



**KTO KARATAY ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ELEKTRİK VE BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
TEZLİ YÜKSEK LİSANS PROGRAMI**

**EL YAZISI ANALİZİ İLE İNSANLARIN KİŞİLİK ÖZELLİKLERİNİN TESPİT
EDİLMESİ**

Hilal MÜSEVİTOĞLU

Yüksek Lisans Tezi

**KONYA
Mayıs 2022**

EL YAZISI ANALİZİ İLE İNSANLARIN KİŞİLİK ÖZELLİKLERİNİN TESPİT
EDİLMESİ

Hilal MÜSEVİTOĞLU

KTO Karatay Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik ve Bilgisayar Mühendisliği Anabilim Dalı
Tezli Yüksek Lisans Programı

Yüksek Lisans Tezi

Tez Danışmanı: Dr. Öğr. Üyesi Ali ÖZTÜRK

Konya
Mayıs 2022

BİLDİRİM

Enstitü tarafından onaylanan Yüksek Lisans/Doktora tezimin tamamını veya herhangi bir kısmını basılı veya dijital biçimde arşivleme ve aşağıda belirtilen koşullar dahilinde erişime açma iznini KTO Karatay Üniversitesine verdiğimi bildiririm. Bu izinle, Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak ve gelecekteki çalışmalar (makale, kitap, lisans, patent vb.) için tezimin tamamının veya bir bölümünün kullanım hakları yalnızca bana ait olacaktır.

Tezimin bütünüyle kendi çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Telif hakkı bulunan ve sahiplerinden yazılı izinle kullanılması zorunlu olan kaynakları, yazılı izin alarak kullandığımı ve istenildiğinde izinlerin suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayımlanan “Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge” kapsamında, tezim, aşağıda belirtilen koşullar haricince, YÖK Ulusal Tez Merkezi ve KTO Karatay Üniversitesi Açık Erişim Sisteminde erişime açılır.

Enstitü / Fakülte Yönetim Kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.¹

Enstitü / Fakülte Yönetim Kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir.²

Tezimle ilgili gizlilik kararı verilmiştir.³⁴

16 Mayıs 2022

Hilal MÜSEVİTOĞLU

¹ MADDE 6(1) Lisansüstü tezle ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.

² MADDE 6(2) Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ay aşmamak üzere tezin erişime açılması engellenebilir.

³ MADDE 7(1) Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

⁴ MADDE 7(2) Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

ETİK BEYAN

KTO Karatay Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Hazırlama ve Yazım Kurallarına uygun olarak Dr. Öğr. Üyesi Ali ÖZTÜRK danışmanlığında tarafımdan üretilen bu tez çalışmasında; sunduğum tüm veri, enformasyon, bilgi ve belgeleri bilimsel etik kuralları çerçevesinde elde ettiğimi, tüm değerlendirme, analiz, bulgu ve sonuçları bilimsel usullere uygun olarak sunduğumu, tez çalışmasında yararlandığım kaynakların tümüne bilimsel normlara uygun biçimde atıfta bulunarak kaynak gösterdiğimi, tezimin/projemin kaynak gösterilen durumlar dışında özgün olduğunu bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

16 Mayıs 2022

Hilal MÜSEVİTOĞLU

TEŐEKKÜR

Tez alıőmamın tım aőamalarında, kıymetli bilgi, birikim ve tecrübeleri ile bana yol gösteren deęerli danıőman Hocam Sayın Dr. Öğr. Üyesi Ali Öztürk'e, tım süreç boyunca hem uzmanlığıyla hem de tım yardımseverliğiyle yanımda olan sevgili Uzman Psikolog Özlem Konađ' a ve baőta annem, babam ve eőim olmak üzere hep yanımda olduklarını hissettiren sevgili aileme ve arkadaşlarıma içten teşekkürlerimi sunarım.

16 Mayıs 2022

Hilal MÜSEVİTOĐLU

ÖZET

Hilal MÜSEVİTOĞLU

El Yazısı Analizi ile İnsanların Kişilik Özelliklerinin Tespit Edilmesi

Yüksek Lisans

Konya, 2022

Teknoloji her geçen gün hızla artıyor olsa da yazının insan hayatındaki yeri başkadır. Bilginin sonraki kuşaklara birikimli bir şekilde aktarılmasında yazının önemi büyüktür. Aynı zamanda duygu ve düşüncelerimizi de yazı ile kalıcı hale getirmek mümkündür. El yazısı ise daha derin bir anlam içermektedir. El yazısı, yazan kişi hakkında birçok bilgiyi barındırır. Beyindeki nörolojik desenler tarafından temsil edilen kişilik özelliklerinin işaretidir el yazısı. Diğer bir deyişle beynimiz ve bilinçaltımız aslında alışkanlıklarımızın bir sonucu olarak karakterimizi biçimlendirmektedir. El yazısı incelenerek bireyin içinde bulunduğu ruh hali hakkında bir fikre varmak mümkündür. Sevinç, hüznün, öfke ve kaygı bunlardan bazılarıdır.

Bu çalışma ile farklı meslek ve yaş gruplarındaki kişilerin kâğıda döktükleri yazılardan veri kümesi oluşturularak makine öğrenmesi algoritmalarına bu veri kümesi uygulanmıştır. Elde edilen kişilik analizi sonuçları ile uzman psikolog tarafından temin edilen kişilik testinden elde edilen sonuçların bir karşılaştırılması yapılmıştır. Bunun sonucunda makine öğrenmesi algoritmalarının (NB, Karar Ağaçları, Rastgele Orman, DVM) performansları karşılaştırılarak kişilik analizi için en iyi sonucu veren algoritma tespit edilmiştir. Doğruluk, kesinlik, hatırlama ve f1 skoru ölçütleri için tüm el yazısı özelliklerine bakıldığında ortalama olarak en başarılı algoritmaların Rastgele Orman ve Karar Ağacı algoritmaları olduğu görülmektedir. Bu algoritmalara en yakın başarı oranına sahip algoritmalar ise sırasıyla Lojistik Regresyon ve Destek Vektör Makineleridir. Naive Bayes algoritmasının başarı oranı ise diğerlerine göre daha düşük olarak belirlenmiştir.

Uzman psikologun uyguladığı kişilik testinden elde edilen sonuçlar ile el yazısı analizinden elde edilen sonuçlar %72 oranında eşleşmiştir.

Anahtar Kelimeler

El Yazısı, El Yazısı Sayısal İnceleme, Kişilik Analizi, Uzman Sistem, Grafoloji

ABSTRACT

Hilal MÜSEVİTOĞLU

Identifying People's Personality Traits With Handwriting Analysis

Master's

Konya, 2022

Although technology is rapidly increasing day by day, the place of manuscript in human life is different. The importance of writing is great in conveying knowledge to the next generations in a cumulative way. At the same time, it is possible to make our feelings and thoughts permanent with writing. Handwriting has a deeper meaning. The handwriting contains a lot of information about the writer. It is an indication of the personality traits represented by the neurological patterns in his brain. In other words, our brains and consciousness actually shape our character as a result of our habits. By examining the handwriting it is possible to get an idea of the mood that the individual is in. Joy, sadness, anger, anxiety are some of them.

In this study, a dataset was created from the writings of people in different professions and age groups, and this dataset was applied to machine learning algorithms. A comparison was made between the results of the personality analysis and the results obtained from the attached personality test. As a result, the performances of machine learning algorithms (NB, Decision Trees, Random Forest, SVM) were compared and the algorithm that gave the best result for personality analysis was determined. When looking at all handwriting features for accuracy, precision, recall and f1 score criteria, it is seen that the most successful algorithms on average are Random Forest and Decision Tree algorithms. The algorithms with the closest success rate to these algorithms are Logistic Regression and Support Vector Machines, respectively. The success rate of the Naive Bayes algorithm was determined to be lower than the others.

The results obtained from the personality test administered by the expert psychologist and the results obtained from the handwriting analysis matched at a rate of 72%.

Keywords

Expert System, Handwriting, Handwriting Numerical Review, Personality Analysis, Graphology

İÇİNDEKİLER

BİLDİRİM	i
ETİK BEYAN.....	ii
TEŞEKKÜR.....	iii
ÖZET	v
ABSTRACT.....	vi
İÇİNDEKİLER	vii
TABLolar DİZİNİ	ix
ŞEKİLLER DİZİNİ.....	x
SİMGELER DİZİNİ	xi
KISALTMALAR DİZİNİ.....	xii
1. GİRİŞ	1
2. İLGİLİ ÇALIŞMALAR.....	4
3. MATERYAL ve YÖNTEMLER.....	8
3.1. Kullanılan Algoritmalar	8
3.1.1. Naive Bayes (NB) Sınıflandırıcı.....	8
3.1.2. Karar Ağaçları	9
3.1.3. Rastgele Orman	10
3.1.4. Destek Vektör Makineleri (DVM).....	11
3.1.5. Lojistik Regresyon.....	12
3.2. Kişilik Testi	14
3.3. Kullanılan Veri Kümesi.....	15
3.4. Yöntemler.....	16
3.5. Metrikler.....	19
4. DENEYSEL SONUÇLAR	20
4.1. Sol, sağ, üst ve alt marjinlerin çıkarılması	20
4.2. Marjinlerden elde edilen sayfanın genel biçimi	21
4.3. Yazının genel eğimi.....	21
4.4. Kelimeler arası ortalama boşluk.....	22
4.5. Satırlar arası ortalama boşluk	23
4.6. Yazı basıssı	24
5. SONUÇ	30
KAYNAKLAR	31

ÖZGEÇMİŞ	33
EK 1. PSİKOLOG KİŞİLİK TESTİ	34
EK 2. KULLANILAN KODLAR.....	36
ETİK KURUL.....	43

TABLULAR DİZİNİ

Tablo 1. Karmaşıklık matrisi.....	19
Tablo 2. Elde edilen veri kümesine ait giriş özellikleri	24
Tablo 3. Kişilik analizi sonucu elde edilecek çıkış değerleri.....	27
Tablo 4. Rastgele Orman algoritması için elde edilen sonuçlar.....	27
Tablo 5. Lojistik Regresyon algoritması için elde edilen sonuçlar	27
Tablo 6. Naive Bayes algoritması için elde edilen sonuçlar	28
Tablo 7. Karar Ağacı algoritması için elde edilen sonuçlar.....	28
Tablo 8. Destek Vektör Makineleri algoritması için elde edilen sonuçlar.....	29

ŞEKİLLER DİZİNİ

Şekil 1. Eski el yazısı örneği.....	2
Şekil 2. Naive Bayes (Taheri S. ve Mammadov M., 2013).....	9
Şekil 3. Rastgele orman algoritmasının şeması (Zhang vd., 2018).....	10
Şekil 4. İki boyutlu uzay için marjin ve destek vektörlerinin gösterimi (Support Vector Machines).....	11
Şekil 5. Lojistik fonksiyon kullanılarak elde edilen sayıların grafiği (Machine Learning Mastery).....	12
Şekil 6. Softmax Aktivasyon Fonksiyonu.....	13
Şekil 7. Kişilerden alınan el yazısı örneği.....	15
Şekil 8. Sağ ve sol kenar boşlukları gösterimi.....	16
Şekil 9. Alt ve üst kenar boşlukları gösterimi.....	16
Şekil 10. Satır arası boşlukların gösterimi.....	17
Şekil 11. Kelimeler arası boşlukların gösterimi.....	17
Şekil 12. (a) Basısı az olan (b) Basısı çok olan el yazısı örneği.....	18
Şekil 13. Yazıdaki eğimin gösterimi.....	18
Şekil 14. Marjin çizgilerinin gösterimi.....	21
Şekil 15. Tespit edilen kelimelerin gösterimi.....	22
Şekil 16. Tespit edilen satırların gösterimi.....	23

SİMGELER DİZİNİ

Simge	Açıklama
Σ	Sigma
p_i	i özelliğın olasılıđı
$P(y_i)$	y_i sınıfının olasılıđı
\in	Eleman
A	Özellik

KISALTMALAR DİZİNİ

Kısaltma	Açıklama
YSA	Yapay Sinir Ağı (Artificial Neural Network)
NB	Naive Bayes
RO	Rastgele Orman
DVM	Destek Vektör Makineleri
KA	Karar Ağaçları
SVM	Support Vector Machines

1. GİRİŞ

Hayatımızın bir parçası olan yazı sadece kağıda dökülenleri ifade etmekle kalmaz bir yandan da bizim iç duygularımız hakkında ipuçları barındırır. Teknolojinin hızlı bir şekilde ilerlediği bu dönemde el yazısının yerini yadsınamaz boyutta internet, medya, elektronik posta almış gibi görünse de el yazısı hala önemini ve varlığını korumaktadır. Herkes sosyal hayatında ya da iş hayatında teknolojinin yanı sıra el yazısı ile de bir takım işlemler yapmaktadır. Kalem ucunun kâğıda dokunduğu andan itibaren bırakılan izler, desenler, şekiller, semboller kişinin kimliğini tıpkı bir parmak izi veya DNA gibi eşsiz bir şekilde yansıtmaya yeteneğine sahiptir.

