

Genetik Algoritma Kullanılarak Deriden Geçiş Özelliklerine Göre Kimyasal Maddelerin Molekül Yapılarının Bulunması

Aysun COŞKUN *, Nursal ARICI**

* Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Bilgisayar Eğitimi Bölümü

**Gazi Üniversitesi Teknik Eğitim Fakültesi Elektronik ve Bilgisayar Eğitimi Bölümü
06500 Teknikokullar, ANKARA

ÖZET

Geçmişte araştırmacılar, problemleri parametreler arasındaki ilişkileri deneme yoluyla çözmeye çalışırken, parametre sayısı artınca çözümsüzlük veya elde edilen çözümü değerlendirememesi problemi ortaya çıkmıştır. Bu tür çok parametrelili problemlerin çözümü için yeni teknikler bulunmaya çalışılmış ve yapay zeka optimizasyon tekniklerinden genetik algoritma uygulamalarda olumlu sonuçlar vermiştir. Eczacılık ve kimya alanlarında optimizasyon problemlerinin çözümünde klasik optimizasyon teknikleri ile geliştirilen programlar yerine yapay zeka optimizasyon teknikleri ile geliştirilen programlar kullanılmaktadır. Böylece molekül tasarımında, molekül tanınmasında, ilaç tasarımında ve protein yapısının tahmini gibi işlemlerde sezgisel yaklaşımları kullanma, doğru analiz yapabilme gibi yöntemlerin geliştirilmesi ile daha hızlı ve doğru çözümler elde edilmektedir.

Yapay zeka optimizasyon tekniklerinden olan genetik algoritma, uygulamalarda sık kullanılmakta ve başarılı sonuçlar vermektedir. Bu çalışmada, kimyasal maddelerin deriden geçiş katsayılarının belirlenmesine yönelik genetik algoritma uygulaması yapılmıştır. Hedeflenen deriden geçiş katsayısına göre kimyasal maddelerin molekül yapıları bulunmuş ve değerlendirilmiştir.

Anahtar Kelimeler: Genetik algoritma, Deriden geçiş katsayısı, Çaprazlama operatörü, Rulet çemberi seçimi

By Using Genetic Algorithm to Find Structure of Molecul of Chemical Materials According to Passing Features of Through Skin

ABSTRACT

Researchers in the past, when working to solve the problems using the manner of trial in the relation between parameters, impassible solution rising parameter number or an evaluation problem of obtained solution has been occurred. In order to solve this kind of problems with many parameters, to fined new techniques have been done and genetic algorithm which is a part of the artificial intelligence optimization techniques has given positive results. To solve the optimization problem in the areas of pharmacology and chemistry, the programs which have been developed by using artificial intelligence optimization techniques instead of the programs which have been developed by using classical optimization techniques have been used. In this way, by developing methods like designing molecule, recognizing molecule, designing medicine and to estimate structure of protein and similar issues using perceptual approach, to be able to make true analysis, quick and true solutions can be obtained.

Genetic algorithm which is a part of the artificial intelligence optimization techniques, which is frequently used in applications and to give successful results. In this study, the application of genetic algorithm has been done to determine the coefficient of passing through skin of chemical materials. According to the passing coefficient of the objective skin, the structure of molecule of chemical materials have been determined and evaluated.

Keywords: Genetic algorithms, Permeability coefficients, Crossover operators, Roulette wheel selection

1.GİRİŞ

Karmaşık optimizasyon problemlerinin çözümünde yapay zeka optimizasyon algoritmalarının kullanımını artmaktadır. Bu optimizasyon algoritmalarının geliştirilmesi çok yeni olsa da problem çözümlerinde başarılı olduklarını görmek mümkündür. Yapay zeka optimizasyon algoritmalarının veya modern sezgisel algoritmalarının başlıcaları genetik algoritma (GA), tabu

araştırmaları, tavlama benzetimidir. Tabu araştırmaları ve tavlama benzetim yöntemlerinin temeli bölgesel ve sezgisel komşuluk araştırmalarına dayanmaktadır. Ancak genetik algoritma biyoloji biliminin prensiplerini temel almaktadır.

GA, genetik biliminin temelleri üzerine kurulan bir araştırma tekniği olarak ortaya atılmıştır (1). Klasik bir tanımlamayla biyolojik bir sistemin çevresine uyu-

munda kullandığı yöntemin modellemesidir. GA bireylerden oluşan bir popülasyon üzerinde işlemektedir.

