir.

$$f(x_i) = (w^m)^T \phi(x_i) + b^m \tag{2.20}$$

$$x = \underset{m=1,\cdots,k}{\operatorname{argmax}}((w^m)^T \phi(x) + b^m) \tag{2.21}$$

Bir birine karşı (one-against-one) Yöntemi

Hsu ve Lin (2002) ve Milgram vd. (2006) çalışmalarında yer alan biri birine karşı yaklaşımında; k adet sınıftan her bir sınıf çifti için SVM kullanılmaktadır. Burada SVM bir sınıfı diğerinden ayırmak için eğitilmektedir. Test aşamasında, x test verisi olmak üzere, x'in hangi sınıfa ait olduğunun bulunmasında oylama stratejisi uygulanmaktadır. Buna göre her bir SVM bir sınıfa oy vermekte ve (sign($f^{ij}(x)$) işleminin sonucuna göre) en fazla oyu alan sınıfa x verisi atanmaktadır. Bu yaklaşımdaki problem her iki sınıfın da eşit sayıda oy alma ihtimalinin bulunmasıdır.

İkili hiyerarşi (binary hierarchical) Yöntemi

Wang ve Casasent (2009) çalışmasında yer alan ikili hiyerarşi yaklaşımında, hiyerarşinin en üst seviyesinde kaba sınıflandırma yapılmaktadır. Daha alt seviyelere gidildikçe daha iyi kararlarla sınıflar ayrıştırılmaktadır. Her bir seviyede SVM kullanılmaktadır. Hiyerarşinin en üst düğümünde k adet sınıf, iki adet makro sınıfa ayrılmaktadır. Bu makro sınıflar küçük sınıfların oluşturduğu gruptur. Test aşamasında sınıf ataması için 0-1 kararı (Hard-decision (binary decision)) veya yumuşak karar (soft-decision) ile yapılmaktadır.

Sınıf ataması için kullanılan 0-1 kararı yönteminde test girişi, her bir düğümde her zaman makro-sınıflardan birine atanmaktadır. Bu karar verme yönteminin dezavantajı, iç düğümlerin birinde yanlış sınıflandırma hatasının meydana gelmesi durumunda bu hatanın daha alt düğümlerde düzeltilememesidir. Çünkü hatanın yapıldığı düğüme, doğru olan ancak seçilmeyen yol göz ardı edilerek bakılmamaktadır (Wang ve Casasent 2009).

Yumuşak karar yönteminde ise hiyerarşideki her bir seviyede her bir düğüme güven skoru atanmaktadır. Bu karar verme yönteminin avantajı, tüm düğümler ve tüm yollar aranmakta ve en alt seviyede test girişi, en yüksek olasılığa sahip olan sınıfa atanmaktadır. Sınıflara ait sonsal olasılık değerleri (posterior probability) SVM çıkışının dönüştürülmesiyle elde edilmektedir (Wang ve Casasent 2009).Yumuşak karar verme, ikili karar verme yapısına sahip olan tüm sınıflandırıcılar için uygulanabilir bir yöntemdir.

İkili hiyerarşik SVM yapısında eldeki sınıflardan hangilerinin ilk iki makro sınıfa ayrılacağı ve en son seviyeye gelene kadar bu sınıf gruplarının oluşturulması ile ilgili temel kümeleme (clustering) yöntemleri uygulanmaktadır (Wang ve Casasent 2009).

Sonuç olarak, SVM sınıflandırıcısı sınıfları birbirinden ayırmak için iki sınıfın birbirine en yakın olan elemanlarından faydalanarak sınırlar oluşturmaktadır. Bu sınırlar özniteliklerin sayısına göre doğrusal çizgiler (tek öznitelik) ya da hiper düzlemler olabilmektedir. Sınırlar belirlendikten sonra test verisinin bu sınırlardan hangisinin içinde kaldığı veya hangisine en olduğu hesabına göre atama yapılmaktadır.

3. MATERYAL VE YÖNTEM

Bu bölümde, tez çalışmasında kullanılan materyalden, materyalin özelliklerinden ve kullanımına ait bilgiler verilmektedir. Araştırma bulgularının elde edilmesinde kullanılan teknikler açıklanmaktadır.

Çalışma kapsamında, materyal olarak radar simülatöründen elde edilen benzetim verileri kullanılmıştır. Bu benzetim verilerinin üretilmesinde kullanılan radar simülatörü özellikleri ve benzetim verisi üretim kısıtları bu bölüm altında sunulmuştur. Üretilen benzetim verileri vasıtasıyla sınıflandırma işleminin gerçekleştirilebilmesi için WEKA adı verilen makine öğrenmesi aracı kullanılmıştır. Bu aracın özellikleri de bu bölüm altında verilmektedir.

3.1 Radar Simülatör Özellikleri

Tez çalışmasında, sahil şeridine yakın bölgelerdeki gemilerin sınıflandırılması problemi incelenmiştir. Gemilerin kıyıya yakın olmalarından dolayı uzakta bulunan gemilerin aksine radarın huzme genişliği geminin tamamını tarayamamaktadır, bu durum radar yankısında çok sayıda menzil ve kerteriz hücresinde veri edinimine neden olmaktadır. Radarın tam bir tur tarama yapması sonucunda hedefin tamamı taranabilmektedir, bu nedenle klasik menzil profillerinin aksine menzil ve kerterizde yayılan iki boyutlu veri elde edilmiştir. SAR görüntüleri geminin tepeden elde edilmiş, gelen enerji seviyelerine göre kodlama yapılarak geminin yüksek çözünürlüklü görüntüsünün elde edilmesine dayandığı için burada elde edilen görüntüler SAR görüntülerinden ayrılmaktadır.

Bu kapsamda, gerçek veri edinimi maliyetli ve zor olduğundan dolayı bir radar simülatörünün çıktılarından elde edilen benzetim verileri kullanılmaktadır. Bu benzetim verileri Elektrik-Elektronik Mühendisliği Anabilim Dalı'nda doktora öğrenimini yapan Bahadır Tevfik Sarıkaya tarafından hazırlanmıştır. Kullanılan simülatörün temel özellikleri çizelge 3.1'de verilmektedir.

Çizelge 3.1 Radar simülatörü ve veri özellikleri

Parametreler	Değerler
Radar Türü	X-bant radar
Menzil Çözünürlüğü	2.5 metre
Açı Çözünürlüğü	0.5 derece
Hedef Modeli	Swerling-1
Tespit Olasılıkları	% 75, % 85, % 95
Hedef Geometrik Modeli	Dikdörtgen
(0,0) metre konumunda bulunan radara göre hedeflerin merkezlerinin konumları	(100, 100), (1000,1000), (5000,5000) metre
Hedeflerin radara göre kerteriz açıları	45˚, 75˚, 105˚, 135˚, 165˚, 195˚, 225˚

Simülatörde modellenen radarın türü, X-bant radar olarak seçilmiştir. X-bant radarlar 8-12 GHz bant genişliğine sahiptir ve kısa mesafe takibi, füze yönlendirme, deniz radarı, hava önleyici (airborne intercept) gibi alanlarda kullanılmaktadır (Skolnik 1990).

Simülatörde hedef modeli Swerling-1 olarak seçilmiştir. Bu model Peter Swerling tarafından, dalgalanma kayıplarının (fluctuation loss) radar menzili üzerindeki etkilerinin belirlenmesi amacıyla oluşturulmuştur. Bu kapsamda dört adet durum modellemesi tanımlanmıştır. Swerling modelini anlatmadan önce bu modeli ortaya koymanın kaynağı olan dalgalanma kayıplarından bahsetmek gerekmektedir. Alınan sinyal üzerinde oluşan dalgalanma kaybı, hedefin yansıtıcı yüzeyinin yapısına yani RCS'ye dayanmaktadır. Bunun nedeni RCS'nin geri yansıma çizgesinin karmaşık yapılı olmasından kaynaklanmaktadır. Radara göre ileri yönde hareket eden bir hedefin, RCS grafiği referans radara göre dönmektedir. Bu durum, hedefin rotasının zamansal değişimlerinin neden olduğu genlikler ve faz değişiklikleri aracılığı ile radar antenindeki alım alanı kuvvetinde güçlü bir dalgalanma yaratmaktadır (Swerling 1954).

Swerling-1 modeline göre, radar art arda taramalar yaparken, bir tarama anında iletilen sinyalin hedefin üzerinde kaldığı anda (dwell time) yansıma sinyallerinin genlikleri sabit kalmaktadır ancak farklı tarama anları için sinyallerin genlikleri değişebilmektedir. Bu nedenle taramalar birbirinden bağımsızdır ve aralarında bir ilinti (correlation) kurulamamaktadır. Bunun bir sonucu olarak bu model taramadan-taramaya dalgalanma

(scan-to-scan fluctuation) modeli olarak adlandırılmaktadır. Daha detaylı bilgi ve diğer durum modelleri için (Swerling 1954) incelenebilir.

Swerling-1 ile modellenmesinin yanında farklı tespit olasılıklarına göre simülatör işletilerek veri edinimi gerçekleştirilmiştir. Tespit olasılığı (probability of detection), radarın, taradığı alandaki tüm hedeflere oranla, yankı sinyalinden tespit ettiği hedef sayısının yüzde cinsinden ifade edilmesidir. Radar gönderdiği sinyalden topladığı yankı sinyalini demodüle ettikten sonra bir eşikleme mantığı kullanarak veriyi işlemektedir. Bu eşikleme mantığı, hedeften yansıyan sinyal ile ortamda bulunan gürültünün birbirinden ayırt edilmesi için kullanılmaktadır. Ancak bu işlem 0/1 kadar basit olmamaktadır çünkü gürültü içinde bastırılması gereken gürültü tepeleri bulunurken, daha düşük genlikli hedefe ait sinyaller de kaybedilmek istenmemektedir. Sinyalleri kaybetmemeye çalışırken de hatalı alarm oramı (false alarm rate) yüzdesinin aşılmaması gerekmektedir. Tespit olasılığının değeri de direkt olarak SNR, eşik değeri gibi parametrelere bağlı olarak değişmektedir. Çalışma kapsamında farklı tespit olasılıkları kullanılarak sınıflandırma başarısına olan etkisi incelenmiştir.

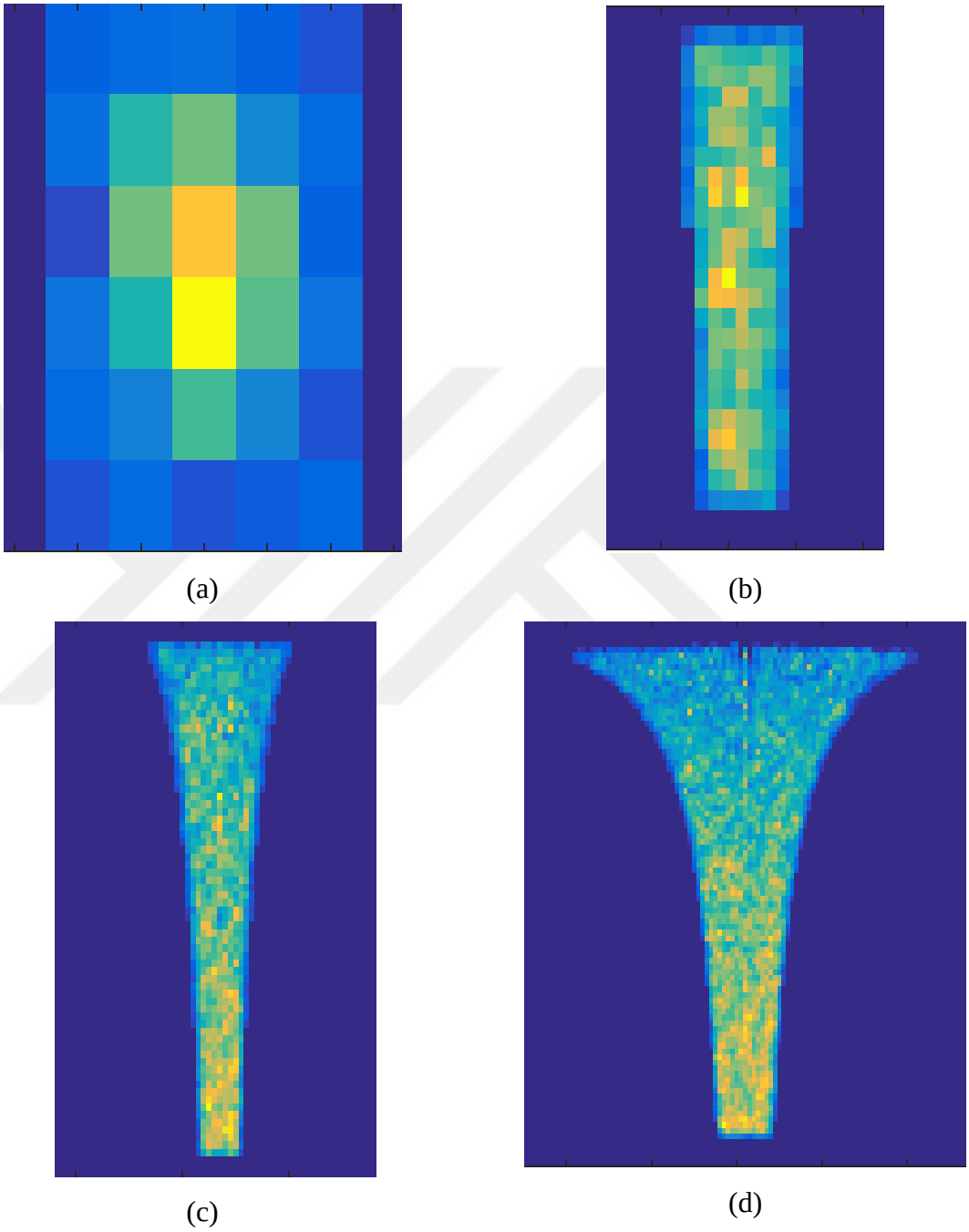
Radar simülatörü kullanılarak dört farklı gemi için benzetim verisi üretilmiştir, bu gemiler, balıkçı teknesi, sahil güvenlik gemisi, kargo gemisi ve orta boy tankerdir. Bu gemiler, küçük, orta, büyük ve çok büyük boy gemi olmak üzere sınıflandırılmaktadır. Seçilen gemilerin boy ve en bilgiler çizelge 3.2'de verilmektedir.

Çizelge 3.2 Sınıflandırma için kullanılan hedef özellikleri

Gemi Türü	Boy	En
Balıkçı Teknesi	9 metre	2 metre
Sahil Güvenlik Gemisi	55 metre	7 metre
Kargo Gemisi	167 metre	12 metre
Orta Boy Tanker	213 metre	25 metre

Çizelge 3.2'de yer alan en-boy bilgisi kullanılarak bir dikdörtgen modellenmiştir. Bu gemi modeli ve çizelge 3.1'de yer alan parametreler radar denklemlerine girdi olarak verilerek, denklemler çözülmüş ve yansıma sinyalleri elde edilmiştir. Şekil 3.1'de

hedeflerin konumlarının (100,100) metre ve baş açılarının 45˚ yani radar görüş hattı üzerinde olduğu durumda elde edilen benzetim verileri sunulmaktadır.



(a) (b) (c) (d)

Şekil 3.1 Radar simülatöründen elde edilen örnek benzetim verileri

a.Balıkçı Teknesi, b. Sahil Güvenlik Gemisi, c. Kargo Gemisi, d. Orta Boy Tanker

Şekil 3.1'de verilen simülatör çıktılarında –x ekseni kerteriz eksenine, -y ekseni ise menzil eksenine denk gelmektedir. Hedefler dikdörtgen şeklinde modellendiği için elde edilen görüntülerde dikdörtgene benzer şekilde elde edilmiştir.

3.2 WEKA Temel Özellikler

Tez çalışmasında sınıflandırıcıların sonuçları, makine öğrenmesi için çalışma ortamı oluşturan WEKA (Waikato Environment for Knowledge Analysis) uygulamasından faydalanılarak elde edilmiştir. WEKA, makine öğrenmesi amacıyla Waikato Üniversitesinde geliştirilmiş, makine öğrenimi algoritmalarını ve metotlarını içeren bir uygulamadır. WEKA yazılımı kullanılarak, veri kümeleri üzerinde görselleştirme, veri analizi, veri madenciliği gibi işlemler yapılabilmektedir. Veri madenciliği alanında temelde üç işlem yapılabilmektedir bunlar sınıflandırma, kümeleme (clustering) ve ilişkilendirmedir.

WEKA kullanılarak analiz yapılması için ilk olarak elde edilen veriler WEKA'nın kendine özgü geliştirilmiş dosya yapısında biçimlendirilmelidir. WEKA'da kullanılan dosya yapısı .arff (attribute relationship file format) olarak adlandırılmıştır, bu biçimde hazırlanan veriler WEKA'ya yüklenmektedir. Veriler yüklendikten sonra, sınıflandırma adımında önce verinin eğitim ve test kümlerine bölünmeleri gerçekleştirilmektedir. WEKA'da eğitim ve test kümelerine veri bölüştürülmesi üç farklı yöntem kullanılarak yapılmaktadır. Bu yöntemlerden ilki, yüklenen verinin sadece eğitim için kullanılması ve test için ayrıca veri kümesi yüklenerek sınıflandırıcının test edilmesidir.