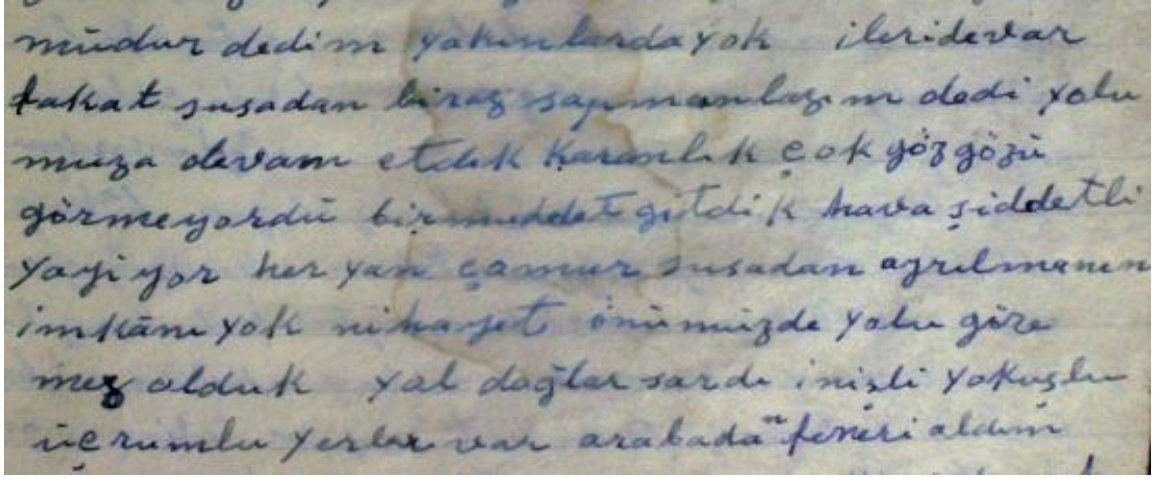
Yaklaşık yüz yıl önce Alman Profesör W. Preyer yazının belli kasların eğitilmesi ile değil, beyin tarafından oluşturulduğunu söylemiştir. Sonraki çalışmalar da yazıda en etkili organın beyin olduğunu ortaya koymuştur (Robertson, 1991). Buradan hareketle insan beynine ve dolayısıyla psikolojisine doğrudan erişmek için el yazısının etkin bir araç olduğu görülmektedir. Dolayısıyla el yazısından karakter analizinin, çok farklı alanlarda uygulanma potansiyeli vardır. Literatüre bakıldığında; adli soruşturmalar, insan kaynakları (işe alım süreci), psikolojik danışmanlık ve rehberlik hizmetleri, eğitim vb. alanlarda kullanıldığı görülmektedir.

El yazısı ile kişilik analizinin ilişkilendirildiği çalışmalar genel olarak Grafoloji bilimi altında incelenir. Grafoloji; bir kişinin el yazısından yola çıkarak, bu kimsenin kişiliği ve karakteri ile ilgili yapılan çıkarımları kapsayan çalışma alanıdır (Sheikholeslami vd., 1997). Karakter analiz etmek için el yazısından elde edilen bazı belirgin özelliklerin kullanıldığı bir bilim dalı olarak da tanımlanmaktadır. Grafoloji terimi, Yunanca "graphein" (yazma) ve "logos" (çalışma) kelimelerinin bir araya gelmesinden oluşmuştur.

Dil biliminde, grafoloji terimi bazen konuşma dilinin yazıya döküldüğü geleneksel yolların bilimsel çalışması olan grafemik ile eş anlamlı olarak kullanılır .

1622 gibi erken bir tarihte grafolojiden bahsedilmiş olsa da (Camilo Baldi, Bir Yazarın Niteliğini ve Niteliğini Mektuplarından Tanımak İçin Bir Yöntem Üzerine İnceleme) grafolojinin pratik kökenleri 19. yüzyılın ortalarında, Jacques-Hippolyte Michon (Fransa) ve Ludwig Klages'in (Almanya) çalışmalarına ve yazılarına dayanmaktadır. Aslında, The Practical System of Graphology (1871 ve yeniden basımlar) adlı kitabının başlığında

kullandığı “grafoloji” terimini kullanan Michon'dur. “Grafoanaliz” teriminin kökeni MN Bunker'a atfedilir. Şekil 1’de gerçek bir eski el yazısı örneği gösterilmiştir.



Şekil 1. Eski el yazısı örneği

Literatür incelendiğinde; genel olarak insan karakteri veya davranış tahmini ve kişilik analizi için el yazısından özellik çıkarımı için aşağıdaki yöntemler kullanılmıştır:

- 1- Yazının eğimi
- 2- Kalem basısı
- 3- Belli harflerinin karakteristiği
- 4- Sayfanın sağ, sol, üst ve alt kenar boşlukları
- 5- Satırlar arası boşluklar
- 6- Kelimeler arası boşluklar
- 7- Harflerin bitişikliği
- 8- Kelimelerin kapladığı dikdörtgen
- 9- Harflerin eğimi
- 10- Harflerin boyutu

Yukarıda bahsedilen özellikleri kullanarak kişilik analizini otomatik olarak yapmak için kullanılan belli başlı makine öğrenmesi algoritmaları ise aşağıdaki gibidir:

- 1- Destek Vektör Makineleri
- 2- Yapay Sinir Ağları

3- Evrişimsel Sinir Ağları (Derin Öğrenme)

4- Saklı Markov Modelleri

Bu çalışmada kullanılan el yazısı özellikleri ise; yazı eğimi, sayfanın sağ, sol, üst ve alt kenar boşlukları, satırlar arası boşluklar, kelimeler arası boşluklar ve kalem basısıdır. Bunlar, grafolojide en yaygın kullanılan kişilik analizi özellikleri olduğu için tercih edilmiştir. Kullanılan makine öğrenmesi algoritmaları olarak da; NB Sınıflandırıcı, Karar Ağaçları, Rastgele Orman, DVM ve Lojistik Regresyon olarak seçilmiştir.

2. İLGİLİ ÇALIŞMALAR

Champa ve Kumar, el yazısı analizinden otomatik insan davranış tahmini üzerine olan çalışmada; eğim, kalem basısı, y ve t harflerinin karakteristiği gibi özellikleri kullanmışlardır. El yazısı genellikle beyin yazısı olarak da adlandırılır. Her kişilik özelliği, nörolojik bir beyin modeli ile temsil edilir. Her nörolojik beyin modeli, o belirli kişilik özelliğine sahip her insan için aynı olan benzersiz bir nöromusküler hareket üretir. Yazarken, bu küçük hareketler bilinçsizce gerçekleşir. Her yazılı hareket veya vuruş, belirli bir kişilik özelliğini ortaya çıkarır. Grafoloji, bu vuruşları el yazısında gördükleri gibi tanımlama ve bunlara karşılık gelen kişilik özelliklerini belirleme bilimidir. Bu çalışmada; bir insanın kişiliğini, el yazısından çıkarılan özelliklerden tahmin etmek için bir yöntem önerilmiştir. Bir kişinin el yazısında bulunan taban çizgisi, kalem basısı, “t” harfi, “y” harfinin alt ilmeği ve yazının eğiminin ortaya koyduğu kişilik özellikleri incelenmiştir. Beş parametre, taban çizgisi, kalem basıncı ve “t” harfinin gövdesindeki t-barın yüksekliği, “y” harfinin alt döngüsü, yazının eğimi, kuralın girdileridir. Taban çizgisinin değerlendirilmesi poligonalizasyon yöntemini kullanır ve kalem basıncının değerlendirilmesi gri seviye eşik değerini kullanır. “t” harfinin gövdesindeki t çubuğunun yüksekliği, şablon eşleştirme kullanılarak hesaplanır. “y” harfinin alt döngüsünün şekli Genelleştirilmiş Hough Dönüşümü (GHT) kullanılarak hesaplanır ve yazı eğimi yine şablon eşleştirme kullanılarak hesaplanır. Bir el yazısından elde edilen bu parametreler, yazar hakkında birçok doğru bilgiyi ortaya çıkarır. MATLAB, amaç için kullanılan araçtır. Performans, birden fazla numune incelenerek ölçülür (Champa ve AnandaKumar, 2010).

Fatimah ve arkadaşları, evrimsel sinir ağları kullanarak el yazısından kişilik özelliklerini belirleme üzerine yaptıkları çalışmada; bir bireyin kişilik, yapı ve sembol özellik analizinin el yazısı görüntüsünden belirlemeyi önermişlerdir. Sembol analize dayalı olarak devam eden çoklu yapı analizi olan CNN yöntemi kullanılarak sınıflandırılmıştır. Kenar boşluğuna, satırlar arası boşluklara, kelimeler arası boşluklara, baskıya ve daha önceki araştırmalarda yapılmayan eğik veya eğim ve belirli harf özelliğine bakılmıştır (Fatimah vd., 2019).

Durga ve Deepu, grafoloji yöntemleri ile el yazısı analizi üzerine 1971 ile 2017 yılları arasında yapılmış çalışmalarını kapsamlı olarak inceleyip kullanılan yapay zekâ ve özellik

çıkarm yöntemlerini deęerlendirmişlerdir. Çalışma, bireysellik onayı, yazar özelliklerini öğrenmek için birkaç örneğin kullanıldığı ve öğrenilen modelleri test etmek için farklı örneklerin kullanıldığı bir makine öğrenimi yaklaşımı kullanılarak tamamlanmıştır (Durga ve Deepu, 2018).

2010 yılında Omar Santana ve arkadaşları tarafından yapılan çalışmada el yazısından kimlik analizi tahmini yapılmıştır. Bunun için 29 kişinin katıldığı bir data seti üzerinde girdi olarak “düşey konum”, “harflerin birleşikliği”, “bası kuvveti”, “inceltme alanı” ve “a harfi” olmak üzere 5 farklı grafolojik parametre kullanılmıştır. Yapay sinir ağıları (YSA) ve Destek Vektör Makinesi (DVM) algoritmalarını kullanmışlardır. Çalışma sonucunda 29 yazardan oluşan veri setinden %99.34 oranında başarı sağlanmıştır. Daha sonra veri seti genişletilerek 70’e çıkarıldığında ise elde edilen başarı oranı %92’dir (Santana vd., 2010).

“Grafometrik Parametrelere Dayalı Çevrimdışı Yazar Kimliğinin Belirlenmesi” adlı 2013 yılında gerçekleştirilen çalışmada yazı stilinden yazarın kimlik tespit tahmini yapılmıştır. Çalışmada, sinir ağıları tarafından bir karar füzyon bloęu ile sayısallaştırılan ve sınıflandırılan Grafoloji veya Adli tekniklere dayalı yeni bir özellik sunulmuştur. Analiz işlemi çevrim dışı bir şekilde el yazısı görüntüleri kullanılarak yapılmıştır. 100 adet örnekten oluşan veri seti üzerinde %94.6 başarı oranı elde edilmiştir (Vásquez vd., 2013).

Prasetyo ve arkadaşları tarafından 2017 yılında yapılan çalışmada grafoloji yöntemleri ile kimlik tespiti için kullanılacak bir mobil uygulama tasarlanmıştır. Uygulamada analiz için yazı yönü, yazının eğimi, yazı genişliği, yazı kenar boşluğu, sivri veya yuvarlak harf ve satır arası boşluk parametreleri kullanılmıştır. Uygulama kullanılarak gerçekleştirilen 25 adet el yazısının test edilmesiyle ulaşılan sonuç ile bir uzman tarafından deęerlendirilen performansın birebir aynı olduğu görülmüştür (Prasetyo vd., 2017).

“El Yazısı ve Bilgisayarlı Grafoloji Üzerine Kapsamlı Bir Anket” adlı makalede Afnan H. Garoot ve arkadaşları bilgisayarlı grafoloji sistemlerindeki en son gelişmeleri ele almışlardır. Bu araştırma makelesinde sayfa kenar boşluğu, satır taban çizgisi, satır ve sözcük aralığı, satır yönü, sözcük ve harflerin eğimi, harflerin boyutu, kalem basıncı, yazma hızı gibi bilgisayarlı grafolojide kullanılan önemli özelliklerin yanında “t” harfi, “i” harfi, “f” harfi, ve “y” harfinin incelenmesinin de yazı analizinde etkili olabileceęi söylenmiştir. Ayrıca grafolojide yaygın olarak destek vektör makinesi (DVM) ve yapay

sinir ağırları (YSA), gibi kural tabanlı farklı analiz yöntemlerinin sıkça kullanıldığından söz edilmiştir. Güçlü bir analiz yöntemi olan derin öğrenmenin bu alanda henüz uygulanmamış olduğu ve yüksek doğruluk oranına sahip tam otomatik bir grafoloji sistemi tasarlamak için derin öğrenme yönteminin uygulanmasının iyi sonuçlar verebileceği vurgusu yapılmıştır. Bunlara ek olarak eğitim verilerinin miktarının daha iyi sonuçlar elde etmede etkisi olduğundan bahsedilmiştir (Garoot vd., 2017).

Apoorva Anand ve arkadaşları 2018 yılında; grafoloji, yetenek testi ve kişilik testi kullanarak Otomatik Kariyer Rehberliği çalışması yapmıştır. Bu çalışma kişilerin çeşitli kariyer alanları ile tanışması ve uygun mesleği seçmesine yardımcı olmak için yapılmıştır. Çalışmada, insan davranışı ve analizinin temel yönleri dikkate alınmaktadır. Anahtar yönler, kişinin tutum ve yeteneğine karar vermeye yardımcı olan bilinçli ve bilinçaltı faktörleri içerir. Kullanılan analizler Yetenek Testi, Psikometrik Test (Myers-Briggs Tip Göstergesi - MBTI) ve el yazısı analizidir. Soru-cevap modülü bireyler tarafından verilen bilinçli cevaplarla kariyer ve kişilik verirken, el yazısı modülü kariyer ve kişiliğin bilinçaltı haritasını çıkarmaya odaklanır. Alınan el yazısı örneklerinden, kelimeler arası boşluk, sol marjın, sağ marjın, yazı basısı, harf eğimi, harf boyutu, “I” harfinin boyutu gibi özellikler hesaplanmıştır. Her üç modülün entegrasyonu kişilere uygun kariyer seçenekleri sunmaktadır (Anand vd., 2018).

