Nanopartikül biyosentezinde ortalama çap tahmini ve optimizasyonu için yanıt yüzeyi yöntemi ve makine öğreniminin birleştirilmesi

Combining response surface methodology and machine learning for harmonic mean diameter prediction and optimization in the nanoparticle biosynthesis

Gülizar Çalışkan Bilgin[1], Ayça Kumluca Topallı[2], Tuğçe Mutaf Kılıç[3], Murat Elibol[4]

[1] İzmir Ekonomi Üniversitesi, Genetik ve Biyomühendislik Bölümü, İzmir
[2] İzmir Ekonomi Üniversitesi, Elektrik ve Elektronik Mühendisliği Bölümü, İzmir
[3] Manisa Celal Bayar Üniversitesi, Biyomühendislik Bölümü, Manisa
[4] Ege Üniversitesi, Biyomühendislik Bölümü, İzmir

Özetçe—Nanopartiküllerin yeşil sentez yöntemi ile biyolojik kaynaklardan sentezlenmesi ve üretim optimizasyonu çalışmaları günümüzde popülerliğini arttırmaktadır. Ancak bu gibi süreçlerde biyolojik kaynak ve çevresel etkilerin değişkenliği son üründe farklı morfoloji ve fonksiyonelliğe yol açmaktadır. Bu çalışmada nanopartikül sentezinde mikroalgler biyoindirgeyici ajan olarak kullanılmış olup parçacık sentezinde $FeSO_4$ konsantrasyonunun ve mikroalg ortamı ile oranının ortalama harmonik partikül çapının analizleri gerçekleştirilmiştir. İki aşamalı olarak yürütülen bu çalışmada önce deney tasarımı yapılmış, veri üretimi ile elde edilen partikül çapları makine öğrenmesi ile geliştirilmiştir. Her iki aşamadaki hata oranları karşılaştırılarak iyileştirmeler kaydedilmiştir. Sonuç olarak veri yetersizliği sorununu çözmek için düşük maliyetli, hızlı, basit ve çevre dostu yeni bir yaklaşım getirilmiş ve partikül çapı tahmininde kullanılmıştır. Elde edilen sonuçlar, önerilen birleşik stratejinin, tek başına istatistiksel yaklaşımdan daha iyi nanoparçacık boyutu tahminleri sağladığını göstermiştir. Önerilen yöntem, önemli ileri bilgi birikimine sahip çok çeşitli biyoteknoloji ve biyomühendislik uygulamalarına uygulanabilecek niteliktedir.

Anahtar Kelimeler—*nanopartikül sentezi, veri üretimi, yanıt yüzey yöntemi, makine öğrenmesi.*

Abstract—The synthesis of nanoparticles from biological sources by green synthesis method and production optimization studies are increasing in popularity today. However, the variability of biological source and environmental effects in such processes leads to different morphology and functionality in the final product. In this study, microalgae was used as a bioreduction agent in nanoparticle synthesis and analyses of the harmonic mean particle diameter of $FeSO_4$ concentration and its ratio with microalgae medium were carried out in particle synthesis. In this two-stage study, the experimental design was carried out first, and the particle diameters obtained by data generation were developed by machine learning. The error rates at both stages were compared and improvements were recorded. As a result, a new low-cost, fast, simple and environmentally friendly approach was introduced to solve the data insufficiency problem and used in particle diameter estimation. The results obtained showed that the proposed combined strategy provides better nanoparticle size estimates than the statistical approach alone. The proposed method is applicable to a wide range of biotechnology and bioengineering applications with significant advanced knowledge.

Keywords—*nanoparticle synthesis, data generation, response surface methodology, machine learning.*

I. GİRİŞ

Nanoteknoloji; biyoteknoloji, doku mühendisliği, enerji, tıp, ilaç endüstrisi, tarımsal uygulamalar, elektronik ve uzay endüstrisi gibi bilimsel alanlarda popülerliğini artıran ve dikkat çeken önemli bir bilim dalıdır. Bu teknoloji ile nano boyutlarda sentezlenen bileşikler yüksek yüzey/hacim oranı, kuantum boyutu etkileri, gelişmiş manyetik özellikler, biyouyumluluk, kimyasal stabilite, biyolojik olarak parçalanabilirlik, düşük toksisite, yapısal çeşitlilik, yüzey atomlarının benzersiz özellikleri ve etkili yüzey etkileşimleri gibi avantajlara sahiptir [1-2].

