



İLAÇ TASARIMINDA YAPAY ZEKÂ UYGULAMALARI

ARTIFICIAL INTELLIGENCE APPLICATIONS IN DRUG DESIGN

Özden TARI^{1*} , Nuray ARPACI¹ 

¹Çukurova Üniversitesi, Eczacılık Fakültesi, Farmasötik Kimya Anabilim Dalı, 01330, Adana, Türkiye

ÖZ

Amaç: Yapay zekâ üzerindeki çalışmaların giderek artması, her alanda olduğu gibi ilaç endüstrisinin de bu çalışmalardan faydalanmasına sebep olmaktadır. Bu çalışmada, yapay zeka uygulamalarının ilaç tasarımı ve geliştirilmesi üzerinde nasıl bir rol aldığına incelenmesi amaçlanmıştır.

Sonuç ve Tartışma: Yeni biyolojik olarak aktif bileşiklere ihtiyacın giderek arttığı günümüzde, yapay zekada sürekli yeni algoritmaların ortaya çıkması, güçlü hesaplama yeteneği, elde edilen kimyasal ve biyolojik verilerin birikmesi, ilaç tasarımı yapay zekâ kullanımına olanak sunmaktadır. İlaç tasarım aşamalarının neredeyse tüm basamaklarında uygulanabilen yapay zekâ yöntemleriyle, yeni ilaç geliştirilmesindeki uzun zaman gereksinimi ve yüksek maliyet gibi zorluklar azaltılmaya çalışılmaktadır. Bu çalışma sonucunda, yapay zekâ teknolojisinin ilaç tasarım sürecindeki uygulamaları ve geleneksel yöntemlere göre avantajları kapsamlı bir şekilde analiz edilerek karşılaştırılmıştır.

Anahtar Kelimeler: Derin öğrenme, ilaç tasarımı, makine öğrenimi, yapay zekâ

ABSTRACT

Objective: The increasing number of studies on artificial intelligence causes the pharmaceutical industry to benefit from these studies, as in every other field. This study is aimed at examining how artificial intelligence applications play a role in drug design and development.

Result and Discussion: In today's world, where the need for new biologically active compounds is increasing, the continuous emergence of new algorithms in artificial intelligence, strong computational ability, and accumulation of obtained chemical and biological data allow the use of artificial intelligence in drug design. With artificial intelligence methods that can be applied at almost all stages of drug design, difficulties such as long time requirements and high costs in developing new drugs are tried to be reduced. As a result of this study, the applications of artificial intelligence technology in the drug design process and its advantages over traditional methods have been extensively analyzed and compared.

Keywords: Artificial intelligence, deep learning, drug design, machine learning

GİRİŞ

İlaçlar; canlılarda mevcut hastalıkların tedavi edilmesini veya semptomların azaltılmasını aynı zamanda hastalıklardan korunmayı amaçlayan maddelerdir. Her hastalık için spesifik etkili olan ilaçlar üzerinde durulması önemli bir kavramdır. Her geçen gün yeni hastalıklarla karşılaşılma ile birlikte bilinen hastalıklar ve tedavileri için mevcut ilaçların her zaman tam olarak istenen özelliklerde ve

* Sorumlu Yazar / Corresponding Author: Özden Tari
e-posta / e-mail: ozdentari@hotmail.com, Tel. / Phone: +903223387334/170

Gönderilme / Submitted : 17.07.2023

Kabul / Accepted : 31.10.2023

Yayınlanma / Published : 20.01.2024

yeterlilikte olmamasından dolayı yeni ilaçlara olan ihtiyaç da artmaktadır. Bu nedenle ilaç tasarımı, yeni ilaç adayları olabilecek yeni biyoaktif maddelerin keşfi ve tasarımı, aynı zamanda mevcut ilaçların geliştirilmesi için kullanılan bir süreç olmaktadır. Bu kapsamda ilaç tasarımı, ilaç-reseptör ve ilaç-enzim etkileşimlerini anlamak ve simüle eden hesaplama yöntemlerini kullanan bir disiplin olarak ilerlemiştir.

İlaç tasarımında tercih edilen metotlar; araştırma ve geliştirme çalışmalarında, pre klinik çalışmalarda ve klinik çalışmalar (faz çalışmaları) aşamalarında büyük önem taşımaktadır. Bu aşamalarda, genellikle bir öncü bileşikle başlanarak ilaçların etki mekanizmasını iyileştirmek için bir dizi *in vitro* ve hücre bazlı test yapılmaktadır. Daha sonra, uygun *in vivo* testler ve fare modelleri kullanılarak ilacın farmakokinetik özellikleri değerlendirilmektedir. İlaçların hastalık hedeflerine yeterince bağlanma afinitesi göstermesi yeterli değildir, aynı zamanda hedefe ulaşmak için fizyolojik bariyerleri aşması gerekmektedir. İlaç tasarımında, istenmeyen proteinlerle hedef dışı etkileşimler ve genetik varyasyonun ilaç yanıtını etkileyen farmakogenetik etkileşimler gibi diğer sorunlar da göz önünde bulundurulmalıdır. Bu nedenle, ilaç tasarımı, araştırmacılar için oldukça zor bir çalışma sürecini içermektedir [1]. Ciddi zaman alan bu zorlu sürecin aynı zamanda yüksek maliyet gerektirmesi nedeniyle, çok daha etkin ilaç tasarımı ve geliştirme yöntemlerine ihtiyaç duyulmakta, bu amaçla çeşitli multidisipliner çalışmalar sürdürülmektedir [2].

Günümüze kadar ilaç tasarımında, ligand bazlı ve yapı bazlı moleküler tasarım metotları olan iki temel yaklaşım ile birlikte, kimyasal moleküllerin üç boyutlu özelliklerini inceleyerek hesaplamalar yapan bilgisayar destekli ilaç tasarım yöntemleri (CADD) karşımıza çıkmaktadır. Gelişen bilgi işleme gücü, hesaplamalı kimya ve biyolojinin hızlı gelişimi sayesinde de CADD, araştırma sürecini hızlandırmakla birlikte, klinik öncesi ve klinik çalışmalarla ilgili maliyet ve riski azaltmak için ilaç tasarım ve geliştirme sürecinin hemen hemen her aşamasında başarıyla uygulanmaktadır [3]. Son yıllarda, ilaç tasarımının çok karmaşık olması ve hibrit tekniklerin kullanılmasını gerektirmesinden dolayı, ilaç tasarım araçları olarak yeni teknolojik gelişmelerden özellikle yapay zekâ ve makine öğrenimi tekniklerinin (MLT) uygulanmasına olan ilgi artmaktadır [4].