Mathur ve Patil 2020 yılında, el yazısı ile kişilik analizi ve yazar tanımlama için makine öğrenimi yaklaşımı konulu çalışma gerçekleştirmişlerdir. Yaptıkları bu çalışma, çıkarılan özniteliklerden bir kişinin kişilik tahmini için bir yöntem önermektedir. Bunlar taban çizgisi, kelimeler arasındaki boşluk, sol kenar boşluğu boyutu, harf eğimi, “t” harfindeki çubuk yüksekliğidir. Temelden öznitelik çıkarmak için kullanılan yöntem polygonization yöntemidir. “t” harfi için şablon eşleme, diğer özellikler için metnin izdüşüm profilleri için bir satır kullanılmıştır. Bu parametreler, yazar hakkında doğru bilgileri ortaya çıkarmaktadır. Çalışma için Python programlama dilini kullanmışlardır. Belirtilen bu özellikler yardımı ile yazarın kimliği geliştirilen araç tarafından belirlenmiştir. Aracın geliştirilmesi için ise IAM veritabanı kullanılmıştır (Mathur ve Patil, 2020).

Prabhu ve Ranjitha, çevrimdışı ve çevrimiçi sistemler için veri kümeleri, veri edinme cihazları ve ön işleme çalışmalarını incelemişlerdir. Ardından yazarın kişilik özelliklerini ortaya çıkaran kenar boşlukları, eğik çizgiler, yazı eğimi, bağlantılı harfler, hız ve baskı

noktaları gibi el yazısı karakterlerini elde etme yöntemlerini önermişlerdir. DVM, Saklı Markov Modelleri, YSA, Dinamik Zaman Paketleme ve Sembolik Gösterimlerin gelecek vadede yöntemler olduğu belirtilmiştir (Prabhu ve Ranjitha, 2020).

3. MATERYAL ve YÖNTEMLER

3.1. Kullanılan Algoritmalar

Farklı makine öğrenmesi algoritmaları uygulanarak performansları karşılaştırılacaktır. Kullanılacak makine öğrenme algoritmaları şunlar olacaktır.

3.1.1. Naive Bayes (NB) Sınıflandırıcı

Naive Bayes sınıflandırıcısı, Bayes teoremine dayanan sade bir yöntemdir. Özellikle veri kümesindeki özellik sayısı çok olduğunda kullanılması durumunda iyi sonuçlar vermektedir (Mitchell, 1997). Bu yüzden genellikle metinlerin sınıflandırılmasında yaygın olarak kullanılır. Sadeliğine rağmen, NB sınıflandırıcısı, karar ağacı ve sinir ağı sınıflandırıcısı gibi karmaşık sınıflandırma yöntemleriyle karşılaştırılabilir performans elde edebilir.

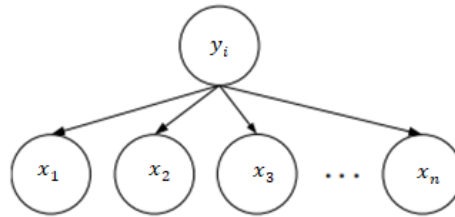
NB sınıflandırıcı hipotezler için olasılıkları açık biçimde hesaplar. Bu açıdan olasılık kullanmayan algoritmalara göre üstün bir yönü bulunmaktadır. Eğitim esnasında her eğitim örneği, hipotez olasılığına katkıda bulunmaktadır.

Şekil 2’de basit bir Naive Bayes sınıflandırıcısının yapısı gösterilmiştir. Bu yapıda sadece iki katman bulunmaktadır. “y” sınıf değişkeni diğer tüm “x” değişkenlerinin üst düğümüdür.

$$P(A|B) = P(B|A)P(A)/P(B) \quad (1)$$

- $P(B|A)$: A verildiğinde, B’nin doğru olma olasılığı.
- $P(A|B)$: B verildiğinde, A’nın doğru olma olasılığı.
- $P(A)$: A’nın doğru olma olasılığı.
- $P(B)$: B’nin doğru olma olasılığı.

$$P(y_i|x_1, x_2, \dots, x_n) = P(x_1, x_2, \dots, x_n|y_i) \times P(y_i)/P(x_1, x_2, \dots, x_n) \quad (2)$$



Şekil 2. Naive Bayes (Taheri S. ve Mammadov M., 2013)

3.1.2. Karar Ağaçları

Karar ağaçları sınıflandırma veya regresyon amaçlı kullanıldığında veri kümesindeki değerlerin istatistiksel sonuçlarına göre çalıştığı için hatalara karşı toleranslıdır. Karar ağacı oluşturulurken veri kümesindeki özelliklerin hangisinin seçileceği Bilgi Kazancı (Information Gain) adı verilen bir hesaplama yöntemi ile belirlenir. Bilgi kazancı değeri bilgi teorisinde yer alan entropi kavramından yararlanılarak bulunur. Bilgi kazancı yüksek olan özellikler ağacın köküne yakın olarak yerleşir. Diğer özellikler de sırasıyla ağaçta yer alır ve karar ağacı bu şekilde oluşturulur.

Karar ağaçları kesişimlerin birleşimi (disjunction of conjunctions) şeklindeki hipotez uzayında arama yaparak veri kümesine en çok uyan hipoteze ulaşmaya çalışır. Kesişimlerin birleşimi şeklinde ifade edilen hipotezlerin ifade gücü yüksektir.

Karar ağaçlarının ilk sınıflandırma uygulaması ID3 algoritması ile olmuştur (Quinlan, 1993). Daha sonra sürekli değerleri de ele alabilen versiyonu C4.5 ortaya çıkmıştır.

Karar ağaçlarının avantajlarından birisi de eğitim sonucunda oluşan ağacın insanlar tarafından anlaşılır biçimde olmasıdır. Diğer makine öğrenmesi algoritmalarının içyapısı insanlar tarafından anlaşılacak durumdadır.

Karar ağacı oluşturulurken eğitim kümesinin tamamını doğru sınıflandıracak şekilde ağaç büyütülür. Bu durum eğitim örneklerinin ezberlenmesi ile sonuçlanabilir. Bu durumun ortaya çıkmasını önlemek için ağaçtaki bazı düğümler sonradan budanabilir. Bu işleme sonradan budama (post pruning) denir. Başka bir yöntem de ağacın çok fazla karmaşık hale gelmeden önce eğitimin durdurulmasıdır. Makine öğrenmesi algoritmalarında ezberlemenin (overfitting) ortaya çıkması yöntemin eğitim kümesine kendini fazla uyduracak şekilde karmaşık hale gelmeleri ile olur.

$$Entropy(S) = - \sum_{i=1}^c p_i \log_2(p_i) \quad (3)$$

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (4)$$

3.1.3. Rastgele Orman

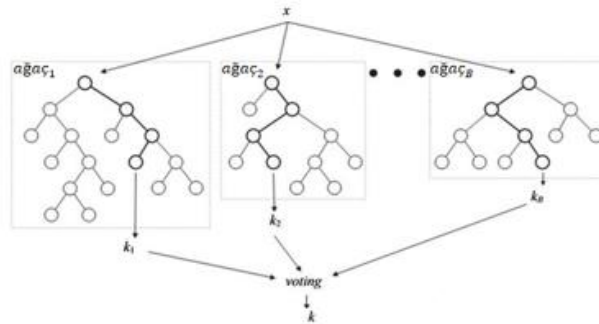
Rastgele orman algoritması tek bir büyük karar ağacı yerine, hepsinin çıktısının birleştirileceği küçük karar ağaçları oluşturma mantığına dayanır (Breiman, 2001). Şekil 3'te küçük karar ağaçlarından oluşan rastgele orman algoritmasının genel yapısı görülmektedir.

Böylece sınıflandırıcıların birleştirilmesi (ensemble) mantığı ile daha güçlü bir sınıflandırıcı elde edilmiş olur. Rastgele ormandaki küçük karar ağaçlarının oluşturulmasında kullanılan algoritma normal karar ağaçlarının oluşturulması ile aynıdır. Rastgele ormandaki fark veri ağaçları oluştururken veri kümesinin nasıl kullanıldığıdır. Öncelikle veri kümesindeki satırlar rastgele biçimde seçilerek ayrı bir veri kümesi oluşturulur. Aynı satır birden fazla kez yeni bir kümesinde yer alabilir. Ayrıca yeni veri kümesinde tüm özellikler değil rastgele seçilen birkaç özellik kullanılır. Böylece ormandaki her bir ağacın eğitimi orijinal veri kümesinden rastgele seçilerek oluşturulmuş yeni bir veri kümesi ile ve ayrıca yeni kümesindeki özelliklerden rastgele seçilmiş özelliklerle yapılır. Bu şekilde veri kümesinin boyutuna ve özelliklerin sayısına göre değişen çok sayıda küçük ağaç oluşur.

Yeni veri kümelerinin oluşumu (bootstapping) rastgele seçimle yapıldığı için bu işlem sırasında bazı satırların hiç seçilmeme durumu olabilir. Buna out-of-bag (torbanın dışında kalma) denir. Bu satırlar daha sonra rastgele ormanın performansının ölçülmesinde kullanılır.

Rastgele ormandaki küçük karar ağaçlarının oluşturulması sırasında düğümlere yerleşecek özelliklerin seçimi için genellikle Gini İndisi ölçütü kullanılır (Wu vd., 2017).

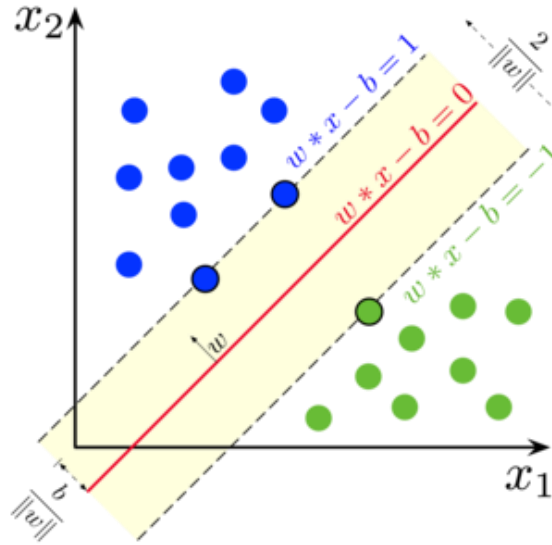
$$Gini = 1 - \sum_{i=1}^C (p_i)^2 \quad (5)$$



Şekil 3. Rastgele orman algoritmasının şeması (Zhang vd., 2018)

3.1.4. Destek Vektör Makineleri (DVM)

Destek vektör makineleri algoritmasının orijinal hali iki farklı sınıfa ait verilerin sınıflandırılmasını doğrusal biçimde sağlar (Corinna ve Vapnik, 1995). Bu algoritmada veri kümesindeki örnekleri birbirinden ayırmak için kullanılan marjın kavramı vardır. Perceptron gibi doğrusal sınıflandırıcılar verileri tek bir doğrusal çizgi ile birbirinden ayırarak sınıflandırır. Dolayısıyla bu yöntem kaba bir sınıflandırma sağlamakta ve yeni gelecek verileri sınıflandırmada başarısız olma ihtimali bulunmaktadır. DVM yönteminde ise verileri birinden ayıran bir çizgi yerine marjın söz konusudur. Marjın, birbirine paralel iki çizgi arasındaki bölgeye verilen isimdir. DVM algoritması bu marjını maksimize ederken sınıflandırma hatasını minimum tutmak zorundadır. Bu klasik bir optimizasyon problemidir. Şekil 4'te marjın sarı ile gösterilmiştir. Kesikli çizgi ile gösterilen birbirine paralel iki sınır çizgisi üzerinde sınıflara ait veriler bulunmaktadır. Bu noktalara destek vektörleri denmektedir. DVM yöntemi adını bu vektörlerden almaktadır.

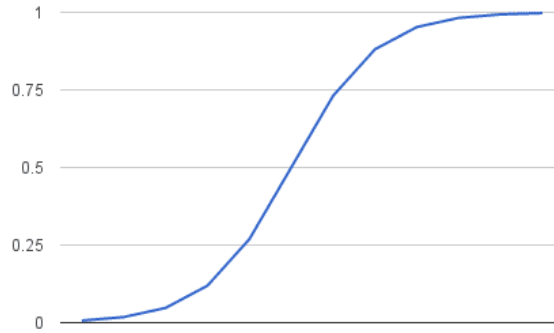


Şekil 4. İki boyutlu uzay için marjın ve destek vektörlerinin gösterimi (Support Vector Machines)

3.1.5. Lojistik Regresyon

Lojistik regresyon, ikili sınıflandırma problemlerinde (hedef kategorik olduğunda) kullanılan makine öğrenmesi algoritmasıdır. Lojistik regresyon; temel olarak, ikili çıktı değişkenini modellemek için fonksiyonda tanımlanan bir lojistik işlevi kullanır (Tolles ve Meurer, 2016). Lojistik Regresyonun amacı, özellikler ile belirli bir sonucun olasılığı arasındaki bağlantıyı keşfetmektir. “Lojistik” adı, bu sınıflandırma yaklaşımında kullanılan “Logit” işlevinden gelmektedir Lojistik fonksiyon denkleminde x girdi değişkenidir. Şekil 5’te lojistik fonksiyon kullanılarak “0 ve 1” aralığına dönüştürülmüş “-5 ile 5” arasındaki sayıların grafiği gösterilmiştir.

$$\text{Lojistik fonksiyon} = \frac{1}{1+e^{-x}} \quad (6)$$



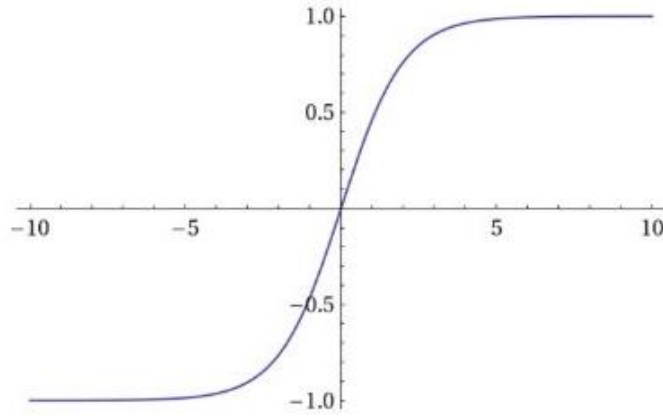
Şekil 5. Lojistik fonksiyon kullanılarak elde edilen sayıların grafiği (Machine Learning Mastery)

Softmax işlevi, “K” gerçek değerlerinin bir vektörünü, toplamı 1 olan bir “K” gerçek değerlerinin vektörüne dönüştüren bir işlevdir. Giriş değerleri pozitif, negatif, sıfır veya birden büyük olabilir, ancak softmax bunları 0 ile 1 arasında değerlere dönüştürür böylece bunlar olasılık olarak yorumlanabilir.