Popülasyon problemin olası çözümlerini temsil eden bireylerden oluşmaktadır. Problemin olası çözümlerinden oluşan başlangıç popülasyonu ile başlanarak bundan daha iyi çözümler oluşturulmasına çalışılır. Popülasyon, zaman içerisinde değişir, ancak aynı sayıda elemana sahiptir (2). GA'nın genel yapısı içerisinde popülasyon içerisindeki her çözüm diğer çözümler ile karşılaştırıldığında ne kadar iyi bir çözüm olduğunu gösteren bir uyum değeri atanır. Uyum değeri ne kadar büyükse bireyin veya organizmanın, sonraki nesilde yaşama şansı ve özelliklerini sonraki nesillere taşıma şansı o kadar fazla olur (3).

GA, doğal olayların modellenmesi üzerine kurulu stokastik ve iteratif (yönlendirilmiş) bir algoritmadır (4). GA global optimum çözümü bulmayı garanti etmez ancak kabul edilebilir hızda, kabul edilebilir ölçüde iyi çözümler bulunmasında genel olarak başarılıdır. Karmaşık problemlerin çözümünde klasik optimizasyon teknikleri hem sonucun doğruluğu hem de zaman açısından GA'ya göre daha iyi çözümlemesi olasıdır. Ancak GA'nın esas alanı, bu tekniklerin uygulanmadığı alanlardır. Bazı problemler, GA ile klasik optimizasyon tekniklerinin birlikte kullanılmasıyla uygun çözümler bulunabildiğini göstermiştir. Eczacılık alanı da hem klasik optimizasyon tekniklerinin hem de GA'nın birlikte uygulanabildiği karmaşık problemlerin bulunduğu bir alandır. Bu alanda literatürde GA ile yapılmış çalışmaların bazıları ise, Illgen ve arkadaşları (5) senteze gerek duymadan büyük sanal molekül kütüphanelerinden, yüksek biyolojik aktiviteye sahip moleküllerin bulunmasına yönelik ve sanal kütüphaneye dahil tüm molekül yapılarının biyolojik faaliyetlerini belirleme problemlerine genetik algoritma uygulamışlardır. Hou ve Xu (6) ise, GA tekniğini kullanarak kişisel bilgisayarlarda işleyen kapsamlı bir molekül simülasyon program paketi oluşturmuşlardır. İki ve üç boyutlu nicel yapı-faaliyet ilişkileri, molekül docking ve veri tabanı kullanarak GA uygulamalı bilgisayar destekli ilaç tasarımı geliştirmişlerdir. Yine Kian-Goh ve Foster (7) GA'yı iki uygulamada kullanmışlardır. İlkinde, kendiliğinden birleşen materyaller için molekül tasarımı yapmaya ve ikincisinde ise mevcut materyaller içinde en uygun molekülün bulunmasına çalışılmıştır. Keser ve Stop (8), genel tanımlanmış kimya yapı bloklarından, kimyasal maddenin molekül yapılarının bulunması probleminde GA kullanmışlardır. Jones (9) kimyadaki bazı kısıtlardan yola çıkarak GA'nın bu tür kısıtlı problemlere uygulanabilirliğini tartışmıştır.

Literatürde, kimya ve eczacılıkta kabul edilen moleküler formüller (sembolik kimyasal grup alfabeti) çerçevesinde uygulamalar yapılmış ve mevcut molekül kütüphanelerinden amaca uygun molekül yapılarının bulunması hedeflenmiştir. Bu çalışmada, bazı kısıtlar kullanılmakla beraber kimyasal maddelerin molekül yapılarının tahmini deriden geçiş katsayılarına göre belirlenmektedir. Uygulamada iki jenerasyonla GA'nın işle-

yişi adım adım gösterilmektedir. Jenerasyon sayılarının arttırılması mümkün olmaktadır ancak işleyişin açıklanması için iki jenerasyon yeterli görülmektedir.

Çalışmanın 2. Bölümünde uygulamada bulunan kimyasal maddelerin deriden geçiş katsayılarının belirlenmesinin önemi verilmektedir. 3. Bölümde sistem ve problemin tanımı ve genetik algoritmanın probleme uygulanması verilirken son bölümde ise iki jenerasyonda bulunan sonuçlar değerlendirilmektedir.