Günümüzde metal nanopartiküllerin biyouyumlu uygulama alanları, nanopartikül sentezinde organik çözücü, indirgeyici madde, stabilizatör gibi toksik kimyasalların kullanımı gerektirmeyen, çevre dostu yeşil sentez yaklaşımını popüler kılmıştır. Yeşil sentez yaklaşımında partikül boyutunu ve fonksiyonelliğini etkileyen sıcaklık, pH, metal konsantrasyonu, kullanılan biyolojik kaynağın çeşidi ve oranı gibi çeşitli faktörler bulunmaktadır. Her faktörün etkinliğini belirlemek için çok sayıda deney düzeneğinin kurulması öngörülmektedir. Ancak herhangi bir istatistiksel tasarım yöntemi ile nesnel

979-8-3315-2981-9/24/$31.00 ©2024 IEEE

sonuçlar çıkarmak, partikül sentezin sürdürülebilirliğini sağlamak için sonuç analizleri yapmak, planlanan deneylerden daha az sayıda ve daha uygun maliyetli deney tasarlamak mümkündür [1-6]. Yanıt Yüzey Yöntemi (RSM) doğrusal, kare polinom fonksiyonları ve diğerleri gibi çeşitli istatistiksel ve matematiksel tekniklere dayalı optimizasyon için kullanılan deneysel tasarım yöntemlerinden biridir. RSM içinde yer alan Merkezi Kompozit Tasarımı (CCD), dış noktalar ve deneysel bölgenin merkezinde bir nokta kullanan beş seviyeli kesirli faktöriyel tasarım olarak tanımlanır. Nanopartikül sentezinde CCD ile daha az sayıda deney uygulanarak, istatistiksel yaklaşımlarla bireysel faktörlerin önemi ve her bir faktörün yanıt üzerindeki etkisi ortaya çıkarılabilmektedir [7]. Ancak CCD'de merkezi noktanın, ikinci dereceden polinomlara uyması için döndürülebilirlik veya ortogonallik gibi çeşitli özelliklere sahip olması beklenir. Ayrıca tanımlanan sınır değerler dışındaki noktalarda modelin yalnızca ikinci dereceden polinomlarla uydurulmasına izin veren kısıtlamalar ve en az iki deneysel faktör kullanma ihtiyacı nedeniyle düşük tahmin kalitesine sahiptir [8]. Ayrıca, süreçteki değişken sayısını artırmak ve deneysel verilerin oluşturulması, özellikle biyoteknoloji ve malzeme bilimlerinde maliyetli ve zaman alıcıdır.

Fonksiyonun bir polinom ile temsil edildiği geleneksel RSM ile karşılaştırıldığında, yapay zekâ yöntemlerinden biri olan Makine Öğrenmesi (ML), veri güdümlü bir model olup var olan veriden kayda değer doğrulukta bilgi çıkarabilen ve daha esnek tahminler sağlayan bir modeldir [9-11]. ML'nin doğası gereği, deney tekrarlarına duyulan ihtiyaç azalır, bu da sürecin hızlanmasına, işçilik ve ekipman maliyetlerinin düşmesine neden olur. Bu anlamda ML, CCD'ye zaman ve maliyet açısından etkin bir alternatif sunmaktadır [12].

Nanopartiküllerin boyutu, çeşitli biyoteknolojik çalışmalarda etkili yüzey etkileşimi için önemlidir. Bu nedenle süreçlerde kullanılan koşullara göre boyutun düşük hata oranı ile tahmin edilmesi önemlidir. Ancak yeşil sentez yaklaşımında değişken sayısının artması ve deneysel verilerin oluşturulması, özellikle biyoteknoloji ve malzeme bilimlerinde maliyetli ve zaman alıcıdır [13].

