



Creditability and Classification Work for Finance Sector with Gray Wolf Optimization (GWO) Algorithm

Ali Tunç

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

December 19, 2019

Gri Kurt Optimizasyonu (GWO) Algoritması İle Finans Sektörü İçin Kredi Verilebilirlik ve Sınıflandırma Çalışması

Ali TUNÇ – Kuveyt Türk Katılım Bankası, Ar-Ge Merkezi Konya
Feritpaşa Mah. Kule Caddesi Kule Plaza Kat: 9 No: 1 - 42060 Selçuklu / KONYA
ali.tunc@kuveytturk.com.tr

Özet.

Bilgi teknolojilerindeki gelişmelerle birlikte, banklar, müşterilerinin kredi taleplerini etkili analitik yöntemler ve risk analizleri ile değerlendirebilmektedirler. Kredi skorlama sistemleri olarak adlandırılan yazılım ürünleri genel olarak daha önce belirlenen kredi faktörlerine göre müşterinin verilerinin toplanması, elde edilen verinin çeşitli istatistiksel veya makine öğrenmesi teknikleriyle işlenmesi ve kredi risk analizinin yapılarak nihai kredi kararının belirlenmesi aşamalarından oluşur. Bu modeller dâhilinde müşterilerin geçmiş banka hareketleri işlenerek kredi kararı verilebilmektedir. Yapılan çalışma ile müşteriye ait değişken kümelerinden oluşan veriler, makine öğrenmesi teknikleriyle işlenerek, müşteriye ait kredi değeri belirlenmeye çalışılmıştır. Gri Kurt Optimizasyonu ile sınıflandırma problemi olarak çalışma odağı oluşturulmuştur. Elde edilen bu bilgilere göre müşteriye kredi verilebilir ya da verilemez kararı ortaya çıkarılmıştır. Bu çalışmada kredi başvurusunda bulunan ve kredi kullanan tüketicilerdeki artışı sağlıklı yönetebilecek yapı ihtiyacının karşılanması, doğru müşteriye, doğru zamanda, doğru miktarda ve doğru vadede kredi vermeyi sağlayacak yapının kurulması, kredi tahsilatlarının verimliliğinin artırılması, riski minimize ederek karlılığın maksimum noktaya getirecek optimum stratejilerin oluşturulması, bankanın kredi skorlamasında ve değerlendirme sisteminde uzman görüş etkisini azaltılması ve maliyetlerin düşürülmesi amaçlanmıştır.

Anahtar Sözcükler: Kredi Risk Analizi, Kredi Skor Modellemesi, Makine Öğrenmesi, Sınıflandırma Algoritmaları, Gri Kurt Optimizasyonu

Creditability and Classification Work for Finance Sector with Gray Wolf Optimization (GWO) Algorithm

Abstract.

With the advances in the Information Technology (IT) field, banks can evaluate the credit requests of the customers via effective analytical methods and risk analysis. The software products, named Credit Scoring Systems, consist of collecting customer data based on pre-determined credit factors, processing the data

with various statistical or machine learning methods, and conducting credit risk analysis to make the final credit decision. These models utilize the previous transactions on the bank accounts of the customers to make a decision on the credit applications. In the proposed work, the information about the customer related to several aspects and processed with machine learning techniques, and finally a credit score will be determined for each customer. Classification problem using Grey Wolf Optimization method was focused in this work. This information will later be used to decide whether the credit application of a customer can be approved or not. In this study, intentions can be summarized as, providing useful tools to manage the increasing number of customers who apply for consume credits, establishing a structure for crediting the right customers at the right time with the right amount and payment plan, increasing the efficiency of collecting credit payments, thus contributing to the national economy by using the resources more effectively, creating optimal strategies for maximizing the profit by minimizing the risk, reducing the effect of an expert for credit scoring and evaluation, and reducing the costs.