2. KİMYASAL MADDELERİN DERİDEN (TRANSDERMAL YOLDAN) GEÇİŞ ÖZELLİKLERİNİN BELİRLENMESİNİN ÖNEMİ

Farmasötik, kozmetik ve tarımsal kimya endüstrilerinde moleküllerin deriden geçiş katsayılarının hesaplayabilme ihtiyacı genelde hissedilmektedir. Bu, hem klinik etkinliğin değerlendirmesi hem de toksikolojik risk açısından önem taşır. Lokal davranış göstermek ya da deriye uygulanmak üzere spesifik ilaçlar tasarlandığında, farklı kimyasal yapıların deriden geçişlerinin ne ölçüde ve nasıl olacağını bilmesi yararlı olur. Örnek olarak UV filtresi gibi kozmetik ürünlerin tasarlanmasında ürünün faydasının maksimuma çıkarabilmek için deriden geçiş özelliğinin bilinmesi yararlı olur. Aynı şekilde yeni pestisidlerin tasarlanmasında da aktif olan ama deri tarafından iyice emilmeyen moleküllere ihtiyaç vardır (10). Deri, vücut sıvılarını sistemimiz içinde ve zararlı maddeleri de sistem dışında tutarak dış çevreye karşı bir duvar oluşturur. Bu yüzden eczacılık, kozmetik ve tarımsal kimya alanlarında kimyasal maddelerin deriden geçiş oranlarının tahmin edilmesi önemli bir tartışma konusudur.

Transdermal yoldan bir etkin maddenin uygulamasının bazı avantajları vardır. Bu sebepten dolayı etkin maddelerin deri yoluyla uygulanmasına ilgi gün geçtikçe artmaktadır. Bilinen alışılmış etkin maddelerin dışında bazı kimyasal maddelerin deriden geçiş uygulamalarına rastlanmaktadır.

Kimyasal maddelerin transdermal yolla uygulanmasına yönelik birçok çalışma yapılmıştır. Özellikle derinin düşük geçirgenliğe sahip olma özelliği, uygulanan kimyasal maddenin alerjik reaksiyonu olup olmaması, ilacın moleküler yapısının deriden geçiş oranına etkisinin belirlenmesi bu araştırmaların temelini oluşturmaktadır.

Son yüzyılda endüstriyel, zirai ve evde kullanılan kimyasal maddelerin kullanımı kişileri, çevreyi ve ekosistemleri ciddi şekilde tehdit etmektedir. İnsanların da bu tehlikeye oldukça sık maruz kalması dikkati çekmiştir ve transdermal yolla bulaşma riskleri araştırılmıştır. Genel olarak insan sağlığını tehlikeye atacak kimyasal maddelerin etkilerinin sınıflandırılmasında ürünün kullanımını açısından önceliği, deri ile temas halinde alınacak toksik maddelerin etkileri oluşmaktadır. Bu etkiler, alerjik, kanserojen, tahriş şeklinde tanımlanabilmektedir. Bu sebeptendir ki, herhangi bir kimyasalın

veya kimyasalların ürün olarak kullanıldığında deriden geçiş özelliğine göre sayılan bu etkilerin minimize edilmesi amaçlanmaktadır.

3. GENETİK ALGORİTMA KULLANARAK DERİDEN EN İYİ GEÇEN KİMYASAL MADDENİN MOLEKÜL YAPISININ BULUNMASI

Bu çalışmada amaç deriden en iyi geçen kimyasal maddenin molekül yapısının GA kullanarak

- Moleküldeki Oksijen(O) atomu sayısı
 - Moleküldeki Halojen atomu sayısı (I,Br,F,Cl)
 - Moleküldeki Kükürt(S) atomu sayısı olarak belirlenmiştir. Tablo 1. de veri olarak alınan rasgele seçilen 10 kimyasal madde ve molekül parametreleri gösterilmektedir.

Verilen kriterlere göre problem modelinin oluşturulması ve GA işleyişinin sağlanması için başlangıç popülasyonundaki her kromozomun

Tablo 1. Rasgele seçilen 10 kimyasal madde ve molekül parametreleri

	Molekül	Aromatik Halka Sayısı	H Sayısı	OH Sayısı	N Sayısı	S Sayısı	C Sayısı	Aromatik Olmayan Halka Sayısı	O Sayısı	Halojen Atom Sayısı(Cl,F,I,Br)
1	Aldosterone -C ₂₁ H ₂₈ O ₅	0	28	2	0	0	21	5	5	0
2	Amobarbital -C ₁₁ H ₁₈ N ₂ O ₃	0	18	0	2	0	11	1	3	0
3	Atropine -C ₁₇ H ₂₃ NO ₃	1	23	1	1	0	17	2	3	0
4	Barbital -C ₈ H ₁₂ N ₂ O ₃	0	12	0	2	0	8	1	3	0
5	Benzylalcohol - C ₇ H ₈ O	1	8	1	0	0	7	0	1	0
6	2,4,6-Tribromophenol -C ₆ H ₃ Br ₃ O	1	3	1	0	0	6	0	1	3
7	Butamben - C ₁₁ H ₁₅ NO ₂	1	15	0	1	0	11	0	2	0
8	Butanedioic Acid - C ₄ H ₆ O ₄	0	6	2	0	0	4	0	4	0
9	Butanolide - C ₄ H ₈ O ₂	0	6	0	0	0	4	1	2	0
10	2,3-Butanolone -C ₄ H ₈ O ₂	0	8	1	0	0	4	0	2	0

