

KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

İMLEÇ HAREKETLERİNDEN KULLANICI CİNSİYETİNİN
MAKİNE ÖĞRENMESİ TABANLI TESPİTİ

EYYÜP SAYILIR

KOCAELİ 2020

KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI

YÜKSEK LİSANS TEZİ

İMLEÇ HAREKETLERİNDEN KULLANICI CİNSİYETİNİN
MAKİNE ÖĞRENMESİ TABANLI TESPİTİ

EYYÜP SAYILIR

Dr.Öğr.Üyesi Orhan AKBULUT
Danışman, Kocaeli Üniversitesi

Dr.Öğr.Üyesi Alev MUTLU
Jüri Üyesi, Kocaeli Üniversitesi

Dr.Öğr.Üyesi Burhan BARAKLI
Jüri Üyesi, Sakarya Üniversitesi

Tezin Savunulduğu Tarih: 27.08.2020

ÖNSÖZ VE TEŞEKKÜR

Bu tez çalışmasında, imleç hareketlerini kullanarak makine öğrenmesi tabanlı cinsiyet tespiti yüksek sınıflandırma başarımları ile gerçekleştirilmiştir. Bu çalışmanın özellikle cinsiyet tespiti tabanlı güvenlik uygulamalarında yaygın uygulanabilir olmasını temenni etmekteyim.

Tez çalışmam boyunca bilgi, birikim ve deneyimleri ile bana yol gösteren ve destek olan çok değerli danışman hocam Dr. Öğr. Üyesi Orhan AKBULUT'a teşekkürü bir borç bilirim.

Çalışmalarım boyunca bana her zaman güvenen ve maddi manevi desteğini eksik etmeyen aileme sonsuz teşekkür ederim.

Veri kümesini çalışmamızda kullanılmak üzere paylaşan Texas A&M Üniversitesinde öğretim üyesi Dr. Takashi Yamauchi'ye teşekkürlerimizi sunarız.

Ağustos - 2020

Eyyüp SAYILIR

İÇİNDEKİLER

ÖNSÖZ VE TEŞEKKÜR	i
İÇİNDEKİLER	ii
ŞEKİLLER DİZİNİ.....	iii
TABLOLAR DİZİNİ	iv
SİMGELER VE KISALTMALAR DİZİNİ	vi
ÖZET.....	vii
ABSTRACT	viii
GİRİŞ	1
1. LİTERATÜRDEKİ ÇALIŞMALAR.....	4
2. TEMEL KAVRAMLAR.....	14
2.1. İmleç Hareketleri	14
2.2. Veri Kümesi	18
3. MAKİNE ÖĞRENMESİ	24
3.1. Denetimli Öğrenme	26
3.2. Denetimsiz Öğrenme	27
3.3. Yarı Denetimli Öğrenme	27
3.4. Takviyeli Öğrenme	28
3.5. Destek Vektör Makinesi	28
4. İMLEÇ HAREKETLERİNİN CİNSİYETE GÖRE SINIFLANDIRILMASI	31
4.1. Yöntem	31
4.2. Eğitim Aşaması	33
4.3. Test Aşaması	33
4.4. Deneysel Sonuçlar	33
5. SONUÇLAR VE ÖNERİLER	46
KAYNAKLAR	47
KİŞİSEL YAYINLAR VE ESERLER	53
ÖZGEÇMİŞ	54

ŞEKİLLER DİZİNİ

Şekil 2.1. İmleç hareket verileri toplama program arayüzü	18
Şekil 2.2. Kullanıcılara gösterilen şekil grupları.....	19
Şekil 2.3. Farklı şekil grupları.....	19
Şekil 2.4. Şekil gruplarındaki şekillerin farklı dizilişleri	20
Şekil 2.5. İmleç hareketinde ideal yoldan ayrılma	21
Şekil 2.6. İmleç hareketinde yön değiştirme.....	21
Şekil 2.7. İmleç hareketi yön sınıfları	22
Şekil 2.8. Eşit zaman dilimindeki özniteliklerin elde edilmesi.....	22



TABLolar DİZİNİ

Tablo 2.1. Zamana baęlı hareket metrikleri	15
Tablo 2.2. Mekânsal hareket ölçüleri	16
Tablo 2.3. Hedefleme kontrol ölçüleri	17
Tablo 4.1. Hata matrisi	32
Tablo 4.2. 16 öznitelik eğitim veri kümesi Lineer çekirdek fonksiyonu hata matrisi	34
Tablo 4.3. 16 öznitelik eğitim veri kümesi Polinomiyal çekirdek fonksiyonu hata matrisi	35
Tablo 4.4. 16 öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu hata matrisi	35
Tablo 4.5. 16 öznitelik test veri kümesi Lineer çekirdek fonksiyonu hata matrisi	36
Tablo 4.6. 16 öznitelik test veri kümesi Polinomiyal çekirdek fonksiyonu hata matrisi	36
Tablo 4.7. 16 öznitelik test veri kümesi Gaussian çekirdek fonksiyonu hata matrisi	36
Tablo 4.8. 8 Hedefleme öznitelik eğitim veri kümesi Lineer çekirdek fonksiyonu hata matrisi	37
Tablo 4.9. 8 Hedefleme öznitelik eğitim veri kümesi Polinomiyal çekirdek fonksiyonu hata matrisi	37
Tablo 4.10. 8 Hedefleme öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu hata matrisi	38
Tablo 4.11. 8 Hedefleme öznitelik test veri kümesi Lineer çekirdek fonksiyonu hata matrisi	38
Tablo 4.12. 8 Hedefleme öznitelik test veri kümesi Polinomiyal çekirdek fonksiyonu hata matrisi	39
Tablo 4.13. 8 Hedefleme öznitelik test veri kümesi Gaussian çekirdek fonksiyonu hata matrisi	39
Tablo 4.14. 8 Kontrol öznitelik eğitim veri kümesi Lineer çekirdek fonksiyonu hata matrisi	40
Tablo 4.15. 8 Kontrol öznitelik eğitim veri kümesi Polinomiyal çekirdek fonksiyonu hata matrisi	40
Tablo 4.16. 8 Kontrol öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu hata matrisi	40
Tablo 4.17. 8 Kontrol öznitelik test veri kümesi Lineer çekirdek fonksiyonu hata matrisi	41
Tablo 4.18. 8 Kontrol öznitelik test veri kümesi Polinomiyal çekirdek fonksiyonu hata matrisi	41
Tablo 4.19. 8 Kontrol öznitelik test veri kümesi Gaussian çekirdek fonksiyonu hata matrisi	42
Tablo 4.20. 16 öznitelik veri kümesi sınıflandırma başarımları ve F-Skor değerleri	43

Tablo 4.21. 8 Hedefleme öznitelik sınıflandırma başarıml ve F-Skor deęerleri	44
Tablo 4.22. 8 Kontrol öznitelik sınıflandırma başarıml ve F-Skor deęerleri	45



SİMGELER VE KISALTMALAR DİZİNİ

Kısaltmalar

ADHD	:Attention-Deficit/Hyperactivity Disorder (Dikkat eksikliği / Hiperaktivite Bozukluğu)
EEG	:Elektroensefalografi
ML-SST	:Machine Learning- Stop-Signal Task (Makine Öğrenimi- Duruş Sinyal Görevi)
PIN	:Personal Identification Number (Kişisel Kimlik Numarası)
RBF	:Radial Basis Function (Radyal Tabanlı Fonksiyon)
SST	:Stop-Signal Task (Duruş Sinyal Görevi)
SSRT	:Stop-Signal Reaction Time (Duruş Sinyal Tepki Zamanı)
SVM	:Support Vector Machine (Destek Vektör Makineleri)
NLP	:Natural Language Processing (Doğal Dil İşleme)
OCR	:Optical Character Recognition (Optik Karakter Tanıma)

İMLEÇ HAREKETLERİNDEN KULLANICI CİNSİYETİNİN MAKİNE ÖĞRENMESİ TABANLI TESPİTİ

ÖZET

Kullanıcıların bilgisayar ile etkileşimleri genellikle bilgisayar çevre birimleriyle (klavye, fare, dokunmatik ekran vb.) gerçekleşmektedir. Bu etkileşimden yola çıkarak kullanıcıların cinsiyet bilgisi tahminine yönelik çıkarımlar elde edilebilmektedir. Daha spesifik bir şekilde, çevresel aygıtlar üzerinden cinsiyet tespiti sayesinde reklam, pazarlama gibi alanlarda cinsiyete göre öneri sistemleri oluşturulabilmektedir. Ayrıca, web sitelerinde kimlik doğrulama ve siber suçların tespitinde de bu aygıtlar üzerinden cinsiyet tespiti tabanlı güvenlik uygulamasını gerçekleştirmek mümkün olmaktadır.

Bu tez çalışmasında, imleç hareketlerini kullanarak makine öğrenmesi tabanlı cinsiyet tespiti yaklaşımı önerilmiştir. İmleç hareketlerinden çıkartılan öznitelikler kullanarak Destek Vektör Makinesi tabanlı ikili bir sınıflandırılma işlemi gerçekleştirilmiştir. Önerilen yöntemin başarımı, farklı çekirdek fonksiyonlarıyla da kıyaslanmıştır. Sınıflandırma sonuçları, imleç hareketleri üzerinden makine öğrenmesi tabanlı cinsiyet tespiti yapılabileceğini açık bir şekilde ortaya koymaktadır.

Anahtar Kelimeler: Cinsiyet Tespiti, İmleç Hareketleri, Makine Öğrenmesi, Sınıflandırma.

MACHINE LEARNING BASED GENDER DETECTION FROM CURSOR MOVEMENTS

ABSTRACT

In Human-Computer interactions, peripheral devices such as keyboard, mouse, and touch screen are generally used. It can be possible to detect the gender of the users in accordance with these interactions. More specifically, thanks to gender detection through environmental devices, it can be devised gender-based suggestion systems in areas such as advertising and marketing. In addition, it is also possible to obtain gender detection based security with these devices for authentication and cybercrime on websites.

In this thesis, machine learning based gender detection approach which uses cursor movements is proposed. Using distinctive features obtained from cursor movements, Support Vector Machine based binary classification process is carried out. The performance of the proposed method is also compared to different kernel functions. Classification results clearly reveal that machine learning based gender detection using cursor movements can be accomplished.

Keywords: Gender Detection, Cursor Movements, Machine Learning, Classification.

GİRİŞ

Bilgisayarların birçok alanda kullanımını son yıllarda önemli bir artış göstermiştir. Bu durum bilgisayarlar ve kullanıcılar arasında bir etkileşim zorunluluğu oluşturmuştur. Bu etkileşim ise çeşitli aygıtlarla sağlanmaktadır (klavye, fare, dokunmatik ekran vb.). Her gün bilişim sistemlerinde kullanıcılar mailleri kontrol etme, çevrim içi sosyal ağlarda paylaşım ve yorum yapma, bilimsel araştırma yapma, anketlere katılma, online kitap arama, satın alma vb. birçok işlem yapmaktadırlar [1]. Bu işlemlerdeki hareket, işaretleme, seçme ve sürüklenme gibi bilgisayar imleci ve fare aktiviteleri bu etkileşimlerin temel parçalarıdır [1]. Ayrıca kullanıcıların cinsiyetini bilmek birçok alanda, örneğin çevrimiçi pazarlamada, öneri sistemlerinde veya siber suçluları tespit etmede faydalı olabilmektedir [2]. Otomatik cinsiyet tanıma olasılığı yıllardır büyük ilgi görmüştür. Bu konuda kapsamlı birçok çalışmalar yapılmış olup, insan-bilgisayar etkileşimi, gözetim sistemleri, içeriğe dayalı indeksleme ve arama, biyometri, demografik veri toplama, hedefli reklamcılık uygulama alanlarında gerçekleştirilmiştir [3]. Yine de kadın ve erkeklerin bilgisayar imleçlerini hareket ettirme biçimini nasıl değiştirdikleri bilinmemektedir [1].

İmleç hareketleri çeşitli alanlarda araştırma konusu olmuştur. Bunlardan bazıları; İnsan-bilgisayar etkileşimi, bilgisayar kullanıcılarının duygularının tespit edilmesi, bilişim güvenlik sistemleri, kullanıcı-modelleme olarak sınıflara ayrılabilir. Son zamanlarda popüler uygulamalardan olan duygu tanıma hem tuş vuruşunu hem de fare hareket özelliklerini içeren sistemlerden birçok makalede bahsedilmiştir [4]. Bununla ilgili yapılan çalışmalarda duygusal durumların veya bu durumların tam olarak belirlenemese bile en azından olumlu mu yoksa olumsuz mu olduğuna karar verilmesinin, bir dereceye kadar bir tanesini veya daha fazlasını tanımanın mümkün olduğu ortaya çıkmıştır [3]. Araştırmacılar, bilgisayar kullanıcısının duyguları hakkında bilgi edinmek için imleç hareketlerinden yararlanmak istemişlerdir. Bir kullanıcının neler hissettiğini belirlemek için birçok yöntem geliştirilmiş ve araştırılmıştır [4]. Bu yöntemler farklı veri kaynaklarına dayanmaktadır. Örneğin; görsel (yüz ifadesi, el hareketleri), ses (konuşma, ses), metinsel (anlamsal), fizyolojik

(kalp atış hızı, sıcaklık, cilt iletkenliği vb.), giriş cihazları (klavye, fare, dokunmatik ekran) ve bunların bir kombinasyonu olacak şekilde kullanılarak daha doğru bir tespit yapılmaya çalışılmıştır [4]. Fakat bu yöntemler pahalı ve zahmetli olabilmektedir [5]. Bundan dolayı, bu yöntemler günlük uygulamalar için her zaman pratik değildir [1]. Son 15 yılda araştırmacılar, imleç hareketlerinin duygu tanımda etkinliklerini incelemişlerdir [5]. Yapılan bir araştırmada, film tabanlı bir duygu çıkarma tekniği kullanmış ve bir çevrimiçi alışveriş görevinde uyarılma ve değerliliğin imleç hareketi üzerindeki etkisini araştırmışlardır [5]. Yapılan diğer bir çalışmada, akıllı bir özel ders sistemi kullanarak cebir öğrenmesi sırasında EEG sinyallerini ve fare aktivitelerini (fare tıklamaları sayısı, kat edilen mesafe, tıklama süresi) değerlendirmiştir [5]. Diğer bir duygu tanıma çalışmasında, duyguların tespit edilmesi için resimler kullanarak kat edilen mesafe, hız ve yön değişikliği gibi ölçülebilen fare imleci hareket özelliklerini kullandılar [5].

İmleç hareketleri, psikolojik bilimin çok çeşitli alanlarında kullanılan bilişsel süreçleri karakterize etmek için önemli bir yöntem haline gelmiştir [6]. İmleç hareketlerinin kullanıldığı bilişsel alanlarda kullanıcıların duygu durumlarının tespit edilmesine yönelik yapılan çalışmalardır. Bu çalışmalarda; kullanıcıların duygusal durumlarının tespiti klavye, fare gibi bilgisayar temel giriş aygıtları kullanılarak yapılabilmektedir [4]. Duygusal durum araştırmalarında kullanıcıların, yüz ifadeleri, sesleri, hareketleri gibi kişisel özelliklerinin fizyolojik sinyaller kullanarak gelişmiş duygu tanıma sistemlerini gerçekleştirmektedir [5]. Bu sistemlerden verileri elde etmede kullanılan yöntemler genellikle pratik değildir. Bundan dolayı imleç hareketlerinden kullanıcıların duyguları hakkında bilgi sağlanması pratiklik sağlamaktadır. [7]'de; kullanıcıların duygu değerlendirmesi için imleç hareketlerini izleme yöntemini kullanmış, duygusal deneyimin imleç hareketlerinin maksimum hızı olarak belirtilen zamansal ve yörüngenin sapması olarak belirtilen uzamsal özelliklerini etkilediğini gösterilmiştir.

Cinsiyet, kişisel kimliğin temel özelliklerinden biridir. Doğal cinsiyet farklılıklarından faydalanarak, fare biyometrisine dayalı cinsiyet sınıflandırması kötü amaçları önlemek için bir yaklaşım önerebilir [8]. Biyometrik veriler, cinsiyet veya yaş dahil olmak üzere insanların fizyolojik özelliklerinden etkilenir. Fare hareketi verilerinden elde edilecek biyometrik veriler cinsiyet ve yaşın tespit edilebileceği bir sistemin web

kullanıcılarının profillemesi için oldukça faydalı olabileceği görülmektedir [9]. İmleç hızlanması, imlecin düz yoldaki hareketi veya duraklatma gibi fare hareket dinamiklerinin sınıflayıcı sistemler tarafından sınıflandırılarak cinsiyet tahmini gerçekleştirilebilir [9]. Böylece cinsiyet tahmini, özellikle bir web sitesi kullanıcı kitlesinin belirlenmesinde istatistiksel analizi için bir yöntem olabileceği sonuçlarını göstermektedir [9].