Yapay zekâ (AI), makinelerin gösterdiği zekâ olarak kabul edilmektedir. Bu terim, bir makine öğrenimi veya problem çözümü gibi insanlarla ilişkili bilişsel davranışlar gösterdiğinde kullanılmaktadır. AI, yeni özelliklerin öğrenilmesi ve tahmin edilmesi için iyi kurulmuş makine öğrenimi gibi teknolojileri içermektedir. Özellikle, derin sinir ağları (DNN) veya tekrarlayan sinir ağları (RNN) gibi yapay sinir ağlarının gelişmesiyle yapay zekanın evrimsel nitelikteki ilerleyişi gerçekleşmiştir. Bu gelişmeler ilaç tasarımında yapay zekanın kullanılmasını kaçınılmaz kılmaktadır. MLT'den elde edilen, örnek olarak destek vektör makineleri (SVM), Rastgele Ormanlar (RF) ve Bayes Öğrenimi modelleri, özellikle sanal tarama çalışmalarında, molekül filtrelerinin geliştirilmesi/keşfedilmesi için kullanılabilir. Bu sayede ilaç tasarımı alanındaki önemli zorluklardan olan farmakokinetik ve toksisite özelliklerinin tahmin edilmesiyle klinik fazlardaki başarısızlıklar önlenmektedir [4].

Son zamanlarda derin sinir ağları gibi yapay sinir ağlarının gelişmesiyle fizikokimyasal ve ADME/T özellikleri gibi özellik veya aktivite tahminlerinde çok sayıda uygulama ortaya çıkmış ve bu teknolojinin nicel yapı-özellik ilişkilerinde (QSPR) veya nicel yapı-aktivite ilişkilerinde (QSAR) hesaplama gücünü ve netliğini arttırmıştır. Aynı zamanda de novo tasarımda yapay zekâ kullanımıyla, yeni ve biyolojik olarak aktif moleküllerin üretimi, istenen özellikler doğrultusunda tasarlanabilmektedir [5].

Sonuç olarak, AI teknolojisi ve makine öğrenimi teorisinin gelişmesiyle birlikte farmakolojik verilerin birikmesi sayesinde, sanal tarama, aktivite puanlama, kantitatif yapı-aktivite ilişkisi (QSAR) analizi, de novo ilaç tasarımı ile absorpsiyon, dağılım, metabolizma, atılım ve toksisite (ADME/T) özelliklerinin *in silico* değerlendirilmesi gibi birçok ilaç tasarımı alanlarında AI tabanlı modeller kullanılmaktadır. AI tabanlı modeller ilaç tasarımını çok yönlü çerçeveler aracılığıyla yönlendirmeye yardımcı olmak için büyük bir güç olarak kullanılmaktadır. Son zamanlarda, güçlü genelleme yeteneği ve güçlü özellik çıkarma yeteneği nedeniyle, moleküler özelliklerin tahmin edilmesinde ve istenen moleküllerin üretilmesinde derin öğrenme yöntemleri kullanılmıştır ve bu da AI teknolojilerinin ilaç tasarımı alanında uygulanmasını daha da teşvik etmektedir [3].

Yaşanan gelişmeler ilaç tasarımı ve yapay zekanın ortak bir noktada birleşmesi zorunluluğunu doğurmaktadır. Bu sebeple bu alanlarda yapılan çalışmaları anlamak ve güncel çalışmaları takip etmek

büyük bir önem taşımaktadır. Yapılan bu çalışmada, yapay zekada uygulanan yöntemler ile son teknolojik gelişmelerle birlikte ilaç tasarımındaki yapay zekâ uygulamalarına ve önemine değinilmektedir.

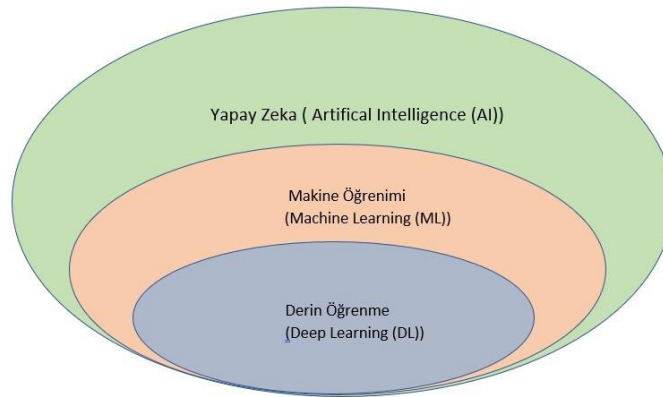
Yapay Zekâ

Yapay zekâ 'normalde insan zekâsı gerektiren görevleri yerine getirebilen bilgisayar sistemlerinin teorisi ve gelişimi' olarak tanımlanmaktadır. Farklı yeteneklerin (yani, öğrenme, akıl yürütme, problem çözme, algılama ve dili kullanma) bir kombinasyonu olan insan zekasını taklit etmek üzere yapay (veya makine) zekâ, aynı yetenekleri makinelerde kullanmak için farklı bilim ve mühendislik disiplinlerinden oluşan yöntem ve teknikleri kullanmaktadır. Yapay zekâ genellikle makine öğrenimi ile karıştırılmaktadır. Aslında, yapay zekada öğrenme (makine/derin öğrenme), makinelerdeki öğrenme yeteneklerini özümsemek için yöntem ve tekniklerle ilgilenen bir alt alandır. Yapay zekâ, makine öğrenimi, matematik, istatistik ve bilgisayar bilimlerinden bazı görevlerle (yani problemin doğası); ilgili deneyimlerden (tarihsel veriler) öğrenmek ve performansı ölçmek (performans matrisi) ve geliştirmek (yeniden uygulama) için yöntem ve teknikleri içeren multidisipliner bir alt alan olarak karşımıza çıkmaktadır [6].

Yapay zekâ, basit bir "koşula bağlı kurallar" olarak başlayıp, yıllar içinde insan beynine benzer şekilde çalışan daha karmaşık algoritmaları içerecek şekilde ilerlemiştir [7].