bulunmasıdır. Uygulamanın adım adım işleyişinin gösterimi açısından rasgele seçilen 10 kimyasal maddenin molekül parametrelerine göre deriden geçiş katsayıları hesaplanmıştır. Kimyasal maddelerin molekül yapısını oluşturan parametreler;

uygunluk değerlerinin belirlenmesi gerekmektedir. Uygunluk fonksiyonunun bulunmasına yönelik, rasgele seçilen 100 kimyasal maddenin molekül yapılarından, SPSS 11.0 ile deriden geçiş katsayısı (LogKp) üzerinde

Tablo 2. Log K_p 'nin bulunmasına yönelik stepwise çoklu regresyon analizi sonuçları

Değişken	B	SH _B	β	t	P	İkili r	Kısmi r
Sabit	,316	,279		1,130	,261		
Aromatik halka sayısı	,342	,210	,133	1,627	,107	,437	,166
Nsayısı	-,669	,100	-,437	-6,679	,000	-,221	-,569
Csayısı	,247	,034	,848	7,261	,000	,380	,601
OHsayısı	-,517	,098	-,351	-5,287	,000	-,189	-,481
Halojen atom Sayısı	,570	,109	,340	5,230	,000	,376	,477
Aromatik olmayan halka sayısı	-,449	,152	-,306	-2,949	,004	,063	-,292
R= 0.807, R ² =0.652							
F= 28.988, P=0.000							

- Moleküldeki Hidrojen(H) atomu sayısı
- Moleküldeki Aromatik halka sayısı
- Moleküldeki OH sayısı
- Moleküldeki Azot(N) atomu sayısı
- Moleküldeki Karbon(C) atomu sayısı
- Moleküldeki Aromatik olmayan halka sayısı

belirlenen molekül parametrelerinin etkisi incelenmiş ve uygulama için seçilen 9 parametreden en fazla etkili olan 6'sı deriden geçiş katsayısının % 65'ini açıkladığı görülmüştür. Tablo 2 de analiz sonuçlarına yer verilmektedir. Regresyon analizi sonucuna göre LogKp'ye ilişkin matematiksel model şu şekildedir.

$$\text{LOG } K_p = ,316\text{SABİT} + ,342\text{AROHALS} - ,669\text{N SAYISI} + ,247\text{CSAYISI} - ,517\text{OHSAYISI} + ,570\text{HALATS} - ,449\text{NONAS}$$

Bu çalışmada uygunluk değerleri tek bir regresyon denklemi şeklinde elde edilmeye çalışılmıştır. Probleme aranan çözüm deriden en iyi geçen kimyasalın bulunması ve molekül yapısının tespitidir. Bu amaçla belirlenen denklem ile uygunluk değeri hesaplanabilmektedir.

Deriden en iyi geçen kimyasal maddenin GA ile bulunmasında bazı kabuller yapılmıştır.

- Deriden en iyi geçişi bulmada kimyasal maddelerin geçiş katsayıları(LogK_p) değerlendirilmektedir.
- Kimyasal maddenin molekül yapısının belirlenmesinde dokuz parametre ele alınmaktadır.
- Molekül yapısının belirlenmesinde kullanılacak parametreler LogK_p hesaplanmasında da kullanılacaktır.
- Kullanılan kimyasal maddelerin miktarı veya oranı çalışmada dikkate alınmamaktadır. Kimyasal maddenin miktarı/oranı analizden bağımsızdır.
- GA uygulamasında seçim yöntemi olarak rulet çemberi seçim yöntemi kullanılmıştır.
- Genetik operatörlerden çaprazlama operatörü iki nokta çaprazlama operatörü olarak kullanılmıştır.
- Deriden geçiş katsayısının maksimum olması beklenmektedir.
- GA uygulamasında istenen kimyasal maddenin molekül yapısının belirlenmesinde H sayısı, O sayısı ve S sayısı esnek bırakılmıştır.

Adım 1. Çözümlerin bir başlangıç popülasyonunu oluştur.

Adım 2. Populasyondaki her çözümün uygunluk değerini hesapla.

Adım 3. Durdurma kriteri sağlanıyorsa araştırmayı durdur.

Yoksa, aşağıdaki adımları gerçekleştir.

3.1. Tabi seçim işlemini uygula (uygunluk değerleri daha yüksek olan çözümler yeni popülasyonda daha fazla temsil edilirler).

3.2. Çaprazlam işlemi uygula (Mevcut iki çözümden yeni iki yapı üretir).