Bu tez kapsamında, imleç hareketleri ile kullanıcıların cinsiyetinin makine öğrenmesi yöntemlerini kullanarak tespit edilmesi hedeflenmektedir. Bu çalışmada doğru cinsiyet tanımlanmasının sağlanması için fare hareketlerinin davranışsal özelliklerinin kullanılmasına odaklanan bir çalışma olacaktır. Hedeflenen fare hareketlerinden türetilmiş hareket profilleri, klavyeden elde edilen verilerden daha zengin bir özellik kümesinin çıkarılmasını sağlar. Bu özellikler, klavye özelliklerinde belirgin olmayan kullanıcının özelliklerini yakalayarak sınıflandırmanın doğruluğunu artırmak için kullanılabilir.

1. LİTERATÜRDEKİ ÇALIŞMALAR

İmleç hareketleri birçok alanda araştırma konusu olmuştur. Literatürde, kullanıcı duygu tanıma, psikolojik bilişsel süreçlerin davranışsal analizi, kullanıcı çevrim içi tercihlerinin belirlenmesi, bilişim sistemlerinde bilgi güvenliği için kullanıcı kimlik doğrulama sistemleri, çevrim içi dolandırıcılık ve sahtekarlık tespiti gibi çalışmalar bulunmaktadır.

İnsan duyguların öğrenme sürecini etkilediğine dair güçlü kanıtlar bulunmaktadır. Bundan yola çıkarak, kullanıcıların fare hareketleri ve tuş vuruş dinamiklerinden elde edilecek verilerin veri madenciliği teknikleriyle modellenerek kullanıcı duygularının öğrenmeye etkisi fare veya klavye ile nasıl etkileşimde buldukları konusunda etkileşim temelli bir model oluşturma çalışması yapılmıştır [10].

Kullanıcı davranışsal analiz yöntemlerinde de imleç hareketleri kullanılmaktadır. Yapılan bir çalışmada akıllı özel ders sistemlerinde öğrencilerin eğitim sırasındaki görev dışı davranışlarını otomatik olarak algılayabilen bir makine öğrenme modeli önermiştir. Bu model, günlük olarak sistemdeki eylemlerin kaydedildiği dosyadaki verileri kullanır. Model, zaman özelliği, performans özelliği ve fare hareketleri özelliğini içeren verilerden oluşturulmuştur. Farklı öğrencilerin farklı davranış türleri olduğundan modelin kişiselleştirilmiş versiyonu oluşturulup, bu modellerin karşılaştırılması ile davranışsal farklılık tespit edilmesine çalışılmıştır [11].

Kullanıcı-modelleme, insan-bilgisayar etkileşimi alanı dahilinde, kullanıcı özelliklerinin bir kullanıcı modeli biçiminde tanımlanmasına odaklanmaktadır. Bir kullanıcı modeli, özniteliklerle veya daha karmaşık yapılarla açıklanan bireysel kullanıcı özellikleri kümesi olarak tanımlanabilir. Model oluşturma'nın temel amacı, belirli bir kullanıcı profili için sistem işlevselliğini, görünümünü veya davranışını uyarlayabilmek ve özelleştirebilmektir. Kullanıcı modelleri, uygulama alanlarına ve sisteme özgü özelliklere bağlı olarak farklı özellikler kullanır [3]. Yapılan çalışmalarda, çevre birimleri kullanımının davranış kalıplarını temel alan otomatik erkek ve kadın tanıma ile ilgilidir. [3]'de fare ve klavyenin kullanımında gözlemlenen

davranış kalıpları daha önce kişi tanımlamasında [12-14] ve duygu tanımda [15-17] kullanılmıştır. Kullanıcı profili madenciliği üzerine yapılan çeşitli araştırmalar, reklam tıklama davranışı, e-öğrenmede günlük dosyası analizi ve çevrimiçi sorgu incelemeleri gibi gerçek bilgilere odaklanmıştır [18].

Web siteleri günümüzde daha dinamik bir hale gelmektedir. Böylece web sitelerinde yer alan veriler artmakta ve artık kullanıcılar benzersiz farklı bilgiler aramaktadır. Böylece her kullanıcıya doğru içeriği sunma zorluğu doğurmaktadır. Bu problem, belirli fare hareketler veya kullanıcının fare seçimlerine dayanarak oluşturulacak içerik ile dinamik web sitelerinin kendilerini şekillendirerek çözülebilir. Fare hareket verilerine dayanarak, bir veya daha fazla makine öğrenme modeli kullanılarak, her kullanıcı için daha fazla içerik tahmin edilebilir ve kişiselleştirilebilir [19]. Diğer taraftan çevrim içi kullanıcı tercihlerinin belirlenmesi, kullanıcıların bir konu hakkında öneri veya görüşleri ile ilgili tespitler için web anketleri oldukça pratik ve kolay bir yöntem olarak tercih edilmektedir. Bu yüzden, web anketleri kısmen standart olan ve genellikle tercih edilen bir anket yönetim biçimi haline gelmiştir. Ankete katılan kullanıcının gerçek bir kullanıcı olduğunun belirlenmesi anket güvenilirliği için önem arz etmektedir. Böylece bireysel katılımcıları tanımlamak için fare hareketlerinin özellikleri kullanılabilir. Hareket edilen toplam mesafeyi, imlecin yörüngesini, ilgi, belirsizlik ve yanıt verilirken karşılaşılan zorlukları ölçmek için fare hareket özelliklerinden belirli hareket kalıplarını kullanmışlardır. Bu yöntem ile ayrıca web anketlerindeki zorluk göstergelerini tespit etmek ve ölçmek için otomatik prosedürler geliştirmekte kullanılabilir [20].

Elektronik ticaret günümüzde en yaygın alışveriş tercihleri arasında yer almaktadır. Elektronik ticaret sitelerindeki çevrimiçi kullanıcı tercih ettiği ürün miktarını tespit edebilmek için, kullanıcıların çevrimiçi davranışlarına ilişkin bilgi gerekir [21]. Elektronik ticaret işletmelerini daha başarılı hale getirmek için uygun müşteri analizi yöntemi satış stratejilerini oluşturmak ve uygulamak için çok önemlidir. Bu tür çevrimiçi satış yöntemlerinde, müşterilerin davranışlarını etkili bir şekilde analiz etmek için çeşitli veri analizi süreçleriyle analiz edilir. Müşterilerin davranışları fare hareket kalıpları ile çok iyi bir şekilde analiz edilebilir. Böylece fare hareket modeline dayalı oluşturulacak bir müşteri davranış analiz modeli elektronik ticaret şirketlerinin satış, pazarlama stratejileri için önemli bir yol gösterici olabilir [22]. Diğer yandan

müşteri yaşı ve cinsiyeti hem satış hem de pazarlama için çok önemli parametrelerdir. Her ikisinin de satın alma alışkanlıklarında çok önemli roller oynadığı bilinmektedir. Çevrimiçi alışveriş yapan bir müşterinin fare hareketlerini analiz ederek cinsiyetini tahmin etmek için bir model oluşturulabilir. Bu modelin müşteri demografik verileri, müşteri ziyaretleri, fare hareketleri, çevrimiçi alışveriş sepeti, harici veriler ve zaman boyutu gibi parametrelerinden kişiye özel alış-veriş öneri sistemleri, kişisel kullanım tercihleri gibi kişiselleştirilmiş sistemler oluşturulabilir [22]. Diğer yapılan bir çalışmada, müşterilerin çevrimiçi davranışlarını analiz etmenin yeni bir yolu olarak fareyle üzerine gelme hareketlerini izlemektir. Fareyle üzerine gelme işlemlerinin, müşterilerin çevrimiçi satış zamanında satın alma olasılığının tahminine ne ölçüde yardımcı olduğunu tespit etmeye yönelik bir çalışmada, müşteri satın alma olasılığını tahmin etmeye büyük katkıda bulunduğu sonucuna varmıştır [21].

İmleç hareketleri ve tuş vuruş analizi yeni bir çalışma alanı değildir. Bilgisayar sistemlerinin güvenliğini sağlamak için kullanılan yaygın biyometrik teknikler tuş vuruşu dinamikleri ve fare hareketlerinin analizidir. Biyometri, kesin bir kimlik oluşturmak veya doğrulamak için insan davranışsal veya fizyolojik özelliklerini tanımlayan bir dizi faktörün otomatik kullanımı olarak tanımlanır. Biyometrik özellikler fizyolojik (yüz, avuç içi, parmak izi, iris) veya davranışsal (ses, el yazısı imzalar, tuş vuruşları dinamiği, fare hareketleri) olabilir [4]. Biyometrik sistemlerin çoğunda en önemli sınırlama, biyometrik veri elde etmek için özel ekipman ve donanım cihazlarına ihtiyaç duyulmasıdır [23]. Fizyolojik biyometrik özellikler için bunları kaydetmek için özel donanımın gerekli olması önemli bir dezavantaj olup davranışsal biyometrik sistemlerde kullanıcı davranışlarını kaydetmek için özel bir donanım gerektirmemesi ayrı bir avantaja sahiptir. Fare dinamiğine dayalı biyometrik sistemlerde tuş vuruşlarına benzer şekilde, fare dinamikleri veri toplama için özel bir donanım cihazı gerektirmez. Mevcut tuş vuruşu dinamiklerinin aksine, fare dinamikleri pasif olarak toplanabilir ve oturum boyunca doğrulanabilir. Fare hareketlerini biyometrik olarak kullanmak, tuş vuruşlarından daha az yaygındır. Fare hızı, hızlanma, yön, tıklama sayısı vb. fare özelliklerini analiz ederek, fare dinamiği biyometrisi ile izinsiz giriş tespiti için erişim kontrolüne ek olan bir güvenlik tedbiri şeklinde uygulanabilir. [4,23]. Davranışsal biyometrik sistemler ile çalışmalar tuş vuruş analizi ile başlamış olup daha sonra bu alana fare-imleç hareketleri eklenmiştir.

Tuş vuruşu sistemin bazı eksiklikleri ile karşılaşılmıştır. Bunlardan bazıları farklı tip ve tür klavyelerin olması, klavye tabanlı kimlik doğrulama sistemlerinin yeni grafiksel ara yüzlerine uygun olmaması olarak gösterilebilir.

Bilişim sistemlerde kullanıcıların bilgi güvenliği hayati önem taşımaktadır ve daha güçlü, daha güvenli kimlik doğrulama sistemlerine gün geçtikçe ihtiyaç duyulmaktadır. Bu konudaki en güvenli yol kullanıcının gerçek bir kişi olduğunun belirlenmesinde kullanılan kimlik doğrulama sistemidir. Kimlik doğrulama, kullanıcının parola veya kişisel kimlik numarasına (PIN) göre benzersiz bir şekilde tanımlanmasını sağlamak için kullanılan güvenli bir yöntem olarak tanımlanabilir [24]. Fakat doğrulama sistemine ait parola veya PIN'in unutulabilme, kaybolma veya çalınma gibi ciddi riskleri bulunmaktadır. Bunu önlemek amacıyla bilişim sistemleri için biyometriye dayanan güvenlik sistemleri geliştirilmiştir. Biyometri; bir bireyi fizyolojik veya davranışsal özelliklere dayanarak doğrulamak için kullanılır [25]. Biyometrik kimlik doğrulama, bir kullanıcıyı doğası gereği benzersiz özelliklerine (kim olduğunuza) göre doğrular [26]. Fizyolojik (fiziksel) biyometri sistemleri iris tanıma, parmak izi tanıma, avuç içi izi tanıma gibi sistemlerdir. Bu sistemleri uygulayacak donanım ve aygıtların maliyetli olmasından dolayı kullanımı ve uygulanması daha az tercih edilmektedir. Böylece daha kolay uygulanabilirliği ve daha az maliyetli olduğundan imleç hareketlerine dayalı davranışsal biyometri kullanılması fikri oluşmuş ve biyometrik tabanlı kimlik doğrulama sistemleri giderek daha fazla kullanıcı adı-parola sistemlerine üstünlük sağlamıştır. Fizyolojik biyometriye ek olarak, davranışsal biyometriklerin bir kullanıcının kimliğini doğrulamak için çok yararlı olduğu kanıtlanmıştır. Fare dinamiği, benzersiz fare hareketi kalıpları ile böyle bir davranışsal biyometriktir [26]. İmleç hareketlerine dayalı biyometri, kullanıcının bir grafik ara yüz ile etkileşimde olduğu sürece imleç hareket özelliklerini yakalanıp, ayıklayarak analiz eden davranışsal bir biyometrik sistemdir [27]. Böylece fare hareket özelliklerini tanımlamak ve uygulamak kimlik doğrulama sisteminin bir parçası olarak kullanılabilir. [28]'de, farklı kullanıcılardan elde edilen fare özellikleri kimlik doğrulama sistemine girdi olarak kullanılarak güçlü bir biyometrik sistem oluşturmuştur. Diğer yapılan çalışmalarda kimlik doğrulama amacıyla davranışsal bir biyometrik olarak imleç hareketleri kullanma fikri son yıllarda artan bir ilgi görülmüş ve literatürde birçok yaklaşım ve umut verici deneysel sonuçlar elde edildiği

görülmüştür [29]. İmleç hareketleri ile kullanıcı kimlik doğrulanması ilk oturum açma sırasında kullanılır. Kullanıcının imleç hareketlerine bağlı olarak koordinat eksenleri ve geçen süre yakalanmış, daha sonra çıkarılan özellikler çeşitli algoritmalarla sınıflandırılarak kullanıcının yetkili veya yetkisiz olduğu tespit edilmiştir [30]. Diğer bir kullanıcı kimlik doğrulama sistemi ise yeniden kimlik doğrulama sistemidir. Bu sistem, kullanıcının oturum açma işleminden sonra belirli bir süre işlem yapmamasından veya sistemde oluşan beklenmeyen bir durumdan dolayı kullanıcıyı sistem atması ve sistemi kilitlemesinden ötürü kullanıcının sistemi tekrar kullanabilmesi için yapması gerek tanımlama işlemidir. Yapılan yeniden kimlik doğrulama çalışmalarında imleç hareketleri ile farklı kullanıcı tespitinin yapılabileceği, kullanıcının fare hareketleri modellenerek, kullanıcının öğrenilmiş normal davranışlarından saptığında uyarı vererek farklı kullanıcı olduğu belirlenmiştir. Fakat yeniden kimlik doğrulama sistemi için sadece fare hareketlerine dayalı analiz sistemin yeterli olmadığını belirtmişlerdir [12]. Diğer taraftan kullanıcının açık oturumu boyunca kimlik doğrulamasını gerçekleştireceği çalışmalarda mevcuttur. [31]'deki çalışmada imleç hareketlerinin davranışsal karakteristiğine ve makine veya derin öğrenmeye dayalı sürekli kimlik doğrulamasını doğru ve verimli bir şekilde gerçekleştiren kullanıcı kimlik doğrulama yöntemi önermiştir. Sürekli kimlik doğrulaması gerçekleştirmek için imleç hareketleri, güvenlik ihlallerine karşı koruma olarak kullanılabilen davranışsal biyometrik bir örnektir [25]. Elde edilen işaretleme, tıklama ve sürükleyip bırakma gibi fare hareketleri makine öğrenimi sınıflandırma yöntemleriyle kimlik doğrulama gerçekleştirilebilir [25]. [32]'deki yapılan kimlik doğrulama çalışmasında, fare hareketlerinin yörüngedeki nokta sayısı, gecikme süresi, gecikme sayısı, eylem sayısı, yörünge uzunluğu, yatay konumdaki değişim sayısı ve dikey konumdaki değişim sayısı gibi özelliklerini kullanarak makine öğrenimi sınıflandırıcısı yardımıyla sınıflandırarak bir kullanıcı kimlik doğrulama sistemi önerilmiştir. [26]'da; kullanımı yeterince doğru ve verimli olan fare dinamiklerini kullanan bir kullanıcı doğrulama sistemi önermiş olup, sistemin temel özelliği, kullanıcı doğrulaması için fare hareketlerinin çok daha ayrıntılı (noktadan noktaya) açık tabanlı metriklerinin kullanılmasıdır. Bu yeni metrikler, kişiden kişiye göreceli olarak benzersiz ve bilgi işlem platformundan bağımsız olduğunu belirtmiştir.