Key words: Classification Algorithms, Credit Risk Analysis, Credit Scores Modeling, Machine Learning, Gray Wolf Optimization, GWO

Giriş

Kredi skorlama sistemleri; müşterilerin kredi marketine erişim imkânını arttırması, kredi ücretlendirmesinin düşürülmesi, karar süreçlerinin kısaltılması ve karar aşamalarına bağlı ortaya çıkan hata ile temerrütlerin azaltılmasında önemli bir rol oynayarak ülkenin ekonomik gelişimine katkı sağlamaktadır. Yapılan çalışma ile bankaların tüm kredi kararlarını otomatik bir skorlama sistemi üzerinden vermesi, oldukça maliyetli olan banka yetkilisinin kredi sonucunu belirlediği yöntemlere göre daha az maliyetli, hızlı ve kesinliği yüksek bir sistemin oluşturma bilirliliği incelenmiştir. Yapılan araştırmalar ve istatistiksel verilere göre banka kredilerine duyulan talepteki artışın ve buna bağlı olarak geliştirilen efektif kredi karar modellerinin oluşturulması ihtiyacının somut bir göstergesidir. Bu pazarda başarılı olabilmek için değişen ve gelişen ihtiyaçlara cevap veren efektif, kesinlik oranı yüksek karar mekanizmalarına sahip olmak gerekmektedir. Bununla birlikte, geliştirilen yeni karar modeli ile elde edilmek istenen en büyük katkı, müşterilere ait yeterli verinin bulunmadığı belirsiz koşullarda ortaya etkin bir sonucun çıkarılmaya çalışılmasıdır. Bu çalışma kapsamında sınıflandırma çalışmaları üzerine yeni geliştirilmiş GWO algoritmasının mevcut problemlerin çözümündeki başarı oranı incelenmeye çalışılmıştır. GWO algoritmasının seçilme nedeni sezgisel olarak sınıflandırma yapan bu algoritmanın finans verilerindeki başarı oranının değerlendirilmesidir. Diğer makine öğrenmesi teknikleri ile ilgili bulunan sonuçlar bu çalışma kapsamında değerlendirilmemiştir. Algoritmaların performans ve başarı oranlarının karşılaştırılması sonraki çalışmalar ile gösterilmeye çalışılacaktır. Bu alanlarda yapılan geliştirmelerle banka şirketlerinin istatistiksel skorlama modellerinden kaynaklanan kısıtlamalara maruz kalmadan kredi kararlarını verebilmeleri sağlanması hedeflenmiştir. Geliştirilen makine öğrenmesi tabanlı yeni teknolojinin kesinliği yüksek kredi skorlama çözümleri

üretmesinde ciddi faydalar sağlayacağı öngörülmektedir. Çalışmanın genel amacı; kredi kararı aşamalarında oluşabilecek hataları önlemeye yardımcı olan, farklı karar faktörlerinin değerlendirilmesinde standart bir çözüm sunan otomatik skortlama araçları ve modellerinin geliştirilmesidir.