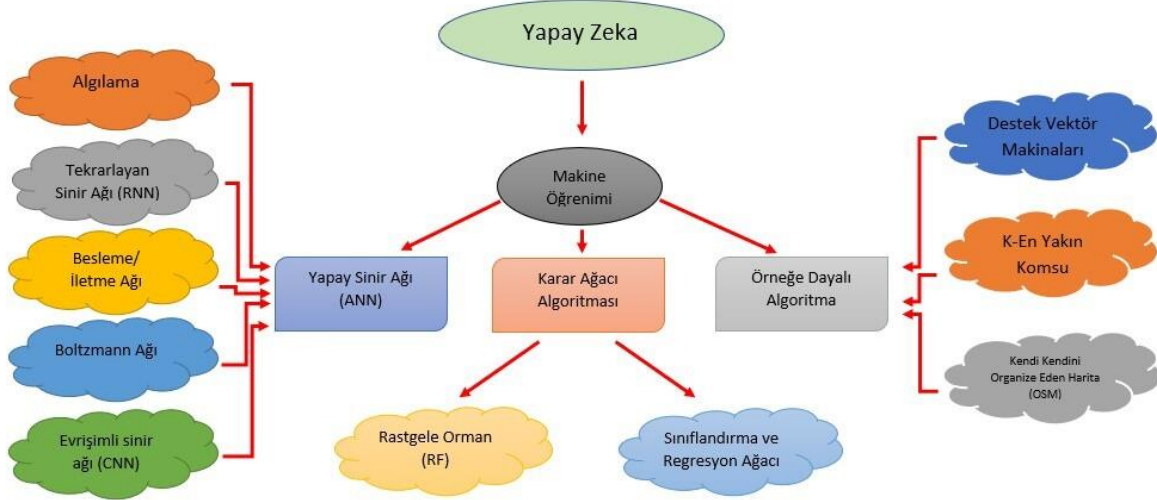
Yapay zekâ teknolojisi, verileri değerlendirmek için bilgisayarlı görüş, derin öğrenme, makine öğrenimi, doğal dil işleme, robotik, konuşma, denetimli öğrenme ve denetimsiz öğrenme gibi uygulamaları kullanmaktadır. Makine öğrenimi (Machine Learning (ML)), belirli bir durumu analiz etmek için kullanılabilir kalıpları tanımlamak için belirli özelliklerin kullanılmasıdır. Makine, tanımlanan bilgiyi "öğrenebilir" ve bu bilgileri gelecekteki benzer senaryolara uygulayabilmektedir. ML yapılandırılmış ve yapılandırılmamış verilerden öğrenen, gizli kalıpları belirleyen, sınıflandırmalar yapan ve ileri adımları öngören bilgisayar algoritmalarından oluşmaktadır. Derin öğrenme (Deep Learning (DL)), konuşma, görüntü ve video gibi kalıpları tanımlamak ve tüm bu verileri ayırt etmek için sinir ağlarından oluşan makine öğrenimi tabanlı bir yaklaşımdır. İnsan beynine benzer şekilde, öğrenebilen ve kendi kendine karar verebilen bir yapay sinir ağı (YSA) oluşturmak için algoritmalarından oluşan DL, ML'nin daha geliştirilmiş şekli olarak karşımıza çıkmaktadır. Doğal dil işleme (Natural Language Processing (NLP)), hesaplama tekniklerinin doğal dil ve konuşmanın analiz ve sentezini yapmaktadır [8-11]. Geleneksel ML, ham verilerden elde edilen özelliklerin manuel olarak oluşturulmasını ve sıralanmasını gerektirirken DL, yüksek boyutlu, çok sayıda ve heterojen olan ham verilerden en önemli soyut özellikleri, sürekli olarak oluşturabilen çok sayıda gizli katman ve nörona sahiptir. Standart veya aşırı zor testlerde üstün sonuçların elde edilebilmesi için pratik olarak hiçbir manuel müdahale gerektirmeyen bu süreç, sadece genellemelerden kaynaklı küçük bir hata payına sahiptir [12].

Yapay zekâ, makine öğrenimi ve derin öğrenme arasındaki kavramsal ilişki Şekil 1'de verilmektedir [13].



Şekil 1. Yapay zekâ, makine öğrenimi ve derin öğrenme arasındaki ilişki [13]

Makine öğreniminin dört temel alt türü bulunmaktadır. Bunlar; denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme ve güçlendirilmiş öğrenme olarak karşımıza çıkmaktadır. Makine öğreniminin alt dallarında yaygın kullanılan sinir ağları ve algoritmalar Şekil 2’de verilmektedir.



Şekil 2. Yapay zekanın alt dallarında yaygın kullanılan sinir ağları ve algoritmalar [13]

Denetimli öğrenme (Supervised learning), görev odaklıdır. Örnek veri noktalarını etiketleri bilinmektedir. Model, test örneklerinin etiketlerini tahmin etmek için bu bilinen örnekler üzerinde eğitilmektedir [14]. Denetimli öğrenmede yapay sinir ağı için hem girdi hem de çıktı değerleri belirlenmektedir. Ağın ürettiği çıktı ile istenen çıktı arasındaki fark sıfır ya da sıfıra yakın bir değere ulaşıncaya kadar yapay sinir ağının ağırlıkları değiştirilmektedir. Algoritma kabul edilebilir doğruluğa geldiğinde ya da eğitim veri kümesi tükendiğinde eğitim sonlandırılmaktadır. Yaygın olarak kullanılan denetimli makine öğrenimi algoritmalarına, Karar Ağaçları, Rastgele Ormanlar (RF), Destek Vektör Makineleri (SVM), Doğrusal ve Lojistik Regresyon, Naive Bayes, Doğrusal Diskriminant Analizi, k-En Yakın Komşu (KNN) algoritmaları ve Sinir Ağları örnekleri verilebilmektedir [15].

Denetimli makine öğreniminde sınıflandırma ve regresyon olmak üzere iki alt alan bulunmaktadır. Öngörülen çıktılarının sonuçlarını değerlendirirken, problem kategorik olduğunda sınıflandırma modelleri kullanılmaktadır. Problem bir aralık içindeki sayısal bir değeri öngörmeyi içerdiğinde regresyon modelleri kullanılmaktadır. Rastgele Orman, Otomatik Kodlayıcılar (AutoEncoder (AE)) ve Evrişimli Sinir Ağları (CNN) gibi çeşitli farklı türde makine öğrenimi uygulamaları bulunmaktadır [16].

Denetimsiz öğrenme (unsupervised learning), verilerle ilişki bulmak için etiketli verilerle eğitime bağlı değildir. Veri odaklı olup, gizli yapıları bulmak için verilerdeki kalıpları öğrenmeyi amaçlamaktadır [14]. Denetimsiz öğrenmede, ağa sadece girdi verileri öğretilir. Girdi değerlerine uygun bir çıktı üretilinceye kadar bağlantı ağırlıkları değiştirilir. Bu metot görüntü işleme, işaret işleme, kontrol problemlerinde etkin olarak kullanılmaktadır. Denetimsiz öğrenmede Hiyerarşik Kümeleme (Hierarchical Clustering), K-Ortalama Kümeleme (k-Means Clustering), Gaussian Karışım Modelleri (Gaussian Mixture Modeling (GMM)), Kendi Kendini Düzenleyen Haritalar (The Self Organizing Map (SOM)) ve Gizli Markov Modelleri (Hidden Markov Model (HMM)) en yaygın kullanılan modellerdir [15].

Makine öğreniminde, etiketli veri, bir veri kümesindeki her örneğin doğru cevabının veya sınıflandırmasının önceden belirlenmiş olduğu verilerdir. Bu tür veriler, makine öğrenimi algoritmalarının eğitiminde kullanılmakta ve bu sayede algoritmaların doğru sonuçlar üretmesi için gerekli olan doğru cevaplar sağlanmaktadır. Örneğin, bir resim sınıflandırma modeli eğitiliyorsa, etiketli veriler resimlerin hangi nesnelere içerdiğini belirten etiketlerdir. Yarı denetimli öğrenme (Semi-

supervised learning), daha küçük bir etiketlenmiş veri kümesi ile bir modelin birkaç veriden tahmin edilmesi gereken ve daha az maliyetli olan daha büyük bir etiketlenmemiş veri kümesinin içindedir. Eğitim için etiketli veri noktalarını kullanmakta ve etiketli verilerden gelen bilgileri kullanarak bilinmeyen örneklerin etiketlerini tahmin etmeyi amaçlamaktadır.