Fare hareketleri çevrim içi uygulamalarda dolandırıcılık, sahtekarlık tespiti için kullanılmıştır. Yapılan çalışmada fare hareketleri, sürükle-bırak gibi fare özelliklerinde toplanan verileri ikili sınıflandırıcı kullanarak sahtekarlık tespiti için uygun özellikler olduğunu belirtmiştir. Yapay fare hareketleri, ters biyometrik model olarak adlandırılan gerçek kullanıcılardan toplanan rastgele örneklerden öğrenerek elde edilen örneklerden oluşturulmuştur [33]. [34]'de gerçek olmayan fare hareketlerini kullanan sahte kullanıcıları tespit edilmesiyle ilgidir. Burada, özellikle fare hareketlerine dayalı güvenlik sistemlerinde veya kimlik doğrulama işlemlerinde yapay üretilen fare hareketlerini kullanan sahte kullanıcılarının tespiti amaçlanmıştır.

Fare hareketlerine dayalı sistemler geliştikçe bu sistemler için tehdit olabilecek yapılar, fare verilerinin açığa çıkması nedeniyle sağlayamayan güvenlik açıkları ve güvenlik tehditleri vardır. Bu sorunu çözmek için, fare veri koruma teknolojisinin ortaya çıkmasına neden olmuştur. Bu teknoloji, herhangi bir zamanda çok sayıda rasgele fare konumu oluşturarak fare verilerini korur. Böylece saldırgan, kullanıcının fare hareketlerini ele geçirse bile koruma teknolojisi tarafından üretilen fare konumlarına erişeceğinden gerçek fare konumlarına ulaşamayacaktır. Bu nedenle, bu fare koruma teknolojisi neredeyse mevcut fare veri saldırı tekniklerine karşı savunabilir [35].

[36]'da ki çalışmada; fare hareket verilerini çalarak oluşturulan fare veri saldırısı, makine öğrenim ile sınıflandırarak tespit edilmeye çalışılmıştır. Bu tespit kullanıcıdan gelen fare koordinatları arasında görünen özellik analiz edilerek sağlanmıştır.

Bilişim sistemlerinde insanlar ve makineler arasında ayırım yapmak için kullanılan yöntemlerden biri de görüntüler ve metin içeren CAPTCHA kullanılmasıdır. Kullanıcın doğrulama için CAPTCHA ile etkileşimde bulunması gerekir, bu işlem doğrulama için gereken süreyi uzatmış olur. Bu doğrulama işlemi daha basit, kısa sürede gerçekleştirmek için fare hareketlerine dayalı imleç yörünge tanıma çalışması yapılmıştır [37].

Herhangi bir web sitesi, kullanıcıların kişiselleştirme amaçları için kullanıcıların ilgi alanları hakkında bilgi edinmek için kullanıcıların web sitesindeki fare etkileşimlerini kaydedebilir. Fare dinamikleri kullanıcıların davranış modelleri ile ilişkilendirilebileceğinden, kötü niyetli web siteleri bu tarz modelleri gizlice elde

etmeye ve böylece kullanıcıların gizliliğini tehlikeye atmaya ve onları hedefli saldırılara eğilimli hale getirmeye çalışabilir. Bununla ilgili yapılan bir çalışmada, kullanıcıların bilişsel yeteneklerine ve davranış modellerini yakalayarak bu fare hareket özelliklerini temel alıp, bilinmeyen demografik özellikleri tahmin eden makine öğrenme yaklaşımı ile kullanıcı tespitinde yardım eder [38].

Yukarıda bahsedilen alanlarda fare imleci hareket analizi, araştırmacıların farklı girdi cihazlarının performansını karşılaştırmaya başladıkları 1970'lerin sonlarında ortaya çıkmıştır [9,39-42]. Fare hareketi özellikleri, bir farenin hareket-yörünge şeklini ve fare düğmelerine nasıl basıldığı şeklinde belirlenebilir. Fare hareketleri genellikle fare hızı, hızlanma ve hareket şekilleriyle bağlantılı bir dizi parametre ile karakterize edilir. Örneğin, fare yolunun düz bir çizgiden nasıl saptığını, yönünün ne sıklıkta değiştiğini, kullanıcının tercih ettiği yönün ne olduğunu saptayabilirler. Tıklama özellikleri, bir düğmeye basma ve bırakma arasındaki veya çift tıklama durumunda iki basma arasındaki süreyi ölçebilir. Bir fareyi durdurmak ile hareket durumunda bir düğmeye basmak veya bırakmak ve sırasıyla olayları tıkladmak veya sürükleyip bırakmak arasındaki gecikmeler de dikkate değerdir. Bu özelliklerin çoğunun, kullanılan uygulamaya büyük ölçüde bağlı olduğu açıktır [3]. Bununla birlikte, bildiğimiz kadarıyla hiçbir çalışma, bilgisayar fareleri ve imleçlerin hareketlerinde cinsiyet farklılıkları üzerine sistematik bir araştırma yapmamıştır; ancak [43]'de yaşlanmanın fare hareketleri üzerindeki etkileri ile ilgili çalışma yapmıştır [1].

Klavye ve fare dinamik özelliklerin hangisinin cinsiyet, yaş gibi kullanıcı temel özelliklerinin doğru tespiti ile ilgili çalışmalar da yapılmıştır.

Bilgisayar kullanıcılarının cinsiyet tespiti için çeşitli çalışmalar yapılmıştır. Bu çalışmalar standart giriş aygıtı olan klavyenin tuş vuruş dinamikleri aracılığıyla yapılmıştır. Tuş vuruş dinamiklerinden türetilen verilerle cinsiyet tespiti yapılmıştır. Tuş vuruş dinamikleri kullanıcının yazma şekli olarak tanımlanabilir [44]. Daha sonraki çalışmalarda fare hareketlerinin de cinsiyet tespiti çalışmaları yapılmış, fare hareket dinamiklerinin, tuş vuruş dinamiklerinden fazla özelliklere sahip olmasından dolayı son yıllarda fare hareketlerinden kullanıcı cinsiyet, duygu, yaş gibi kişisel özellik tespiti çalışmaları yapılmıştır [44].

Kullanıcı özelliklerinden olan yaş, cinsiyet hakkında standart giriş donanımlarından olan klavye ve fare kullanımları, birçok bilgi ve ipucu içerir. Yapılan çalışmada kullanıcıdan toplanan fare ve klavye verileri çeşitli makine öğrenimi modelleri kullanarak, kullanıcı yaşı ve cinsiyeti hakkında bilgi verir [2]. Ayrıca imleç hareketlerinden kullanıcı profilleri olan cinsiyet, duygu ve duygusal durum gibi kullanıcı özelliklerinin ne ölçüde değerlendirilebileceği araştırılmış, imleç hareketlerinin dinamik kullanıcı profili için zengin bilgiler sağlanabildiği önermiştir [18].

Yapılan çalışmalarda; erkeklerin ve kadınların cinsiyetler arasındaki, biyomekanik ve algısal-motor kontrol farklılıkları nedeniyle doğal hedefleme hareketlerinde farklılık göstermesidir [16,45]. Bundan dolayı doğru cinsiyet tanımlaması için fare biyometri kullanımına odaklanarak fare hareketlerin özelliklerinin veri olarak kullanarak cinsiyet sınıflandırması ile kişisel tercihlerin belirlenmesinde kolaylık sağlayabilecektir [45]. [1]'de; imleç hareketlerinin cinsiyet göre değiştiği, bir seçim görevi uygulamasında imleç yörüngelerinin erkek ve kadın katılımcılar arasında kontrol ve hedefleme ile ilgili hareket özelliklerinin ne ölçüde farklı olduğunu araştırmış, sonuç olarak imleç hareketlerinde önemli cinsiyet farklılıkları göstermiştir.

Diğer yandan yapılan çalışmalarda imleç hareketlerinin davranışsal sağlık problemlerin tespitinde yardımcı olarak kullanıldığı belirtilmiştir. Bu sağlık problemlerinde biri olan dikkat eksikliği / hiperaktivite bozukluğu (ADHD) dünya çapında yaşam kalitesini etkileyen bir sağlık sorunudur. Genellikle özel anketler ve davranış testleri ile teşhis edilir ve incelenir. Davranış testlerinin duyarlılığını artırmak için [46]'de, makine öğrenimini ve fare imleci izlemeyi (ML-SST) entegre eden Duruş-Sinyal görevinin (SST) işlemi gerçekleştirilmiştir. Bu çalışma sonucunda fare imleci izleme ve ridge regresyon yönteminin kullanılmasının, anket verileri ve davranışsal önlemler arasında çok güçlü ve istikrarlı ilişkileri ürettiğini göstermektedir. Bu konudaki diğer bir çalışmada, dikkatsizlik / hiperaktivite bozukluğu (ADHD) semptomlarının dikkatsizlik ve davranışsal disinhibisyon durumlarının doğru tespiti, zamanında yardım ve tedavi sağlamak için uygulanan davranış testlerinin duyarlılığını artırmak için, imleç hareketlerinin izlenmesi gibi sistemlerin kullanıldığı belirtilmiştir. Ayrıca bu hastalıkla ilgili tespitlerde fare imleci hareketinde hızlanma ve seçime ulaşmadaki hızlanma gibi davranışların izlenmesi

kullanılmaktadır [47,48]. Diğer yapılan bir çalışmada; yukarıdan aşağıya dikkatin bilinçsiz semantik işlemeyi kolaylaştırdığını göstermektedir. Bilinçdışı semantik işlemede dikkatin rolünü açıklığa kavuşturmak için, bilgisayar faresinin yörüngelerinin analizinde faydalandıkları görülmüştür [49].

Fare hareketleri yaygın kullanılan alanlar dışında da tespit sistemi olarak kullanılır. Örneğin çevrimiçi sitelerin ara yüz elemanlarını düzenlemek, video oyunlarında oyun kullanıcılarını, tanımlamak, bilişsel işlemlerin sürekliliğini tespit gibi alanlarda da imleç hareketleri kullanılmaktadır. Bilişsel işleme sürekliliği çalışmasında sürekli işleme ve hareket arasındaki varsayılan etkileşimin ölçümün metodolojik kurulumundan etkilenip etkilenmeyeceği önemli bir sorudur. Sonuçlar, metodolojik düzenin sürekli işleme ve fare hareketi arasındaki etkileşimi önemli ölçüde etkilediğini göstermektedir [50]. Diğer taraftan çevrimiçi sitelerin tasarım arabirimler ve kontrol panelleri için en iyi düzen yapılandırmasını oluşturmak için çeşitli yöntemler geliştirilmiştir. Bu ara yüzler için geçerli olan düzenler elde etmek manuel olarak ayarlanması gereken bir tasarım durumudur. Yazılım tabanlı ara yüzler için optimum düzenler elde etmek için önerilen yöntem, Ara yüz kullanımı sırasında toplanan göz izleme ve fare izleme verileri, tasarım elemanları için en iyi yapılandırmayı elde etmek için kullanılabilir [51]. Fare hareketi analizindeki son çalışmalar, yeterli veri ile kullanıcıların yalnızca fare hareketleriyle benzersiz bir şekilde tanımlanabileceğini belirlemiştir. Buradan yola çıkarak yapılan bir çalışmada; video oyun oynaya oyuncuları tanımlamak için fare hareketlerini kullanmaya çalışmıştır. Çalışmada yöntem olarak oyunculara öncelikle temel görev verilip bunları yerine getirdikten sonra fare hareketleri ham verisi elde edilir. Daha sonra kullanıcılara iki popüler video oyununu oynatarak kullanıcıların doğru oyuncu olduğu fare hareketlerinden çıkarılan özelliklerin sınıflandırması makine öğrenimi teknikleriyle belirlemeye çalışılmıştır [51].

Diğer bir uygulama alanı ise yalan tespit teknolojilerinde davranış analizi ve uygulamaları için insan-bilgisayar etkileşimini kullanan makine öğrenmesinin uygulanmasına ilişkin araştırmalar bulunmaktadır [52].

Son yıllarda dokunma ve çoklu dokunma hareketleri, akıllı telefonlar, tabletler ve diğer mobil cihazlar gibi teknolojiyle etkileşime girmenin en yaygın yolu haline geliyor. En

son dokunmatik ekran giriř kapasiteleri, mevcut hareket verilerinin miktarını ve kalitesini önemli ölçüde artırarak psikolojiden biyometriye kadar çoklu disiplinlerde kullanımının araştırılmasına yol açmıştır. [53]'te tuř vuruřu ve fare kullanım biyometrisi gibi benzer yöntemlerle yapılan araştırma çalışmalarını takiben, mevcut çalışma yumuřak biyometri, özellikle kullanıcının cinsiyeti tahmini için kaydırma hareketi verilerinin kullanılmasını önermektedir. Çalışma sonucunda iki farklı yönden kaydırma hareketi verilerini kullanarak, hareketi verilerinden cinsiyet tahmini olasılığını dođruladıđında yumuřak biyometri özelliklerinin, dokunmatik ekranlı cihazlara dayalı kimlik dođrulama işlemlerinin geliştirilmesinde önemli bir rol oynayabileceđi belirtilmiştir



2. TEMEL KAVRAMLAR

Bu bölümde imleç hareketlerinin özellikleri ve tez kapsamında kullanılan veri kümesinin nasıl oluşturulduğu hakkında bilgi verilecektir.

2.1. İmleç Hareketleri

İmleç hareket analizi; imleç yörüngeleri ile duygu, cinsiyet, kişilik özellikleri gibi kişi profilleri arasındaki ilişkiyi saptama için yapılan çalışmalara odaklanmıştır [18]. Kullanıcıların imleç hareketleri analiz edilirken fare dinamik özelliklerinden yararlanılır. Fare dinamiği, kullanıcının grafik arabirimi ile etkileşime girerken giriş aygıt olarak kullanılan farenden elde edilen eylemlerin özellikleri olarak tanımlanır. Fare eylemleri; genel fare hareketleri, sürükle ve bırak, işaretle ve tıkla ve de sessizlik olarak sınıflandırılabilir. Bu eylemlerden, genel fare hareket eylemleri; standart, serbest şekilde genel fare hareketi, sürükle ve bırak; fare düğmesine basıldığı andan bırakılma anına kadar olan eylem, işaretle ve tıkla; fare hareketi ardından tek tıklama veya çift tıklama eylemi, sessizlik ise herhangi bir hareket olmadığı eylemler olarak açıklanabilir. Kullanıcıların kaydedilen fare dinamikleri analiz edilerek imleç hareket modeli oluşturulabilir. Model, hareket hızı, hareket yönü, eylem türü, kat edilen mesafe ve geçen süre gibi kategorilerden ve bunların kombinasyonları kullanılarak elde edilebilir. Bu analiz fare dinamik özelliklerinde gerçekleşen ortalama hız, ortalama hareket mesafesi, belirli hareket yönlerindeki ortalama hız ve belirli hareket yönlerindeki hareket mesafesi, bir konumdan başka bir konuma geçtiği hareket sonrasında sessizlik dönemleri, tek tıklama veya çift tıklama durumları hesaplanarak elde edilir [23].

Diğer yapılan çalışmada kullanıcının imleç hareketlerinde mesafe ve hız profilleri çıkarılmıştır. Bu profiller aşağıda tabloda belirtilen özelliklerin kullanılması ile elde edilmiştir [45]. Bunlar İmleç hareketlerinden elde edilen verilerin analiz sürecindeki metrikler olarak tanımlanır.

Zamana baęlı hareket ve fare tuşlarına basma metrikleri olarak kullanabilecek özellikler Tablo 2.1.' de belirtilmiştir. Bu özellikler imleç hareket hızı ile ilgili ve başlama veya bitiş işlemleri sırasındaki tıklama veya çift tıklama özellikleri ile ilgilidir.