3.3. Mutasyon işlemi uygula (Çözümlerde rasgele değişim meydana getirilir).

Adım 4. Adım 2 ye git

GA'nın başarılı olabilmesi için yapılması gereken ilk iş, başlangıç popülasyonunu oluşturan kromozomların kodlanmasıdır. Kromozomlar genlerle ifade edilir ve her bir gen problemin bağımsız parametrelerini temsil eder. Kromozomlar, belirli bir çözüm hakkında tüm bilgileri taşır. Kodlamanın çeşitli yolları vardır. En yaygın olarak kullanılan kodlama ikili kodlamadır. Basit bir GA, ikili dizilerde oluşan bir popülasyon üzerinde çalışır. Burada kromozomlar, 0' lar ve 1' lerden oluşan dizilerdir. Kromozomların kodlanmasında tamsayı dizisi ya da kurallar dizisi olarak çeşitli gösterim biçimleri vardır.

GA'nın uygulanmasında ilk adım, kromozomların bir başlangıç popülasyonunun oluşturulmasıdır. Başlangıç popülasyonu çoğunlukla rasgele olarak oluşturulur. Ancak, özellikle kısıtlı optimizasyon problemlerinde, rasgelelik uygun olmayan çözümlere neden

Tablo 3. Kromozomlardan oluşan popülasyonun temel yapısı

	1.Gen	2.Gen	3.Gen	4.Gen	5.Gen	6.Gen	7.Gen	8.Gen	9.Gen
1.Kromozom	0	28	2	0	0	21	5	5	0
2.Kromozom	0	18	0	2	0	11	1	3	0
3.Kromozom	1	23	1	1	0	17	2	3	0
4.Kromozom	0	12	0	2	0	8	1	3	0
5.Kromozom	1	8	1	0	0	7	0	1	0
6.Kromozom	1	3	1	0	0	6	0	1	3
7.Kromozom	1	15	0	1	0	11	0	2	0
8.Kromozom	0	6	2	0	0	4	0	4	0
9.Kromozom	0	6	0	0	0	4	1	2	0
10.Kromozom	0	8	1	0	0	4	0	2	0

- Kromozomların oluşturulmasında genlerin temsili 10 tabanına göre yapılmıştır.

3.1 Genetik Algoritma'nın Probleme Uygulanması

Bir problemin GA ile çözümünde takip edilecek işlem adımları aşağıda verilmiştir (11).

olabilir. Bu durumdan kaçınmak için genellikle incelenen probleme özgü sezgisel yöntemlerden yararlanır.

Kullanılan her bir kromozom, veri olarak seçilen kimyasal maddenin molekül yapısından elde edilmiştir. Kimyasal maddenin molekül yapısını belirleyen parametrelere bağlı olarak kromozom uzunluğu 9 dur. Bunun anlamı kromozom 9 genden oluşmaktadır ve bu çalışmada 10 kromozom göz önüne alınmıştır. Bu da

popülasyon genişliğini vermektedir. Kromozom üzerindeki her bir gen kimyasal maddenin molekül parametrelerini temsil eder.

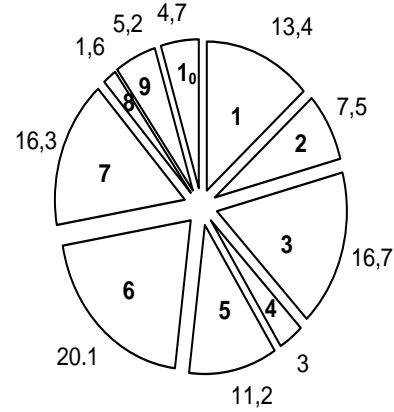
Çözümlerin değerlendirilmesinde kullanılan uygunluk değeri GA'nın performansında oldukça etkilidir. Popülasyondaki bireylerin uygunluk değerlerinin bulunması bir sonraki neslin oluşturulmasında popülasyon içerisindeki hangi kromozomların seçileceğinin belirlenmesinde önemli olmaktadır. Çalışmada her bir kromozomun değerlendirilmesi için kullanılan uygunluk değeri LogK_p ile gösterilen deriden geçiş katsayısıdır.

Teoriye göre iyi olan bireyler yaşamını sürdürmeli ve bu bireylerden yeni bireyler oluşmalıdır. Bu seçim çeşitli kriterlere göre yapılabilir. Rulet çemberi seçimi, elitizm, turnuva seçimi, sıralı seçim bunlardan bazılarıdır. Rulet seçiminde kromozomlar uygunluk fonksiyonuna göre bir rulet etrafına gruplanır. Uygunluk fonksiyonu herhangi bir kritere uyan bireylerin seçilmesi için kullanılır. Bu rulet üzerinden rasgele bir birey seçilir. Daha büyük alana sahip bireyin seçilme şansı daha fazla olacaktır. Rulet seçimi eğer uygunluk çok fazla değişiyorsa sorun çıkartabilir. Bu çalışmada molekül parametrelerine göre uygunluk çok fazla değişiklik göstermediğinden rulet çemberi seçim yöntemi kullanılmıştır.