Güçlendirilmiş öğrenme (Reinforced learning), çevre ile etkileşim kurarak sürekli öğrenebilen ve belirli bir veri setini kullanmayı amaçlayan bir sistemdir. Yarı denetimli öğrenmede, denetimli öğrenmeye benzer algoritmalar kullanılmıştır. Güçlendirilmiş öğrenme için kullanılan algoritmalara Q-Öğrenme (Q-Learning) ve Derin Öğrenme örnek olarak verilebilmektedir [17]. Derin Öğrenme, veri modellemelerini öğrenmek için birçok doğrusal olmayan işlem birimi katmanıyla yapay sinir ağlarını (Artificial Neural Networks (ANN)) kullanan makine öğreniminin alt dallarındandır. İnsan merkezi sinir sistemini taklit etmek amacıyla, bağlı yapay nöronlardan oluşan ‘yapay sinir ağı’ adı verilen geleneksel makine öğrenimi algoritmalarının yeniden tasarlanıp kullanılmasıdır [18,19]. Derin öğrenme yöntemleri, temel olarak derin sinir ağlarını kullanır ve bu arın farklı türleri vardır. Örneğin, Evrişimli Sinir Ağları (Convolutional Neural Networks, CNN), görsel verileri işlemek için kullanılırken, tekrarlayan sinir ağları (Recurrent Neural Networks, RNN) sıralı verileri işlemek için kullanılmaktadır. Otomatik kodlayıcılar (Autoencoders) veri boyutunu azaltmak ve özellik çıkarmak için kullanılabilir. Kısıtlı Boltzmann Makineleri (Restricted Boltzmann Machines) ise yapay sinir ağı modellerinin eğitimi için kullanılmaktadır.

Derin Sinir Ağları (Deep Neural Networks (DNN)), birden çok gizli katmana sahip birbirine tamamen bağlı sinir ağlarıdır. Her gizli katman, birden fazla doğrusal olmayan işlem birimini içermektedir. Böylece DNN'ler, istenen özellikteki verileri otomatik olarak çıkarabilmek için birçok farklı katmanda birden çok nöron kullanabilmektedirler. Bu özellikleri sayesinde DNN'ler, görüntü işleme, doğal dil işleme, ses tanıma ve diğer pek çok uygulama alanında oldukça etkili bir şekilde kullanılabilir [14].

CNN, derin öğrenmedeki en temel yöntemlerden biridir. Evrişimli Sinir Ağı, görüntü işleme, nesne sınıflandırması ve doğal dil işleme gibi alanlarda kullanılan ileri beslemeli bir sinir ağıdır. CNN'ler, özellikle görüntü işlemede oldukça etkilidirler ve çok boyutlu diziler biçiminde olan verileri işlemek için tasarlanmıştır. Bu nedenle, gen ekspresyon veri analizi gibi diğer uygulama alanlarında da kullanılmaktadırlar. CNN'lerin temel özelliği, evrişim ve havuzlama katmanlarından oluşmasıdır. Evrişim katmanları, özellikle görüntülerdeki özellikleri çıkarmak için kullanılırken, havuzlama katmanları ise boyut azaltma işlemi yaparak hesaplama yükünü azaltmaktadırlar [20].

CNN, 1980'lerde Fukushima tarafından, Hubel ve Wiesel'in bir kedinin görsel korteksindeki alıcı alan (receptive field) araştırmalarından esinlenen, çok katmanlı yapay sinir ağı olan Neocognitron'un geliştirilmesiyle oluşturulmuştur [21,22].

CNN, görsel sinyalleri işlerken, yerel nöron kalıpları duyuşal boşluktaki belirli bölgeleri algılamaktadır. Aynı zamanda evrişimli katmanlarda seyrek bağlanabilirlik ve paylaşılan ağırlıklar olmak üzere iki ana karakter geliştirerek bunların özelliklerini taklit edebilmektedir [19].

RNN, derin öğrenmedeki diğer bir modelleme türüdür. Özellikle sekans verilerini işlemeyi amaçlayan RNN, NLP'de yaygın olarak kullanılmış ve büyük başarılar elde edilmiştir. RNN'ler tahmin modellerinde ve dil modellemelerinde yaygın olarak kullanılmaktadır. RNN, bir katmanın çıktısını kaydetme ve katmanın çıktısını tahmin etmek için girişi geri besleme prensibiyle çalışmaktadır [19].

1950'lerden 1970'lere Yapay Zekanın Hayatımıza Girişi

Akıllı davranışı ve eleştirel düşüncüyü simüle etmek için bilgisayarları kullanma fikri ilk olarak 1950'de Alan Turing tarafından ortaya konulmuştur [23]. Bilgisayarlar ve Zekâ kitabında Turing tarafından, bilgisayarların yetenekli olup olmadığını belirlemek için ‘‘Turing testi’’ olarak bilinen basit bir testi tanımlanmıştır [9]. Bundan altı yıl sonra ise John McCarthy yapay zekâ (AI) terimini ‘‘akıllı makineler yapma bilimi ve mühendisliğı’’ olarak nitelendirmiştir [24,25].

Yapay zekâ, erken dönemdeki AI çalışmalarında yalnızca insanların yapabildiğı çıkarımlar yapma veya karar verme yeteneğine sahip makinelerin geliştirilmesine odaklanmıştır. İlk endüstriyel robot kolu (Unimate; Unimation, Danbury, Conn, ABD) 1961'de General Motors'taki montaj hattına katılmış ve otomatik döküm gerçekleştirilmiştir [26]. Birkaç yıl sonra (1964) Joseph Weizenbaum tarafından, bir makine olan Eliza tanıtılmıştır. Doğal dil işlemeyi kullanan Eliza, kalıp eşleştirme ve

değiştirme kullanarak, insan konuşmasını taklit etme metodolojisiyle sohbet robotları olma görevini görmüştür [27]. 1966 yılında ise “ilk elektronik insan” olan Shakey geliştirilmiştir. Shakey, Stanford Araştırma Enstitüsü’nde talimatları yorumlayabilen ilk mobil robottur. Sadece 1 adımlık komutları takip etmek yerine, daha karmaşık talimatları işleyerek uygun eylemleri saptamıştır. Bu gelişme yapay zekada bir dönüm noktası olmuştur [28].