Tablo.2.1. Zamana baęlı hareket metrikleri [45]

Reaksiyon Süresi	Kullanıcının bir görevi gerçekleştirirken başlama ve bitiş işlemi arasındaki geçen zaman
Hız	Bir hareketin toplam mesafesinin hareket süresine bölümü
Tepe Hızı	İmleç hareketindeki maksimum hız
Tepe Hızına kadar geçen süre	Hareket başlangıcından tepe hızına ulaşılan kadar geçen zaman aralığı
Hızlanma	Tüm hareket yörüngelerindeki hızlanma ortalaması
Hareket hızı	Hareket yörüngesi üzerinde ortalama hız
Yavaşlama	Tüm hareket yörüngelerindeki yavaşlama ortalaması
Hareket Süresi	Hareketin başlangıcından hareketin bitiş noktasına kadar geçen zaman aralığı
Tıklama Zamanı	Hareketin bitiş noktasına geldiği zaman ile fare düğmesine basma arasındaki zaman aralığı
Tıklama Sıklığı	Bir zaman birimindeki tıklama sayısı
Tıklama Süresi	Fare tuşlarındaki basma ve bırakma olayları arasındaki süre
Basılı Tutma Süresi	Hareketin bitiş noktasına ulaşıldıktan sonra kullanıcının fare düğmesini basılı tutma süresi

Mekânsal hareket ölçüleri ise imleç hareket yörüngeleri ile ilgili özellikler Tablo 2.2. de belirtilmiştir. Burada imleç hareket sırasındaki katledilen mesafe, ideal yol olarak da adlandırılan görev eksen uzunluğu ise imleç hareketinin başlama noktası ile hareketin bitiş noktası arasındaki en kısa yol uzunluğu, imleç hareket yön değişiklikleri, görev eksen geçişleri gibi metriklerden oluşturulur. Hareket doğruluğu metrikleri olarak kullanabilecek özellikler Tablo 2.3. de belirtilmiştir. Bu özellikler kullanıcının hedefleme kontrol eylemlerinin belirlendiği metrikler olan mutlak hata, yatay hata dikey hata gibi metriklerden oluşturulur. Hareket doğruluk metrikleri, kullanıcıların imleç hareketinin bitiş noktası ile hedef merkezinin arasındaki isabet farkı olarak da açıklanabilir. Bazı araştırma çalışmaları, bilgisayar kullanıcılarının ilgisi ve hissi gibi

Tablo 2.2. Mekânsal hareket ölçüleri [45]

Yol Uzunluğu	Başlangıç ve bitiş arasındaki Toplam kat edilen yol uzunluğunun
En İyi Yol Uzunluk Oranı	Toplam kat edilen yol uzunluğunun, başlangıç ve bitiş arasındaki en kısa yolun uzunluğuna bölümüdür.
Eksen Geçişleri	Hareketin yönünün görev eksenini geçme sayısı. Görev eksenini, başlangıç noktası ile bitiş noktası arasındaki düz çizgi olarak tanımlanır.
Hareket Yönü Değişiklikleri	Hareketin yönünün görev eksenine dikey olarak değiştirme sayısı
Hareket Hedefleme	Bir hareket yolunun ideal (düz) hareketten standart sapması
Ortogonal Hareket Değişiklikleri	Hareketin görev eksenine paralel olarak yön değiştirme sayısı
Hareket Değişkenliği	Hareket yolunun görev eksenine olan mesafenin standart sapması

insan davranışlarını anlamak için imleç hareketlerini analiz etmiştir. Yapılan çalışmalarda kullanıcıların imleç hareket modelleri çıkarılarak, kullanıcıların kişisel özellikleri ile ilgili tahminler yapılmaya çalışılmıştır. Bu hareket modelleri imleç yörüngelerinin, ortalama yörünge uzunluğu, kat edilen mesafe, tıklama süresi gibi imleç etkinliklerinden oluşturulmuştur. [1]'de imleçleri kontrol etme ve hedefleme ile ilgili yaklaşık 20 yörünge özelliğinin içeren bir analiz yapılmıştır. Hedefleme ve kontrol etme iki tür imleç yörünge özelliği ölçülmüştür. Hedefleme özelliğinde atraksiyon, en kısa yoldan ayrılma alanı olarak tanımlanır. Bu, bir imleç yörüngesinin ideal yoldan sapma derecesi ile belirlenir. Bu değerlendirme başlangıç konumu olarak belirlenen konumdan, bitiş olarak belirlenen konuma olan en kısa düz bir doğru olarak tanımlanır. Bu en kısa yol olan doğru, ideal yol olarak da adlandırılır. Özellikle, ideal yoldan uzaklaşan bir yörüngesinin alanını ve ideal yola yaklaşan bir yörüngesinin alanının hesaplanması ile kullanıcının imleç hedefleme özelliği belirlenebilir [54]. Diğer yörünge özelliği imleç kontrolüdür. İmleç kontrolü sırasındaki kullanıcının yaptığı yön değiştirmelere zikzak adı verilir [54]. Bu başlangıç konumundan son konuma doğru düz çizgiye göre değişen yönlerin sayısı şeklinde ifade edilir Kullanıcının imleç hareketlerindeki yörüngesinin belirli bir zaman dilimindeki yön değiştirmesinin temel dört yönlü bir açı şeklinde sınıflandırılarak ve bu her açı sınıfında kümülatif mesafe hesaplanarak imleç kontrol değişimleri belirlenir. İmlecini başlangıç noktasından bitiş

Tablo 2.3. Hedefleme kontrol ölçüleri [45]

Mutlak Hata	Hareket uç noktası ile hedef merkezinin arasındaki Öklid uzaklığıdır.
Yatay Hata	Hareket uç noktası ile hedef merkezinin arasındaki yatay koordinat (x koordinat eksenini) arasındaki farktır. Sonucun negatif olması hareket bitiş noktasının hedef konumun altında kalması, pozitif olması ise hedef konumun üstünde kalmasını anlamına gelmektedir.
Dikey Hata	Hareket bitiş noktası ile hedef merkezinin arasındaki dikey koordinat (y koordinat eksenini) arasındaki farktır. Sonucun negatif olması hareket bitiş noktasının hedef konumun solunda kalması, pozitif olması ise hedef konumun sağında kalmasını anlamına gelmektedir.
Mutlak Yatay Hata	Hareket uç noktası ile hedef merkezinin arasındaki yatay koordinatlardaki farkın mutlak değeri.
Mutlak Dikey Hata	Hareket uç noktası ile hedef merkezinin arasındaki dikey koordinatlardaki farkın mutlak değeri.

konumuna kadar olan gezinme yolunu analiz ederek eylem (imleci hareket ettirme) arasında sürekli bir etkileşim ortaya çıkardığını buldular imleç yörünge özelliklerinin gözlemcinin algısal, bilişsel ve sosyal özelliklerini yansıttığı gösterilmiştir [5].

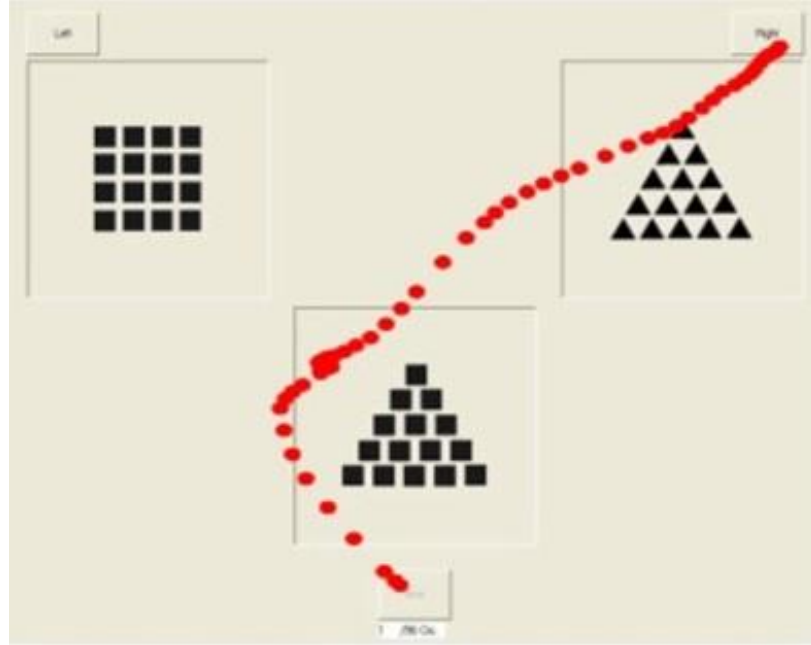
İmleç hareketleri, fare aktiviteleri olan fare tıklamaları sayısı, kat edilen mesafe, tıklama süresi gibi özellikleri içerir [5]. Bu özellikleri kullanarak mesafe, hız ve yön değişikliği gibi imleç hareket modelleri (hedefleme hassasiyeti, tıklama sayısı, hareket verimliliği, hız ve tıklama süresi) çıkarılır [5].

[5]'de kullanıcıların üç fare hareket görevi gerçekleştirir. Bunlar işaretleme, sürükler-bırak ve yönlendirme. Bu hareket ve kontrol özelliklerinin seyahat edilen mesafe ile birbirini takip eden iki butona basma olayı arasındaki en kısa mesafenin standart sapmasının katılımcıların değerlik derecelendirmeleri ile anlamlı bir şekilde ilişkili olduğunu bulmuşlardır. Yine bu çalışmada iki yörünge özelliği olan atraksiyon ve zikzağa odaklanmıştır. Bu özellikler atraksiyon, başlangıç konumunu bitiş konuma bağlayan düz çizgiden hareket alanı olarak tanımlanır; zikzak, düz çizgiye göre yön

değiştirme sayısıdır. Bilişsel karar verme ve motor kontrolünde bu yörünge özellikleri önemli olduğu belirlenmiştir.

2.2. Veri Kümesi

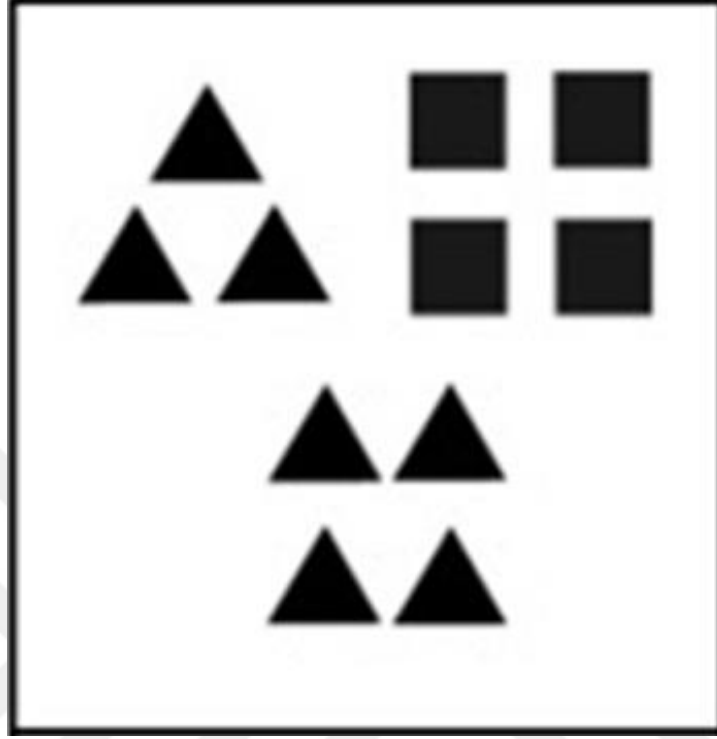
Bu tez kapsamında kullanılan veri kümesi [1]'de yapılan çalışmadan sağlanmıştır. Çalışmada kullanıcılardan imleç hareket verilerinin elde edilmesi, basit şekillerin kullanıcılara gösterilerek bu şekillerin benzer olanlarının yakalanması ile ilgili görsel algı görevlerinden oluşmaktadır. Kullanıcıların imleç hareket verileri, görsel şekilleri eşleme işlemi uygulanan bir program vasıtasıyla elde edilmiştir. Bu görev, bilgisayar ekranında üzerinde 3'lü geometrik şekillerin olduğu Şekil 2.1'de ki gibi bir yazılım ile gerçekleştirilmiştir. Yazılımda, temel şekil ekranın altında gösterilmekte ve kullanıcıdan seçmesi için ekranın üstünde 2 adet şekil gösterilmektedir. Kullanıcıdan, üstte gösterilen şekillerden temel şekle benzeyenini seçerek belirlemesi beklenmektedir. Kullanıcı en fazla 6 saniyede seçim işlemi tamamlamalıdır. Kullanıcılar tercihlerini şekillerin üstüne yerleştirilmiş olan butonlar (sağ ve sol) vasıtasıyla belirtirler.



Şekil 2.1. İmleç hareket verileri toplama program arayüzü [1]

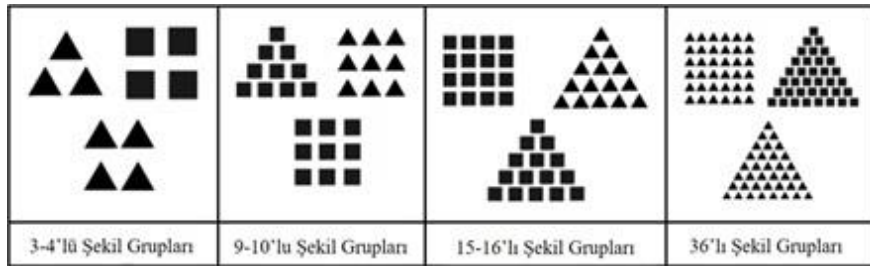
Kullanıcı her denemeyi Şekil 2.1.'de verilen ekranın altında bulunan ileri butonuna basarak başlatır. Kullanıcılara gösterilen şekil grupları, Şekil 2.2.'de gösterilen her bir

şekil, daha küçük kareler veya üçgenler içeren dört tür şekil içeren genel bir şekilden oluşmaktadır.



Şekil 2.2. Kullanıcılara gösterilen şekil grupları [1]

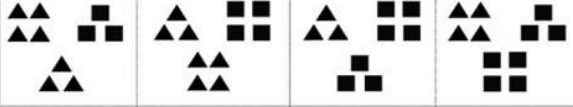
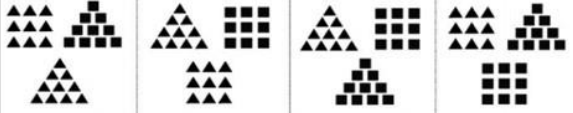
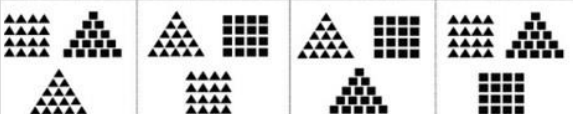
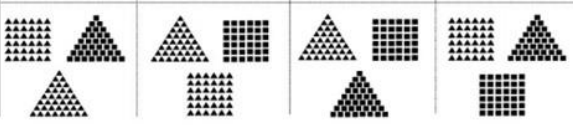
Şekil grupları, 3'lü geometrik şekillerin diziliş ve sayıları değiştirilerek oluşturulmuştur. Bu 3'lü veya 4'lü şekil grupları Şekil 2.3.'te gösterildiği gibi 3-4, 9-10, 15-16 ve 36'lık küçük şekillerden oluşturulmuştur.



Şekil 2.3. Farklı şekil grupları [1]

Bu şekil gruplarındaki şekillerin farklı dizilişleri ile Şekil 2.4.'te gösterildiği gibi toplam 16 tane temel üçlü şekil grubu oluşturulmuştur.

Kullanıcıya görev sırasında bu temel şekil gruplarının yerleri değiştirilerek gösterildiğinden her bir görev setinde 32 tane şekil tanımlama görevi yaptırılır.