1 Literatür Çalışması

Yapay öğrenme tekniklerinden olan Grey Wolf Optimization (GWO) algoritması hakkında yapılan araştırmalara göre GWO algoritmasının geçmişi 2010 yılına dayanmaktadır. GWO Algoritması için literatür taramasında ulaşılan kaynaklar şunlardır; Mirjalili ve arkadaşları GWO üzerine detaylı bir çalışma yapmışlar ve Parçacık Sürü Optimizasyonu (Particle Swarm Optimization - PSO), Yerçekimi Arama Algoritması (Gravity Search Algorithm - GSA), Diferansiyel Gelişim Algoritması (Differential Evolution Algorithm - DEA) gibi birçok algoritmaya yakın sonuç verdiğini ortaya koymuşlardır. Canis lupus ismi verilen gri kurtların liderlik hiyerarşisi ve av mekanizmasını taklit ederek meta sezgisel bir yaklaşım ortaya çıkarmışlardır. Av ararken, av çevrelerken ve ava saldırırken gösterdikleri yaklaşımı modellemeye çalışmışlardır. 29 tanınmış test fonksiyonu ve test verileri üzerinde benchmarked testlerine tabi tutmuşlar ve PSO, GSA, DEA gibi diğer algoritmalarla rekabetçi sonuçlar ortaya koyduğu bilgisine ulaşmışlardır[1]. Kamboh ve arkadaşları, GWO kullanarak ekonomik yük dağıtım sorununun çözümü ile çalışmışlar, GWO algoritmasını elektrik güç sistemi olmayan dışbükey ve dinamik bir ekonomik yük dağıtım problemi (ELDP) çözümü için kullanmışlardır. GWO algoritması diğer iyi algoritmaları ile karşılaştırıldığında rekabetçi sonuçlar sağlayabildiğini göstermişlerdir[3]. Yusof ve Mustaffa GWO algoritması kullanarak hammadde enerji zaman serisi tahmini üzerine bir çalışma yapmışlardır [4]. Muro, av stratejileri için hesaplama simülasyonlarına basit kurallar ortaya çıkarmıştır [5]. Jayapriya ve Arock, birden fazla moleküler dizileri hizalanması için paralel GWO tekniği kullanarak bir model geliştirmeye çalışmışlar, GWO algoritmasının diğer mevcut olanlara göre hesaplama süresini azaltır olduğunu göstermişlerdir[6]. Mirjalili, çok katmanlı algılayıcıların eğitiminde GWO algoritmasının ne kadar iyileştirici olduğunu göstermeye çalışmış ve Multi Layer Perceptron (MLP) ların eğitiminde GWO yu denemiş elde ettiği sonuçları PSO, ACO, GA, ES ve PBIL ile karşılaştırmıştır. GWO ile elde ettiği değerlerin çok rekabetçi olduğunu ve yakınlıktırma yüksek değerler elde ettiğini sunmuştur[2]. Korayem, GWO algoritması ile araçların seyahatteki toplam maliyet ve mesafeyi minimize etmek için GWO algoritması kullanılarak bir çalışma yapılmıştır. Elde ettikleri sonuçları ve sorunları diğer algoritmalarla kıyaslamışlar GWO'nun bu problemin çözümünde yakın sonuçlar ortaya koyduğunu göstermeye çalışmıştır[7]. Aghaee, çok katmanlı algılayıcı için GWO algoritması, bir eğitim algoritması olarak kullanılmış optimizasyon problemini çözmek için geri beslemeli sinir ağlarının eğitiminde GWO algoritmasından yararlanmışlardır[8]. Shakarami ve Davoudkhani, zaman gecikmesi dikkate alarak GWO algoritmasına dayalı geniş alan güç sistemi dengeleyicisi (WAPSS) tasarımı için bir yöntem önermişlerdir[9]. Sharma ve Saikia, klasik kontrol tabanlı termik güç santrallerinde GWO algoritması kullanarak Multi Area ST'ların otomatik üretim kontrollerini yapmaya çalışmışlardır[10].

2 Gri Kurt Optimizasyonu

GWO algoritması sezgisel optimizasyon yöntemidir. Sezgisel optimizasyon; bir problemin çözümünü, doğruluğunun kanıtlanabilir olup olmadığını önemsenmeden iyiye yakın olarak ele alma yöntemidir. GWO algoritması doğadaki gri kurtların liderlik ve avlanma mekanizmasını taklit eder. Gri kurt hiyerarşisinde Alfa, Beta, Delta ve Omega olmak üzere 4 tip uygulanır[1,2]. Gri kurt hiyerarşisinde lider Alfa kurttur ve sürünün avlanmaya karar vermesi, uyku yeri, kalkma vakti vs. gibi olaylarda sorumluluk sahibidir. Hiyerarşide ikinci sırada bulunan Beta kurt karar verme ve diğer sürü aktivitelerinde Alfa kurdun yardımcısı olarak bilinir. Beta kurt, Alfa kurdun emirlerini diğer kurtlara iletir ve geri dönüş sağlar. Beta kurt, Alfa kurdun yaşlanmasıyla ya da uzaklaşmasıyla onun yerine geçer. Hiyerarşide en düşük seviyeli Omega kurt her zaman yemekte diğer kurtların doymasını ve ona sıra gelmesini bekler. Omega kurdu, her zaman baskın kurt seçer[1,2]. Hiyerarşide üçüncü sırada bulunan Delta kurdu ise eğer kurt Avlanma hiyerarşisinde gri kurtlar öncelikle avlanma davranışına girerler, ardından sırasıyla takip, yaklaşma ve avını izlerler, son olarak takip, rahatsız etme ve çevreleme işlemini yaparlar. Bu izleme, çevreleme ve ava saldırmadır[1].