Faktörlerin etkisinin incelenmesi için yeterli veri sunulması da önemlidir. Hem RSM hem de ML, girdiler ve çıktılar arasındaki matematiksel ilişki hakkında kesin bilgi bulunmayan sistemler için güçlü veri modelleme araçları olarak kabul edilmektedir. Her ikisi de karmaşık doğrusal olmayan ilişkileri yakalayabilmekte ve temsil edebilmektedir [8].

ML ve nanopartiküllerle ilgili literatür incelendiğinde, deneysel sonuçların uygulanmasıyla ilgili az sayıda makale olduğu görülmüştür. Nanopartikül sentezleme için, eczacılık ve ilaç taşınımı gibi alanlarda mevcut makaleler ya sadece RSM kullanmış ya da sadece ML ile boyut, şekil ve renk ilişkisi hakkında tahmini veriler sağlamıştır [6,13-21].

Bu bildirinin yazarlarının bir önceki çalışmasında, , gümüş nanopartiküllerin bitki ekstresi kullanılarak yeşil sentezi, Box Behnken Tasarımı (BBD) ve ML'nin iki aşamalı kullanımıyla daha az maliyet ve zaman ile, çevreci bir yaklaşımla gerçekleştirilmiştir [22].

Burada ise, benzer bir yaklaşım, CCD ve ML yöntemlerinin iki aşamada demir nanopartiküllerinin biyosentezinde kullanılması için benimsenmiştir. Çalışmada demir nanopartikülün yeşil sentezi için son zamanlarda tercih edilen mikroalgin potansiyeli araştırılmıştır. Erlenmeyer ölçeğinde demir nanopartiküllerin biyosentez optimizasyonunda mikroalgin büyüme ortamı ve metal içeriği oranı ile metal konsantrasyonunun ortalama harmonik partikül çapı üzerine etkisi öncelikle, istatistiksel deneysel tasarım kullanılarak gerçekleştirilmiştir. Ardından, girdi/çıktı ilişkisini öğrenmek ve genelleme yapmak için, veri üretimi ve makine öğrenimi uygulanmıştır. Süreçlerde kullanılan koşullara göre boyutun düşük hata oranı ile tahmin edilmesi önemlidir. Bu da daha az deneyle ekonomik bir yaklaşıma olanak sağlamaktadır. Mikroalg ile partikül sentezinde CCD ve ML kullanılarak ortalama harmonik partikül çapının tahminleme çalışmaları literatürde ilk kez uygulanmıştır. Bu açıdan bu çalışma, ML kullanarak yeşil sentezlenmiş nanopartiküllerin ticari üretimi için yenilikçi bir yaklaşım getirmektedir.

II. YÖNTEM

Bu çalışmada, biyoindirgeyici ajan olarak kullanılan mikroalglerden sentezlenen demir nanopartiküllerin ortalama harmonik partikül çapını belirlemek için iki aşamalı bir yaklaşım benimsenerek CCD deneyleri ve makine öğrenmesi birleştirilmiştir.

A. CCD deneyleri

Partikül sentezi için logaritmik fazdaki mikroalgal ortam (süpernatant), farklı konsantrasyonlarda hazırlanan $FeSO_4$ çözeltisinin içine 1 damla/s hızında damla damla eklenmiştir. Bu süreç, sıcaklık, inkübasyon süresi ve karıştırma hızı sırasıyla 25°C, 1 saat ve 200 rpm'de sabit tutularak 25 mL çalışma hacminde optimize edilmiştir. CCD ile deney tasarımında $FeSO_4$ / süpernatant oranı (0.2-8.56), $FeSO_4$ konsantrasyonu (1-12.95) ortalama harmonik partikül çapının optimizasyonunda çalışılmıştır (Tablo 1). Optimizasyon süreçleri için ilgili deneysel tasarımlar Design-Expert® yazılımı (sürüm 7.0.0, Stat-Ease Inc., Minneapolis, MN) kullanılarak gerçekleştirilmiştir.

Ortalama harmonik partikül çapının belirlenmesinde Malvern ZetaSizer Analizörü kullanılmıştır. Sonuçlar, birçok taramanın üç döngüsünün ortalamasıyla belirlenmiş (cihaz tarafından otomatik olarak hesaplanmıştır) ve ortalama boyut, ortalama değer ± ortalamanın standart hatası olarak sunulmuştur [23].