1970’lerden 2000’lere AI Kışı

Yapay zekâ adına daha az önemli gelişmelerin yaşanması nedeniyle, fonlamanın azaltılarak faizlerin artırıldığı bu yıllar ‘AI kışı’ dönemi olarak tanımlanmıştır [9]. 1970’lerin sonlarında başlayan ve 1990’ların başlarına kadar uzanan, uzman dijital bilgi veri tabanlarının geliştirilmesi ve sürdürülmesindeki aşırı maliyet nedeniyle bu dönemlerde AI ile ilgili gelişmeler yavaşlamıştır. Buna rağmen, 1971’de Saul Amarel tarafından Rutgers Üniversitesi’nde ‘Biyotıpta Bilgisayarlar Araştırma Kaynağı’ üzerine çalışılmış ve 1973’te Stanford Üniversitesi Tıp Bölümünde, çeşitli kurumlardan biyomedikal araştırmacılar tarafından, gelişmiş ağ oluşturma yeteneği olan paylaşımlı bir bilgisayar sistemi ‘Tıpta Deneysel-Yapay Zekâ’ oluşturulmuştur [29].

Bunların dışında 2000’li yıllara kadar kayda değer başka gelişme yaşanmamıştır.

2000’den Günümüze; Çığır Açan Gelişmeler

2007 yılında IBM (International Business Machines) tarafından geliştirilen ve doğal dilde sorulan sorulara cevap vermek için tasarlanan, bir yapay zekâ programı olan ‘Watson’ geliştirilmiştir. Watson telesekreter sistemi, 2011’de insanlarla yarışan ‘Jeopardy!’ adlı oyun programında birinciliği kazanmıştır. DeepQA adı verilen bu teknoloji, ileriye dönük akıl yürütme (verilerden sonuçlara kadar takip etme) ve geriye dönük akıl yürütme (takip etme) kullanılan geleneksel sistemlerin aksine, olası cevaplar oluşturmak için yapılandırılmamış içerik üzerinde verileri analiz etmek için doğal dil işlemeyi ve çeşitli veri tabanlarını kullanmıştır [30]. Ardından yüzeysel iletişimi kullanan sohbet robotlarından (Eliza) sonra anlamlı konuşma tabanlı arayüzleri kullanan Apple’ın sanal asistanı Siri (2011) ve Amazon’un sanal asistanı Alexa’ya (2014) doğal dil işleme teknolojisi uygulanmıştır [31,32].

Mart 2016’da ise AlphaGo, dünyanın en iyi Go oyuncularından biri olan Lee Sedol’u yenerek yapay zekayı (AI) yeniden kamuoyunun dikkatine sunmuş ve büyük ilgi uyandırmıştır [33]. IBM tarafından geliştirilen ve satranç oynayan bir bilgisayar olan Deep Blue ile karşılaştırıldığında, AlphaGo, en başarılı bilgisayarlardan biri olan evrimsel sinir ağı (CNN) adı verilen sinir ağlarında (NN’ler) derin öğrenme (DL) algoritmalarının uygulamalarını kullanan gelişmiş ve yenilikçi bir mimariyi entegre etmiştir [34].

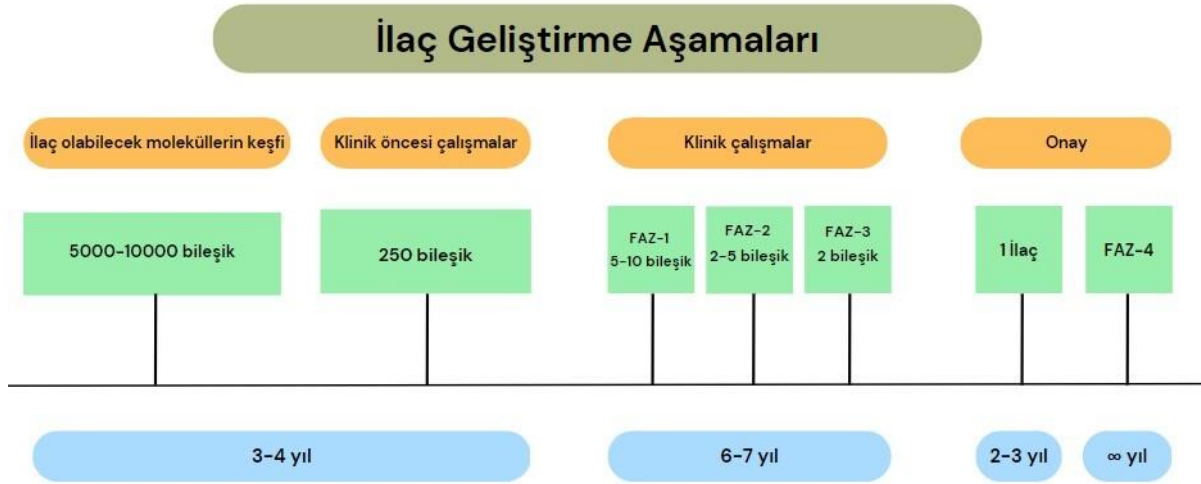
Ardından, 2017 yılında Fabien Beckers tarafından, bulut/derin öğrenme girişimi olan, görüntü tabanlı hassas tıp cihazları geliştiren Arterys’in kurulumunun ardından ilk Gıda ve İlaç Dairesi (FDA) onaylı bulut tabanlı yapay zekâ destekli kardiyak görüntüleme ilk kez ABD’de sağlık sektöründe kullanılmaya başlanmıştır.

Yine 2017 yılında yapay zekayla geliştirilen, otomatik olarak hasta alımı yapan ‘Chatbot Mandy’ teknolojisinden sonra, 2018-2020 yılları arasında karar destek sistemi, görüntüleme ve tanı koyma gibi konularda birçok AI destekli cihaz ve teknoloji pediatri, gastroenteroloji, kardiyoloji, oftalmoloji gibi alanlarda kullanılması yaygınlanmıştır [7,35,36]

OpenAI tarafından 2022 yılında, metin tabanlı sorgulara yanıt vermek ve doğal dilde yanıtlar oluşturmak için tasarlanmış bir dil modeli ‘ChatGPT’ geliştirilmiştir. ChatGPT, dildeki kalıpları öğrenmeyi, tutarlı ve insan benzeri metinler oluşturmayı sağlayan Transformer adlı derin bir öğrenme mimarisi kullanılarak oluşturulmuştur. Çok büyük bir metin verisi dizileri üzerinde eğitilmiştir ve bu nedenle genel bilgi sorularından daha karmaşık konuşma konularına kadar çok çeşitli taleplere yanıtlar üretebilmektedir. Genel olarak, ChatGPT, NLP alanında önemli bir ilerlemeyi temsil etmekle birlikte bilgisayarlar ve dijital sistemlerle etkileşim biçimimizde devrim yaratma potansiyeline sahiptir [37]. Mart 2023’e gelindiğinde ise OpenAI’nin derin öğrenmeyi ölçeklendirme çabalarındaki en son kilometre taşı olan GPT-4 geliştirilmiştir. GPT-4, insanlardan daha az yetenekli olmasına rağmen, çeşitli profesyonel ve akademik kriterlerde insan düzeyinde performans sergileyen, görüntü ve metin girişlerini kabul edip metin çıktılarını yayan, büyük birçok modlu model olarak karşımıza çıkmaktadır [38].