Softmax işlevi bazen softargmax işlevi veya çok sınıflı lojistik regresyon olarak adlandırılır. Bunun nedeni, softmax'ın çok sınıflı sınıflandırma için kullanılabilen bir lojistik regresyon genellemesi olması ve formülünün lojistik regresyon için kullanılan sigmoid işlevine çok benzer olmasıdır. Softmax işlevi, yalnızca sınıflar birbirini dışladığında bir sınıflandırıcıda kullanılabilir. Şekil 6'da Softmax'ın aktivasyon fonksiyonu gösterilmiştir.

$$\sigma\left(\frac{\rightarrow}{z}\right)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (7)$$

- $\frac{\rightarrow}{z}$: (z_0, \dots, z_K) 'den oluşan softmax işlevine giriş vektörü.
- z_i : Tüm z_i değerleri, softmax işlevine giriş vektörünün öğeleridir ve pozitif, sıfır veya negatif herhangi bir gerçek değeri alabilirler.
- e^{z_i} : Giriş vektörünün her elemanına standart üstel fonksiyon uygulanır. Bu, girdi negatifse çok küçük ve girdi büyükse çok büyük olacak olan 0'ın üzerinde pozitif bir değer verir. Ancak yine de bir olasılık için gerekli olan $(0, 1)$ aralığında sabit değildir.
- K : Çok sınıflı sınıflandırıcıdaki sınıf sayısı.



Şekil 6. Softmax Aktivasyon Fonksiyonu

3.2. Kişilik Testi

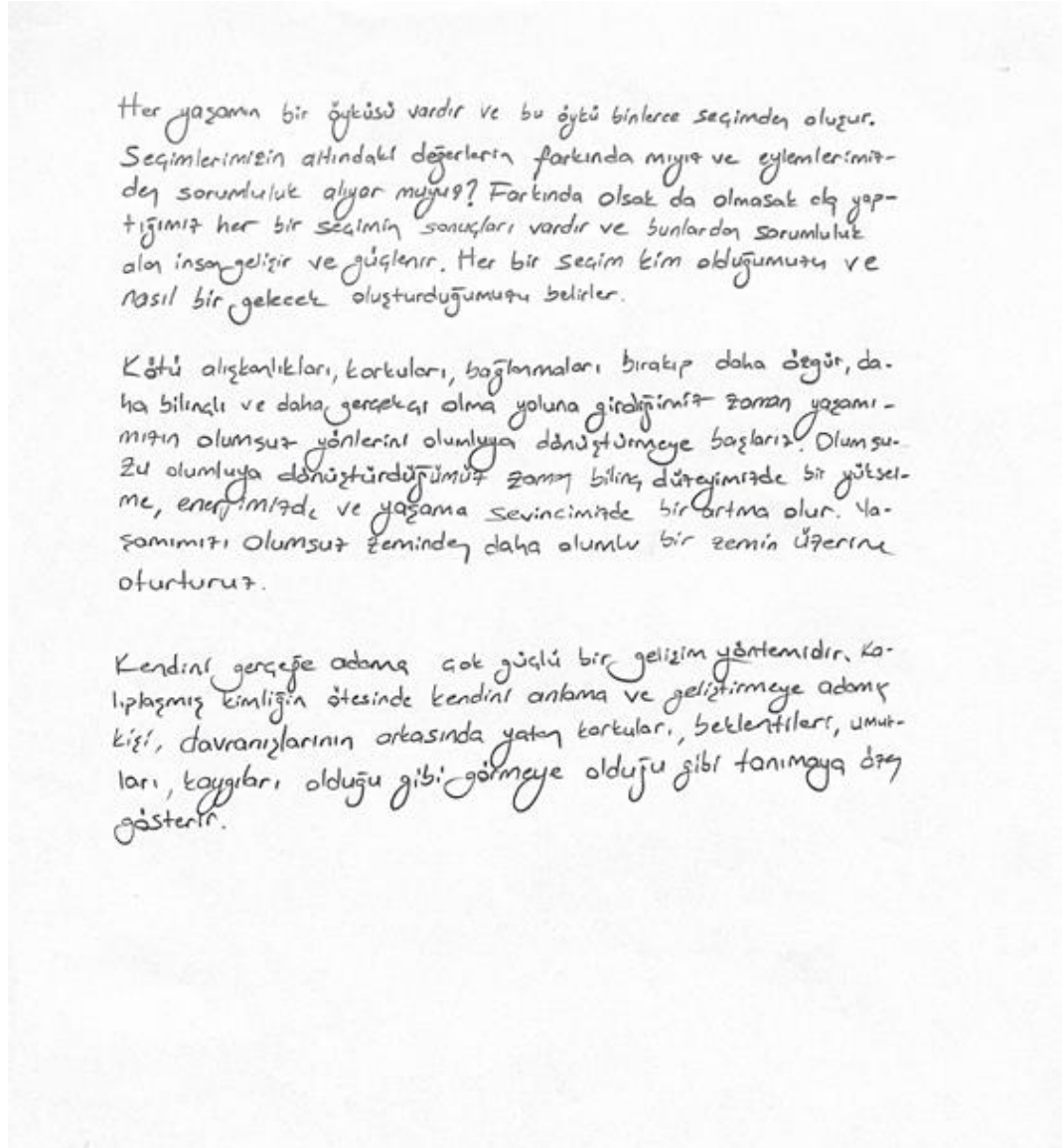
Araştırmada kullanılan Sıfatlara Dayalı Kişilik Testi, Beş Faktör Kişilik kuramı doğrultusunda Bacanlı ve arkadaşları (2009) tarafından geliştirilmiştir. Ölçek, 40 maddeden oluşmakta ve katılımcılar tarafından 10-15 dakika içerisinde tamamlanabilmektedir. Uygulamada; kullanıcılardan, iki zıt kutup olarak verilmiş 40 tane sıfattan kendilerine en yakın olanlarını işaretlemeleri beklenmektedir. Seçenekler (1) Çok uygun, (2) Oldukça uygun, (3) Biraz uygun, (4) Ne uygun, ne uygun değil, (5) Biraz uygun, (6) Oldukça uygun ve (7) Çok uygun şeklinde verilmiştir. Sıfatlara Dayalı Kişilik Testinin orijinali (1) Nevrotiklik-7 madde, (2) Dışa Dönüklük-9 madde, (3) Gelişime Açıklık-8 madde, (4) Uyumluluk-9 madde ve (5) Sorumluluk-7 madde olmak üzere beş boyuttan oluşmaktadır. Ancak, bu araştırmada, el yazısından elde edilen kişilik analizi ışığında maddeler bir uzman psikolog tarafından ayrı ayrı ilişkilendirilerek kişilik tiplerine uygun olup olmadığı, maddelerin ortalama değerleri alınarak değerlendirilmiştir. Sonunda, elde edilmiş kişilik tipleri ile yapılan uzman değerlendirmesinin korelasyonuna bakılmıştır.

Kısıtlamalar

- Gelecek araştırmalarda, sıfatlara dayalı kişilik testi maddelerinin bu teori için uygunluğunun madde ve faktör analizi yapılarak değerlendirilmesi gerekmektedir, yapılmamış olması araştırmanın kısıtlarındandır.
- Sonuçlar sadece bir uzman tarafından değerlendirilmiştir.
- Örneklem sayısının daha çok olması gerekmektedir.
- Sonuçlar betimsel olup, herhangi bir genelleme yapmak mümkün değildir.

3.3. Kullanılan Veri Kümesi

Bu çalışmada kullanılan veri seti kümesi, farklı yaş ve çalışma grubuna mensup kişilerden alınan yazı örnekleriyle tarafımdan oluşturulmuştur. Kişilere hazır bir metin verilmiş ve aynı metnin boş bir A4 kâğıdına kendi el yazıları ile yazılması istenmiştir. Şekil 7’de bir kişiden alınan gerçek el yazısı örneği gösterilmiştir.



Şekil 7. Kişilerden alınan el yazısı örneği

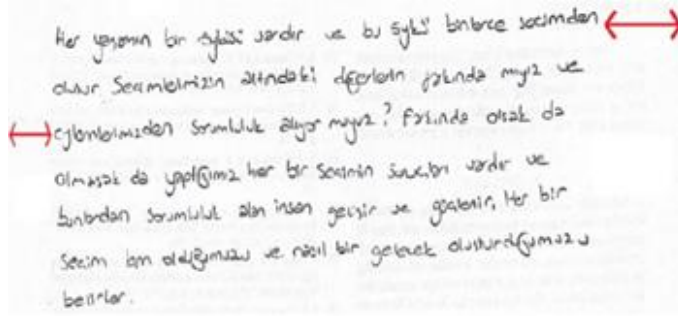
3.4. Yöntemler

El yazısı örneklerinden aşağıdaki özellikler çıkartılacaktır.

1. Sayfanın sağ, sol kenar boşlukları

Şekil 8’de sağ ve sol kenar boşlukları oklarla gösterilmiştir.

Sayfanın sağ kısmında genişlik çok ise; canlı ve hareketli, genişlik az ise; ciddi ve temkinli bir kişiliği, sol kısmında genişlik çok ise; riske karşı korkulu, genişlik az ise; sabırsız bir kişiliği gösterir.

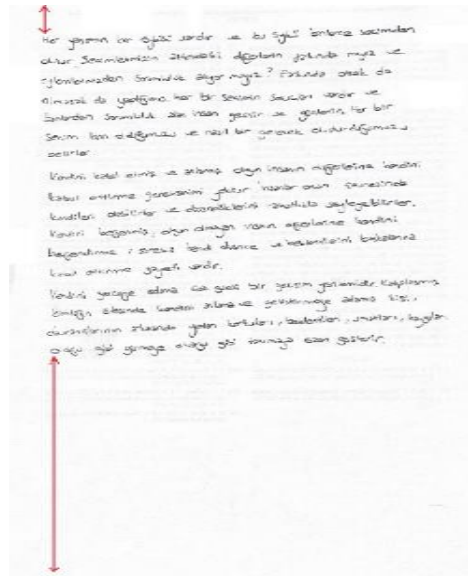


Şekil 8. Sağ ve sol kenar boşlukları gösterimi

2. Sayfanın üst ve alt kenar boşlukları

Şekil 9’da sayfanın üst ve alt boşlukları oklarla gösterilmiştir.

Sayfanın üst kısmında genişlik çok ise; temkinli, genişlik az ise; sabırsız bir kişiliği, alt kısmında genişlik çok ise duygu ve davranışlarda ince, genişlik az ise; tedirgin, ekonomik kaygısı yüksek bir kişiliği gösterir.



Şekil 9. Alt ve üst kenar boşlukları gösterimi

3. Genel sayfa düzeni

Formül 6'nın sonucuna göre sayfanın belli bir düzende olup olmadığı kontrol edilecektir. Sayfa kenarlarında ortalama boşluk;

%10'dan az ise; çekingen, içe dönük aileye bağlı kişiliği

%10-%25 arası ise; dengeli ve uyumlu kişiliği

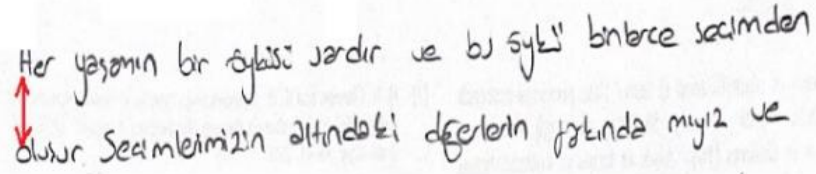
%25'den fazla ise; cömert, dışa dönük kişiliği gösterir.

$$\text{Sayfa Düzeni} = \left(\frac{(\text{Sol Marjin} + \text{Sağ Marjin})}{2} + \frac{(\text{Üst Marjin} + \text{Alt Marjin})}{2} \right) / 2 \quad (6)$$

4. Satırlar arası boşluk miktarı

Şekil 10'da iki satır arası boşluk ok ile gösterilmiştir.

Satırlar arası boşluk çok ise; sakın ve geniş perspektiften bakma eğilimi yüksek, az ise; hareketi ve kalabalığı seven bir kişiliği gösterir.