1.Seviye Şekil Grupları	
2.Seviye Şekil Grupları	
3.Seviye Şekil Grupları	
4.Seviye Şekil Grupları	

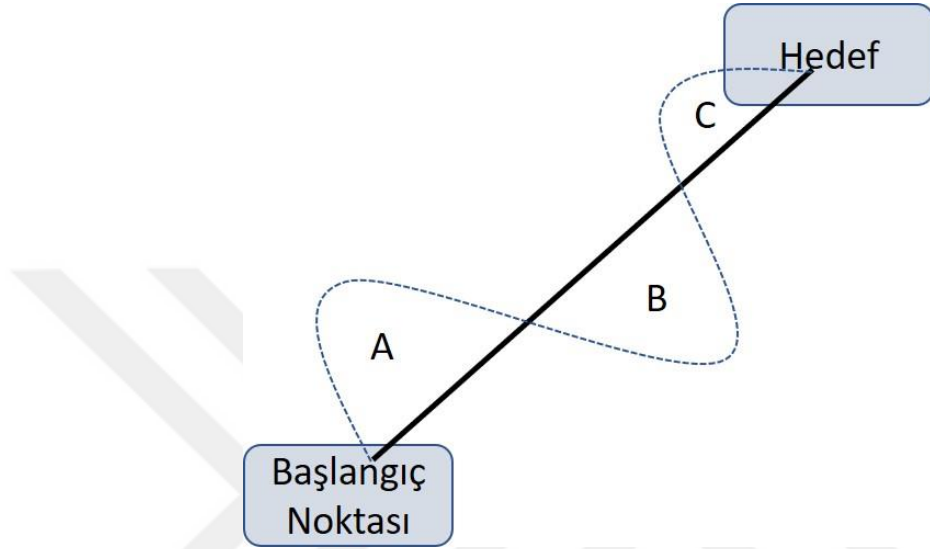
Şekil 2.4. Şekil gruplarındaki şekillerin farklı dizilişleri [1]

Toplamda yerel şekillerin sayısı değiştirilerek 16 temel deneme platformu oluşturulur. Çalışmada kullanılan şekiller 16 farklı şekil kümesinden 3'lü veya 4'lü şekil gruplarının rastgele gösterilmesiyle başlar. Deneyde seçim şekillerinin yerlerini değiştirerek 16 temel üçlü şekil grubundan 32 adet ekranda gösterilecek 3'lü şekil grubu üretilir. Bu üretilen 32 adet şekil gurubu her kullanıcı için 3'er kez gösterilmiştir. Böylece, her kullanıcı için 96 adet seçim denemesi elde edilmiştir. Kullanıcı her denemeyi başlatmak için ileri butonuna basarlar ve ekranda yeni bir 3'lü şekil grubu gösterilir. Kullanıcı tabanda gösterilen benzer şekil için seçimini, Şekil-1'de gösterilen şekillerin üstünde bulunan sağ-sol butonuna basarak belirlerler. Bu döngü her kullanıcı için 96 kez tekrar edilir. Bu görevde, kullanıcılara doğru/yanlış cevap bulunmadığı ve seçimlerini kişisel algılarına göre yapmaları söylenmiştir. Her denemede, program imleç konumunun x-y koordinatlarını veri noktası olarak, denemenin başlangıcından kullanıcının seçimini yaptığı sağ veya sol tuşuna basana kadar her 20-30 milisaniyede kaydeder. Böylece bu veri kümesinden, 16 imleç hareket özelliği çıkarılmıştır.

İmleç hareket verileri kullanıcının her göreve başlaması (ileri butona basarak) ile bitirmesi (sağ veya sol butonun seçilmesi) işleminin gerçekleşmesine kadar 100 eşit zaman diliminde toplanmıştır. Her imleç yörüngesi için, 100 zaman dilimi 4 eşit parçaya ayrılmıştır ve her 4 zaman dilimi için Şekil 2.8.'de gösterildiği gibi 2 adet öznitelik çıkarılmıştır.

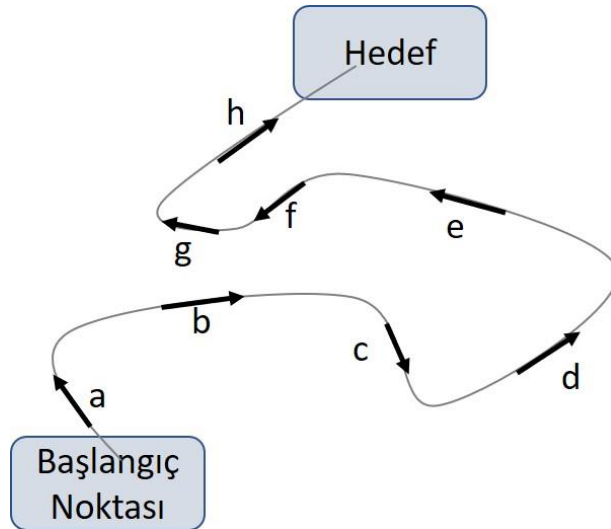
Bu çalışmada 2 tür imleç hareket özelliği incelenmiştir; hedefleme ve kontrol. Hedefleme, imleç yörüngesinin ideal yoldan ne kadar saptığı değerlendirilmiştir. İdeal

yol; başlangıç noktasından bitiş noktasına olan düz bir çizgidir (noktalar arası en kısa mesafe). Bu sapma ölçümü ideal yoldan uzaklaşmış ile ideal yola yaklaşmış bölgelerin alanlarının hesaplanması ile gerçekleşmiştir. Diğer bir ifadeyle hedefleme Şekil 2.5.'te gösterildiği gibi; başlangıç ve hedef arasındaki en kısa yoldan ayrılma alanı olarak belirtilebilir.



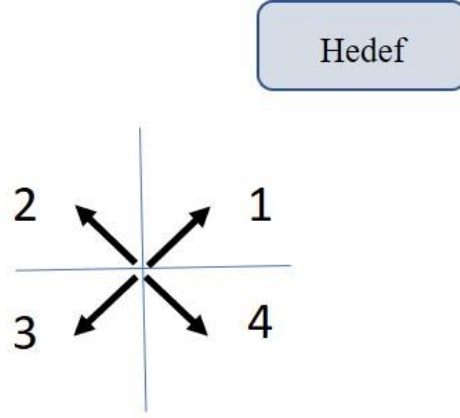
Şekil 2.5. İmleç hareketinde ideal yoldan ayrılma [1]

Kontrol ise; Şekil 2.6.'da gösterildiği gibi yine başlangıç noktası ile bitiş noktasına doğru yörüngedeki yön değiştirme miktarı olarak tanımlanır.



Şekil 2.6. İmleç hareketinde yön değiştirme [1]

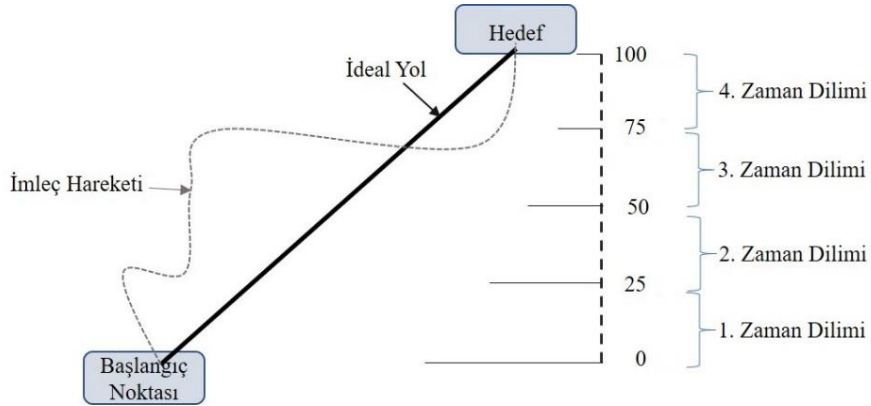
Kullanıcıların hareketindeki yön değiştirmeler Şekil 2.7.'de verildiği gibi 4 yön sınıfı olarak sınıflandırılmıştır.



Başlangıç
Noktası

Şekil 2.7. İmleç hareketi yön sınıfları [1]

Bu yön deęiřtirmeler, belirlenen 4 yön sınıfı için kümülatif mesafe olarak hesaplanmıřtır. Her kullanıcı için bu eřit zaman dilimlerindeki deęerlerin ortalaması ve standart sapması hesaplanmıřtır. Daha sonra bu ortalama ve standart sapma deęerleri 0 ile 1 arasında normalize edilmiřtir. Böylece her kullanıcı için 16 öznitelik çıkarılmıř olur (2 öznitelik x 4 zaman dilimi x 2 istatistiksel özellik (ortalama-standart sapma)). Bu öznitelikler, Şekil 2.8.'de gösterildięi gibi 4 eřit zaman dilimindeki 2 öznitelięin (hedefleme ve kontrol) ortalama ve standart sapma deęerlerinden oluřmuřtur.



Şekil 2.8. Eřit zaman dilimindeki özniteliklerin elde edilmesi [18]

Yapılan alıřmalarda bu imle yörünge özellikleri biliřsel karar vermede önemli olduęunu göstermiřtir [1]. Bu görev ortamında doęru veya yanlıř cevap olarak bir deęerlendirme yapılmamıř, kullanıcıların kiřisel tercihlerine göre seim yapmaları

beklenmiştir. Bunun nedeni kullanıcıların tercihlerinde karşılaştırma ve seçme gibi daha karmaşık ve gerçekçi durumlarla ilgili olmasıdır.



3. MAKİNE ÖĞRENMESİ

Makine öğrenimi son yıllarda oldukça yaygın kullanılan bilgisayar bilimlerinin bir dalı olmuştur. Bu popülerlik büyük verilere kolay ulaşma, bu verilerin işlenmesinde kullanılan donanımların güçlü özellikleri ve insanların büyük verilerden bilgi keşfi merakı oldukça etkili olmaktadır. Makine öğrenimi, veriden bilgi çıkarma işi olarak kısaca tanımlanabilir. İstatistik, yapay zeka ve bilgisayar biliminin kesişim noktasında bir araştırma alanıdır ve aynı zamanda tahmini analitik veya istatistiksel öğrenme olarak da bilinir [55]. Diğer bir tanımlamada; makine öğrenmesi, matematiksel ve istatistiksel yöntemler kullanarak eldeki verilerden bir tahmin veya anlamlı sonuç çıkarımları yapan, yöntemler bütünü olarak tanımlanabilir. Bu alan bilgisayar görme, işitme, doğal dil işleme, görüntü işleme ve örüntü tanıma, bilişsel hesaplama, bilgi gösterimi vb. bilgisayar bilimleri konularına odaklanmaktadır. Bu araştırma eğilimleri, makinelere insan duyularına benzer duyular yoluyla veri toplama ve daha sonra tahminleri yürütme ve insanlarla aynı seviyede kararlar vermek için hesaplama zekâsı araçlarını ve makine öğrenme yöntemlerini kullanarak toplanan verileri işlemek için yetenekler sağlamayı amaçlamaktadır. Makine öğrenmesinin amacı, makinelerin tahminlerde bulunmalarını, kümelenmeyi gerçekleştirmelerini, ilişkilendirme kurallarını çıkarmalarını veya belirli bir veri kümesinden kararlar almalarını sağlamaktır. Makine öğrenimi yöntemlerinin uygulanması son yıllarda günlük yaşamda her yerde yaygınlaşmıştır.

Makine öğrenimindeki güncel araştırmalar çok çeşitli pratik uygulamaları bulunmaktadır. Bunlardan bazıları;

- Metin veya belge sınıflandırması: Bu, bir metnin veya belgenin konusunu belirlemek veya bir web sayfasının içeriğinin uygunsuz olup olmadığını otomatik olarak belirlemek ayrıca spam tespitini de içermektedir.
- Doğal dil işleme (NLP): Bu alanda, konuşma parçası etiketlemesinde, bir cümle tahmini, her kelimeyi etiketleyen konuşma parçası etiketlenmesini içermektedir.

- Konuşma işleme uygulamaları: Bu alanda, konuşma tanıma, konuşma sentezi, konuşmacı doğrulama, konuşmacı tanımlamanın yanı sıra dil modelleme ve akustik modelleme gibi alt sorunları da içerir.
- Bilgisayarlı görme uygulamaları: Buna nesne tanıma, nesne tanıma, yüz tanıma, OCR, içerik tabanlı görüntü alma veya poz tahmini içerir.
- Hesaplamalı biyoloji uygulamaları: Bu alanda, protein fonksiyon tahmini, anahtar alanların tanımlanması veya gen ve protein ağlarının analizini içerir.
- Kredi kartı, telefon veya sigorta şirketleri için sahtekarlık tespiti, ağa izinsiz giriş, satranç, tavla gibi oyunları oynamayı öğrenme, robot veya otomobil gibi araçların yardımsız kontrolü, tıbbi teşhis, öneri tasarımı gibi diğer birçok sorun sistemler, arama motorları veya bilgi çıkarma sistemleri, makine öğrenme teknikleri kullanılarak ele alınır [56].

Makine öğrenimi çalışmalarında, standart makine öğrenimi işlevleri bulunmaktadır. Bunlar kısaca aşağıda açıklanmıştır.

- Sınıflandırma: Bu işlev, her bir öğeye bir kategori atama işlemidir. Örneğin, belge sınıflandırması her belgeye siyaset, işletme, spor veya hava durumu gibi bir kategori atamaktan, görüntü sınıflandırması ise her görüntüye araba, tren veya uçak gibi bir kategori atamaktan oluşur.
- Regresyon: Bu işlev, her bir veri için gerçek bir değer öngörme problemidir. Regresyon örnekleri, hisse senedi değerlerinin veya ekonomik değişkenlerin varyasyonlarının tahminini içerir. Regresyonda, çeşitli kategoriler arasında tipik olarak yakınlık kavramının olmadığı sınıflandırma probleminin aksine, gerçek ve öngörülen değerler arasındaki farkın büyüklüğüne bağlıdır.
- Sıralama: Bu işlev, belirlenen bazı kriterlere göre verinin veya içeriğin elde edilmesinin öğrenme problemidir. Web araması, bu arama sonucu alakalı web sayfalarının getirilmesi standart sıralama örneğidir. Diğer birçok benzer sıralama problemi, bilgi çıkarma veya doğal dil işleme sistemlerinin tasarımı bağlamında ortaya çıkmaktadır.
- Kümeleme: Bu işlev, bir dizi verinin homojen alt kümelere bölünmesi durumudur. Kümeleme genellikle çok büyük veri kümelerini analiz etmek için kullanılır. Örneğin, sosyal ağ analizi bağlamında, kümeleme algoritmaları büyük insan gruplarındaki olası toplulukları belirlemeye çalışır.

- Boyutsal küçültme: Bu işlev, verilerin ilk temsilini daha düşük boyutlu bir temsile dönüştürürken, ilk temsil durumundaki bazı özelliklerini koruyarak oluşturulur. Yaygın bir örnek, bilgisayar görüntü görevlerinde dijital görüntülerin önceden işlenmesini içerir.

En başarılı makine öğrenme algoritmaları, bilinen örneklerden genelleştirerek karar alma süreçlerini otomatikleştiren algoritmalar [55]. Makine öğrenimi algoritmalarında öğrenmek için veri gerektirdiğinden, verinin doğruluğu ve işlenmesi elde edilmesi istenen sonucunun doğruluğuyla doğrudan ilişkilidir [57]. Veriler, sayısallaştırılmış etiketli eğitim setleri veya çevre ile etkileşim yoluyla elde edilen diğer bilgi türleri şeklinde olabilir. Her durumda, kalitesi ve büyüklüğü öğrencinin yaptığı tahminlerin başarısı için çok önemlidir [56]. Makine öğrenimi algoritmaları makinelerin öğrenmesini ve sonuçların tahmini gibi görevleri yerine getirmelerini sağlayan bilgisayar programlarıdır. Makine öğrenimi algoritmasının amacı, girdiyi alan ve istenen sonucu üreten bir model oluşturmaktır. Bir makine öğrenme algoritmasının performans ölçüsü sonuçların doğruluğu yanı sıra işlem hızı ve bellek kullanımı metrikleri de çok önemli parametrelerindedir.

Dört genel makine öğrenme yöntemi vardır:

- Denetimli,
- Denetimsiz,
- Yarı Denetimli ve
- Takviyeli öğrenme yöntemleri.

3.1. Denetimli Öğrenme

Denetimli öğrenmede hedef, etiketlenmiş eğitim verilerinden bir benzerlik veya eşleme çıkarmaktır. Etiketlenmiş verilerden öğrenme süreci olarak da tanımlanabilir. Etiketler, algoritmaya verileri nasıl etiketlemesi gerektiğini öğretir [57]. Kullanıcı algoritmaya girdi ve istenen çıktı çiftlerini verir ve algoritma girdi verildiğinde istenen çıktıyı algoritmalar ile tahmin edilmesini sağlar. Bu durumda, algoritma daha önce hiç görmediği bir girdi için bir çıktı oluşturur [55]. Giriş / çıkış çiftlerinden öğrenen makine öğrenme algoritmalarına denetimli öğrenme algoritmaları da denir [55]. Sistem, eğitim verisi olarak bir dizi etiketli örnek alır ve diğer tüm noktalar için

tahminlerde bulunur. Bu sistem, sınıflandırma, regresyon ve sıralama problemleriyle ilişkili en yaygın kullanılan öğrenme yöntemlerindedir [56].

Denetimli öğrenme algoritmaları bir dizi örneğe dayalı tahminler yapar. Örneğin, geçmiş satışlar gelecekteki fiyatları tahmin etmek için kullanılabilir. Denetimli öğrenme ile, etiketlenmiş eğitim verileri ve istenen çıktı değişkeninden oluşan bir giriş değişkeni vardır. Girdi ile çıktıyı eşleştiren işlevi öğrenmek üzere eğitim verilerini analiz etmek için bir algoritma kullanılır. Bu çıkarımsal işlev, eğitim verilerinden genelleme veya bir çıkarım yaparak bilinmeyen yeni örnekleri eşleştirir.