2.1. Sosyal Hiyerarşi: Matematiksel modellemeyi yaparak GWO tasarımını sağlar. Alfa en iyi çözüm olarak kullanılır. Ardından sırasıyla Beta ve Delta, ikinci ve üçüncü çözüm olarak adlandırılır. Çözüme aday olarak da kalan Omega olarak kabul edilir[1].

2.2. Avı Çevreleme: Yukarıda belirtildiği gibi, gri kurtlar av sırasında avı çevreler. Çevreleme davranışında sırasıyla matematiksel modellemede aşağıdaki denklem 2.1 ve denklem 2.2 gibidir [1,2]. Burada t geçerli yinelemeyi gösterir. A ve C vektör katsayısını, X_p avın konumunun vektörü, X bir gri kurdun pozisyon vektörünü gösterir. A ve C Vektörü aşağıdaki denklem 2.3 ve denklem 2.4 deki gibi hesaplanır.

$$\vec{D} = | \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) | \quad (2.1)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad (2.2)$$

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (2.3)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (2.4)$$

(x, y) pozisyonunda gri kurt av pozisyonuna göre konumunu günceller (X^* , Y^*). En iyi ajan A ve C vektör değerlerini güncel konumuna göre değiştirerek farklı yerlere ulaşabilir. Aynı kavram n boyutlara sahip bir arama alanında uzatılabilir. Ve gri kurtlar şimdiye kadar elde edilen en iyi çözüm etrafında hiper küp şeklinde hareket edecektir.

2.3. Avlanma: Gri kurt yerini tanıma ve orayı kuşatma yeteneğine sahiptir. Av genellikle Alfa tarafından yönlendirilir, Beta ve Delta da bazen ava katılabilir. Alfa (en iyi aday çözüm) Beta olduğunu varsayalım ve Delta avın potansiyel konumu hakkında daha iyi bilgi sahibidir. Bu nedenle, şimdiye kadar elde edilen en iyi ilk üç sonucu kaydedilir ve en iyi arama ajanlarının konumuna göre diğer arama ajanlarının (Omega'lar dâhil) konumu güncellenir. Bu güncellemelere ait denklemler aşağıda denklem 2.5, denklem 2.6 ve denklem 2.7 ile gösterilmektedir [1,2].

$$\vec{D}_\alpha = | \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} | \quad \vec{D}_\beta = | \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} | \quad \vec{D}_\delta = | \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} | \quad (2.5)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \quad (2.6)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (2.7)$$

Son konum arama alanındaki Alfa, Beta ve Delta pozisyonlara göre tanımlanan bir daire içinde rasgele bir yerde olacağı görülmektedir. Alfa, Beta, Delta avın konumunu tahmin eder ve diğer kurtlar av etrafında rastgele konumlarını günceller.

2.4. Ava Saldırma (Sömürü): Gri kurt av durduğunda ava saldırarak avı bitirir. Sırayla matematiksel modelle yaklaşan avın a vektörünün değerini azaltıyoruz. Böylelikle a vektörüne bağlı A vektörü de azalır. Başka bir deyişle A vektörü [-2a, 2a] değeri arasında rastgele bir sayıdır ve (a) yineleme boyunca 2'den 0 a düşürülmüştür. A vektörü [-1,1] içinde rastgele değer olduğunda, arama ajanının bir sonraki pozisyonu geçerli pozisyonu ve avın pozisyonu arasında herhangi bir pozisyonda olabilir.