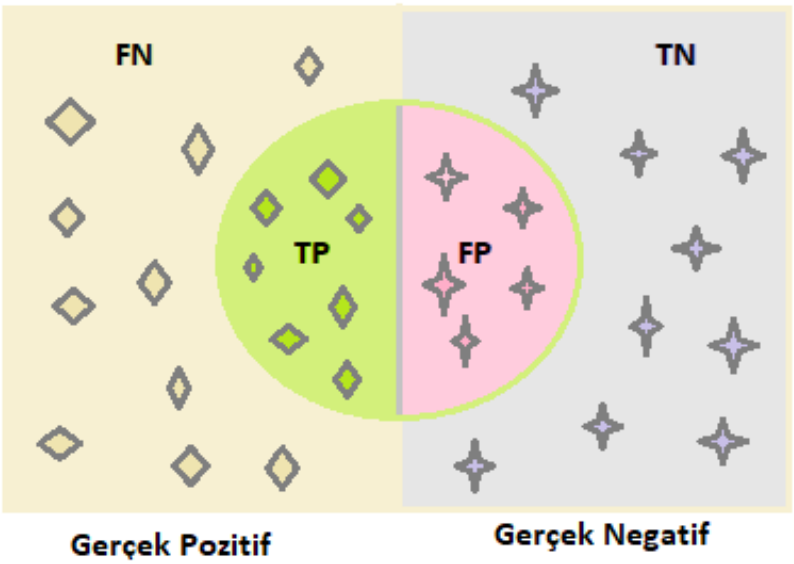
İkinci yöntem çapraz geçerleme (cross-validation) olarak adlandırılmaktadır, bu yöntemde eğitim-test kümesi rasgele seçilmektedir. Bunu gerçekleştirebilmek için veriler, belirlenen k kadar kümeye eşit sayıda ayrılmaktadır. Daha sonra bu kümelerden biri test geriye kalan $k-1$ küme eğitim için seçilmektedir. Bu kümelerin hangi sıra ile seçildiğinin bir önemi yoktur çünkü bu işlem k defa tekrarlanmaktadır ve her bir küme bir kez test kümesi olarak seçilmektedir. Her bir test kümesi için elde edilen sonuçların ortalaması alınarak sınıflandırıcının genel başarısı elde edilmektedir. Bu yöntemdeki amaç kümenin nihai olarak tamamının test olarak kullanılarak rasgele elde edilen

başarının veya başarısızlığın önüne geçilmesini sağlamaktadır. Araştırmacılar genel olarak $k = 10$ olarak seçmektedir bu çalışmada da aynı değer seçilmiştir.

Son yöntem belirlenen yüzde oranı (percentage split) olarak adlandırılmaktadır, bu yöntemde yüzde oranına göre eğitim-test kümesi seçilmesinde, yüklenen verinin yüzde kaçının eğitim kümesi için seçileceği belirlenmektedir. Örneğin %70 olarak belirlenmişse ve veri sayısı 1000 ise 700 veri eğitim için rasgele olarak seçilmektedir geriye kalan 300 veri ise eğitilen sınıflandırıcının testi için kullanılmaktadır. Genellikle verinin %66'sı eğitim için kullanılmaktadır.

3.2.1 WEKA sınıflandırıcı başarı parametreleri

WEKA'da sınıflandırıcıların başarılarının analizi için birçok parametre üretilmektedir. WEKA'nın sınıflandırıcıların başarıları tanımlayan temel parametreler aşağıda verilmiştir. Bu parametrelerin iki sınıf için gösterimi şekil 3.2'de verilmektedir. Şekil 3.2'de iki sınıfa ait veri bulunmaktadır bunlar şekildeki dikdörtgenin sol tarafında bulunan pozitif ve sağ tarafında bulunan negatif sınıflardır. Şekilde yer alan daire pozitif olarak sınıflandırılmış olan verileri gösterirken daire dışında kalan negatif olarak sınıflandırılmış verileri temsil etmektedir. Bu şekle göre oranlar aşağıda verildiği gibi tanımlanmaktadır:



Şekil 3.2 Başarı oranlarının temsili gösterimi

1) TP = Doğru pozitif: Gerçekte pozitif olan ve pozitif olarak tahmin edilmiş örnek sayısı
2) TN = Doğru negatif: Gerçekte negatif olan ve negatif olarak tahmin edilmiş örnek sayısı
3) FP = Yanlış pozitif: Gerçekte negatif olan ancak pozitif olarak tahmin edilmiş örnek sayısı
4) FN = Yanlış negatif: Gerçekte pozitif olan ancak negatif olarak tahmin edilmiş örnek sayısı

Sınıflandırma işlemi sonucunda, yukarıda verilen parametrelere göre bir hata matrisi elde edilmektedir. Örneğin, iki sınıftan oluşan bir problem için sınıflar a ve b olarak adlandırılsın. Problem ise tümörün kötü huylu veya iyi huylu olarak sınıflandırılması olsun, a sınıfı iyi huylu tümörü (pozitif sınıf), b sınıfı da kötü huylu tümörü (negatif sınıf) temsil etsin bu durumda WEKA'nın hata matrisi aşağıda verildiği gibi olmaktadır.

$$\begin{array}{lcl} & a & b \leftarrow sınıfına\ atanmaktadır \\ gerçek\ a = iyihuylu & TP & FN \\ gerçek\ b = kötühuylu & FP & TN \end{array} \tag{3.1}$$

Temel parametrelere bağlı olarak WEKA'da hesaplanan ve ayrıntılı kesinlik (Detailed Accuracy) başlığı altında verilen parametrelerin matematiksel anlamları çizelge 3.3'te verilmektedir.

Çizelge 3.3 WEKA'da tanımlı parametrelerin açıklaması

Değişken Adı	Matematiksel İfade
Kesinlik *ACC (*Accuracy*)*	$ACC = \frac{TP + TN}{TP + TN + FP + FN}$
Pozitif tahmin edilen değer (Precision)	$TP/(TP + FP)$
Hassasiyet (Recall)	$TP/(TP + FN)$
F-Ölçümü	$F = 2 \times \frac{precision \times recall}{precision + recall}$
Matthews korelasyon katsayısı (MCC)	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$
ROC (Receiver Operating Analysis)	TP oranına karşılık FP oranının çizdirilmesidir. ROC sınıflandırıcıların karşılaştırılmasına ve (0,1) en yakın ve $TP_{rate} = FP_{rate}$ en uzak olan modelin seçilmesine olanak sağlamaktadır.
PRC Alanı (Precision Recall Curve Area)	Sınıflandırıcının performansının dengeli ve dengesiz veri kümelerinde nasıl davrandığını ölçmektedir. (Dengeli veri kümesi TP ile TN ve FP ile FN eşit oranda dağıldığı veri kümesi)

Özet bilgi kapsamında verilen değişkenlerin matematiksel ifadeler çizelge 3.4'te verilmektedir. Burada θ gerçek değeri, $\hat{\theta}$ algoritma sonucunda tahmin edilen değeri ve $\acute{\theta}$ gerçek değerin ortalama değerini temsil etmektedir.

Çizelge 3.4 WEKA kullanılan değişkenlerin matematiksel ifadesi

Değişken Adı	Matematiksel İfade
Kappa İstatistiği	Gözlemlenen kesinlik (*observed accuracy*) ile beklenen kesinlik (*expected accuracy (random chance)*) karşılaştırılmasından elde edilen bir değerdir. $kappa = \frac{(observed_{ACC} - expected_{ACC})}{1 - expected_{ACC}}$
Ortalama Mutlak Hata (Mean Absolute Error)	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert\hat{\theta}_i - \theta_i\rvert$
Ortalama Kare Hatanın Karekökü (Root Mean Square)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta_i)^2}$
Görece Mutlak Hata (Relative Absolute Error)	$RAE = \frac{\sum_{i=1}^{N}\lvert\hat{\theta}_i - \theta_i\rvert}{\sum_{i=1}^{N}\lvert\grave{\theta} - \hat{\theta}_i\rvert}$
Görece Hatanın Karekökü (Root Relative Error)	$RRSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{\theta}_i - \theta_i)^2}{\sum_{i=1}^{N}(\grave{\theta} - \hat{\theta}_i)^2}}$

4. ARAŞTIRMA BULGULARI

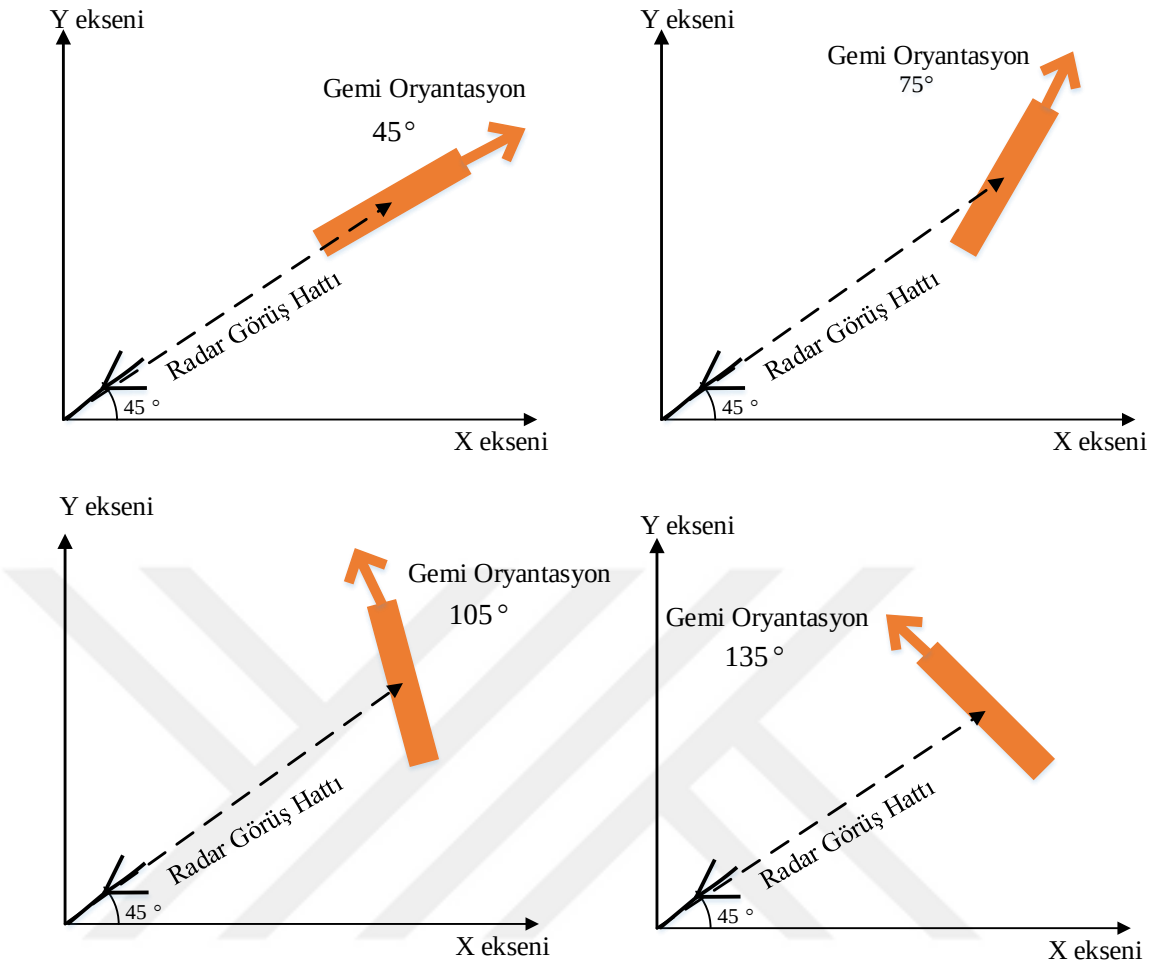
Bu bölümde tez kapsamında yapılan çalışmalar ve elde edilen sonuçlar alt bölümlerde paylaşılmaktadır.

4.1 Özniteliklerin Çıkartılması ve Sınıflandırma Sonuçları

Bu bölümde verilen çalışma kapsamında, Bölüm 2 Kuramsal Temellerde anlatılan RCS, SNR, uzunluk, dağıtıcı ve simetri özniteliklerinin radar simülatöründen elde edilen benzetim verisinden çıkartılması işlemi uygulanmıştır. Öznitelikler çıkartıldıktan sonra özniteliklerin sınıflara göre değerlerinin nasıl değiştiği gözlenmiştir ve son olarak WEKA aracı kullanılarak yine Bölüm 2'de verilen sınıflandırıcılar koşturularak başarımlar incelenmiştir.

Şekil 4.1'de radar hedef geometrisi verilmektedir. Şekil incelendiğinde, geminin farklı oryantasyonlarına (kerteriz açıları) göre radara ve radar görüş hattına göre nasıl durduğu görülmektedir. Gemi her durumda radara göre 45°'de durmaktadır ancak geminin oryantasyonu, geminin pruva-pupa hat çizgisinin X ekseni ile yaptığı açı olarak hesaplanmaktadır. Bu durumda, geminin 45° oryantasyonu için, gemi radar görüş hattı üzerinde bulunmaktadır ve menzil ekseninde en uzun göründüğü durumdur. Geminin oryantasyonu 135° olduğu durumda ise gemi ile radar görüş hattı arasında 90° açı oluşmaktadır ve gemi kerteriz ekseninde en geniş menzil ekseninde ise kısa şekilde elde edilmektedir.

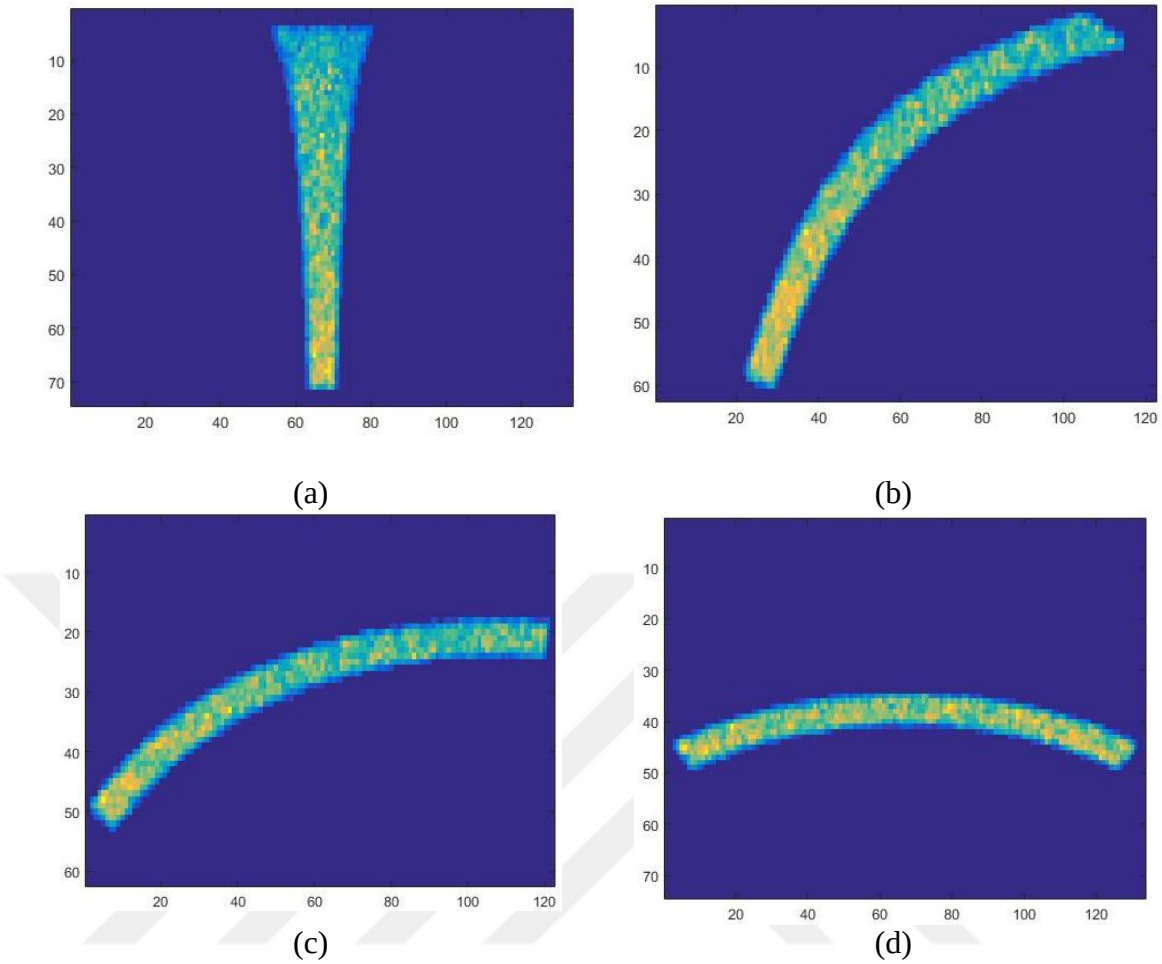
Şekil 3.1'de verilen simülatör çıktılarında, gemilerin oryantasyonlarının 45° olduğu en uygun durum verilmiştir çünkü bu oryantasyonda, hedef ile radar görüş hattı arasında 0° bulunmaktadır. Bu koşullar altında, gemilerin menzilde en uzun ve doğru ölçüm alınmaktadır. Geminin 45° kerterizde durduğu durumda tüm öznitelikler denklemler çözülerek doğru bir şekilde bulunabilmektedir.