Bu tasarımları uyguladıktan sonra, CCD tarafından çıkarılan denklem aşağıdaki gibidir:

$$C = 142{,}22185 + 148{,}19296A + 403{,}41249B - 39{,}05248AB + 20{,}01249A^2 - 5{,}63265B^2 \quad (1)$$

Burada A $FeSO_4$/Süpernatant, B $FeSO_4$, C ise ortalama harmonik partikül çapıdır

TABLO I. DENEYLER

Deney	$FeSO_4$/ Süpernatant	$FeSO_4$	Ölçülen Çap	Hesaplanan Çap	Hata (%)
1	5,10	8,00	2.429,00	2.692,00	10,83
2	5,10	8,00	2.770,00	2.692,00	2,82
3	5,10	1,00	1.703,00	1.617,00	5,04
4	5,10	8,00	2.949,00	2.692,00	8,71
5	5,10	8,00	2.788,00	2.692,00	3,44
6	5,10	8,00	2.524,00	2.692,00	6,66
7	8,56	3,05	3.136,00	3.036,00	3,20
8	8,56	12,95	3.100,00	2.828,00	8,78
9	10,00	8,00	3.121,00	3.368,00	7,91
10	5,10	15,00	3.055,00	3.215,00	5,23
11	0,20	8,00	3.150,00	2.977,00	5,49
12	1,64	3,05	1.223,00	1.422,00	16,25
13	1,64	12,95	3.866,00	3.889,00	0,60
				Ortalama Hata	**6,54**

B. Makine öğrenmesi

Bu aşamada, ilk olarak makine öğrenmesi için gerekli olan veriler üretilmiştir. Tablo 2'de verildiği gibi, sistemin girdileri olan $FeSO_4$/süpernatant ve $FeSO_4$ değişkenlerinin alabilecekleri sınır değerler belirlenmiş, bu değişkenlerin, belirlenen sınırlar arasında kaç farklı değerde olabilecekleri öngörülmüş ve (2) numaralı denklem kullanılarak, değerler arasındaki aralık hesaplanmıştır.

$$Aralık = \frac{(En\ Yüksek - En\ Düşük)}{(Miktar - 1)} \quad (2)$$

Her iki girdi için de sınır değerler arasında 100 farklı değer olacağı kabul edildiğinden toplam 10.000 veri (ML için uygun) üretilmiştir. Daha sonra, girdiler için üretilen her bir veri için ilk aşamada ortaya çıkan formül kullanılarak 10.000 adet sistem çıktısı hesaplanmıştır. Fakat bu çıktılar, Tablo 1'de verildiği üzere, ortalama % 6,54 hata içermektedir. Bu hatayı düzeltmek için, söz konusu ortalama hata etrafında, standart sapması 1 olan, (3) numaralı denklemde verilen normal kümülatif dağılımın tersi ile 10.000 sayı üretilmiştir.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3)$$

Burada μ dağılımın ortalaması ve σ^2varyans anlamındadır. Bu sayılar, rasgele olarak yukarıdaki çıktılarla toplanarak ya da onlardan çıkartılarak yeni çıktılar elde edilmiştir. Makine öğreniminde, üretilen girdiler ve bu yeni çıktılar kullanılmıştır. Oluşturulan veri kümesinden bir kesit Tablo 3'te gösterilmiştir.

TABLO II. GİRDİLERİN SINIR DEĞERLERİ, MİKTARLARI VE ARALIKLARI

Girdiler	En Düşük	En Yüksek	Miktar	Aralık
$FeSO_4$/süpernatant	0,20	10,00	100	0.10
$FeSO_4$	1,00	15,00	100	0.14

TABLO III. MAKİNE ÖĞRENMESİ İÇİN ÜRETİLEN GİRDİ VE ÇIKTILAR

Sıra	Girdiler		Hata İçeren Çıktı	Hatası Düzeltilmiş Çıktı
	Fe	$FeSO_4$	Çap	Çap
1	0,20	1,00	562,630.00	532,933.00
2	0,20	1,14	616,868.00	653,933.00
...				
101	0,30	1,00	574,423.00	538,027.00
...				
9.999	10,00	14,86	2.573,33	2.784,03
10.000	10,00	15,00	2.551,37	2.734,88