2.5. Av Arama (Keşif): Gri kurtlar kurtların konumuna göre arama yaparlar. Gri kurtlar arama yapmak için dağılır ve avı bulduğunda saldırmak için toplanırlar [1]. Matematiksel olarak sapma modelinde, 1 den büyük rasgele değerler için A vektörü kullanılır. Tanımlanan C vektörü ise [0,2] rasgele değerlerini içerir. Her zaman keşif esnasında rastgele değerleri sağlayabilmek için C gereklidir. Genel olarak özetleyecek olursak, doğada engeller kurtların avlanma durumunda ortaya çıkar, kurdun hızlı ve rahat bir şekilde ava yaklaşmasını engeller, C vektörünün yaptığı da tam olarak budur. Alfa, Beta, Delta ve kurtların ve yırtıcının muhtemel konumunu tahmin edilir. Her aday çözüm yolu için av mesafesi güncellenir. Son olarak, GWO algoritması son bir kriter ile sonlandırılır. Bu algoritmanın sahte kodunu şu şekilde özetleyebiliriz. Gri Kurt Popülasyonunu başlat, Başlangıç değerlerini ata, Herbir temsilci için uygunluk fonksiyon metodu hesapla, X_α en iyi arama temsilcisini ata, X_β en iyi 2. arama temsilcisini ata, X_δ en iyi 3. arama temsilcisini ata, TEKRARLA (HER BİR TEMSİLCİ İÇİ TEKRARLA, Temsilcinin pozisyon bilgisini güncelle, Başlangıç değerlerini güncelle, Herbir temsilci için uygunluk fonksiyonu metodu hesapla, $X_\alpha - X_\beta - X_\delta$ değerlerini güncelle, $t = t+1$), Çevrim sayısı maksimum çevrim sayısından küçük olduğu sürece, Geri çevir X_α . Bu işlem sırasına göre algoritma çalışmaktadır.

3 Sonuç

Çalışma kapsamında şu adımlar izlenmiştir. Veri seti oluşturulmuş, oluşturulan veri seti veri ön işleme teknikleri ile temizlenmiş, normalize edilmiş ve kullanılabilir hale getirilmiştir. Sonrasında özellik seçimi algoritmaları kullanılarak veri özelliği yüksek kullanılabilir bir veri seti haline getirilmiştir. Elde edilen bu veri kümesindeki belirlenmiş özelliklere GWO algoritması uygulanış ve algoritma ile her bir kaydın sonucu kredi verilebilir, kredi verilemez olarak bulunmuştur. Bulunan sonuçlara göre çalışmanın başarı oranları ve doğruluk değerleri hesaplanmış bu hesaplamalara göre de algoritmanın başarısına ait gerekli yorumlamalar yapılmıştır. Bu çalışmada aşağıdaki performans metrikleri kullanılmıştır. Pozitif (P): 'Başarılı' olarak doğru etiketlenmiş müşteri sayısını, Negatif (N): 'Başarısız' olarak doğru etiketlenmiş müşteri sayısını, Doğru Pozitif (TP): 'Başarılı' olarak doğru etiketlenmiş müşterilerin tahmin edilme

sayısını, Yanlış Negatif (FN): ‘Başarısız’ olarak yanlış etiketlenen müşterilerin tahmin edilme sayısını (Gerçekte ‘Başarılı’ olan müşterinin ‘Başarısız’ olarak etiketlenme durumu), Yanlış Pozitif (FP): ‘Başarılı’ olarak yanlış etiketlenen müşterilerin tahmin edilme sayısını (Gerçekte ‘Başarısız’ olan müşterinin ‘Başarılı’ olarak etiketlenme durumu), Doğru Negatif (TN): ‘Başarısız’ olarak doğru etiketlenmiş müşterilerin tahmin edilme sayısını göstermektedir. Burada “Başarılı” terimi “kredi verilebilirliği”, “Başarısız” terimi ise “kredi verilemezliği” ifade etmektedir. Aslında problem iki sınıflı bir sınıflandırma problemidir. Performans matrikleri gerçek sonuç ve sınıflandırma sonucuna göre karışıklık Matrisi (Class Confusion Matrix) olarak gösterilmektedir. Matrisdeki sütunlar tahmin sonuçlarına göre, satırlarsa gerçek sonuçlara göre elde edilen sonuçların gösteriminde kullanılmaktadır. Bu değerler literatürde karşılaştırma ölçütü olarak kullanılacak değerlerin üretilmesinde kullanılmaktadır.