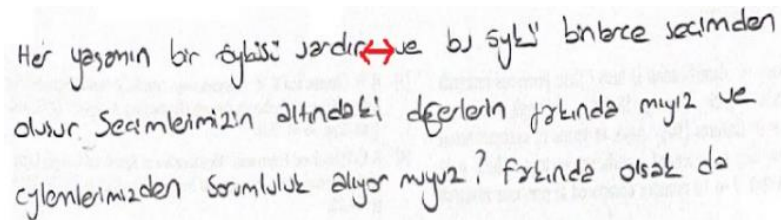
Her yaşamın bir sırası vardır ve bu sıralar birbirine sıralardan oluşur. Seçimlerimizin altındaki değerlerin farkında mıyız ve

Şekil 10. Satır arası boşlukların gösterimi

5. Kelimeler arası boşluk miktarı

Şekil 11'de iki kelime arası boşluk ok ile gösterilmiştir.

Kelimeler arası boşluk çok ise; içe dönük, özgürlüğüne ve kişisel alanına düşkün, az ise; sosyal ve hareketli bir kişiliği gösterir.



Her yaşamın bir sırası vardır ve bu sıralar birbirine sıralardan oluşur. Seçimlerimizin altındaki değerlerin farkında mıyız ve çözümlerimizden sorumluluk alıyor muyuz? farkında olsak da

Şekil 11. Kelimeler arası boşlukların gösterimi

6. Kaleme uygulanan baskı

Kaleme uygulanan baskı fazla ise; stresli, gergi ve alıngan, baskı az ise; kırılğan ve zarif bir kişiliği gösterir. Şekil 12'de iki farklı kişiden alınan el yazısı örneği arasındaki baskı farkları gösterilmiştir.

Her yaşamın bir öyküsü vardır ve bu öykü binlerce seçimden oluşur. Seçim altındaki değerlerin farkında mıyız ve eylemlerimizden sorumluluk alıyor muyuz? Farkında olsak da olmasak da yaptığımız her bir seçimin sonuçları vardır ve bunlardan sorumluluk alan insan gelişir ve güçlenir. Her bir seçim kim olduğumuzu ve nasıl bir gelecek oluşturduğumuzu belirler.

Kötü alışkanlıkları, korkuları, bağlanmaları birleştiren daha özgür, daha bilinçli ve daha gerçekçi olma yoluna çıktığımız zaman yaşamımızı olumsuz yönlere dönüştürmeye başlarız. Olumsuz olumluya dönüştürdüğümüz zaman bilinçli yaşamımızda bir yükselme, enerjimizde ve yaşam seviyamızda bir artış olur. Yaşamımızı olumsuz zeminlerden daha olumlu bir zemine oturturuz.

Kendini geçiğe adanmış çok güçlü bir gelişim yöntemidir. Kalıplaşmış kimliğinin etkisinde kendini tanıma ve geliştirmeye adanmış kişi, davranışlarının etkisinde yatan korkuları, beklentileri, umutları, kaygıları olduğu gibi görmeye olduğu gibi tanımayacağı gibi gösterir.

(a)

Her yaşamın bir öyküsü vardır ve bu öykü binlerce seçimden oluşur. Seçim altındaki değerlerin farkında mıyız ve eylemlerimizden sorumluluk alıyor muyuz? Farkında olsak da olmasak da yaptığımız her bir seçimin sonuçları vardır ve bunlardan sorumluluk alan insan gelişir ve güçlenir. Her bir seçim kim olduğumuzu ve nasıl bir gelecek oluşturduğumuzu belirler.

Kötü alışkanlıkları, korkuları, bağlanmaları birleştiren daha özgür, daha bilinçli ve daha gerçekçi olma yoluna çıktığımız zaman yaşamımızı olumsuz yönlere dönüştürmeye başlarız. Olumsuz olumluya dönüştürdüğümüz zaman bilinçli yaşamımızda bir yükselme, enerjimizde ve yaşam seviyamızda bir artış olur. Yaşamımızı olumsuz zeminlerden daha olumlu bir zemine oturturuz.

Kendini geçiğe adanmış çok güçlü bir gelişim yöntemidir. Kalıplaşmış kimliğinin etkisinde kendini tanıma ve geliştirmeye adanmış kişi, davranışlarının etkisinde yatan korkuları, beklentileri, umutları, kaygıları olduğu gibi görmeye olduğu gibi tanımayacağı gibi gösterir.

(b)

Şekil 12. (a) Basısı az olan (b) Basısı çok olan el yazısı örneği

7. Yazının Eğimi

Şekil 13'te eğimi yukarı doğru olan bir yazı çizgileriyle gösterilmiştir.

Yazının eğimi aşağı doğru ise; kötümser, yukarı doğru ise; iyimser, düz, dengeli bir kişiliği gösterir.

Her yaşamın bir öyküsü vardır ve bu öykü binlerce seçimden oluşur. Seçim altındaki değerlerin farkında mıyız ve eylemlerimizden sorumluluk alıyor muyuz? Farkında olsak da

Şekil 13. Yazıdaki eğimin gösterimi

3.5. Metrikler

Performans analizi yapılırken kullanılan metrikler; Accuracy, F-measure, Recall, precision ve confisuion matris metrikleri kullanılmıştır. Kullanılan metrikler aşağıda tanımlanmıştır ve metriklerde bulunan parametreler ise Tablo 1’de gösterilmiştir.

Tablo 1. Karmaşıklık matrisi

	Beklenen Negatif	Beklenen Pozitif
Gerçek Negatif	TN	FP
Gerçek Pozitif	FN	TP

- TN (True Negatif): Doğru sınıflandırılmış negatif vakaların oranıdır.
- FN (False Negatif): Yanlışlıkla negatif olarak sınıflandırılan pozitif vakaların oranıdır.
- FP (False Pozitif): Yanlışlıkla pozitif olarak sınıflandırılan negatif vakaların oranıdır.
- TP (True Pozitif): Doğru şekilde tanımlanan pozitif vakaların oranıdır (Büyükoflaz ve Öztürk, 2018).

4. DENEYSEL SONUÇLAR

Kişilerden alınan el yazısı görüntülerinden kişilik analizinde yararlanılacak özellikler Python dilinde Open CV kütüphaneleri kullanılarak yazılan kodlar ile çıkarılmıştır. Hazır kodlardan yararlanılmış olup gerekli özellik çıkarımlarının yapıldığı kod parçacıkları üzerlerine eklenmiştir.

Kişilerden alınan el yazısı görüntüleri üzerinde aşağıdaki özellikler çıkarılmıştır:

- Sol marjin
- Sağ marjin
- Üst marjin
- Alt marjin
- Marjinlerden elde edilen sayfanın genel düzeni
- Yazının genel eğimi
- Kelimeler arası ortalama boşluk
- Satırlar arası ortalama boşluk
- Kalem basısı

4.1. Sol, sağ, üst ve alt marjinlerin çıkarılması

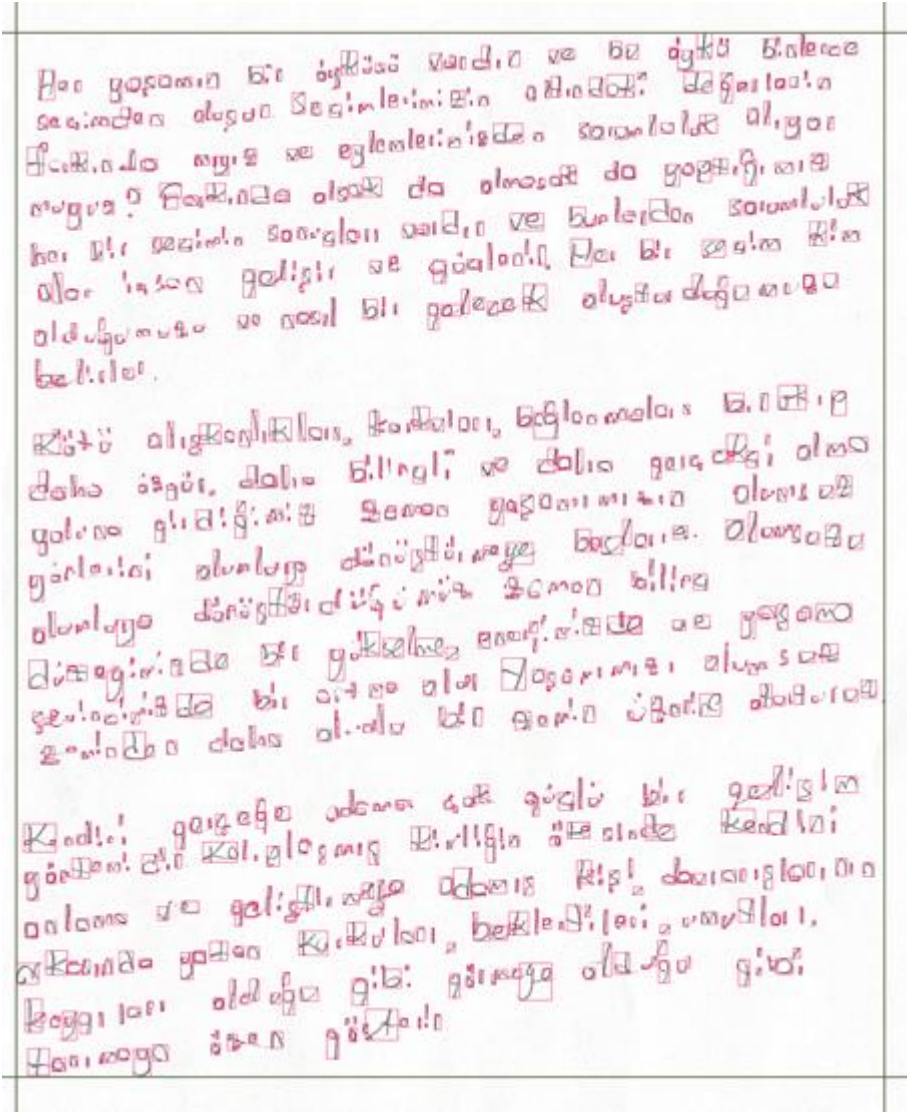
Bunun için el yazısı görüntüsünün kelime ve harf gibi temel bileşenlerinin çıkarıldığı görüntüde, temel bileşenlerin etrafına çizilen dikdörtgenlerden yararlanılmıştır. Şekil 14'te bu dikdörtgenlere göre; sağ, sol, üst ve alt marjinlerin çizilmesi gösterilmiştir.

Sol marjin: Elde edilen dikdörtgenlerin en solda olanın sol kenar koordinatı

Sağ marjin: Elde edilen dikdörtgenlerin sağ kenarlarından en sağ olan koordinatı

Üst marjin: Elde edilen dikdörtgenlerin üst kenarının en düşük olduğu koordinat

Alt marjin: Elde edilen dikdörtgenlerin alt kenarının en yüksek olduğu koordinat



Şekil 14. Marjin çizgilerinin gösterimi

4.2. Marjinlerden elde edilen sayfanın genel biçimi

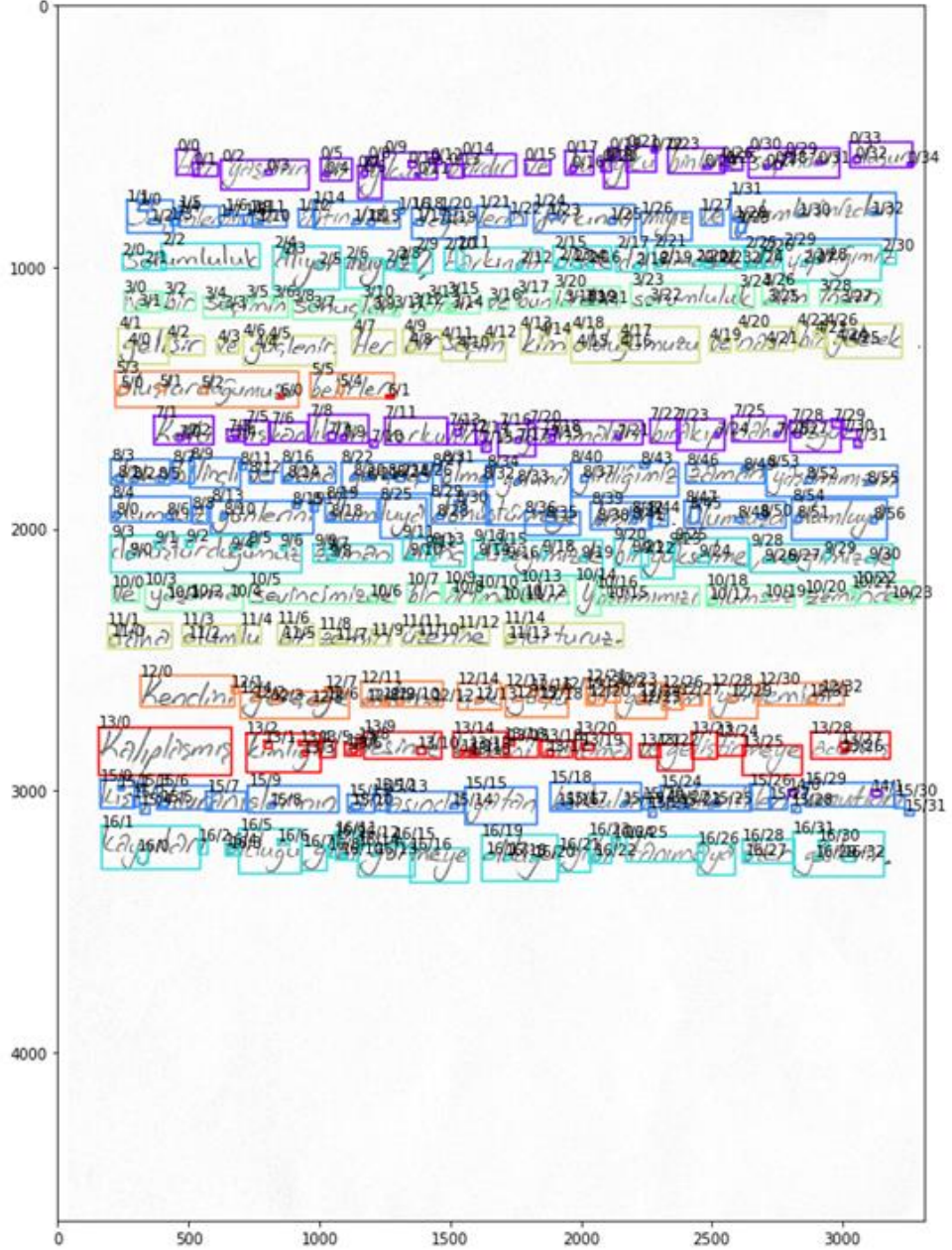
Sol ve sağ marjinin ortalaması ile üst ve alt marjinlerin ortalamasının toplamının ikiye bölünmesi ile elde edilen değere göre yapılan değerlendirme ile sayfanın genel düzeni elde edilmiştir.