Denetimli öğrenme çatısı altında iki grup veya algoritma kategorisi bulunur. Bunlar regresyon ve sınıflandırma kategorileridir.

3.2. Denetimsiz Öğrenme

Denetimsiz öğrenmede etiketlenmemiş verilerden kümeleme veya gruplama ile bir benzerlik elde etme işlemidir [57]. Diğer bir deyişle, etiketlenmemiş verilerden öğrenme sürecidir. Algoritmanın kendi kendine tahminler yapması, tanımlanamayan örüntüleri tahmin etmesi beklenir. Denetimsiz öğrenmede, yalnızca girdi verileri bilinir ve algoritmaya bilinen hiçbir çıktı verisi verilmez [55]. Sistem yalnızca etiketlenmemiş eğitim verilerini alır ve bilinmeyen tüm noktalar için tahminlerde bulunur. Gözetimsiz öğrenme yaparken, makineye tamamen etiketlenmemiş veriler sunulur. Genel olarak bu ortamda etiketli bir örnek bulunmadığından, bir sistemin performansını nicel olarak değerlendirmek zor olabilir. Kümeleme ve boyutsallık azalması denetimsiz öğrenme problemlerine bir örnektir [56].

3.3. Yarı Denetimli Öğrenme

Yarı denetimli öğrenme yapısal olarak denetimli öğrenmeye benzemektedir. Buradaki en büyük fark hem etiketli hem de etiketsiz verilerin kullanılmasıdır. Bu öğrenme yönteminde etiketlenmiş verilerden etiketlenmemiş yani bilinmeyen verileri tahmin etmeye dayalı bir sistemdir. Etiketleme işlemi maliyetli bir işlem olduğundan yarı denetimli öğrenme yöntemi denetimli öğrenmeden daha az maliyetli bir yöntem olarak pratiklik sağlamaktadır [58].

3.4. Takviyeli Öğrenme

Takviye öğrenme, bir makine öğrenmesi algoritmasına bir dizi eylem, parametre ve son değer içeren, düzenli öğrenme süreçlerine odaklanmaktadır. Kuralları belirleyerek, makine öğrenme algoritması farklı seçenekleri ve olasılıkları keşfetmeye çalışır. Hangisinin en uygun olduğunu belirlemek için her bir sonucu izleyerek ve değerlendirmektedir. Takviyeli öğrenme; makine denemesini ve hatayı öğretir. Bu öğrenme türü geçmiş deneyimlerden öğrenerek gerçekleşir ve böylece mümkün olan en iyi sonucu elde etmek için duruma cevap olarak yaklaşımını adapte etmeye başlar [59].

Takviye öğrenme yöntemi, ödülü en üst düzeye çıkaracak veya riski en aza indirecek eylemler yapmak için çevre ile etkileşimden elde edilen gözlemleri kullanmayı amaçlamaktadır. Takviye öğrenimi akıllı programlar üretmek için, aşağıdaki adımlardan geçer:

1. Giriş durumu aracı tarafından gözlenir.
2. Karar verme işlevi, aracının bir eylem gerçekleştirmesini sağlamak için kullanılır.
3. Eylem gerçekleştirildikten sonra, temsilci çevreden ödül veya takviye alır.
4. Ödülle ilgili durum-eylem çifti bilgileri saklanır.

Saklanan bilgileri kullanarak, eylem açısından belirli bir durum için tekrar ayar yapılabilir böylece temsilci için en uygun kararların alınmasına yardımcı olur.

Eğitim ve test aşamaları takviye öğreniminde karıştırılır. Bilgi toplamak için, çevre ile aktif olarak etkileşime girer ve bazı durumlarda çevreyi etkiler böylece her eylem için anında bir ödül alır. Amaç çevreyle ilgili bir eylem ve yineleme boyunca ödülünü en üst düzeye çıkarmaktır. [56].

3.5. Destek Vektör Makinesi

Makine öğreniminde, destek vektör makineleri, sınıflandırma ve regresyon analizi için kullanılan, verileri analiz eden ve örüntüleri tanıyan ilişkili öğrenme algoritmalarına sahip denetimli öğrenme modelleridir. Her biri iki kategoriden birine ait olarak işaretlenmiş bir dizi eğitim örneği göz önüne alındığında, bir SVM eğitim algoritması bir kategoriye veya diğerine yeni örnekler atayarak onu olasılıksal olmayan bir ikili doğrusal sınıflandırıcı haline getirir. SVM modeli, örneklerin uzaydaki noktalar olarak temsil edilmesidir, böylece ayrı kategorilerin örnekleri mümkün olduğunca geniş bir

açıklığa bölünür. Daha sonra aynı alana yeni örnekler eşleştirilir ve boşluğun hangi tarafına düştüklerine göre bir kategoriye ait olduğu tahmin edilir.

Destek Vektör Makineleri, iki farklı sınıfı ayıran bir ayırma çizgisi, düzlem veya hiper düzlemlerden oluşur [60].

Destek Vektör Makineleri, denetlenen öğrenme yöntemlerinde kullanılan etiketlenmiş verilerden yeni verileri sınıflandırmak için kullanılan öğrenme modelidir. Verileri sınıflandırmak için temel yaklaşım, veri noktalarını belirlenen etiketlere mümkün olan en az hata miktarıyla veya mümkün olan en geniş kenar boşluğu. Bunun nedeni, bölme işlevinin yanındaki daha büyük boş alanların daha az hataya neden olmasıdır, çünkü etiketler birbirinden daha iyi ayırt edilir. Böylece bir veri kümesinin hatasız çoklu fonksiyonlarla çok iyi ayrılabilir. Bu nedenle, ayırma işlevinin etrafındaki kenar boşluğu, ayırmanın kalitesini değerlendirmek için ek bir parametre olarak kullanılmaktadır. Bu durumda, iki sınıfı daha kesin bir şekilde ayırt ettiği için A ayrımı daha iyidir [61]. Destek Vektör Makineleri n boyutlu bir alanda bir veya daha fazla hiper düzlem oluşturur. Verileri bölme işlemindeki ilk girişim, verileri ilgili etiketlere doğrusal olarak ayırmaya çalışmaktır. Veriler doğrusal bir şekilde tamamen ayrılabilirse, ortaya çıkan işlev gelecekteki olayları sınıflandırmak için kullanılabilir. [61].

Veriler iyi bir şekilde doğrusal olarak ayrılamayabilir, eğer sınıflar doğrusal olarak ayrılamıyor ise bu durumda girdi verilerini yüksek boyutlu özellik alanlarına dolaylı olarak eşleyerek çekirdek değer değişimi adı verilen doğrusal olmayan bir sınıflandırma gerçekleştirebilir. Diğer bir deyişle doğrusal olmayan ayrılabilir bir alanı daha yüksek boyutlu doğrusal olarak ayrılabilir bir alana dönüştürülerek bu sorun giderilir [56,61,62].

Bu çekirdek fonksiyonları lineer olarak ayrılamayan veriyi, daha yüksek bir boyuta taşıyarak bir düzlem ile ayırım yapılmasını sağlamaktadır. Çekirdek fonksiyonları sınıflama ve regresyon problemlerinde bir benzerlik fonksiyonu olarak görev yapmaktadır. Bu fonksiyonlar makine öğrenmesi algoritmalarına çeşitli esnek özellikler kazandırır. Birçok çekirdek fonksiyonu vardır. Bunlara Lineer, Polinomial, Gaussian, Gaussian RBF, Sigmoidal çekirdek fonksiyonları örnek olarak verilebilir.

Polinomial çekirdek fonksiyonu sabit olmayan bir çekirdektir. Polinomial çekirdeği, tüm eğitim verilerinin normalleştirildiği problemlerde kullanılan bir fonksiyondur. Bu fonksiyonun parametreleri olan, alfa eğimi, c sabit terimi ve d polinom derecesini ifade etmektedir İlgili çekirdek fonksiyonu Denklem (3.1)'deki formül ile ifade edilir;

$$k(x_i, x_j) = (\alpha x_i \cdot x_j + c)^d \quad (3.1)$$

Gaussian çekirdek fonksiyonu Denklem (3.2)'deki formül ile ifade edilir;

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (3.2)$$

Gaussian çekirdeği, radyal taban fonksiyonu çekirdeğine bir örnektir. İlgili çekirdek fonksiyonu Denklem (3.3)'deki formül ile ifade edilir;

$$k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \quad (3.3)$$

Bu fonksiyon da değiştirilebilir parametre olan sigma ve gama dikkatli seçilmelidir. Çünkü bu parametre çekirdeğin performansını doğrudan etkilemektedir. Aşırı tahmin meydana geldiğinde, üstel ifade doğrusallaşarak, doğrusal olmayan yapısını yitirmiş bir yüksek boyutlu projeksiyon meydana gelir.

4. İMLEÇ HAREKETLERİNİN CİNSİYETE GÖRE SINIFLANDIRILMASI

Tez kapsamında kullanıcılardan toplanan imleç hareketleri denetimli öğrenim makine öğrenmesi yöntemine göre sınıflandırılma işlemi yapılmıştır.

4.1. Yöntem

Katılımcılar lisans öğrencilerinden oluşmakta ve bu kullanıcılardan 177 adet veri toplanmıştır. Bu katılımcıların 111 kadın, 66 tanesi erkeklerden oluşmaktadır. Çalışmada kullanılan veriler bölüm 1.1 de açıklandığı şekilde, basit şekillerin kullanıcılara gösterilerek bu şekillerin benzer olanlarının yakalanması ile ilgili görsel algı görevleri sırasındaki imleç hareketlerinin kaydedilmesi ile elde edilmiştir. İmleç hareketlerinden çıkartılan öznitelikler kullanarak Destek Vektör Makinesi tabanlı ikili bir sınıflandırılma işlemi gerçekleştirilmiştir. Önerilen yöntemin başarımı, farklı çekirdek fonksiyonlarıyla da kıyaslanmıştır. Veri Kümesi 111 Kadın ve 66 Erkek katılımcılarından olmak üzere toplam 177 imleç hareketlerinden oluşmuştur. Bu hareket verileri hedefleme ve yön özelliklerinin 4 eşit zaman dilimindeki, ortalama ve standart sapma değerlerinden oluşmaktadır. Böylece her bir cinsiyet için toplam 16 tane öznitelik vektörü elde edilmiştir. Bu 16 öznitelik destek vektör makinesi kullanılarak ikili bir sınıflandırma yapılmıştır. Sınıflandırma işleminde Destek Vektör Makinesi için seçilecek çekirdek fonksiyon önemli rol oynadığından dolayı bu çalışmada Lineer, Polinomial ve Gaussian tabanlı çekirdek fonksiyonları kullanılarak modelin başarımı kıyaslanmıştır. Bu çekirdek fonksiyonlarında Polinomial çekirdek fonksiyonu için iki farklı parametre olan Quadratic ve Cubic değerleri kullanılmıştır. Aynı zamanda Gaussian çekirdek fonksiyonu için dört farklı çekirdek parametresi olarak Fine, Medium, Coarse ve Radyal çekirdek (RBF) değerleri kullanılmıştır. Makine öğrenmesi modellerinin değerlendirilmesinde kullanılan metriklerden olan hata matrisi, çalışmada kullanılan her çekirdek fonksiyonu için ayrı ayrı çıkarılmıştır. Tablo 4.1’de gösterilen hata matrisi; oluşturulan sınıflandırma modeli için kullanılan veri kümesinde gerçek sonuçlar ile yapılan doğru ve yanlış tahminlerin sayısını gösteren bir matristir.

Tablo 4.1. Hata matrisi

		GERÇEK DURUM	
		Pozitif	Negatif
TAHMİN	Pozitif	TP	FP
	Negatif	FN	TN

Hata matrisinde sınıflandırma tahminleri dört değerlendirmeden birine aittir. Bunlar; TP: Doğru sonuç, doğru olarak tahmin edilmişse, TN: Doğru sonuç, yanlış olarak tahmin edilmişse, FP: Yanlış sonuç, doğru olarak tahmin edilmişse, FN: Yanlış sonuç, yanlış olarak tahmin edilmişse. Hata matrisinde kullanılan ölçü metrikleri; Doğruluk (Accuracy), Hassasiyet (Recall), Kesinlik (Precision) ve F-Skor'dur. Doğruluk (Accuracy); modelde doğru tahmin edilen veri sayısının toplamının veri kümesindeki veri sayısına oranı ile hesaplanır. Formülü Denklem (4.1)'deki gibi;

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

ifade edilir. Kesinlik (Precision), doğru tanımlanmış pozitif sonuçların sayısının, doğru tanımlanmayanlar da dahil olmak üzere tüm pozitif sonuçların sayısına bölünmesidir. Diğer bir ifadeyle; pozitif olarak tahmin ettiğimiz değerlerin gerçekten kaç adedinin pozitif olduğunu göstermektedir. Formülü Denklem (4.2)'deki gibi;

$$\text{Kesinlik} = \frac{TP}{TP + FN} \quad (4.2)$$

ifade edilir. Hassasiyet (Recall) ise; doğru olarak tanımlanmış pozitif sonuçların sayısının, pozitif olarak tanımlanmış olması gereken tüm örneklerin sayısına bölünmesiyle elde edilir. Pozitif olarak tahmin etmemiz gereken işlemlerin ne kadarını pozitif olarak tahmin ettiğimizi gösteren bir metriktir. Formülü Denklem (4.3)'deki gibi;

$$\text{Doğruluk} = \frac{TP}{TP + FP} \quad (4.3)$$

ifade edilir. F-Skor, ikili sınıflandırma analizi için testin doğruluğunun bir ölçüsüdür. F-Skor değeri, kesinlik (Precision) ile hassasiyet (Recall) değerlerinin harmonik

ortalaması ile elde edilir [64,65]. Sınıflandırma analizinde doğruluk (Accuracy) ile F-Skor'un kullanılmasının amacı eşit dağılmayan veri kümeleri için doğru model seçimi yapmak içindir [63]. Doğru tasarlanmış modelin F-Skor değeri 1'dir. F-Skor hesaplanmasına ait formül Denklem (4.4)'deki gibi;

$$F\text{-Skor} = \frac{2 \times (\text{Kesinlik} \times \text{Hassasiyet})}{(\text{Kesinlik} + \text{Hassasiyet})} \quad (4.4)$$

ifade edilir. Sınıflandırma işlemi hedefleme ve kontrol özniteliklerinden oluşan toplam 16 öznitelik veri kümesi, sadece hedefleme özniteliğinden oluşan 8 öznitelik kümesi ve sadece kontrol özniteliğinden oluşan 8 öznitelik kümesi için ayrı ayrı yapılmıştır. Bütün yazılımsal işlemler Matlab R2018b 64 bit üzerinde gerçekleştirilmiştir.

4.2. Eğitim Aşaması

Eğitim aşaması, sistemin öğrenme algoritmaları kullanarak problemin çözümünü öğrendiği aşamadır. Bu çözüm aşaması sistemin oluşturduğu öğrenme modeli üzerinden gerçekleşir. Bu tez kapsamında 177 kullanıcıdan toplanan hedefleme ve kontrol özniteliklerinden oluşan toplam 16 öznitelik veri kümesinin %80'i eğitim (142 veri) için kullanılmıştır.

4.3. Test Aşaması

Test aşamasında, sistemin öğrenmiş olduğu model test edilir. Böylece oluşturulan modelin sistemin problem çözümünde ne kadar başarılı olduğu gözlemlenir: Bu tez kapsamında, 177 kullanıcıdan toplanan hedefleme ve kontrol özniteliklerinden oluşan toplam 16 öznitelik veri kümesinin %20'si test (35 veri) için kullanılmıştır.

4.4. Deneysel Sonuçlar

Tez kapsamında yapılan çalışmada, 16 öznitelik eğitim veri kümesi için Tablo 4.2'de lineer çekirdek fonksiyonu hata matrisi, Tablo 4.3'de Polinomial çekirdek fonksiyonun Quadratic ve Cubic parametrelerine ait hata matrisleri ve Tablo 4.4'de Gaussian çekirdek fonksiyonu Fine, Medium, Coarse ve RBF parametrelerine ait hata matrisleri verilmiştir. 16 öznitelik test veri kümesi için Tablo 4.5'de lineer çekirdek fonksiyonu hata matrisi, Tablo 4.6'da Polinomial çekirdek fonksiyonun Quadratic ve

Cubic parametreleri hata matrisleri ve Tablo 4.7’de Gaussian çekirdek fonksiyonu Fine, Medium, Coarse ve RBF parametrelerine ait hata matrisleri verilmiştir.