Accuracy-Doğruluk: $(TP + TN) / (P + N)$ – Doğru Sonuçların Genel Popülasyona Oranı, Presisyon – Pozitif: $TP / (TP + FP)$ – Kesinlik – Gerçek Pozitiflerin, tüm pozitiflere oranı, Presisyon – Negatif: $TN / (TN + FN)$ – Kesinlik – Gerçek Negatiflerin, tüm negatiflere oranı, Type I Accuracy: $FN / (P + N)$ – Doğruluk – Yanlış Negatiflerin Genel Popülasyona oranı, Bir müşteriyi gerçekte “Başarılı” olmasına rağmen “Başarısız” olarak etiketlemek “Type I Accuracy” terimi ile ifade edilir. ‘Başarılı’ olmasına rağmen ‘Başarısız’ olarak ifade edilen bir kaydın finansal açıdan yapılacak çalışmalarda daha kötü sonuçlar doğurabilme olasılığı bulunmaktadır. Bu sebeple başarı oran kriteri olarak “Type I Accuracy” ölçütü tercih edilmiştir.

Bir finans kurumuna kredi talebinde bulunan ve değerlendirilmesi yapılmış 16088 müşteriye ait 157 niteliği olan kayıtlar üzerinde gerekli çalışma gerçekleştirilmiştir. Nitelikler arasında yaş, meslek, evlilik durumu, çocuk sayısı, gelir bilgisi, maaş bilgisi, eş çalışma durumu, araç ve konut bilgileri, toplam kredi bilgileri, ödenen kredi bilgileri, borç bilgileri, ödenmeyen kredi bilgileri gibi kişisel ve finansal veri içeren 157 nitelik bulunmaktadır. 16088 kayıttan oluşan veri kümesinde 13718 “Başarılı” olarak sonuçlanmış, 2370 “Başarısız” olarak sonuçlanmış kredi başvuru bilgisi içermektedir. Çalışılan veri kümesi MS SQL SERVER veri tabanı üzerinde tutulmak için gerekli alanlara uygun tablo tasarımı yapılmıştır. Veri tabanında yer alan alanlar üzerinde veri ön işleme teknikleri kullanılarak aykırı veriler atılmış gürültüler temizlenmiş ve gerekli alanlara normalizasyon işlemleri uygulanmıştır. Veri kaynaklarına erişecek uygulama ara yüzlerin geliştirilmesi C# programlama diliyle geliştirilmiştir. C# ile yazılan GWO algoritması çalıştırılarak sınıflandırma çalışması yapılmış elde edilen ‘Başarılı’ ve ‘Başarısız’ kayıt sayılarına ait sonuç değerleri bulunmuştur. GWO algoritmasının seçilme nedeni sezgisel olarak sınıflandırma yapan bu yeni algoritmanın finans verilerindeki başarı oranının değerlendirilmesidir. Eğitim seti 5 parçaya bölünmüştür. K-Katlamalı Çapraz Doğrulama (K-Fold Cross Validation) yöntemi ile K değeri 5 seçilerek 4 parça (%80) eğitim verisi olarak, 1 parça (%20) ise test verisi olarak kullanılmıştır.

