GÜFBED/*GUSTIJ* (2021) 11 (1): 169-178
DOI: 10.17714/gumusfenbil.776329 Araştırma Makalesi / *Research Article*

İkili parçacık sürü optimizasyonu ve destek vektör makinelerinin hibrit kullanımı ile ilaç keşfi için özellik seçimi

Feature selection for drug discovery with hybrid usage of binary particle swarm optimization and support vector machines

Nilay SUBAŞ[1,a], Ayça ÇAKMAK PEHLİVANLI*[2,b]
[1]*İdea Teknoloji Çözümleri, Proje Yöneticisi, 34398, Maslak, İstanbul*
[2]*Mimar Sinan Güzel Sanatlar Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, 34380, Şişli, İstanbul*

• Geliş tarihi / *Received*: 01.08.2020 • Düzeltilerek geliş tarihi / *Received in revised form*: 02.12.2020 • Kabul tarihi / *Accepted*: 10.12.2020

Öz
Hastalıkların tedavisini ve önlenmesini sağlayan yeni bir ilacın keşif süreci oldukça maliyetli, karmaşık ve zaman alan bir süreç olduğu için ilaç endüstrisinde kritik bir konudur. Bu çalışma, ilaç keşif sürecinde klinik öncesi aşamayı in silico olarak da anılan hesaplamalı yöntemler ile kısaltmayı hedeflemektedir. Çalışma kapsamında potansiyel ilaç moleküllerini belirlemekte etkin ve ilgili olan özelliklerin seçimi için destek vektör makineleri ile iki sezgisel algoritma -sürekli ve ikili parçacık sürü optimizasyonu- hibritlenmiştir. İlaç molekülleri ve ilgili 161 özellikten oluşan ayrık iki veri seti eğitim ve sınama setleri olarak kullanılmış, uygun parametreler seçilerek farklı parçacık sayıları ile hem sürekli hem de ikili olarak karşılaştırmalı özellik seçimleri gerçekleştirilmiştir. İkili parçacık sürü optimizasyonunda 30 parçacık sayısıyla 49 özellik seçilmiş ve %92,54 doğruluk oranı elde edilmiştir. Diğer taraftan, doğruluk oranı sürekli parçacık sürü optimizasyonunda 50 parçacık ve 82 özellik sayısıyla %94.03 olarak bulunmuştur.

Anahtar kelimeler: Destek vektör motorları, İlaç keşfi, İstatistiksel öğrenme, Özellik seçimi, Sürekli/ikili parçacık sürü optimizasyonu

Abstract
The discovery process of a new drug that provides treatment and prevention of diseases is a critical issue in the pharmaceutical industry, as it is a costly, complex and time-consuming process. This study aims to shorten the preclinical stage in the drug discovery process with computational methods, also called in silico. Within the scope of this study, support vector machines have been hybridized with two heuristic algorithms -binary and continues particle swarm optimizations- in order to select the most relevant and informative properties for determining potential drug molecules. Two distinct datasets which consist of drug molecules with related 161 features were used as train and test sets, and both continuous and binary particle swarm optimizations were conducted with tuned parameters and different particle numbers for comparative feature selections. In binary particle swarm optimization, 49 features had been selected with 30 particles and an accuracy rate of 92.54% was obtained. On the other hand, the accuracy rate was found as 94.03% with 50 particles and 82 features by continuous particle swarm optimization.

Keywords*: Support vector machines, Drug discovery, Statistical learning, Feature selection, Continuous/binary particle swarm optimizations*

*[b] Ayça ÇAKMAK PEHLİVANLI; ayca.pehlivanli@msgsu.edu.tr, Tel: (533) 712 63 03, orcid.org/0000-0001-9884-6538
[a] orcid.org/0000-0002-3173-4942

1. Giriş

Hastalıkların tedavisini ve önlenmesinde etkin kullanıma sahip ilaçların keşif süreci oldukça maliyetli ve zaman alıcı aşamaları içermektedir. İlaç endüstrisindeki gelişmelerle, yeni ilaç keşfi aşamalarının bu denli zahmetli olması nedeni ile farmasötik çalışmalarda kullanılan deneme-yanılma gibi geleneksel yöntemlerin yerini istatistiğin ve istatistiksel düşüncenin temel alındığı, matematik ve bilgisayar bilimleri ile desteklenen "akıllı ilaç geliştirme" yöntemleri almıştır (Arciniegas vd., 2000; Rockhold, 2000). Özellikle istatistik temelli makine öğrenmesi algoritmaları ve son yıllarda sezgisel arama teknikleriyle elde edilen umut verici sonuçlar nedeni çeşitli eniyileme algoritmaları bilgisayar bilimi, tıp, finans ve mühendislik gibi birçok önemli alanda karmaşık sorunları gidermek amacı ile kullanılmıştır (Tretea, 2003). Genel olarak in-silico adı verilen bu yöntemler canlı organizma dışında yapılan in-vitro ve/ya canlı organizma üzerinde yapılan in-vivo testlere geçmeden önce aday ilaç moleküllerine yönelik öngörme, önbilgi verebilme yetkinliğindedir. Doğru bir in-silico yaklaşım, moleküle ait elde edilen bilginin laboratuvar deneylerine geçilip geçilmemesi konusunda yönlendirici olmasının yanında yapılacak testlerin tasarımında daha az deney hayvanı kullanılması, kullanılacak konsantrasyonun önceden belirlenebilmesi, zaman ve maliyetin azaltılabilmesi gibi avantajlar da sağlayabilir (Pehlivanlı ve Gümüştaş, 2019).

Literatür incelendiğinde, Ajay ve vd. 1998 yılında ilaç ve ilaç olmayan molekülleri ayırabilmek için Bayes sinir ağını kullanmış, ilaç ve ilaç olmayan molekülleri %80 doğruluk oranı ile sınıflayabilmiştir (Ajay vd., 1998). 2000 yılında Wagener ve ark. ilaç ve ilaç olmayan molekülleri sınıflandırabilmek için karar ağaçlarını kullanarak bir model geliştirmiştir. Available Chemicals Directory (ACD) ve World Drug Index (WDI) veritabanlarından gelen bileşikleri eğitim verisi olarak kullanmışlardır (Wagener, 2000). Benzer biçimde Byvatov ve ark. 2003 yılında destek vektör makineleri ve yapay sinir ağları sınıflandırıcılarının ilaç ve ilaç olmayan moleküllerin sınıflandırma başarılarını karşılaştırmıştır. Destek vektör makineleri ile ilaç ve ilaç olmayan molekülleri %82, yapay sinir ağları ile %80 doğruluk oranında sınıflayabilmiştir (Byvatov vd., 2003). Pehlivanlı (2008), Cherkasov ve Murcia-Soler verilerini düzenleyip ilaç ve ilaç olmayan molekülleri ayırt edebilmek için yapay sinir ağları, genel regresyon sinir ağları (General Regression Neural Network-GRNN), uyarlanmış GRNN (Adaptive GRNN), genetik algoritma (GA), kendi kendini düzenleyen haritalar (Self Organizing Map-SOM) ve kendi kendini global olarak düzenleyen haritalar (Self Organizing Global Ranking- SOGR) yöntemlerini kullanarak ortak karar ile karşılaştırmalı bir çalışma yapmıştır. Önerilen yöntem ile ilaç ve ilaç olmayan moleküller için sırasıyla %86,54 ve %82,67 başarı oranı elde etmiştir (Pehlivanlı, 2008).