Şekil 4.1 Radar-Gemi geometrisi temsili gösterim

Geminin baş açısı değiştikçe, elde edilen menzil profili değişmekte ve öznitelik çıkartılması zorlaşmaktadır. Şekil 4.2'de (100,100) metrede konumlanmış kargo gemisinin farklı kerteriz değerleri için elde edilmiş simülatör çıktıları görülmektedir. Kerteriz açısı 135˚ olduğunda gemi radar görüş hattına dik bir şekilde durmaktadır bu durumda kerterizde en uzun, menzilde en kısa görüntü elde edilmektedir.

Bölüm 2'den hatırlanacağı üzere, dağıtıcı ve simetri öznitelikleri yankı sinyalinin üzerindeki merkez referans alınarak sinyal dağılımının incelenmesi ile üretilen özniteliklerdir şekil 4.2'de görüldüğü üzere, 45˚ harici kerterizler için bunun uygulanması oldukça güçtür. Bu nedenle, elde edilen ham veriler işlenerek simetri ve dağıtıcı özniteliklerinin hesaplanabilir hale getirilmiştir.



Şekil 4.2 Kargo gemisi için farklı kerteriz açıları için elde edilen simülatör çıktıları
Gemi Oryantasyonu a. 45˚, b. 75˚, c. 105˚, d. 135˚

Dağıtıcı ve simetrinin elde edilebilmesi için, enerji hücreleri aralarındaki korelasyon ilişkisine bağlı olarak, hücreler kaydırılarak düz bir hat üzerinde dağılmaları sağlanmıştır ve hesaplama Bunu gerçekleştirebilmek için Bölüm 2 Kuramsal Temellerde yer alan ön işleme yöntemlerinden faydalanılmıştır. Liu vd. (2011b) uyguladığı menzil hizalama yöntemi referans alınarak enerji hücrelerinin doğrusal bir hat üzerinde sıralanmaları sağlanmıştır.

Liu vd (2011b), çalışmasında gemiler radardan yeteri kadar uzakta olduklarından dolayı radarın bir huzmesi ile geminin tamamı taranabilmektedir. Bu nedenle radarın tam bir tur taraması sonucunda N tane tek boyutlu menzil profili elde edilmektedir. Çalışmada, radarın her bir huzmesinden toplanan profiller menzil ekseninde kayma göstermektedir

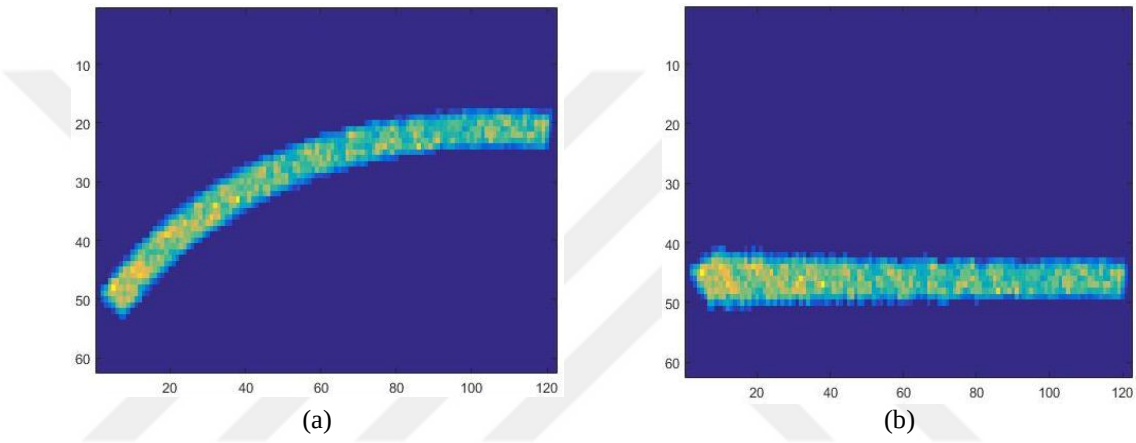
ve profiller arasındaki ilinti hesaplanarak, profillerin menzil hücrelerinin başlangıç bitiş hücreleri aynı ya da çok yakın olacak şekilde kaydırılması sonucunda aynı hizada olmaları sağlanmaktadır.

Bu çalışmada ise radarın tam bir tur dönüşünden iki boyutlu bir sinyal üretilmektedir. Bu nedenle bu iki boyutlu veri içerisinde ilinti aranarak, hedef modeli olan dikdörtgenin doğru bir biçimde çıkartılması amaçlanmıştır. İlintiyi araştırmak için bir referans noktasına ihtiyaç duyulmaktadır. Gemi hakkında en çok bilgiyi menzil ekseni sunmaktadır, bu nedenle her bir kerteriz açısı için elde edilmiş menzil sütunlarının sahip olduğu hücre sayısına bakılarak en fazla hücreye sahip olan, referans sinyal olarak seçilmiştir.

Geri kalan kerteriz açıları için elde edilmiş menzil sütunları bu referans sinyale göre hizalanmaktadır. Hizalama ölçütünün belirlenmesi için referans sinyal ile diğer kerteriz sütunları arasındaki benzerlik ilişkisi, çapraz korelasyon (cross correlation) sonucunda elde edilen değere göre belirlenmektedir. Çapraz korelasyon işlemi sonrasında gecikme vektörü ve korelasyon matrisi elde edilmektedir. Korelasyon matrisinin maksimum değere sahip olan parametresinin indis değeri, gecikme vektörüne atanarak ilgili kerteriz sütunun, menzil hücreleri boyunca ne kadar kayması gerektiği bulunmaktadır. Buna göre, tüm profil hizalanmaktadır. İşlemin uygulanmasına ait yalancı kod (pseudocode) aşağıda verilmektedir.

```
referans_sinyal = seç_max_menzil_hücresi (menzil_profili)
k_num =  menzil_profili yer alan toplam kerteriz hücresi sayısı
m_num = menzil_profili yer alan toplam menzil hücresi
for i=1'den k_num'a kadar
     gecikme = çapraz_korelasyon(referans_sinyal, menzil_profili(:,i))
    gecikme_vektörü(i) = gecikme;
end
for i=1'den k_num'a kadar
     for j=1'den m_num'a kadar
          if  gecikme_vektörü(i)==0
             hizalanmış_profil(1:m_num,i) = menzil_profili(:,i)
         else
              hizalanmış_profil(j+gecikme_vektörü(i),i) = menzil_profili(j,i)
        end
   end
end
```

Uygulamanın örnek bir sonucu şekil 4.3'te sunulmaktadır, bu örnekte Kargo gemisinin 105˚ oryantasyonu için hizalama işlemi uygulamasının öncesi ve sonrasında verinin değişimi verilmektedir. Şekilde 4.3'te görüldüğü üzere yapılan işlem aslında hücrelerin menzil ekseni üzerinde kaç hücrelik kaydırılması gerektiğinin bulunması üzerinedir. Bu işlem sonrasında hedef yine aynı sayıda kerteriz hücresi ile temsil edilmektedir. Veri bu duruma getirildikten sonra, Kuramsal Temellerde verilen dağıtıcı ve simetri özniteliklerinin bulunmasına ilişkin eşitlik (2.3-2.4) kullanılarak öznitelikler çıkartılmaktadır.



(a) (b)

Şekil 4.3 Hizalama işlemi uygulaması sonrasında profilin değişimi

a.Simülatör çıktısı Gemi Oryantasyonu 105˚, b. Simülatör çıktısı: Hizalama işlemi sonrası Gemi Oryantasyonu 105˚

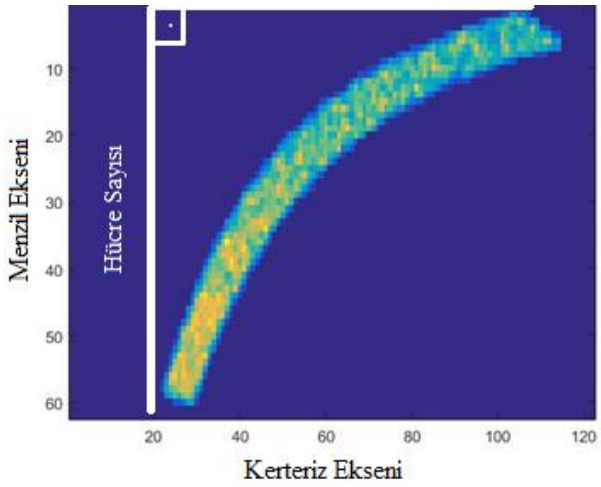
Uzunluk özniteliği, geminin baş açısına bağlı olarak değişmektedir hizalama işlemi sonrası kerteriz eksenine paralel profiller elde edilmiştir. Hizalanmış profilden, uzunluk verisinin elde edilebilmesi için yay uzunluğunun hesaplanması gerekmektedir. Ancak yay uzunluğu, geminin radara göre olan uzaklığına bağlı olduğundan dolayı değişmez bir sonuç elde edilememektedir. Bu nedenle uzunluk özniteliği simülatör çıktısı üzerinden hesaplanmaktadır.

Bu çalışmada, menzil profillerinden uzunluk özniteliği çıkarımı, menzil ekseninde sıfırdan farklı enerji seviyesine karşılık gelen hücrelerin sayısına bağlı olarak yapılmaktadır. Menzil ekseni boyunca sıfırdan farklı hücre sayısı hesaplandıktan sonra bu değer menzil çözünürlüğü ile çarpılarak uzunluk bulunmaktadır.

Bu yöntemde, menzil hücre sayısının bulunması ile ilgili matrisin satır sütun sayısının bulunması arasında bir benzerlik kurulabilmektedir. Problem, eşitlik (4.1)'de verilen 7x5'lik matrisin sıfırdan farklı satır sayısının bulunması olsun. Burada tüm satırlar, sütunlar boyunca taranmaktadır ve mühim olan ilgili satırın, kaç adet sütununda sıfırdan farklı eleman olduğu değil, ilgili satırda sıfırdan farklı eleman olup olmadığıdır yani eşitlikte verilen matrisin toplamda sıfırdan farklı 17 elemanı vardır ancak sıfırdan farklı 6 satırı bulunmaktadır.

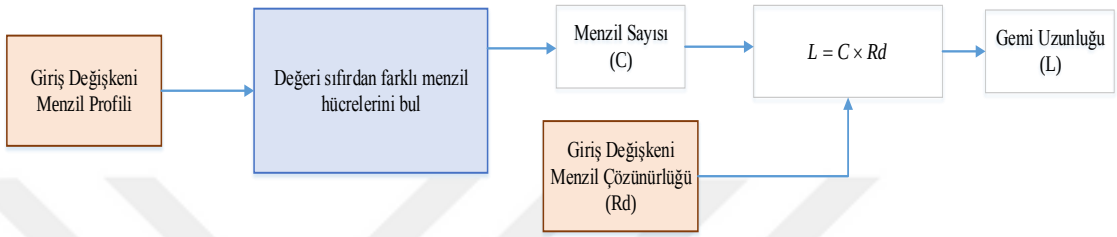
$$A = \begin{bmatrix} 0 & 0 & 3 & 4 & 1 \\ 0 & 3 & 7 & 6 & 0 \\ 2 & 8 & 5 & 3 & 1 \\ 0 & 7 & 4 & 2 & 0 \\ 0 & 0 & 5 & 3 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.1}$$

Burada uygulanan yaklaşım ile menzil ekseni boyunca kaç tane sıfırdan farklı hücre olduğu hesaplanmaktadır, yani bir matris olarak düşünülürse ve menzil ekseni satırları, kerteriz ekseni sütunları temsil etmek üzere kaç tane sıfırdan farklı satır olduğu hesaplanmaktadır. Bu yaklaşım ile gemilerin uzunluk değeri özellikle geminin baş açısı değiştikçe % 100 başarı ile bulunamamaktadır çünkü bu yöntemde şekil 4.4'te verildiği gibi her zaman, oluşturulan üçgenin, menzil ekseni ile paralel olan dik kenarı hesaplanmaktadır.



Şekil 4.4 Gemi uzunluğunun bulunmasının temsili gösterimi

Şekil 4.5'te uzunluk özniteliğinin çıkartılma adımları blok şema halinde verilmektedir. Profil üzerinde yapılan ilk işlem değeri sıfırdan farklı hücrelerin bulunmasıdır. Arama işlemi, ilk olarak kerteriz ekseninde sonrasında ise menzil ekseninde iç içe iki döngü içerisinde sorgulama ile yapılmaktadır. Bu sayede menzil ekseni boyunca hedefin kaç tane hücreye yayıldığı bulunmaktadır. Daha sonra bu değer menzil çözünürlüğü ile çarpılarak uzunluk değeri bulunmaktadır.



Şekil 4.5 Uzunluk özniteliğinin çıkarımı blok şeması

SNR özniteliği bilindiği üzere hedeften alınan sinyalin gücünün gürültünün gücüne oranıdır. RCS özniteliği, hedefin büyüklüğü ile ilgili olan bilgiyi radara geri yansıttığı enerji miktarının ölçülmesi ile hesaplanmaktadır. Uzunluk bilgisinde sonra hedefin boyutu ile ilgili bilgi verdiği için önemli bir özniteliktir.

Simülatörde hedefe ait RCS, 0 dB (Decibel) RCS sahip bir hedefe oranla dönen, hedef radar sinyalinin değeri olarak belirlenmektedir. SNR ise evre uyumsuz entegrasyon kazancı eklendikten sonraki dB cinsinden sinyal gücünün gürültü gücüne oranı olarak değeri olarak hesaplanmaktadır.

RCS ve SNR öznitelikleri de simülatör çıktısı olan veriden hesaplanmaktadır. Bu iki özelliğe ait olarak öznitelikler ortalama RCS, RCS standart sapması, ortalama SNR, SNR standart sapması olarak hesaplanmaktadır.

Elde edilen özniteliklerin aldıkları değer aralığı özellikle uzunluk, RCS ve SNR için çok geniştir. Değer skalasının geniş olması sınıflandırma işlemi için negatif etki oluşturacağı için özniteliklere Theodoridis ve Koutroumbas (2009) referans alınarak ve

yumuşak ölçeklendirme (softmax scaling) uygulanmış ve değerlerin [0,1] arasında değişmesi sağlanmıştır.

Yumuşak ölçeklendirme işlemi eşitlik (4.2-4.4) aralığında verilen işlem dizisi uygulanarak yapılmaktadır. Normalleştirme işlemi gerçekleştirilmek istenen öznitelik kümesi $x = [.]_{lxK}$ boyutlu bir vektör olarak temsil edilmektedir. İlk olarak eşitlik (4.2)'de verildiği gibi ortalama değer bulunmaktadır. Bunu takiben varyans eşitlik (4.3)'te verildiği şekilde hesaplanmaktadır. Son olarak, eşitlik (4.4)'te verildiği üzere normalleştirme yapılmaktadır.