4.3. Yazının genel eğimi

Yazının tamamının y eksenine göre yaptığı açı bulunmuştur.

4.4. Kelimeler arası ortalama boşluk

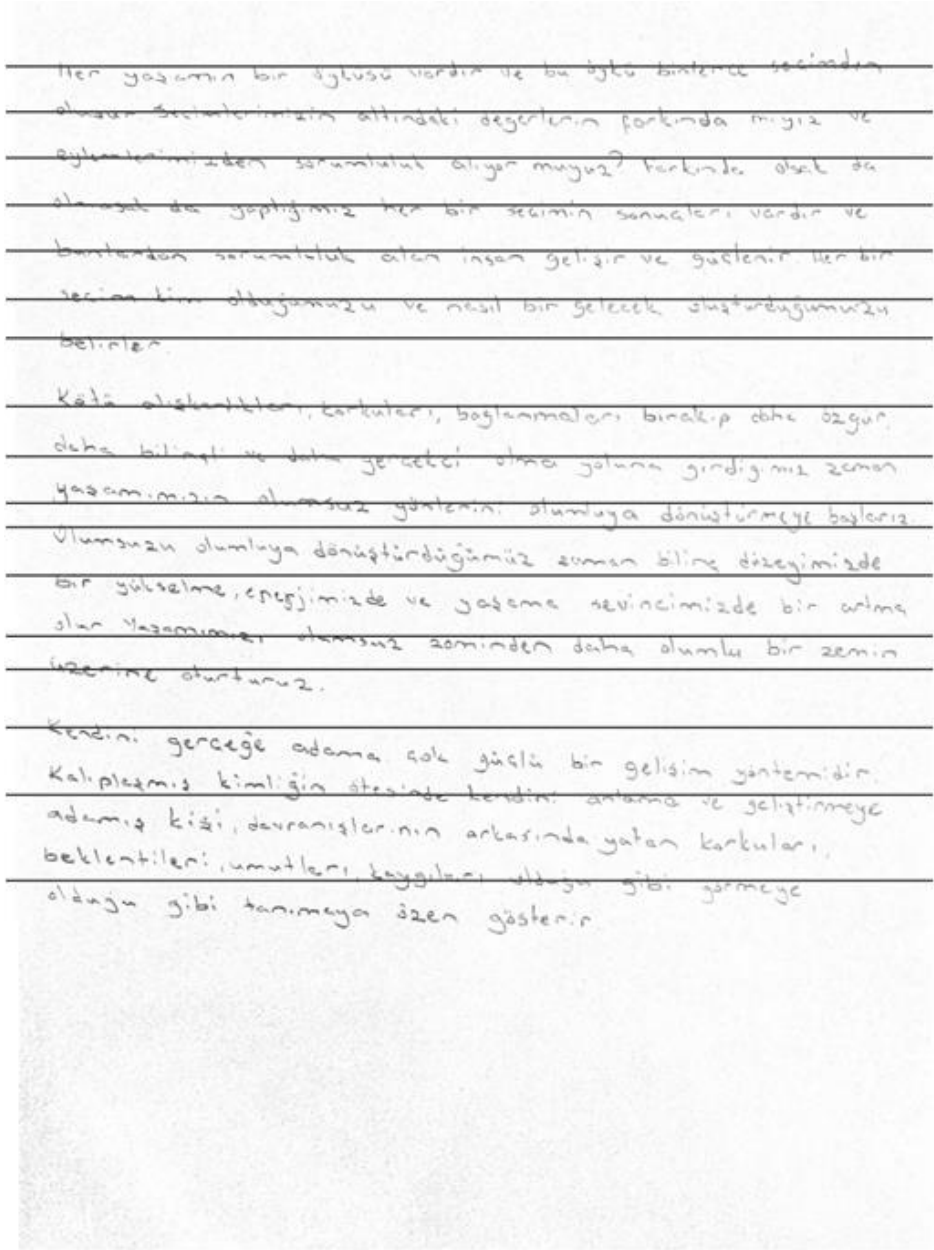
El yazısında kelimeler tespit edildikten kelimelerin arasındaki boşluklar bulunarak bunların ortalaması bulunmuştur. Şekil 15’te tespit edilen kelimeler ve bunların satırlara göre numaralandırılması gösterilmiştir.



Şekil 15. Tespit edilen kelimelerin gösterimi

4.5. Satırlar arası ortalama boşluk

El yazısındaki kelimeler tespit edildikten sonra kelimelerin oluşturduđu satırlar arasındaki boşluklar bulunarak bunların ortalaması hesaplanmıřtır. Őekil 16'da her çizginin bir satırı ifade ettiđi gösterilmiřtir. Bu çizgiler arasındaki boşluklar satırlar arası boşlukları ifade etmektedir.



Őekil 16. Tespit edilen satırların gösterimi

4.6. Yazı basısı

Yazıdaki gri seviye tonlamaların ortalamasının siyaha yakınlığı hesaplanarak bulunmuştur.

Tablo 2’de tüm giriş özelliklerinin her bir kişi için tek tek hesaplanmasından elde edilen veri kümesi gösterilmiştir.

Tablo 2. Elde edilen veri kümesine ait giriş özellikleri

	Eğim	Bası	Sol Marjin	Sağ Marjin	Üst Marjin	Alt Marjin	Kelimler Arası Boşluk	Satırlar Arası Boşluk	Sayfa Düzeni
Kişi 1	2,19	247,25	266	146	147	1316	20	179	468,75
Kişi 2	1,83	246,80	77	155	174	214	42	465	155
Kişi 3	1,89	248,57	588	345	742	3120	54	106	1198,75
Kişi 4	0,60	245,01	229	22	731	1817	13	154	699,75
Kişi 5	0,61	245,21	53	111	42	800	1	32	251,5
Kişi 7	90,00	247,06	13	7	24	633	6	47	169,25
Kişi 8	1,21	248,20	693	0	466	2220	38	179	844,75
Kişi 9	12,37	249,09	0	146	0	2418	3	97	641
Kişi 10	0,53	250,05	549	254	426	3363	35	105	1148
Kişi 11	2,38	243,86	62	0	25	166	53	79	63,25
Kişi 12	88,23	250,33	58	96	145	876	7	30	293,75
Kişi 13	0,04	249,55	0	78	0	2207	18	116	571,25
Kişi 14	0,73	249,29	0	249	0	2415	38	103	666
Kişi 15	0,62	249,23	58	0	221	707	11	36	246,5

Kişi 16	0,51	248,94	431	9	163	3882	70	237	1121,25
Kişi 17	1,92	246,12	0	9	0	746	48	354	188,75
Kişi 18	0,65	248,73	338	296	624	1708	41	118	741,5
Kişi 19	0,93	246,83	532	112	723	1296	11	154	665,75
Kişi 20	0,82	244,93	174	14	355	1621	10	102	541
Kişi 21	0,99	244,55	2	41	22	598	8	40	165,75
Kişi 22	0,72	247,63	0	237	6	1287	14	119	382,5
Kişi 23	0,59	247,58	169	114	229	1362	7	153	468,5
Kişi 24	0,25	248,51	316	490	232	840	22	139	469,5
Kişi 25	0,24	247,96	438	157	350	2742	7	197	921,75
Kişi 26	2,43	244,47	0	197	2	194	22	185	98,25
Kişi 27	0,88	245,01	151	296	139	9	44	151	148,75
Kişi 28	1,56	247,25	340	324	134	752	42	717	387,5
Kişi 29	0,31	247,15	140	376	338	376	11	143	307,5
Kişi 30	88,94	248,87	40	15	21	841	27	50	229,25
Kişi 31	4,32	242,81	340	0	0	232	42	294	143
Kişi 32	6,80	246,47	304	78	475	1859	12	98	679
Kişi 33	0,65	245,37	113	6	180	521	5	38	205
Kişi 34	1,41	244,10	147	159	467	1917	16	97	672,5
Kişi 36	2,07	245,81	368	24	0	1739	4	97	532,75
Kişi 37	0,60	244,72	267	137	384	2103	9	143	722,75

Kişi 40	0,00	244,42	267	0	0	1965	33	202	558
Kişi 41	0,14	243,39	180	0	87	990	7	171	314,25
Kişi 42	2,12	246,28	41	0	45	725	44	50	202,75
Kişi 43	1,13	242,73	430	60	0	763	45	201	313,25
Kişi 44	0,21	246,59	234	142	319	1314	20	122	502,25
Kişi 48	0,00	244,61	191	0	191	937	16	147	329,75
Kişi 49	0,16	245,55	680	85	883	862	39	189	627,5
Kişi 51	0,00	244,20	78	1212	109	136	46	302	383,75
Kişi 52	0,21	244,71	148	221	178	102	29	348	162,25
Kişi 53	0,84	244,77	226	325	342	2152	14	115	761,25
Kişi 54	0,18	245,46	193	21	80	753	34	277	261,75
Kişi 57	2,39	246,25	526	73	0	1194	20	113	448,25
Kişi 58	1,96	246,09	675	284	349	2835	34	89	1035,75
Kişi 59	0,57	245,24	4	92	5	875	11	146	244
Kişi 61	0,46	238,38	446	529	326	2198	33	117	874,75
Kişi 62	2,12	243,55	83	74	182	640	40	263	244,75
Kişi 63	4,79	244,65	303	3	505	0	38	285	202,75
Kişi 64	0,63	244,76	3	0	0	880	32	218	220,75
Kişi 65	5,57	247,86	0	17	0	151	10	101	42
Kişi 69	3,52	247,52	84	10	68	665	51	68	206,75
Kişi 70	0,66	247,35	88	85	101	950	14	32	306

Tablo 3’de, yapılan kişilik analizi sonucunda elde edilecek çıkış değerleri gösterilmiştir. Çıkış değerleri sayısal “1” ya da “0” şeklindedir.

Tablo 3. Kişilik analizi sonucu elde edilecek çıkış değerleri

İyimser	Dengeli	Stresli	Sabırsız	Temkinli	Tedirgin	Duygusal	İçe Dönük	Dışa Dönük	Sakin
---------	---------	---------	----------	----------	----------	----------	-----------	------------	-------

Tablo 4, 5, 6, 7 ve 8’de kullanılan makine öğrenmesi algoritmaları için elde edilen sonuçlar verilmiştir. Algoritmalar genel olarak, el yazısı özelliklerinden az sayıda veri içerenler için daha başarısız olmuştur.

Tablo 4. Rastgele Orman algoritması için elde edilen sonuçlar

	Rastgele Orman			
	Doğruluk	Kesinlik	Hatırlama	F1 Skoru
İyimser	0,80	0,73	0,72	0,68
Dengeli	0,80	0,88	0,85	0,86
Stresli	0,93	0,83	0,50	0,61
Sabırsız	0,98	0,97	1,00	0,98
Temkinli	0,89	0,91	0,95	0,93
Tedirgin	0,91	0,83	0,72	0,75
Duygusal	0,87	0,88	0,89	0,87
İçe Dönük	0,85	0,96	0,80	0,87
Dışa Dönük	0,84	0,84	0,80	0,81
Sakin	0,91	0,92	0,78	0,84

Tablo 4’e göre “sabırsız” ve “sakin” özellikleri için “Rastgele Orman” algoritmasının sonuçlarının diğer özelliklere göre daha başarılı olduğu görülmektedir.

Tablo 5. Lojistik Regresyon algoritması için elde edilen sonuçlar

	Lojistik Regresyon			
	Doğruluk	Kesinlik	Hatırlama	F1 Skoru
İyimser	0,59	0,24	0,59	0,34
Dengeli	0,59	0,68	0,77	0,72
Stresli	0,93	0,83	0,50	0,61
Sabırsız	0,89	0,88	0,93	0,90
Temkinli	0,71	0,80	0,79	0,78
Tedirgin	0,95	0,83	0,92	0,86
Duygusal	0,80	0,76	0,85	0,79
İçe Dönük	0,85	0,93	0,82	0,87
Dışa Dönük	0,62	0,70	0,44	0,51
Sakin	0,89	0,90	0,78	0,80

Tablo 5'e göre "tedirgin" özelliği için "Lojistik Regresyon" algoritmasının diğer özelliklere göre daha başarılı, "duygusal" ve "dışa dönük" özellikleri için daha başarısız olduğu görülmektedir.

