T.C.

BATMAN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANABİLİM DALI

İSTENMEYEN E-POSTALARIN FİLTRELEMESİNDE AÇI DÖNÜŞÜMÜ TABANLI İÇERİK BAĞIMSIZ BİR YAKLAŞIM

YÜKSEK LİSANS

Tuncay ÖZER

Danışman
Doç. Dr. Yılmaz KAYA

Kasım-2023
BATMAN

TEZ KABUL VE ONAYI

Tuncay ÖZER tarafından hazırlanan “İstenmeyen e-postaların filtrelenmesinde açı dönüşümü tabanlı içerik bağımsız bir yaklaşım” adlı tez çalışması …/…/… tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Batman Üniversitesi Lisanüstü Eğitim Enstitüsü Bilgi Teknolojileri Anabilim Dalı’nda YÜKSEK LİSANS olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Başkan Unvanı Adı SOYADI	……..
Danışman Doç. Dr. Yılmaz KAYA	……..
Üye Unvanı Adı SOYADI	……..
Üye Unvanı Adı SOYADI	……..
Üye Unvanı Adı SOYADI	……..

Yukarıdaki sonucu onaylarım.

Doç. Dr. ……. ……..
Lisansüstü Eğitim Enstitüsü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Tuncay ÖZER

Tarih:

ÖZET

YÜKSEK LİSANS

İSTENMEYEN E-POSTALARIN FİLTRELEMESİNDE AÇI DÖNÜŞÜMÜ TABANLI İÇERİK BAĞIMSIZ BİR YAKLAŞIM

Tuncay ÖZER

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü
Bilgi Teknolojileri Anabilim Dalı

Danışman: Doç. Dr. Yılmaz KAYA

2023, … Sayfa

Jüri
Prof. Dr. Necmettin SEZGİN
Doç. Dr. Yılmaz KAYA
Doç. Dr. Fatma KUNCAN

Bu çalışmada, spam olarak tanımlanan istenmeyen e-postaların tespiti için geliştirilen açı yaklaşımı incelenmiştir. E-postaların metin içerikleri Unikodlara dönüştürülerek bir boyutlu sinyal olarak ele alınmış ve bu sinyal üzerindeki değerler arasındaki açı bilgileri hesaplanmıştır. Elde edilen açı sinyali, her e-postaya özgü bir histogram öznitelik vektörü olarak kullanılmıştır. Bu yaklaşımın başarısını test etmek amacıyla çeşitli makine öğrenimi yöntemleri kullanılmıştır, bunlar arasında Naive Bayes (NB), Karar Destek Vektörleri (SVM), K-En Yakın Komşu (Knn) ve Random Forest (RF) bulunmaktadır. Bu sınıflandırma işlemleri açık kaynak kodlu Weka programı ile gerçekleştirilmiştir ve 10-fold çapraz doğrulama ile değerlendirilmiştir. Sonuçlara bakıldığında Knn yöntemi ile %94,2'lik bir başarı elde edildiğini göstermektedir. Diğer yöntemler de kabul edilebilir başarılar göstermiştir. Ayrıca, uL ve uR gibi parametrelerin farklı değerlerinin kullanılması, açı yaklaşımının esnekliğini ve farklı örüntüler elde etme kapasitesini vurgulamıştır. Özellikle, uR=1 ve uL=1 parametre değerleri ile yüksek bir başarı elde edilmiştir. Ancak, bu parametrelerin farklı veri setlerinde farklı değerlendirmeler gerektirebileceği vurgulanmıştır. Bu çalışma ile, spam tespiti için karakterlerin Unikod değerleri arasındaki açı bilgilerini kullanan içerik bağımsız bir yaklaşımın etkili bir yol olduğunu ortaya koymaktadır. Bu yaklaşım, spam ile mücadelede geleneksel metin analizine alternatif bir yöntem sunmaktadır.

Anahtar Kelimeler: Açı Metodu, Derin Öğrenme, Destek Vektör Makinaları, Eposta, Makine öğrenmesi, Naive Bayes, Spam, Spam Filtre

ABSTRACT

MS THESIS

A CONTENT-INDEPENDENT APPROACH BASED ON ANGLE TRANSFORMATION FOR FILTERING SPAM EMAILS

Tuncay ÖZER

INSTITUTE OF GRADUATE STUDIES
OF BATMAN UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE
IN INFORMATION TECHNOLOGIES

Advisor: Assoc. Prof. Yılmaz KAYA

Year, … Pages

Jury
Advisor Prof. Dr. Necmettin SEZGİN
Assoc. Prof. Yılmaz KAYA
Assoc. Prof. Fatma KUNCAN

In this study, an angle-based approach developed for the detection of unwanted emails defined as spam has been examined. The text contents of the emails were transformed into Unicodes and treated as one-dimensional signals. Angle information between the values on this signal has been calculated. The obtained angle signal has been used as a histogram feature vector specific to each email. Various machine learning methods, including Naive Bayes (NB), Support Vector Machines (SVM), K-Nearest Neighbors (Knn), and Random Forest (RF), were used to test the success of this approach. These classification processes were conducted using the open-source Weka program and evaluated through 10-fold cross-validation. The results indicate that the Knn method achieved a success rate of 94.2%. Other methods also showed acceptable success rates. Furthermore, the use of different values for parameters such as uL and uR emphasized the flexibility of the angle approach in obtaining different patterns. Particularly, high success was achieved with parameter values uR=1 and uL=1. However, it was emphasized that these parameters might require different evaluations in various datasets. This study demonstrates that an angle-based approach using character Unicode values is an effective way for spam detection. This approach offers an alternative method to traditional text analysis in the fight against spam.

Keywords: Angle Method, Deep Learning, E-mail, Knn, Learning Machine, Naive Bayes, Spam, Spam Filter

ÖNSÖZ

Bu tezin oluşturulması esnasında içeriği ve yüzlerce ufak ayrıntının birbirleriyle ahenk içerisinde tezde yer almasında önemli katkılarda bulunan, bilgi, birikim ve engin tecrübeleri ile bana yol gösterici ve destek olan, çalışmalarımı yönlendiren ve bütün içtenliğiyle yardımcı olan değerli tez danışmanım sayın Doç. Dr. Yılmaz KAYA'ya en içten saygılarımı ve teşekkürlerimi sunuyorum.

Hiçbir zaman maddi ve manevi desteklerini esirgemeyen sevgili eşim Pervin ISLANMAZ ÖZER'e ve çocuklarım Toprak Efe, Muhammed Enes ve Mirsad Ali'ye teşekkür ediyor ve onlara minnettar olduğumu bir kez daha dile getiriyorum.

Tuncay ÖZER
BATMAN-2023

İÇİNDEKİLER

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	**Açıklama**
SPF	Sender Policiy Framework, Gönderen Politikası Çerçevesi
SMTP	Simple Mail Transfer Protocol, Basit Mail Gönderme Protokolü
POP3	Post Office Protocol, Postane Protokolü 3
IMAP	Internet Message Access Protocol, Internet Mesaj Erişim Protokolü
DKIM	Domain Keys Identified Mail - Etki Alanı Anahtarları Tanımlı Posta
DMARC	Domain-based Message Authentication, Reporting and Conformance - Etki Alanı Tabanlı Mesaj Kimlik Doğrulaması, Raporlama ve Uyumluluk
ISS	İnternet Servis Sağlayıcı
IP	Internet Protokol, İnternet Protokol
DCC	Direct Client Connection, Doğrudan İstemci Bağlantısı
DNS	Domain Name System, Alan Adı Sistemi
K-NN	K-Nearest Neighbors, K-en Yakın Komşu
SVM	Support Vector Machine, Destek Vektör Makinası
N-BAYES	Navie Bayes
URL	Uniform Resource Loader, Tekdüzen Kaynak Bulucu
SPAM	İstenmeyen e-posta, İstenmeyen e-mail
Non-SPAM	İstenen e-mail, spam olmayan e-mail
RF	Random Forest, Rasgele Orman
BD	Dirichlet Bayesian, Diric
BIC	Bayesian Information
ML	Machine Learning, Makine Öğrenmesi

ÇİZELGE LİSTESİ

ŞEKİL DİZİNİ

1. GİRİŞ

Günümüzde teknolojideki ilerlemelerin etkisiyle iletişim, yazılı bir format olan e-posta aracılığıyla gerçekleştirilmektedir. Bu tercihin arkasında, ekonomik açıdan uygun maliyeti yanı sıra yazılı iletişim kültürünün de etkisi bulunmaktadır. E-posta, günümüz yaşamında belirgin bir öneme sahiptir ve küresel ölçekte yaygın bir iletişim aracı haline gelmiştir. İnternet teknolojilerinin ilerlemesiyle birlikte, e-posta kullanımı dünya genelinde hızla artmış ve milyarlarca kişi tarafından benimsenmiştir. Bu noktada, e-postanın ticaretin temel bir unsuruna dönüştüğü gözlenmektedir (Laorden ve diğerleri, 2012). Ancak, e-posta iletişiminin bu yaygın kullanımı, bazı olumsuz sonuçları da beraberinde getirmektedir. Özellikle, istenmeyen postalar ve virüs saldırıları, e-posta iletişimi içinde potansiyel tehditler oluşturur. İstenmeyen postaların başında gelen spam, e-posta trafiğinin büyük bir bölümünü oluşturmakta ve günlük yaşantımızın ayrılmaz bir parçası haline gelerek, hem kullanıcılar hem de internet trafiği için önemli bir sorun haline gelmektedir. İstenmeyen postalar, genellikle alıcıların bu tür iletileri istemediği durumlar için gönderildiğinden, rahatsız edici bir uygulama olarak kabul edilirler. Bu e-postalar, alıcıların posta kutularını gereksiz yere doldurabilir, zaman kaybına neden olabilir ve ara sıra güvenlik risklerini beraberinde getirebilir. E-posta kullanıcılarının kaynaklarını tüketen, zamanlarını boşa harcatan, finansal ve duygusal açıdan zarara yol açabilen ve iletişim ağını meşgul eden bu tür istenmeyen e-postalar, bu nedenle kullanıcılar tarafından pek istenmemektedir (Carpinter ve Hunt, 2006; Guzella ve Caminhas, 2009). Bu bağlamda, istenmeyen postaların etkili bir şekilde engellenmesi ve kontrol altına alınması, e-posta iletişiminin verimli ve güvenli bir şekilde sürdürülmesi açısından büyük bir öneme sahiptir.

Spam e-postalar, genellikle düşük maliyetleri ve kontrol dışında gerçekleşen doğası nedeniyle, çeşitli ticari reklam ve pazarlama faaliyetlerini, kampanya ve duyuruları, yanıltıcı ve gerçek dışı vaatleri, politik ve ideolojik görüşlerin propagandasını, belirli bir konuda kamuoyu oluşturma girişimlerini, yasa dışı ve yarı yasa dışı faaliyetleri, servis duyurularını, kimlik avı saldırılarını, sahte ödeme sistemleri çağrılarını, talep edilmemiş kitlesel e-postaları ve tanıtım amaçlı iletileri içermektedir. Bu tür e-postaların gönderenlerinin, alıcı adreslerini kamuya açık kaynaklardan, müşteri listelerinden, haber gruplarından, web sitelerinden, sohbet odalarından, sosyal medya bilgilerinden ve kullanıcı adres defterlerinden topladığı gözlemlenmektedir. Aynı zamanda, çeşitli

otomatik programlar ve tahminler kullanarak alıcı bilgilerini toplamaktadırlar. Bu, kullanıcıların izni olmadan kişisel bilgilerin toplanmasını içerir ve bazen zararlı yazılımlar ve virüsler aracılığıyla bilgisayar sistemlerine sızılarak bilgi çalınmasına neden olabilir. Spam e-postalar, alıcıları rahatsız eden, zamanlarını israf eden, maddi ve manevi kayıplara neden olan ve iletişim ağını meşgul eden bir yapıya sahiptir. Kullanıcılar, bu istenmeyen e-postaların neden olduğu sorunları deneyimledikçe, bu tür iletileri kabul edilemez bulmaktadır. Özellikle kişisel kullanıcılar için, spam ve spam olmayan e-postaların ayırt edilmesi zaman alıcı bir süreç olabilir ve sınırlı bant genişliği kaynaklarının gereksiz tüketilmesine yol açabilir. Bu dezavantajları ele almak ve spam e-postaların etkilerini en aza indirmek için teknik ve yasal önlemler alınmaktadır. Ancak, spam e-postalar genellikle para kazanma fırsatları, kilo verme ürünleri, iş kuluçka hizmetleri, arkadaşlık platformları, yetişkin eğlence ürünleri ve hızlı para kazanma vaatleri gibi konuları içerdiği için bu sorun gün geçtikçe daha da ciddi bir hale gelmektedir (Çıltık ve Güngör, 2008).

Bu nedenle, spam e-postalarını tanımlamak, tespit etmek ve filtrelemek amacıyla kullanılabilecek çeşitli tekniklerin geliştirilmesi, hala popüler bir araştırma alanını oluşturmaktadır. Spam e-postaların tehdidine karşı koymak için bir dizi yaklaşım benimsenmiş ve bazıları, dünya genelindeki internet kullanıcılarına yönelik spam saldırılarının yoğunluğunu önemli ölçüde azaltmıştır. Bu yöntemlerden bazıları, gelen e-postaların adreslerini kara listeye alırken, diğerleri e-posta içeriğini belirli anahtar kelimelere göre incelemektedir. Ancak, anahtar kelime tabanlı yöntemler, spam tespitinde tatmin edici sonuçlar vermekle birlikte, her zaman yeterli olmamıştır. Bu bağlamda, bağlamsal spam e-posta tespit yöntemleri, bir kelimenin (w1) bir mesajdaki başka bir kelimenin (w2) görünmesini nasıl etkilediği üzerine odaklanmaktadır. Yani, w1 kelimesinin mesajda bulunup bulunmaması, bir iletiyi spam olarak sınıflandırmada w2 kelimesine dayalı bir sınıflandırmayı nasıl etkileyeceğini belirlemektedir. Temel amaç, spam olmayan iletileri spam iletilerinden ayırmak ve tek taraflı spam iletişimini mümkün olduğunca engellemektir. Spam tespiti, aslında bir sınıflandırma problemi olarak ele alınmaktadır, yani gelen e-postaları "spam" veya "spam değil" olarak sınıflandırmak gerekmektedir. Bu sınıflandırma problemine çözüm üretmek amacıyla bir dizi farklı yöntem denenmiştir ve bu yöntemlerin geliştirilmesi, spam e-postalarının tanımlanması ve filtrelenmesi konusundaki araştırmaların önemli bir parçasını oluşturur (Nakov ve Dobrikov, 2004).