2018 yılında Majarja ve vd. tarafından yapılan ve UCI veri bankasındaki 12 bilinen verisetine uygulanan özellik seçim yaklaşımında ikili parçacık sürü optimizasyonu kullanılmıştır. Elde edilen sonuçlar benzer algoritmalar ile karşılaştırılmış ve oldukça iyi sonuçlar elde edilmiştir (Mafarja, 2018). Aynı yıl yapılan benzer çalışmalarda parçacık sürü optimizasyon algoritması değişken seçimi için kullanılmış ve sonuçlar lojistik regresyon, en yakın k komşuluk algoritması, naive bayes gibi yapay öğrenme algortimaları ile karşılaştırmalı olarak verilmiştir (Qasim ve Algamal, 2018; Sakri vd., 2018) Al-Thanoon ve ark. ateş böceği ve parçacık sürü optimizasyon algoritmalarını hibritleyerek destek vektör motorları algoritmasında kullanılan parametreleri en iyilemeyi amaçlamışlardır. Önerdikleri yaklaşımı kemoinformatiğin en önemli konularından olan kantitatif yapı-aktivite ilişkisi alanında uygulayarak oldukça iyi sonuçlar elde etmişlerdir (Al-Thanoon vd., 2019).

Bu çalışmada, teknolojinin çok hızlı ilerlemesi ile son yıllarda oldukça önem kazanan ilaç geliştirme çalışmalarına yönelik, özellik kümelerinin indirgenerek sınıflandırma başarısını arttırmak ve potansiyel ilaç aday moleküllerinin belirlenmesi için evrimsel algoritmalardan biri olan ikili parçacık sürü optimizasyonu (BPSO) ile makine öğrenmesi sınıflandırıcı algoritmalarından destek vektör makine yöntemleri hibritlenerek, esnek yapıya sahip bir yaklaşım sunulmuştur. Söz konusu yaklaşım ilaç ve ilaç olmayan moleküllerin sınıflandırmasında en etkili değişkenlerin seçimi için ilk defa uygulanarak bu alanda da etkin olarak kullanılabileceği gösterilmiştir. Sonuçları karşılaştırmalı olarak değerlendirebilmek adına benzer model sürekli parçacık sürü optimizasyonu (PSO) ile de uygulanmıştır.

2. Materyal ve metot

2.1. Özellik seçimi

İstatistiksel öğrenme yöntemleri sınıflama ve kümeleme olmak üzere iki ana grupta incelenir. Eğitmenli öğrenme olan sınıflama, veri setinde

bulunan her bir örneği, özellikleri tarafından açıklanan bilgilere ve sınıf bilgisine dayanarak etiketlemeyi amaçlayan bir yaklaşımdır. Birçok gerçek yaşam probleminin çözümü için kullanılan verilerin gereksiz, ilgisiz, gürültülü özellikler içermesi sınıflandırma başarısını olumsuz etkilemektedir. Özellik seçim yöntemleri ilgisiz ve gereksiz özellikleri ortadan kaldırarak ya da etkilerini azaltarak boyut indirgemesi ile algoritma hızının artmasını, veri setinin daha anlaşılabilir ve görselleştirilebilir hale gelmesini ve sınıflandırma performansının iyileşmesini sağlar (Dash ve Liu, 1997; Ünler ve Murat, 2010).

Sınıflandırmaya dayalı çalışmalarda özellik seçme algoritmaları istatistiksel bilgiye dayalı olan filtreleme (filter) yöntemleri, özellikler üzerinde arama işlemleri gerçekleştiren sarmal (wrapper) yöntemler ve en iyi bölen ölçütünü bulmaya dayalı olan gömülü (embedded) yöntemler olmak üzere üç grupta toplanırlar (Vashishtha ve Vashishtha, 2016).

Filtreleme yöntemleri ile herhangi bir sınıflandırıcı algoritma kullanmadan istatistiksel ölçütlere dayalı fonksiyonlar yardımıyla özellik seçimi yapılırken, ***sarmal yöntemlerde*** çeşitli öğrenme algoritmaları kullanılarak en iyi tahmin performansını gösteren özellikler seçilmektedir. ***Gömülü yöntemler*** ise yapısında hem sınıflandırma hem de özellik seçimi algoritması barındırır (Guyon ve Elisseeff, 2003).

2.2. Parçacık sürü optimizasyonu (PSO)

1995 yılında psikolog James Kennedy ve elektrik mühendisi Russel Eberhart tarafından geliştirilen PSO kuş, balık, arı gibi sürü halinde yaşayan hayvanların sosyal davranışlarını esas alan popülasyon tabanlı bir eniyileme algoritmasıdır (Kennedy ve Eberhart, 1995). Sürü halinde hareket eden hayvanlar aralarındaki bilgi paylaşımı sayesinde, zengin besin kaynaklarına ulaşmak için sürünün hedefe en yakın bireyini takip ederek hız ve konumlarını bu bireye göre güncellerler (Der vd., 2008).

PSO'da, parçacık (kuş) olarak isimlendirilen birden fazla çözüm adayı bulunur. Bu parçacıklardan oluşan popülasyona sürü (swarm) denir. PSO eniyi ya da en iyiye yakın çözüm bulmak için her biri çözüm adayı olan parçacıklar oluşturur. Başlangıçta bu parçacıkların konumu, belirlenen sınırlar altında rastgele seçilerek başlangıç sürüsü rastgele oluşturulur. Sürüyü oluşturan her bir parçacık, konum ve hız bilgisine sahiptir. Hız bilgisi, parçacığın mevcut konumundan bir sonraki konuma geçişini sağlarken, konum bilgisi de çözümü temsil etmektedir.

PSO'da her bir parçacık ayrıca uygunluk fonksiyonu ile bulunan uygunluk değerlerine sahiptir. Uygunluk fonksiyonu, her bir parçacığın en iyi çözüme olan uzaklığını değerlendirir. Parçacık pozisyonunu güncelledikçe, pozisyonunun koordinatlarını bu uygunluk fonksiyonuna gönderir ve böylece parçacığın uygunluk değeri (en iyi çözüme olan uzaklık) hesaplanır (Der vd., 2008).