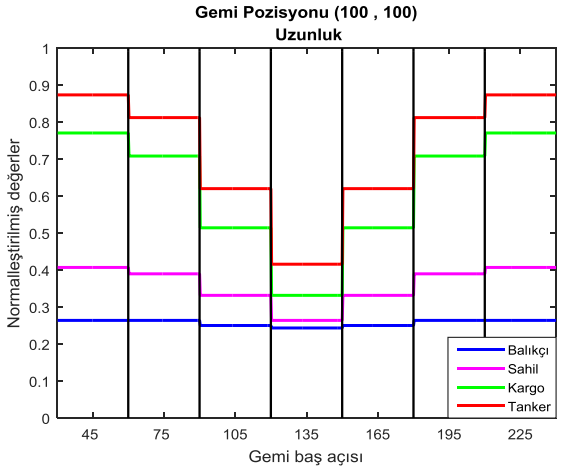
$$\acute{x}_k = \frac{1}{N}\sum_{i=1}^{N} x_{ik} k = 1,2,\cdots,l \tag{4.2}$$

$$\sigma_k^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_{ik} - \acute{x}_k)^2 \tag{4.3}$$

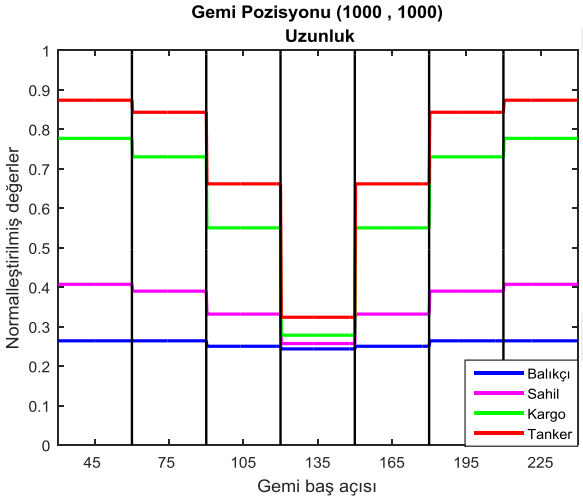
$$y = \frac{x_{ik} - \acute{x}_k}{r\sigma_k}$$

$$\hat{x}_{ik} = \frac{1}{1 + exp(-y)} \tag{4.4}$$

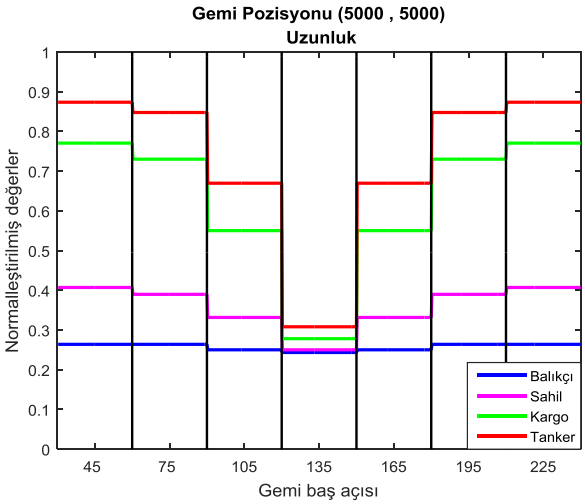
Radar simülatöründen her bir gemi için, yedi farklı baş açısı (45°, 75°, 105°, 135°, 165°, 195°, 225°), radardan üç farklı uzaklıkta konumlandırılmış halde toplamda 84 farklı senaryo simülatörde koşturulmuştur ve % 95 tespit olasılığı için simülatör çıktıları üretilmiştir. Üretilen görüntülerden, uzunluk, simetri, dağıtıcı, ortalama/standart sapma RCS ve SNR öznitelikleri çıkartılmıştır. Çıkartılan özniteliklere, yumuşak ölçeklendirme uygulanması sonrası gemilere türlerine, baş açılarına ve radara göre olan uzaklıklarına göre olan dağılımları şekil 4.6-4.12'de verilmiştir.



(a)



(b)



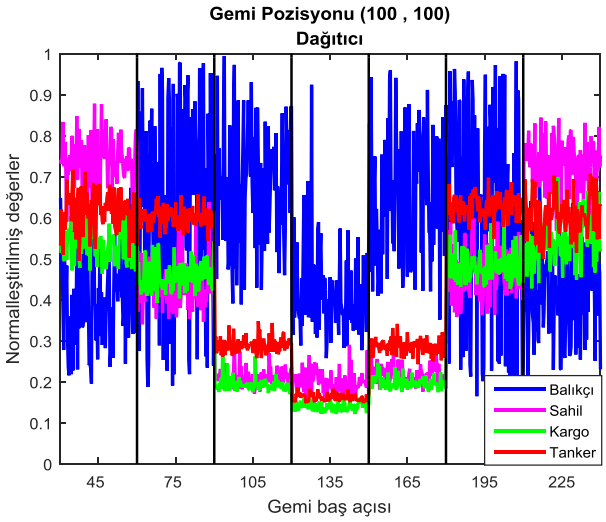
(c)

Şekil 4.6 Farklı konumlar ve baş açıları için uzunluk özniteliğinin değişimi

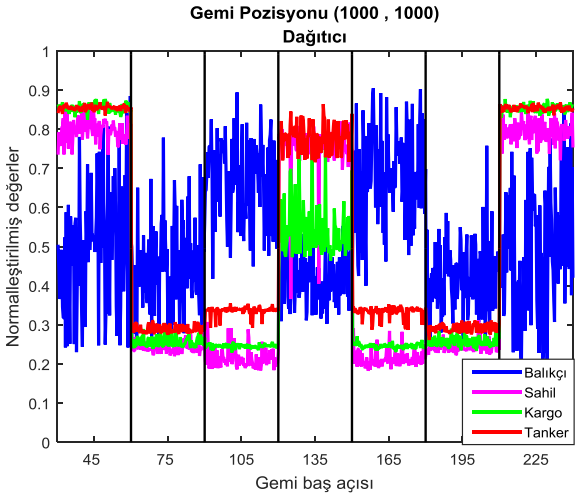
Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m

Şekil 4.6'da verilen sonuçlar incelendiğinde, uzunluk özniteliğinin 135° baş açısı etrafında simetrik bir şekilde değiştiği gözlenmektedir. Gemiler en uzun 45˚ ve 225˚'de hesaplanmıştır çünkü bu açı değerlerinden gemiler ile radar görüş hattı arasındaki açı 0 derecedir. Her bir baş açısı için farklı uzunluk değerleri hesaplanmıştır bunun nedeni gemilerin radar görüş hattına göre aralarındaki açı arttıkça kerteriz ekseninde daha fazla hücre ile temsil edilmektedir bu da menzil ekseninde kısalmalarına sebep olmaktadır. Bu nedenle 135° baş açısında gemiler en kısa şekilde hesaplanmıştır bu durum şekilde de görüleceği üzere sınıfların birbirine karışmasına neden olmaktadır. Bu durumlara rağmen, uzunluk özniteliğinin iyi bir ayırıcı özellik gösterdiği görülmektedir.

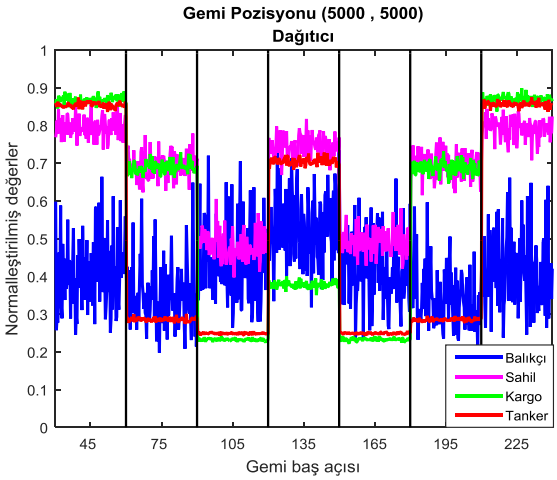
Şekil 4.7'de verilen sonuçlar incelendiğinde dağıtıcı özniteliğinin değerlerinin sınıflar arası karıştığı gözlenmektedir. Dağıtıcı özniteliği özellikle balıkçı teknesinde büyük salınımlar göstermektedir. Bunun sebebinin balıkçı teknesinin çok daha az sayıda hücre ile temsil edilmesinden kaynaklı olarak merkezden uzaklaştıkça hücrelerden birinin yüksek değerli olması eşitliği sonucu değişmesidir. Dağıtıcı özniteliği şekillerden de görüleceği üzere geminin baş açısından da etkilenmektedir. Gemi radardan uzaklaştıkça elde edilen dağıtıcı sonuçları arasındaki dalgalanma azalmaktadır.



(a)



(b)

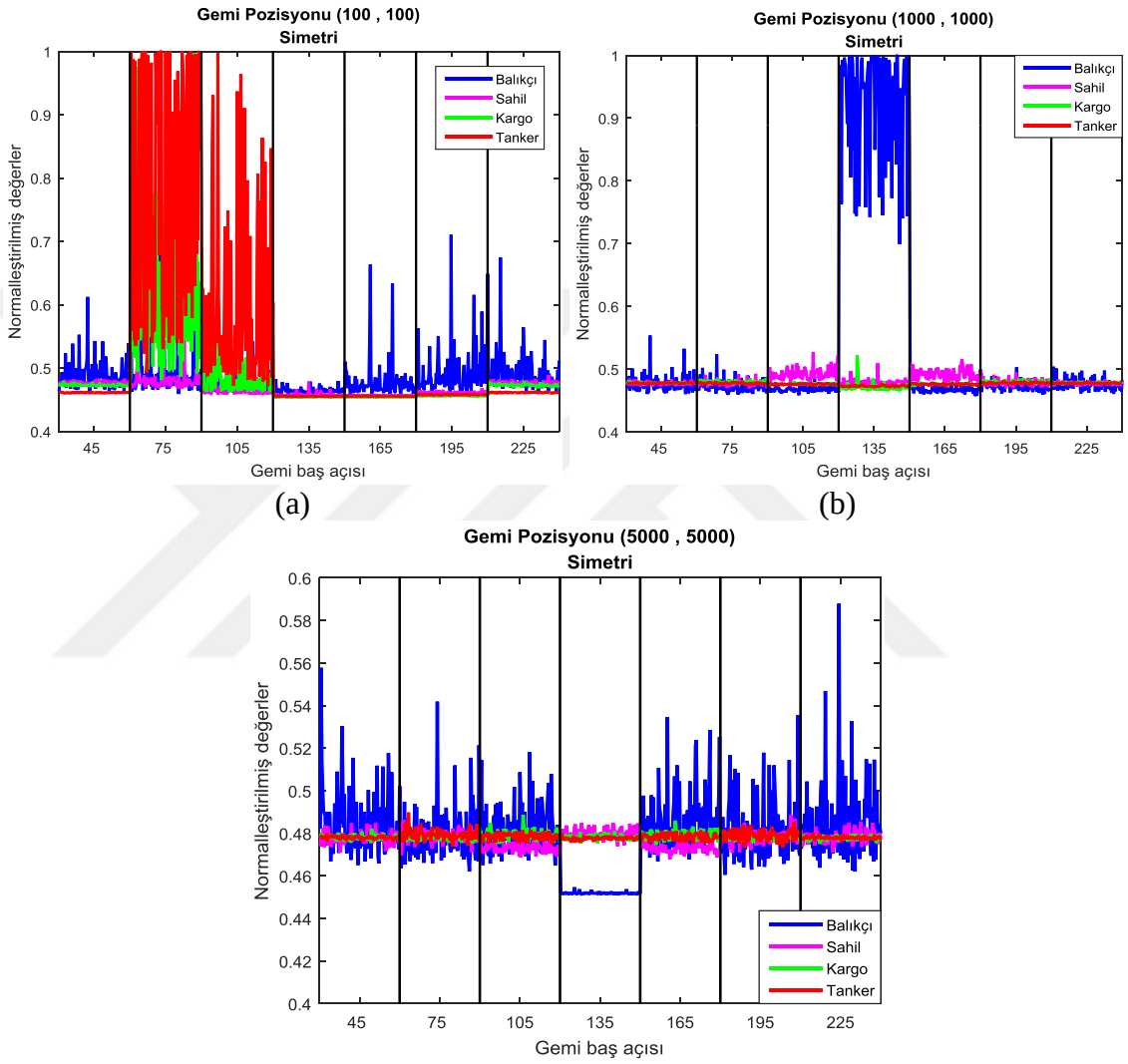


(c)

Şekil 4.7 Faklı konum ve baş açıları için dağıtıcı özniteliği değişimi

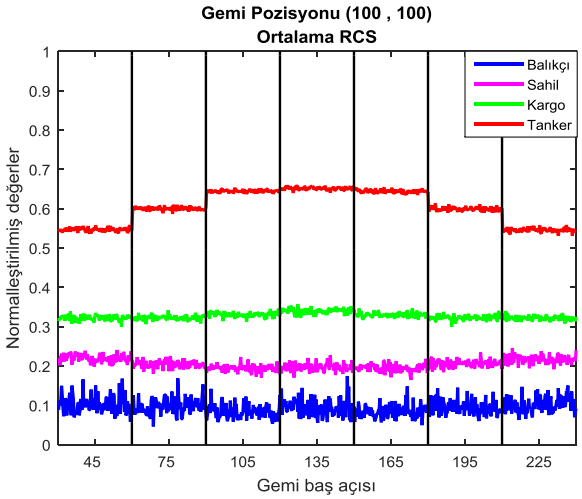
Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m

Şekil 4.8'de verilen simetri özniteliğinin değişim grafikleri incelendiğinde simetri özniteliğinin neredeyse tüm gemiler için aynı değere sahip olduğu görülmektedir. Bu durum, gemilerin dikdörtgen şeklinde modellenmesi ve gemiye karakteristik olan, anten, baca, direk gibi özelliklerin modellenememesinden kaynaklanmaktadır.

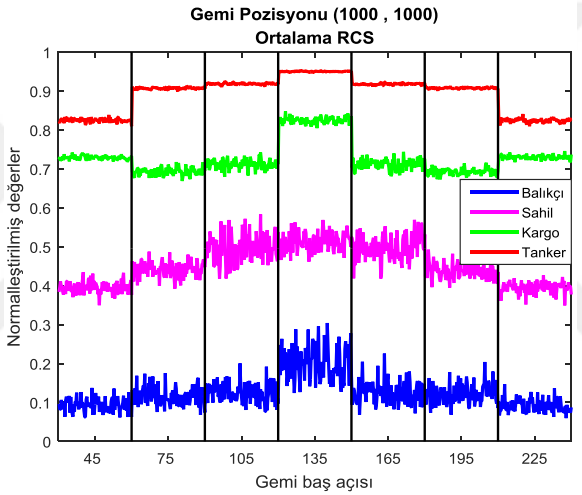


Şekil 4.8 Faklı konum ve baş açıları için simetri özniteliği değişimi

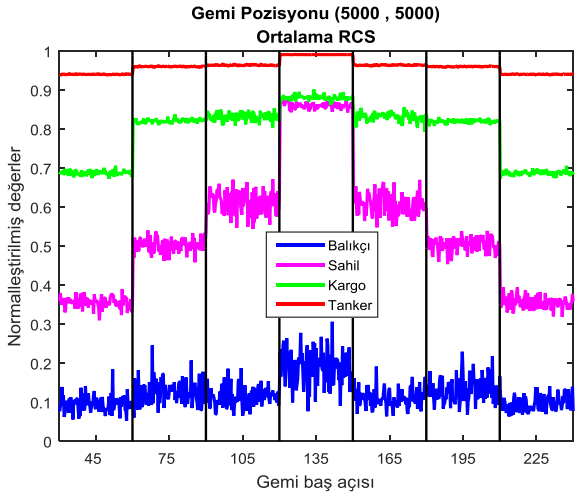
Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m



(a)



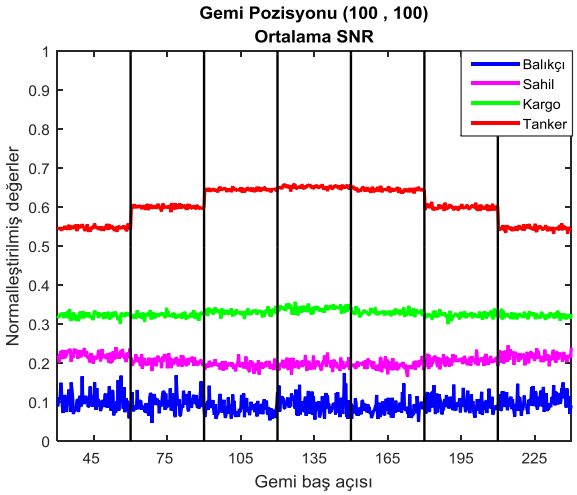
(b)



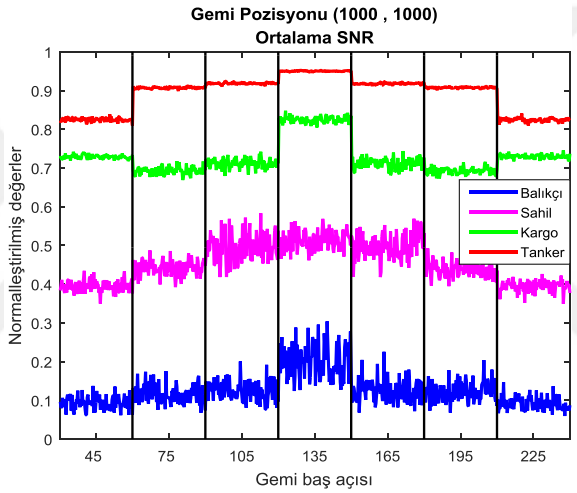
(c)

Şekil 4.9 Faklı konum ve baş açıları için ortalama RCS özniteliği değişimi

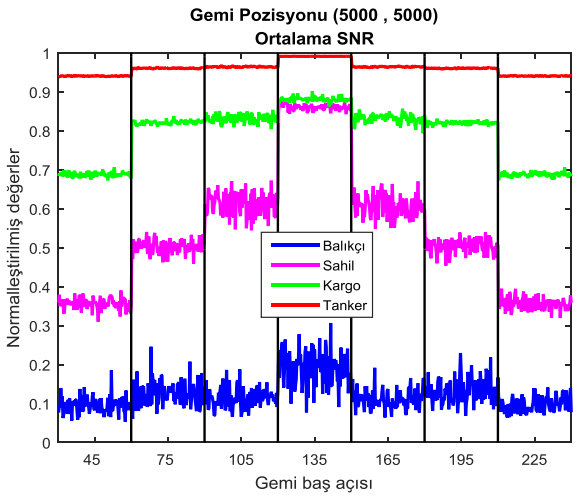
Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m