Başlangıçta, spam e-postaların önlenmesi konusunda, organizasyonların yapısı ve teknolojinin gelişimi doğrultusunda geleneksel yöntemlerle spam göndericileriyle mücadele edilmiştir. Ancak, bu geleneksel yöntemlerin istenen sonuçları elde edememesi, farklı tekniklerin keşfedilmesi gerekliliğini ortaya koymuştur. İletişimde kullanılan e-posta standardı, başlangıçta SMTP veya "Simple Mail Transfer Protocol" adı verilen basit bir ileti aktarma protokolü olarak tanımlandı. SMTP, oldukça basit bir protokoldür ve iletileri herhangi bir türden bir filtrelemeye tabi tutmadan gönderir. Bu basitlik, hem iyi niyetli hem de kötü niyetli kullanıcılar tarafından istismar edilmektedir. Daha sonra POP3 ve IMAP gibi protokollerin geliştirilmesi ve sunucu standartlarının kullanılmaya başlanması, e-posta iletişiminin daha güvenli hale getirilmesine yardımcı olmuştur. Ancak, spam e-postaların önlenmesi için bu tür teknikler yeterli olmamıştır. Özellikle geçmiş yıllarda, klasik kurumsal ve kişisel güvenlik önlemleri kullanılarak spam e-postaların engellenmeye çalışıldığı görülmüştür. Bu geleneksel önlemler arasında kişisel bilgisayarlarda antivirüs yazılımı kullanımı, spam e-postaların engelleme listelerine eklenmesi, e-posta adreslerinin dijital ortamlarda açıkça paylaşılmaması, gelen e-postaların içeriklerinin dikkatli bir şekilde incelenmesi ve ilgisiz e-postaların açılmaması gibi yöntemler bulunmaktadır. Kurumsal düzeyde ise, hosting firmalarından özel spam ayarlarının talep edilmesi, kendi web sunucusunun kurulması veya bu hizmetin satın alınması, spam filtreleme sistemlerine sahip hosting hizmeti sunan firmaların tercih edilmesi, domain alınırken whois (kimlik) bilgilerinin gizlenmesi, web sitelerinde güvenlik önlemlerinin kullanılması (örneğin, güvenlik kodları, robotların tespiti için önlemler, bal kapları, beyaz listeler, kara listeler, gri listeler, itibar tabanlı filtreler, optik karakter tanıma filtreleri) gibi çözümler denendi. Ancak, spam e-postaların giderek daha karmaşık hale gelmesi, bu geleneksel önlemlerin yetersiz kaldığını göstermektedir (Bhowmick ve Hazarika, 2016).

Yapay zekâ teknolojisinin kullanılmadığı sistemler, genellikle bağımsız yazılımlar veya çevrimiçi tabanlı çözümler gibi farklı platformlarda yaygın olarak kullanılan spam e-postaları engelleme çerçeveleri sunarlar. Bu çözümler arasında sunucu yetkilendirme kimlik doğrulama sistemleri, işbirlikçi yöntemler, sezgisel filtreleme teknikleri ve içeriğe dayalı yaklaşımlar bulunmaktadır (Karim ve ark., 2019). Özellikle, SPF (Sender Policy Framework - Gönderen Politika Çerçevesi), DKIM (Domain Keys Identified Mail - Etki Alanı Anahtarları Tanımlı Posta) ve DMARC (Domain-based Message Authentication, Reporting & Conformance - Alan Adı Esaslı İleti Kimlik

Doğrulaması, Raporlama ve Uyumluluk) gibi önemli e-posta kimlik doğrulama ve güvenliği standartları, gönderenin kimliğini doğrulamanın ve ISS'lerin, e-posta hizmet sağlayıcılarının ve gönderenlerin gerçekten yetkili olduğu diğer alıcı posta sunucularına bildirilmesinin bir yolunu sunar. Bu üç yöntem, kurulumları doğru bir şekilde yapıldığında, gönderenin meşru olduğunu, kimliklerinin taklit edilmediğini ve başka birinin adına e-posta göndermediğini kanıtlar. DKIM, tüm giden e-postaların üstbilgisine şifreli bir imza ekler. Bu imzalı iletileri alan posta sunucuları, ileti üstbilgisinin şifresini çözmek ve gönderildikten sonra ileti içeriğinin değiştirilmediğini doğrulamak için kullanır. SPF ise, e-posta gönderenin hangi IP adreslerinden gönderebileceğini belirtir. SPF, dolandırıcıların e-postaları başka birinin adına dağıtmasını, yalnızca yetkilendirilmiş bir IP adresinden geldiğini doğrulayarak engeller. Alan adı esaslı ileti kimlik doğrulaması, raporlama ve uyumluluk anlamına gelen DMARC, alanınızın şüpheli e-postaları nasıl ele alacağını belirler. DMARC, gönderen alanın e-posta trafiğini nasıl ele alacağını belirler ve alanınıza gelebilecek şüpheli e-postaları tanımlar. Bu kimlik doğrulama ve güvenlik standartları, spam e-postaların sınırlı tutulmasına ve daha güvenli bir e-posta iletişimi sağlanmasına yardımcı olur (Karim ve ark., 2019; Hameed ve ark., 2013).

SPF ve DKIM gibi kimlik doğrulama yöntemleri geniş bir kabul görmüş olsa da bu tür yöntemlerin sorunları da göz ardı edilmemelidir. Özellikle, şifreleme işlemleri, zaman zaman e-posta sunucularının performansını düşürebilir ve bu da iletişim hızını etkileyebilir. Bu tür kriptografik hatalar veya zayıflıklar, güvenliği zayıflatabilir ve sorunlara yol açabilir. Sunucu yetkilendirme ve kimlik doğrulama sistemlerinin yanı sıra, işbirlikçi modellere dayalı yaklaşımlar da yaygın bir şekilde kullanılmaktadır. İşbirlikçi modeller, mesajların alınması ve değerlendirilmesi aşamasında birden fazla kullanıcının iş birliği yapmasını içerir. Bu modeller, kararların önceden tanımlanmasını, kayıt altına alınmasını ve sorgulanmasını içeren bir süreç yürütür. İstenmeyen e-postaların filtrelenmesinde işbirlikçi modeller, kriptografik hash, bulanık hash, DCC (Distributed Checksum Clearinghouse), gri liste, DNS kara liste-beyaz liste ve sosyal güven temelli çözümler gibi çeşitli yaklaşımları kullanır. Bu yöntemler, kullanıcıların bir arada çalışması ve istenmeyen e-postaların daha etkili bir şekilde tespit edilmesine yardımcı olur (Francisco ve diğerleri, 2012). Bu çeşitli yöntemler, spam e-postaların engellenmesi ve e-posta iletişiminin güvenliğini artırma amacıyla kullanılan önemli araçlardır. Mevcut anti-spam teknikleri, genellikle iki büyük kategori altında sınıflandırılabilir: önleyici ve

iyileştirici önlemler. Önleyici çözümler, istenmeyen e-posta gönderenleri caydırarak istenmeyen e-postaların gönderilmesini engellemeyi amaçlar. İyileştirme önlemleri ise istenmeyen e-postaları gönderildikten sonra algılamak, silmek veya engellemek için tasarlanmıştır. Her iki grup da farklı yaklaşımları içerir ve iletişim süreci açısından farklı avantajlara sahiptir. Önleyici önlemler, anti-spam yasalarını, e-posta protokolü değişikliklerini, ödemeye dayalı sistemleri ve meydan okuma yanıt (Captcha) filtrelerini içerebilir. Bu tür çözümler, istenmeyen e-postaların üretilmemesini sağlamayı amaçlar. Bu, iletişim sürecinin optimal olmasını sağlar çünkü istenmeyen e-postalar oluşturulmaz ve dolayısıyla iletmeye, depolamaya veya istenmeyen e-postayı işleme gibi süreçler gerektirmez. Ancak, bu tür önleyici önlemlerin etkili olabilmesi için genellikle ülkeler veya ISS'ler (Internet Service Providers) ile kullanıcılar arasında geniş bir fikir birliği gerekmektedir. İyileştirici önlemler, e-postaların aktarım sırasında veya varış noktasında istenmeyen e-postaları filtrelemeyi amaçlar. Bu tür çözümler, ileti sınıflandırma stratejisine bağlı olarak iki ana kategoriye ayrılabilir. Bazıları makine öğrenimi (ML) tekniklerine dayanırken, diğerleri çeşitli sınıflandırma stratejilerini içerir. Bu tür yöntemler, e-postaların içeriğine dayalı olarak spam veya spam olmayan sınıflarına ayrılmasını sağlar. Özellik çıkarımı bu süreçte kritik bir rol oynar çünkü yanlış özellik seçimi sınıflandırma doğruluğunu olumsuz etkileyebilir. İstatistiksel yöntemler, kelime frekansı temel alınarak kelime geçişlerine dayanır. Makine öğrenimi yöntemleri ise başlangıç hipotezine ihtiyaç duymadan spam e-postalarını filtreleme konusunda başarılı olmuşlardır. Bu nedenle, son yıllarda daha yaygın bir şekilde kullanılmaktadır. Makine öğrenimi yöntemleriyle spam e-postalarının tespit veya filtreleme süreci, e-postaların içeriğine göre spam veya spam olmayan sınıflarına ayrılmasını içerir. Özelliklerin seçimi bu sürecin önemli bir bileşenidir, çünkü yanlış özellik seçimi sınıflandırma doğruluğunu etkileyebilir (Su ve ark., 2012; Lin, 2009; Idris ve ark., 2014).

Bu tez çalışmasının temel motivasyonu, bağlamsal ve istatistiksel temelli özellik çıkarım yöntemlerinin sunduğu avantajların, makine öğrenimi yöntemlerinin sınıflandırma yetenekleri ile birleştirilebileceği yeni bir yaklaşımla spam tespitini gerçekleştirmektir. Bu çalışmada, bir e-postanın spam olup olmadığını belirlemek için karakterler arasındaki açı bilgisine dayalı etkili bir yaklaşım kullanılmıştır, ki bu yaklaşım açı dönüşüm örüntülerini kullanır. Açı dönüşümü örüntüleri, e-postalardan temsil edici özellikler çıkarmak için istatistiksel bir yöntemdir. Önerilen yaklaşım, karakterlerin Unicode değerlerini kullanır ve bunları sıralanmış karakterler arasındaki açı bilgilerini

çıkararak bu bilgileri öznitelik vektörleri halinde makine öğrenimi yöntemlerine sunar. Açı dönüşümü, özellik çıkarımının hızlı ve etkili bir şekilde gerçekleştirilebilmesinin önemli bir avantajını sunar. Ayrıca, hesaplama basitliği açısından da büyük bir avantaja sahiptir, bu da gerçek zamanlı metin işleme uygulamalarında kullanılabilmesini mümkün kılar. Bu yeni yaklaşım, spam tespiti konusunda geleneksel yöntemlere göre daha etkili ve hızlı sonuçlar elde etmek amacıyla geliştirilmiştir. Bu çalışma, spam e-postaları tanımlama ve filtreleme konusunda önemli bir katkı sağlamayı amaçlamaktadır.

1.1. Amaç

Bu tez çalışmasının temel amacı, spam e-postaların etkili bir şekilde tespit edilmesini sağlayacak bir metin analizi yaklaşımı geliştirmektir. Bu geliştirilen yaklaşım, e-postaların içeriklerine odaklanmaksızın çalışabilen içerik bağımsız bir yöntem olarak konumlandırılmıştır. Önerilen metot, istatistiksel özelliklerin çıkarılması yoluyla spam tespiti gerçekleştirir ve bu özellikler, geleneksel makine öğrenme yöntemleri ile analiz edilir.

1.2. Kapsam

Bu çalışma kapsamında, spam e-postaların tespiti için farklı yaklaşımlar incelenmiş, özellikle açı dönüşümü yöntemi ve çeşitli sınıflandırma algoritmaları (Random Forest, Knn, SVM, Naive Bayes) üzerinde detaylı bir inceleme gerçekleştirilmiştir. Mevcut literatürdeki benzer çalışmaların incelenmesi ve bu çalışmanın bu literatüre katkı sağlaması amaçlanmıştır. İçerik bağımsız bir yaklaşım olan açı dönüşümü tekniği kullanılarak, e-postaların içeriğinden bağımsız bir şekilde spam tespiti hedeflenmiş ve bu amaç doğrultusunda elde edilen sonuçlar detaylı bir şekilde incelenmiştir. Bu çalışma, spam tespiti konusunda yeni bir perspektif sunmayı ve mevcut yaklaşımlarla karşılaştırmayı amaçlamaktadır. Ayrıca, spam e-postaların etkili bir şekilde sınıflandırılması için istatistiksel ve makine öğrenme yöntemlerinin birleştirilmesinin potansiyelini araştırarak, bu alandaki araştırmalara önemli bir katkı sağlamayı hedeflemektedir. Sonuç olarak, bu çalışma, dijital iletişimde spam tespiti ve filtreleme konusundaki önemli bir problemi ele alarak, yeni bir yaklaşım sunma amacını taşır.

2. KAYNAK ARAŞTIRMASI

Bu bölümde, istenmeyen e-postaların filtrelenmesi için gerçekleştirilen çeşitli çalışmaların özetleri sunulmuştur. Bu çalışmalar, genellikle iki temel aşamadan oluşan bir yaklaşımı benimsemektedir. İlk aşamada, e-postalardan öznitelik çıkarımı işlemleri gerçekleştirilirken, ikinci aşamada ise çıkarılan öznitelikler kullanılarak makine öğrenmesi metotları ile sınıflandırma işlemi gerçekleştirilmektedir.

Singh ve Batra (2018) yaptıkları çalışmada, Sosyal Nesnelerin İnterneti'nde (Social IoT) istenmeyen e-posta tespiti için bir yarı denetimli teknik önermek amacıyla olasılık veri yapılarını kullanmıştır. Bu çalışmada, Twitter'da spam tespiti için dört sınıflandırıcıdan oluşan topluluk tabanlı bir çerçeve kullanılmıştır. Çerçeve, URL veritabanı, spam kullanıcıları, istenmeyen e-posta kelimeleri veritabanları ve benzerlik araması için olasılıksal veri yapıları (PDS) olarak Quotient Filter (QF) ve Locality Sensitive Hashing (LSH) gibi hızlı sonuçlar sağlayan ve az bilgi işlem çabası gerektiren yapıları içermektedir. Çalışmanın performansı, benzer veri yapılarıyla karşılaştırmalı analiz ve hassasiyet, geri çağırma ve F-skor gibi standart değerlendirme parametreleri kullanılarak değerlendirilmiştir (Singh ve Batra, 2018).

Makkar ve Kumar (2021) yaptıkları çalışmalarında, görüntü spam tespiti ve önlemi için optimize edilmiş bir derin öğrenme tabanlı yöntem geliştirmişlerdir. Bu çalışma, geleneksel metin tabanlı istenmeyen e-posta filtrelerini atlatmak amacıyla istenmeyen e-posta metin içeriğini grafiksel görüntülere entegre eden bir istenmeyen e-posta tespiti tekniğine karşı görüntü istenmeyen e-postalarını etkili bir şekilde tespit etmek için görüntü verilerini analiz etmiştir. Çalışmada, görüntü istenmeyen e-posta tespiti ve önlemi için görüntünün bağlantı özellikleri analiz edilmiştir. Bu amaçla, "Protector" olarak adlandırılan optimize edilmiş bir yöntem geliştirilmiştir. Protector, görüntünün bağlantı bilgisi, metin bilgisi ve meta verileri kullanılarak bir sıralama puanı oluşturur. Bu sıralama puanı, bir görüntünün ilgisini gösterir ve yapay sinir ağı eğitimi için doğru öğrenme yönteminin tasarımında kullanılır. Çalışma, çeşitli performans değerlendirme metrikleri kullanılarak detaylı bir şekilde değerlendirilmiştir (Makkar ve Kumar, 2021). Bu çalışmalar, istenmeyen e-postaların tespiti ve filtrelenmesi konusunda çeşitli yöntemlerin ve tekniklerin kullanılabileceğini göstermektedir. İstenmeyen e-postaların tespiti, spam e-postaların hızla arttığı bir çağda, dijital iletişimde önemli bir

sorun olarak karşımıza çıkmaktadır. Bu nedenle, farklı disiplinlerden gelen araştırmacılar, bu alandaki yeni ve etkili çözümler geliştirmek için çalışmalarını sürdürmektedirler.