Her bir parçacık konumunun değişim miktarını belirleyen hız vektörünü, kendi tecrübelerinden faydalanarak yaptığı hareket (bilişsel hareket), sürü ile bilgi paylaşımından (sosyal hareket) faydalanarak yaptığı hareket ve aynı yönde sabit hızla yapılan hareket (eylemsizlik hareketi) ile Eşitlik 1'e göre hesaplar.

$$v_{ij}^{(t+1)} = w v_{ij}^{(t)} + c_1 r_1 \left(pbest_{ij}^{(t)} - x_{ij}^{(t)} \right) + c_2 r_2 (gbest_j^{(t)} - x_{ij}^{(t)}) \quad (1)$$

Eşitlik 1'de; t iterasyon sayısını, $v_{ij}^{(t+1)}$ ve $v_{ij}^{(t)}$ sırası ile i. parçacığın (t+1) anındaki yeni ve t anındaki eski hız değerini, $pbest_{ij}^{(t)}$ t anında parçacığın yerel en iyi değerini, $gbest_j^{(t)}$ t anında sürünün en iyi değerini, $x_{ij}^{(t)}$ i. parçacığın t anındaki konum değerini, c_1 ve c_2 hızlandırma katsayılarını, r_1 ve r_2 $[0-1]$ aralığında rasgele üretilen sayıları, w ise eylemsizlik ağırlığını ifade etmektedir.

$pbest_{ij}^{(t)}$, t anında parçacığın kendi ulaştığı konumlar arasındaki en iyi uygunluk değerine sahip konum (yerel en iyi), $gbest_j^{(t)}$, t anına kadar sürü tarafından bulunan en iyi uygunluk değerini veren konum (global en iyi) olarak tanımlanır. N adet parçacıktan oluşan sürü ile çalışan bir PSO'da; her iterasyonda N adet $pbest$ mevcut iken sadece bir adet $gbest$ mevcuttur.

Hız güncellemesinden sonra her bir parçacık eski konum vektörüne (t zamanındaki x vektörü) yeni hız vektörünü ekleyerek konumunu (t+1) zamanında x vektörü olarak Eşitlik 2'ye göre günceller.

$$x_{ij}^{(t+1)} = x_i^{(t)} + v_{ij}^{(t+1)} \quad (2)$$

Eşitlik 2'de; t iterasyon sayısını, $v_{ij}^{(t+1)}$ $i.$ parçacığın yeni hız değerini, $x_{ij}^{(t+1)}$ ve $x_{ij}^{(t)}$ sırası ile $i.$ parçacığın (t+1) anındaki yeni ve t anındaki eski konumunu belirtmektedir.

Böylece yeni konum bilgisini uygunluk fonksiyonuna göndererek yeni uygunluk değerini elde eder. En küçükleme problemlerinde uygunluk değeri küçük olan parçacıklar büyük olana tercih edilirken, en büyükleme problemlerinde uygunluk değeri büyük olan parçacıklar küçük olana tercih edilir (Ortakcı ve Güloğlu, 2012).

2.3. İkili parçacık sürü optimizasyonu (BPSO)

PSO, sürekli uzayda eniyileme problemlerini çözmek için geliştirilmiştir. Bununla birlikte, birçok eniyileme problemi, özellik seçimi, zamanlama ve yönlendirme gibi ayrık alanlarda tanımlanır (Al-Thanoon vd., 2019). PSO'nun uygulanabilirliğini geliştirmek ve bu gibi ayrık problemlerin üstesinden gelmek için Kennedy ve Eberhardt tarafından 1997 yılında ikili parçacık sürü optimizasyonu (BPSO) tanımlanmıştır (Kennedy ve Eberhart, 1997).

İkili parçacık sürü optimizasyonunda yerel yani kişisel en iyi ($pbest$) ve global en iyi ($gbest$) değerleri sürekli versiyondaki gibi güncellenir. İkili PSO ile sürekli PSO arasında temel iki farklılık vardır. Birincisi, BPSO'da, her parçacığın konumu 0 ve 1 değerlerinden yani ikili değerlerden oluşmaktadır. İkinci farklılık ise hız tanımında olup hız, parçacığın 0 ya da 1 değerini alma olasılığı olarak tanımlanmaktadır (Khanesar vd., 2007). Diğer yandan BPSO'da hız güncellemesi sürekli PSO'da olduğu gibidir.

Parçacıkların konumu ikili sayı (0 ve 1) değeri alacağından, hız değerini [0,1] aralığına dönüştürmek ve her x_{ij}'nin 1 değerini alma olasılığını belirlemek amacıyla Eşitlik 3 ile verilen Sigmoid fonksiyonu kullanılır:

$$S(v_{ij}) = \frac{1}{1+e^{-v_{ij}}} \quad (3)$$

Eşitlik 3'te verilen v_{ij} hızı ifade ederken, S Sigmoid fonksiyondur. Sigmoid fonksiyonu ile hız değeri [0, 1] aralığında belirlendikten sonra parçacığın konum güncellemesi (0 ya da 1 değerlerinden hangisini alacağının belirlenmesi) Eşitlik 4'e göre yapılır:

$$x_{ij} = \begin{cases} 1, & rand() < S(v_{ij}) \\ 0, & diğer\ durumda \end{cases} \quad (4)$$

x_{ij} i parçacığın hız değeri olup, *rand()* fonksiyonu ile [0, 1] aralığında düzgün dağılımdan rastgele bir sayı seçilir (Cervante vd., 2012).

2.4. Destek vektör motorları (DVM)

Vapnik tarafından 1963'te ortaya atılan istatistiksel öğrenme teorisine dayanan destek vektör makineleri, çok boyutlu verilerde sınıfları birbirinden ayıran hiper düzlemi belirlemeye çalışan istatistik tabanlı bir makine öğrenmesi yöntemidir (Vapnik, 1995). Bu ayırıcı hiper düzlem, farklı sınıflardaki veriler arasındaki geometrik anlamdaki genişliği en büyükleyecek şekilde bir karar yüzeyi olarak tanımlanabilir (Pehlivanlı, 2016). Bulunan bu ayırıcı hiper-düzleme komşu olan, ait olduğu sınıfın sınırını belirleyen noktalara ise *destek vektörleri* adı verilir. Destek vektörleri, ayırıcı hiper düzleme en yakın olan, düzlemin konumunu ve yönünü etkileyen örnekler olarak da tanımlanabilirler.

w ağırlıklar vektörü olmak üzere, $2/\|\vec{w}\|$ olarak verilen bu genişliği en büyükleyecek hiper düzlemi bulmak ile $\|\vec{w}^2\|/2$ 'yi en küçük yapmak eşdeğerdir. DVM, *J (w)* tarafından verilen bir hata fonksiyonunu en aza indirecek şekilde Eşitlik 5'te ilgili kısıtlar altında verildiği gibi formüle edilmiştir.

$$J(w) = \frac{1}{2}w^T w + C\sum_i \xi_i \quad (5)$$

$d_i[w^T\varphi(x_i) + b] \geq 1 - \xi_i\ ve\ \xi_i \geq 0, i = 1,2,\dots,n$ kısıtı altında

burada b sabit, ξ_i gevşek değişkenler olup sınıflandırma hatasına izin veren parametrelerdir. C parametresi, sınıflar arasındaki geometrik genişliğin boyutu ile ayrıştırılamaz nokta sayısı arasındaki dengeyi düzenler. x_i 'ler, eğitim veri setindeki her bir gözlemi temsil eden bağımsız değişkenlerden oluşan vektörler olup, sınıf etiketleri d_i ile temsil edilmişlerdir. $\varphi(.)$ doğrusal olmayan karar sınırı oluşturmak için bir çekirdek işlevidir.