(a)



(b)



(c)

Şekil 4.10 Faklı konum ve baş açıları için ortalama SNR özniteliği değişimi

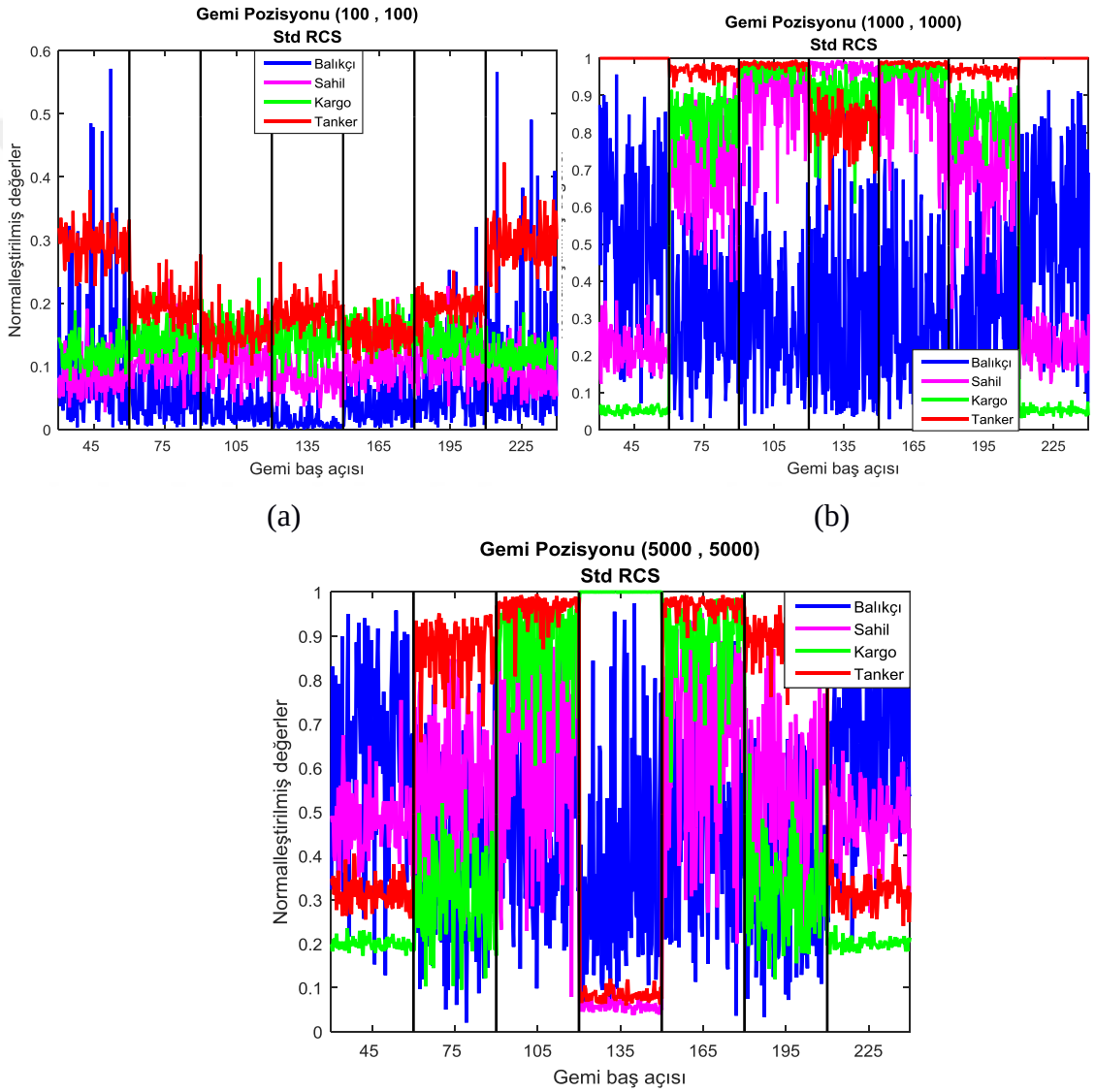
Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m

Şekil 4.9 - 4.10'da verilen sonuçlar incelendiğinde, RCS ve benzer şekilde SNR ortalama değerlerinin, gemi, radarın konuşlandığı kıyıdan uzaklaştıkça arttığı görülmektedir. Bunun nedeni eşitlik (2.5) incelendiğinde anlaşılmaktadır, RCS değeri, radar-hedef arasındaki mesafenin karesi ile değişmektedir. Radar-hedef arasındaki mesafe arttıkça, radar huzmesi genişlemekte ve hedef üzerinde oluşturduğu etkin alan artmaktadır. Ortalama SNR değerinin artmasının sebebi ise eşitlik (2.6-2.7) incelendiğinde görülmektedir, istenen sinyalin gücü hesaplanmasında RCS yani etkin alan değeri kullanılmaktadır, dolayısı ile uzaklık arttıkça etkin alan arttığı için istenen sinyalin gücü artmakta bu da SNR artmasına neden olmaktadır.

RCS ve SNR ortalama değerleri incelendiğinde, değerlerin bir simetriye sahip oldukları görülmektedir. Merkezde yani baş açısının 135˚ olduğu konumda gemiler kendi içlerinde en yüksek RCS değerine sahip olmuşlar ve daha sonra değerler merkeze göre azalarak devam etmiştir. Geminin baş açısının 135˚ olduğu konumda, gemiler radar görüş hattı üzerinde maksimum alanı radar huzmesi içerisinde yer almaktadır bunun tam tersi olarak da baş açısının 45˚ veya 225˚ olduğu durumda minimum alan yer aldığı için RCS'de en küçük değerini bu konumda almaktadır. Gemilerin baş açılarına göre duruşları simetrik olduğundan dolayı (45˚, 225˚), (75˚, 195˚), (105˚, 165˚), elde edilen ortalama RCS ve SNR değerleri de simetrik özellik göstermektedir.

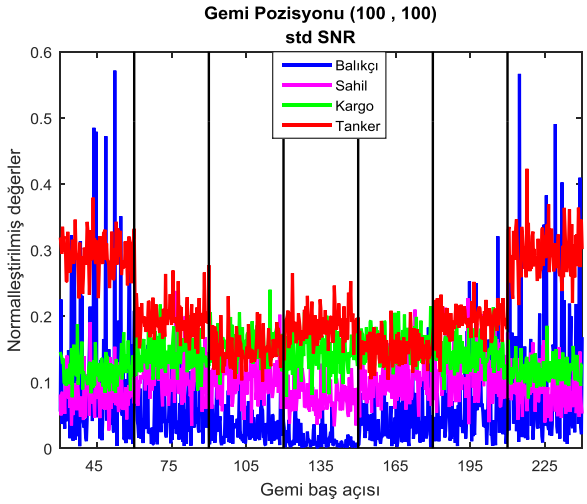
Şekil 4.9.a, yani gemilerin kıyıya yakın oldukları durum incelendiğinde ortalama RCS değerlerinin çok değişmediği görülmektedir. Bunun nedeni, yakın bölgede radar huzmesinin kısıtlı bir alana etki etmesidir, bu durum aynı zamanda ortalama RCS değerinin gemilerin baş açılarından etkilenmemesini sağlamaktadır. Gemiler kıyıdan uzaklaşmaya başladığında ise, şekil 4.9.b, özellikle büyük boy gemilerde (Tanker ve Kargo gemisi) baş açısının değişimine bağlı olarak ortalama RCS değerinin değiştiği görülmektedir. Sahil güvenlik gemisi ve balıkçı teknesinde ise bu durum çok daha az seviyededir. Gemilerin, kıyıya göre konumlarının en uzak olduğu durumda ise, şekil 4.9.c, gemilerin baş açılarının ortalama RCS değerine etkisi net bir şekilde görülmektedir. Gemilerin baş açısının 135˚ olduğunda gemiler maksimum alanı gösterecek şekilde konumlandığı için ortalama RCS değerleri en yüksek değeri almıştır, en büyük gemi olan Tanker için bu değer 1'e çok yakın çıkmıştır. Balıkçı teknesi, diğer

gemilere göre en uzak mesafede bile baş açısının oluşturduğu değişimi net olarak gösterememektedir bunun en büyük nedeni çizelge 3.2'de verilen gemilerin en-boy oranından kaynaklanmaktadır, kendisine en yakın büyüklükte olan sahil güvenlik gemisi, balıkçı teknesi 1.7 kat daha büyüktür. Geminin boyutları küçüldükçe baş açısına göre duruştan etkilenme seviyeleri düşmektedir. Sonuç olarak şekil 4.9-4.10'da verilen tüm grafikler incelendiğinde yer yer gemiler arası karışmalar gözlense de ortalama RCS ve SNR ayırıcı özellik gösterebileceği görülmektedir.

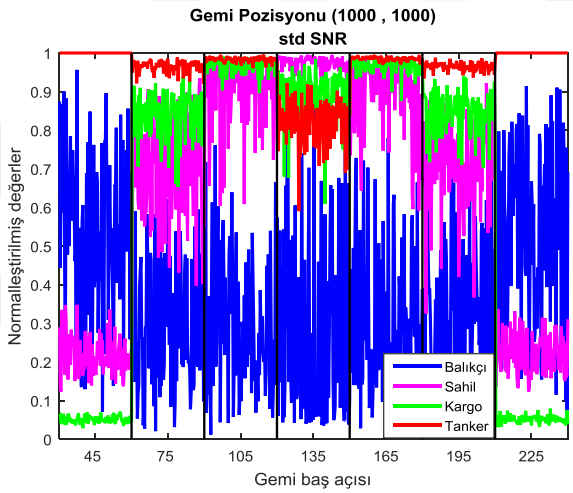


Şekil 4.11 Faklı konum ve baş açıları için standart sapma RCS özniteliği değişimi

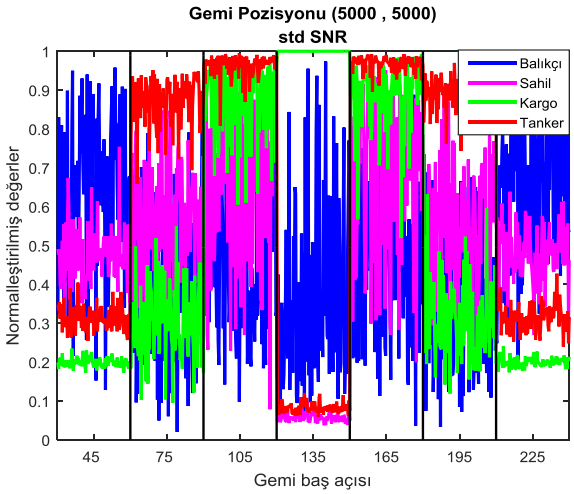
Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m



(a)



(b)



(c)

Şekil 4.12 Faklı konum ve baş açıları için standart sapma SNR özniteliği değişimi Gemi Konum a. (100,100) m, b. (1000,1000) m, c. (5000,5000) m

Standart sapma, veri kümesindeki her bir elemanın kümenin ortalama değeri etrafında nasıl dağıldığının ölçüsüdür (Bertsekas ve Tsitsiklis, 2008). Standart sapmanın, [0,1] değer aralığında değiştiği düşünüldüğünde, standart sapmanın sıfır veya sıfıra yakın çıkması veri kümesindeki elemanların ortalama değere yüksek yakınlıkta çıktığı anlamına gelmektedir. Standart sapmanın bire yakın çıkması ise kümedeki elemanların geniş bir aralıkta saçıldığı anlamına gelmektedir. Şekil 4.11-4.12 incelendiğinde standart sapması en çok değişen geminin balıkçı teknesi olduğu görülmektedir, bunun iki sebebi bulunmaktadır; birincisi ortalama değerlerinin çok salınım göstermesi, ikincisi ise balıkçı teknesi küçük boyutlu bir gemi olduğu için menzil profilinde çok daha az hücre ile temsil edilmektedir bu da ortalamaya göre uç değerler alan elemanların etkisinin yüksek gözlenmesine sebep olmaktadır.

Gemilerin baş açılarına göre de standart sapma değeri değişmektedir, örneğin şekil 4.9.c ve 4.11.c incelendiğinde, 135˚ baş açısında, tanker ve sahil güvenlik gemisinin ortalama değerleri çok minimal değiştikleri için standart sapmaları da o baş açısı için minimum olmuştur. Sonuç olarak, ortalama değerler baş açısına, uzaklığa göre değiştiği için şekil 4.11 ve 4.12'den görüleceği üzere, standart sapma değerleri de buna bağlı olarak oldukça karışık, birbirine karışan şekilde elde edilmiştir. Bu nedenle, RCS ve SNR standart sapma değerlerinin sınıflandırma için uygun bir öznitelik olamayacağı öngörülmektedir.

Grafiklerde verilen özniteliklerden, 8400 elemandan oluşan veri kümesi elde edilmiştir, WEKA'da % 66'sı eğitim (5544 adet veri) % 34'ü test kümesi (2856 adet veri) olarak kullanılmıştır. WEKA'da koşturulan sınıflandırıcıların performans sonuçları çizelge 4.1-4.2'de verilmektedir.

Özniteliklerin grafiklerine bakıldığında dağıtıcı, simetri ve RCS ile SNR standart sapma değerlerinin sınıflar arası karıştığı gözlenmekteydi ancak sınıflandırma sonuçlarına bakıldığında yüksek başarılar elde edildiği görülmektedir. En yakın komşu (KNN) sınıflandırıcısı % 100 başarı, MLP, Karar Ağacı ve Bayes ağı % 90'dan fazla başarı göstermiştir.

Çizelge 4.1 Tüm öznitelikler kullanıldığında elde edilen sınıflandırma başarıları

Sınıflandırıcı	Hata Matrisi	Başarı
En Yakın Komşu	$\begin{bmatrix} a & b & c & d \\ 734 & 0 & 0 & 0 \\ 0 & 737 & 0 & 0 \\ 0 & 0 & 717 & 0 \\ 0 & 0 & 0 & 668 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
Naive Bayes	$\begin{bmatrix} a & b & c & d \\ 726 & 8 & 0 & 0 \\ 28 & 661 & 48 & 0 \\ 0 & 45 & 351 & 321 \\ 0 & 0 & 207 & 461 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 76.99
Bayes Ağı	$\begin{bmatrix} a & b & c & d \\ 731 & 2 & 1 & 0 \\ 1 & 735 & 1 & 0 \\ 0 & 3 & 713 & 0 \\ 0 & 1 & 1 & 666 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 99.65
Karar Ağacı	$\begin{bmatrix} a & b & c & d \\ 730 & 4 & 0 & 0 \\ 0 & 736 & 1 & 0 \\ 0 & 0 & 717 & 0 \\ 0 & 0 & 0 & 668 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 99.83
MLP	$\begin{bmatrix} a & b & c & d \\ 734 & 0 & 0 & 0 \\ 1 & 719 & 17 & 0 \\ 0 & 16 & 696 & 1 \\ 0 & 0 & 165 & 503 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 92.86
SVM	$\begin{bmatrix} a & b & c & d \\ 734 & 0 & 0 & 0 \\ 0 & 730 & 7 & 0 \\ 0 & 74 & 409 & 234 \\ 0 & 28 & 36 & 604 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 86.73

Bu başarıları sırasıyla % 86 ile SVM ve % 77 ile Bayes sınıflandırıcısı takip etmiştir. Yanlış sınıflandırma yoğunlukla orta-büyük-çok-büyük sınıfları arasında gerçekleşmiştir. Sınıflandırma başarısına paralel olarak çizelge 4.2'de sınıflandırıcıların, hassasiyet değerleri, doğru yanlış tahminler, dengeli-dengesiz veriye karşı sergiledikleri performans değerleri de yüksek çıkmıştır. Sınıflar arası karışan özniteliklerin yanında, ayırma özelliği yüksek olan diğer özniteliklerin baskın olmalarının sınıflandırma başarısını getirmiştir. Bu durumda, yedi özniteliğin hepsinin kullanılması yerine ayırma performansı yüksek daha az sayıda öznitelik kullanılması düşünülmüştür ve bu kapsamda bir sonraki bölümde yer alan özniteliklerin gürbüzlükleri araştırılmıştır.