III. BULGULAR

Tablo 3'teki girdiler ve hatası düzeltilmiş çıktı kullanılarak yapılan makine öğrenmesi için üç katmanlı bir sinir ağı kullanılmıştır. İlk katmanın girdileri $FeSO_4$/süpernatant ve $FeSO_4$ değerleridir. Gizli katmanlardaki düğüm sayısı olarak 10 seçilmiştir. Son katmanın çıktısı, ortalama harmonik partikül çapı değeridir. Tüm katmanlarda ReLU aktivasyon fonksiyonu kullanılmıştır. Eğitimde 0,01 faktörlü L2 regülasyonu ve Adam optimizasyonu uygulanmıştır. Optimize edilecek kayıp fonksiyonu, ortalama kare hatası olarak belirlenmiştir. Veri kümesinin rastgele %90'ı eğitime, %10'u sınamaya ayrılmıştır. Döngü sayısı olarak 100 seçilmiştir. Python dili ile Tensorflow ve Keras kütüphaneleri kullanılarak Google Colaboratory ortamında kodlama yapılmıştır.

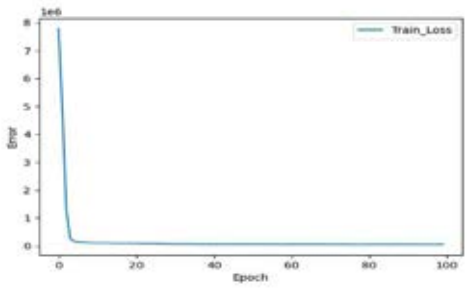
Önerilen yaklaşımı doğrulamak için, Tablo 1'de verilen gerçek deneylere ait girdi değerleri, eğitilmiş sinir ağına verilmiş ve gerçek çıktılar ile kıyaslanarak ortalama hata hesaplanmıştır. On kez tekrarlanan eğitim sonrasında bulunan hatalar Tablo 4'te verilmiştir.

TABLO IV. MAKİNE ÖĞRENMESİ SONUÇLARI

Eğitim	Eğitim Hatası (%)	Sınama Hatası (%)	Doğrulama Hatası (%)
1	4,20	4,01	5,32
2	3,80	3,79	5,51
3	3,92	4,00	5,36
4	4,19	4,10	5,60
5	4,19	4,18	5,39
6	4,13	4,35	5,63
7	4,22	4,27	5,84
8	3,68	3,87	5,90
9	3,88	3,80	4,84
10	4,34	4,52	5,63
Ortalama	**4,06**	**4,09**	**5,50**

Eğitimlerden birinde elde edilen kayıp grafiği Şekil 1'deki gibidir. Rassal ağırlıklardan dolayı yüksek başlayan kayıp eğitim süresince düşmüş ve 100 döngü sonunda stabil bir değere ulaşmıştır. Gerçek deneylerden oluşan doğrulama setine Denklem (1) uygulandığında ortalama 6.54% hata alınırken, sinir ağı kullanılınca ortalama hata %5.50'ye düşmüştür. Bu da,

hem CCD hem de makine öğrenmesini içeren iki aşamalı yaklaşımı kullanmanın avantajını göstermektedir.



Şekil. 1. Makine Öğrenmesi Eğitim Kayıp Grafiği.

IV. SONUÇLAR

Bu çalışma hem istatistiksel deneysel tasarım hem de ML ile oluşturulan ve etkili yeşil sentez nanopartikül üretiminin potansiyelini gösteren model bir çalışmadır. Nanopartikül sentezinde optimizasyon çalışmalarında oluşan verilerin ML için yeterli sayıda olmaması veri üretimi ve ML yaklaşımı ile umut verici sonuçlar açığa çıkarmıştır. İlk aşamada, deneysel kurulum, ikinci dereceden model formülasyonu ve veri üretimi için CCD kullanılmıştır. İkinci aşamada, gerçek deney sonuçlarına dayanan CCD formülasyonu ile üretilen verilerle otomatik ve hızlı bir veri analiz yöntemi olan makine öğrenimi modeli kullanılarak, bir yapay sinir ağı eğitilmiştir. Böylece ML, nanoparçacık boyutu ile çeşitli etkileyen faktörler arasındaki ilişkiyi öğrenmek için bu verilerle beslenmiştir.