2.5. Veri seti

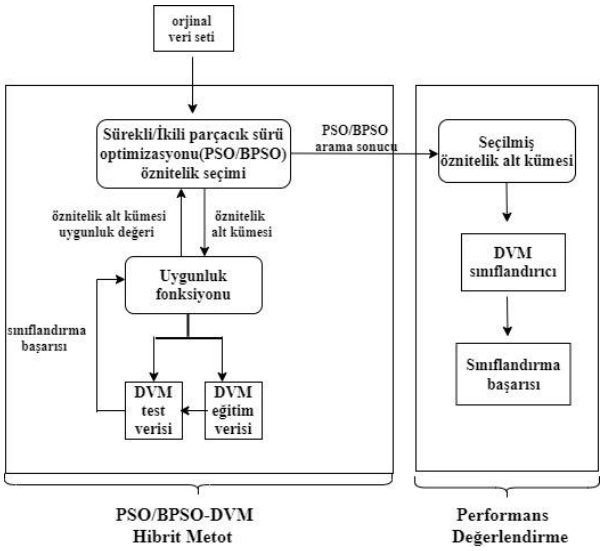
Çalışmada ilaç ve ilaç olmayan moleküllerden oluşan tamamen ayrık iki veri seti kullanılmıştır. Söz konusu moleküller Cherkasov ve Murcia-Soler'in çalışmalarından elde edilerek, her bir moleküle ait açıklayıcı özellikler Molecular Operating Environment (MOE, 2006) programı ile hesaplanmıştır (Cherkasov, 2006; Murcia-Soler vd., 2003). Murcia-Soler veri seti BPSO-DVM ve

PSO-DVM biçimindeki hibrit yaklaşım ile özellikleri seçmek için, seçilen bu özelliklerin sınıflandırma başarısına etkisini ölçmek için ise Cherkasov verisi kullanılmıştır.

Cherkasov veri seti 523 adet onaylanmış antimikrobiyal, 959 adet onaylanmış ilaç, 1202 adet ilaç benzeri moleküller olmak üzere toplam 2684 bileşik içermektedir. Murcia-Soler veri seti farmakolojik aktivitesi onaylanmış 416 bileşik ve farmakolojik aktivitesi olmayan 225 bileşik olmak üzere toplamda 641 bileşik içermekte olup her iki veri seti de MOE ile hesaplanan 161 değişkene sahiptir (Pehlivanlı, 2008; Pehlivanlı vd., 2008).

3. Hibrit model

Bu çalışmada, ilaç (aktif) ve ilaç olmaya aday (aktif olmayan) moleküllerden oluşan veri setinde ilaç olmayı belirlemede etkili özellikleri seçerek sınıflandırma başarısını arttırmak ya da daha az özellik ile sınıflama yapmak, böylece binlerce molekül arasından potansiyel ilaç aday moleküllerinin seçilmesini sağlamak amacıyla, temeli BPSO/PSO ve DVM'ye dayanan hibrit bir özellik seçim yöntemi uygulanmıştır. Hibrit yöntem ile ilaç olmayı belirlemede en etkili özellikler seçilerek, ilaç olarak etiketlenmiş molekülleri destek vektör makineleri yardımı ile sınıflayıp ilaç olma potansiyeline sahip aday moleküller belirlenmeye çalışılmıştır. Amaca yönelik hibrit yöntem Şekil 1'deki akış diyagramına göre uygulanmıştır (Subaş, 2019).



Şekil 1. PSO/BPSO-SVM Hibrit Yöntem Akış Diyagramı

Şekil 1'de gözlenen uygunluk fonksiyonu, her bir parçacığın en iyi çözüme olan uzaklığının belirlenmesini sağlar. Parçacıklar konum değerlerini uygunluk fonksiyonuna göndererek uygunluk değeri alırlar. Bu uygunluk değeri, hız güncellemede kullanılan, sürüdeki iki en iyi değer olan ve Bölüm 2.2.de açıklanan *pbest* ve *gbest*'in belirlenmesini sağlarken, parçacığın bileşeninin 0 ve 1 değerini almasını etkiler. 0 ve 1 değerleri, sırasıyla sınıflandırmada etkisiz (seçilmemiş) ve etkili (seçili) özellikleri temsil eder.

Seçilen özellik alt kümesi altında her bir parçacığın uygunluk değeri Eşitlik 6 ile hesaplanır.

$$f(x) = w_1 A(x_i) + \left(w_2 \left(d - S(x_i)\right)/n\right) \qquad (6)$$

Belirlenen uygunluk fonksiyonunun iki temel amacı vardır. İlaç olmayı belirlemede etkin özellik alt kümesinin seçilmesini sağlamak ve indirgenmiş veri kümesi ile sınıflandırma başarısını arttırmaktır (Subaş, 2019).

$A(x_i)$, seçilen özellik alt kümesinin DVM tarafından sağlanan sınıflandırma doğruluğudur. Seçilen özellik alt kümelerinin sınıflandırma başarısını sınamak için k kat çapraz doğrulama yöntemi kullanılmıştır. Veri seti k alt kümeye bölünmüştür. Her defasında k alt kümeden 1 tanesi test verisi olarak *k-1* tanesi eğitim verisi olarak kullanılmıştır ve tüm k deneme için ortalama doğruluk hesaplanmıştır.

$S(x_i)$, parçacık x_i de seçilen özellik sayısı iken (parçacıkta 1 değerini alan bitlerin sayısı), d veri setinde bulunan toplam özellik sayısıdır. Parçacığın bileşeninin 1 değerini alması ilgili özelliğin seçildiği, 0 değerini alması ilgili özelliğin seçilmediği anlamına gelmektedir.

Eşitlik 6'da verilen w_1 ve w_2 sırasıyla sınıflandırma başarısı ve seçilen özellik sayısı için ağırlıklandırma katsayılarıdır.