Tablo 4. Rasgele oluşturulmuş başlangıç popülasyonu ve rulet çemberindeki payı

Bireyler (i)	Kromozomlar	Uygunluk Değerleri (LogK_p)	Toplam (%)
1	0 28 2 0 0 21 5 5 0	2.224	13,4
2	0 18 0 2 0 11 1 3 0	1.246	7,5
3	1 23 1 1 0 17 2 3 0	2.773	16,7
4	0 12 0 2 0 8 1 3 0	0.505	3,0
5	1 8 1 0 0 7 0 1 0	1.870	11,2
6	1 3 1 0 0 6 0 1 3	3.333	20,1
7	1 15 0 1 0 11 0 2 0	2.706	16,3
8	0 6 2 0 0 4 0 4 0	0.270	1,6
9	0 6 0 0 0 4 1 2 0	0.855	5,2
10	0 8 1 0 0 4 0 2 0	0.787	4,7

Başlangıç popülasyonundaki bireylerin uygunluk değerlerine göre üreme işleminde kullanılacak rulet çemberi Şekil 1. de gösterilmiştir.



Şekil 1. Seçim işlemi için kullanılacak rulet çemberi

Gelecek nesilde, rulet tekerleğindeki paylara göre kromozomların seçilme olasılıkları

$$P_i = \frac{f_i}{\sum f_i} \text{ eşitliğinden hesaplanmaktadır.}$$

Seçilme olasılıklarına göre kromozomların bir sonraki nesil'e aktarılması Tablo 5' de gösterilmektedir. Bu çizelgeden de görüleceği gibi bir sonraki nesil'e 3., 6. ve 7. kromozomların 2 kopyası, 1., 2., 5. ve 9. kromozomların birer kopyası alınmıştır. 4., 8. ve 10. kromozomlar bir sonraki nesile aktarılmamıştır.

GA'nın bir sonraki adımı çaprazlama operatörünün uygulanmasıdır. Bu operatör iki adımda gerçekleştirilir. Öncelikle popülasyondaki yeni bireyler rasgele eşleştirilir. Daha sonra eşlenen bireyler karşılıklı gen değişimine uğrar. Tablo 6'da 1. ile 3., 2. ile 3., 5. ile 6., 6. ile 7. ve 7. ile 9. kromozomlar eşleştirilmiştir. Ayrıca çaprazlama pozisyonları gösterilmiştir.

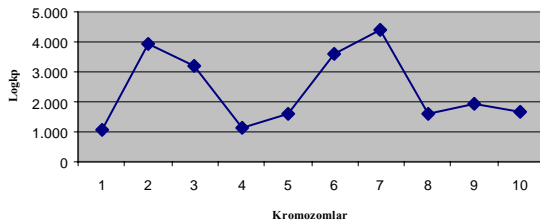
Tablo 5. Genetik algoritma ile üreme işlemi

Bireyler (i)	Kromozomlar	Uygunluk Değerleri (LogK_p)	Seçim Değeri	Beklenen	Kopya Sayısı
			$P_i = \frac{f_i}{\sum f_i}$	$\frac{f_i}{f_{ort}}$	
1	0 28 2 0 0 21 5 5 0	2.224	0.134	1.340	1
2	0 18 0 2 0 11 1 3 0	1.246	0.075	0.752	1
3	1 23 1 1 0 17 2 3 0	2.773	0.167	1.673	2
4	0 12 0 2 0 8 1 3 0	0.505	0.030	0.304	0
5	1 8 1 0 0 7 0 1 0	1.870	0.112	1.128	1
6	1 3 1 0 0 6 0 1 3	3.333	0.201	2.011	2
7	1 15 0 1 0 11 0 2 0	2.706	0.163	1.633	2
8	0 6 2 0 0 4 0 4 0	0.270	0.016	0.162	0
9	0 6 0 0 0 4 1 2 0	0.855	0.051	0.516	1
10	0 8 1 0 0 4 0 2 0	0.787	0.047	0.471	0
Toplam		16.569	1.000		
Ort.		1.6569	0.100		
Max.		3.333	0.201		