Tablo 6. Naive Bayes algoritması için elde edilen sonuçlar

	Naive Bayes			
	Doğruluk	Kesinlik	Hatırlama	F1 Skoru
İyimser	0,59	0,41	0,72	0,51
Dengeli	0,59	0,82	0,54	0,64
Stresli	0,93	0,67	0,67	0,67
Sabırsız	0,68	0,83	0,52	0,61
Temkinli	0,64	0,84	0,64	0,70
Tedirgin	0,75	0,45	0,75	0,52
Duygusal	0,86	0,87	0,85	0,85
İçe Dönük	0,85	0,91	0,85	0,88
Dışa Dönük	0,68	0,83	0,41	0,48
Sakin	0,68	0,52	0,72	0,57

Tablo 6'ya göre "içe dönük" özelliği için "Naive Bayes" algoritmasının diğer özelliklere göre daha başarılı, "dengeli", "sabırsız", "temkinli", "tedirgin" ve "sakin" özellikleri için daha başarısız olduğu görülmektedir.

Tablo 7. Karar Ağacı algoritması için elde edilen sonuçlar

	Karar Ağacı			
	Doğruluk	Kesinlik	Hatırlama	F1 Skoru
İyimser	0,86	0,78	0,83	0,78
Dengeli	0,79	0,93	0,85	0,88
Stresli	0,89	0,76	0,83	0,70
Sabırsız	0,96	0,94	1,00	0,97
Temkinli	0,89	0,89	0,97	0,93
Tedirgin	0,82	0,63	0,64	0,57
Duygusal	0,93	0,97	0,89	0,92
İçe Dönük	0,78	0,83	0,94	0,87
Dışa Dönük	0,82	0,96	0,68	0,79
Sakin	0,84	0,70	0,77	0,73

Tablo 7'ye göre "içe dönük" özelliği için "Karar Ağacı" algoritmasının sonuçlarının diğer özelliklere göre biraz daha başarısız, "iyimser", "dengeli", "stresli", "temkinli", "duygusal" ve "dışa dönük" özellikleri için daha başarılı olduğu görülmektedir.

Tablo 8. Destek Vektör Makineleri algoritması için elde edilen sonuçlar

	Destek Vektör Makineleri			
	Doğruluk	Kesinlik	Hatırlama	F1 Skoru
İyimser	0,55	0,19	0,17	0,18
Dengeli	0,55	0,67	0,72	0,69
Stresli	0,89	0,78	0,67	0,61
Sabırsız	0,86	0,80	0,97	0,87
Temkinli	0,64	0,73	0,77	0,74
Tedirgin	0,86	0,62	0,83	0,68
Duygusal	0,80	0,76	0,85	0,80
İçe Dönük	0,85	0,96	0,80	0,87
Dışa Dönük	0,68	0,76	0,53	0,58
Sakin	0,88	0,85	0,78	0,78

Tablo 8'e göre "iyimser" özelliği için "Destek Vektör Makineleri" algoritmasının sonuçlarının diğer özelliklere göre daha başarısız olduğu görülmektedir.

5. SONUÇ

Uzman bir psikolog tarafından değerlendirilen ve Ek-1'de gösterilen kişilik testi 40 sorudan oluşmaktadır. Bu test profesyonel danışmanlar tarafından kişilik analizini belirlemek için kullanılmaktadır. Python üzerinde Open CV kütüphanesinin ilgili fonksiyonlarının el yazısı görüntüleri üzerine uygulanması ile Tablo 2'de verilen giriş özellikleri elde edilmiştir. Bu giriş özelliklerine karşılık gelen ve Tablo 3'te gösterilen çıkış kişilik özellikleri de eklenerek veri kümesi oluşturulmuştur.

Bölüm 4.1'de verilen makine öğrenmesi algoritmalarına veri kümesi uygulanmış ve kişilik özelliklerinden elde edilen kişilik analizi sonuçları ile ektaki testten elde edilen sonuçların bir karşılaştırılması yapılmıştır. Bunun için Python'da makine öğrenmesi algoritmalarının performansları karşılaştırılarak kişilik analizi için en iyi sonucu veren algoritma tespit edilmiştir.

Tablo 4, 5, 6, 7 ve 8'e bakıldığında Doğruluk, Kesinlik, Hatırlama ve F1 Skoru ölçütlerine göre genel ortalamaya bakıldığında; Naive Bayes algoritması en başarısız algoritma olarak bulunmuştur. Naive Bayes algoritmasını sırasıyla, Destek Vektör Makineleri ve Lojistik Regresyon algoritmaları takip etmiştir. Rastgele Orman ve Karar Ağacı algoritmaları ise yaklaşık oranda en başarılı algoritmalar olarak bulunmuştur.

Uzman psikolog tarafından verilen kişilik testi toplam 39 kişiye uygulanmıştır. Bu testlerden elde edilen sonuçlar ile el yazısı analizinden elde edilen sonuçlar karşılaştırıldığında %72 oranında eşleşme olduğu görülmüştür. Bu testlerden 28 tanesinin sonucu el yazısı analizinden elde edilen sonuçlarla aynıdır.

KAYNAKLAR

- Breiman, L. (2001) Random Forests. *Machine Learning*, 45: 5–32.
- Büyükoflaz, F. N. and Öztürk, A. (2018). "Early autism diagnosis of children with machine learning algorithms," 2018 26th Signal Processing and Communications Applications Conference (SIU), pp. 1-4, doi: 10.1109/SIU.2018.8404223.
- Champa, H. N. and AnandaKumar, K. R. (2010). "Automated Human Behavior Prediction through Handwriting Analysis," 2010 First International Conference on Integrated Intelligent Computing, pp. 160-165, doi: 10.1109/ICIIC.2010.29.
- Corinna, C. and Vapnik, V. (1995). "Support Vector Networks", *Machine Learning*. 20 (3): 273–297.
- Durga, L. and Deepu, R. (2018). "Handwriting Analysis Through Graphology: A Review," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1160-1166, doi: 10.1109/ICACCI.2018.8554416.
- Fatimah, S. H., Djamal, E. C., Ilyas, R. and Renaldi, F. (2019). "Personality Features Identification from Handwriting Using Convolutional Neural Networks," 2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), pp. 119-124, doi: 10.1109/ICITISEE48480.2019.9003855.
- Garoot, A. H., Safar, M. and Suen, C. Y. (2017). "A Comprehensive Survey on Handwriting and Computerized Graphology," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 621-626, doi: 10.1109/ICDAR.2017.107.
- "Machine Learning Mastery", erişim adresi: <https://machinelearningmastery.com/logistic-regression-for-machine-learning>, erişim tarihi: 19.04.2022.
- Mitchell, T. (1997). *Machine Learning*. McGraw-Hill.
- Nichol, M. A. (1991). *Handwriting analysis*, Mc Graw Hill.
- Prasetyo, K. A., Ramadijanti, N. and Basuki, A. (2017). "Mobile application for identifying personality of person using graphology," 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), pp. 212-219, doi: 10.1109/KCIC.2017.8228589.
- Quinlan, J.R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.
- Robertson, E.W. (1991). *Fundamentals of Document Examination*. Chicago, Nelson-Hall Publishers.
- Santana, O., Travieso, C. M., Alonso, J. B. and Ferrer, M. A. (2010). "Writer identification based on graphology techniques," in *IEEE Aerospace and*

Electronic Systems Magazine, vol. 25, no. 6, pp. 35-42, doi: 10.1109/MAES.2010.5525319.

Sheikholeslami, G. S., Srihari, N. and Govindaraju, V. (1997). *Center of Excellence for Document Analysis and Recognition State University of New York at Buffalo Amherst*, New York 14228-2567, USA.

“Support Vector Machines”, erişim adresi: <https://medium.com/@skilltohire/support-vector-machines-4d28a427ebd>, erişim tarihi: 10.02.2022.

Taheri, S. and Mammadov, M. (2013). Learning the naive Bayes classifier with optimization models. *International Journal of Applied Mathematics and Computer Science*, 23(4), pp.787-795.

Tolles, J. and Meurer, WJ. (2016) Logistic Regression: Relating Patient Characteristics to Outcomes. *JAMA*, 316(5):533–534. doi:10.1001/jama.2016.7653

Vásquez, J. L., Travieso, C. M. and Alonso, J. B. (2013). "Using calligraphies features for off line writer identification," 2013 47th International Carnahan Conference on Security Technology (ICCST), pp. 1-6, doi: 10.1109/CCST.2013.6922062.

Wu, Z., Lin, W., Zhang, Z., Wen, A. and Lin, L. (2017). "An Ensemble Random Forest Algorithm for Insurance Big Data Analysis," 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), pp. 531-536, doi: 10.1109/CSE-EUC.2017.99.

Yurtoğlu, H. (2005). Yapay Sinir Ağları Metodolojisi ile Önörü Modellemesi: Bazı Makroekonomik Değişkenler için Türkiye Örneği. Uzmanlık Tezi, Devlet Planlama Teşkilatı.

Zhang C., Cao, L. and Romagnoli, A. (2018). On the feature engineering of building energy data mining, *Sustainable Cities and Society*, 39: 508-518.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Hilal MÜSEVİTOĞLU

EĞİTİM DURUMU

Lisans Öğrenimi :2010-2014, Selçuk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü

Yüksek Lisans Öğrenimi :2018-2022, KTO Karatay Üniversitesi, Mühendislik Fakültesi, Elektrik – Bilgisayar Mühendisliği Bölümü

Bildiği Yabancı Diller :İngilizce

İŞ DENEYİMİ

Stajlar :2012, Stajyer Mühendis, AE Kod Teknolojisi

:2013, Stajyer Mühendis, Kuveyt Türk Bilgi Teknolojileri

:2014, Stajyer Mühendis, Kuveyt Türk Bilgi Teknolojileri

Çalıştığı Kurumlar :2014-2017, Yazılım Mühendisi, Bizpark Business Solutions

:2018-2019, Yazılım Mühendisi, Post Ajans A.Ş.

:2020- , Yazılım Mühendisi, Havelsan A.Ş.

EK 1. PSİKOLOG KİŞİLİK TESTİ

No :

EK 3- SIFATLARA DAYALI KİŞİLİK TESTİ

Açıklama: Aşağıda bireyleri tanımak için kullanılan sıfat çiftleri verilmektedir. Sizden istenen, her bir sıfat çiftini okuyarak size uygunluk derecesine karar vermenizdir. Her sıfat çifti için bir tek daireyi doldurunuz. Doğru cevap yoktur, size uygun cevap vardır. Bunu dikkate alarak cevaplamaya çalışınız. Cevaplarınızı aşağıdaki örneklere göre belirtiniz.

	Çok uygun	Oldukça uygun	Biraz uygun	Ne uygun, ne uygun değil	Biraz uygun	Oldukça uygun	Çok uygun	
İçedönük	0	●	0	0	0	0	0	Dışadönük
İçedönük	0	0	0	0	0	0	●	Dışadönük
İçedönük	0	0	0	●	0	0	0	Dışadönük

Bu örneklerde;

1. Ömek kişi kendini oldukça içedönük olarak görmektedir,
2. Ömek ise kişi kendini çok dışadönük olarak görmektedir,
3. Ömek ise kişi bu boyutlarda kararsızdır veya her iki sıfatı da kendine uzak veya yakın görmektedir, anlamına gelmektedir.

		Çok uygun	Oldukça uygun	Biraz uygun	Ne uygun, ne uygun değil	Biraz uygun	Oldukça uygun	Çok uygun	
1	Sakin	0	0	0	0	0	0	0	Sinirli
2	Yalnızlığı tercih eden	0	0	0	0	0	0	0	Sosyal (topluluğu seven)
3	Sanata ilgisiz	0	0	0	0	0	0	0	Sanata ilgili
4	Kindar	0	0	0	0	0	0	0	Affedici
5	Düzensiz	0	0	0	0	0	0	0	Düzenli
6	Sabırlı	0	0	0	0	0	0	0	Sabırsız
7	Silik	0	0	0	0	0	0	0	Atak
8	Hayal gücü zayıf	0	0	0	0	0	0	0	Hayal gücü kuvvetli
9	(Başkalarına kayıtsız)	0	0	0	0	0	0	0	Yardımsaver

10	Sorumsuz	0	0	0	0	0	0	0	Sorumluluk sahibi
11	Rahat	0	0	0	0	0	0	0	Tedirgin
12	Uyuşuk, eli ağır	0	0	0	0	0	0	0	Canlı
13	Dargörüştü	0	0	0	0	0	0	0	Geniş görüşlü
14	Rekabetçi	0	0	0	0	0	0	0	İşbirliği yapan
15	Hırslı değil	0	0	0	0	0	0	0	Hırslı
16	Tutarlı	0	0	0	0	0	0	0	Tatarsız
17	Durgun	0	0	0	0	0	0	0	Delidolu
18	Alışılmış	0	0	0	0	0	0	0	Yenilikçi
19	Kibirli	0	0	0	0	0	0	0	Alçakgönüllü
20	Dikkatsiz	0	0	0	0	0	0	0	Dikkatli
21	İyimser	0	0	0	0	0	0	0	Karamsar
22	Neşesiz	0	0	0	0	0	0	0	Neşeli
23	Meraksız	0	0	0	0	0	0	0	Meraklı
24	Asi	0	0	0	0	0	0	0	Uysal, yumuşakbaşlı
25	Gayretsiz	0	0	0	0	0	0	0	Gayretli
26	Huzurlu	0	0	0	0	0	0	0	Huzursuz
27	Arkaplanda kalan	0	0	0	0	0	0	0	Öne çıkan
28	Tutucu	0	0	0	0	0	0	0	Liberal
29	Acımasız	0	0	0	0	0	0	0	Merhametli
30	Hazırlıksız	0	0	0	0	0	0	0	Hazırlıklı
31	Kaygısız	0	0	0	0	0	0	0	Kaygılı
32	Dikkat çekmeyen	0	0	0	0	0	0	0	Baskm, belirgin
33	İlgileri dar	0	0	0	0	0	0	0	İlgileri geniş
34	Bencil	0	0	0	0	0	0	0	Fedakar (diğer gam)
35	Disiplinsiz	0	0	0	0	0	0	0	Disiplinli
36	Yeni ilişkilere kapalı	0	0	0	0	0	0	0	Yeni ilişkilere açık
37	Etkisiz	0	0	0	0	0	0	0	Etkili
38	Hoşgörüsüz	0	0	0	0	0	0	0	Hoşgörülü
39	Donuk	0	0	0	0	0	0	0	Coşkulu
40	İnatçı	0	0	0	0	0	0	0	Uzlaşmacı