GWO algoritmasının probleme uygulanması şu şekilde açıklanabilir. Her bir özelliğin birbirine ve sonuca olan etkisi olduğu için öncelikle kolon değerleri sonuçlara göre sınıflandırılmıştır. Her bir özelliğin sonuca olan etkisi GWO algoritması yardımı ile hesaplanarak özellik etki değerleri bulunmuştur. Burada her bir niteliği bir kurt indexi olarak yorumlayabiliriz. Bu işlem yapılarak eğitim kümesinde sonucu “Başarılı” olarak kümelenen verilerin her bir özellik etki değeri hesaplanmış ve niteliklerin sonuca etki

katsayısı çıkarılmıştır. Aynı işlem eğitim setindeki sonucun “Başarısız” olduğu veriler içinde yapılmıştır. Test edilecek verilerin var olan değerleri sütun etki değerleri ile çarpılarak sonuca ulaşmaya hedefleyen bir uygunluk yöntemi metodu yazılmaya çalışılmıştır. Bu metod sonucunda bulunan değerlere göre “Başarılı” değer “Başarısız” değerden daha büyük bulunursa test verisinin sonucu “Başarılı”, daha küçük bulunursa test verisinin sonucu “Başarısız” olarak işaretlenir. Bu işaretleme sonucu gerçek değeri ile bulunan değeri ayrı ayrı kaydedilir. Bu kaydedilen bilgilere göre P, N, TP, FN, FP, TN bilgileri çıkartılır. Bu bilgiler üzerinden Accuracy, Precision-P, Precision-N, Type I Accuracy değerleri hesaplanır. GWO algoritması kullanarak elde edilen sonuçlar aşağıdaki gibi Tablo 2’de sunulmuştur. Bu veriler ile gerekli kazanım hesapları yapılmış elde edilen değerlere göre bulunan doğruluk sonuçları Tablo 3’de sonuç performans karşılaştırma çizelgesinde sunulmuştur.

Eğitim Oranı	Test Oranı	KS	P	N	TP	FN	TN	FP
80%	20%	16088	13718	2370	12077	1641	1226	1144
75%	25%	16088	13718	2370	11797	1921	1331	1039

Tablo 2. GWO Algoritması Sınıflandırma Değerleri

Veri Seti	Performans Metrikleri			
	Accuracy	Precision-P	Precision-N	TypeI Accur.
16088 kayıt	0.8268	0.9134	0.4276	0.1020

Tablo 3. GWO Sınıflandırma Ölçütlerinin Performans Tablosu

Tablodaki değerleri yorumlamak gerekirse veri setindeki alanlar GWO algoritması yardımıyla sınıflandırılmak istenirse 13718 kredi verilebilir kayıttın 12077 tanesi bu algoritma sayesinde de kredi verilebilir olarak bulunmuştur. Kredi verilemez denilen 2370 kayıttın 1226 tanesi kredi verilemez olarak bulunmuştur. Bu verilere göre Accuracy değeri hesaplanırsa 0.8268 lik doğruluk oranı ile bu algoritma sınıflandırma işlemini yapabiliyor demektir. Precision-P ve N değerlerini yorumlamak gerekirse P (kredi verilebilir) sonuçlanmış işlemlerdeki başarı oranı N (kredi verilemez) işlemlerdeki başarı oranına göre çok daha yüksek doğrulukta çalışmıştır. Yani veri seti üzerinde kredi verilebilir sonucunda olan kayıt sayısı ve özellik sayısı daha baskın gelmektedir denilebilir. Başarısız (kredi verilemez) olarak değerlendirilmiş kayıt sayısının artırılması ya da veri setine olumsuzluk gerekçesi olabilecek yeni nitelikler, alanların dahil edilmesi ile bu çalışmanın sınıflandırma sonuçlarının ve doğruluk derecelerinin daha yüksek doğrulukta çalışmasını sağlayabiliriz.