Salcedo-Campos ve ark. (2012) tarafından gerçekleştirilen istenmeyen e-posta tespiti çalışması oldukça dikkat çekici ve etkili bir yaklaşım sunmuştur. Bu çalışmada, sadece e-posta başlıklarında bulunan bilgilere dayalı yeni bir istenmeyen e-posta filtreleme tekniği benimsenmiştir. İşte bu yaklaşımın detayları: Bu yaklaşımda, e-posta başlıkları dinamik bir süreç sonucunda üretilen karakterler olarak kabul edilmiştir. Bu karakterler, sinyal olarak kabul edilmiş ve standart sinyal ön işleme tekniklerine uygun bir şekilde parametrelendirilmiştir. Bu parametreler, başlık bilgilerinden çıkarılmıştır.Daha sonra, istenmeyen e-posta tespiti sistemi için Gizli Markov Modelleri (HMM'ler) kullanılmıştır. HMM'ler, başlıkların belirli özelliklerini ve karakteristiklerini modellemek için kullanılan istatistiksel bir yöntemdir. Bu sayede, istenmeyen e-postalarının belirli desenleri ve davranışları tanımlanmış ve tespit edilebilmiştir. Bu çalışma, sadece başlık bilgileri üzerinden ilerleyerek istenmeyen e-posta tespiti yapma konusunda yeni bir yaklaşım sunmuştur ve HMM'lerin etkin kullanımıyla başarılı sonuçlar elde edilmiştir. Elde edilen performans, istenmeyen e-posta filtreleme için kullanılan diğer desen sınıflandırma paradigmaları ile karşılaştırılarak değerlendirilmiştir. Salcedo-Campos ve ekibinin çalışması, SpamAssassin, TREC05 ve CEAS 2008 Lab Değerlendirmesi için elde edilen deneysel sonuçları iyileştirmiş ve yalnızca başlık bilgilerinden bilgi kullanmanın ve e-postanın hangi dilde yazıldığından bağımsız olmanın ek avantajlarıyla % 98,42'lik bir istenmeyen e-posta tespit başarısı elde etmiştir (Salcedo-Campos, ve ark. 2012). Bu çalışma, istenmeyen e-postaların tespiti için geleneksel metin içeriği analizinden farklı bir yaklaşım sunarak, istenmeyen e-posta filtreleme alanında yeni bir perspektif sunmuş ve önemli bir başarı elde etmiştir.

Laorden ve ark. (2012) yaptıkları çalışmada, istenmeyen e-postaların filtrelenmesinde anormallik tespiti temelli bir yaklaşımın etkinliğini incelemiştir. Bu çalışma, etiketleme gerekliliğini azaltan ve yalnızca meşru veya istenmeyen e-posta gibi tek bir sınıfın temsiline dayanan bir anormallik temelli istenmeyen e-posta filtreleme sistemi geliştirmiştir. Ayrıca, bu sistem, etiketli veri kümesine veri azaltma algoritması uygulayarak işlem süresini azaltırken tespit oranlarını koruma yeteneği sunmuştur. Ayrıca, meşru e-postaların veya istenmeyen e-postaların temsilinin uygunluğu analiz edilmiştir (Laorden ve ark. 2012). Özetle, bu çalışma, anormallik tespiti temelli bir

yaklaşımın istenmeyen e-postaların etkili bir şekilde filtrelenmesi için kullanılabilirliğini vurgulamış ve etiketleme gerekliliğini azaltarak sürecin daha verimli hale getirilmesine katkıda bulunmuştur.

Xu ve ark. (2015) yılında yaptıkları çalışmada ise güvenli ve akıllı otonom çoklu robot sistemleri için görüş sahtekarlığı tespitine odaklanmıştır. Bu çalışma, görüş sahtekarlığı tespiti için topluluk keşfi yaklaşımını kullanmayı amaçlamış ve tahmin hassasiyetini artırmayı hedeflemiştir. Bu amaçla, "SPClique" adını verdikleri yeni bir yaklaşım önerilmiştir. SPClique yaklaşımı, inceleme veri kümesini yansıma grafiği olarak modellenen bir topluluk oluşturma yöntemi olan Clique Percolation Method (CPM) temel alarak görüş sahtekarlığı gruplarını tespit etmektedir. Bu süreçte, yaklaşık hesaplamaları tanıtarak hesaplama gücünü genişleten bir metodoloji benimsenmiştir.Önerilen yöntem, görüş sahtekarlığı gruplarını tespit etmek ve şüphelilik düzeyini ölçmek için grup tabanlı ve bireysel tabanlı spam göstergelerini kullanmaktadır. Sonuç olarak, görüş sahtekarlığı gruplarının şüpheli sıralamasını çıkartmaktadır.

Xu ve ekibinin çalışması, tahmin hassasiyeti açısından diğer karşılaştırma yöntemlerini geride bırakmış ve özellikle büyük ölçekli inceleme veri kümelerinde daha fazla gerçek görüş sahtekarını tespit etmiştir. Bu iki çalışma, istenmeyen e-postaların filtrelenmesi ve görüş sahtekarlığı tespiti gibi güncel güvenlik sorunlarına yönelik yenilikçi ve etkili yaklaşımlar sunarak bilim dünyasına önemli katkılarda bulunmuştur.

Sokhangoee ve Rezapour (2022) yılında yaptıkları çalışmada, Çevrimiçi Sosyal Ağlar (CSA) platformlarında spam tespiti için ilgi çekici bir yaklaşım sunmaktadır. Bu çalışma, denetimli bir yöntem kullanarak spam tespiti konusunda iki temel faktörün etkisini araştırmıştır. Bu faktörler, istenilen özelliklerin seçimi ve uygun bir sınıflandırıcı kullanımıdır. İlk faktör için benimsenen yenilikçi bir yöntem, ilişkilendirme kuralı madenciliği ve genetik algoritma kombinasyonunu kullanmaktadır. Bu yöntem, çeşitli özellikler arasından istenilen özelliklerin seçilmesini kolaylaştırmaktadır. İkinci faktörde ise bir dizi popüler sınıflandırıcı kullanılmıştır. Önerilen yöntem, üç farklı veri kümesi üzerinde değerlendirilmiş ve sonuçlar, önerilen özellik seçimi yönteminin sınıflandırıcıların doğruluğu üzerinde olumlu bir etkisi olduğunu göstermektedir. Bu çalışmada kullanılan sınıflandırma algoritmaları J48, K-NN, Naive Bayes, Random Forest, Random Tree ve Neural Network (Yapay Sinir Ağı) gibi çeşitlilik göstermektedir. Ayrıca, derin öğrenme sınıflandırmalarında Long Term Memory (LSTM) da

kullanılmıştır. Üç farklı veri kümesi kullanılarak çeşitlilik sağlanmış ve yöntemin farklı senaryolardaki performansı değerlendirilmiştir. SAC'13 veri kümesi, konum tabanlı sosyal medya verilerini içermekte ve 2762 kayıttan oluşmaktadır. Spambase veri kümesi ise 4601 kayıttan oluşmakta ve spam tespiti için geniş bir veri havuzunu temsil etmektedir. ICC veri kümesi ise Twitter sosyal ağına ait verileri içerir ve 10.000 kayıt içerir.İlişki kuralları madenciliği ve genetik algoritmaların kullanıldığı bu çalışma, özellik seçimi ve sınıflandırma açısından oldukça etkili sonuçlar sunmuş, her iki yaklaşımın ortalama doğruluk oranlarının sırasıyla %87,99 ve %95,24 olduğunu göstermiştir. Bu sonuçlar, spam tespiti alanında yeni bir yöntem arayan araştırmacılara önemli bir rehberlik sağlamaktadır (Sokhangoee ve Rezapour, 2022).

Saidani ve ark. (2020) tarafından geliştirilen spam tespiti yöntemi, anlam temelli bir sınıflandırma yaklaşımı ve iki farklı anlamsal düzey analizi üzerine kurulmuştur. Bu yaklaşım, istenmeyen e-postaların tespiti için e-postaları belirli kategorilere (örneğin sağlık, eğitim, finans vb.) göre sınıflandırarak ayrı kavramsal görünümler elde etmeyi amaçlamaktadır. İlk seviyede, her alandaki istenmeyen e-postalar için ayrı kavramsal görünümler oluşturulmuş ve bu görünümler, e-postaların içeriğini belirli kategorilere göre gruplandırarak elde edilmiştir. İkinci seviyede ise, her etki alanında spam tespiti için belirlenmiş ve otomatik olarak ayıklanmış semantik özellikler bir araya getirilmiştir. Bu özellikler, e-postaların içeriğini etkili bir şekilde istenmeyen e-postalardan ayıran ve özetleyen semantik konuları hedeflemektedir. Önerilen yöntem, mevcut kelime torbası (BoW) ve semantik içeriğe dayalı yöntemlere göre daha iyi bir istenmeyen posta tespiti sağladığı ve daha anlaşılır sonuçlar elde edildiği gösterilmiştir. Yapılan deneysel sonuçlar, K-NN, Naive Bayes, Karar Ağacı, AdaBoost ve Random Forest gibi sınıflandırma yöntemleri kullanılarak değerlendirilmiş ve bu yöntemler arasında Naive Bayes ve AdaBoost'un ortalama olarak %98 başarı elde ettiği görülmüştür. (Saidani ve ark., 2022). 2021 yılında Neisari ve ark. tarafından gerçekleştirilen spam incelemeleri tespiti çalışmasında, spam incelemelerini gerçek incelemelerden ayırt etmek için dil tabanlı özellikler kullanarak yeni bir yaklaşım sunulmuştur. Bu çalışmada, evrişimli sinir ağları (Convolutional Neural Networks veya CNN) ve kendiliğinden örgütlenmiş haritalar (Self-Organizing Maps veya SOM) kullanılarak unsupervised öğrenme yöntemiyle incelemelerin sınıflandırılması amaçlanmıştır. İncelemeler, semantik olarak benzer kelimelerin bir pikselin etrafına veya bir SOM ızgara hücresine düzenlenmesiyle görüntülere dönüştürülmüştür. Bu dönüşüm süreci, incelemelerin görsel bir temsiline

dönüştürmektedir. Elde edilen inceleme görüntüleri daha sonra CNN'e beslenerek gözetimli eğitim süreci gerçekleştirilmiş ve ardından sınıflandırma için kullanılmıştır. Bu çalışma, tek ve çok alan bağlamlarında önerilen yöntemin etkinliğini değerlendirmek için iki altın standart veri kümesi üzerinde kapsamlı testler içermektedir. Yapılan testler, önerilen yöntemin %88 ve %87 doğruluk oranlarıyla sırasıyla tek ve çok alan bağlamlarında etkili olduğunu göstermiştir. Bu sonuçlar, dil tabanlı özelliklerin ve evrişimli sinir ağlarının spam incelemeleri tespitinde başarılı bir şekilde kullanılabileceğini göstermektedir (Neisari, 2021). Bu yaklaşımlar, spam tespiti konusunda gelecekteki çalışmalar için önemli bir temel oluşturmaktadır.

Rosita ve Jacob (2022) tarafından gerçekleştirilen bir çalışmada, Twitter spam tespiti için geliştirilen Çok Amaçlı Genetik Algoritma ve CNN Tabanlı Derin Öğrenme Mimarisi (MOGA-CNN-DLAS) yöntemi sunulmuştur. Bu yöntem, dil tabanlı özellikleri kullanarak spam incelemelerini gerçek incelemelerden ayırt etmeye odaklanmaktadır. MOGA-CNN-DLAS, kendiliğinden örgütlenmiş haritalar (SOM) ve evrişimli sinir ağları (CNN) kullanarak spam incelemelerinin sınıflandırılmasını sağlar. İncelemeler, semantik olarak benzer kelimelerin bir pikselin etrafına veya SOM ızgara hücresine düzenlendiği görüntülere dönüştürülür. Elde edilen görüntüler daha sonra CNN'e beslenir, eğitilir ve sınıflandırma için kullanılır. MOGA-CNN-DLAS, iki temel motivasyona dayanmaktadır. İlk olarak, öğrenilmemiş veri kümesinden gizli anlamların öğrenilmesine yardımcı olmayı hedefler. İkinci olarak, sınırlı kaynakları verimli bir şekilde kullanmayı amaçlar. Önerilen yöntem, çoklu amaçlı optimizasyon süreci olan Multi-Objective Optimization (MOGA) ile en uygun özellik kümesini seçer. MOGA-CNN-DLAS, çok katmanlı evrişimli, havuzlama ve tam bağlantılı katmanlardan oluşan bir yapı kullanır. Bu yöntem, Twitter spamının etkin bir şekilde tespitine yardımcı olur ve tweet'leri normal ve zararlı spam tweetleri olarak başarılı bir şekilde sınıflandırır. Yapılan deneysel sonuçlar, MOGA-CNN-DLAS yönteminin diğer referans yöntemlere göre ortalama doğruluk, hassasiyet, geri çağırma, F-skoru ve hata ölçütleri açısından önemli bir iyileştirme sağladığını göstermektedir. Ayrıca, bu çalışma çoklu amaçlı optimizasyonun sosyal medya spam tespiti gibi uygulamalarda etkili bir şekilde kullanılabileceğini vurgulamaktadır (Rosita ve Jacob, 2022). Bu araştırma, spam tespiti konusunda gelecekteki çalışmalar için önemli bir katkı sağlamaktadır.

Liu ve ark. (2021) tarafından geliştirilen çalışmada, çok-granülar anlamsal bilgilerin yakalanması için iki katmanlı farklı dikkat mekanizmalarının kullanıldığı bir hiyerarşik dikkat ağı önerilmiştir. Bu yöntem, metin belgelerindeki önemli anlamsal özellikleri çıkarmayı amaçlamaktadır. İlk katmanda, cümlelerin çok-granülar anlamsal özelliklerini çıkarmak için N-gram CNN kullanılmıştır. N-gram CNN, metin içindeki dilbilgisi yapısını anlamaya yardımcı olan bir özellik çıkarım tekniğidir. Ardından, ikinci katmanda belgedeki önemli ve kapsamlı anlamsal özellikleri çıkarmak için evrişim yapısı ve Bi-LSTM'in birleşimini kullanılmıştır. Bu yöntem, farklı alanlarda yapılan deneylerde diğer yöntemlere göre üstün tespit performansı göstermiştir. Örneğin, karışık bir alan için F1 skorunu %89,3'e (4,8 puan mutlak iyileştirme ile) yükseltirken, doktor domain için F1 skorunu %92,8'e (9,9 puan mutlak iyileştirme ile) çıkarmıştır. Benzer şekilde, otel alanında %86,1'e (2.4 puan mutlak iyileştirme ile) ve çapraz alanlarda %84,7'ye (10,4 puan mutlak iyileştirme ile) kadar F1 skorları arttırılmıştır. Saeed ve ark. (2022) tarafından gerçekleştirilen bir çalışmada, Arapça metinlerinde spam tespiti için dört farklı yöntem sunulmuş ve özellikle birleşik bir yaklaşımın geliştirilmesi ve değerlendirilmesine odaklanılmıştır. Önerilen birleşik yöntem, kural tabanlı sınıflandırıcıyı makine öğrenme teknikleriyle bütünleştirmekte ve metin içeriğinde N-gram özelliklerine ve olumsuzluk işleme özelliğine dayanan içerik tabanlı özellikleri kullanmaktadır. Bu dört önerilen yöntem, farklı boyutlarda iki veri kümesi üzerinde değerlendirilmiştir. Sonuçlar, bu birleşik yaklaşımın her iki veri kümesi için sırasıyla %95,25 ve %99,98 sınıflandırma doğruluğu elde ettiğini göstermiştir. Bu sonuçlar, mevcut ilgili çalışmaları %25 oranında geride bırakarak spam tespiti konusunda etkili bir yaklaşımın mümkün olduğunu göstermektedir (Saeed ve ark., 2022). Bu çalışma, Arapça metinlerde spam tespiti alanında önemli bir katkı sağlamıştır.