Tablo 6. Genetik algoritma ile çaprazlama işlemi

Çaprazlama Öncesi	Çaprazlama Sonrası	Uygunluk Değerleri (LogK _p)	Seçim Değeri $P_i = \frac{f_i}{\sum f_i}$	Beklenen $\frac{f_i}{f_{ort}}$	Kopya Sayısı
0 28/ 2 0 0 21/5 5 0	0 28 1 1 0 17 5 5 0	1.084	0.044	0.44	0
0 18/ 0 2 0 11/1 3 0	1 23 2 0 0 21 2 3 0	3.913	0.159	1.59	2
1 23/ 1 1 0 17/2 3 0	1 18 1 1 0 17 1 3 0	3.222	0.131	1.31	1
0 12/ 0 2 0 8/1 3 0	1 23 0 2 0 11 2 3 0	1.139	0.046	0.46	0
1 8/ 1 0 0 7/0 1 0	1 8 1 0 0 6 0 1 0	1.623	0.066	0.66	1
1 3 / 1 0 0 6/0 1 3	1 3 1 0 0 7 0 1 3	3.580	0.145	1.45	1
1 15/ 0 1 0 11/0 2 0	1 3 0 1 0 11 0 1 3	4.416	0.179	1.79	2
0 6 / 2 0 0 4/0 4 0	1 15 1 0 0 6 0 2 0	1.623	0.066	0.66	1
0 6/ 0 0 0 4/1 2 0	0 6 0 1 0 11 1 2 0	1.915	0.079	0.79	1
0 8/ 1 0 0 4/0 2 0	1 15 0 0 0 4 0 2 0	1.649	0.067	0.67	1
Toplam		24.541			
Ort		2.4541			
Max		4.416			

Yeni neslin oluşturulmasında görüldüğü gibi deriden geçiş katsayısının maksimum değeri değişmiştir ve 4.416 olarak bulunmuştur. Ayrıca toplam ve ortalama değerde de bir artış olmuştur. Şekil 2’de ilk jenerasyon için rasgele seçilen 10 kimyasal maddenin maksimum deriden geçiş katsayısı görülmektedir.



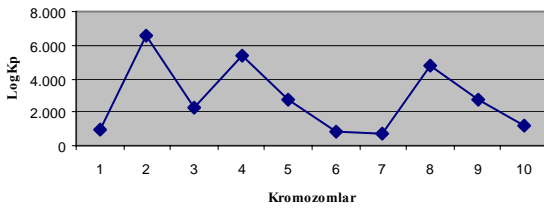
Şekil 2. İlk jenerasyon için rasgele seçilen 10 kimyasal maddenin maksimum deriden geçiş katsayısı

İkinci jenerasyonu gerçekleştirdiğimizde sadece 2. ve 7. kromozomlardan iki kopya alınacak, 1. ve 4. kromozomlardan kopya alınmayacak diğer kromozomlardan ise birer kopya alınacaktır. 2. ile 6., 3. ile 7., 5. ile 10., 8. ile 2., 7. ile 9. eşleştirilmiş ve çaprazlama operatörü uygulanmıştır. Tablo 7’de ikinci jenerasyona ait çaprazlama ve çaprazlamadan sonra oluşan yeni bireylerin deriden geçiş değerleri gösterilmektedir.

Tablo 7. Genetik algoritma ile çaprazlama işlemi

Çaprazlama Öncesi	Çaprazlama Sonrası	Uygunluk Değerleri (LogK _p)
0 28 1 1 0 17 5 5 0	1 23 1 0 0 7 2 3 0	0.972
1 23 2 0 0 21 2 3 0	1 3 2 0 0 21 0 1 3	6.521
1 18 1 1 0 17 1 3 0	1 18 0 1 0 11 1 3 0	2.257
1 23 0 2 0 11 2 3 0	1 3 1 1 0 17 0 1 3	5.381
1 8 1 0 0 6 0 1 0	1 8 0 1 0 11 0 1 0	2.706
1 3 1 0 0 7 0 1 3	0 6 1 0 0 6 1 2 0	0.832
1 3 0 1 0 11 0 1 3	1 23 1 0 0 6 2 3 0	0.725
1 15 1 0 0 6 0 2 0	1 5 2 0 0 21 0 2 0	4.811
0 6 0 1 0 11 1 2 0	1 3 0 1 0 11 0 1 3	2.706
1 15 0 0 0 4 0 2 0	1 15 0 0 0 4 1 2 0	1.197
Toplam		28.108
Ort		2.8108
Max		6.521

Son GA operatörü, mutasyon operatörü olup gen değişimi işlemi ile uygulanır. Bu çalışmada mutasyon olasılığının 0.001 olduğu varsayılmıştır. Her biri dokuz genden oluşan 10 kromozom için $9 \cdot 10 \cdot 0.001 = 0.09$ genin değişime uğrayacağı beklenir. Bu olasılık değeri için bu adımda hiçbir gen değişime uğramayacaktır. Oysa ki, P_m mutasyon olasılığı 0.05 olduğunu varsayılırsa $9 \cdot 10 \cdot 0.05 = 4.5$ genin değişime uğrayacağı belirlenir.