4. Değerlendirme ölçütleri

Çalışmada DVM sınıflandırma başarısını değerlendirmek için doğruluk oranına ek olarak, karmaşıklık matrisi kullanılarak bilinen ve yaygın olarak kullanılan doğruluk, F-ölçütü, kesinlik, hassaslık, özgüllük ölçütlerine ek olarak ROC(Receiver Operating Characteristics) eğrisi altında kalan alan hesaplanmıştır. ROC eğrisi sınıflandırıcının tüm olası değerler üzerinde performansını özetlemek için kullanılan bir grafiktir. ROC eğrisinin altında kalan alan AUC (Area Under Curve) olarak ifade edilmekte olup yüksek olması istatistiksel olarak daha anlamlı bir

sonuç elde edildiği anlamına gelir (Sokolova vd., 2006).

5. Uygulama tasarımı

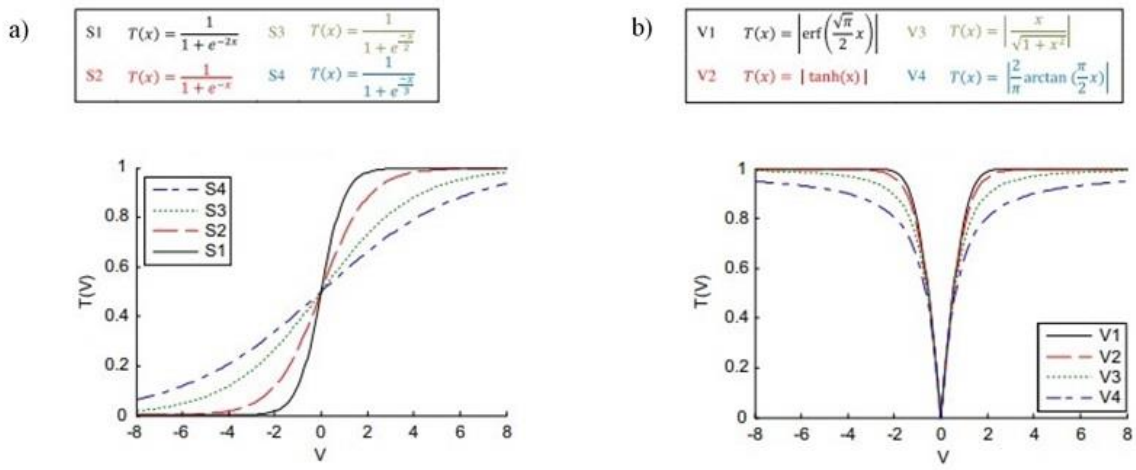
5.1.Parametre seçimi:

Bu çalışmada, hibrit yöntemin performansını etkileyen parametre değerleri için çekirdek fonksiyonu olarak radyal tabanlı fonksiyonlar kullanılmıştır. Değişken parametre C ve çekirdek parametresi σ'nın optimal değeri, şebeke arama ve çapraz geçerleme ile belirlenmiştir. Yapılan çalışma sonucunda, C ve σ sırasıyla 1.5 ve 3 olarak seçilmiştir.

Problemin uygunluk fonksiyonu olan Eşitlik 6'da, seçilen özellikler ile elde edilen sınıflandırma doğruluğu olan $A(x_i)$'yi elde etmek için kullanılan k kat çapraz doğrulama yönteminde k=10 olarak alınmıştır. Veri seti 10 alt kümeye bölünmüştür. Her defasında 10 alt kümeden 1 tanesi test verisi olarak 9 tanesi eğitim verisi olarak kullanılmıştır ve tüm 10 deneme için ortalama doğruluk ($A(x_i)$) hesaplanmıştır. Murcia-Soler verisinde 161 tanımlayıcı özellik bulunduğundan Eşitlik 6'da d, 161 olarak alınmıştır. Sınıflandırma doğruluğu daha ön planda tutulduğundan w_1 değeri [0.6, 0.9] aralığında seçilirken, w_2 değeri 1-w_1 olacak şekilde seçilmiştir.

5.2. Transfer fonksiyonu seçimi:

Hız değerini [0, 1] aralığına dönüştüren ve her bir parçacığın bitlerinin 0 ya da 1 değerini alma olasılığını belirleyen sigmoid (transfer) fonksiyonu için, daha iyi sonuç alabilmek adına, geleneksel ikili parçacık sürü optimizasyonunda kullanılan sigmoid fonksiyonuna ek olarak yedi farklı sigmoid (transfer) fonksiyonu daha kullanılmıştır. Şekil 2'de ilk dört fonksiyonun eğrisi S şeklinde olduğundan S-biçimli, diğer dört fonksiyonun eğrisi V şeklinde olduğundan V-biçimli transfer fonksiyon ailesi olarak ifade edilir (Mirjalili ve Lewis, 2013). Geleneksel ikili parçacık sürü optimizasyonunda kullanılan sigmoid (transfer) fonksiyonu Şekil 2a'da *S2* ile ifade edilmektedir.



Şekil 2. (a) S-biçimli ve **(b)** V-biçimli transfer(sigmoid) fonksiyon ailesi (Mirjalili ve Lewis, 2013)

Seçilen parametreler kullanılarak yapılan hesaplamalara göre sınıflandırma başarısı açısından en iyi sonuç geleneksel ikili parçacık sürü optimizasyonunda kullanılan sigmoid fonksiyonu (S2) ile elde edildiğinden çalışmada bu transfer fonksiyonu kullanılmıştır.

Parçacık sürü optimizasyonunun ana parametreleri olan parçacık sayısı (sürü büyüklüğü) ve iterasyon sayısı probleme bağlı değerler olduğundan Murcia-Soler veri setine uygulanan parçacık sürü optimizasyonu algoritmasında farklı parçacık ve iterasyon sayılarına göre denemeler yapılarak seçilen özellikler kullanılarak Cherkasov veri seti üzerinden destek vektör makineleriyle elde edilen doğruluk kesinlik, hassaslık, F-ölçütü, özgüllük ölçütlerine ilişkin sonuçlar elde edilip değerlendirilmiştir.

Çalışmanın ana amacı ikili parçacık sürü optimizasyonu ve destek vektör makinesi algoritmasının hibritlenmesi ile ilaç ve ilaç olmayan molekülleri ayırmada en etkili değişken setini bulmaktır. Elde edilen indirgenmiş veri seti üzerine destek vektör makinaları uygulanarak sonuçlar elde edilmiştir. Sonuçları daha net yorumlayabilmek adına ikili parçacık sürü optimizasyonuna ek olarak sürekli parçacık sürü optimizasyonu ile destek vektör makinaları hibritlenerek elde edilen değişken seti ile de sonuçlar elde edilmiş ve yaklaşımlar karşılaştırmalı olarak verilmiştir.