Idris ve Selamet (2014) tarafından yapılan çalışmada, e-posta spam tespiti konusundaki önemli bir araştırmayı temsil etmektedir. Bu çalışma, gelişmiş spam filtreleme teknikleri geliştirmeyi amaçlamış ve Negatif Seçim Algoritması ve Parçacık Sürü Optimizasyonu tabanlı bir model sunmaktadır. E-posta spamı, modern iletişimde yaygın bir sorundur ve bu tür spam mesajları, kullanıcıların posta kutularını dolmaktadır. Bu nedenle, e-posta sağlayıcıları ve kullanıcılar, spam e-postalarını tanımlamak ve filtrelemek için etkili yöntemlere ihtiyaç duyarlar. Negatif Seçim Algoritması, bir yapay bağışıklık sistemi modeline dayanır ve spam e-postalarını "tanıdık" ve "yabancı" olarak ayırmak için kullanılır. Bu yaklaşım, spam tespitinde büyük bir potansiyele sahiptir ancak

daha etkin hale getirilmesi gerekmektedir. İşte bu noktada, Parçacık Sürü Optimizasyonu devreye girmektedir. Bu optimizasyon tekniği, dedektörlerin rastgele üretilmesini iyileştirmek için kullanılır. Bu sayede, daha iyi performans gösteren dedektörler seçilir ve kullanılır. Çalışmanın özgünlüğü, adaptif yapısı ve spam konusuna özgü özellikleri içeren Negatif Seçim Algoritması'nın Parçacık Sürü Optimizasyonu ile birleştirilmiş olmasıdır. Ayrıca, mesafe ölçüleri gibi farklı bileşenler kullanılarak spam olmayan ve spam aday dedektörleri arasındaki fark artırılmıştır. Bu sayede, spam e-postalarının daha etkili bir şekilde tanımlanması mümkün olmuştur. Deneysel sonuçlar, önerilen NSA-PSO modelinin standart Negatif Seçim Algoritması'na göre daha yüksek bir tespit oranına sahip olduğunu göstermektedir. Özellikle, 0,4 eşik değeriyle yapılan doğruluk karşılaştırmasında, Negatif Seçim Algoritması %68,86 doğruluk sağlarken, önerilen NSA-PSO modeli %91,22 doğruluk sağlamıştır. Bu sonuçlar, e-posta spamı tespiti için yeni ve etkili bir yöntemin başarıyla uygulanabileceğini göstermektedir (İdris ve Selamet, 2014). 2020 yılında Murugavel ve Santhi tarafından gerçekleştirilen çalışma, e-posta spam tespiti ve spam tehditlerinin sınıflandırılmasına odaklanmıştır. Spam e-postaların ve spam tehditlerinin artması, bu konunun önemini artırmıştır. Önerilen algoritma, yoğun metin verilerini kullanarak spam e-postalarını tanımlamayı ve sınıflandırmayı hedeflemektedir. Bu işlem, spam kelimelerinin içerik tabanlı metin analiziyle kontrol edilmesi ve spamın farklı türlerine göre sınıflandırılmasıyla gerçekleşir. Çalışma, spam korpusu veritabanında bulunan spam kelimelerini inceleyerek spam tehditlerini kataloglamış ve farklı türdeki spam tehditlerini tanımlamıştır. Ayrıca, sıkça kullanılan işlem adımları ve metotlar kullanılarak spam tehditlerini sınıflandırmıştır.

Elde edilen sonuçlar, önerilen yöntemin spam e-postalarını ve spam tehditlerini sınıflandırmada oldukça etkili olduğunu göstermektedir. Bu çalışma, sahte e-postalar, kötü amaçlı yazılımlar, gereksiz postalar, zincir mektuplar, görüntü spamı, pornografi spamı gibi farklı spam tehditlerini azaltmak ve tanımlamak için farklı yöntemler ve işlem adımları kullanmıştır. Özellikle, farklı veri işleme yaklaşımları ve sınıflandırma yöntemleri sayesinde spam tehditlerinin tespit edilmesi ve kullanıcıların e-posta ile ilgili güvenlik sorunlarına karşı daha iyi korunması mümkün olmuştur. Bu çalışma, e-posta spamıyla mücadelede yeni bir yaklaşım sunmakta ve spam tehditlerinin etkili bir şekilde azaltılmasına katkıda bulunmaktadır (Murugavel ve Santhi, 2020).

3. MATERYAL VE METOT

3.1. Materyal

Çalışmada kullanılan veri seti Kaggle veri tabanından (https://www.kaggle.com/) indirilmiştir. Veri seti kaynağı belirsiz e-postalardan oluşmaktadır. Veri setinde 497 spam ve 2499 non-spam olarak etiketlenmiş toplamda 2996 adet e-posta bulunmaktadır. Veri setinde spam e-postaların oranı %0,1666 olarak gözlemlenmiştir. Bu da veri setinin dengesiz dağıldığını göstermektedir. Veri seti bir excel dosyasında bulunmaktadır. Veri setinden özniteliklerin elde edilmesi için Matlab programı kullanılmıştır. Örnek bir spam ve non-spam mesajlar Çizelge 3.1. 'de gösterilmiştir.

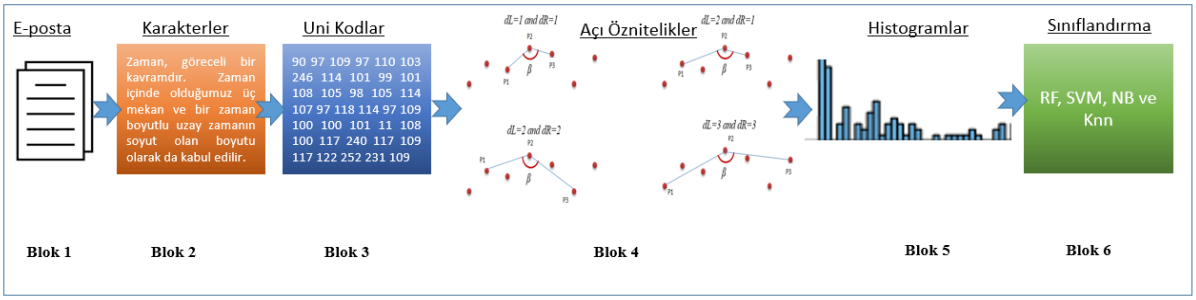
Çizelge 3. 1. Örnek e-postalar

Spam	Non-Spam
the famous ebay marketing e course learn to sell with the complete ebay auction marketing e course here s your chance to join the online selling revolution and earn a full time income our ebay marketing e course will show you how to create huge profits selling on ebay do you sell on ebay if so you could be making up to NUMBER NUMBER per month this is no hype and no scam receiving over NUMBER NUMBER billion page views per month ebay is the ultimate venue for selling virtually anything and making huge profits with almost no effort but you have to know what to sell and how to sell that s where i come in as a leading expert in internet marketing and the owner of several profitable auction ……	in adding cream to spaghetti carbonara which has the same effect on pasta as making a pizza a deep pie i just had to jump in here as carbonara is one of my favourites to make and ask what the hell are you supposed to use instead of cream i ve never seen a recipe that hasn t used this personally i use low fat creme fraiche because it works quite nicely but the only time i ve seen an supposedly authentic recipe for carbonara it was identical to mine cream eggs and lots of fresh parmesan except for the creme fraiche stew stewart smith scottish microelectronics centre university of edinburgh URL yahoo groups sponsor NUMBER dvds free s p join now URL to unsubscribe from this group send an email to forteana unsubscribe URL your use of yahoo groups is subject to URL

3.2. Metot

3.2.1. Spam tespiti için blok diyagram

İstenmeyen e-postaların tespiti için önerilen yaklaşıma ait blok diyagram Şekil 3.1'de verilmiştir. Önerilen yaklaşım 6 bloktan oluşmaktadır. Her blokta gerçekleştirilen işlemler aşağıda özetlenmiştir.



Şekil 3. 1. Önerilen Yaklaşım Blok Diyagramı

Örnek bir e-posta Çizelge 3.2'de görülmektedir.

Çizelge 3. 2. Örnek non-spam bir e-posta

in adding cream to spaghetti carbonara which has the same effect on pasta as making a pizza a deep pie i just had to jump in here as carbonara is one of my favourites to make and ask what the hell are you supposed to use instead of cream i ve never seen a recipe that hasn t used this personally i use low fat creme fraiche because it works quite nicely but the only time i ve seen an supposedly authentic recipe for carbonara it was identical to mine cream eggs and lots of fresh parmesan except for the creme fraiche stew stewart smith scottish microelectronics centre university of edinburgh URL yahoo groups sponsor NUMBER dvds free s p join now URL to unsubscribe from this group send an email to forteana unsubscribe URL your use of yahoo groups is subject to URL

Blok 1: E-postalar spam ve non-spam olarak etiketlenmiştir. Gelen e-postalarda noktalama, sayılar gibi gereksiz karakterler silinmiştir.

Blok 2: Bu blokta ön işlemden geçen e-postaların içindeki metinleri ifade eder. Metinlerden boşluklarda silinmektedir. Yukarıdaki örnek e-postadan gerekli ön işlemlerden sonra Çizelge 3.3.'te görüldüğü metin elde edilmiştir.

Çizelge 3. 3. Ön işlemlerden sonra elde edilen metin

inaddingcreamtospaghetticarbonarawhichhasthesameeffectonpastaasmakingapizzaadeeppieijusthadtojumpinhereascarbonaraisoneofmyfavouritestomakeandaskwhatthehellareyousupposedtouseinsteadofcreamiveneverseenarecipethathasntusedthispersonallyiuselowfatcremefraichebecauseitworksquitenicelybuttheonlytimeiveseenansupposedlyauthenticrecipeforcarbonaraitwasidenticaltominecreameggsandlotsoffreshparmesanexceptforthecremefraichestewstewartsmithscottishmicroelectronicscentreuniversityofedinburghURLyahoogroupssponsorNUMBERdvdsfreespjoinnowURLtounsubscribefromthisgroupsendanemailtoforteanaunsubscribeURLyouruseofyahoogroupsissubjecttoURL

Blok 3: Bu blokta metinlerin unikodları elde edilmektedir. Metin içindeki her karakterin unikodları elde edilir. Yukarıdaki örnek e-postaya ait unikodlar Çizelge 3.4'de verilmiştir.

Çizelge 3. 4. E-postaya ait unikodlar

105	110	97	100	100	105	110	103	99	114	101	97
	109	116	111	115	112	97	103	104	101	116	116
	105	99	97	114	98	111	110	97	114	97	119
	104	105	99	104	104	97	115	116	104	101	115
	97	109	101	101	102	102	101	99	116	111	110
	112	97	115	116	97	97	115	109	97	107	105
	110	103	97	112	105	122	122	97	97	100	101
	101	112	112	105	101	105	106	117	115	116	104
	97	100	116	111	106	117	109	112	105	110	104
	101	114	101	97	115	99	97	114	98	111	110
	97	114	97	105	115	111	110	101	111	102	109
	121	102	97	118	111	117	114	105	116	101	115
	116	111	109	97	107	101	97	110	100	97	115
	107	119	104	97	116	116	104	101	104	101	108
	108	97	114	101	121	111	117	115	117	112	112
	111	115	101	100	116	111	117	115	101	105	110
	115	116	101	97	100	111	102	99	114	101	97
	109	105	118	101	110	101	118	101	114	115	101
	101	110	97	114	101	99	105	112	101	116	104
	97	116	104	97	115	110	116	117	115	101	100
	116	104	105	115	112	101	114	115	111	110	97
	108	108	121	105	117	115	101	108	111	119	102
	97	116	99	114	101	109	101	102	114	97	105
	99	104	101	98	101	99	97	117	115	101	105
	116	119	111	114	107	115	113	117	105	116	101
	110	105	99	101	108	121	98	117	116	116	104
	101	111	110	108	121	116	105	109	101	105	118
	101	115	101	101	…						
……											

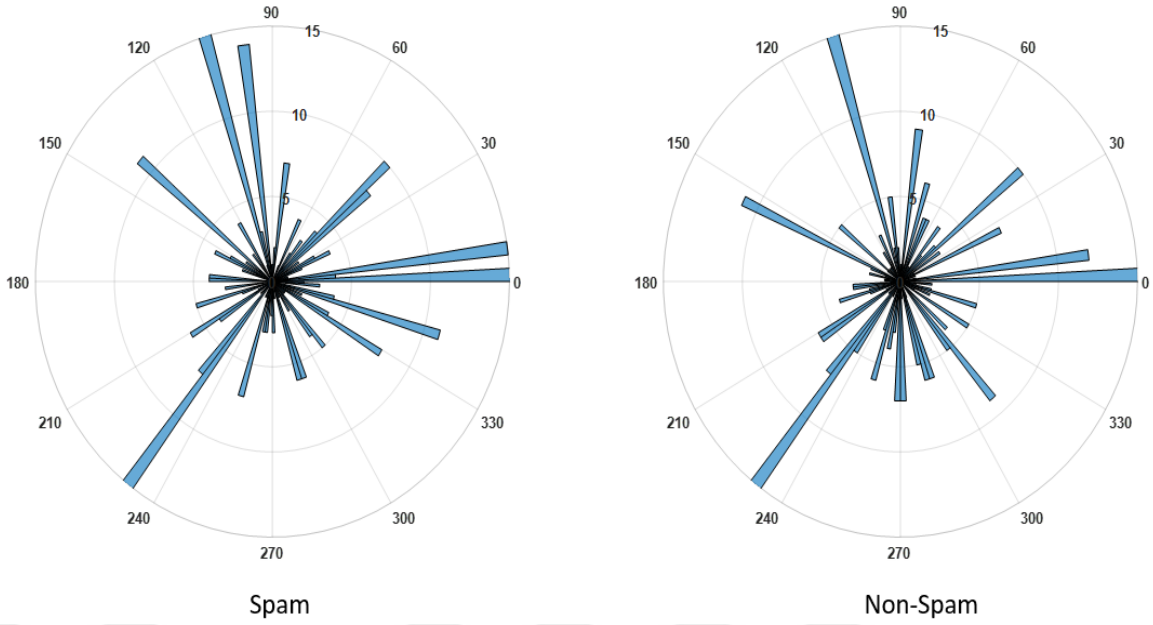
Blok 4: Bu aşamada uygun hale getirilmiş e-postalara açı dönüşümü uygulanmaktadır. Açı dönüşümü aşağıda detaylıca anlatılmıştır. Açı dönüşümü unikodlar arasındaki açı bilgilerinden elde edilmesidir. Örnek e-postadan elde edilen açı değerleri Çizelge 3.5'te verilmiştir.

Çizelge 3. 5. Hesaplanan açı değerleri

15	337	108	258	180	19	185	342	8	189	341	176
	19	334	32	165	346	144	63	337	93	95	184
	197	330	6	352	49	139	352	6	354	6	311
	54	339	101	98	348	138	49	193	337	7	352

11	262	225	135	135	161	330	14	213	288	30
353	138	48	266	266	12	175	349	32	322	19
181	346	11	348	93	92	267	251	153	135	264
95	98	185	331	149	219	31	288	49	183	333
194	14	180	343	12	334	26	340	20	188	337
8	189	342	6	202	330	6	352	49	139	352
6	349	181	19	210	141	347	12	345	183	7
188	345	10	342	27	167	348	9	352	139	56
195	158	349	15	184	341	10	192	338	10	348
8	184	348	93	94	193	323	36	333	98	95
351	7	352	8	344	36	306	37	258	135	300
18	220	311	14	339	36	157	341	182	180	146
48	190	327	193	11	192	337	8	189	341	18
341	7	350	12	350	6	352	139	49	265	263
10	352	7	202	323	181	13	350	8	183	348
7	183	348	14	339	144	71	157	220	311	8
310	219	24	166	350	139	59	210	139	350	95
265	7	351	31	157	347	169	191	10	187	345
6	352	8	348	14	307	220	8	349	16	350
360	180	0	360	180	0	360	180	0	180	180
360	0	360	0	360	0	360	0	360	0	360
180	0	180	180	360	0	360	180	180	180	0
360	180	0	360	180	0	360	0	180	360	0
360	180	180	360	0	180	180	360	180	180	0
360	180	0	180	360	0	360	180	180	0	180
360	180	0	360	180	0	180	360	0	180	180
360	180	0	360	0	360	180	0	360	0	180
180	360	0	360	180	0	180	180	360	180	180
0	……									

Blok 5: Açı dönüşümü uygulanmış unikodların histogramları bu blokta bulunur. Histogramlar sınıflandırma algoritmalarına verilen öznitelikleri ifade eder. Örnek spam ve spam olmayan e-postalara ait gül histogramları Şekil 3.2'de verilmiştir. Şekilden görüldüğü gibi öznitelik dağılımların farklı olduğu görülmektedir.