Çizelge 4.2 Tüm öznitelikler kullanıldığında sınıflandırıcıların doğruluk analizi

Sınıflandırıcı	TP Oranı	FP Oranı	Hassasiyet	Recall	F-Ölçümü	MCC	ROC Alanı	PRC Alanı
En Yakın Komşu	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0
Naive Bayes	0.770	0.074	0.770	0.770	0.768	0.695	0.952	0.867
Bayes Ağı	0.996	0.001	0.997	0.996	0.997	0.995	1.0	1.0
Karar Ağacı	0.998	0.001	0.998	0.998	0.998	0.998	0.999	0.997
MLP	0.929	0.024	0.940	0.929	0.928	0.910	0.987	0.965
SVM	0.867	0.042	0.879	0.867	0.860	0.829	0.954	0.839

Sınıflandırıcıların başarıları karşılaştırıldığında ise ortaya ilginç bir tablo çıkmaktadır en yakın komşu % 100 sınıflandırma başarısı göstermiştir. Çizelge 4.1'de verilen KNN sınıflandırıcısının komşu sayısı 1 seçilmiştir ve uzaklık hesabı Öklit uzaklığı kullanılarak yapılmaktadır. KNN sınıflandırıcısı, test verisi ile eğitim kümesindeki veriler arasındaki uzaklığı hesaplamaktadır. Bu hesaplamadan test verisi, hangi eğitim verisine en yakın çıkıyorsa onun ait olduğu sınıfa atanmaktadır. Şekil 4.6-4.12 arasında özniteliklerin değişimine bakıldığında bazı özniteliklerin iyi ayırıcı özellik gösterdiği görülmektedir. Bu durum, KNN'in sahip olduğu yaklaşım dikkate alındığında yüksek başarıya sahip olmasına neden olmaktadır. KNN'in komşu sayısı $k = 3, 5, 7, 9$ için elde edilen sonuçlarda da % 100 sınıflandırma başarısı elde edilmiştir.

Karar ağacı sınıflandırıcısı ise bölüm 2.4.4'te verildiği üzere ikili ağaç yapısına sahiptir ve karar kuralları dilendiği kadar sıkı konulabilmektedir. Kuralların sıkı konması da başarıyı beraberinde getirmektedir lakin genelleştirme ve uyarlanır olma yönünü köreltmektedir. Ağ yapıları, girdilere belirli ağırlık katsayısı ile çarparak, seçilen karar fonksiyonun en uygun sonucu üretmesini sağlamaktadır. Bu nedenle bu sınıflandırıcılarda eğitim esnasında, ayırıcı özelliği yüksek olan öznitelikler için seçilen ağırlık katsayısı ile diğer özniteliklerin katsayıları farklı olmaktadır. Bu durum, ayırıcı özelliği yüksek olan özniteliklerin karar verilmesinde baskın olmasını sağlamakta ve sınıflandırma başarısı artmaktadır.

SVM sınıflandırıcısı, genel bir çözüm ürettiği ve geometrik tabanlı olduğu için sınıflandırma performansı biraz geride kalmıştır. SVM sınıfları geometrik sınırlarla ayırdığı için sinir ağları gibi fonksiyon çıktısına bağlı daha esnek bir karar verme mekanizmasına sahip değildir bu nedenle performansı diğer sınıflandırıcılara göre daha geride kalmasına rağmen yüksek bir başarıya sahiptir. Naive Bayes sınıflandırıcısı bu grupta en kötü performansı sunmuştur. Naive Bayes sınıflandırıcısı, olasılık dağılım fonksiyonu çıkartmaktadır, sınıflandırmanın doğru yapılabilmesi dağılım fonksiyonun doğru olmasına bağıdır bu fonksiyon veri sayısı ile ilişkilidir. Burada performansının düşük olması eğitim veri sayısının yetersiz kalmasından kaynaklanıyor olabileceği düşünülmektedir.

4.2 Gürbüz Özniteliklerin Tespit Edilmesi ve Sınıflandırma Sonuçları

Sınıflandırma başarısının arttırılması, uygun model seçilmesi, verinin özellikleri ve yeterli sayıda olmasının yanı sıra çıkartılan özniteliklerin ayırma performansının yüksek olmasına bağlıdır. Bu nedenle seçilen özniteliklerin bireysel başarılarının değerlendirilmesi sınıflandırma işleminin başarısında büyük öneme sahip olmaktadır.

Şekil 4.6-4.12'de gemilerin üç farklı konumu için elde edilen verilerden çıkartılan özniteliklerin sınıflara göre değişimleri verilmiştir. Bu şekillerde, özniteliklerin değişimleri analiz edildiğinde bazı özniteliklerin karıştığı gözlenmiştir. Bu nedenle tez çalışması kapsamında çıkartılan özniteliklerden hangilerinin sınıflandırma başarısına yüksek katkı sağladığı, gürbüz olduğu belirlenmeye çalışılmıştır. Bunun sonucu olarak da daha az sayıda daha yüksek sınıflandırma başarısı elde edilmesi amaçlanmıştır.

Sınıflandırma işleminde öznitelik sayısının azaltılmasının olumlu sonuç getireceği Jain vd. (2000) çalışmalarında yer almıştır. Buna göre, sınıflandırma eğitim aşamasının görece az veri kullanılıyor ise öznitelik sayısının arttırılmasının sınıflandırma başarısına katkıdan çok zarar getireceğini vurgulamışlardır. Öznitelik sayısının fazla olmasının sınıflandırma başarısına etki edebilecek negatif yan etkilerinin yanı sıra, işlem yükünü, hafıza kapasitesini arttırması gibi etkileri de bulunmaktadır. Özniteliklerin gürbüzlüklerinin tespit edilmesi adına Slomka vd. (1999) ve Jain vd. (2000) çalışmaları

incelenmiştir. Jain vd (2000) önerilen ve Bölüm 2'de yer alan öznitelik seçilmesi algoritmaları arasında bulunan SFS algoritması uygulanmıştır.

SFS algoritması, açgözlü (greedy) arama algoritma ailesine ait, $k < d$ olmak üzere d boyutlu öznitelik uzayını k boyutuna indirmeyi amaçlamaktadır. Böylelikle problem ile daha çok ilintili olan özniteliklerin oluşturduğu bir alt küme elde edilmektedir. Bu da, daha az işlem ve hata getirmektedir. En iyi alt kümeyi seçmek için öncelikle alt küme, boş küme olarak seçilmektedir. Daha sonra, özniteliklerin bireysel başarıları araştırılmaktadır. Bireysel başarıları en yüksek olan öznitelikten başlanarak, alt kümeye teker teker öznitelikler eklenerek istenen başarıya ulaşılıp ulaşılmadığı kontrol edilmektedir. Bu algoritmada geriye dönüş olmadığı için tüm ihtimaller değerlendirilememektedir bu sebeple bu çalışmada genelleştirilmiş SFS kullanılmıştır ve tüm ihtimaller değerlendirilmiştir.

SFS algoritması uygulanırken kullanılan başarı ölçütü olarak SVM, KNN ve Naive Bayes (NB) sınıflandırıcıları seçilmiştir. SVM'nin seçilmesinin nedeni, ortalama bir başarı göstermiş olmasıdır ve öznitelik kümesinin değişiminin başarıyı arttırıp arttırmayacağı araştırılmıştır. Ek olarak SVM, geometrik tabanlı bir sınıflandırıcı olduğu için sınıflar arası ayrımda özniteliklerin etkisinin daha net tespit edileceği düşünülmektedir. En yakın komşu sınıflandırıcısı ise en yüksek başarıyı, Naive Bayes ise en kötü başarıyı gösterdiği için, sınıflandırma başarısını azaltıp/arttıracak öznitelikler tespitinin daha iyi yapılacağı düşünülmektedir. KNN sınıflandırıcısında k değeri arttırıldığında sınıflandırma performansı arttığı için en düşük performans gösterdiği $k = 1$ için elde edilen sonuçlar çizelge 4.3'te verilmektedir.

Çizelge 4.3'te sunulan sonuçlar değerlendirildiğinde bireyselde en başarısız sınıflandırma sonucu veren simetri özniteliği olduğu görülmektedir. Bu durum, simülatörde gemilerin dikdörtgen şeklinde modellenmesinden kaynaklanmaktadır. Şekil 2.3'te geminin inşasına ilişkin özelliklerinin HRRP nasıl etkilediği verilmiştir, simetri özniteliği de profildeki bu bölgesel değişimlerden faydalanmaktadır ve geminin boyutsal özelliklerinden etkilenmemektedir. Benzer şekilde, simetri özniteliğini başarısız sınıflandırmada RCS ve SNR standart sapma değerleri takip etmektedir. En

yakın komşu sınıflandırıcısı bu özniteliklerde en düşük performansı göstermiştir. Bu değerlerde sınıflar arası karışmanın çok olması ve KNN sınıflandırıcısının en yakın olduğu birden fazla komşu varsa burada kullanılan algoritmada ilkini seçmesi gibi bir yaklaşım olması bu başarısızlığa neden olmuştur.

Çizelge 4.3 Çeşitli öznitelik kümelerinin sınıflandırma performansları

Ort. RCS	Ort. SNR	Uzunluk	Std RCS	Std SNR	Dağıtıcı	Simetri	SVM Başarı	KNN Başarı	NB Başarı
X							% 63.76	% 89.88	% 59.38
	X						% 63.76	% 89.88	% 59.38
		X					% 68.59	% 96.39	% 68.59
			X				% 31.72	% 39.25	% 32.21
				X			% 31.72	% 39.25	% 32.21
					X		% 26.93	% 47.69	% 39.74
						X	% 23.88	% 40.19	% 28.39
X		X					% 74.01	% 99.75	% 75.21
X	X	X					% 75.28	% 99.75	% 73.17
X	X						% 63.62	% 89.88	% 60.08
X	X	X	X				% 82.24	% 99.96	% 73.17
X		X	X				% 82.28	% 99.96	% 73.07
X	X	X	X	X			% 82.4	% 99.96	% 75.98
X	X	X			X		% 79.48	% 99.93	% 74.02
X	X	X	X	X	X		% 86.34	% 100	% 75.91
X	X	X	X	X		X	% 86.34	% 99.93	% 75.67
X	X	X	X	X	X	X	% 86.73	% 100	% 76.99

RCS ve SNR öznitelikleri bireyselde aynı başarıyı üretmişlerdir. Şekil 4.6 hatırlanacağı üzere ürettikleri değerler ve grafikler oldukça benzerlik göstermektedir bu nedenle bireysel başarıları da aynı çıkmıştır. En başarılı öznitelik ise her üç sınıflandırıcıda da en yüksek başarıyı elde eden uzunluk özniteliği olmuştur. Uzunluk özniteliğinin çıkartıldığı senaryolarda başarının hemen düştüğü çizelge 4.3'te görülebilmektedir.

Özniteliklerin bireyselde en çok hangi gemileri/sınıfları bir biri ile karıştırdığı çizelge 4.4'te verilen hata matrisinden anlaşılabilmektedir. Bu hata matrisi SVM sınıflandırıcısının sonuçlarını içermektedir. SVM sınıflandırma yaklaşımı, özniteliklerin gürbüzlüğünün tespit edilebilmesi için daha uygun olduğu düşünülmüştür.

Ortalama RCS ve SNR büyük boy gemi sınıfına ait olan kargo gemisini orta boy ve çok büyük boy olarak yanlış sınıflandırmıştır. RCS gemilerin boyutları göre değeri değişen bir öznitelik olduğundan dolayı bazı baş açılarında veya uzak mesafelerde gemilerin olduğundan büyük veya küçük görünmesinden dolayı bu karışmanın yaşandığı düşünülmektedir. Şekil 4.6'da oluşan karışmalar incelenebilmektedir.

Uzunluk özniteliği en iyi performansı gösteren özniteliktir buna rağmen orta boy sınıf (sahil güvenlik gemisi) ile küçük boy sınıf (balıkçı teknesi) veya kargo gemisinin yani büyük boy sınıf (kargo gemisi) orta boy sınıf (sahil güvenlik gemisi) ait gemilerin bir kısmı yanlış sınıflandırılmıştır. Bunun nedeni gemilerin baş açılarının 135° olduğu durumdur. Çizelge 4.5'te gemilerin gerçek boyları ve baş açılarına göre yukarıda bahsedilen yönteme göre hesaplanan uzunluk bilgileri verilmektedir.

Çizelge 4.4 Özniteliklerin bireysel sınıflandırma performansları hata matrisleri

Öznitelik	Hata Matrisi
Ortalama RCS	$\begin{bmatrix} a & b & c & d \\ 684 & 50 & 0 & 0 \\ 2 & 608 & 81 & 46 \\ 0 & 242 & 0 & 475 \\ 0 & 0 & 139 & 529 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy
Ortalama SNR	$\begin{bmatrix} a & b & c & d \\ 684 & 50 & 0 & 0 \\ 2 & 608 & 81 & 46 \\ 0 & 242 & 0 & 475 \\ 0 & 0 & 139 & 529 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy
Uzunluk	$\begin{bmatrix} a & b & c & d \\ 734 & 0 & 0 & 0 \\ 100 & 637 & 0 & 0 \\ 67 & 38 & 212 & 400 \\ 0 & 100 & 192 & 376 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy
Std RCS	$\begin{bmatrix} a & b & c & d \\ 511 & 40 & 12 & 171 \\ 407 & 42 & 8 & 280 \\ 440 & 4 & 0 & 273 \\ 315 & 0 & 0 & 353 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy
Std SNR	$\begin{bmatrix} a & b & c & d \\ 511 & 40 & 12 & 171 \\ 407 & 42 & 8 & 280 \\ 440 & 4 & 0 & 273 \\ 315 & 0 & 0 & 353 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy
Dağıtıcı	$\begin{bmatrix} a & b & c & d \\ 0 & 230 & 435 & 69 \\ 0 & 348 & 353 & 36 \\ 0 & 253 & 417 & 47 \\ 0 & 316 & 348 & 4 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy
Simetri	$\begin{bmatrix} a & b & c & d \\ 33 & 0 & 0 & 701 \\ 0 & 0 & 0 & 737 \\ 4 & 0 & 0 & 414 \\ 19 & 0 & 0 & 649 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy

-

Çizelge 4.5 Gemilerin baş açılarına göre elde edilen profillerden çıkartılan uzunluk değerleri

Gemi	45°	75°	105°	135°	165°	195°	225°	Gerçek Değer
Balıkçı Teknesi	15	15	10	7.5	10	15	15	9
Sahil Güvenlik Gemisi	60	55	37.5	15	37.5	55	60	55
Kargo Gemisi	170	147.5	90	37.5	90	147.5	170	167
Orta Boy tanker	220	187.5	120	62.5	120	187.5	220	213

Çizelge 4.5'te hesaplanan değerler göz önüne alındığında sınıflar arasında karışma olmaması kaçınılmazdır, 135° baş açısında radar geminin eni boyunca veri topladığı için gemilerin boyu ile ilgili doğru çıkarım yapılamamaktadır bu nedenle 135° baş açısı verileri çıkartılmış ve uzunluk özniteliğinin sınıflandırma performansının değişimi çizelge 4.6'da verilmektedir.

Sonuçlar incelendiğinde, 135° baş açısı verisinin çıkartılması ile uzunluk özniteliğinin sınıflandırma başarısı % 10 artmıştır, hata matrisleri incelendiğinde orta boy sınıfa ait geminin küçük sınıfa atanması, büyük boy sınıfına ait geminin orta veya küçük boy sınıfına atanması hatası düzelmiştir.

Tüm bu sonuçlar doğrultusunda öznitelikler değerlendirildiğinde, uzunluk, RCS'nin ortalama değeri ve standart sapma değeri kullanılarak başarının arttığı gözlenmiştir. Dağıtıcı ve simetri bireysel çok başarısız olduklarından dolayı öznitelik kümesinden çıkartılmıştır. SNR özniteliği RCS ile aynı performansı gösterdiği için, SNR öznitelik kümesinden çıkartılmıştır.

Çizelge 4.6 Baş açısının uzunluk özniteliğinin sınıflandırma başarısına etkisi

Yöntem	Hata Matrisi	Başarı
Tüm baş açıları	$\begin{bmatrix} a & b & c & d \\ 734 & 0 & 0 & 0 \\ 100 & 637 & 0 & 0 \\ 67 & 38 & 212 & 400 \\ 0 & 100 & 192 & 376 \end{bmatrix} \begin{matrix} a \rightarrow Küçük \\ b \rightarrow Orta \\ c \rightarrow Büyük \\ d \rightarrow ÇokBüyük \end{matrix}$	% 68.59
135° baş açısı çıkartılmış	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 293 & 332 \\ 0 & 0 & 196 & 423 \end{bmatrix} \begin{matrix} a \rightarrow Küçük \\ b \rightarrow Orta \\ c \rightarrow Büyük \\ d \rightarrow ÇokBüyük \end{matrix}$	% 78.43

Yeni öznitelik kümesi kullanılarak ve 135° baş açısı çıkartılmış olan verilerin, % 66'sının eğitim kümesi olarak kullanıldığı, WEKA aracı kullanılarak elde edilen sınıflandırma başarıları çizelge 4.7'de verilmektedir.

Çizelge 4.7 incelendiğinde, çizelge 4.1'de sonuçlara göre iyileşme olduğu görülmektedir bundaki en büyük pay 135° baş açısı çıkartılarak uzunluk özniteliğinin ayırma gücünün arttırılmasıdır. Bir diğer önemli durum öznitelik kümesinin boyutu 7'den 3'e düşürülmesi sınıflandırma performansında bir düşüşe neden olmamıştır. Bundan sonraki çalışmalarda azaltılmış öznitelik kümesi kullanılmıştır.

Bu çalışmanın bir diğer önemli sonucu, uzunluk özniteliğinin gemilerin boyutlarına göre sınıflandırılmasında önemli bir etkisi olduğu görülmüştür. Uzunluk özniteliğinin yüksek doğrulukta elde edilmesi sınıflandırma performansı bakımından önem taşımaktadır.