6. Bulgular

6.1. İkili parçacık sürü optimizasyonu

İkili parçacık sürü optimizasyonunda 20 ile 100 arası farklı parçacık sayıları için denemeler yapılmıştır. Parçacık ve iterasyon sayısına göre yapılan denemeler incelendiğinde seçilen özellik sayısı ve sınıflandırma başarısı açısından en iyi sonuç parçacık sayısı 30, iterasyon sayısı 400 olarak alındığında elde edilmiştir. 30 parçacık ve 400 iterasyonda seçilen 49 özellik için doğru sınıflandırma oranı %92,54 olarak saptanmıştır

Tablo 1. 30 parçacık ile seçilmiş özellikler ile destek vektör makineleri ile elde edilen sınıflandırma başarı oranları (%)

İterasyon	Doğruluk	Kesinlik	Hassaslık	F-ölçütü	Özgüllük	AUC	Özellik Sayısı
25	89.55	89.38	89.67	89.48	89.67	89.38	61
50	91.79	91.76	91.62	91.69	91.62	91.76	62
75	91.42	91.25	91.52	91.35	91.52	91.25	64
100	91.42	91.35	91.28	91.32	91.28	91.35	57
200	92.16	92.27	91.88	92.04	91.88	92.27	64
300	**92.91**	**92.77**	**92.95**	**92.85**	**92.95**	**92.77**	65
400	**92.54**	92.41	92.53	92.47	92.53	92.41	**49**
500	92.54	92.52	92.38	92.44	92.38	92.52	66

Tablo 1 incelendiğinde çok az farkla en iyi sonuç 300 iterasyon ile elde edilmesine rağmen, özellik sayısı 400 iterasyon ile elde edilen sonuçlara göre daha fazladır. Amaç çok daha az özellik ile iyi sonuçlar elde etmek olduğundan 49 özellikle elde edilen başarı oranı tercih edilmiştir. Tablo 1 ve 2'te verilen iterasyon sayılarına karşılık gelen özellik sayıları Murcia-Soler veri seti ile hibrit yaklaşım kullanılarak elde edilmiştir. Tablolarda özetlenen diğer ölçütler Murcia-Soler veri setinden tamamen ayrık olan Cherkasov veri seti üzerine hibrit modelin seçtiği değişkenler ile elde edilen sonuçları içermektedir.

6.2. Sürekli parçacık sürü optimizasyonu

Sürekli parçacık sürü optimizasyonunda ise yine 20-100 arası parçacık sayısı için yapılan denemeler sonucunda 50 parçacık sayısıyla 300 iterasyon en iyi sonucu vermiştir. Tablo 2'te koyu renk ile gösterilen değerler iterasyon sayısı 300 ve 82 adet değişkene karşılık gelmekte olup diğerlerine oranla yaklaşık %3 düzeyinde daha iyi sonuçlar vermiştir.

Tablo 2. 50 parçacık ile seçilmiş özellikler ile destek vektör makineleri ile elde edilen sınıflandırma başarı oranları(%)

İterasyon	Doğruluk	Kesinlik	Hassaslık	F-ölçütü	Özgüllük	ROC	Özellik sayısı
25	90.67	90.52	90.69	90.59	90.69	90.52	78
50	91.04	91.01	90.87	90.93	90.87	91.01	82
75	92.54	92.45	92.45	92.45	92.45	92.45	84
100	91.79	91.63	91.86	91.72	91.86	91.63	79
200	91.04	90.88	91.1	90.97	91.1	90.88	85
300	**94.03**	**93.88**	**94.28**	**94**	**94.28**	**93.88**	**82**
400	90.67	90.6	90.53	90.56	90.53	90.6	88
500	90.30	90.13	90.35	90.22	90.35	90.13	89

Elde edilen tüm sonuçlar Tablo 3'te özetlenmiştir. Tabloda, ikili sürü optimizasyonu ile destek vektör makinelerinin hibritlenmesi ile elde edilen değişken sayıları ve karşılık gelen değerlendirme ölçütleri (BPSO-DVM), sürekli sürü optimizasyonu ile destek vektör makinelerinin hibritlenmesi ile elde edilen değişken sayıları ve karşılık gelen değerlendirme ölçütleri (PSO-DVM) ile herhangi bir değişken seçimi yapılmadan (DVM) Cherkasov sınama verisi ile elde edilen sonuçlar karşılaştırılmıştır.

Tablo 3. Kullanılan yöntemlere ilişkin değerlendirme oranları

Metod	Özellik Sayısı	Doğruluk	Kesinlik	Hassaslık	F-ölçütü	Özgüllük	AUC
BPSO-DVM	49	92.54	92.41	92.53	92.47	92.53	92.41
PSO-DVM	82	94.03	93.88	94.28	94	94.28	93.88
DVM	161	91.42	91.26	91.6	91.36	91.6	91.26

BPSO-DVM: ikili sürü optimizasyonu-destek vektör makineleri hibrit modeli, PSO-DVM: sürekli sürü optimizasyonu-destek vektör makineleri hibrit modeli, DVM: destek vektör makineleri

Cherkasov veri setinde özelliklerin tamamı kullanıldığında destek vektör makineleri algoritması ile doğru sınıflandırma oranı %91,42 olarak saptanmış olup, bu oran ikili parçacık sürü optimizasyonu kullanıldığında seçilmiş 49 özellik için %92,54 olarak elde edilmiştir. BPSO-DVM ile özellik sayısı yaklaşık %70 oranında indirgenirken sınıflandırma başarısı %1,12 artmıştır. Sürekli parçacık sürü optimizasyonu ile seçilen 82 özellik için doğru sınıflandırma oranı %94,03 olarak bulunmuştur. PSO ile doğru sınıflandırma başarısında DVM'ye göre yaklaşık %3'lük artış elde edilirken özellik sayısı yaklaşık %50 oranında indirgenmiştir. Cherkasov veri setinde özelliklerin tamamının kullanıldığı duruma göre ikili ve sürekli parçacık sürü optimizasyonu ile seçilen çok daha az sayıda özellik ile daha iyi sınıflandırma başarısının elde edildiği açıkça gözlemlenmiştir. Öte yandan, veri setinde bulunan sınıf sayıları eşit olmamasına rağmen hassaslık ve özgüllük ölçütlerinin oldukça dengeli olması ve AUC değerinin yüksek olması modelin başarısını ortaya koyan önemli göstergelerdir.

Her iki yaklaşım için de tablolar incelendiğinde genel olarak iterasyon sayısı yükseldikçe başarı oranındaki artış miktarındaki değişkenlik azalmıştır. Hibrit yaklaşımda değişkenlerin seçim aşaması yüksek iterasyon sayılarında uzun süre almasına ve başarı oranında belirgin değişiklik göstermemesine rağmen değişken sayılarında düşüş sağlayabilmektedir. Bu bağlamda, değişken sayılarındaki değişimi görmek ve bütünlük sağlaması açısından yüksek iterasyon sayısı ile elde edilen sonuçlar ilgili tablolarda verilmiştir.