İlaç Tasarımı

İlaç tasarımı, daha önceden tanımlanmış yapı-aktivite ilişkilerinden yararlanarak, farmakolojik aktivitesi öngörülebilir potansiyel ilaç moleküllerinin tasarlanmasıdır. Yeni ilaç molekülleri geliştirmenin amacı, var olanlardan daha güçlü, daha az toksik ve yan etkileri en aza indirilmiş, yararlı terapötik bileşikler geliştirmektir. Var olan ilaçların hemen hemen hiçbiri her yönüyle istenilen özelliklere sahip olmadığından, araştırma kurumlarında ve ilaç endüstrisinin araştırma-geliştirme laboratuvarlarında, sürekli olarak ilaç olabileceği düşünülen yeni bileşikler geliştirilerek elde edilmekte ve birçok farmakolojik aktivite testlerine tâbi tutulmaktadır. Günümüzde tedaviye girmiş 14.000'in üzerinde ilaç bulunmakta ve yeni ilaç molekülleri geliştirmek üzere yapılan çalışmaların sayısı giderek artmaktadır [39]. İlaç geliştirme yöntemlerindeki ilerlemelerle, terapötik etkileri iyileştirilmiş yeni ilaçlar tedaviye sunulabilmektedir. Geliştirilen bu ilaçlar, 20-30 yıl önce yaygın olarak kullanılan birçok ilacın yerini alabilmektedir. Ayrıca, tıp ve biyolojik bilimlerdeki gelişmelerle, nedeni bilinmeyen bazı hastalıkların mekanizmaları açıklanmakta ve henüz tedavisi olmayan hastalıklarda kullanılmak üzere geliştirilen yeni ilaçlar da tedaviye girmektedir. Ancak yeni bir ilacın geliştirilme aşamasından klinik kullanıma sunulmasına kadar uzun ve zorlu bir süreç gerekmektedir. Bu zorlu sürecin Şekil 3'te görüldüğü gibi farklı aşamaları bulunmaktadır [40].



Şekil 3. İlaç geliştirme aşamaları [40]: Klinik öncesi çalışmalarda yeni bileşiğin etkinliği ve güvenliğiyle ilgili çalışmalar, klinik çalışmalarda faz 1, faz 2 ve faz 3 aşamaları (Faz 1 de az sayıda sağlıklı kişide yan etkiler ve güvenliğin doğrulanması ile ilgili çalışmalar yapılırken faz 2 de az sayıda hastada etkililik, güvenli dozaj ve uygulama yönteminin doğrulanması yapılır. Faz 3 aşamasında çok sayıda hastada mevcut ilaçlara kıyasla etkinlik ve güvenlik doğrulanması yapılır. Faz 4 aşamasında ise onaylanmış ilacın toplumda uygulanmasıyla uzun dönemdeki etki ve güvenlik sonuçları toplanır)

İlaç geliştirmede farmasötik tarama testlerinden önce etkin olabilecek aday bileşiklerin geliştirilmesinde iki önemli aşama söz konusudur. İlk olarak, belli bir farmakolojik aktivite için model (prototip) olan öncü bileşiğin bulunması gerekmektedir. Genellikle bu bileşik istenen farmakolojik aktiviteyi göstermekle birlikte etkisi yetersiz olabilmekte ve/veya istenmeyen farmasötik, farmakokinetik ve toksik özelliklere sahip olabilmektedir. Bu nedenle ikinci aşamada, bu öncü bileşiklerden yola çıkılarak optimizasyon çalışmaları yapılmaktadır. Öncü bileşiği bulabilmek için pek çok yol izlenebilmektedir. Bunlar geçmişten beri bilinen, doğal kaynaklardan molekül elde etme, tesadüfen bulma, klinik gözlemler, farmakolojik tarama testleri, ilaç metabolizma çalışmaları ve etkin endojen moleküllerin yapılarını taklit etme ve akılcı ilaç tasarımı gibi yöntemlerdir. Günümüzde bu yöntemler içinden akılcı (rasyonel) ilaç tasarımı daha çok ön plana çıkmıştır. Akılcı ilaç tasarımında, yeni ilaç molekülleri geliştirmek üzere hem öncü bileşiğin bulunması hem de optimizasyon aşamalarında çeşitli yöntemler kullanılmaktadır. Yöntemlerden biri olan bilgisayar destekli ilaç

tasarımı, ilaç aktivitesinde önemli fizikokimyasal parametreleri, kantitatif yapı-aktivite ilişkilerini ve kuantum kimyasal modelleri kullanmaktadır. Diğer yöntem olan moleküler modellemede ise öncü molekül veya analog tasarımı için çeşitli bilgisayar programları, X-ışınları kristalografisi gibi yöntemlerle ilacın moleküler şekli ve konformasyonu araştırılmaktadır. Bunlarla birlikte, ilacın hedefle olan etkileşimin gücü ve uyumu değerlendirilerek etkin maddelerin tasarlanması için reseptör tanımlama yöntemleri kullanılmaktadır. İlacın geliştirilmesinde tüm aşamalarda etkin olarak kullanılan bilgisayar yardımlı ilaç tasarımı, günümüzde en yaygın kullanılan teknik olmuştur [40].

Bir ilaç molekülünün tasarımı ve geliştirilmesi uzun zaman alan, yüksek maliyetli ve kompleks aşamalar içeren zorlu bir süreçtir. Son zamanlarda, bilgisayar destekli teknolojilerde ve yapay zekanın uygulama alanlarındaki gelişmelerin ilaç tasarımı ve geliştirme alanına da yansımalarıyla daha başarılı ilaç adayları molekülünün, daha kısa zamanda ve daha az maliyetle elde edilmesine olanak sağlamaktadır.

İlaç Tasarımında Yapay Zekâ Uygulamaları

Yapay zekâ (AI), çeşitli sektörlerde en yeni uygulamaları kullanıma sunmaktadır. Bilgisayar destekli ilaç tasarım teknolojisi, yüksek verimlilik, düşük maliyet ve hızlı sonuç verme gibi avantajlarıyla ilaç keşfinde ve tasarımında önemli bir rol oynamaktadır. Bilgi işlemedeki önemli ilerlemeler, üst düzey algoritmaların geliştirilmesi, biyoloji ve kimya alanlarındaki verilerin çoğalması da bunu desteklemektedir. Son yıllarda, makine öğrenimi (ML) algoritmalarındaki hızlı ilerleme nedeniyle, AI, ilaç tasarımında çeşitli aşamalarda yaygın olarak kullanılmaktadır. Makine öğrenimi algoritmalarının genelleme yeteneği güçlenerek, daha etkili ve daha büyük veri işleme yeteneğine sahip olmasıyla derin öğrenme tekniğine dönüşmüştür; bu durum, yapay zekâ teknolojisi ile bilgisayar destekli ilaç tasarım teknolojisinin entegrasyonunu daha da teşvik ederek en yeni ilaçların tasarımını ve keşfini hızlandırmaktadır [41].