Sınıflandırma algoritmalarında en önemli kısım eğitim aşaması ve doğru modeli ortaya koymaktır. Bu anlamda dikkat edilmesi gereken öncelikli konu doğru bir veri seti üzerinde sistemi eğitmektir. Çalışmada ortaya konulan en önemli önerilerden birisi yapılacak sınıflandırma çalışmalarında veri setinin gerçekçi veriler içermesini sağlayacak gerekli ön işlemler uygulanmış doğru bir veri setinin kullanılmasıdır. Uygun veri seti sağlandıktan sonra performansı artırmak için sonuca en çok etki gösterecek özelliklerin

tespiti için özellik seçim yöntemleri uygulanabilir. Bu adımda başarısı kanıtlanmış algoritmalarından Kazanım Oranı, Bilgi Kazanımı, ID3 Karar Ağaçları, KN komşuluk gibi algoritmalar kullanılabilir. Özellik seçimi algoritmaları yardımıyla bir alt küme üzerinde çalışmak hem hesaplama maliyetlerini azaltacak, hem işlem sürelerini kısaltacak hem de performans artıracaktır. Az sayıda nitelikten oluşan modellerde istatistiksel sınıflandırma algoritmaları, farklı bağıllık değerlerinden ve çok sayıda nitelik barındıran veri kümelerine ait modellerde ise sezgisel algoritmaların daha başarılı sonuçlar ortaya koyduğu yapılan çalışmalar ile gözlemlenmiştir. Testlerde elde edilen modellerde ortaya çıkan sonuca göre GWO yönteminin sınıflandırma çalışmalarında başarılı olduğu tespit edilmiştir. GWO algoritmasının sonuçlarının başarı oranlarının iyi olduğu fakat çalışma süresinin daha yüksek olduğu, geliştirilebilecek metotlar sayesinde sınıflandırma işleminde kullanılacağı tespit edilmiştir. GWO algoritmasının sınıflandırma işlemi gibi özellik seçimi aşamasında da kullanılacağı sonucu ortaya çıkarılmıştır. İlerleyen çalışmalarda GWO algoritmasının başarı sonuçları, bayesian, ANN, SWM, SOM gibi algoritmaların sonuçları ile karşılaştırılacak bulunan sonuçlar üzerinde performans ve başarı oranı yorumlamaları yapılacaktır.

4 Kaynakça

1. Mirjalili, S., Mirjalili, S. M. ve Lewis, A., 2014, Grey Wolf Optimizer, *Advances in Engineering Software*, 69, 46-61.
2. Mirjalili, S., 2015, How effective is the Grey Wolf optimizer in training multi-layer perceptrons, *Applied Intelligence*, 43 (1), 150-161.
3. Kamboj, V. K., Bath, S. K. ve Dhillon, J. S., 2016, Solution of non-convex economic load dispatch problem using Grey Wolf Optimizer, *Neural Computing & Applications*, 27 (5), 1301-1316.
4. Yusof, Y. ve Mustafa, Z., 2015, Time Series Forecasting of Energy Commodity using Grey Wolf Optimizer, *Proceedings of the International MultiConference of Engineers and Computer Scientists 2015*, 1.
5. Muro, C., Escobedo, R., Spector, L. ve Coppinger, R. P., 2011, Wolf-pack (Canis lupus) hunting strategies emerge from simple rules in computational simulations, *Behavioural Processes*, 88 (3), 192-197.
6. Jayapriya, J. ve Arock, M., 2015, A Parallel GWO Technique for Aligning Multiple Molecular Sequences, *2015 International Conference on Advances in Computing, Communications and Informatics (Icacci)*, 210-215.
7. Korayem, L., Khorsid, M. ve Kassem, S. S., 2015, Using Grey Wolf Algorithm to Solve the Capacitated Vehicle Routing Problem, *3rd International Conference on Manufacturing, Optimization, Industrial and Material Engineering (Moime 2015)*, 83.
8. Aghaee, N., Hayati, M. ve Valian, E., 2014, Feedforward neural network training using Grey Wolf Optimizer, *National Conference on Technology, Energy and the Electrical and Computer Engineering Approach*.
9. Shakarami, M. R. ve Davoudkhani, I. F., 2016, Wide-area power system stabilizer design based on Grey Wolf Optimization algorithm considering the time delay, *Electric Power Systems Research*, 133, 149-159.
10. Sharma, Y. ve Saikia, L. C., 2015, Automatic generation control of a multi-area ST - Thermal power system using Grey Wolf Optimizer algorithm based classical controllers, *International Journal of Electrical Power & Energy Systems*, 73, 853-862.