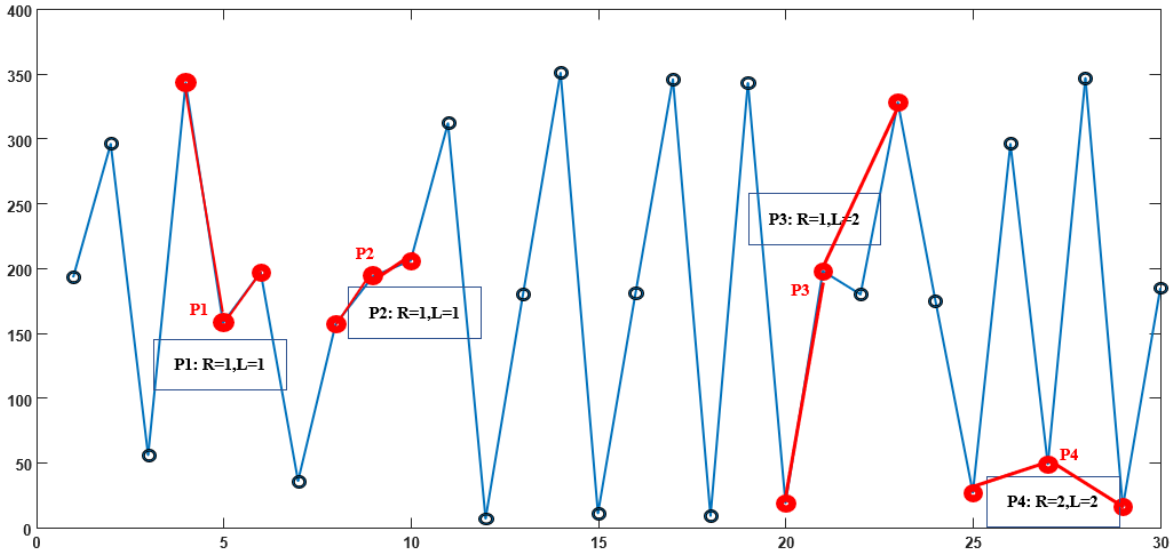


Şekil 3. 2. Örnek epostalara ait gül histogramları

Blok 6: Son blokta çıkarılan açı örüntülerin kullanılması ile makine öğrenmesi sınıflandırma algoritmaları kullanılır. Sınıflandırma işlemi RF, SVM, Knn ve NB yöntemleri ile gerçekleştirilmiştir.

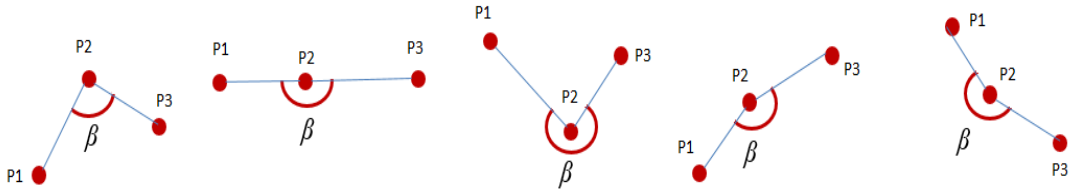
3.2.2. Açı Dönüşüm Metodu

Bu çalışmada istenmeyen e-postaların tespiti için farklı bir yaklaşım önerilmiştir. Açı dönüşümü, metin içerikli e-postaların tek boyutlu sinyal olarak düşünülmesi ve bu sinyal üzerindeki her karakterin Unikod değerleriyle ifade edilmesi esasına dayanır. Açı dönüşüm yöntemi tek boyutlu verilerden etkin özelliklerin elde edilmesi için işaretler üzerindeki unikod değerlerin birbirleri ile oluşturdukları açı bilgilerini kullanan istatistiksel yeni bir yaklaşımdır. Şekil 3.3 ’te örnek bir epostaya ait karakterlerin unikod değerlerini gösteren noktalar gösterilmiştir. Açı dönüşüm yöntemini bu örnek üzerinde anlatmış olalım.



Şekil 3. 3. Bir e-postaya ait Unikod değerleri ve açı örnekleri

Açı dönüşüm yönteminde, ilk olarak e-posta içindeki metne ait karakterler arasındaki açı değerleri hesaplanır. Komşu 3 nokta arasındaki aşağı yöne bakan açı hesaplanır. Bu işlem bir e-postaya ait tüm metin için elde edilen Unikod değerleri için gerçekleştirilir. Unikod değerleri yerine tümüyle açı bilgilerinden oluşmuş yeni bir vektör elde edilmiş olur. Daha sonra her bir açıdan kaç tane olduğunu gösteren frekansları hesaplanır. Yani oluşan açı vektörüne karşılık gelen bir histogram elde edilir. Şekil 3.4'te örnek açılar gösterilmiştir.



Şekil 3.4. Örnek açı örüntüler

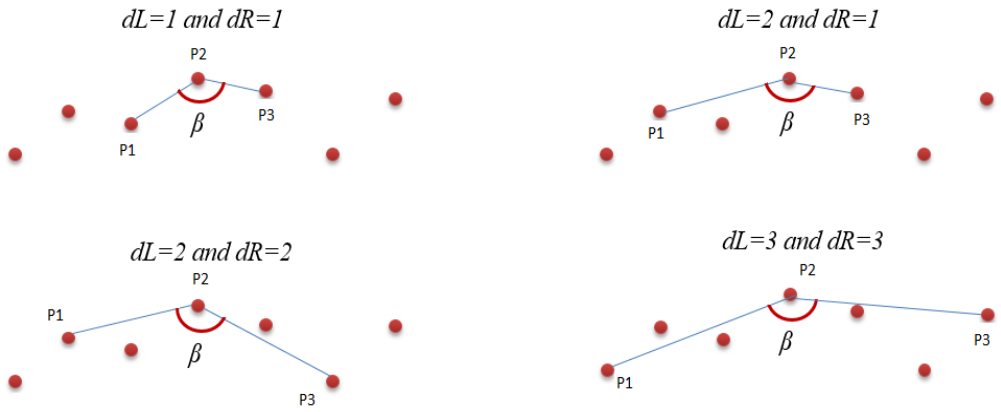
Yukarıdaki şekillerde görüldüğü gibi komşu 3 nokta arasında aşağı bakan β açıları bizim elde edeceğimiz açı değerlerimizi oluşturur.

Zaman serisi şeklinde ede edilen Unikod değerlerinden herhangi 3 nokta arasında kalan $P_i^{x,y}$ noktası için açı değeri aşağıdaki eşitlik ile hesaplanmaktadır. ($P_i^x > 0$ olmak üzere):

$$\theta_i = \arctan\left(\frac{\left|det\begin{vmatrix} P_{i-1}^x - P_i^x & P_{i-1}^y - P_i^y \\ P_i^x - P_{i+1}^x & P_i^y - P_{i+1}^y \end{vmatrix}\right|}{(P_{i-1}^x - P_i^x)(P_i^x - P_{i+1}^x) + (P_{i-1}^y - P_i^y)(P_i^y - P_{i+1}^y)}\right) \quad (3.1)$$

Burada Bağıntı 3.1'de θ_i radyal olarak elde edilmektedir.

Açı dönüşüm yönteminde bir açı bilgisi hesaplanırken referans noktasının solundan ve sağından kaçıncı komşuluğa bakacağını ifade eden iki parametre bulunmaktadır. Bunlar uR(dR, Right(sağ)'dan Uzaklık) ve uL(dL, Left(sol)'dan uzaklık) isimli uzaklık parametreleridir. uR parametresi referans noktasına sağdan uzaklığını belirtir, uL parametresi ise referans noktasına soldan uzaklığını belirtir. Yani uR parametresi P3 noktasının P2 noktasına sağdan olan uzaklığını belirtir. uL parametresi ise P1 noktasının P2 noktasına soldan olan uzaklığını belirtir. uL ve uR değerlerine göre P1, P2 ve P3 arasında farklı açı değerleri oluşmaktadır. Oluşan farklı açı değerleri farklı örüntülerin oluşmasını sağlamaktadır. Örnekler Şekil 3.5 'te gösterilmiştir.



Şekil 3. 5. L ve R parametrelerine göre oluşan örüntüler

E-postalara ait metinlere açı dönüşüm yöntemi uygulandıktan sonra açı değerleri 0 ile 359 arasındaki bilgilere dönüşmektedir. Her açı değerinin frekansı bir açı örüntüsü olarak ele alınmaktadır. Diğer bir deyişle yeni oluşan açı değerlerine ait histogram öznitelik vektörü olarak ele alınmaktadır.

3.2.3. Navie Bayes (NB)

Naive Bayes Ağları (NB'ler), bir problemin değişkenleri arasındaki bağımlılıkları ve bağımsızlıkları temsil etmek için yaygın olarak kullanılan olasılıksal grafik modellerdir. Bir NB, bir ortak olasılık dağılım fonksiyonunun kompakt bir temsili olarak

düşünülebilir. Bir NB, Yönlendirilmiş Asiklik Grafik olan bir yapıdan ve bir dizi değişken arasındaki bağımlılıklar hakkında nicel bilgileri temsil eden bir parametre setinden oluşur. Bu ağlar, yapay görme, biyoinformatik, veri birleştirme ve karar destek sistemlerinde yaygın olarak kullanılmaktadır. Sürekli bir NB'yi öğrenmek için birçok yaklaşım geliştirilmiş olsa da ayrık olan için bir eksiklik vardır. Bunun nedeni, ayrık bir NB 'yi öğrenmenin, geniş parametre alanı ve etkin bir yapı aramanın zorluğu nedeniyle zorlu bir problem olmasıdır. Ancak, yüksek boyutlu ayrık bir veriden seyrek bir NB öğrenmemiz gerektiğinde problem daha da zorlaşır. Beyin bilimlerinden biyoloji bilimlerine kadar geniş bir problem koleksiyonunda seyrek yapılara büyük talep vardır. Örneğin, işlevsel beyin ağlarının tanımlanması veya mikro dizi gen ifade verilerinden genler arasındaki etkileşim modellerinin modellenmesi, numune sayısının değişken sayısına eşit veya daha az olduğu yüksek boyutlu verileri temsil eder. Ayrıca, gen ilişkilendirme ağları ve beyin bağlantı ağları gibi birçok gerçek dünya ağı seyrektir. Bu nedenle, bu tür veri kümelerinden seyrek bir yapının doğru bir şekilde öğrenilmesi büyük önem taşımaktadır (Ren ve ark., 2022).

Naive Bayes yapısı için önerilen öğrenme yöntemi üç kategoriye ayrılır:

(1) Temel Bileşen (Personal Component), Maks-Min Parents and Children (Ebeveynler ve Çocuklar) ve Hızlı Nedensel Çıkarım gibi kısıtlamaya dayalı yöntemler,

(2) Puana dayalı yöntemler,

(3) MAX-Min Hill-Climbing gibi hibrit yöntemler. Seyrek modellemeye artan eğilimle birlikte, skor tabanlı yöntemler, skor fonksiyonlarına kısıtlamalar uygulama yeteneklerinden dolayı daha fazla dikkat çekmiştir. Bu yöntemler, her yapıya bir skor atar ve ardından en iyi skora sahip yapıyı arar. Farklı skor fonksiyonları puana dayalı yapı öğreniminde Dirichlet (BD) metriği, Bayes Bilgi Kriteri (BIC), Minimum Açıklama Uzunluğu ve entropi tabanlı metrikler kullanılır. Puanları atadıktan sonra, optimum yapıyı bulmak için bir arama algoritması ile optimum puan kullanılır.

Naive Bayes, hedef değişkenle bağımsız değişkenler arasındaki ilişkiyi analiz eden tahminci ve tanımlayıcı bir sınıflama algoritmasıdır. Kullanım alanlarına örnek olarak gerçek zamanlı tahmin, duyarlılık analizi, çok sınıflı tahmin, metin sınıflandırması, öneri sistemleri ve spam filtreleme verilebilir. Bayes filtreleme, istenmeyen e-posta tespiti için muhtemelen en yaygın şekilde uygulanan makine öğrenimi yöntemidir. İlgili filtreleme işlemi, e-postaların gövdesindeki tüm ilgili terimlerin aranmasını ve genel sınıflandırma olasılığının hesaplanmasını içerir (Sahami ve ark., 1998; Al-Kadhi ve Mishaal-Abdullah, 2011; Androutsopoulos ve ark., 2000). Naive Bayes yaklaşımı, genellikle gelecek

olasılıkları hesaplamakta kullanılan ve iki rastgele olayın koşullu ve marjinal olasılıklarını ilişkilendiren bir teoremdir. Maksimum Olabilirlik ilkesi üzerine kurulu bir teoremdir. Bu durumda Bayes Teoremi, konulan olasılıkların doğruluk oranını hesaplamak için kullanılabilir.

A ve B rastgele olaylar olsun;

$$P\left(\frac{A}{B}\right) = P\left(\frac{B}{A}\right)\frac{P(A)}{P(B)} \tag{3.2}$$

P(A) : A olayının bağımsız olasılığı

P(B) : B olayının bağımsız olasılığı

P(B | A) : A olayının olduğu bilindiğinde B olayının olasılığı (Likelihood, Şartlı Olasılık)

P(A | B) : B olayının olduğu bilindiğinde A olayının olasılığı (Posterior, Artçıl Olasılık)

NB elektronik postaların SPAM postalardan ayrıştırılması konusunda yapısı itibariyle uygun özellikler taşıyan bir yapıya sahiptir. Bu nedenle elektronik postaların ayrıştırılmasında etkisi büyüktür. Naive Bayes, belirli bir verinin sınıfını tahmin etmek için olasılık hesaplamalarını temel alır. Temel bir varsayım, kullanılan değişkenlerin birbirinden bağımsız olması gerekliliğidir. Yani, bu yöntemin gücü, farklı özelliklere sahip değişkenlerin karşılıklı olarak bağımsız olduğunu varsayar. Bu, Naive Bayes sınıflandırıcının çok sayıda veri noktası ile etkili bir şekilde çalışabilmesini sağlar. Ayrıca, bu yöntemin önemli bir avantajı, az miktarda eğitim verisi ile bile başarılı bir şekilde çalışabilmesidir. Hızlı, etkili ve doğru bir sınıflandırma algoritması olarak kabul edilen Naive Bayes, genellikle pratik uygulamalarda tercih edilir. Naive Bayes yöntemi dört ana adımdan oluşur (Bhat ve ark., 2022; Gopalsamy ve Radha, 2022; Kachhia ve Rathod, 2022):

Adım 1: Belirli bir sınıfa ait etiketlerin öncül olasılığı hesaplanır.

Adım 2: Her sınıf için her özellikle ilgili olabilirlik olasılığı hesaplanır.

Adım 3: Bayes teoremi kullanılarak ardıl olasılık hesaplanır.

Adım 4: En yüksek olasılığa sahip sınıf tespit edilir.

3.2.4. Support Vector Machine (SVM)

Support Vector Machine (SVM) algoritması, ilk kez Vladimir Vapnik tarafından 1963 yılında tanıtılmıştır (Vapnik, 1963). SVM, bir veri kümesini bir uzay düzleminde temsil ederek, bu verileri optimal bir şekilde ayıracak bir hiperdüzlemi bulmaya çalışan denetimli bir öğrenme modeli ve doğrusal bir sınıflandırıcıdır. SVM'nin birçok avantajı bulunmaktadır; özellikle küçük veri setlerinde yüksek performans sergiler. Ayrıca, SVM aşırı uydurma (overfitting) eğiliminde değildir ve iyi bir genelleme yeteneği sunar. Ancak, büyük veri setlerinde hesaplama ve bellek maliyeti nedeniyle performansı azalabilir. İkili sınıflandırma durumunda, SVM, veri noktalarını iki sınıf arasında en büyük marj ile bölen bir hiperdüzlem bulmaya çalışır. Marj, iki sınıf arasındaki boşluk olarak düşünülebilir ve SVM, bu marjı maksimize ederek en iyi ayrımı yapmaya çalışır. Her iki sınıfın en yakın veri noktalarından en uzak olanlar hiperdüzlemi tanımlar. Bu tür veri noktalarına "destek vektörleri" denir. SVM, çoğu zaman doğrusal modelleri ifade eden hiperdüzlemlerle tanınır, ancak doğrusal olmayan problemleri çözmek için "kernel" adı verilen yapıları kullanabilir. Kernels, yüksek boyutlu ve doğrusal olmayan modellerin oluşturulmasını sağlar ve işlenmemiş verileri yüksek boyutlu uzayda temsil ederek doğrusal bir problemle ele alır. Kernel fonksiyonları, belirli hesaplamaların daha hızlı bir şekilde yapılmasına yardımcı olacak şekilde tasarlanmıştır. SVM, birçok uygulama alanında kullanılır ve özellikle veri madenciliği ve sınıflandırma problemlerinde etkilidir. Karar sınırları veya hiperdüzlemler, sınıfları ayıran önemli sınırları temsil eder ve SVM bu sınırları tanımlamak için kullanılır.

Özellik uzayının boyutu olan d'nin $x = \in R^d$ öznitelik vektörü, yüksek boyutlu bir uzaya eşlenir. Genellikle, doğrusal olmayan ayırıcı veri kümeleri nihai uzayda ayrılabilir. Eşleme, bir fonksiyon $K = (x, x^{'})$ işlevi ve SVM karar fonksiyonu olarak uygulanır. Bağıntı 3.3 ile ifade edilir.

$$f(x) = sgn\left(\sum_{i=1}^{N} y_i * \alpha_i * K(x, x_i) + b\right) \tag{3.3}$$

Burada;

N: öğrenme kümesinin asal sayısıdır,

$y_i \in \{-1,1\}$: öğrenme kümesinin i'inci üyesi için sınıflandırma kararıdır.

a: katsayıları öğrenme sürecinde elde edilir.

x: karar fonksiyonu, çekirdek uzayında tanımlanan hiper düzleme göre sınıflandırılan x noktasının konumunu hesaplar.

Nihai sınıflandırma, sonucun işaretiyle tanımlanan x noktasını içeren yarı uzaya bağlıdır. (Marcin ve ark., 2014) Öğrenme süreci, Bağıntı 3.4'te dış bükey ikinci dereceden programlama problemini maksimize eder.

$$\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \quad \sum_{j=1}^{N} ai * aj * y_i * y_j * K = \left(x_j * x_i\right) \tag{3.4}$$

Problemin şekli, hiperdüzlem parametrelerine göre minimize ve α'ya göre maksimize etmemizin sonucudur. Problem Bağıntı 3.5'te sınırlamalara göre çözülmüştür.

$$(\forall i \in N) 0 \leq a_i \leq C \wedge \sum_{i=1}^{N} a_i * y_i = 0 \tag{3.5}$$

3.2.5. K-En Yakın Komşu Algoritması (kNN)

K-En Yakın Komşu (kNN) algoritması, makine öğrenmesi alanında sınıflandırma ve regresyon problemlerini çözmek için kullanılan basit, ancak etkili bir yaklaşımdır. Bu algoritma, bir veri noktasının sınıfını veya değerini belirlemek için, bu noktanın en yakın komşularının sınıfı veya değeri ile ilişkilendirme prensibine dayanır. K-En Yakın Komşu algoritması, farklı benzerlik metrikleri kullanarak, öklidyen mesafe veya Manhattan mesafesi gibi, bir veri noktasının en yakın komşularını belirler. Bu algoritma, literatürde "kNN" olarak da bilinir ve sıkça kullanılan bir sınıflandırma algoritmasıdır (Cover ve Hart, 1967). K-En Yakın Komşu, basitliği ve anlaşılabilirliği ile öne çıkar ve modelin eğitilmesini gerektirmez. Bunun yerine, veri noktaları arasındaki uzaklıkları hesaplar ve her tahminde bu uzaklıkları kullanır. Ayrıca, k-NN, hem doğrusal hem de doğrusal olmayan veri kümelerinde etkili bir şekilde çalışabilir. Ancak, k-NN algoritmasının bazı dezavantajları vardır. Özellikle büyük veri setlerinde hesaplama maliyeti yüksek olabilir, çünkü tüm veri noktalarıyla uzaklık hesaplamaları yapmak gerekebilir. Ayrıca, belirli bir K değeri seçmek gerekir. K, en yakın komşuların sayısını belirler ve bu hiperparametre deneme yanılma veya çapraz doğrulama ile seçilir. Eşit uzaklık durumunda, çoğunluk oylaması ile sınıf veya değer tahmini yaparken belirsizlik olabilir. Ayrıca, dengesiz veri setleri ve gürültü, k-NN algoritmasının performansını olumsuz etkileyebilir. K-En Yakın Komşu algoritması ayrıca regresyon problemlerini çözmek için de kullanılabilir ve

bağımsız değişkenlerin sayısal olduğu zamanlarda etkilidir. Ancak, hesaplama maliyetini minimize etmek için boyut azaltma teknikleri veya arama ağaçları gibi daha gelişmiş veri yapıları ile birleştirilebilir. Bu yöntem, özellikle spam filtreleme gibi uygulamalarda kullanılır (Francisco ve ark., 2012). Bu tür uygulamalarda, e-posta mesajları başlıklarına ve diğer özelliklere dayalı olarak sınıflandırılır. Bu, özellikle spam ve ham (jambon) e-postaları ayırt etmek için etkili bir yaklaşımdır.

K-En Yakın Komşu (kNN) algoritmasının temel adımları şunlardır:

- Veri kümesi toplanır veya hazırlanır. Bu veri kümesi, veri örneklerinin özellik vektörlerini ve bunların sınıf etiketlerini içerir.
- İki veri örneği arasındaki benzerliği veya uzaklığı ölçmek için bir metrik belirlenir.
- Uzaklıklar küçükten büyüğe sıralanır ve K-En Yakın Komşu'yu belirlemek için K değeri seçilir.
- Tahmin yapmak istenen veri örneği için, bu örneğe en yakın K komşuyu bulun ve bu komşuların sınıf etiketlerini kullanarak tahminde bulunun.

K-En Yakın Komşu algoritması, basitliği ve etkililiği ile bilinir ve birçok uygulama alanında kullanılabilir.

3.2.6. Random Forest (RF)

Random Forest, veri madenciliği ve makine öğrenimi alanlarında yaygın olarak kullanılan bir ensemble (birleştirilmiş) öğrenme algoritmasıdır. Bu algoritma, birden çok karar ağacını bir araya getirerek daha güçlü ve kararlı bir tahmin modeli oluşturmayı amaçlar. Random Forest, Leo Breiman ve Adele Cutler tarafından geliştirilmiştir. Temel yapı taşı karar ağaçlarıdır. Karar ağaçları, veri noktalarını sınıflandırmak veya regresyon yapmak için kullanılan ağaç benzeri yapılar oluştururlar. Her düğümde bir özellik seçilir ve bu özellik kullanılarak veri bölünür. Bu bölünme işlemi, veri kümesini daha homojen alt gruplara bölmeyi amaçlar. Karar ağacının yapısı, veriye ve probleme göre otomatik olarak öğrenilir. Random Forest, ensemble öğrenme modelini kullanır, yani birden çok modelin birleştirilmesini içerir ve bu modellerin tahminlerini birleştirerek daha güçlü bir model oluşturur. Her bir karar ağacı, kendi veri alt kümesi üzerinde eğitilir ve bu ağaçların tahminleri birleştirilir. Bu eğitim verisi, rastgele seçilen alt kümeler halinde kullanılarak her bir karar ağacı için farklı bir eğitim seti oluşturulur. Bu, ağaçların birbirinden bağımsız olmasını sağlar ve çeşitliliği artırır (Nizam ve Akın, 2014).

Random Forest'in avantajları şunlar:

- Yüksek performans ve genelleme yeteneği: Random Forest, birden çok ağacın birleştirilmesiyle daha güçlü ve kararlı tahminler yapabilir. Ayrıca, aşırı uyumu azaltır ve daha iyi genelleme yeteneği sağlar.
- Eksik veri ve gürültü ile başa çıkma: Random Forest, eksik veri ile başa çıkmada etkilidir ve gürültülü veriye karşı daha dayanıklıdır.
- Özellik önem sıralamaları: Random Forest, hangi özelliklerin tahminde daha etkili olduğunu anlamak için özellik önem sıralamalarını hesaplayabilir, bu da veri anlayışını artırabilir.

Random Forest algoritmasının adımları şu şekildedir:

- Giriş veri setinden rastgele örnekler seçilir.
- Random Forest algoritması, her seçilen örnek için bir karar ağacı oluşturur.
- Her bir karar ağacı tahminlerde bulunur.
- Tahmin edilen sonuçlar kullanılarak, çoğunluk oylaması veya ortalama değer hesaplaması yapılır ve nihai tahmin elde edilir.

Random Forest, sınıflandırma ve regresyon problemleri için kullanılabilir. Her bir ağaç rastgele bir alt kümesi üzerinde eğitildiği için farklı görüşlere sahip olabilirler. Ancak birleştirildiklerinde, doğru tahminlerin bir araya gelmesi ve hataların düzeltilmesi sağlanır. Bu algoritma, genellikle yüksek doğruluk gerektiren uygulamalarda tercih edilir ve birçok endüstri alanında kullanılır.

3.2.7. Performans Ölçütleri

İstenmeyen e-postaların tespiti için önerilen yaklaşımın performansını değerlendirmek için karışıklık matrisi kullanılır. Veri setimizdeki gerçek çıkış etiketleri ile modellerin tahmin etiketleri yanlış ve doğru sayılarını Çizelge 3.6’da gösterilmiştir.

Çizelge 3. 6. Karışıklık matrisi gösterimi

	Gerçek		
Tahmin		**Spam**	**Non-Spam**
	Spam	**True Pozitif (TP)**	**True Negatif (TN)**
	Non-Spam	**False Negatif (FN)**	**False Pozitif (FP)**

Burada;

TP (True Pozitif): Spam olarak etiketlenmiş e-postaların model tarafından da Spam olarak tahmin edilmesidir. Bu durumda doğru sınıflandırma sayısıdır.

TN (True Negatif): Spam olarak etiketlenmiş görüntülerin model tarafından Non-Spam olarak tahmin edilmesidir. Bu durumdaki yanlış sınıflandırma sayısıdır.

FP (False Pozitif): Non-Spam olarak etiketlenmiş görüntülerin model tarafından Spam olarak etiketlenmiş e-postaların sayısını belirtir.

FN (False Negatif) : Non-spam olarak etiketlenmiş e-postaların model tarafından Spam olarak etiketlenmiş e-postaların sayısını belirtir.

İstenmeyen e-postaların tespiti için önerilen yaklaşımın performansını değerlendirmek için Bağıntı (3.6), Bağıntı (3.7), Bağıntı (3.8) ve Bağıntı (3.9) ile ifade edilir.

$$Başarı = \frac{TP + TN}{TP + TN + FN + FP} \tag{3.6}$$

$$Kesinlik = \frac{TP}{TP + FP} \tag{3.7}$$

$$Hatırlatma = \frac{TP}{TP + FN} \tag{3.8}$$

$$F - \text{Ölçütü} = 2 * \frac{Kesinlik * Hatırlatma}{Kesinlik + Hatırlatma} \tag{3.9}$$

4. SONUÇLAR

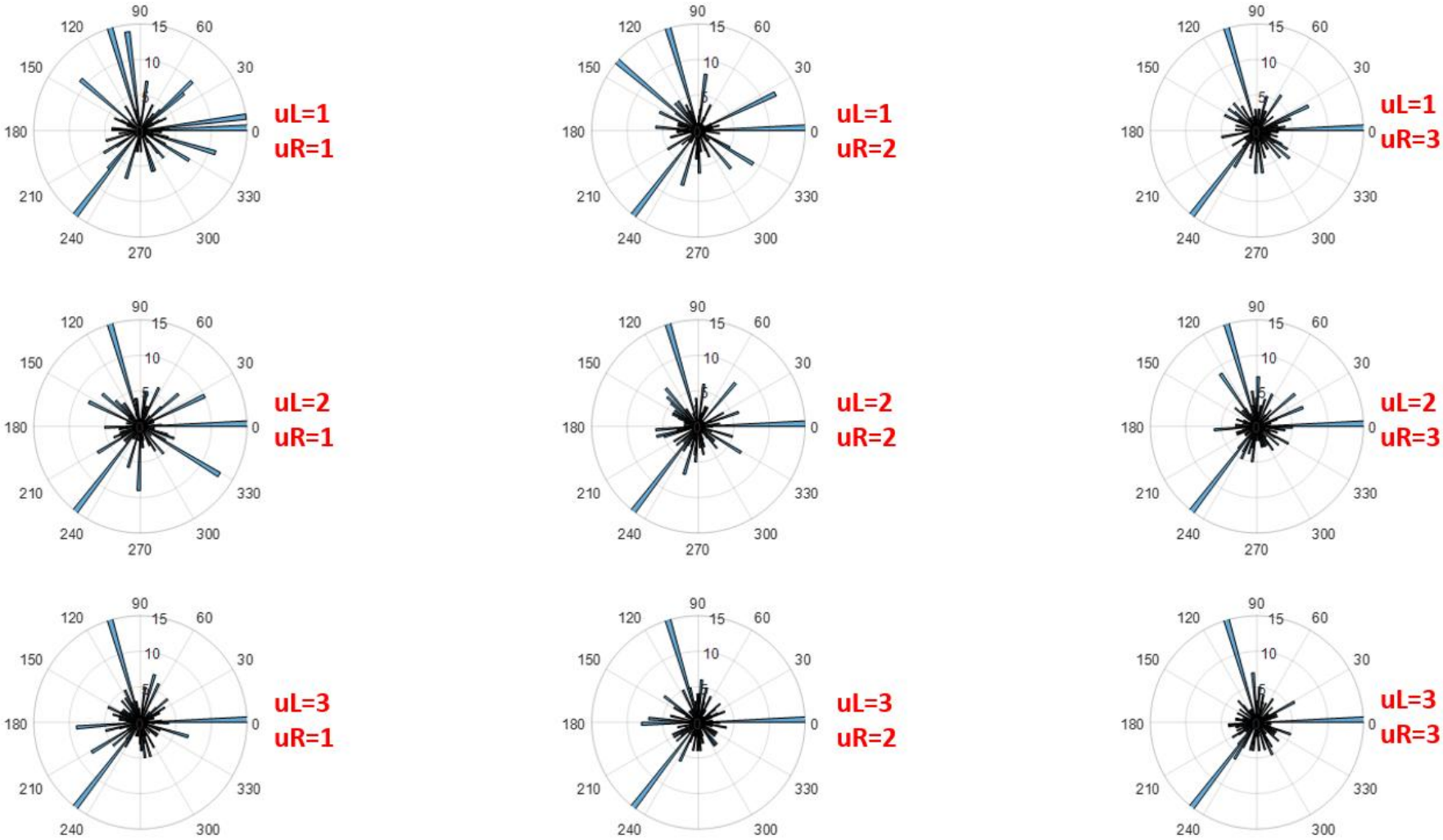
Bu çalışmada, istenmeyen e-postaların filtrelenmesi için karakterlerin UTF-8 Unicode değerleri arasındaki açı bilgilerini kullanarak içerik bağımsız bir yöntem önerilmiştir. Gelen e-postaların metin içerikleri öncelikle Unicode karakterlere dönüştürülmüş ve bu dönüşüm sonucu oluşan Unicode karakterler, bir boyutlu bir sinyal olarak ele alınmıştır. Daha sonra, bu sinyal üzerindeki her bir değerin, etrafındaki komşu değerlerle oluşturduğu açı bilgileri hesaplanmıştır. Bu açı yaklaşımı, e-posta içeriğinin özelliklerinden bağımsız bir istatistiksel yöntemdir. Açı dönüşümü sonucunda yeni bir sinyal oluşturulur ve bu açı sinyaline ait histogram, öznitelik vektörü olarak kullanılmıştır. Bu öznitelik vektörü, e-posta içeriklerinin Unicode karakterlerinin açısal dağılımını temsil eder. Bu önerilen yaklaşımın etkinliğini test etmek amacıyla Kaggle veri tabanından indirilen bir veri seti kullanılmıştır. Bu veri seti içerisinde spam ve non-spam (istenen) e-postalar bulunmaktadır. Bu e-postalar, açı örüntülerini kullanarak farklı makine öğrenme yöntemleri ile birbirinden ayrıştırılmıştır. Bu sınıflandırma işlemi için Naive Bayes (NB), Karar Destek Vektörleri (SVM), K-En Yakın Komşu (KNN) ve Random Forest (RF) gibi makine öğrenme yöntemleri kullanılmıştır. Sınıflandırma işlemleri, açık kaynak kodlu ve ücretsiz olarak kullanılabilen Weka programı ile gerçekleştirilmiştir. Ayrıca, sınıflandırma yöntemlerinin performansını değerlendirmek için 10-fold çapraz geçerlilik testi yapılmıştır. Bu test sonuçları, başarı sonuçlarını değerlendirmek için kullanılmıştır. Çizelge 4.1'de elde edilen başarı sonuçları ayrıntılı olarak sunulmuştur. Bu çalışma, gelen e-postaların içeriğinden bağımsız bir yöntem kullanarak spam ve non-spam e-postaları başarıyla sınıflandırmanın mümkün olduğunu göstermektedir. Açı yaklaşımı, e-posta filtreleme işlemlerinde potansiyel olarak etkili bir araç olarak değerlendirilebilir ve farklı makine öğrenme yöntemleri ile kullanılarak istenmeyen e-postaların tespitinde başarılı sonuçlar elde edilebilir.

Çizelge 4. 1. Başarı değerleri (%)

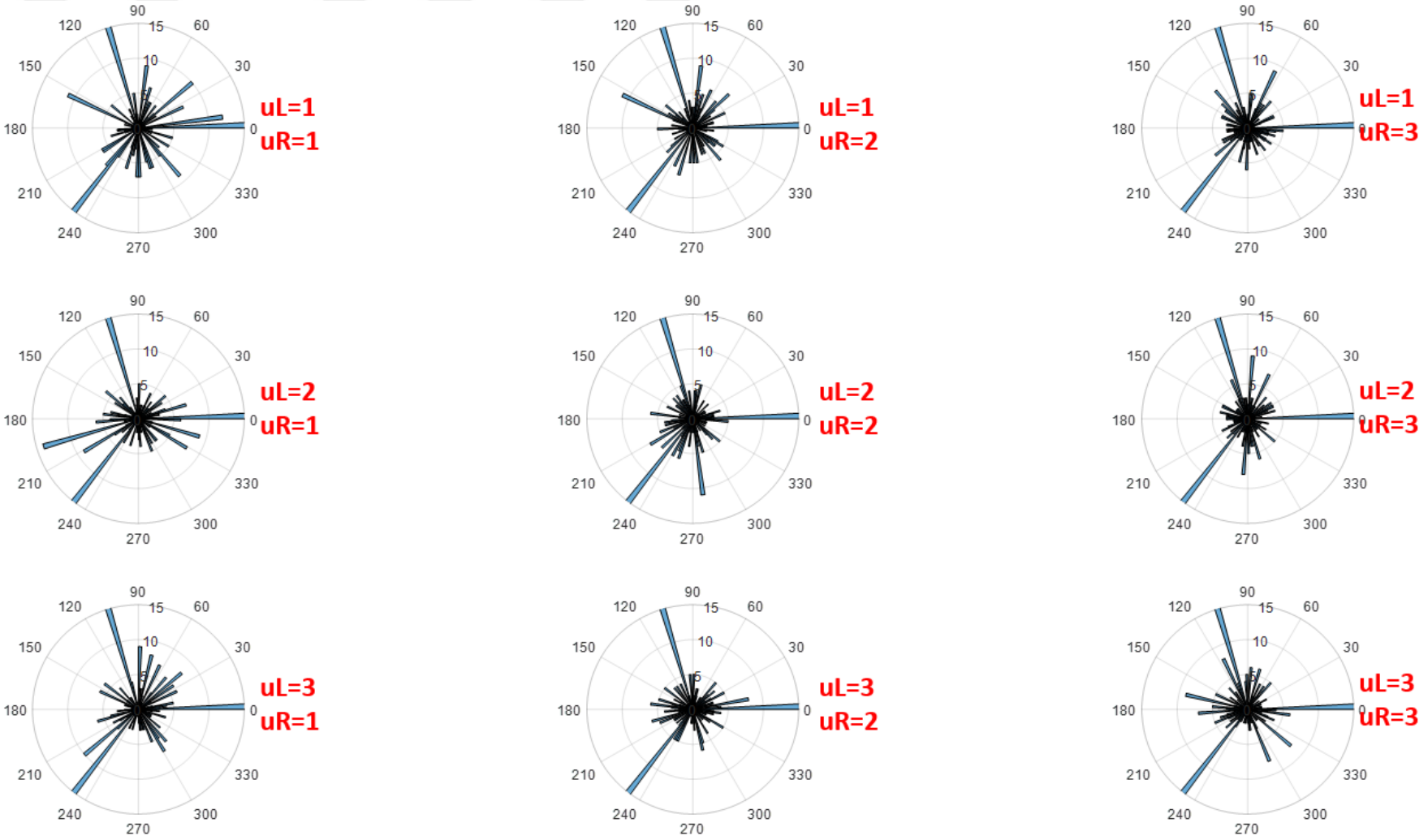
Ölçüt	**NB**	**SVM**	**KNN**	**RF**
Başarı (Accuracy)	0,889	0,933	0,942	0,924
Kesinlik (Precision)	0,885	0,931	0,940	0,931
Hatırlama (Recall)	0,890	0,933	0,942	0,924
F-Ölçütü (F-measure)	0,887	0,932	0,940	0,915

Çizelge 4.1'deki sonuçlara göre, K-En Yakın Komşu (Knn) modelinin en yüksek başarı elde ettiği görülmektedir, bu model %94,2 başarı oranına ulaşmıştır. Diğer yandan, Naive Bayes (NB) modeli en düşük başarıyı %88,9 oranında elde etmiştir. Genel olarak, diğer makine öğrenme yöntemleri ile de kabul edilebilir başarılar elde edilmiştir.

Önerilen açı yaklaşımı, tek boyutlu bir yapıdadır ve iki önemli parametre olan uR (Right) ve uL (Left) tarafından karakterizedir. Bu parametreler, e-postaların metin içeriklerinden farklı örüntülerin çıkarılmasına olanak tanır. Örneğin, uL ve uR parametrelerinin farklı değerlerine sahip e-postalar için elde edilen örnek örüntüler, Şekil Şekil 4.1 ve Şekil 4.2'de gösterilmiştir. Şekillerden anlaşıldığı gibi, uL ve uR parametrelerinin farklı değerlerine sahip olmaları hem spam hem de non-spam e-postalar için elde edilen açı örüntülerinin farklılık gösterebileceğini göstermektedir. Bu açı örüntülerinin uL ve uR parametrelerinin farklı değerlerine bağlı olarak farklı dağılımlar sergilediği gözlemlenmektedir. Bu durum, önerilen yöntemin e-postaları etkili bir şekilde sınıflandırmak için açı örüntülerini kullanarak farklı özelliklerin yakalanmasını sağladığını işaret etmektedir.



Şekil 4. 1. uL ve uR parametrelerin bir non-spam e-posta örneği için farklı değerlerine göre elde edilen örüntülerin dağılımları



Şekil 4. 2.uL ve uR parametrelerin bir spam e-posta örneği için farklı değerlerine göre elde edilen örüntülerin dağılımları

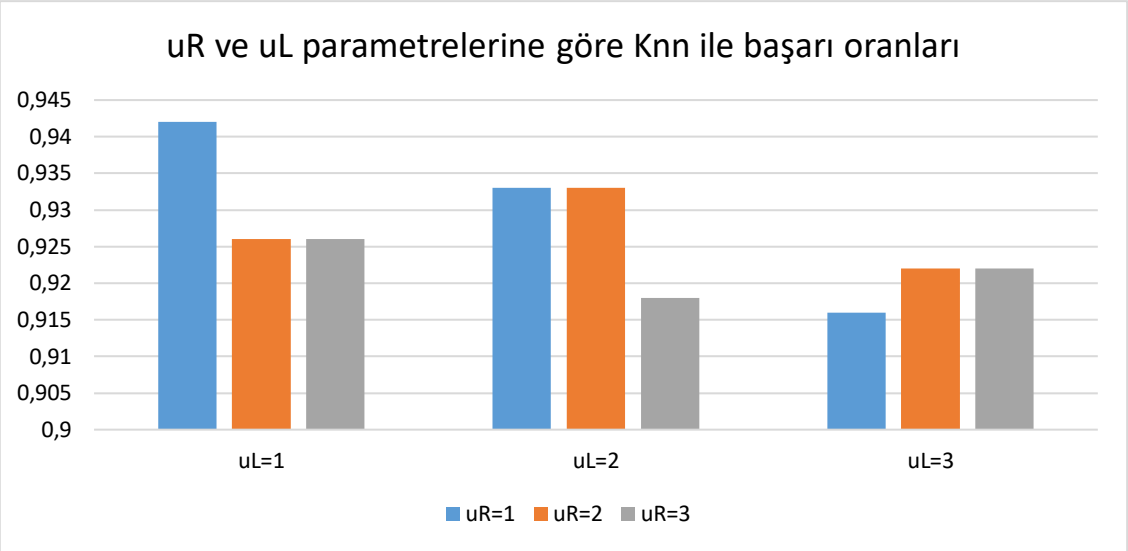
Açı yönteminin uR ve uL parametreleri ile elde edilen farklı örüntüler dikkate alınmıştır. Bu farklı örüntülerin performansını değerlendirmek amacıyla, farklı parametre değerleri ile elde edilen örüntüler kullanılarak K-En Yakın Komşu (Knn) ile sınıflandırma işlemi gerçekleştirilmiştir. Elde edilen başarı oranları Çizelge 4.2'de sunulmuştur.

Çizelge 4. 2. uL ve uR parametrelerin farklı değerlerine göre başarılar

uL, uR	Başarı	Kesinlik	Hatırlama	F-Ölçütü
uR=1,uL=1	0,942	0,940	0,942	0,940
uR=1,uL=2	0,933	0,930	0,933	0,930
uR=1,uL=3	0,916	0,913	0,916	0,914
uR=2,uL=1	0,926	0,923	0,926	0,923
uR=2,uL=2	0,933	0,931	0,933	0,930
uR=2,uL=3	0,922	0,919	0,922	0,919
uR=3,uL=1	0,926	0,923	0,926	0,923
uR=3,uL=2	0,918	0,915	0,918	0,915
uR=3,uL=3	0,922	0,918	0,921	0,918

Çizelge 4.2'ye göre, K-En Yakın Komşu (Knn) ile en yüksek başarı oranı uR=1 ve uL=1 olduğunda elde edilmiştir ve bu başarı oranı %94,2 olarak gözlemlenmiştir. uL ve uR parametrelerinin artmasıyla başarı oranlarında hafif bir azalma gözlenmiştir. Bu nedenle, uL ve uR parametrelerinin farklı değerleriyle farklı veri setlerinde denemeler yapılması gerektiği düşünülmektedir.

uL ve uR parametrelerinin farklı değerleriyle elde edilen Knn başarı oranlarına ait grafik Şekil 4.3'te sunulmuştur. Bu grafik, parametrelerin değişimi ile başarı oranlarının nasıl etkilendiğini göstermektedir.



Şekil 4. 3. uR ve uL parametrelerine göre Knn ile başarı oranları

5. TARTIŞMA

Spam, istenmeyen, genellikle ticari veya kötü niyetli içerik taşıyan iletileri ifade eder ve e-posta, mesajlaşma platformları, sosyal medya, web siteleri ve diğer iletişim kanalları üzerinden kullanıcılara gönderilir. Spam, kullanıcıların posta kutularını doldurarak istenmeyen içeriği eleme ve gerçek önemli iletilere erişme zorluğu yaratır. Bu nedenle spam tespiti büyük bir öneme sahiptir. Spam tespiti, kullanıcı deneyimini iyileştirmenin yanı sıra veri güvenliği risklerini azaltmayı hedefler. Spam içeriği, sahtekarlık, kimlik avı (phishing), zararlı yazılım bulaştırma gibi tehlikeleri barındırabilir. Ayrıca, spam mesajlarının fazla olması, iletişim altyapısının gereksiz yere yüklenmesine neden olabilir. Spam tespiti için kullanılan genel yaklaşımlar arasında kural tabanlı yöntemler (belirli kurallara dayalı olarak spam'ı tanımlama), Bayes teoremi (özellikle Naive Bayes sınıflandırıcı), makine öğrenimi yöntemleri (örneğin, karar ağaçları, destek vektör makineleri, k-en yakın komşu), DNS tabanlı yaklaşımlar (kara liste gibi) ve benzeri teknikler yer alır. Gün geçtikçe spamcılar da yöntemlerini geliştirerek spam ile mücadele zorlaşmaktadır. Bu nedenle spam tespiti alanında sürekli araştırma ve geliştirme önemlidir.

Spam e-postaları tespit etme ve filtreleme, modern iletişimde karşılaşılan önemli bir sorundur. Bu çalışmada, spam e-postaları sınıflandırmak için gelen e-postaların içeriğindeki karakterlerin UTF-8 Unicode (Unikod) değerleri arasındaki açı bilgilerini kullanarak karakter odaklı bir yaklaşım önerilmiştir. E-postaların içeriği Unikodlara dönüştürülerek, bu Unikod sinyalleri üzerinde açı hesaplamaları yapılmış ve açı yaklaşımı ile yeni bir öznitelik vektörü elde edilmiştir. Bu öznitelik vektörü, e-postaların spam veya spam olmayan kategorilere ayrılmasında kullanılmıştır. Bu yaklaşım, e-posta içeriğine bağımsız bir metot olup, farklı dillerdeki ve içeriklerdeki e-postaları etkili bir şekilde işleyebilme yeteneği sunmaktadır. Sonuçlar, farklı makine öğrenimi yöntemleri ile yapılan sınıflandırma işlemleri sonucunda değerlendirilmiştir. Bu sınıflandırma işlemleri, açı yaklaşımını kullanarak elde edilen Unikod verileri üzerinde gerçekleştirilmiştir. En yüksek başarı oranı %94,2 ile K-En Yakın Komşu (Knn) algoritması ile elde edilmiştir. Diğer makine öğrenimi yöntemleri de kabul edilebilir başarılar göstermiştir. Bu çalışma, spam e-postaların tespiti için karakter odaklı bir yaklaşımın etkili olabileceğini göstermektedir. Ancak, bu yaklaşımın daha fazla test edilmesi ve iyileştirilmesi gerekmektedir. Ayrıca, farklı dillerdeki e-postalar ve çok dilli veri kümeleri üzerindeki performansı daha fazla araştırma gerektirmektedir. Sonuç

olarak, bu çalışma spam e-postaların filtrelenmesi alanında yeni bir yaklaşım sunmaktadır. Gelecekteki çalışmalar, bu yaklaşımın daha da geliştirilmesi ve gerçek dünya uygulamalarındaki başarısının daha fazla test edilmesi yönünde olabilir.

Hedefler ve Kısıtlamalar

Bu çalışmanın temel hedefi, e-posta filtreleme için karakter odaklı bir yaklaşımın etkinliğini test etmek ve bu yaklaşımı gelen e-postaları sınıflandırmak için kullanabilecek makine öğrenimi yöntemleriyle değerlendirmekti. Bu hedefi gerçekleştirmek için, Unikod değerleri arasındaki açı bilgileri kullanılarak karakter odaklı açı yaklaşımı geliştirildi ve bu yaklaşımın performansı çeşitli makine öğrenimi yöntemleri ile test edildi. Ancak bu çalışmanın bazı kısıtlamaları bulunmaktadır. İlk olarak, Unikod değerleri arasındaki açı yaklaşımı, yalnızca karakterlere odaklanmaktadır ve diğer e-posta özelliklerini (örneğin, gönderen, konu, metin uzunluğu) dikkate almamaktadır. Bu, bazı durumlarda eksik veya yanıltıcı sonuçlara yol açabilir. Ayrıca, bu yaklaşımın farklı dillerdeki e-postalara uygulanabilirliği ve çok dilli veri kümeleri üzerindeki performansı daha fazla çalışma gerektirebilir. Diğer bir kısıtlama, bu çalışmanın sadece belirli bir veri kümesi üzerinde gerçekleştirilmiş olmasıdır. Bu, yaklaşımın genelleme yeteneğini ve çeşitli veri tipleri üzerindeki uygulanabilirliğini sınırlayabilir. Daha büyük ve farklı veri setleri üzerinde yapılan testler, yaklaşımın gerçek dünya uygulamalarındaki başarısını daha iyi yansıtabilir. Sonuç olarak, bu çalışma, karakter odaklı açı yaklaşımının e-posta filtrelemesi için potansiyel bir seçenek olduğunu göstermektedir, ancak daha fazla araştırma ve test gerektirmektedir. Bu yaklaşımın geliştirilmesi ve iyileştirilmesi, istenmeyen e-postaları daha etkili bir şekilde tespit etme amacıyla gelecekteki çalışmaların bir odak noktası olabilir.

6. KAYNAKLAR

A. Çıltık, and T. Güngör, "Time-efficient spam e-mail filtering using n-gram models," Pattern Recognition Letters, vol. 29, pp. 19-33, Jan. 2008.

Al-Kadhi, Mishaal Abdullah, Assessment of the status of spam in the Kingdom of Saudi Arabia, Journal of King Saudi University – Computer and Information Sciences 23 (2011) 45–58.

Amayri O., N. Bouguila, A study of spam filtering using support vector machines, Artificial Intelligence Review 34 (2010) 73–108.

Amir Herzberg, "DNS-based email sender authentication mechanisms: A critical review", 2009.

Androutsopoulos, I.I., Koutsias, J., Chandrinos, K.V., Spyropoulos, C.D., 2000. An experimental comparison of naive Bayesian and keywordbased anti-spam filtering with personal e-mail messages. In: Proceedings of the 23rd annual inter- national ACM SIGIR conference on Research and development in information retrieval. ACM, pp. 160–167.

Androutsopoulos,I.I, J. Koutsias, K.V. Chandrinos, G. Paliouras, C. Pyropoulos, An evaluation of naive Bayesian anti-spam filtering; in: Proceedings of the workshop on Machine Learning in the New Information Age, G. 11th European Conference on Machine Learning, Barcelona, Spain, 2000.

Ali-Vehmas Timo, Economical Impact of SPAM in the Internet; Raimo Kantola's Technical Report, 2003, pp. 147–158.

Bhat, S., Saritha, M., Yatakunta, P. R., Naik, P. S., & Bhat, P. (2022). Chronic Kidney Disease Prediction Using Naive Bayesian Classifier and K-NN Machine-Learning Algorithms. Research & Review: Machine Learning and Cloud Computing, 1(2), 1- 5.

Bhowmick, A. and Hazarika, S. M. 2016. Machine learning for E-mail spam filtering: review, techniques and trends. arXiv preprint arXiv:1606.01042.

Bhuiyan, H., Ashiquzzaman, A., Juthi, T.I., Biswas, S., 2018. A survey of existing e-mail spam filtering methods considering machine learning techniques. Global J. Comput. Sci. Technol.(C) 18 (1), 20–29.

Bozkir, A.S., Sahin, E., Aydos, M., Sezer, E.A., Orhan, F., 2017. Spam e-mail classification by utilizing n-gram features of hyperlink texts. In: 2017 IEEE 11th International Conference on Application of Information and Communication Technolo- gies (AICT), pp. 1–15.

B. Burton, Spamprobe-bayesian spam filtering tweaks, in: Proceedings of the Spam Conference, 2003.

Caruana, G., Li, M., 2012. A survey of emerging approaches to spam filtering. ACM Comput. Surv. 44 (2), 1–27.

C. Leung, Z. Liang, An Analysis of the Impact of Phishing and Anti-Phishing Related Announcements on Market Value of Global Firms, HKU Theses Online (HKUTO), 2009.

C.C. Lai, An empirical study of three machine learning methods for spam filtering, Knowledge-Based Systems 20 (3) (2007) 249–254.

C.C. Wang, S.Y. Chen, Using header session messages to anti-spamming, Computers & Security 26 (5) (2007) 381–390.

Carpinter, J, Hunt, R. Tightening the net: A review of current and next generation spam filtering tools. Computers & security, 25(8),2006, 566-578.

Cormack, G.V., 2007. Email spam filtering: a systematic review. Found. Trends Inf. Retriev. 1 (4), 335–455.

Dada, E.G., Bassi, J.S., Chiroma, H., Adetunmbi, A.O., Ajibuwa, O.E., 2019. Machine learning for email spam filtering: review, approaches and open research prob- lems 5 (6), 1–23.

D. Zhang, H. Xu, Z. Su, Y. Xu, "Chinese comments sentiment classification based on word2vec and SVMperf", Expert Systems with Applications, 42(4), 1857-1863, 2015.

E-Mail Usage in the United States, Statistics & Facts, Statista, 2022.

Ezpelet, E., Zurutuza, U., Hidalgo, J.M.G., 2016. Does sentiment analysis help in Bayesian spam filtering? In: International Conference on Hybrid Artificial Intelligence Systems, volume 9648. Springer International Publishing, pp. 79– 90.

Francisco Salcedo-Campos, Jesús Díaz-Verdejo, Pedro García-Teodoro (2012), Segmental parameterisation and statistical modelling of e-mail headers for spam detection. Information Sciences Volume 195 (2012) 45-61.

G. Şahin, "Turkish document classification based on Word2Vec and SVM classifier", Signal Processing and Communications Applications Conference, Antalya, Turkey,1-4,15-18 Mayıs, 2017.

Gansterer, W.N., Janecek, A.G.K., Neumayer, R., 2008. Spam filtering basedon latent semantic indexing. In: Berry, M.W., Castellanos, M. (Eds.), Survey of Text Mining II - Clustering, Classification, and Retrieval, vol. 2, pp. 165–185.

Gopalsamy, A., & Radha, B. (2022). Machine Learning-Based Ensemble Classifier Using Naïve Bayesian Tree with Logit Regression for the Prediction of Parkinson's Disease.

Guzella, TS, Caminhas, WM, A review of machine learning approaches to spam filtering. Expert Systems with Applications, 36(7), 2009, 10206-10222.

Hameed, S., Kloht, T. and Fu, X. (2013). Identity based email sender authentication for spam mitigation. Eighth International Conference on Digital Information Management (ICDIM 2013), IEEE, 14-19.

Idris I., Selamat A., Improved email spam detection model with negative selectionalgorithm and particle swarm optimization, Applied Soft Computing 22 (2014) 11-2.

J. Mason, Filtering spam with spamassassin, in: HEANet Annual Conference, 2002.E. Raymond, Bogofilter: A Fast Open Source Bayesian Spam Filters, 2005.

Kachhia, P., & Rathod, D. (2022). Kidney Disease Detection Using Supervised Machine Learning Techniques. Içinde Y.-D. Zhang, T. Senjyu, C. So-In, & A. Joshi (Ed.), Smart Trends in Computing and Communications (ss. 357-365). Springer.

Kadam, S., Gala, A., Gehlot, P., Kurup, A., Ghag, K., 2018. Word embedding based multinomial naive Bayes algorithm for spam filtering. In: 2018 Fourth International Conference on Computing Communication Control and Automation (IC- CUBEA). IEEE, pp. 1–5.

Karim, A., Azam, S., Shanmugam, B., Kannoorpatti, K. and Alazab, M. 2019. A Comprehensive Survey for Intelligent Spam Email Detection. IEEE Access, 7, 168261-168295.

Kaspersky, 2018. Spam and phishing in Q3 2018, Spam and phishing reports, Kaspersky.

Laorden, C., Santos, I., Sanz, B., Alvarez, G., Bringas, P.G., 2012. word sense disam- biguation for spam filtering. Electr. Comm. Res. Appl. 11 (3), 290–298.

Laorden, C., Santos, I., Sanz, B., Bringas, P.G., 2012. Enhanced topic-based vector space model for semantics aware spam filtering. Expert Syst. Appl. 39 (1), 437–444.

Liu Y., Wang L. Shi T. Li J., Detection of spam reviews through a hierarchical attention architecture with N-gram CNN and Bi-LSTM, Information Systems 103, 101865 (2022).

M. Sahami, S. Dumais, D. Heckerman, E. Horvitz, A Bayesian approach to filtering junk e-mail, in: Learning for Text Categorization: Papers from the 1998 Workshop, Madison, Wisconsin, AAAI Technical Report WS-98-05, 1998.

Makkar A., Kumar N., PROTECTOR: An optimized deep learning-based framework for image spam detection and prevention, Future Generation Computer Systems Volume 25, December 2021, 41-58.

Marcin Luckner, Michał Gad, Paweł Sobkowiak, “Stable web spam detection using features based on lexical items” , Faculty of Mathematics and Information Science, Warsaw University of Technology, Koszykowa 75, 00-662 Warszawa, Poland, Computer and Security 46 (2014) 79-93.

Messaging Anti-Abuse Working Group, MAAWG Email Security Awareness and Usage Report, 2010.

Murugavel U., Santhi R., Detection of spam and threads identification in E-mail spam corpus using content based text analytics method, Materials Today: Proceedings 33 (2020) pp.3310-3323

N. Mostafa Raad, G. Alam, B. Zaidan, A. Zaidan, Impact of spam advertisement through e-mail: a study to assess the influence of the anti-spam on the email marketing, Afri. J. Bus. Manage. 4 (2010) 2362–2367

Nakov, PI, Dobrikov, PM, Non-parametric SPAM filtering based on kNN and LSA. In Proceedings of the 33th National Spring Conference of the Bulgarian Mathematicians Union, pp. 1-4, 2004.

Neisari A., Rueda L., S. Sherif, Spam review detection using self-organizing maps and convolutional neural networks, Science Direct Computer &Security Volume 106, 102274 (2021).

Nizam, H., & Akın, S. S. (2014). Sosyal medyada makine öğrenmesi ile duygu analizinde dengeli ve dengesiz veri setlerinin performanslarının karşılaştırılması. XIX. Türkiye’de İnternet Konferansı, 1(6).

Pérez-Díaz, N., Ruano-Ordás, D., Mendez, J.R., Galvez, J.F., Fdez-Riverola, F., 2012. Rough sets for spam filtering: Selecting appropriate decision rules for boundary e-mail classification. Appl. Soft Comput. 12 (11), 3671–3682.

Peter, I. (2004). The history of email. Internet History Project. Retrieved from http://www.nethistory.info/History of the Internet/email.html.

Radwa M.K. Saeed, Sherine Rady, Tarek F. Gharib, “An ensemble approach for spam detection in Arabic opinion texts”,Journal of King Saud University – Computer and Information Sciences 34 (2022) 1407-1416.

Ren, Y., Ji, D., 2017. Neural networks for deceptive opinion spam detection:an empirical study. Inf. Sci. 385, 213–224.

Rosita J.D., Jacob W.S., Multi-Objective Genetic Algorithm and CNN-Based Deep Learning Architectural Scheme for effective spam detection, International Journal of Intelligent Networks, Volume 3, 2022, Pages 9-15.

Saidani N, Adi K. Allili M.S., A semantic-based classification approach for an enhanced spam detection, Computer & Security Volume 94. 101716 (2020)

Sharma, S. K. ve Sharma, V. (2012). Time Series Prediction Using Knn Algorithms Via Euclidian Distance Function: A Case Of Foreign Exchange Rate Prediction. Asian Journal of Computer Science and Information Technology, 2 /7, 219 – 221.

Singh A. and Batra S., "Ensemble based spam detection in social IoT using probabilistic data structures," Future Generation Computer Systems, vol. 81, pp. 359–371, 2018.

Singh, V.K., Bhardwaj, S., 2018. Spam mail detection using classification techniques and global training set. In: Intelligent Computing and Information and Communication, 673, pp. 623–632.

Sokhangoee Z. F., Rezapour A., 2022. A novel approach for spam detection based on association rule mining and genetic algorithm.

Song, L., Lau, R.Y.K., Kwok, R.C.W., Mirkovski, K., Dou, W., 2017. Who are the spoilers in social media marketing? Incremental learning of latent semantics for social spam detection. Electr. Comm. Res. 17 (1), 51–81.

Symantec, 2018. Internet Security Threat Report, Technical report, Symantec.

Torabi, Z.S., Nadimi-Shahraki, M.H., Nabiollahi, A., 2015. Efficient support vector machines for spam detection: a survey. Int. J. Comput. Sci. Inf. Secur. 13 (1), 11.

T Cover, P Hart - IEEE transactions on information theory, 1967 - ieeexplore.ieee.org

Vapnik, V. (1963). Pattern recognition using generalized portrait method. Automation and remote control, 24, 774-780.

Vorakulpipat, C., Visoottiviseth, V., Siwamogsatham, S., 2012. Polite sender: a resource-saving spam email countermeasure based on sender responsibilities and recipient justifications. Comput. Secur. 31 (3), 286–298.

Wu, T., Wen, S., Xiang, Y., Zhou, W., 2018. Twitter spam detection: Survey of new approaches and comparative study. Comput. Secur. 76, 265–284.

Y. Ren, L. Wang, X. Li, M. Pang J. Wei, "Stochastic optimization for bayesian network classifiers," Applied Intelligence, pages 1–21, 2022.