Elde edilen sonuçlar, önerilen birleşik CCD ve ML stratejisinin, geliştirilmiş verilerle, tek başına istatistiksel CCD yaklaşımından daha iyi nanoparçacık boyutu tahminleri sağladığını göstermiştir. İki aşamalı CCD ve ML stratejisi aynı zamanda istatistiksel yöntemlerin modelleme için kullanıldığı ancak veri kullanılabilirliğinin sınırlı olduğu değerli biyoteknolojik ürünlerin (enzimler, polisakkaritler, bakteriyel selüloz, glikan vb.) üretimi, kirleticilerin giderilmesi, mikroorganizmaların sentezi, hidrojen üretim verimlerinin hesaplanması, çeşitli biyomedikal ürün optimizasyonları gibi herhangi bir alana da düşük maliyetli, hızlı ve çevre dostu bir yaklaşım ile uygulanabilme potansiyeline sahiptir.

KAYNAKLAR

[1] Nayak, S., Sajankila, S. P., Goveas, L. C., Rao, V. C., Mutalik, S., Shreya, B. A. "Two fold increase in synthesis of gold nanoparticles assisted by proteins and phenolic compounds in pongamia seed cake extract: response surface methodology approach". *SN Appl. Sci.*, 2(4), 2020, p1–12. https://doi.org/10.1007/s42452-020-2348-5.

[2] Ba-Abbad, M. M., Chai, P. V., Benamour, A., Ewis, D., Mohammad, A. W., Mahmoudi, E. "Optimizing and control of effective synthesize parameters for Fe_3O_4 nanoparticles using response surface methodology." *Chem. Pap.*, 76 (10), 2022, 6359–6370. https://doi.org/10.1007/s11696-022-02320-y.

[3] Caliskan, G., Mutaf, T., Agba, H. C., Elibol, M. "Green synthesis and characterization of titanium nanoparticles using microalga, *Phaeodactylum Tricornutum*". *Geomicrobiol. J.*, 39 (1), 2022, 83–96. https://doi.org/10.1080/01490451.2021.2008549.

[4] Salari, M. "Optimization by Box-Behnken Design and synthesis of magnetite nanoparticles for removal of the antibiotic from an aqueous phase." *Adsorpt. Sci. Technol.*, 2022. https://doi.org/10.1155/2022/1267460.

[5] Nayak, S., Manjunatha, K. B., Goveas, L. C., Rao, C. V., Sajankila, S. P. "Investigation of nonlinear optical properties of AgNPs synthesized using *Cyclea Peltata* leaf extract" *Post OVAT Optimization. Bionanoscience*, 11 (3), 2021, 884–892. https://doi.org/10.1007/s12668-021-00875-w.

[6] Abhishek, S. R., Sneha, N., Karthik, P. B. H. "An emerging treatment technology". *Contam. Emerg. Concerns Reigning Remov. Technol.*, No. August, 2022, 403–428. https://doi.org/10.1201/9781003247869-20.

[7] Keskin Gündoğdu, T., Deniz, I., Çalışkan, G., Şahin, E. S., & Azbar, N. "Experimental design methods for bioengineering applications". *Critical reviews in biotechnology*, 36(2), 2016, 368-388.

[8] Witek-Krowiak, A., Chojnacka, K., Podstawczyk, D., Dawiec, A., Pokomeda, K. "Application of response surface methodology and artificial neural network methods in modelling and optimization of biosorption process. *Bioresour. Technol.*, 160, 2014, 150–160. https://doi.org/10.1016/j.biortech.2014.01.021.

[9] Jordan, M. I., Mitchell, T. M. "Machine learning: trends, perspectives, and prospects." *Science,* 349(6245), 2015, 255–260. https://doi.org/10.1126/science.aaa8415.