7. Tartışma ve sonuçlar

İlaç geliştirme süreci uzun zaman alan, karmaşık ve maliyetli bir süreçtir. Özellikle, on binlerce molekülün incelendiği in-vivo ve in-vitro aşamalarını kısaltmak, yapılacak testlerin tasarımında hayvan deneylerini en aza indirgemek, kullanılacak konsantrasyonun önceden belirleyebilmek, zaman ve maliyeti azaltabilmek adına in-silico yaklaşımlar önerilmiştir. İlaç veri setlerindeki ilgisiz özelliklerin varlığı, sınıflandırma başarısını olumsuz etkilemekte ve dolayısıyla ilaç olabilecek moleküllerin gözden kaçmasına sebep olmaktadır. Bu problemi çözmek için tek başına özellik seçimi ya da tek başına sınıflandırma algoritmaları kullanmak yerine, bu iki yaklaşımın hibrit kullanımı ile daha etkin bir özellik seçimi gerçekleştirerek ilaç geliştirme sürecini, karmaşıklığını ve maliyetini azaltmak mümkün olmaktır.

Bu çalışmanın amacı, ilaç geliştirmek için kullanılan veri setlerinde ilgisiz özellikleri eleyip, ilaç olmaya aday moleküllerin hangi özelliklerinin ilaç olmayı belirlemede etkili olduğunu tespit ederek, daha az sayıda özellik ile daha iyi ya da eş değer sınıflandırma başarısı elde etmektir. Bu amaç doğrultusunda bir eniyileme algoritması olan ikili parçacık sürü optimizasyonu ile sınıflama algoritması olan destek vektör makineleri hibritlenerek etkin bir özellik seçme yöntemi uygulanmıştır. Çalışmada, ilaç ve ilaç olmayan moleküllerden oluşan, Molecular Operating Environment programı ile hesaplanan 161 açıklayıcı özellik içeren Cherkasov ve Murcia-Soler olmak üzere iki veri seti kullanılmıştır. Murcia-Soler veri seti hibrit yaklaşımda ikili ve sürekli parçacık sürü optimizasyonu ile özellikleri belirlemek için kullanılmış ve 161 özellik arasından ilaç olmayı belirlemede en etkili 49 özellik BPSO-DVM ile, 82 tanesi ise PSO-DVM ile seçilmiştir. Hibrit yaklaşımın seçtiği özelliklerin sınıflandırma başarısını değerlendirmek için Cherkasov verisi kullanılmıştır. Sonuçlar, özellik sayısı ve sınıflandırma başarısı açısından birbiri ile ve tüm değişkenler kullanılarak elde edilen DVM sonuçları ile karşılaştırılmıştır.

Bu çalışmada elde edilen sonuçlar doğrultusunda ilaç geliştirme sürecinin erken evrelerinde, etkin özellik seçimi yöntemi sonucunda daha az sayıda özellik ile sınıflandırma başarısında iyileşme elde edilebilmektedir. Bunun sonucunda potansiyel ilaç aday moleküllerin seçilmesiyle, laboratuvar ortamında incelenecek aday molekül veri seti minimize edilmekte ve böylece ilaç geliştirme süresinin, maliyetinin ve hayvan deney sayısının azaltılabileceği düşünülmektedir.
Ayrıca çalışmada gerçekte ilaç olmayıp modelimiz tarafından ilaç olarak sınıflanan yani yanlış pozitif

(False Positive) olarak sınıflanan moleküller ilaç moleküllerine benzerlik göstermeleri nedeni ile ilaç-benzeri (drug-like) olarak değerlendirilebilirler. Bu da onbinlerce ilaç olmadığı bilinen molekül yerine ilaç-benzeri bu moleküllerin daha az boyut ve molekül sayısı ile in-vitro ve in-vivo aşamalarında öncelikli olarak değerlendirilmelerinin daha etkili olabileceğini öneri olarak ortaya koymaktadır.

Uygulanan hibrit çerçevenin, kullanılan yöntem ve yaklaşımlar açısından esnek bir yapıya sahip olması nedeniyle, sınıflandırma performansını arttıracak farklı eniyileme ya da sınıflama algoritmalarının kullanımı mümkündür. Aynı zamanda farklı en iyileme algoritmalarının birden fazla sınıflama algoritmasıyla hibrit kullanımı ile birlikte öğrenme gerçekleştirilip daha etkin özellik seçme yöntemi geliştirilebilir. Eniyileme algoritmalarının gücüyle sınıflama algoritmalarını birleştiren, hibrit ve esnek yapıya sahip bu in-silico yaklaşım, yalnızca ilaç verisi ile sınırlı kalmayıp, benzer tasarım çerçevesinde farklı alanlardaki verilere de uygulanabilir.

İleride yapılacak çalışmalarda, önerilen hibrit yaklaşım, farklı algoritmalar ile farklı veri setleri üzerine uygulanarak başarı değerlendirmesi yapılacaktır. Ancak çalışmada ortaya çıkan en önemli dezavantaj hesaplama süresinin fazla olmasıdır. Ayrıca kullanılan yöntemdeki parametre sayısının çokluğu ve eniyileme aşaması da önemli bir etkendir. Tüm bunlara karşın model tarafından önerilen çok daha düşük sayıdaki değişken seti özellikle in-vitro ve in-vivo çalışmaları azaltmak yönünde olumlu bir adımdır.

Teşekkür

Bu çalışma, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı Yüksek Lisans Programı'nda, Nilay Subaş tarafından, Doç. Dr. Ayça Çakmak Pehlivanlı danışmanlığında tamamlanan "Sürekli/ikili parçacık sürü optimizasyonu ve destek vektör makinelerinin hibrit kullanımı ile özellik seçimi" başlıklı Yüksek Lisans tezinden üretilmiştir. Tezin inceleme ve değerlendirme aşamasında yapmış oldukları katkılardan dolayı juri üyelerine teşekkür ederiz.

Kaynaklar

Ajay, W., Walters, P. and Murcko, M. A. (1998). Can we learn to distinguish between "drug-like" and "nondrug-like" molecules? *Journal of Medicinal Chemistry, 41*, 3314-3324. https://doi.org/10.1021/jm970666c

Al-Thanoon, N. A., Qasim, O. S. and Algamal, Z. Y. (2019). A new hybrid firefly algorithm and particle swarm optimization for tuning parameter estimation in penalized support vector machine with application in chemometrics. *Chemometrics and Intelligent Laboratory Systems, 184*, 142-152. https://doi.org/10.1016/j.chemolab.2018.12.003

Arciniegas, F., Bennett, K., Breneman, C. and Embrechts, M.J. (2000). Molecular database mining using self-organizing maps for the design of novel pharmaceuticals. *Intelligent Engineering Systems through Artificial Neural Networks: Smart Engineering System Design, 10*, 477-481. St. Louis, MO.