EK 2. KULLANILAN KODLAR

1. Marjinlerin Bulunması

```
import numpy as np
import cv2
import os
import re
# construct the argument parse and parse the arguments
for root, dirnames, filenames in os.walk(".\\tez _metin_img\\"):
    for filename in filenames:
        if re.search("\.(jpg|jpeg|png|bmp|tiff)$", filename):
            filepath = os.path.join(root, filename)
            image = cv2.imread(filepath)

#gri seviye resme çevir
    gray = cv2.cvtColor(image,cv2.COLOR_BGR2GRAY)
    cv2.imwrite('gray.png',gray)

#siyah beyaz hale resme çevir
    ret,thresh = cv2.threshold(gray,180,255,cv2.THRESH_BINARY_INV)
    cv2.imwrite('binary.png',thresh)

#genişletme (kopuk pikselleri birleştirme)
    kernel = np.ones((5,100), np.uint8)
    img_dilation = cv2.dilate(thresh, kernel, iterations=1)
    cv2.imwrite('dilated.png',img_dilation)

#Kelime ve harflerin çevrelerini tespit etme
contours,
hierachy=cv2.findContours(thresh,cv2.RETR_TREE,cv2.CHAIN_APPROX_SIMPLE)

#sort contours
    sorted_ctrs = sorted(contours, key=lambda ctr: cv2.boundingRect(ctr)[0])

    min_x = 1000000
    max_x = 0
    min_y = 1000000
    max_y = 0

    for i, ctr in enumerate(sorted_ctrs):
# Get bounding box
        x, y, w, h = cv2.boundingRect(ctr)
        if x<min_x:
            min_x=x
        if x+w>max_x:
            max_x=x+w
        if y<min_y:
            min_y=y
        if y+h>max_y:
            max_y=y+
```

```

# Getting ROI
roi = image[y:y+h, x:x+w]

cv2.rectangle(image,(x,y),( x + w, y + h ),(90,0,255),2)

cv2.imwrite('final_bounded_box_image.png',image)
#cv2.imshow('marked areas',image)

height, width, channels = image.shape

print(filename, " ")
print("sol margin ", min_x, "sağ margin ", width-max_x, "üst marjin", min_y, "alt
marjin", height-max_y)

```

2. Yazı Basısının Bulunması

```

def rgb2gray(rgb):

    r, g, b = rgb[:, :, 0], rgb[:, :, 1], rgb[:, :, 2]
    gray = 0.2989 * r + 0.5870 * g + 0.1140 * b

    return gray

import cv2
img = cv2.imread("..\tez _metin_img\Metin-22-1.jpg")
gray = rgb2gray(img)

toplam = 0
height = img.shape[0]
width = img.shape[1]

for i in range (height):
    for j in range (width):
        toplam=toplam+gray[i][j]

avg_gray = toplam / (img.shape[0]*img.shape[1])
print("Ortalama Grilik", avg_gray)
'''

from PIL import Image

import os
import re
# construct the argument parse and parse the arguments
for root, dirnames, filenames in os.walk("..\tez _metin_img\"):
    for filename in filenames:
        if re.search("\.(jpg|jpeg|png|bmp|tiff)$", filename):
            filepath = os.path.join(root, filename)
            im = Image.open(filepath)
            im_grey = im.convert('LA') # convert to grayscale

```

```

width, height = im.size

total = 0
for i in range(0, width):
    for j in range(0, height):
        total += im_grey.getpixel((i,j))[0]

mean = total / (width * height)
print(filename, " ", mean)

```

3. Yazı Eğiminin Bulunması

```

# import the necessary packages
import numpy as np
import cv2
import os
import re
# construct the argument parse and parse the arguments
for root, dirnames, filenames in os.walk(".\\tez _metin_img\\"):
    for filename in filenames:
        if re.search("\.(jpg|jpeg|png|bmp|tiff)$", filename):
            filepath = os.path.join(root, filename)
            image = cv2.imread(filepath)

            gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
            gray = cv2.bitwise_not(gray)
            thresh = cv2.threshold(gray, 0, 255, cv2.THRESH_BINARY |
cv2.THRESH_OTSU)[1]

            coords = np.column_stack(np.where(thresh > 0))
            angle = cv2.minAreaRect(coords)[-1]
            if angle < -45:
                le = -(90 + angle)
            # otherwise, just take the inverse of the angle to make it positive
            else:
                le = -angle

            # show the output image
            print(filename, " [INFO] angle: {:.3f}".format(angle))

```

4. Kelime ve Satır Boşluklarının Bulunması

```

import matplotlib.pyplot as plt
import os
import re
# construct the argument parse and parse the arguments
for root, dirnames, filenames in os.walk(".\\tez _metin_img\\"):
    for filename in filenames:
        if re.search("\.(jpg|jpeg|png|bmp|tiff)$", filename):
            filepath = os.path.join(root, filename)
            img = cv2.imread(filepath)

```

```

img.shape
boyut = tuple(img.shape[1::-1])
height=0
width=0
height=boyut[0]
width=boyut[1]
img = prepare_img(img, height)
detections = detect(img, kernel_size=25, sigma=11, theta=7, min_area=100)
lines = sort_multiline(detections)

#plt.figure(figsize=(45,15))
#plt.imshow(img, cmap='gray')
num_colors = 7
colors = plt.cm.get_cmap('rainbow', num_colors)
total_word_spaces = 0
word_cnt = 0
total_line_spaces = 0
for line_idx, line in enumerate(lines):
    c = colors(line_idx % num_colors)
    line_y_avg = 0;
    line_word_cnt=0;
    det = line[0] # ilk word'u bulmak için
# print(det.bbox.x, det.bbox.x + det.bbox.w)
    previous_word_y = det.bbox.x + det.bbox.w
    for word_idx, det in enumerate(line):
        xs = [det.bbox.x, det.bbox.x, det.bbox.x + det.bbox.w, det.bbox.x +
det.bbox.w, det.bbox.x]
        ys = [det.bbox.y, det.bbox.y + det.bbox.h, det.bbox.y + det.bbox.h,
det.bbox.y, det.bbox.y]
        #plt.plot(xs, ys, c=colors(line_idx % num_colors))
        #plt.text(det.bbox.x, det.bbox.y, f'{line_idx}/{word_idx}')

# print(det.bbox.x, det.bbox.x + det.bbox.w) # kelime kutuları
    if line_word_cnt > 0:
        total_word_spaces = total_word_spaces + det.bbox.x - previous_word_y
        previous_word_y = det.bbox.x + det.bbox.w
        line_y_avg = line_y_avg + det.bbox.y
        line_word_cnt = line_word_cnt + 1
    word_cnt = word_cnt + line_word_cnt
    cizgi = line_y_avg / line_word_cnt
    cizgi = round(cizgi)
# print(line_idx, cizgi)
    if line_idx == 0:
        prev_line=cizgi
    else:
        total_line_spaces = total_line_spaces + (cizgi-prev_line)
        prev_line = cizgi
    #cv2.line(img, (0, cizgi), (width,cizgi), (60, 90, 80), 5)
#plt.show()
word_averages = round(total_word_spaces / word_cnt)
print(filename)
print("Average word spaces ", abs(word_averages))

```

```

#cv2.imwrite('.\\cizgiler.png',img)
total_line_spaces = round(total_line_spaces/(line_idx+1))
print("Average line spaces ", total_line_spaces)
colors = plt.cm.get_cmap('rainbow', num_colors)
total_word_spaces = 0
word_cnt = 0
total_line_spaces = 0
for line_idx, line in enumerate(lines):
    c = colors(line_idx % num_colors)
    line_y_avg = 0;
    line_word_cnt=0;
    det = line[0] # ilk word'u bulmak için
# print(det.bbox.x, det.bbox.x + det.bbox.w)

```

5. Makine Öğrenmesi Algoritmalarıyla Sonuçların Bulunması

```

import warnings
warnings.filterwarnings('ignore')

import pandas as pd
import numpy as np

def clean_dataset(df):
    assert isinstance(df, pd.DataFrame), "df needs to be a pd.DataFrame"
    df.dropna(inplace=True)
    indices_to_keep = ~df.isin([np.nan, np.inf, -np.inf]).any(1)
    return df[indices_to_keep].astype(np.float64)

data = pd.read_excel('./sample_data/Datalar_Kisilik.xlsx')

data = data.reset_index()
data = clean_dataset(data)

#print(data.head())
#data.info()
#print(data.describe())

from sklearn.ensemble import RandomForestClassifier
model = RandomForestClassifier(random_state=1)

modeller = [model]

from sklearn.linear_model import LogisticRegression
model = LogisticRegression(multi_class="multinomial", solver="lbfgs", C=10)

modeller.append(model)

from sklearn.naive_bayes import GaussianNB
model = GaussianNB()

modeller.append(model)

```

```

from sklearn import tree
model = tree.DecisionTreeClassifier()

modeller.append(model)

from sklearn.svm import SVC
model = SVC(kernel='poly', degree = 2)

modeller.append(model)

input_attributes =
["Angle", "Press", "Left", "Right", "Top", "Bottom", "Words", "Lines", "Page"]

output_attributes =
["iyimser", "dengeli", "stresli", "sabirsiz", "temkinli", "tedirgin", "duygusal", "icedonuk", "di
sardonuk", "sakin"]

def display_scores(measure, scores):
#             print("Scores:", scores)
#             print(measure, "Mean:", scores.mean())
#             print("Standard deviation:", scores.std())

i=1
for model in modeller:
print("\n=====")
if (i==1):
print("Random Forest")
if (i==2):
print("Logistic")
if (i==3):
print("Gaussian NB")
if (i==4):
print("Decision Tree")
if (i==5):
print("SVC")
print("=====")
i=i+1
for output in output_attributes:
X = data[input_attributes]

Y = data[output]

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler(feature_range=(0.1, 0.9))
scaled_train = scaler.fit_transform(X)

scaled_train = np.nan_to_num(scaled_train)

X=scaled_train

from sklearn.model_selection import cross_val_score

```

```
print("\nFeature is -->", output)
print("-----")

#print("Accuracy")
#print("-----")
accuracy_scores = cross_val_score(model, X, Y, scoring="accuracy", cv=3)
display_scores("Accuracy",accuracy_scores)

#print("Precision")
#print("-----")
precision_scores = cross_val_score(model, X, Y, scoring="precision", cv=3)
display_scores("Precision",precision_scores)

#print("Recall")
#print("-----")
recall_scores = cross_val_score(model, X, Y, scoring="recall", cv=3)
display_scores("Recall", recall_scores)

#print("F1_score")
#print("-----")
f1_scores = 2 * (precision_scores * recall_scores) / (precision_scores + recall_scores)
display_scores("F1_score",f1_scores)
```

ETİK KURUL

Evrak Tarih ve Sayısı: 30.12.2021-23642



T.C.
KTO KARATAY ÜNİVERSİTESİ REKTÖRLÜĞÜ
İnsan Araştırmaları Etik Kurulu

Sayı : E-46409256-300-23642
Konu : Dr. Öğr. Üyesi Ali Öztürk Hk.

30.12.2021

Sayın Dr. Öğr. Üyesi Ali ÖZTÜRK
Öğretim Üyesi

El Yazısı Analizi ile İnsanların Kişilik Özelliklerinin Tespit Edilmesi isimli ekte başvuru evrakları verilen araştırma projesi çalışmasının Dr. Öğr. Üyesi Ali ÖZTÜRK'un sorumluluğunda, Yardımcı Araştırmacı olarak Hilal MÜSEVİTOĞLU 'nun katılımları ile yürütülmesi ile ilgili İnsan Araştırmaları Etik Kurulumuza yapmış olduğunuz başvurunuz 30.12.2021 tarihli 2021/12/01 sayılı kararımızda değerlendirilmiştir. İlgili çalışmanızda Üniversitemizin adının geçmemesi, Covid-19 virüsü nedeniyle Ülkemizde yaşanan salgın sürecinde salgın için alınan kararlara uyarak ve araştırmanın yapılacağı kurum ve kuruluşlardan idari izin alınarak çalışmanız şartı ile kurulumuzca uygun bulunmuştur.

Çalışmalarınızda başarılar diler gereğini saygılarımla rica ederim.

Prof. Dr. Çağatay ÜNÜSAN
İnsan Araştırmaları Etik Kurul Başkanı

Ek:Dr. Öğr. Üyesi Ali ÖZTÜRK (12 sayfa)

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu :BS996ZTC6

Adres: Akabe Mahallesi Alaaddin Kapı Caddesi No:130 Karatay/Konya
Telefon: 444 1251 Faks: 0332 202 00 44
e-Posta: rektorluk@karatay.edu.tr Web: www.karatay.edu.tr
Kep Adresi: ktokaratayuniversitesi@is01.kep.tr

Belge Takip Adresi : <https://turkiye.gov.tr/ktokaratay-universitesi-ebys>

Bilgi için: Celaleddin ÇİBİK
Unvan: Sekreter

Tel No: 444 1251-7258