İlaç tasarımı ve keşfinin tüm işlemleri, hedef tanıma, hedef bulma, potansiyel hasta grupların optimizasyonu, klinik öncesi ilaç adaylarının tanınması, klinik öncesi çalışmalar ve klinik araştırmayı içermektedir. Raporlara göre, daha yeni bir ruhsatlı ilacın standart araştırma ve geliştirme döngüsü yaklaşık 10-17 yıl süreye ve ortalama 2.558 milyar dolar bütçeye ihtiyaç duymaktadır [42,43]. Bununla birlikte, uzun zaman ve yüksek maliyet harcanmasına rağmen, küçük moleküllü ilaçlar için onaylanarak ruhsat alma oranları %10'dan azdır. Bilgisayar destekli ilaç tasarımı yaklaşımları ile ilaç tasarım çalışmalarında önemli akılcı ipuçları elde edilerek daha etkin bileşiklere daha uygun maliyetlerle ulaşılabilmektedir [41].

Günümüzde teknolojinin ve yapay zekâ çalışmalarının ilerlemesi ile geliştirilen çeşitli algoritmalar ve sinir ağları bilgisayar destekli ilaç tasarım çalışmalarında da kullanılmaya başlanmıştır. Makine öğrenimi ve derin öğrenmedeki ilerlemelerle, teknolojinin her alanında çarpıcı gelişmeler ortaya konulmasıyla, yapay zekâ disiplinleri arası ortak bir alan olmuştur. AI tabanlı hesaplamalı modelleme, bileşikler için biyolojik hedef aranması ve toksisiteleri açısından değerlendirilmesi için umut verici bir yöntem olarak karşımıza çıkmaktadır. Kantitatif yapı-aktivite ilişkisi (QSAR) tabanlı yaklaşımlar da dahil olmak üzere mevcut hesaplama modelleri, çeşitli biyolojik hedefler için çok sayıda yeni molekülün hızlı bir şekilde ön görülmesinde kullanılabilmektedir [41].

Yapay zekanın alt gruplarından olan makine öğrenimiyle birlikte, yıllardır, Naive Bayes Sınıflandırıcısı (NB), Lojistik Regresyon (LR), En Yakın K Komşusu (KNN) algoritması, Destek Vektör Makineleri (SVM), Çoklu Doğrusal Regresyon (MLR), Gauss Süreci (GP), Rastgele Ormanlar (RF), Boosting ve Karar Ağacı, ilaç tasarımında yaygın olarak kullanılan diğer uygulamalardır [44-49].

Umut verici gelişmelerin ardından ilaç tasarımında günümüze kadar aktif kullanılan programlar ile kullanım amaçlarına ve ilaç şirketleriyle yapay zekâ şirketlerinin ortak çalışmalarını içeren tablolar (Tablo 1 ve Tablo 2) aşağıda verilmektedir.

Tablo 1. İlaç tasarımında aktif kullanılan programlar ve kullanım amaçları [41]

Programlar	Kullanım Amaçları
PPICurator (https://ppicurator.hupo.org.cn)	Bu araç, geniş kapsamlı PPI'leri araştırmak için kullanılmaktadır.
BioRAT (http://bioinfadmin.cs.ucl.ac.uk/biorat/docs/index)	Tüm metin madenciliği çalışmaları için kullanılmaktadır.
DeepChem (https://github.com/deepchem/deepchem)	Phyton tabanlı, AI sistemi kullanan MLP modelleme yöntemiyle yeni ilaç adayı bulunmaktadır.
DeepNeuralNetQSAR (https://github.com/Merck/DeepNeuralNet-QSAR)	Phyton tabanlı bu sistem, hesaplama araçlarıyla birlikte, bileşiklerin moleküler aktivitesinin saptanmasına yardımcı olmaktadır.
DeepTox (www.bioinf.jku.at/research/DeepTox)	12.000 ilacın toksisitesini tahmin eden bir yazılımdır.
GeneWays (http://geneways.genomeleft.columbia.edu)	Bu araç, biyolojik yolları saptamaktadır.
PotentialNet (https://pubs.acs.org/doi/full/10.1021/acscentsci.8b00507)	NN'leri kullanarak ligandların bağlanma afinitelerini tahmin etmektedir.
ORGANIC (https://github.com/aspuru-guzik-group/ORGANIC)	İstenen özelliklere sahip moleküllerin üretilmesine yardımcı olan bir molekül oluşturma aracıdır.
AlphaFold (https://deepmind.com/blog/alphafold)	Proteinlerin 3 boyutlu yapılarını tahmin etmektedir.
CancerDR (http://crdd.osdd.net/raghava/cancerdr/)	148 antikanser ilacı ve bunların yaklaşık 1000 kanser hücre hattına karşı etkinliğini kapsayan kanser ilaç direnci veri tabanıdır.
BRENDA (http://www.brenda-enzymes.org)	Enzim ve enzim-ligand veri tabanıdır.
UniProt (http://www.uniprot.org)	Proteinlerle ilgili bilgi merkezidir.
InterPro (http://www.ebi.ac.uk/interpro)	Protein bölgesiyle ilgi verileri bulunduran bilgi merkezidir.
ChEMBL (https://www.ebi.ac.uk/chembl)	Küçük ilaç benzeri moleküllerin birincil yayınlanmış literatürüne ve bunların biyoaktif özelliklerine dayanan veri tabanıdır.
TDR targets (http://tdrtargets.org/)	İhmal edilen tropikal hastalıklar için geliştirilmiş kemogenomik bir veri tabanıdır.
MATADOR (http://matador.embl.de/)	Protein-kimyasal etkileşimlerle ilgili veri tabanıdır.