8 hedefleme öznitelik eğitim veri kümesi için Tablo 4.8’de lineer çekirdek fonksiyonu hata matrisi, Tablo 4.9’da Polinomiyal çekirdek fonksiyonun Quadratic ve Cubic parametreleri hata matrisleri ve Tablo 4.10’da Gaussian çekirdek fonksiyonu Fine, Medium, Coarse ve RBF parametrelerine ait hata matrisleri verilmiştir. 8 hedefleme öznitelik test veri kümesi için Tablo 4.11’de lineer çekirdek fonksiyonu hata matrisi, Tablo 4.12’de Polinomiyal çekirdek fonksiyonun Quadratic ve Cubic parametreleri hata matrisleri ve Tablo 4.13’de Gaussian çekirdek fonksiyonu Fine, Medium, Coarse ve RBF parametrelerine ait hata matrisleri verilmiştir.

8 kontrol öznitelik eğitim veri kümesi için Tablo 4.14’de lineer çekirdek fonksiyonu hata matrisi, Tablo 4.15’de Polinomiyal çekirdek fonksiyonun Quadratic ve Cubic parametreleri hata matrisleri ve Tablo 4.16’da Gaussian çekirdek fonksiyonu Fine, Medium, Coarse ve RBF parametrelerine ait hata matrisleri verilmiştir. 8 kontrol öznitelik test veri kümesi için Tablo 4.17’de lineer çekirdek fonksiyonu hata matrisi, Tablo 4.18’de Polinomiyal çekirdek fonksiyonun Quadratic ve Cubic parametreleri hata matrisleri ve Tablo 4.19’da Gaussian çekirdek fonksiyonu Fine, Medium, Coarse ve RBF parametrelerine ait hata matrisleri verilmiştir. Tablo 4.20, Tablo 4.21 ve Tablo 4.22’de imleç hareketlerinin sınıflandırma başarımları ve F-Skor değerleri verilmiştir.

Tablo 4.2. 16 öznitelik eğitim veri kümesi lineer çekirdek fonksiyonu hata matrisi

	KADIN	ERKEK
KADIN	80	5
ERKEK	38	19

Tablo 4.3. 16 öznitelik eğitim veri kümesi Polinomiyal çekirdek fonksiyonu (a) Quadratic, (b) Cubic parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	81	4
ERKEK	40	17

(a)

	KADIN	ERKEK
KADIN	85	0
ERKEK	2	55

(b)

Tablo 4.4. 16 öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	85	0
ERKEK	0	57

(a)

	KADIN	ERKEK
KADIN	85	0
ERKEK	22	35

(b)

	KADIN	ERKEK
KADIN	84	1
ERKEK	52	5

(c)

	KADIN	ERKEK
KADIN	85	0
ERKEK	0	57

(d)

Tablo 4.5. 16 öznitelik test veri kümesi için lineer çekirdek fonksiyonu hata matrisi

	KADIN	ERKEK
KADIN	25	1
ERKEK	9	0

Tablo 4.6. 16 öznitelik test veri kümesi için Polinomiyal çekirdek fonksiyonu (a) Quadratic, (b) Cubic parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	22	4
ERKEK	8	1

(a)

	KADIN	ERKEK
KADIN	19	7
ERKEK	8	1

(b)

Tablo 4.7. 16 öznitelik test veri kümesi için Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	26	0
ERKEK	9	0

(a)

	KADIN	ERKEK
KADIN	24	2
ERKEK	9	0

(b)

Tablo 4.7. (Devam) 16 öznitelik test veri kümesi için Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	26	0
ERKEK	9	0

(c)

	KADIN	ERKEK
KADIN	26	0
ERKEK	9	0

(d)

Tablo 4.8. 8 hedefleme öznitelik eğitim veri kümesi lineer çekirdek fonksiyonu hata matrisi

	KADIN	ERKEK
KADIN	90	1
ERKEK	47	4

Tablo 4.9. 8 hedefleme öznitelik eğitim veri kümesi Polinomiyal çekirdek fonksiyonu (a) Quadratic, (b) Cubic parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	90	1
ERKEK	31	20

(a)

	KADIN	ERKEK
KADIN	90	1
ERKEK	7	44

(b)

Tablo 4.10. 8 hedefleme öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	91	0
ERKEK	6	45

(a)

	KADIN	ERKEK
KADIN	91	0
ERKEK	43	8

(b)

	KADIN	ERKEK
KADIN	91	0
ERKEK	49	2

(c)

	KADIN	ERKEK
KADIN	91	0
ERKEK	11	40

(d)

Tablo 4.11. 8 hedefleme öznitelik test veri kümesi için lineer çekirdek fonksiyonu hata matrisi

	KADIN	ERKEK
KADIN	20	0
ERKEK	14	1

Tablo 4.12. 8 hedefleme öznitelik test veri kümesi için Polinomial çekirdek fonksiyonu (a) Quadratic, (b) Cubic parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	13	7
ERKEK	13	2

(a)

	KADIN	ERKEK
KADIN	11	9
ERKEK	12	3

(b)

Tablo 4.13. 8 hedefleme öznitelik test veri kümesi için Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	20	0
ERKEK	15	0

(a)

	KADIN	ERKEK
KADIN	20	0
ERKEK	14	1

(b)

	KADIN	ERKEK
KADIN	20	0
ERKEK	15	0

(c)

	KADIN	ERKEK
KADIN	19	1
ERKEK	15	0

(d)

Tablo 4.14. 8 kontrol öznitelik eğitim veri kümesi lineer çekirdek fonksiyonu hata matrisi

	KADIN	ERKEK
KADIN	85	0
ERKEK	57	0

Tablo 4.15. 8 kontrol öznitelik eğitim veri kümesi Polinomiyal çekirdek fonksiyonu (a) Quadratic, (b) Cubic parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	77	8
ERKEK	25	32

(a)

	KADIN	ERKEK
KADIN	82	3
ERKEK	7	50

(b)

Tablo 4.16. 8 kontrol öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	85	0
ERKEK	0	57

(a)

	KADIN	ERKEK
KADIN	82	3
ERKEK	41	16

(b)

Tablo 4.16. (Devam) 8 kontrol öznitelik eğitim veri kümesi Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	85	0
ERKEK	57	0

(c)

	KADIN	ERKEK
KADIN	85	0
ERKEK	2	55

(d)

Tablo 4.17. 8 kontrol öznitelik test veri kümesi için Lineer çekirdek fonksiyonu hata matrisi

	KADIN	ERKEK
KADIN	26	0
ERKEK	9	0

Tablo 4.18. 8 kontrol öznitelik test veri kümesi için Polinomiyal çekirdek fonksiyonu (a) Quadratic, (b) Cubic parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	19	7
ERKEK	5	4

(a)

	KADIN	ERKEK
KADIN	14	12
ERKEK	6	3

(b)

Tablo 4.19. 8 kontrol öznitelik test veri kümesi için Gaussian çekirdek fonksiyonu (a) Fine, (b) Medium, (c) Coarse, (d) RBF parametre değerleri için hata matrisleri

	KADIN	ERKEK
KADIN	25	1
ERKEK	9	0

(a)

	KADIN	ERKEK
KADIN	24	2
ERKEK	8	1

(b)

	KADIN	ERKEK
KADIN	26	0
ERKEK	9	0

(c)

	KADIN	ERKEK
KADIN	25	1
ERKEK	9	0

(d)

Tablo 4.20.'de imleç hareketlerinin 16 öznitelik veri kümesinden elde edilen sınıflandırma başarımları ve F-Skor değerleri verilmiştir.

Sınıflandırma doğruluk değerleri incelendiğinde 16 öznitelik veri kümesi için; lineer çekirdek fonksiyonu kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %69, test veri kümesinin %71 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 0,78 ve 0,83 olarak elde edilmiştir. Polinomial çekirdek fonksiyonu kullanarak yapılan cinsiyet sınıflandırmasında, Quadratic parametre değeri ile eğitim veri kümesinin %69, test veri kümesinin %65 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 0,78 ve 0,78 olarak elde edilmiş, Cubic parametre değeri ile eğitim veri kümesinin %98, test veri kümesinin %57 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 0,98 ve 0,71 olarak elde edilmiştir. Gaussian çekirdek fonksiyonunun Fine parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %100, test veri kümesinin ise %74 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 1 ve 0,85 olarak elde edildiği, Medium parametre değeri

Tablo 4.20. 16 öznitelik veri kümesi sınıflandırma başarımları ve F-Skor değerleri

Öznitelik Sayı/Tür		16 / Hedfleme + Kontrol			
Çekirdek Fonksiyonu		Eğitim		Test	
		Doğruluk (%)	F-Skor	Doğruluk (%)	F Skor
Linear Çekirdek		69,7	0,78	71,4	0,83
Polynomial Çekirdek	Quadratic	69,0	0,78	65,7	0,78
	Cubic	98,5	0,98	57,1	0,71
Gaussian Çekirdek	Fine	100	1	74,2	0,85
	Medium	84,5	0,88	68,5	0,81
	Coarse	62,6	0,76	74,2	0,85
	Radyal Çekirdek (RBF)	100	1	74,2	0,85

kullanılarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %84, test veri kümesinin %68 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 0,88 ve 0,81 olarak elde edildiği, Coarse parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %62, test veri kümesinin %74 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 0,76 ve 0,85 olarak elde edildiği, radyal tabanlı parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %100, test veri kümesinin ise %74 doğrulukla yapıldığı, F-Skor değerlerinin ise sırasıyla 1 ve 0,85 olarak elde edilmiştir. Tablo 4.21.'de 8 hedefleme öznitelik veri kümesinden elde edilen sınıflandırma doğruluk değerleri ve F-Skor değerleri verilmiştir.

Bu değerler incelendiğinde; lineer çekirdek fonksiyonu kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %66, test veri kümesinin ise %60 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,78 ve 0,74 olarak elde edilmiştir. Polinomial çekirdek fonksiyonu kullanarak yapılan cinsiyet sınıflandırmasında Quadratic parametre değeri ile eğitim veri kümesinin %77, test veri kümesinin ise %42 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,84 ve 0,56 olarak elde edilmiş, Cubic parametre değeri ile eğitim veri kümesinin %94, test veri kümesinin ise %40 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,95 ve 0,51 olarak elde edilmiştir. Gaussian çekirdek fonksiyonun Fine parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %95, test veri kümesinin ise %57

Tablo 4.21. 8 Hedefleme öznitelik sınıflandırma başarımları ve F-Skor değerleri

Öznitelik Sayı/Tür		8 / Hedefleme			
Çekirdek Fonksiyonu		Eğitim		Test	
		Doğruluk (%)	F-Skor	Doğruluk (%)	F-Skor
Linear Çekirdek		66,1	0,78	60,0	0,74
Polynomial Çekirdek	Quadratic	77,4	0,84	42,8	0,56
	Cubic	94,3	0,95	40,0	0,51
Gaussian Çekirdek	Fine	95,7	0,96	57,1	0,72
	Medium	69,7	0,80	60,0	0,74
	Coarse	65,4	0,78	57,1	0,72
	Radyal Çekirdek (RBF)	92,2	0,94	54,2	0,70

doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,96 ve 0,72 olarak elde edilmiş, Medium parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %69, test veri kümesinin ise %60 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,80 ve 0,74 olarak elde edilmiş, Coarse parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %65, test veri kümesinin ise %57 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,78 ve 0,72 olarak elde edilmiş, radyal tabanlı parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %92, test veri kümesinin ise %54 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,94 ve 0,70 olarak elde edilmiştir.

Tablo 4.22.'de 8 kontrol öznitelik veri kümesinden elde edilen için sınıflandırma doğruluk değerleri ve F-Skor değerleri verilmiştir.

Bu değerler incelendiğinde; lineer çekirdek fonksiyonu kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %59, test veri kümesinin ise %74 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,74 ve 0,85 olarak elde edilmiştir. Polinomiyal çekirdek fonksiyonu kullanarak yapılan cinsiyet sınıflandırmasında Quadratic parametre değeri ile eğitim veri kümesinin %76, test veri kümesinin ise %65 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,82 ve 0,76 olarak elde edilmiş, Cubic parametre değeri ile eğitim veri kümesinin %92, test veri kümesinin ise %48 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,94 ve 0,60 olarak elde edilmiştir. Gaussian çekirdek fonksiyonunun Fine parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %100, test veri kümesinin ise %71

Tablo 4.22. 8 Kontrol öznitelik sınıflandırma başarımları ve F-Skor değerleri

Öznitelik Sayı/Tür		8 / Kontrol			
Çekirdek Fonksiyonu		Eğitim (%)		Test (%)	
		Doğruluk (%)	F-Skor	Doğruluk (%)	F-Skor
Lineer Çekirdek		59,8	0,74	74,2	0,85
Polynomial Çekirdek	Quadratic	76,7	0,82	65,7	0,76
	Cubic	92,9	0,94	48,5	0,60
Gaussian Çekirdek	Fine	100	1	71,4	0,83
	Medium	69,0	0,78	71,4	0,82
	Coarse	59,8	0,74	74,2	0,85
	Radyal Çekirdek (RBF)	98,5	0,98	71,4	0,83

doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 1 ve 0,83 olarak elde edilmiş, Medium parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %69, test veri kümesinin ise %71 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,78 ve 0,82 olarak elde edilmiş, Coarse parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %59, test veri kümesinin ise %74 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,74 ve 0,85 olarak elde edilmiş, radyal tabanlı parametre değeri kullanarak yapılan cinsiyet sınıflandırmasında eğitim veri kümesinin %98, test veri kümesinin ise %71 doğrulukla yapılmış, F-Skor değerlerinin ise sırasıyla 0,98 ve 0,83 olarak elde edilmiştir.

Tüm sınıflandırma başarımları değerleri incelendiğinde, hedefleme öznitelik veri kümesi için, Gaussian çekirdek fonksiyonun Fine ve RBF parametre değerlerinin sınıflandırma performansının başarılı olduğu, kontrol öznitelik veri kümesi için, Gaussian çekirdek fonksiyonun Fine ve RBF parametre değerlerinin sınıflandırma performansının başarılı olduğu, hedefleme ve kontrol öznitelik veri kümesi için, Gaussian çekirdek fonksiyonun RBF parametre değerlerinin sınıflandırma performansının başarılı olduğu değerlendirilmiştir. Bu bilgiler ışığında imleç hareketlerinden cinsiyet tespiti için destek vektörleri sınıflandırma yönteminin Gaussian çekirdek fonksiyonun RBF parametre değerinin kullanılması önerilmektedir.

5. SONUÇLAR VE ÖNERİLER

Bu konudaki çalışmalar, kullanıcıların cinsiyet farklılıklarını belirlemek için imleç hareket özelliklerinin önemli bir faktör olduğunu göstermiştir. Bilgisayar güvenlik araçlarının geliştirilmesi, kimlik doğrulama prosedürlerini bu farklılıklara odaklayarak bu doğal farklılıklardan yararlanabilir. Bu çalışmalar ışığında cinsiyet, online ortamlarda kötü niyetli amaçları engellemek için kişisel tanımlama olarak kullanılan temel özelliklerden birisi olduğundan özellikle imleç hareketleri ile cinsiyet tespiti, veri kümesinin arıtılması doğrultusunda daha yüksek sınıflandırma performansı ile bilişim sistemlerinde kişisel güvenlik sistemlerine yardımcı bir sistem olarak kullanılabilir. Kullanıcıların imleç hareketlerinde belirgin ve güçlü cinsiyet farklılıkları ve imlecin hedeflemesi ve kontrol edilmesinde önemli cinsiyet farklılıklarının tespit edildiğinden dolayı fare-imleç hareketleri, katılımcının cinsiyetini ayırt etmek için kullanılan bir hareket imzası şeklinde tanımlanabilir. Kadın ve erkeklerin fare hareketleri doğal olarak farklılık gösterdiğinden dolayı imleç hareketlerinin özneliklerinden olan hedefleme ve kontrol özelliklerini kullanarak, bilgisayar kullanıcılarının yalnızca fare hareketlerine dayanarak cinsiyet tespitinin makine öğrenimi algoritmaları ile yüksek doğrulukla sınıflandırılmasını yapılabilmektedir.