Çizelge 4.7 Öznitelik alt kümesi ile elde edilen sınıflandırma başarıları

Sınıflandırıcı	Hata Matrisi	Başarı
En Yakın Komşu	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
Naive Bayes	$\begin{bmatrix} a & b & c & d \\ 618 & 3 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 75 & 366 & 184 \\ 0 & 0 & 212 & 407 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 80.63
Bayes Ağı	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
Karar Ağacı	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
MLP	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 490 & 135 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 94.48
SVM	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 474 & 151 \\ 0 & 28 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 93.83

4.3 Uzunluk Özniteliğinin Gemilerin Baş Açısına Göre Hesaplaması ve Sınıflandırma Sonuçları

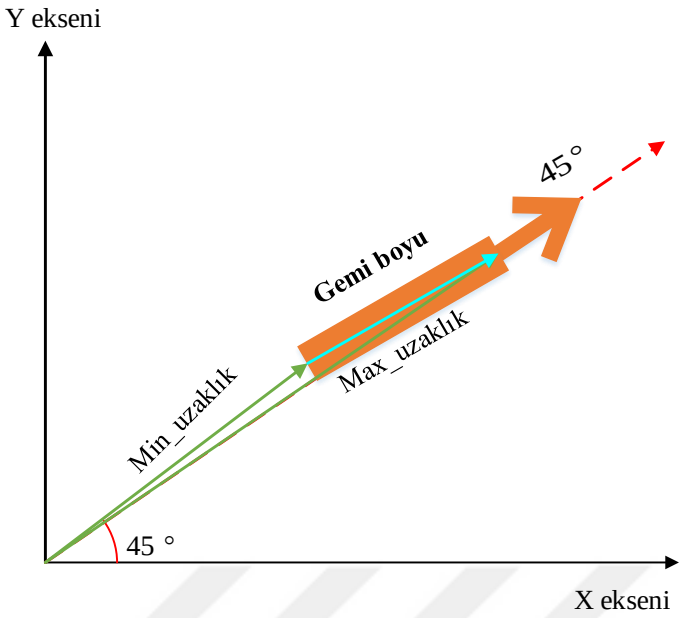
Özniteliklerin sınıflandırma üzerine bireysel performansları değerlendirildiğinde, gemilerin boyutlarına göre sınıflandırılmasındaki en önemli ve ayırma kuvveti yüksek olan özniteliğin uzunluk olduğu bulunmuştur. Uzunluk özniteliğinin yüksek doğrulukta hesaplanamamasının önündeki en büyük engel gemilerin baş açılarına bağlı olarak menzil ekseninde kısalmalarıdır. Bu problemi ortadan kaldırmak için gemilerin

kinematik özelliklerinden faydalanılmıştır. Gözetleme radarları hedeflerin hızları konumları, baş açıları ve benzeri kinematik özellikleri hakkında bilgi sağlamaktadırlar.

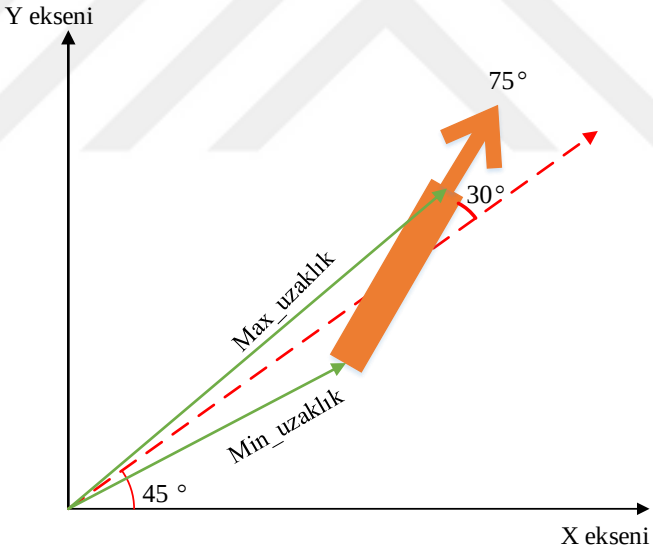
Bu bölümde, hedefe ait menzil profiline ek olarak hedefin baş açısına ve radara göre olan uzaklık bilgisine sahip olunduğu kabul edilerek hesaplamalar gerçekleştirilmiştir. İlgilenilen problemde, radarın görüş hattı ile x ekseni arasındaki açı 45°'dir bu sebeple geminin menzilde en uzun olduğu baş açısı 45°'dir. Baş açısı 45° olduğunda gemi ile radar arasındaki ilişki ve geminin uzunluğunun nasıl bulunduğu şekil 4.7'de verilmektedir. Şekilde, merkezde radarın var olduğu kabul edilmektedir, radarın merkezinden geçen görüş hattı kırmızı kesikli çizgi ile gösterilmektedir, gemi bu hat üzerinde konumlandığında baş açısı 45°'dir ve geminin boyu geminin radara olan en uzak noktasından en yakın noktasının çıkartılması ile bulunabilmektedir, bu durum eşitlik (4.5)'te verilmiştir.

$$gemi_{uzunluğu} = max_{uzaklık} - min_{uzaklık} \tag{4.5}$$

Eşitlik (4.5)'te yöntem ile eşitlik (2.2) verilen menzil hücresi ve çözünürlüğüne dayalı hesap kullanılmadan doğrudan geminin uzunluğu bulunabilmektedir. Diğer baş açıları için ise uzunluk hesaplanması 45° referans alınarak yapılmaktadır. Şekil 4.8'de 75° baş açısına sahip geminin radar görüş hattına göre nasıl durduğu görülmektedir. Şekil 4.9'da baş açısı 75° için uzunluk değerinin nasıl hesaplandığı verilmektedir. İlk olarak geminin radara olan maksimum ve minimum uzaklık değerleri bulunmaktadır.



Şekil 4.7 Baş açısı 45°için geminin radar göre duruşu



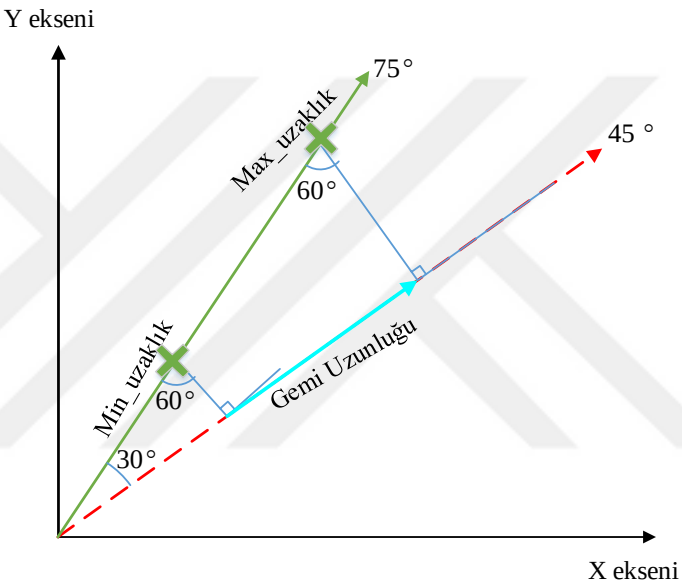
Şekil 4.8 Baş açısı 75°için geminin radar göre duruşu

Maksimum ve minimum uzaklıklar bulunduktan sonra şekil 4.9'da verildiği üzere hesaplanan uzunluklar radarın görüş hattı üzerine yani 45° hattı üzerinde iz düşümü alınmaktadır. Alınan iz düşüm sonrası 45° baş açısına sahip geminin uzunluğunun bulunması adımı işletilerek geminin uzunluğu bulunmaktadır. İz düşümü bulmak için kullanılan açı geminin baş açısının 45° hattı ile yaptığı açıdır (θ), bu değer 75° ve 195°

baş açıları için 30°, 105° ve 165° baş açıları için 60° olmaktadır ve eşitlikler aşağıda verilmiştir.

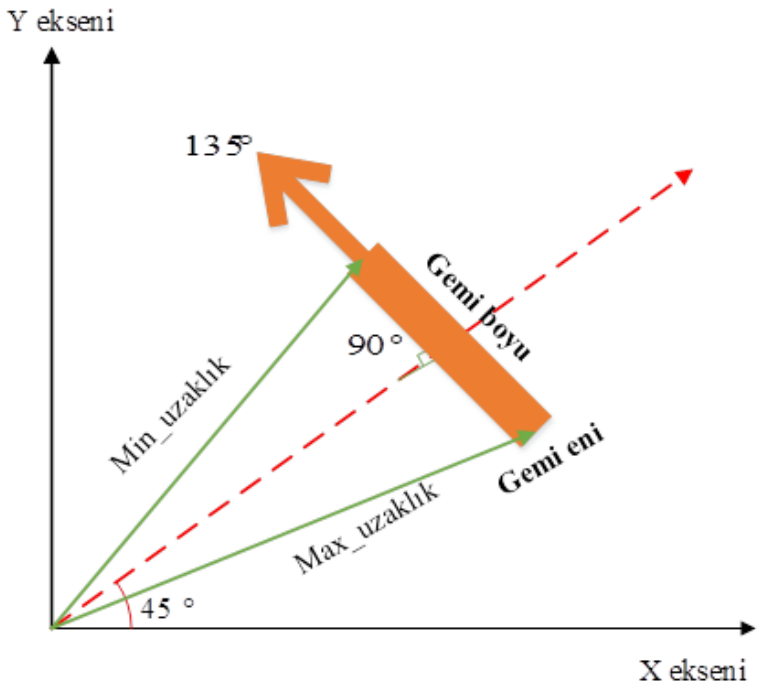
$$min45_{uzaklık} = min_{uzaklık} \cos\theta$$
$$max45_{uzaklık} = min_{uzaklık} \cos\theta \quad (4.6)$$

$$gemi_{uzunluğu} = max45_{uzaklık} - min45_{uzaklık} \quad (4.7)$$



Şekil 4.9 Baş açısı 75° için örnek uzunluk hesaplama gösterimi

Bu yöntem baş açısı 135° oluğu durum için uygulanamamaktadır çünkü bu duruşta gözlenebilirlik problemi meydana gelmektedir. Şekil 4.10'da, baş açısının 135° olduğu durum verilmektedir, burada hedefe olan minimum ve maksimum uzaklık birbirine eşittir ve radar tarama sırasında geminin enini menzilde, boyunu ise kerteriz ekseninde taramaktadır. Bu senaryoda, geminin boyunun hesaplanabilmesi için kosinüs teoremi uygulanabilir fakat gemi radardan uzaklaştıkça aradaki açı küçülüp, yayın kolları uzunluğu artacağı için geminin boyu olduğundan çok daha büyük hesaplanacaktır. Bu nedenlerden dolayı baş açısının 135° olduğu durum için çözüm üretilememektedir.



Şekil 4.10 Baş açısı 135°için geminin radar göre duruşu

Yeni yöntem uygulanarak hesaplanan uzunluk değerleri ile bölüm 4.2'de hesaplanan ve çizelge 4.5'te sunulan uzunluk değerleri, gemilerin gerçek uzunlukları ile karşılaştırmalı olarak çizelge 4.8'de verilmektedir.

Çizelge 4.8 Kinematik bilgi varlığında gemilerin baş açılarına göre uzunluk hesaplama

Gemi	45°	75°	105°	165°	195°	225°	Gerçek Değer
Balıkçı Teknesi	12.5	14.43	12.46	12.46	14.43	12.5	9
Sahil Güvenlik Gemisi	57.5	60.62	62.76	61.74	60.2	57.5	55
Kargo Gemisi	167.5	167.43	176.24	176.24	167.43	167.5	167
Orta Boy tanker	217.5	213.61	227.28	227.28	213.61	217.5	213

Çizelge 4.5 ve çizelge 4.8 incelendiğinde kinematik bilgi varlığında geliştirilen yeni yöntemin açı değişimine daha gürbüz karşılık verdiği görülmektedir. Sınıflandırma başarıları, bölüm 4.2'de belirlenmiş olan alt öznitelik kümesi kullanılarak, toplamda 7200 verinin % 66'sının eğitim, % 34'nün test verisi olarak WEKA'da koşturulması sonucunda elde edilen sonuçlar çizelge 4.9'da verilmektedir.

Çizelge 4.9 Kinematik bilgilerden hesaplanan uzunluğun yer altığı öznitelik alt kümesinin sınıflandırma sonuçları

Sınıflandırıcı	Hata Matrisi	Başarı
En Yakın Komşu	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
Naive Bayes	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 620 & 5 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 99.80
Bayes Ağı	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
Karar Ağacı	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
MLP	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100
SVM	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 625 & 0 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 100

Çizelge 4.9, incelendiğinde sınıflandırma başarılarında çizelge 4.7'ye göre büyük bir iyileşme olduğu görülmektedir buradan da uzunluk özniteliğinin doğru hesaplanmasının sınıflandırma başarısına etkisi doğrudan görülmektedir. Bir diğer önemli nokta ise azaltılan öznitelik kümesi ile bu sonuçlara erişilmesidir. Bunların sonucu olarak sınıflandırıcı başarısı için çok öznitelik değil ayırıcı özelliği yüksek, bağımsız

öznitelikler gerektiği sonucu çıkarılmaktadır. Çizelge 4.7'de sınıflandırıcıların güçlü ve zayıf yönlerine bağlı olarak sınıflandırma performansları değişkenlik göstermekteydi, burada elde edilen sonuçlara bakıldığında ise sınıflandırma başarısının ilk olarak doğru öznitelik kümesinden geçtiği anlaşılmaktadır.

Uzunluk iyileşmesinin etkisinin daha iyi analiz edilebilmesi için RCS ve SNR ortalama ve standart sapma özniteliklerinin dahil olmadığı bir öznitelik kümesi için sınıflandırma başarımı çalışması yapılmıştır ve sonuçlar çizelge 4.10'da sunulmaktadır. Çizelge 4.10 sunulan sonuçlarda öznitelik kümesi içerisinde uzunluk, simetri ve dağıtıcı öznitelikleri bulunmaktadır.

Elde edilen sonuçlar incelendiğinde, uzunluk özniteliğinin daha doğru hesaplandığı kinematik verilere dayanan yöntemden elde edilen öznitelik kümesinin sınıflandırma başarısını arttırdığı görülmektedir. Özellikle Naive Bayes sınıflandırıcısının başarısı yaklaşık % 13 kadar artmıştır. Çok katmanlı algılayıcı ve SVM sınıflandırıcılarının başarıları artarak % 100 başarılı sınıflandırma yapmışlardır.

Çizelge 4.10 Simetri, Dağıtıcı ve Uzunluk öznitelik kümesinin farklı uzunluk hesaplama yöntemlerine göre sınıflandırma başarısının değişimi

Sınıflandırıcı	Uzunluk Özniteliği Menzil Profiline göre hesaplanmış Başarı	Uzunluk Özniteliği Baş Açısına göre hesaplanmış Başarı
En Yakın Komşu	% 99.98	% 100
Naive Bayes	% 86.83	% 99.62
Bayes Ağı	% 99.87	% 99.98
Karar Ağacı	% 100	% 100
MLP	% 93.5	% 100
SVM	% 91.61	% 100

4.4 Radar Tespit Olasılığının Sınıflandırma Başarısına Etkisi

Radar tespit olasılığı, radarın, taradığı alandaki tüm hedeflere oranla, yankı sinyalinden tespit ettiği hedef sayısının yüzde cinsinden ifade edilmesidir. Eğer bir radarın tespit olasılığı % 80 ise radarın görüntülediği 100 noktadan 80'ini tespit ettiği anlamına gelmektedir. Özetlemek gerekirse, radar tespit olasılığı arttıkça taranan hedefe ait daha çok bilgi elde edilmektedir.

Bu bölüme kadar yapılan tüm çalışmalarda radarın % 95 tespit olasılığına sahip olduğu senaryo üzerinden sonuçlar paylaşılmıştır. Bu bölümde ise farklı tespit olasılıklarının sınıflandırma başarısı üzerinde bir etkisi olup olmadığı araştırılmıştır.

Radar simülatöründe, tespit olasılıkları % 75, % 85 ve % 95 olmak üzere veriler üretilmiştir. Bu veriler üzerinde bölüm 4.3'te verilen gemi uzunluğu bulma yöntemi ile elde edilen uzunluk öznitelikleri ile bölüm 4.2'de oluşturulan öznitelik alt kümesinin elemanları olan RCS ortalama ve standart sapma değerleri kullanılmıştır.

Altı farklı baş açısı ve üç farklı konum için elde edilen öznitelikler, WEKA'da % 66'sı eğitim geri kalanı test kümesi olacak şekilde seçilmiştir. Elde edilen sonuçlarda, En yakın komşu, Bayes, Bayes Ağı, Karar Ağacı, SVM, MLP sınıflandırıcılarından sadece Bayes sınıflandırıcısında radar tespit olasılığına tepki verdiği gözlenmiştir. Diğer sınıflandırıcılar çizelge 4.9'da verilen değerlerini almışlardır. Bayes sınıflandırıcısının radar tespit olasılığına göre sınıflandırma başarımı çizelge 4.11'de verilmektedir.