Byvatov, E., Fechner, U., Sadowski, J. and Schneider, G. (2003). Comparison of support vector machine and artificial neural networks systems for drug/nondrug classification. *Journal of Chemical Information and Computer Sciences, 43(6)*, 1882-1889. https://doi.org/10.1021/ci0341161

Cervante, L., Xue, B. and Zhang, M. (2012). Binary particle swarm optimization for feature selection: a filter-based approach. *IEEE Congress on Evolutionary Computation*, 1-8. Brisbane, QLD. https://doi.org/10.1109/CEC.2012.6256452

Cherkasov, A. (2006). Can bacterial-metabolite-likeness model improve odds of in-silico antibiotic discovery? *Journal of Chemical Information and Modeling, 46(3)*, 1214-1222. https://doi.org/10.1021/ci050480j

Dash, M. and Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis, 1(1-4)*, 131–150. https://doi.org/10.1016/S1088-467X(97)00008-5

Der, O., Vural, A. ve Yıldırım, T. (2008). Parçacık sürü optimizasyonu tabanlı evirici tasarımı. *Elektrik-Elektronik ve Biyomedikal Mühendisliği Konferansı*, 1-4. Bursa.

Guyon, I. and Elisseeff, A., (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3(7-8)*, 1157-1162. https://doi.org/10.1162/153244303322753616.

Kennedy, J. and Eberhart, R. C. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks 4*, 1942–1948. Piscataway, NJ. https://doi.org/10.1109/ICNN.1995.488968

Kennedy, J. and Eberhart, R. C. (1997). A discrete binary version of the particle swarm algorithm. *IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, 5*, 4104–4108. Orlando, FL. http://doi.org/10.1109/ICSMC.1997.637339

Khanesar, M, A., Tavakoli, H., Teshnehlab, M. and Shoorehdeli, A., M. (2007). A novel binary particle swarm optimization. *Mediterranean Conference on Control & Automation*, 1-6. Athens. https://doi.org/10.1109/MED.2007.4433821

Mafarja, M., Jarrar, R., Ahmad, S. and Abusnaina, A. A. (2018). Feature selection using binary particle swarm optimization with time varying inertia weight strategies. *Proceedings of the 2nd International Conference on Future Networks and Distributed Systems Association for Computing Machinery, 18*, 1–9. New York, NY. https://doi.org/10.1145/3231053.3231071

Mirjalili, S. and Lewis, A. (2013). S-shaped versus V-shaped transfer functions for binary particle swarm optimization. *Swarm and Evolutionary Computation, 9*, 1–14. https://doi.org/10.1016/j.swevo.2012.09.002

MOE, Molecular Operational Environment, (2006). Chemical Computing Group Inc., Montreal, Canada.

Murcia-Soler, M., Pe´Rez-Gimenez, F., Garcia-M., J., Salabert-Salvador, M. T., Diaz-Villanueva, W. and Castro-Bleda, M. J. (2003). Drugs and nondrugs: an effective discrimination with topological methods and artificial neural networks. *Journal of Chemical Information and Computer Sciences, 43(5)*, 1688-1702. https://doi.org/10.1021/ci0302862

Ortakcı, Y. ve Güloğlu, C., (2012). Parçacık sürü optimizasyonu ile küme sayısının belirlenmesi. *Akademik Bilişim Konferansı*, 335-342. Uşak.

Pehlivanlı, A.Ç. and Gümüştaş, E. (2019). *Mutajenisite tahmininde in-silico istatistiksel öğrenme modeli*. Mimar Sinan Güzel Sanatlar Üniviversitesi. Bilimsel Araştırma Projesi, BAP 2018-30.

Pehlivanlı, A.Ç., (2008). *Consensual classification of drug/nondrug compounds for drug design*. Doktora Tezi, Çukurova Üniversitesi Fen Bilimleri Enstitüsü, Adana.

Pehlivanlı, A.Ç., Ersoy, O.K. and Ibrikci, T. (2008). Drug/nondrug classification with consensual Self-Organising Map and Self-Organising Global Ranking algorithms. *International Journal of Computational Biology and Drug Design, 1(4)*, 436. https://doi.org/10.1504/ijcbdd.2008.022212

Pehlivanlı, A.Ç. (2016). A novel feature selection scheme for high-dimensional data sets: four-Staged Feature Selection. *Journal of Applied Statistics, 43(6)*, 1140-1154. https://doi.org/10.1080/02664763.2015.1092112

Qasim, O.S. and Algamal, Z.Y. (2018). Feature selection using particle swarm optimization-based logistic regression model. *Chemometrics and Intelligent Laboratory Systems, 182*, 41-46. https://doi.org/10.1016/j.chemolab.2018.08.016

Rockhold, F. W. (2000). Strategic use of statistical thinking in drug development. *Statistics in Medicine, 19*, 3211–3217. https://doi.org/10.1002/1097-0258(20001215)19:23<3211::aid-sim622>3.0.co;2-f

Sakri, S. B., Abdul Rashid, N. B. and Zain, Z. M. (2018). Particle swarm optimization feature selection for breast cancer recurrence prediction. *IEEE Access, 6*, 29637-29647. https://doi.org/10.1109/ACCESS.2018.2843443

Sokolova, M., Japkowicz, N. and Szpakowicz, S. (2006). Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. *Australasian Joint Conference on Artificial Intelligence, Springer*. 1015-1021. Berlin Heidelberg. https://doi.org/10.1007/11941439_114

Subaş, N. (2019). *Sürekli/İkili parçacık sürü optimizasyonu ve destek vektör makinelerinin hibrit kullanımı ile özellik seçimi*. Yüksek Lisans Tezi, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.

Tretea, I.C. (2003). The Particle swarm optimization algorithms: Convergence analysis and parameter selection. *Information Processing Letters, 85*, 317-325. https://doi.org/10.1016/S0020-0190(02)00447-7

Ünler, A. and Murat, A. (2010). A discrete particle swarm optimization method for feature selection in binary classification problems. *European Journal of Operational Research, 206(3)*, 528–534. https://doi.org/10.1016/j.ejor.2010.02.032

Vapnik, V. (1995). *The nature of statistical learning theory*. New York, Inc. Springer-Verlag

Vashishtha, N. and Vashishtha, J. (2016). Particle swarm optimization-based feature selection. *International Journal of Computer Applications, 146(6)*, 11-17. https://doi.org/10.5120/ijca2016910789

Wagener, M. and Van Geerestein, V. J. (2000). Potential drug and non-drugs: prediction and identification of important structural features. *Journal of Chemical Information and Computer Sciences, 40(2)*, 280-292. https://doi.org/10.1021/ci990266t