Şekil 3. İkinci jenerasyon için rasgele seçilen 10 kimyasal maddenin maksimum deriden geçiş katsayısı

Bu çalışma için sonuçlar incelenirse; maksimum ve ortalama değer yeni popülasyonda artmıştır. Sadece iki nesil işlemden popülasyonun ortalama değeri 2.4541'den 2.8108'e yükselmiştir. Maksimum uygunluk değeri 4.416'dan 6.521'e artmıştır. Maksimum deriden geçiş değeri 6.521 olan kimyasal maddenin kromozom yapısı incelendiğinde, bu kimyasal maddenin molekül yapısı $C_{21}H_3O$ 'dur. Molekül parametrelerine göre 1 aromatik halka ile bağlanan, 3 halojen atomu ve OH sayısı 2 olan kimyasal madde belirlenmiştir. N sayısı ise bulunduğumuz kimyasal yapıda yer almamaktadır. Belirlenen molekül parametrelerine göre kimyasal maddenin molekül özellikleri Tablo 8'de gösterilmektedir.

Yapılan kabullerde kimyasal yapıların deriden geçiş katsayılarını hesaplarırken H, O ve S sayısını regresyon denkleminde dikkate almadığımızdan bu üç parametreyi belirlediğimiz kimyasal yapı formülizasyonunda esnek bırakmak, yapının tahmini için kolaylık sağlamaktadır. Bu durumda bulunan kimyasal yapıyı $C_{21}H_xO_y$ olarak göstermek mümkündür.

Tablo 8. Uygulamada bulunan maksimum LogK_p 'ye sahip kimyasal maddenin molekül özellikleri

$C_{21}H_xO_y$	
Aromatik Halka 1	Aromatik Hidrokarbon özelliği var
OH 2	Asit ve türevi
N 0	Azot özelliği yok
Aromatik Olmayan Halka 0	Heterosiklik bileşik özelliği yok
Halojen Atom 3	Organik halojen

4. SONUÇ

Yapılan çalışmada, kimyasal maddelerin deriden geçiş katsayılarının hesaplanması veya istenen deriden geçiş özelliğine bağlı olarak kimyasal maddelerin belir-

lenmesi problemine GA uygulanmıştır. Rasgele seçilen birbirinden farklı 10 kimyasal maddenin molekül parametrelerinden oluşan kromozomlara GA operatörlerinden çaprazlama, iki noktada uygulanmıştır. Geliştirilen iki jenerasyonda, başlangıç popülasyonundan farklı olarak çeşitli kimyasal maddeler belirlenmiştir.

5. KAYNAKLAR

- Murty, K.G., Operations Research Deterministic Optimal Models, Prentice Hall, USA, 581p, 1995.
- Srinivas, M., Patnaik, L.M., Genetic Algorithms: A Survey, Computer, Vol:26,27, 1994.
- Michalewicz, Z., Genetic Algorithms +Data Structures = Evolutions Programs, Springer-Verlag, Berlin, 340p., 1994.
- Holland, J. H., Adaptation in Naturel and Artificial System an Introductory Analysis with Applications to Biology, Control and Artificial Intelligence, Second Edition, Cambridge, Mass:MIT Pres, 211p., 1992
- Illgen, K., and Others, Simulated Molekuler Evolution in A Full Combinatorial Library, Chemistry&Biology, Vol7, 443p, 2000.
- Hou, T., and Xu, X., A New Molecular Simulation Software Package--Peking University Drug Design System (PKUDDS) for Structure-Based Drug Design, Medline, vol:19, 455p, 2001.
- Kian, J., Goh, M., and Foster, A., Evolution Molecules for Drug Disang Using Genetic Algorithms, Proc. Int. Conf. on Genetic and Evol. Comp. (GECCO), Morgan Kaufmann, 33p, 2000.
- Keser, M. and Stupp, S.I. A Genetic Algorithm for Conformational Search of Organic Molecules: Implications for Materials Chemistry. *Computers and Chemistry*, Vol: 22, 345p, 1998
- Jones, G., Genetic and Evolutionary, Encyclopedia of Computational Chemistry. Wiley, Chichester, 1998
- Pugh, J.W., Hadgraft, J., Ab Initio Prediction Of Human Skin Permability Coefficients, International Journal of Pharmaceutics, IJP03378, 1-16, 1993
- Karaboğa, D., Yapay Zeka Optimizasyon Algoritmaları, Atlas Yayınevi, Yayın No:38, İstanbul, 2004