Çizelge 4.11 incelendiğinde radar tespit olasılığı arttıkça Bayes sınıflandırıcısının da başarısının arttığı gözlenmiştir. Gemiden toplanan veri sayısının artması, hesaplanan olasılık dağılım fonksiyonun iyileşmesini sağlamış olabileceği düşünülmektedir. Sinir ağlarında birçok katmandan geçerek süzülmüş veri ile karar verilmesi, SVM'de geometrik sınırların en büyük aralık olacak şekilde seçilmesi, karar ağaçlarında kuraların her yeni koşul için yeniden oluşturulması gibi nedenlerden dolayı bu sınıflandırıcılarda tespit olasılığının etkisi gözlenmemiş olabilmektedir.

Çizelge 4.11 Bayes sınıflandırıcısının radar tespit olasılığına göre sınıflandırma başarısı

Tespit Olasılığı	Hata Matrisi	Başarı
% 75	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 553 & 72 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 97.05
% 85	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 564 & 61 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 97.5
% 95	$\begin{bmatrix} a & b & c & d \\ 621 & 0 & 0 & 0 \\ 0 & 583 & 0 & 0 \\ 0 & 0 & 620 & 5 \\ 0 & 0 & 0 & 619 \end{bmatrix}$ a: Küçük Boy b: Orta Boy c: Büyük Boy d: Çok Büyük Boy	% 99.80

5. TARTIŞMA VE SONUÇ

Bu tezde, kıyıya yakın gemilerin yüksek menzil çözünürlüklü radar kullanılarak sınıflandırılması problemi incelenmiştir. Bu çalışma kapsamında radar simülatöründen elde edilen benzetim verileri kullanılarak balıkçı teknesi, sahil güvenlik gemisi, kargo ve tanker gemileri boyutlarına göre küçük, orta, büyük ve çok büyük boy olmak üzere sınıflandırılması problemi üzerinde çalışılmıştır. Radar tarafından iletilen sinyalin gemiden geri yansımasından elde edilen yüksek çözünürlüklü menzil profilinin (HRRP) analizi yapılarak gemilere ait karakteristik bilgi sunacak veriler hesaplanmaya çalışılmıştır. Bu kapsamda HRRP'den, gemilere ait uzunluk, simetri, dağıtıcı, RCS ve SNR öznitelikleri çıkartılmıştır. Gemilerin sınıflandırılması amacıyla dört farklı yaklaşıma sahip altı adet sınıflandırıcı kullanılmış ve performansları değerlendirilmiştir. Bu kapsamda olasılık tabanlı yaklaşım modeli altında Bayes ve Bayes ağı, geometrik tabanlı yaklaşımda destek vektör makineleri, çok katmanlı algılayıcı, benzetim tabanlı yaklaşımda k-en yakın komşu ve kural tabanlı yaklaşımda karar ağacı sınıflandırıcıları kullanılmıştır.

Özniteliklerin sınıflandırmaya olan etkileri araştırılmıştır ve elde edilen sonuçlarda simetri ve dağıtıcı özniteliklerinin bireysel performanslarının düşük olduğu gözlenmiştir. Simetri ve dağıtıcı öznitelikleri gemilerin direk, baca, kamara vb. gibi yapısal özelliklerine ilişkin yansıyan sinyal seviyelerinin dağılımına göre bir oran sunmaktadır. Çalışma kapsamında gemilerin yapısal özelliklerini içeren bilgisayar benzetim teknolojisi içeren bir tasarım kullanılmadığı için, beklenen sinyal seviyesi değişimleri elde edilememiştir. Bu durum simetri ve dağıtıcı özniteliklerinden beklenen performansın altında kalmalarına sebebiyet vermiştir. Özellikle askeri ve sivil gemilerin ayırt edilmesinde simetri ve dağıtıcı öznitelikleri büyük önem taşıdıkları için bu özniteliklerin sınıflandırma problemlerinde göz ardı edilmemesi gerekmektedir.

Uzunluk özniteliği, bireysel performansı en yüksek olan özniteliktir bunun en büyük etkenlerinden biri gemilerin boy-en oranının yüksek olmasıdır. Bu durum gemilerin boylarıyla birbirinden farklılaşmalarını sağlamaktadır. Bu nedenle, gemilerin özellikle boyutlarına göre sınıflandırılmasında uzunluk özniteliğinin doğru bir şekilde elde

edilmesi büyük önem taşımaktadır. Gemilerin baş açılarına göre radar sinyalinden her durumda doğru uzunluk bilgisinin çıkartılabileceği HRRP elde edilememiştir. Uzunluk özniteliğinin daha doğru bir şekilde hesaplanabilmesi için gemilerin kinematik özelliklerinden faydalanılmıştır. Bu kapsamda HRRP verisinin yanında, geminin baş açısı ve geminin baş ve kıç noktalarının radara olan uzaklıklarının bilindiği kabul edilmiştir. Gözetleme radarları tarafından bu tip bilgiler elde edilebilmektedir. Kinematik veri bilgisi varlığında uzunluk özniteliğinin çok daha doğru hesaplandığı ve bu sayede sınıflandırma başarısının arttığı sonucu elde edilmiştir.

Sınıflandırıcılar farklı yaklaşımlara sahip olsalar da benzer başarılar ortaya koymuşlardır. Özellikle uzunluk özniteliğinin kinematik bilgiler varlığında hesap edilmesi sonucunda elde edilen veriler girdi olarak kullanıldığında, tüm sınıflandırıcılar % 100 veya % 100'e yakın başarılar kaydetmişlerdir. Bu nedenle sınıflandırıcıları, performansları anlamında ayırt etmek güç olmasına karşın sahip oldukları yaklaşımlarına bağlı olarak bazı avantajlara ve dezavantajlara sahiptirler. Tez çalışmasında ilgilenilen problemde Karar Ağacının yüksek performans göstermesinin arkasında yatan sebep çok sıkı bir kural dizisi oluşturmasıdır. Karar ağacı sınıflandırıcısı kural tabanlı bir sınıflandırıcı olmasından dolayı yeni veri girişi yapılmasına, sınıf sayısının arttırılmasına uyarlanır bir şekilde cevap verememektedir ve modelin en baştan oluşturulması gerekmektedir. Bu nedenle, koşulları ve gereksinimleri güncellenebilir sistemler için kullanışlı bir sınıflandırıcı değildir.

Çok katmanlı algılayıcı sınıflandırıcısının ağ yapısının karmaşıklığı, kullanılan öznitelik vektörünün boyutuna, sınıf sayısına ve seçilen karar fonksiyonuna göre değişebilmektedir. Ağ ne kadar karmaşıklaşırsa işlem süresi o kadar uzun olmaktadır. Kullanılan sınıflandırıcılar içerisinde en yavaş olan sınıflandırıcı çok katmanlı algılayıcıdır ancak ağın eğitimi sürecinde öznitelikleri performanslarına göre analiz ederek bağlantı katsayılarını buna göre atamaktadır. Bu durum daha çok fayda gösteren özniteliğin karar fonksiyonuna etkisinin artmasını daha az fayda gösteren özniteliğin karar fonksiyonuna etkisinin azalmasını sağlamıştır. Bu nedenle MLP, tez çalışmasındaki en kötü senaryoda bile % 90'ın üzerinde bir başarı kaydetmiştir.

Bayes ile Bayes ağı sınıflandırıcısı azaltılmış veri kümesi için benzer performanslar göstermişlerdir. Radar tespit olasılığının değişiminden sadece Bayes sınıflandırıcısı etkilenmiştir. Tespit olasılığı düştükçe sınıflandırma performansı düşmüştür. Buna karşın Bayes ağı bu durumdan etkilenmemiştir, bunun nedeni bir ağ yapısına sahip olması ve ağın son katmanına ulaşıldığında süzülmüş veri üzerinden karar mekanizmasının işletilmesidir. Bayes ağı da, MLP sınıflandırıcısının sahip olduğu ağın karmaşıklaşması ve işlem süresinin artması problemini yaşamaktadır.

Bayes sınıflandırıcısı, ilk veri kümesinde en kötü performansı gösterirken, gürbüz özniteliklerin seçildiği senaryoda başarısını arttırmıştır. Fazla sayıda öznitelik ve bunların sınıflar arası karışması doğru olasılık dağılım fonksiyonun hesaplanmasını mümkün kılmadığı düşünülmektedir. Bayes sınıflandırıcısını tercih edilebilir kılan en büyük avantajı, kritiklik gerektiren durumlarda riskin derecesinin belirlenip buna göre hesaplamaların yapılabilmesidir ancak bu durum sınıfların veya özniteliklerin güncellenmesi durumunda riskin de tekrar değerlendirilip modelin yenilenmesi gerektirebilir.

K-en yakın komşu sınıflandırıcısı parametrik olmayan bir model olduğundan dolayı özniteliklere bağlı bir karar oluşturmamaktadır. Mevcut verilerin dağılımına bakılarak maksimum benzerliğe sahip olunan sınıfa atama yapılmaktadır. Bu durumda eğer veri kümesi içerisinde, sınıflar arası ayrıklaşmayı sağlayan öznitelikler bulunuyorsa yüksek başarı gösterirken, sınıflar arası karışma başladığı durumda ise yanlış sınıflandırmaya başlamaktadır. Test aşaması için tüm eğitim kümesi kullanıldığı için hafıza etkin bir yöntem değildir.

Destek vektörleri makinesi, sınıfları geometrik olarak bölerek sınıfları birbirlerinden olabildiğince uzak kalacak şekilde sınırlandırmaya çalışmaktadır. Destek vektörleri makinesinin en büyük avantajı değişimlere uyarlanır cevap verebilmesi ve model üzerinde bir değişim yapılmasına ihtiyaç duyulmamasıdır.

Sonuç olarak, gemilerin sınıflandırılması probleminde sınıflandırma başarısını arttıran temel parametrenin özniteliklerin yüksek doğrulukta elde edilmesi ve doğru

özniteliklerin seçilmesi olduğu gözlenmiştir. Sınıflandırma için uygun özniteliklerin yüksek güvenirlikte elde edildiği sürece sınıflandırma başarısının artacağı gözlenmiştir. Kurgulanan sistemden beklenen, işlem hızı, hafıza kapasitesi, uyarlanabilirlik vb. isteklere göre hangi sınıflandırıcının kullanılacağına karar verilmelidir.

Bu tez çalışmasının geliştirilebileceği gelecek çalışmalar içerisinde, özniteliklerin rasgele değişken olarak ele alınıp olasılık dağılımlarının hesaplanarak sınıflandırma işleminin gerçekleştirilmesi yer almaktadır. Bu çalışmada, dört farklı boyuta sahip dört gemiden elde edilen veriler kullanılmıştır, gemi sayısı arttırılarak daha detay bir sınıflandırma işlemi gerçekleştirilebilir. Bilgisayar simülasyon teknolojisi kullanılarak, gemilerin üç boyutları modelleri çıkartılıp, daha gerçekçi benzetim verileri üzerinden sınıflandırma işlemi gerçekleştirilebilir.

KAYNAKLAR

Aha, D.W. Kibler, D. and Albert, M.K. 1991. Instance-Based Learning Algorithms. Machine Learning, 6:37-66.

Alpaydın E. 2010. Introduction to Machine Learning. The MIT Press, 515, USA

Balanis, C.A. 1982. Antenna Theory Analysis and Design. John Wiley & Sons Inc, 959, USA.

Bertsekas, D.P. and Tsitsiklis, J.N. 2008. Introduction to Probability. Athena Scientific, Belmont, Massachusetts, 528, USA.

Bishop, C.M. 2006. Pattern Recognition and Machine Learning. Springer, 749, USA.

Chan, S. and Lee K. 2013. Radar Target Recognition by MSD Algorithm on Angular-Diversity RCS, IEEE Transactions on Antennas and Wireless Prop. Letters, 12: 937-940.

Du, L. Liu, H. Bao Z. and Xing M. 2005. Radar HRRP Target Recognition Based on Higher Order Spectra, IEEE Transactions on Signal Processing Vol.53, No.7, 2359-2368.

Duda, R.O. Hart, P.E. and Stork, D.G. 2001. Pattern Classification. John Wiley & Sons Inc. 738, USA.

Frank, E. Hall, M.A. and Witten, I. H. 2016. Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 4th edition, 654.

Hsu, C. and Lin, C. 2002. A comparison of methods for multi-class support vector machines. IEEE Transactions on Neural Networks. 13; 1045-1052.

Guo, G. Zhang, W. and Wenxian Y. 1989. An Intelligence recognition method of ship targets, Aerospace and Electronics Conference.

Liu, X. Wang, B. Xu, X. Liang, J. Ren, J. and Wei, C. 2011a. Modified Nearest Neighbor Fuzzy Classification Algorithm for Ship Target Recognition, 6th IEEE Conf. Industrial Electronics and Applications, 2254-2258.

Liu, X. Wang, B. Xu, X. Liang, J. Ren, J. and Wei, C. 2011b. Multi-Stage Classification Algorithm of Ship Target Based on HRRP Multi-Features. 6th IEEE Conference on Industrial Electronics and Applications, 2249-2253

Lu, J. Xi, Z. Yuan, X. Yu, G. and Zhang M. 2011. Ship Target Recognition Using High Resolution Range Profiles based on FMT and SVM. IEEE Transactions, 1299-1302

Inggs, M.R and Robinson, A.D. 1999. Ship Target Recognition Using Low Resolution Radar and Neural Networks, IEEE Transactions on Aerospace and Electronic Systems, Vol. 35, No.2, 386-392.

Jain, A.K. Duin, R.P.W. and Mao J. 2000. Statistical Pattern Recognition: A Review, IEEE Transaction on Pattern Analysis and Machine Intelligence, 22(1):4-37.

Junping, S. and Yi, D. 2006. Research on Ship Target Auto-Recognition Technique for Low Resolution Radar, CIE International Conference on Radar, Shangai, China.

Mahafza, B. R. 2000. Radar Systems Analysis and Design Using MATLAB. Chapman & Hall/CRC, 531, USA.

Manning, C.D. Raghavan, P. and Schütze, H. 2009. An Introduction to Information Retrieval. Cambridge University Press, 581, Cambridge, England.

Mayoraz, E. and Alpaydın, E. 1999. Support Vector Machines for Multi-Class Classification. IWANN'99, Alicante, Spain.

Mengdao, X. and Zheng, B. 2001. The properties of range profile of aircraft, CIE International Conference on Radar Proceedings, Beijing, China.

Segeran, T. 2007. Programming Collective Intelligence, O'Reilly Media, 360, USA.

Shafer G. 1976. A Theory of Statistical Evidence. Foundation of Probability Theory, Statistical Inference, and Statistical Theories of Science, Vol. 2, 365-436.

Shafer G. 1992. The Dempster-Shafer Theory, In: Encyclopedia of Artficial Intelligence. Stuart C. Shapiro (eds), Wiley, 330-331.

Skolnik, M. I. 1981. Introduction to Radar Systems. McGraw-Hill, 2nd edition, 570, USA.

Skolnik, M. I. 1990. Radar Handbook. McGraw-Hill, 846, USA.

Slomka, S. Gibbins, D. Gray, D. and Haywood, B. 1999. Features for high resolution radar range profile based ship Classification, *5th International Symposium on Signal Processing and its Applications,* ISSPA, 329-332, Australia.

Swerling, P. 1954. Probability of Detection for Fluctuating Targets, Rand Research Memorandum, RM-1217.

Tait, P. 2005. Introduction to Radar Target Recognition. IET Radar, Sonar and Navigation Series 18, The Institution of Engineering and Technology, 428, London, United Kingdom.

Theodoridis, S. and Koutroumbas K. 2009. Pattern Recognition, Academic Press, 967, USA.

Wang, Y.F. and Casasent, D. 2009. A Support Vector Hierarchical Method for Multi-class Classification and Rejection. IEEE Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA, June 14-19.

Zwicke, P. E. and Kiss, I. 1983. A New Implementation of the Mellin Transform and its Application to Radar Classification of Ships, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No.2, 191-199

ÖZGEÇMİŞ

Adı Soyadı : Özlem ERGÜN
Doğum Yeri : ANKARA
Doğum Tarihi : 22.02.1992
Medeni Hali : Bekar
Yabancı Dili : İngilizce

Eğitim Durumu (Kurum ve Yıl)

Lise : Hacı Ömer Tarman Anadolu Lisesi (2010)
Lisans : Ankara Üniversitesi Mühendislik Fakültesi Elektrik-Elektronik Mühendisliği (2015)
Yüksek Lisans : Ankara Üniversitesi Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı (Eylül 2015 – Ekim 2018)

Çalıştığı Kurum/Kurumlar ve Yıl

STM Savunma Teknolojileri ve Mühendislik A.Ş., Yazılım Mühendisi (2017- …)

Ulusal Kongre Sunum

Ergün Ö. Sarıkaya T. B. Soysal G. and Efe M. 2018. Ship Classification with High Range Resolution Radar. 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey.