

DIYABETİK RETİNOPATİ TESPİTİ İÇİN ATOM ARAMA OPTİMİZASYONU İLE ÖZELLİK SEÇİMİ YÖNTEMİ

Feyza ALTUNBEY ÖZBAY¹, Erdal ÖZBAY^{2*}

¹Fırat Üniversitesi, Mühendislik Fakültesi, Yazılım Mühendisliği Bölümü, Elazığ, 23119, Türkiye

²Fırat Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Elazığ, 23119, Türkiye

Geliş Tarihi/Received Date: 10.11.2021 Kabul Tarihi/Accepted Date: 08.02.2022 DOI: 10.54365/adyumbd.1021738

ÖZET

Diyabetik Retinopati (DR), dünya genelinde milyonlarca insanı etkileyen görme kaybı ve körlüğün başlıca nedenidir. DR tespiti için retinal fundus görüntülerini kullanan birçok farklı bilimsel ve tıbbi yaklaşımlar bulunmaktadır. Bu yaklaşımların çoğunda, özellik seçimi aşaması yapılmadan diyabetik retinopati veri kümelerine çeşitli makine öğrenimi ve derin öğrenme yaklaşımları uygulanmıştır. Çalışmada UCI makine öğrenmesi deposundan elde edilen DR veri kümesi kullanılmıştır. Bu makalede, atom dinamiklerinden faydalanılarak önerilmiş popülasyon temelli yeni bir metasezgisel yöntem olan Atom Arama Optimizasyon (AAO) algoritması, ilk kez DR veri kümesi için bir özellik seçim yöntemi olarak kullanılmıştır. Normalize edilen veri kümesine AAO algoritmasının uygulanmasının ardından elde edilen yeni veri kümesi Bagging, CvR, IBk, JRip, Kstar ve SimpleCart olmak üzere altı sınıflandırma algoritması ile test edilmiştir. Aynı sınıflandırma algoritmaları, orijinal DR veri kümesine de uygulanmıştır. Elde edilen sonuçlar AAO algoritması ile özellik seçimi yapılmış veri kümesi ile karşılaştırılmıştır. Önerilen modelin performansı doğruluk, duyarlılık, özgüllük, kesinlik, f-ölçütü ve roc alanı değerleri bakımından değerlendirilmiştir. Elde edilen sonuçlar, AAO algoritması ile özellik seçimi yapılmış veri kümesi üzerinde Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmaları ile daha iyi değerler elde edildiğini göstermektedir. Bu bakımdan önerilen özellik seçimi ile algoritmaların özellik seçimi olmadan elde edilen sınıflandırma oranlarında doğruluk için ortalama %2.7, duyarlılık için %3.5, özgüllük için %2'lik bir artış sağlanmıştır.

Anahtar Kelimeler: Atom arama optimizasyon algoritması, Diyabetik Retinopati, Özellik seçimi, Sınıflandırma

FEATURE SELECTION METHOD WITH ATOM SEARCH OPTIMIZATION FOR DETECTION OF DIABETIC RETINOPATHY

ABSTRACT

Diabetic Retinopathy (DR) is the leading cause of vision loss and blindness, affecting millions of people worldwide. There are many different scientific and medical approaches that use retinal fundus images for DR detection. In most of these approaches, various machine learning and deep learning approaches have been applied to diabetic retinopathy datasets without the feature selection step. The DR dataset obtained from the UCI machine learning repository was used in the study. In this article, Atom Search Optimization (ASO) algorithm, a new population-based metaheuristic method proposed by utilizing atom dynamics, is used for the first time as a feature selection method for the DR dataset. Applied the ASO algorithm to the normalized dataset, the new dataset was tested by six classification algorithms: Bagging, CvR, IbK, JRip, Kstar, and SimpleCart. The same classification algorithms were applied to the original DR dataset. The results obtained were compared with the data set that was selected with the ASO algorithm. The performance of the proposed model was evaluated in terms of accuracy, sensitivity, specificity, precision, f-measure, and roc curve values. The results show that better values were obtained with Bagging, CvR, IbK, JRip, Kstar, and SimpleCart algorithms on the dataset selected with the ASO algorithm. In this regard, an increase of 2.7% for the average accuracy, 3.5% for the sensitivity, and 2% for the specificity were achieved in the classification rates obtained without feature selection of the algorithms with the proposed feature selection.

Keywords: Atom search optimization algorithm, Diabetic Retinopathy, Feature selection, Classification

*¹ e-mail: faltunbey@firat.edu.tr ORCID ID: <https://orcid.org/0000-0003-0629-6888>,

² e-mail: erdalozbay@firat.edu.tr ORCID ID: <https://orcid.org/0000-0002-9004-4802> (Sorumlu Yazar),

1. Giriş

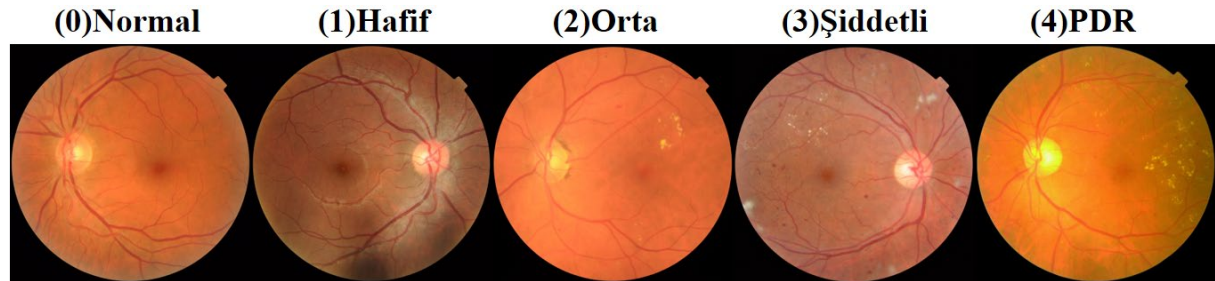
Diyabet, pankreasın yetersiz insülin üretmesi veya vücudun bunu gerektiği gibi kullanamaması sonucu kandaki glikoz miktarını artıran kronik bir hastalıktır. İnsülin, vücudumuzdaki hücrelerin yiyeceklerde glikoz kullanmasına ve vücudun şeker seviyesini kontrol etmesine yardımcı olan bir hormon türüdür. Yetersiz miktarda insülin, vücuttaki glikoz seviyesinde artışa neden olmaktadır. Bu durum da kişinin diyabetik olmasına sebep olmaktadır [1]. Uluslararası Diyabet Federasyonu verilerine göre 2017 yılında dünya genelinde 451 milyon yetişkinin diyabetle yaşadığı tahmin edilmektedir [2]. Çin, Hindistan ve ABD dünyanın en yüksek diyabet popülasyonuna sahip ilk üç ülkesidir [3]. Diyabet süresinin uzaması ile diyabet retina dâhil tüm vücudu etkilemektedir. Diyabetik Retinopati (DR), diyabetin ileri evrelerinde körlüğe neden olan en yaygın ve tehlikeli yan etkilerinden biridir. Dünya çapında, körlüğün %2.6'sının nedeni DR'dir [4]. Uzun süredir hastalıktan muzdarip diyabet hastalarında DR varlığı olasılığı artmaktadır. Diyabet hastalarının körlük riskinden korunmak için, DR'nin erken bir aşamada teşhis ve tedavi edilmesi gerekmektedir. Bu nedenle düzenli retina taraması yapılmalıdır [5]. DR, bir retina görüntüsünde mikroanevrizmalar, kanamalar, yumuşak ve sert eksüdalar gibi farklı tip lezyonların ortaya çıkmasıyla saptanmaktadır [6].

Mikroanevrizmalar (MA): Perisitlerin (kasılma hücreleri) kaybı, mikroanevrizmalar olarak bilinen retinal kılcal damarların üzüm şeklinde veya iğ şeklinde genişlemesiyle vasküler hücreler arası temaslarda çeşitli değişikliklere yol açabilmektedir. Mikroanevrizmalar, DR'nin klinik olarak görülebilen en erken belirtisidir. Muayenede mikroanevrizmalar, retinanın iç nükleer tabakasında bulunan küçük kırmızı noktalar olarak görünmektedir [7].

Retina içi kanamalar: Retinada daha büyük noktalar olarak görünmektedir. Alev (yüzeysel) ve leke (daha derin) olmak üzere iki tür kanama bulunmaktadır [7].

Sert eksüdalar: Merkezi retinadaki anormal kan damarlarından sızan ve merkezi görmede bulanıklığa neden olan ve retinada sarı (lipoproteinli tortular) lekeler olarak görünen lipidlerdir [7].

Yumuşak eksüdalar (Pamuk yünü): Sinir lifinin şişmesi nedeniyle retinada beyaz lekeler olarak görünmektedir [7].



Şekil 1. DR'nin farklı evreleri

Kırmızı lezyonlar mikroanevrizma ve kanama iken parlak lezyonlar yumuşak ve sert eksudalardır. DR'nin iki ana tipi vardır: Proliferatif Olmayan Diyabetik Retinopati (NonProliferative Diabetic Retinopathy - NPDR) ve Proliferatif Diyabetik Retinopati (Proliferative Diabetic Retinopathy - PDR). Erken aşamalarda NPDR olarak adlandırılan DR, hafif, orta ve şiddetli olmak üzere 3 aşamadan oluşmaktadır [8]. Hafif evrede, kan damarlarının ucunda küçük dairesel kırmızı noktalar olarak bilinen MA'lar bulunmaktadır. Orta evrede, MA'lar daha derin katmanlara geçer ve retinada alev şeklinde bir kanama oluşturur. Şiddetli evre, dört kadranın her birinde, belirgin intraretinal mikrovasküler anormalliklerle birlikte kesin venöz kanamaya sahip 20'den fazla intraretinal kanama içermektedir. PDR, retinanın iç yüzeyinde büyüyen fonksiyonel mikrovasküler ağlar şeklinde yeni kan damarlarının doğal bir oluşumu olan neovaskülarizasyona yol açan DR'nin ileri aşamasıdır [9]. DR'nin farklı evreleri Şekil 1'de gösterilmiştir. Şekil incelendiğinde, normal ve hafif evrelerin görsel olarak benzer olduğu görülmektedir. Bu nedenle, hafif evreyi tespit etmek zordur. Bu aşamalar ve barındırdıkları çeşitli lezyonlar Çizelge 1'de kısaca özetlenmiştir.

Çizelge 1. DR'nin seviyeleri ve ilişkili lezyonları.

DR Şiddet Düzeyi	Lezyonlar
DR Yok	Lezyon yok
Hafif DR	Yalnızca mikroanevrizmalar
Orta DR	Mikroanevrizmalardan daha fazla fakat şiddetli DR'den daha az
Şiddetli DR	Aşağıdakilerden herhangi biri: <ul style="list-style-type: none"> • 4 retinal kadranın her birinde 20'den fazla intraretinal HM, • Venöz boncuklanmalar, • 1'den fazla kadranda belirgin intraretinal mikrovasküler anormallikler • Proliferatif DR belirtisi yok
Proliferatif DR	Aşağıdakilerden biri veya birkaçı: <ul style="list-style-type: none"> • Vitreus/preretinal HM, • Neovaskülarizasyon

Tıbbi verilerin boyutları farklı birçok özellikler içermektedir ve her biri farklı değer türlerinden oluşmaktadır. Gürültü, aykırı değerler, eksik veya yinelenen veriler veri kalitesiyle ilgili sorunlara neden olmaktadır. Ham verileri daha sonra yapılacak analizler için uygun hale getirmek için veri hazırlamaya odaklanarak ön işleme adımları uygulamak gerekmektedir. Bu problemlerin bazılarının üstesinden gelmek için özellik seçme yöntemi kullanılabilir. Özellik seçimi, sınıflandırıcının performansını artırmak ve hesaplama yükünü azaltmak için önemli bir adımdır. Ayrıca, boyutun azaltılması alakasız özellikleri ortadan kaldırabilirken, gürültünün azaltılması, daha az özelliğin ilişkilendirilmesi nedeniyle daha sağlam öğrenme modelleri üretilmesini sağlayabilmektedir.

Bu çalışmada, DR tespiti için DR veri kümesinden özellik seçimi yapmak amacıyla atomik hareketi taklit eden, yakın zamanda önerilen fizikten ilham alan Atom Arama Optimizasyonu (AAO) algoritmasını kullanılmıştır. Bu bakımdan popülasyon tabanlı bir metasezgisel bir algoritma olan AAO algoritmasını kullanan yeni bir yöntem önerilmektedir. Önerilen bu yöntemde, AAO algoritması ile özellik seçimi yapılarak elde edilen yeni veri kümesine ve orjinal veri kümesine Bagging, Classification via Regression (CvR), IBk, JRip, Kstar ve SimpleCart olmak üzere altı farklı sınıflandırma algoritması uygulanmıştır. Algoritmalarından elde edilen sonuçlar doğruluk, duyarlılık, özgüllük, kesinlik, f-ölçütü ve roc alanı olmak üzere birçok farklı performans değerlendirme kriterine göre karşılaştırılmıştır. Karşılaştırma sonuçlarına göre, DR tespiti için AAO algoritmasının özellik seçimi yöntemi olarak kullanılmasının Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmalarının performansları üzerindeki etkisi değerlendirilmiştir.

Çalışmanın geri kalanı şu şekilde tasarlanmıştır: İkinci bölüm literatürde DR teşhisi için önerilen farklı yöntemleri sunmaktadır. Özellik seçimi için kullanılan AAO algoritması, sınıflandırma algoritması olarak kullanılan Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmaları ve sınıflandırma algoritmalarını analiz etmek için kullanılan performans değerlendirme metrikleri üçüncü bölümde detaylıca incelenmiştir. Dördüncü bölümde, çalışma kapsamında kullanılan DR veri kümesi açıklanmıştır. Aynı zamanda, önerilen yöntemin performans değerlendirmesi bu bölümde yapılmıştır. Son olarak, beşinci bölümde çalışmanın önemi ve sonuçları vurgulanmıştır.

2. Literatür Taraması

Körlüğe neden olan DR'nin erken evrelerde saptanması DR'nin önlenmesi önemli bir esastır. Bu nedenle DR'nin tespit edilebilmesi için alanında uzman araştırmacılar tarafından araştırma ve ilgi alanlarına göre pek çok farklı çalışma yapılmıştır. Tıp alanında önerilen yöntemlerin dışında, DR'nin tespiti için araştırmacılar birçok farklı makine öğrenmesi ve derin öğrenme yöntemleri önermiş ve uygulamışlardır.

Sahlsten ve diğerleri, DR'nin belirlenmesi için yüksek doğruluğa sahip derin öğrenme temelli bir yaklaşım geliştirmişlerdir [10]. Analizlerini, yüksek çözünürlüklü görüntülerin üzerinde yapmışlardır. Elde edilen sonuçlar, maliyetteki sınırlamalar göz önünde bulundurulduğunda derin öğrenme modellerinin hastalığı teşhis etmekte kabiliyetinin yüksek olduğunu göstermektedir.

Gondal ve diğerleri, DR'nin tespiti için (Evrışimli Sinir Ağı) ESA modeli kullanan bir yöntem önerilmiştir. Kaggle veri kümesini eğitim ve DiaretDB1 veri kümesi test kümesi olmak üzere genel erişime açık iki veri kümesi kullanılmıştır. DiaretDB1 veri kümesinde, önerilen ESA modeli %93.6 duyarlılık ve %97.6 özgüllük ile sonuçlanmıştır [11].

Li ve diğerleri, farklı bir çalışmada, renkli fundus görüntüleri işlenerek DR'nin tespiti için veri odaklı derin öğrenme modeli önerilmiştir. Önerilen model, hastalıklı görüntüleri sağlıklı görüntülerden ayırmayı sağlamıştır [12].

Retinal fundus görüntülerinden DR ve makula ödemi saptamak için derin öğrenme modeli tasarlanmıştır [13]. 128,175 retinal görüntüden veri setini eğitmek için derin bir evrışimsel sinir ağı kullanılmıştır [14]. Çalışmadan elde edilen duyarlılık ve özgüllük oranları, derin bir sinir ağı modelinin, diyabetik hastalar arasında yönlendirilebilir diyabetik retinopatinin saptanmasındaki başarılarını göstermektedir.

Raman ve diğerleri, retinal fundus görüntülerinden anormal özelliklerin varlığını tespit ederken, retinal görüntülemedeki anormalliği bulmak için bilgisayar destekli algılama mekanizmasını geliştirmişlerdir. Önerilen yöntem, görüntülerin iyileştirilmesi, gürültünün filtrelenmesi, kan damarlarının tespiti ve optik diskin tanımlanması, eksüdaların ve mikroanevrizmaların çıkarılması, özelliklerin çıkarılması ve DR'nin hafif, orta, şiddetli NPDR ve PDR olmak üzere çeşitli aşamalarının makine öğrenmesi yöntemleri ile sınıflandırılmasına odaklanmıştır [15].

Porwal ve diğerleri, lezyon segmentasyonunun önceki aşamasından kaçınarak hasta ve sağlıklı vakaları ayırt etmek için renkli fundus görüntülerinin dokusunu tanımlamışlardır. Bu çalışmada, AdaBoost, rastgele orman, lojistik regresyon ve destek vektör makinelerini içeren farklı birçok sınıflandırıcı kullanılmıştır [16].

Hemanth ve diğerleri, retinal fundus görüntülerinden diyabetik retinopati tanısı için alternatif hibrit bir çözüm yöntemi önerilmiştir. Hibrit yöntemde, görüntü işleme ve derin öğrenmeyi daha iyi hale getirmek için kullanılmıştır. Histogram eşitleme ve kontrast sınırlı uyarlanabilir histogram eşitleme teknikleri ile görüntü işleme yapılmıştır. Daha sonra, tanı konvolüsyonel bir sinir ağı ile gerçekleştirilmiştir. Önerilen yöntemin etkinliği, MESSIDOR verikümesindeki 400 retinal fundus görüntüsü kullanılarak test edilmiştir [17].

Shahin ve diğerleri, farklı bir çalışmada, retinal fundus görüntülerini proliferatif diyabet retinopati olan ve olmayan olarak otomatik olarak sınıflandırmak için bir sistem geliştirilmiştir. Entropi ve homojenlik dâhil olmak üzere iki indeksin yanı sıra kan damarları alanı ve eksüda alanı gibi patolojik özellikleri çıkarmak için morfolojik işlemeyi benimsenmişlerdir. Bu özellikler sık bir sinir ağını beslemektedir [18].

Eksüda lezyonunu bölümlere ayırmak için yukarıdan aşağıya bölümlenme ve sert eksüdaların şiddetini derecelendirmek için fovea merkezli bir kutupsal koordinat sistemi olmak üzere iki bölümden oluşan otomatik bir algoritma öneren farklı bir yöntem 236 fundus görüntüsünden oluşan küçük bir veri kümesi üzerinde test edilmiştir [19].

Diyabetik retinopati olan ve olmayan insanları %90'dan fazla doğrulukla ayırt etmek için rastgele orman algoritmasını kullanan bir yöntem önerilmiştir. Önerilen yöntem dereceli fundus fotoğrafları ve sistemik verilere dayanarak DR riskini değerlendirmektedir [20].

Gargeya ve Leng, derin evrışimli sinir ağını geleneksel bir makine öğrenme algoritması ile birleştiren bir yöntem önermişlerdir. Çalışmada, fundus görüntüleri ön işlemden sonra kalıntı bir ağı beslemekte ve daha sonra ağın son havuzlama katmanından elde edilen görüntülerin karakterizasyonu,

birkaç meta veri değişkeni ile birlikte, sağlıklı ve sağlıklı olmayan arasında ayrım yapmak için bir karar ağacı sınıflandırıcısına gönderilmektedir [21].

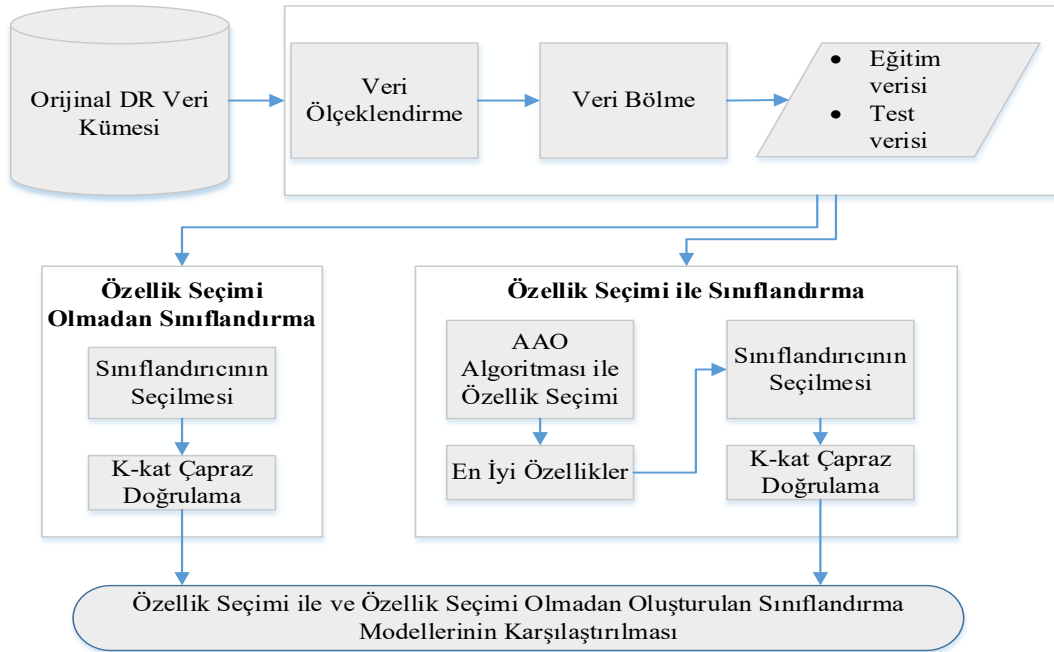
Akram ve diğerleri, retina görüntülerinde mikroanevrizmaları, kanamaları, sert eksüdalı, pamuk yünü lekelerini veya diğer lezyon türlerini saptamak için yeni bir hibrit sınıflandırıcı teknik tasarlamışlardır. Önerilen tekniğin performansını ve sınıflandırmadaki doğruluğunu geliştirmek için araştırmacılar m-medoids ile Gauss karışım modelini birleştirmişlerdir. Yöntemin etkinliğini değerlendirmek için standart fundus görüntü veritabanları kullanılmışlardır [22].

Welikala ve diğerleri yeni bir modifiye hat operatörü ve standart bir hat operatörü kullanarak proliferatif DR'yi sınıflandırmak için neovaskülarizasyonu tespit etmek için bir yaklaşım geliştirmişlerdir. Destek vektör makinesi, her bir özellik kümesi üzerinde son sınıflandırmayı gerçekleştirmek için kullanılmıştır. Önerilen tekniğin performansı değerlendirmek için yalnızca 60 görüntüden oluşan küçük bir veri tabanı kullanılmıştır [23].

Waheed ve diğerleri, retinal kan damarı segmentasyonu için ortam sınıflandırmasına dayalı bir hibrit özellik veri kümesi kullanan bir yöntem sunmuştur. Araştırmacılar, normal kan damarlarını ve diyabetin bir sonucu olarak büyüyen sahte kan damarlarını ayırt etmek için yeni bir bölge tabanlı hibrit özellik seti kullanmışlardır [24].

Zhang ve diğerleri DR'yi saptamak için eksüdalın saptanması için renkli retinal görüntüler üzerinde araştırma yapmışlardır. Bu tekniğin performansını belirlemek için yeni bir klinik veri tabanı olan E-optha EX kullanılmıştır. Bu veri seti, konturlu eksüdalı içermektedir. Eksüda tespiti için önerilen bu yöntem, aynı hastanın iki farklı muayenesi arasında hastalığın büyümesinin belirlenmesi gibi karmaşık problemler için geliştirilmiştir. Ön işleme adımı gürültü giderme, normalleştirme ve eksüdalın yanlış saptanmasına yol açabilecek görüntüdeki artefaktların saptanmasına dayanmaktadır. Aday segmentasyon tekniği, klasik ve bağlamsal özellikleri kullanan matematiksel morfolojiye dayalı olarak tanıtılmıştır. Son bir adım olarak, eksüdalı segmentlere ayırmak için rastgele orman algoritması uygulanmıştır [25].

3. Önerilen Yöntem



Şekil 2. Önerilen yöntemin mimarisi

Bu bölümde, DR tespiti için önerilen yöntem detaylı olarak incelenmiştir. Yöntemin ilk adımında, StandardScaler yöntemi kullanılarak özellik kümesi ölçeklendirilmiştir. Veri kümelerinin ölçeklendirilmesi, aykırı değerlerin kaldırılmasını, verilerin dönüştürülmesini ve verilerin normalleştirilmesini içermektedir [26]. DR tespiti için verilere bir sınıflandırıcı algoritması uygulamadan önce veri kümesinin dikkatli bir şekilde değerlendirilmesi önemlidir. Gereksiz ve alakasız birçok özelliği kullanmak yerine sınıflandırma sürecini daha verimli hale getirmek için sadece gerekli özellikleri dikkate alınabilir. Bu nedenle, ilgili ve önemli özelliklerin seçebilmesi için özellik seçimi yöntemleri kullanılmaktadır. Ayrıca, sınıflandırma uygulamalarında özellik seçimi yöntemlerinin kullanılarak önemli özelliğin bulunması sınıflandırıcının iş yükünün azaltılmasına yardımcı olmaktadır. Bu sayede sınıflandırma algoritmalarının doğruluğu artmaktadır. Bu çalışmada DR veri kümesinden özellik seçimi için atom dinamiklerinden faydalanılarak önerilmiş popülasyon temelli yeni bir algoritma olan Atom Arama Optimizasyon (AAO) algoritması kullanılmıştır. Veri normalizasyon adımından sonra, sınıflandırma algoritmaları için en uygun özelliklerin belirlenmesi için normalize edilmiş veri kümesine AAO algoritması uygulanarak özellik seçimi yapılmıştır. En uygun özellik kümesi seçildikten sonra elde edilen yeni DR veri kümesine Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmaları uygulanmıştır. Aynı zamanda orijinal veri kümesine de aynı sınıflandırma algoritmaları uygulanmıştır. Özellik seçimi ile oluşturulan yeni veri kümesi ve orijinal veri kümesinden elde edilen sonuçlar birçok farklı metriğe göre karşılaştırılarak AAO algoritması ile özellik seçimi yapmanın DR tespiti üzerindeki etkiliği test edilmiştir. Önerilen yöntemin mimarisi Şekil 2’de verilmiştir.

3.1. Atom Arama Optimizasyon (AAO) Algoritması

AAO algoritması, küresel optimizasyon problemleri için etkili bir arama mekanizması tasarlamak için etkileşim ve kısıtlama kuvvetleri tarafından kontrol edilen atomik hareketi taklit eden yakın zamanda önerilen fizikten ilham alan, popülasyon tabanlı bir sezgisel algoritmadır [1]. d . boyutta i . atoma etki eden toplam etkileşim kuvvetleri, aslında dinamik olarak değişen komşu atomların i . atomuna uyguladığı itme ve çekimin vektör toplamıdır. Bu durum denklem (1) ile ifade edilmiştir [27].

$$F_i^d(t) = \sum_{j \in K_{best}} rand_j F_{ij}^d(t) \quad (1)$$

$rand_j$ [0,1] aralığında rastgele bir sayıdır. K_{best} en iyi uygunluk fonksiyonu değerlerine sahip ilk K atomdan oluşan atom popülasyonunun bir alt kümesidir. AAO algoritmasının ilk iterasyonlarda daha fazla keşif yapmasını sağlamak için, her atomun K komşuları kadar daha iyi uygunluk değerlerine sahip çok sayıda atomla etkileşime girmesi ve algoritmanın son iterasyonlarda daha fazla sömürü yapmasını sağlamak için, her atomun K komşuları kadar daha iyi uygunluk değerlerine sahip birkaç atomla etkileşime girmesi gerekir. Bu nedenle, K değerinin denklem (2)’deki gibi yinelemeler boyunca kademeli olarak azaltılması gerekmektedir [27].

$$K(t) = N - (N - 2) \times \sqrt{\frac{t}{T}} \quad (2)$$

N atomik sistemdeki atomların toplam sayısıdır. t mevcut iterasyon ve T maksimum iterasyon sayısıdır. Denklem (1)’deki F_{ij}^d j . en iyi atomun d . boyutta i . atoma uyguladığı etkileşim kuvvetidir. Etkileşim kuvveti Lennard-Jones (L-J) potansiyelinin gradyanıdır ve optimizasyon problemlerini ele almak için, bu modelin pozitif çekim ve negatif itme kuvvetleri ile yeniden düzenlenmiş versiyonu [28] denklem (3)’te verilmiştir.

$$F_{ij}^d = -\eta(t) \left[2 \left(h_{ij}(t) \right)^{-13} - \left(h_{ij}(t) \right)^{-7} \right] \frac{\vec{r}_{ij}}{r_{ij}} \quad (3)$$

$\eta(t)$, itme ve çekim bölgelerini ayarlamak için kullanılan bir derinlik fonksiyonudur. $h_{ij}(t) = r_{ij}/\sigma(t)$ iki atom arasındaki uzaklığın uzunluk ölçeğine oranını ifade etmektedir.

Konum farkı vektörü denklem (4) ile ifade edilmiştir.

$$r_{ij} = \vec{x}_j - \vec{x}_i \quad (4)$$

$\vec{x}_j = (x_{j1}, x_{j2}, x_{j3})$ j . atomun konum vektörüdür. $\vec{x}_i = (x_{i1}, x_{i2}, x_{i3})$ ise i . atomun konum vektörünü ifade eder. Burada r_{ij} , denklem (5)'te verildiği gibi i . ve j . atomlar arasındaki Öklid uzaklığıdır.

$$r_{ij} = \|\vec{x}_j - \vec{x}_i\| = \sqrt{(x_{j1} - x_{i1})^2 + (x_{j2} - x_{i2})^2 + (x_{j3} - x_{i3})^2} \quad (5)$$

Derinlik fonksiyonu ise denklem (6)'da verilmiştir.

$$\eta(t) = \alpha \left(1 - \frac{t-1}{T}\right)^3 e^{-\frac{20t}{T}} \quad (6)$$

α , derinlik ağırlığıdır ve bu çalışmada 50 olarak belirlenmiştir.

İki atom arasındaki ölçeklenmiş mesafe denklem (7) ile hesaplanmaktadır.

$$h_{ij}(t) = \begin{cases} h_{min} \frac{r_{ij}(t)}{\sigma(t)} < h_{min} \\ \frac{r_{ij}(t)}{\sigma(t)} \quad h_{min} \leq \frac{r_{ij}(t)}{\sigma(t)} \leq h_{max} \\ h_{max} \frac{r_{ij}(t)}{\sigma(t)} > h_{max} \end{cases} \quad (7)$$

Ölçeklenmiş uzaklığın (h) alt ve üst sınıflarını ifade eden h_{min} ve h_{max} denklem (8) ile hesaplanmaktadır.

$$\begin{cases} h_{min} = g_0 + g(t) \\ h_{max} = u \end{cases} \quad (8)$$

g_0 , 1.1 olarak ayarlanmış en düşük limit ve u ise 1.24 olarak ayarlanmış en üst limittir. $g(t)$ denklem (9) ile hesaplanan ve algoritmanın keşiften sömürüye geçişini sağlayan sürüklenme faktörüdür.

$$g(t) = 0.1 \times \sin\left(\frac{\pi}{2} \times \frac{t}{T}\right) \quad (9)$$

Denklem (7)'deki uzunluk ölçeği $\sigma(t)$ çarpışma çapını gösterir ve denklem (10) ile hesaplanmaktadır.

$$\sigma(t) = \left\| x_{ij}(t), \frac{\sum_{j \in K_{best}} x_{ij}(t)}{K(t)} \right\|_2 \quad (10)$$

AAO algoritmasında her atomun, en iyi atomla kovalent bağa sahip olduğu kabul edilir. Böylece, her bir atomun en iyi atom ile arasındaki ağırlıklı konumun farkı olan geometrik kısıtlama kuvveti denklem (11) ile hesaplanmaktadır.

$$G_i^d(t) = \lambda(t)(x_{best}^d(t) - x_i^d(t)) \quad (11)$$

x_{best}^d d . boyutta en iyi atomun pozisyonu gösterir. $\lambda(t)$ ise Lagrange çarpanını ifade eder ve denklem (12) ile hesaplanmaktadır:

$$\lambda(t) = \beta e^{-\frac{20t}{T}} \quad (12)$$

β , bu çalışmada 0.2 olarak ayarlanmış çarpan ağırlığıdır. Sırasıyla etkileşim ve kısıtlama kuvvetinin tanımlanmasının ardından denklem (13)'te d . boyutundaki i . atomunun t . iterasyondaki ivmesi hesaplanmaktadır;

$$a_i^d(t) = \frac{F_i^d(t)}{m_i^d(t)} + \frac{G_i^d(t)}{m_i^d(t)} = -\alpha \left(1 - \frac{t-1}{T}\right)^3 e^{-\frac{20t}{T}} \quad (13)$$

$$\times \sum_{j \in K_{best}} \frac{rand_j \left[2 \left(h_{ij}(t)\right)^{-13} - \left(h_{ij}(t)\right)^{-7} \right]}{m_i(t)}$$

$$\cdot \frac{\left(x_j^d(t) - x_i^d(t)\right)}{\|\vec{x}_i(t), \vec{x}_j(t)\|_2} + \beta e^{-\frac{20t}{T}} \frac{x_{best}^d(t) - x_i^d(t)}{m_i(t)}$$

buradaki $m_i(t)$ t . iterasyondaki i . atomun kütleini belirtir ve denklem (14) ve denklem (15)'te verildiği üzere atomun uygunluk fonksiyonu hesaplanmaktadır;

$$M_i(t) = e^{\frac{Fit_i(t) - Fit_{best}(t)}{Fit_{worst}(t) - Fit_{best}(t)}} \quad (14)$$

$$m_i(t) = \frac{M_i(t)}{\sum_{j=1}^N J_j(t)} \quad (15)$$

Denklemden $Fit_i(t)$ i . atomun t . iterasyondaki uygunluk fonksiyonu değerini gösterir. $Fit_{best}(t)$ ve $Fit_{worst}(t)$ sırasıyla t . iterasyondaki en iyi atom ile en kötü atom arasındaki uygunluk fonksiyonu değerlerini ifade eder ve denklem (16) ile hesaplanmaktadır.

$$Fit_{best}(t) = \min_{i \in \{1, 2, \dots, N\}} Fit_i(t); \quad (16)$$

$$Fit_{worst}(t) = \max_{i \in \{1, 2, \dots, N\}} Fit_i(t);$$

i . atomun $(t+1)$. iterasyondaki konum ve hız güncellemeleri denklem (17) ve denklem (18) ile hesaplanmaktadır.

$$v_i^d(t+1) = rand_i^d \cdot v_i^d(t) + a_i^d(t) \quad (17)$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \quad (18)$$

AAO algoritmasının sözde kodu Algoritma-1 de verilmiştir.

Algoritma-1: AAO algoritmasının sözde kodu

- 1: Rastgele bir dizi X atomu (çözümleri) hızları v , $Fit_{Best} = Inf$ şeklinde başlatılır.
 - 2: **while** durdurma kriteri sağlanmadığında **do**
 - 3: **for** her atom X_i **do**
 - 4: Fit_i uygunluk değeri hesaplanır;
 - 5: **if** $Fit_i < Fit_{Best}$ ise o zaman
 - 6: $Fit_{Best} = Fit_i$;
 - 7: $X_{Best} = X_i$;
 - 8: **end if**
-

-
- 9: Denklem (14) ve denklem (15) kullanılarak kütle hesaplanır;
 10: Denklem (2) ile K komşusu belirlenir;
 11: Denklem (1) ve Denklem (11) ile etkileşim kuvveti F_i ve kısıtlama kuvveti G_i hesaplanır;
 12: Denklem (13) ile ivme hesaplanır;
 13: Denklem (17) kullanılarak hız güncellenir;
 14: Denklem (18) kullanılarak pozisyon güncellenir;
 15: **end for**
 16: **end while**
 17: Şimdiye kadarki en iyi çözüm olan X_{Best} bulunur.
-

3.2. Sınıflandırma Algoritmaları

Önerilen yöntemde, DR teşhisi yapmak için AAO algoritması özellik seçimi yapılan ve özellik seçimi yapılmayan orijinal veri kümesine Bagging, CvR, IbK, JRip, Kstar, and SimpleCart olmak üzere altı adet sınıflandırma algoritması uygulanmıştır. Bu algoritmalar ile ilgili bilgiler aşağıda verilmiştir:

Bagging: Bagging algoritması 1994 yılında L. Breiman tarafından önerilmiştir. Mevcut bir eğitim setinden yeni eğitim setleri türeterek temel öğrenciyi yeniden eğitme yöntemidir. Eğitim sırasında ikame örnekleme yapılmaktadır. Bu yöntemde, n adet numuneden oluşan eğitim seti yerine n adet numune bulunan bir numune seti ile eğitim seti rastgele üretilmektedir. Seçilen her örnek eğitim kümesine döndürülür. Bazı örnekler yeni eğitim setinde yer almazken bazıları birden fazla yer alabilir. Rastgele seçilen eğitim kümeleri ile başarılı temel öğrenciler, uygunsuzluğu sağlamak için eğitilmektedir. Böylece kolektif bir başarı sağlanmaktadır. Temel öğrenenler karar ağaçları olmak zorunda değildir. Bu açıdan herhangi bir makine öğrenimi algoritması temel öğrenci olabilir. Ancak, Bagging için minimum eğitim setinin değiştirilmesinde bile sonucu maksimuma etkileyecek temel öğrencilerin seçilmesi başarıyı artıracaktır [29].

CvR: Regresyon yoluyla sınıflandırma, problemleri regresyon fonksiyonlarına dönüştürebilen bir sınıflandırma yöntemidir [30]. Bu yöntem, oluşturulan birkaç alt ağaç üzerinde karar ağacı algoritması ve doğrusal regresyon ilkelerini birleştirmektedir. Bu yöntemin iki ana adımı vardır: Öncelikle, hedef/çıktı değerlerine göre kriterler/parametreler/nitelikler ve bunların varyasyonlarının ayrılmasını maksimize ederek sıradan bir karar ağacı yapılmaktadır. Karar ağacı yapılırken sapma azaltma standardı hesaplanmaktadır. Sonra, karar ağacının birkaç olası alt ağaç üzerinde budanmakta ve genellikle yapraklar üzerinde buna göre regresyon fonksiyonu ile doldurulmaktadır.

IbK: IbK algoritması, en yakın komşu algoritmasından (kNN) esinlenmiştir. IbK algoritması, sınıflandırma sırasında bir model oluşturmaz, ancak tam zamanında bir test örneği için bir tahmin üretmektedir. IbK algoritması, her test örneği için eğitim verilerinde k yakın örneği bulmak için bir mesafe ölçüsü kullanır ve bir tahmin yapmak için seçilen örnekleri kullanmaktadır [31].

JRip: JRip bir önerme kuralı öğrencisidir. Bu algoritmada, veri kümesinin örnekleri artan sırada işlenir ve veri kümesi için bir dizi kural oluşturulur. Daha sonra bir sonraki sınıf işlenir ve bu işlem tüm sınıflar kaplanana kadar devam etmektedir [32].

Kstar: Kstar, örnek tabanlı bir sınıflandırıcıdır, yani bir test örneğinin sınıfı, bazı benzerlik işlevleriyle belirlendiği gibi, kendisine benzer eğitim örneklerinin sınıfını temel almaktadır. Diğer örnek tabanlı öğrencilerden, entropi tabanlı bir uzaklık işlevi kullanması bakımından farklıdır [32].

SimpleCart: Yöntem, Leo Breiman [33] tarafından veri keşfi ve tahmini için geleneksel yöntemlere bir alternatif olarak önerilmiştir. SimpleCart, sınıflandırma için ikili karar ağacını oluşturur. Budama aşamasında, en iyi ağacı seçmek için çapraz doğrulama veya büyük test verileri kullanılır. Bu

yöntemde veriler farklı sonuçlara sahip iki alt gruba ayrılmaktadır. Alt grup boyutu minimum olduğunda bu işlem sonlandırılmaktadır.

3.3. Modelin Değerlendirilmesi için Metrikler

Önerilen yöntemi değerlendirmek için aşağıdaki metrikler kullanılmıştır.

Doğruluk: Test aşamasında etiketin gerçek değeri ile karşılaştırıldığında bir sınıflandırıcının yaptığı doğru tahminlerin yüzdesini ifade etmektedir. Diğer bir ifadeyle, doğru değerlendirme sayısının tüm değerlendirme sayısına oranı olarak da söylenebilir. Doğruluk denklem (19) kullanılarak hesaplanabilir.

$$\text{Doğruluk} = \frac{(DN + DP)}{(DN + DP + YN + YP)} \quad (19)$$

Denklemden, DP doğru pozitifleri, DN doğru negatifleri, YP yanlış pozitifleri ve YN yanlış negatifleri göstermektedir. Bir veri kümesindeki bir örneğin sınıf etiketi pozitifse ve sınıflandırıcı da sınıf etiketini pozitif olarak tahmin ederse, buna doğru pozitif denir. Bir veri kümesindeki bir örneğin sınıf etiketi negatifse ve sınıflandırıcı sınıf etiketini negatif olarak tahmin ederse, buna doğru negatif denir. Bir veri kümesindeki bir örneğin sınıf etiketi pozitifse, ancak sınıflandırıcı o örneğin sınıf etiketini negatif olarak tahmin ederse, buna yanlış negatif denir. Bir veri kümesindeki bir örneğin sınıf etiketi negatifse, ancak sınıflandırıcı o örneğin sınıf etiketini pozitif olarak tahmin ediyorsa, buna yanlış pozitif denir.

Duyarlılık: Test sırasında sınıflandırıcı tarafından doğru olarak tanımlanan doğru pozitiflerin yüzdesidir. Denklem (20) kullanılarak hesaplanmaktadır.

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (20)$$

Özgüllük: Test sırasında sınıflandırıcı tarafından doğru olarak tanımlanan doğru negatiflerin yüzdesidir. Denklem (21) kullanılarak hesaplanmaktadır.

$$\text{Özgüllük} = \frac{DN}{DN + YP} \quad (21)$$

Kesinlik: Önerilen yöntemin hatasızlığını belirlemek için önemli bir ölçüdür ve denklem (22)'de gösterildiği gibi hesaplanır. Toplam tahmine dayalı pozitif örneklere göre sınıflandırıcının pozitif olarak etiketlediği örneklerin yüzdesini ifade etmektedir.

$$\text{Kesinlik} = \frac{DP}{DP + YP} \quad (22)$$

Hassasiyet: Sınıflandırıcı tarafından pozitif olarak tanımlanan pozitif örneklerin yüzdesini belirlemektedir. Hassasiyet, denklem (23) ile hesaplanmaktadır.

$$\text{Hassasiyet} = \frac{DP}{DP + YN} \quad (23)$$

F-ölçütü: F-ölçütü, denklem (24)'te verildiği gibi kesinlik ve hassasiyet metriklerinin harmonik ortalamasını temsil etmektedir. Kesinlik ve hassasiyet arasında bir denge bulmak için F-ölçüsü gereklidir. Doğruluk esas olarak çok sayıda Gerçek Negatif tarafından sağlanırken, Yanlış Negatif ve Yanlış Pozitif genellikle işletme maliyetlerinin artmasına neden olur. Bu bakımdan, eşit olmayan bir

sınıf dağılımı kullanıldığında kesinlik ve hassasiyet arasında bir denge gerektiğinde F-ölçütü ile daha iyi bir ölçüm sağlanmış olabilir.

$$F - \text{ölçütü} = \frac{2 * \text{Kesinlik} * \text{Hassasiyet}}{\text{Kesinlik} + \text{Hassasiyet}} \quad (24)$$

ROC eğrisi: Eksiksiz bir hassasiyet / özgüllük raporu oluşturmayı amaçlayan bir metriktir. ROC eğrisinde, hassasiyet, bir parametrenin farklı kesme noktaları için özgüllük işlevinde çizilmektedir. ROC eğrisinde her nokta belirli bir karar eşğine karşılık gelen bir duyarlılık / özgüllük çiftini temsil etmektedir. ROC eğrisi altında kalan alan (AUC), bir parametre için iki grubun ne kadar iyi ayırt edilebildiğinin gösterildiği bir ölçüttür.

4. Deneysel Sonuçlar ve Tartışma

Bu çalışmadaki deneyler UCI makine öğrenimi deposundaki Diyabetik Retinopati veri kümesinde gerçekleştirilmiştir. Veri kümesinde 1151 örnek ve 20 özellik bulunmaktadır [34]. Özelliklerin açıklamaları Çizelge 2’de listelenmiştir.

Çizelge 2. Veri Kümesi açıklaması.

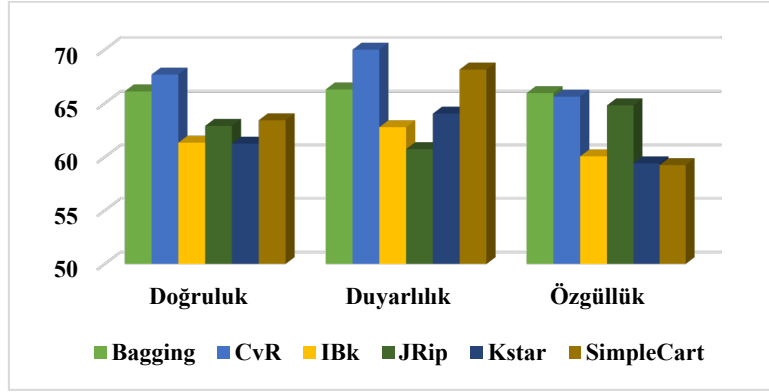
Özellik numarası	Özelliğın açıklaması
1	0 - Kötü kaliteyi temsil eder
2	1 - Ciddi seviyede retina anormalliğini, 0 - retinal anormallik eksikliğini temsil eder.
3-8	Mikroanevrizma (MA) tespiti retinaya kan sızıntısına neden olmaktadır.
9-16	Eksüdalar için 3-8 ile aynı. Fakat farklı görüntü boyutları için normalize edilir.
17	Bu sütun, optik diskin merkezi ile merkez arasındaki Öklid mesafesini temsil eder.
18	Optik Disk Çapı
19	Genlik - frekans modülasyonu (GM / FM) tabanlı sınıflandırmanın ikili değeri.
20	Sınıf etiketi 1 - DR’yi ve 0 - DR olmadığını işaret eder.

Bu bölümde özellik seçme aşaması olmadan sınıflandırma değerlendirmesi, özellik seçimi ile sınıflandırma değerlendirmesi ve karşılaştırmalı analiz aşaması olmak üzere üç aşamada elde edilen deneysel sonuçlar kısaca açıklanmaktadır.

4.1. Özellik Seçimi Aşaması Olmadan Sınıflandırma Değerlendirmesi

Bu aşamada, Bagging, CvR, IBk, JRip, Kstar, SimpleCart sınıflandırıcıları ile altı farklı sınıflandırma modeli oluşturulmuştur. Modeller için test seçeneği olarak on kat çapraz doğrulama kullanılmıştır. Modeller, herhangi bir özellik seçimi süreci olmaksızın DR veri kümesinde yer alan tüm özellikler kullanılarak oluşturulmuştur.

Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmaları için oluşturulan sınıflandırma modellerinin doğruluk, duyarlılık ve özgüllük oranları Şekil 3’te gösterilmiştir. Şekil 3 incelendiğinde, doğruluk ve duyarlılık bakımından CvR algoritmasının diğer algoritmalarından daha iyi sonuç elde ettiği görülürken, özgüllük bakımından Bagging algoritmasının en iyi değeri elde ettiği görülmüştür.



Şekil 3. Özellik seçimi olmadan sınıflandırma modellerinin performansları

Tüm sınıflandırma algoritmaları için elde edilen detaylı istatistiksel sonuçlar Çizelge 3'te verilmiştir. Çizelge 3'teki deney sonuçları incelendiğinde CvR algoritmasını kullanan modelin DR teşhisi için elde ettiği doğruluk yüzdesinin Bagging, IBk, Jrip, Kstar ve SimpleCart algoritmalarından daha yüksek olduğu gözlenmiştir. Bagging algoritması 0.661, CvR 0.677, IBk 0.613, JRip 0.629, Kstar 0.612 ve SimpleCart algoritması 0.634 doğruluk oranı elde etmiştir.

Ayrıca bu modellerin performansını değerlendirmek için kesinlik, f-ölçütü ve ROC alanı olmak üzere üç performans değerlendirme ölçütleri için de elde edilen değerler Çizelge 4'te verilmiştir. Çizelge 4'e göre, kesinlik ve f-ölçütü bakımından CvR algoritmasının sırasıyla 0.712 ve 0.683 değerleri ile diğer algoritmalarından daha yüksek değer elde ettiği gözlemlenmiştir. ROC alanına göre 0.726 değeri ile Bagging ve CvR algoritmaları ile en yüksek değer elde edilmiştir.

Çizelge 3. Özellik seçimi olmadan sınıflandırma modellerinin detaylı istatistikleri.

	Bagging	CvR	IBk	JRip	Kstar	SimpleCart
Doğru Sınıflandırılmış Örnekler	0.661	0.677	0.613	0.629	0.612	0.634
Yanlış Sınıflandırılmış Örnekler	0.339	0.323	0.387	0.371	0.388	0.366
Kappa İstatistiği	0.321	0.354	0.227	0.255	0.226	0.271
Karekök Ortalama Hata	0.4581	0.459	0.621	0.485	0.563	0.481
Bağıl Mutlak Hata	0.811	0.828	0.776	0.907	0.791	0.859

Çizelge 4. Özellik seçimi olmadan sınıflandırma modellerinin performans ölçütleri.

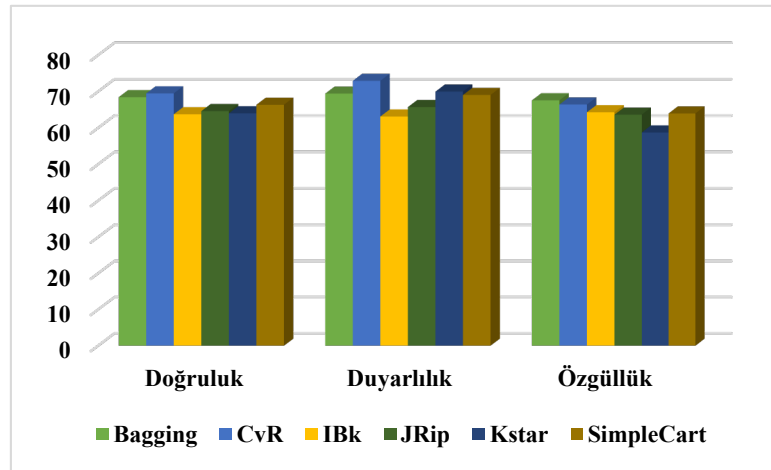
	Kesinlik	F-ölçütü	ROC Alanı	Sınıf Etiketi
Bagging	0.689	0.674	0.726	DR
	0.633	0.647	0.726	DR Değil
CvR	0.712	0.683	0.726	DR
	0.643	0.670	0.726	DR Değil
IBk	0.646	0.623	0.608	DR
	0.581	0.604	0.608	DR Değil
JRip	0.651	0.650	0.637	DR
	0.604	0.606	0.637	DR Değil
Kstar	0.647	0.619	0.665	DR
	0.580	0.605	0.665	DR Değil
SimpleCart	0.678	0.636	0.668	DR
	0.596	0.632	0.668	DR Değil

4.2. Özellik Seçimi Aşaması ile Sınıflandırma Değerlendirmesi

Çizelge 5. AAO ile özellik seçimi için başlangıç parametreleri ve değerleri.

Parametre	Değer
α	50
β	0.2
Popülasyon sayısı	10
Maksimum iterasyon sayısı	500

Çalışmanın bu aşaması iki adımdan oluşmaktadır. İlk adımda, DR veri kümesinden AAO ile özellik seçimi yapılmaktadır. AAO ile özellik seçimi yapmak için algoritmanın başlangıç parametreleri ve değerleri Çizelge 5’te verildiği gibi ayarlanmıştır. Yöntemin ikinci adımında ise, birinci adım sonunda seçilen en iyi özelliklere göre sınıflandırma algoritmalarının performansları değerlendirilmiştir. AAO ile özellik seçimi yapıldıktan sonra 20 özelliği olan DR veri kümesinden Çizelge 2’de açıklamaları verilen 3, 6, 8, 11 ve 14 numaralı beş özellik seçilmiştir.



Şekil 4. Özellik seçimi ile sınıflandırma modellerinin performansları

AAO algoritması ile DR veri kümesinden özellik seçimi yapıldıktan sonra Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmaları ile elde edilen doğruluk, duyarlılık ve özgüllük değerlerinin grafiksel gösterimi Şekil 4’te verilmiştir. Şekil 4 incelendiğinde, doğruluk ve duyarlılık bakımından CvR algoritmasının diğer algoritmalarından daha iyi sonuç elde ettiği görülürken, özgüllük bakımından Bagging algoritmasının en iyi değeri elde ettiği görülmüştür.

Çizelge 6. Özellik seçimi ile sınıflandırma modellerinin detaylı istatistikleri.

	Bagging	CvR	IBk	JRip	Kstar	SimpleCart
Doğru Sınıflandırılmış Örnekler	0.684	0.695	0.637	0.646	0.640	0.663
Yanlış Sınıflandırılmış Örnekler	0.316	0.305	0.363	0.354	0.360	0.337
Kappa İstatistiği	0.369	0.391	0.274	0.292	0.284	0.328
Karekök Ortalama Hata	0.448	0.442	0.601	0.476	0.472	0.474
Bağlı Mutlak Hata	0.785	0.887	0.728	0.888	0.834	0.823

Çizelge 7. Özellik seçimi ile sınıflandırma modellerinin performans ölçütleri.

	Kesinlik	F-ölçütü	ROC Alanı	Sınıf Etiketi
Bagging	0.715	0.695	0.753	DR
	0.654	0.674	0.753	DR Değil
CvR	0.736	0.698	0.766	DR
	0.658	0.692	0.766	DR Değil
IBk	0.664	0.653	0.642	DR
	0.610	0.621	0.642	DR Değil
JRip	0.678	0.636	0.670	DR
	0.615	0.657	0.670	DR Değil
Kstar	0.689	0.634	0.696	DR
	0.600	0.646	0.696	DR Değil
SimpleCart	0.701	0.669	0.691	DR
	0.629	0.658	0.691	DR Değil

Tüm sınıflandırma algoritmaları için elde edilen detaylı istatistiksel sonuçlar Çizelge 6'da verilmiştir. Çizelge 6'daki deney sonuçları incelendiğinde CvR algoritmasını kullanan modelin DR tespiti için elde ettiği doğruluk yüzdesinin Bagging, IBk, Jrip, Kstar ve SimpleCart algoritmalarından daha yüksek olduğu gözlenmiştir. Bagging algoritması 0.661, CvR 0.677, IBk 0.613, JRip 0.629, Kstar 0.612 ve SimpleCart algoritması 0.634 doğruluk oranı elde etmiştir.

Ayrıca bu modellerin performansını değerlendirmek için kesinlik, f-ölçütü ve ROC alanı olmak üzere üç performans değerlendirme ölçütleri için de elde edilen değerler Çizelge 7'de verilmiştir. Çizelge 7'ye göre, kesinlik, f-ölçütü ve ROC alanı bakımından CvR algoritmasının sırasıyla 0.736, 0.698 ve 0.766 değerleri ile diğer algoritmalarından daha yüksek değer elde etmiştir.

4.3. Performans Analizi

Çalışmanın bu aşaması, önceki aşamalarda gerçekleştirilen iki deneye dayalı olarak performans değerlendirme kriterlerinin karşılaştırılmasını içermektedir. Çizelge 8'de AAO algoritması ile özellik seçimi yapılan ve özellik seçimi yapılmadan Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmaları ile doğruluk, duyarlılık ve özgüllük metrikleri için elde edilen sonuçları göstermektedir.

Çizelge 8. Sınıflandırma yöntemlerinin performans karşılaştırılması

Sınıflandırıcı		Değerlendirme Metrikleri		
		Doğruluk %	Duyarlılık %	Özgüllük %
Bagging	Özellik seçimi ile	68.462	69.444	67.594
	Özellik seçimi olmadan	66.116	66.296	65.957
CvR	Özellik seçimi ile	69.504	72.963	66.448
	Özellik seçimi olmadan	67.680	70	65.630
IBk	Özellik seçimi ile	63.771	63.148	64.321
	Özellik seçimi olmadan	61.338	62.778	60.065
JRip	Özellik seçimi ile	64.639	65.741	63.666
	Özellik seçimi olmadan	62.902	60.741	64.812
Kstar	Özellik seçimi ile	64.031	70	58.756
	Özellik seçimi olmadan	61.251	64.045	59.411
SimpleCart	Özellik seçimi ile	66.377	69.074	63.993
	Özellik seçimi olmadan	63.423	68.148	59.247

Çizelge 8'deki sonuçlar, AAO algoritması ile özellik seçimi yapıldığında, tüm sınıflandırma yöntemlerinin doğruluk, duyarlılık ve özgüllük metriklerine bakımından özellik seçimi yapılmayan duruma göre daha iyi performans elde ettiğini göstermektedir. AAO algoritması ile özellik seçimi yapılan durumda, doğruluk bakımından en iyi sonuç 69.504 değeri ile CvR algoritması ile elde edilmiştir. CvR algoritmasından sonra ikinci en iyi doğruluğu 68.462 değeri ile Bagging algoritması vermiştir. AAO algoritması ile özellik seçimi yapıldıktan sonra duyarlılık bakımından en iyi sınıflandırma algoritması 72.963 değeri ile CvR iken özgüllük bakımından en iyi algoritma 67.594 değeri ile Bagging algoritmasıdır. Bu bakımdan önerilen özellik seçimi ile algoritmaların özellik seçimi olmadan elde edilen sınıflandırma oranlarında doğruluk için ortalama %2.7 duyarlılık için %3.5 özgüllük için %2'lik bir artış sağlanmıştır.

AAO algoritması ile özellik seçimi olmadan, doğruluk ve duyarlılık metriğine göre en iyi sonucu veren algoritmalar sırasıyla 67.680 ve 70 değerleri ile CvR'dir. Özgüllük bakımından en iyi algoritma ise 65.957 değeri ile Bagging'dir. Deney sonuçları, uygun bir özellik seçim algoritması belirlenmenin sınıflandırma sonuçlarında iyileşme olabileceğini göstermiştir.

5. Sonuçlar ve Gelecek Çalışmalar

Bu çalışmada, diyabetik retinopati tespiti problemi için atom dinamiklerinden faydalanarak önerilmiş popülasyon temelli yeni bir metasezgisel yöntem olan Atom Arama Optimizasyon algoritması ilk kez özellik seçim yöntemi olarak kullanılmıştır. Bagging, CvR, IBk, JRip, Kstar ve SimpleCart sınıflandırma algoritmaları, AAO algoritması ile özellik seçimi yapılmış veri kümesi ve orjinal veri kümesine uygulanarak elde edilen deney sonuçları doğruluk, duyarlılık, özgüllük, kesinlik, f-ölçütü ve roc alanı olmak üzere farklı performans metriklerine göre karşılaştırılmıştır. Yapılan karşılaştırmalar AAO algoritması ile özellik seçimi yapılması durumunda Bagging, CvR, IBk, JRip, Kstar ve SimpleCart algoritmalarının başarı oranının özellik seçimi olmayan durumdan daha iyi olduğunu göstermiştir.

Çıkar Çatışması Beyanı

Makale yazarları herhangi bir kurum, kuruluş, kişi ile kişisel ve finansal çıkar çatışması olmadığını beyan etmektedirler.

Kaynaklar

- [1] Taylor R, Batey D. Handbook of retinal screening in diabetes: diagnosis and management. John Wiley & Sons.
- [2] Lin X, Xu Y, Pan X, Xu J, Ding Y, Sun X, Shan PF. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Scientific reports 2020; 10(1): 1-11.
- [3] Murray CJ, Lopez A. Mortality by cause for eight regions of the world: Global Burden of Disease Study. The Lancet 1997; 349(9061): 1269-1276.
- [4] Bourne RR, Stevens GA, White RA, Smith JL, Flaxman SR, Price H. Causes of vision loss worldwide, 1990–2010: a systematic analysis. The lancet global health 2013;1(6): 339-e349.
- [5] Harper CA, Keeffe JE. Diabetic retinopathy management guidelines. Expert review of ophthalmology 2012; 7(5): 417–39.
- [6] Khurana DAK. Comprehensive ophthalmology: With supplementary book. Review of Ophthalmology. JP Medical Ltd, editor 2015.
- [7] Patel S, Lohakare M, Prajapati S, Singh, S, Patel N. DiaRet: A browser-based application for the grading of Diabetic Retinopathy with Integrated Gradients. arXiv preprint arXiv:2103.08501, 2021.

- [8] Qummar S, Khan FG, Shah S, Khan A, Shamshirband S, Rehman ZU, Jadoon W. A deep learning ensemble approach for diabetic retinopathy detection. *IEEE Access* 2019; 7: 150530-150539.
- [9] Haneda S, Yamashita H. International clinical diabetic retinopathy disease severity scale. *Nihon rinsho. Japanese journal of clinical medicine* 2010; 68: 228-235.
- [10] Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, Kaski K. Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Scientific reports* 2019; 9: 1–11.
- [11] Gondal WM, Köhler JM, Grzeszick R Fink GA, Hirsch M, Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images. Paper presented at: IEEE international conference on image processing (ICIP) 2017: 2069-2073.
- [12] Li X, Pang T, Xiong B, Liu W, Liang P, Wang T, Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. Paper presented at: 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE 2017: 1–11.
- [13] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama* 2016; 316: 2402–2410.
- [14] Castellano G, Castiello C, Mencar, C.; Vessio, G. Crowd Detection for Drone Safe Landing Through Fully-Convolutional Neural Networks. *International Conference on Current Trends in Theory and Practice of Informatics*. Springer, 2020, pp. 301–312.
- [15] Raman V, Then P, Sumari P. Proposed retinal abnormality detection and classification approach: Computer-aided detection for diabetic retinopathy by machine learning approaches. Paper presented at: 8th IEEE International Conference on Communication Software and Networks (ICCSN) 2016: 636-641.
- [16] Porwal P, Pachade S, Kokare M, Giancardo L, Meriaudeau F. Retinal image analysis for disease screening through local tetra patterns. *Computers in biology and medicine* 2018; 102: 200-210.
- [17] Hemanth DJ, Deperlioglu O, Kose U. An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. *Neural Computing and Applications* 2020; 32(3): 707-721.
- [18] Shahin EM, Taha TE, Al-Nuaimy W, El Rabaie S, Zahran OF, El-Samie FEA. Automated detection of diabetic retinopathy in blurred digital fundus images. Paper presented at: 8th International Computer Engineering Conference (ICENCO) 2013; 20–25.
- [19] Jaafar HF, Nandi AK, Al-Nuaimy W. Automated detection and grading of hard exudates from retinal fundus images. Paper presented at: 19th European signal processing conference 2011: 66–70.
- [20] Casanova R, Saldana S, Chew EY, Danis RP, Greven CM, Ambrosius WT. Application of random forests methods to diabetic retinopathy classification analyses. *PLoS One* 2014; 9(6).
- [21] Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 2017; 124(7): 962–969.
- [22] Akram MU, Shahzad K, Shoaib AK. Identification and classification of micro aneurysms for early detection of diabetic retinopathy. *Pattern Recogn* 2013; 46:107–116.
- [23] Welikala RA, Dehmeshki J, Hoppe A, Tah V, Mann S, Williamson TH. Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification. *Computer methods and programs in biomedicine* 2014; 114(3): 247-261
- [24] Waheed A, Waheed Z, Usman Akram M, Shaukat A. Removal of false blood vessels using shape based features and image inpainting. *Journal of Sensors* 2015.
- [25] Zhang X, Thibault G, Decencièrre E, Marcotegui B, Lay B, Danno R. Exudate detection in color retinal images for mass screening of diabetic retinopathy. *Medical image analysis* 2014; 18(7): 1026-1043.

- [26] Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Srivastava G. (2020). Deep neural networks to predict diabetic retinopathy. *Journal of Ambient Intelligence and Humanized Computing* 2020; 1-14.
- [27] Zhao W, Wang L, Zhang, Z. Atom search optimization and its application to solve a hydrogeologic parameter estimation problem. *Knowledge-Based Systems* 2019; 163: 283-304.
- [28] Zhao W, Wang L, Zhang Z. A novel atom search optimization for dispersion coefficient estimation in groundwater, *Future Generation Computer Systems* 2019; 91: 601-610.
- [29] Breiman L. Bagging predictors, *Machine Learning* 1996; 24: 123-140.
- [30] Ruan Y, Lin H, Tsai M. Improving Ranking Performance with Cost-Sensitive Ordinal Classification Via Regression. *Information retrieval* 2014; 17(1): 1-20.
- [31] Kaladhar DSVGK, Pottumuthu BK, Rao PVN, Vadlamudi V, Chaitanya AK, Reddy RH. The elements of statistical learning in colon cancer data sets: data mining, inference and prediction. *Algorithms Research* 2013; 2: 8-17.
- [32] Cohen WW. Fast effective rule induction. Paper presented in: Twelfth International Conference on Machine Learning 1995: 115-123.
- [33] Breiman L, Friedman JH, Olshen RA, Stone C, *Classification and regression trees*. Wadsworth Books, 1984.
- [34] University of California Irvine Machine Learning Repository. Accessed: Kasım. 5, 2021. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set>