İmleç hareketlerine dayalı cinsiyet sınıflandırmasında Destek Vektör Makineleri kullanılmıştır. Sınıflandırma farklı çekirdek fonksiyonları altında gerçekleştirilmiştir. Deneysel sonuçlara göre, eğitim kümesi için elde edilen sınıflandırma doğruluğu, test kümesi için elde edilen sınıflandırma doğruluğuna göre yüksektir. Bu durum, modelin aşırı öğrenme evresine girmesi olarak yorumlanabilmektedir. Veri kümesindeki gözlem sayısının az olması, aşırı öğrenmeyle doğrudan ilişkili olup, gözlem sayısının artırılması durumunda, daha doğru yüksek sınıflandırma sonuçlarına ulaşılabileceği öngörülmektedir.

KAYNAKLAR

- [1] Yamauchi T., Seo J. H., Jett N., Parks G., Bowman C., Gender Differences in Mouse and Cursor Movements, *International Journal of Human-Computer Interaction*, 2015, **31**(12), 911-921.
- [2] Pentel A., Predicting Age and Gender by Keystroke Dynamics and Mouse Patterns, *25th Conference on User Modeling, Adaptation and Personalization*, Bratislava, Slovakia, 9-12 July 2017.
- [3] Kolakowska A., Landowska A., Jarmolkowicz P., Jarmolkowicz M., Sobota K., Automatic Recognition of Males and Females Among Web Browser Users Based on Behavioural Patterns of Peripherals Usage, *Internet Research*, 2016, **26**(5), 1093-1111.
- [4] Kolakowska A., A Review Of Emotion Recognition Methods Based on Keystroke Dynamics and Mouse Movements, *6th International Conference on Human System Interactions (HSI)*, Sopot, Poland, 6-8 June 2013.
- [5] Yamauchi T., Xiao K., Reading Emotion From Mouse Cursor Motions: Affective Computing Approach, *Cognitive Science*, 2018, **42**(3), 771-819.
- [6] Schoemann M., O'Hora D., Dale R., Scherbaum S., Using Mouse Cursor Tracking to Investigate Online Cognition: Preserving Methodological Ingenuity While Moving Towards Reproducible Science, *PsyArXiv*, DOI: 10.31234/osf.io/4ku26.
- [7] Yamauchi T., Leontyev A., Razavi M., Assessing Emotion by Mouse-cursor Tracking: Theoretical and Empirical Rationales, *8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Cambridge, United Kingdom, 3-6 September 2019.
- [8] Van Balen N., Ball C. T., Wang H., A Behavioral Biometrics Based Approach to Online Gender Classification, Security and Privacy in Communication Networks, *Springer International Publishing*, 2017, **198**, 475-495.
- [9] Kratky P., Chuda D., Estimating Gender and Age of Web Page Visitors from the Way They Use Their Mouse, *25th International Conference Companion on World Wide Web - WWW '16 Companion*, Montreal, Quebec, Canada, 11-15 April 2016.
- [10] Salmeron-Majadas S., Baker R. S., Santos O. C., Boticario J. G., A Machine Learning Approach to Leverage Individual Keyboard and Mouse Interaction Behavior From Multiple Users in Real-World Learning Scenarios, *IEEE Access*, 2018, **6**, 39154-39179.

- [11] Cetintas S., Si L., Ping Xin Y. P., Hord C., Automatic Detection of Off-Task Behaviors in Intelligent Tutoring Systems with Machine Learning Techniques, *IEEE Trans. Learn. Technol.*, 2010, **3**(3), 228-236.
- [12] Pusara M., Brodley C. E., User Re-Authentication Via Mouse Movements, *Proceedings of the 2004 ACM workshop on Visualization and Data Mining for Computer Security - VizSEC/DMSEC '04*, Washington DC, USA, 29 October 2004.
- [13] Gunetti D., Picardi C., Keystroke Analysis Of Free Text, *ACM Trans. Inf. Syst. Secur.*, 2005, **8**(3), 312-347.
- [14] Clarke N. L., Furnell S. M., Authenticating Mobile Phone Users Using Keystroke Analysis, *International Journal of Information Security*, 2006, **6**(1), 1-14.
- [15] Vizer L.M., Zhou L., Sears A., Automated Stress Detection Using Keystroke And Linguistic Features, *International Journal of Human-Computer Studies*, 2009, **67**(10), 870-886.
- [16] Epp C., Lippold M., Mandryk R. L., Identifying Emotional States Using Keystroke Dynamics, *CHI Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada, 7-12 May 2011.
- [17] Khanna P., Sasikumar M., Recognising Emotions from Keyboard Stroke Pattern, *International Journal of Computer Applications*, 2010, **11**(9), 1-5.
- [18] Yamauchi T., Bowman C., Mining Cursor Motions to Find the Gender, Experience, and Feelings of Computer Users, *2014 IEEE International Conference on Data Mining Workshop*, Shenzhen, China, 14 December 2014.
- [19] Gupta S., Web Content Recommendation Using Machine Learning on User Mouse Tracking Data, Master's Thesis, Master of Science in Computer Science, University of Oxford, Pembroke College- Computing Laboratory, Oxford, 2006.
- [20] Horwitz R., Brockhaus S., Henninger F., Kieslich P.J., Schierholz M., Keusch F., Kreuter F., Learning from Mouse Movements: Improving Questionnaires and Respondents' User Experience Through Passive Data Collection, *IAB-Discussion Paper*, 2017, **34**(16), 27.
- [21] Mevissen K., Using Mouse-Overs to Predict Customer Conversion Probability, Master's Thesis, Tilburg University, Department of Communication and Information Sciences, Tilburg, Netherlands, 2017.
- [22] Raj Kannan J., Sabitha R., Karthik S., Shanthini J., Mouse Movement Pattern Based Analysis of Customer Behavior (CBA-MMP) Using Cloud Data Analytics, *Wireless Personal Communications*, 2020, **110**(2) DOI: 10.1007/s11277-020-07055-1.

- [23] Ahmed A. A. E., Traore I., A New Biometric Technology Based on Mouse Dynamics, *IEEE Transactions on Dependable and Secure Computing*, 2007, **4**(3), 165-179.
- [24] Muda R., Hamid N. A., Satar S. D. M., Mohamad M., Mahadi N. A., Ghazali F., Mouse Movement Behavioral Biometric for Static User Authentication, *Advanced Science Letters*, 2017, **23**(6), 5050-5053.
- [25] Almalki S., Chatterjee P., Roy K., Continuous Authentication Using Mouse Clickstream Data Analysis, *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, 2019, **11637**, 76-85.
- [26] Zheng N., Paloski A., Wang H., An Efficient User Verification System via Mouse Movements, *18th ACM Conference on Computer and Communications Security - CCS '11*, Chicago Illinois USA ,17-21 October 2011.
- [27] Sayed B., Traore I., Woungang I., M. Obaidat S., Biometric Authentication Using Mouse Gesture Dynamics, *IEEE Systems Journal*, 2013, **7**(2), 262-274.
- [28] Oliveira P.X., Sinha B., Londhe T., Bakelman N., Monaco J., Tappert C., Mouse Movement Biometric System, *Proceedings of Student-Faculty Research Day*, Seidenberg School of CSIS, Pace University, White Plains NY, 4 May 2013.
- [29] Jorgensen Z., Yu T., On Mouse Dynamics as a Behavioral Biometric for Authentication, *6th ACM Symposium on Information, Computer and Communications Security- ASIACCS '11*, Hong Kong, China, 22-24 March 2011.
- [30] Muthumari G., Shenbagaraj R., Blessa Binolin Pepsi M., Mouse Gesture Based Authentication Using Machine Learning Algorithm, *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies*, Ramanathapuram, India, 8-10 May 2014.
- [31] Hu T., Niu W., Zhang X., Liu X., Lu J., Liu Y., An Insider Threat Detection Approach Based on Mouse Dynamics and Deep Learning, *Security and Communication Networks*, 2019, **2019**, 1-12.
- [32] Karim M., Heickal H., Hasanuzzaman Md., User Authentication from Mouse Movement Data Using Multiple Classifiers, *9th International Conference on Machine Learning and Computing- ICMLC 2017*, Singapore, Singapore, 24 – 26 February 2017.
- [33] Antal M., Egyed-Zsigmond E., Intrusion Detection Using Mouse Dynamics, *IET Biometrics*, 2019, **8**(5), 285-294.
- [34] Nazar A., Traoré I., A. Ahmed A. E., Inverse Biometrics for Mouse Dynamics, *International Journal of Pattern Recognition and Artificial Intelligence*, 2008, **22**(03), 461-495.

- [35] Lee K., Esposito C., Lee S. Y., Vulnerability Analysis Challenges of the Mouse Data Based on Machine Learning for Image-Based User Authentication, *IEEE Access*, 2019, **7**, 177241-177253.
- [36] Lee K., Lee S. Y., Improved Practical Vulnerability Analysis of Mouse Data According to Offensive Security Based on Machine Learning in Image-Based User Authentication, *Entropy*, 2020, **22**(3), 355.
- [37] Yu H., Xiao S., Yu Z., Li Y., Zhang Y., ImCAPTCHA: Imperceptible CAPTCHA Based on Cursor Trajectories, *IEEE Consumer Electronics Magazine*, 2020, **9**(1), 74-82.
- [38] Shrestha P., Saxena N., Neupane A., Satvat K., CATCHA: When Cats Track Your Movements Online, *Information Security Practice and Experience*, 2019, **11879**, 172-193.
- [39] Accot J., Zhai S., Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1997)*, Atlanta Georgia, USA, 22-27 March 1997.
- [40] Accot J., Zhai S., Performance Evaluation of Input Devices in Trajectory-Based Tasks: An Application Of The Steering Law, In M. G. Williams & M. W. Altom (Ed.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1999)*, Pittsburgh Pennsylvania, USA, 15-20 May 1999.
- [41] Card S. K., English W. K., Burr B. J., Evaluation of Mouse, Rate-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT., *Ergonomics*, 1978, **21**(8), 601-613.
- [42] MacKenzie I. S., Kauppinen T., Silfverberg M., Accuracy Measures For Evaluating Computer Pointing Devices. In M. Tremaine (Ed.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Seattle Washington, USA, 31 March - 5 April 2001.
- [43] Hertzum M., Hornbaek K., How Age Affects Pointing with Mouse and Touchpad: A Comparison of Young Adult, and Elderly Users, *International Journal of Human-Computer Interaction*, 2010, **26**(7), 703-734.
- [44] Tsimperidis I., Keystroke Dynamics Features for Gender Recognition, *Digital Investigation*, 2018, **24**, 4-10.
- [45] Van Balen N., Ball C., Wang H., Analysis of Targeted Mouse Movements for Gender Classification, *ICST Trans Security Safety*, 2017, **4**(11), 153.
- [46] Leontyev A., Yamauchi T., Razavi M., Machine Learning Stop Signal Test (ML-SST): ML-based Mouse Tracking Enhances Adult ADHD Diagnosis, *8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, Cambridge, United Kingdom, 3-6 September. 2019.

- [47] Leontyev A., Sun S., Wolfe M., Yamauchi T., Augmented Go/No-Go Task: Mouse Cursor Motion Measures Improve ADHD Symptom Assessment in Healthy College Students, *Frontiers in Psychology*, 2018, **9**, 496.
- [48] Leontyev A., Yamauchi T., Mouse Movement Measures Enhance the Stop-Signal Task in Adult ADHD Assessment, *PLOS ONE*, 2019, **14**, 11.
- [49] Xiao K., Yamauchi T., The Role of Attention in Subliminal Semantic Processing: A Mouse Tracking Study, *PLOS ONE*, 2017, **12**, 6.
- [50] Schoemann M., Lüken M., Grage T., Kieslich P. J., Scherbaum S., Validating Mouse-Tracking: How Design Factors Influence Action Dynamics in Intertemporal Decision Making, *Behavior Research Methods*, 2019, **51**, 2356-2377.
- [51] Diego-Mas J. A., Garzon-Leal D., Poveda-Bautista R., Alcaide-Marzal J., User-Interfaces Layout Optimization Using Eye-Tracking, Mouse Movements and Genetic Algorithms, *Applied Ergonomics*, 2019, **78**, 197-209.
- [52] Sartori G., Zangrossi A., Monaro M., Deception Detection With Behavioral Methods: The Autobiographical Implicit Association Test, Concealed Information Test-Reaction Time, Mouse Dynamics and Keystroke Dynamics, Editor: Rosenfeld J. P., *Detecting Concealed Information and Deception: Recent Developments*, Elsevier Academic Press, 215-241, 2018.
- [53] Miguel-Hurtado O., Stevenage S. V., Bevan C., Guest R., Predicting Sex as A Soft-Biometrics from Device Interaction Swipe Gestures, *Pattern Recognition Letters*, 2016, **79**, 44-51.
- [54] Yamauchi T., Seo H., Choe Y., Bowman C., Xiao K., Assessing Emotions by Cursor Motions: An Affective Computing Approach, *The 37th Annual Conference of the Cognitive Science Society*, Pasadena, California, 23-25 July 2015.
- [55] Mueller A. C., Guido S., Guido S., Introduction to Machine Learning with Python, First Edition, O'Reilly Media, Sebastopol, 2017.
- [56] Mohri M., Rostamizadeh A., Talwalkar A., Foundations Of Machine Learning, Second edition, The MIT Press, Cambridge, 2018.
- [57] Mohammed M., Khan M., Bashier E., Machine Learning Algorithms and Applications, 1st Edition, CRC Press, Boca Raton, 2016.
- [58] Nikki Castle, What is Semi-Supervised Learning?, Oracle AI and Data Science Blog, <https://www.datascience.com/blog/what-is-semi-supervised-learning> (Ziyaret tarihi: 22 Haziran 2020).
- [59] prowmesadmin, Makine Öğrenmesi, Prowmes, <http://www.prowmes.com/blog/makine-ogrenmesi/> (Ziyaret tarihi: 22 Haziran 2020).

- [60] Thulin J., Machine Learning-Based Classifiers In The Direkt Profil Grammatical Profiling System, Master's Thesis, Lund University, Lund Institute of Technology, Lund, Sweden, 2007.
- [61] Luckert M., Using Machine Learning Methods for Evaluating the Quality of Technical Documents, Master's Thesis, Linnaeus University, Department of Computer Science and Media Technology, Kalmar, Sweden, 2016.
- [62] Li H., Which Machine Learning Algorithm Should I Use?, The SAS Data ScienceBlog,
<https://blogs.sas.com/content/subconsciousmusings/2017/04/12/machine-learning-algorithm-use/> (Ziyaret tarihi: 22 Haziran 2020).
- [63] Ögündür G., Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall) yada F1Score?,
<https://medium.com/@gulcanogundur/do%C4%9Fruluk-accuracy-kesinlik-precision-duyarl%C4%B1%C4%B1k-recall-ya-da-f1-score-300c925feb38> (Ziyaret tarihi:01 Eylül 2020).
- [64] Sasaki Y., The Truth of the F-measure, <https://www.toyota.ac.jp/Lab/Denshi/COIN/people/yutaka.sasaki/F-measure-YS-26Oct07.pdf> (Ziyaret tarihi: 01 Eylül 2020).
- [65] Derczynski L., F-score, and NLP Evaluation, *The Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Portorož, Slovenia, 23-28 May 2016.

KİŞİSEL YAYINLAR VE ESERLER

- [1] **Sayılır E.**, Akbulut O., İmleç Hareketlerinden Kullanıcı Cinsiyetinin Makine Öğrenmesi Tabanlı Tespiti, *Uluslararası Marmara Fen ve Sosyal Bilimler Kongresi*, Karatepe, Kocaeli, 01-03 Kasım 2019.



ÖZGEÇMİŞ

İlk, orta ve lise eğitimini Gaziantep’te tamamladı. 2001 yılında Gaziantep Meslek Yüksek Okulundan mezun oldu. Bitirme projesi olarak Delphi programlama dili ile robot kolu kontrolü projesini gerçekleştirdi. 2008 yılında Girne Amerikan Üniversitesi, Bilgisayar Mühendisliği bölümünden mezun oldu. Mezuniyet tezi olarak Ses tanıma (Speech recognition) projesini başarıyla tamamladı. 2010 yılında Anadolu Üniversitesi Kamu yönetiminden mezun oldu. 2008 yılından itibaren yazılım mühendisi olarak birçok firmada çalıştı. 2018 yılında Kocaeli Üniversitesi, Bilgisayar Mühendisliği bölümünde Yüksek Lisans eğitimine başladı.