[10] Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., Walsh, A. "Machine learning for molecular and materials science". *Nature*, 559 (7715), 2018, 547–555. https://doi.org/10.1038/s41586-018-0337-2.

[11] Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., Sarkar, S. "An explainable deep machine vision framework for plant stress phenotyping". *Proc. Natl. Acad. Sci. U. S. A.*, 115 (18), 2018, 4613–4618. https://doi.org/10.1073/pnas.1716999115.

[12] Zhang, Y. and Wu, Y. "Introducing machine learning models to response surface methodologies." *Intech,* 11, 2020, 13.

[13] Mekki-Berrada, F., Ren, Z., Huang, T., Wong, W. K., Zheng, F., Xie, J., Tian, I. P. S., Jayavelu, S., Mahfoud, Z., Bash, D., et al. "Two-step machine learning enables optimized nanoparticle synthesis." *Comput. Mater.*, 7 (1), 2021, 1–10. https://doi.org/10.1038/s41524-021-00520-w.

[14] Tao, H., Wu, T., Aldeghi, M., Wu, T. C., Aspuru-Guzik, A., Kumacheva, E. "nanoparticle synthesis assisted by machine learning." *Nat. Rev. Mater.*,6 (8), 2021, 701–716. https://doi.org/10.1038/s41578-021-00337-5.

[15] Ban, Z., Yuan, P., Yu, F., Peng, T., Zhou, Q., Hu, X. "Machine learning predicts the functional composition of the protein corona and the cellular recognition of nanoparticles." *Proc. Natl. Acad. Sci. U. S. A.,* 117 (19), 2020, 10492–10499. https://doi.org/10.1073/pnas.1919755117.

[16] Chen, X., Lv, H. "Intelligent control of nanoparticle synthesis on microfluidic chips with machine learning". *NPG Asia Mater.*, 14 (1), 2022, 1–20. https://doi.org/10.1038/s41427-022-00416-1.

[17] Schuhmacher, A., Gatto, A., Hinder, M., Kuss, M., Gassmann, O. "The upside of being a digital pharma player." *Drug Discov. Today*, 25 (9), 2020, 1569–1574. https://doi.org/10.1016/j.drudis.2020.06.002.

[18] He, Y., Ye, Z., Liu, X., Wei, Z., Qiu, F., Li, H. F., Zheng, Y., Ouyang, D. "Can machine learning predict drug nanocrystals?" *J. Control. Release*, 322, 2020, 274–285. https://doi.org/10.1016/j.jconrel.2020.03.043.

[19] Zhao, Q., Ye, Z., Su, Y., Ouyang, D. "Predicting complexation performance between cyclodextrins and guest molecules by ıntegrated machine learning and molecular modeling techniques." *Acta Pharm. Sin. B*, 9 (6), 2019, 1241–1252. https://doi.org/10.1016/j.apsb.2019.04.004.

[20] Gao, H., Wang, W., Dong, J., Ye, Z., Ouyang, D. "An ıntegrated computational methodology with data-driven machine learning, molecular modeling and pbpk modeling to accelerate solid dispersion formulation design." *Eur. J. Pharm. Biopharm.,* 158 (November 2020), 2021, 336–346. https://doi.org/10.1016/j.ejpb.2020.12.001.

[21] Gao, H., Jia, H., Dong, J., Yang, X., Li, H., Ouyang, D. "Integrated in silico formulation design of self-emulsifying drug delivery systems." *Acta Pharm. Sin. B*, 11 (11), 2021, 3585–3594. https://doi.org/10.1016/j.apsb.2021.04.017.

[22] Caliskan, G., and Kumluca Topalli, A. "Green synthesized silver nanoparticles in two stages: Box Behnken Design to machine learning." *Inorganic and Nano-Metal Chemistry*, 2021, 1-9.

[23] Ilhan-Ayisigi, E., Ulucan, F., Saygili, E., Saglam-Metiner, P., Gulce-Iz, S., and Yesil-Celiktas, O. "Nano-vesicular formulation of propolis and cytotoxic effects in a 3D spheroid model of lung cancer". *Journal of the Science of Food and Agriculture*, 100(8), 2021, 3525-3535.