Tablo 2. İlaç şirketleriyle yapay zekâ şirketlerinin ortak olarak yürüttüğü çalışmalar [41]

İlaç Şirketleri	Yapay Zekâ Şirketleri	Ortak Çalışmaları
Pfizer	IBM Watson	İmmüno onkolojide yeni ilaç hedeflerini, kombinasyon terapilerini ve hasta seçim stratejilerinin belirlenmesi
Pfizer	XtalPi	Moleküllerin mekanik ve kimyasal özelliklerinin yanı sıra 3B yapıyı tahmin etmek için bulut bilgi işlem çerçevesiyle kuantum mekaniği (QM) ve makine öğrenimi algoritmalarının kullanılması, Moleküllerin proteinlerle bağlanmasının incelenmesi

Tablo 2 (devamı). İlaç şirketleriyle yapay zekâ şirketlerinin ortak olarak yürüttüğü çalışmalar [41]

İlaç Şirketleri	Yapay Zekâ Şirketleri	Ortak Çalışmaları
BAYER	Exscientia	Exscientia'nın, onkolojik ve kardiyovasküler hastalıkları tedavisi ve olası ilaç adayları için en yeni öncü yapılarını optimize etmek amacıyla Centaur Chemist™ AI ilaç bulma platformunu kullanması
BAYER	Sensyne Health	Kardiyovasküler hastalıklar için Sensyne Health'in tescilli, klinik yapay zeka teknoloji platformunu kullanarak yeni tedavilerin geliştirilmesi
NOVARTIS	Microsoft	Yeşil kimyanın, hücre ve gen tabanlı tedavilerin, görüntü bölümlendirmenin, akıllı ve özelleştirilmiş tedavi sunumunun araştırılması
NOVARTIS	IBM Watson	Meme kanseri hastaları için tedavi çalışmalarının yürütülmesi
SANOFI	Exscientia	Diyabet ve komorbiditeleri için bispesifik küçük molekülün bulunması
AstraZeneca	BenevolentAI	İdiyopatik pulmoner fibrozis ve kronik böbrek hastalıklarının tedavisi için sinir ağları çerçeveleri temelinde yeni tedavilerin geliştirilmesi
Janssen	BenevolentAI	BenevolentAI, Janssen'in daha önce Parkinson hastalarında Bavisant'ın faz IIb denemelerinde klinik veri sağlamak için kullanılan yenilikçi klinik faz ilaç adaylarını geliştirmesi
Takeda	Numerate	Gastroenteroloji, onkoloji ve merkezi sinir sistemi temelli bozukluklarda tedavi için ilaç moleküllerinin hedef yerinin belirlenmesi
Eli Lilly	Atomwise	Henüz bilinmeyen protein hedeflerine yönelik ilaçların geliştirilmesi
Roche	OWKIN	ML ağlarına dayalı klinik deneyler ve ilaç tasarımının yapılması

Tarihsel Gelişim

1950 yıllarında AI kavramının doğuşunun ardından 1960'lı yıllarda QSAR üzerine ilk çalışmalara başlanmış ve ilaç tasarımı için kullanılmıştır [50]. 1990'lardan önce, ilaç tasarımı modeller tasarlamak için olağan hesaplama yöntemleri, yani doğrusal regresyonlar kullanılmıştır. Bu erken çalışmalarda, modelleme için kullanılan kimyasal tanımlayıcılar, atomik tip ve parçalı tanımlayıcılar gibi kimyasal ve yapısal özellikleri tanımlamakla sınırlı olmaktadır [51].

Yapay zekanın ilaç tasarımı ve geliştirmesindeki ilerlemesi, ilk olarak, son teknoloji kimyasal tanımlayıcıların yani moleküler parmak izlerinin oluşturulmasıyla ve eğitim setlerinde tanımlayıcı kategorileri önemli ölçüde arttıran topolojik tanımlayıcılarla hızlanmıştır [12,52-54]

İlaç tasarımı ilk olarak 1989'da sinir ağlarının kullanılması üzerine çalışılmıştır [55]. Kabul edilen ilk yaklaşım, yani yapay sinir ağı (YSA), değişkenlerin seçim sürecini temel almaktadır [56,57]. YSA'lar, değişkenler arasında doğrusal olmayan bağlantılar kurmak ve biyolojik etkinliklere odaklanmak için çok iyi bir makine öğrenimi yaklaşımını temsil etmektedir [58]. YSA'lar da dahil olmak üzere çeşitli makine öğrenimi yaklaşımlarını kullanan modern hesaplama modelleri, sağlam bilgisayarlar edinmiş ve 1990'larda bilgisayar donanımlarındaki gelişmelerden yararlanmış [59]. Bu gelişmelerle birlikte 1980 yıllarında YSA ile birlikte temel olarak DL kavramı da tanıtılmıştır [60].

1990'lardan 2000'lere kadar modelleme değerlendirmelerinde, doğrusal regresyon kullanmak yerine, doğrusal olmayan modelleme algoritmaları, yani KNN, RF ve SVM temelinde oluşturulan gelişmiş ML algoritmaları geleneksel olarak kullanılmıştır. Yine aynı dönemde, model doğrulamanın önemi vurgulanmış ve modellemenin olmazsa olmaz bir parçası olarak kabul edilmiştir. Yapay zekanın ilerlemesiyle birlikte, 2000'li yılların başında QSAR modellemelerinde köklü değişimler yaşanmıştır. Bu gelişmenin ardından hesaplama kapasitesi ve modellemelerde verilerin donanım ve erişilebilirliğinde önemli derecede iyileşmeler yaşanmıştır [59]. Merck tarafından desteklenen QSAR çalışmaları sırasında, ilaç tasarımı için kullanılan DL algoritmaları, kullanılan ML algoritmalarından daha güçlü ve üstün bulunmuştur [61].

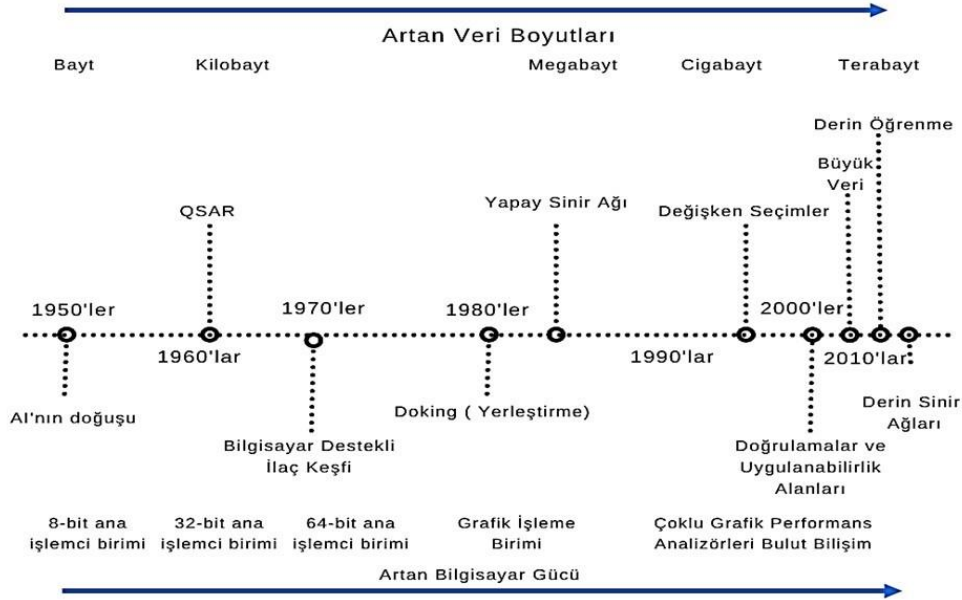
Donanımdaki hızlı gelişmeler, yani 2010'larda grafik işlem birimi olan GPU'ların ve bulut bilişimin kullanımı, NN modelleme araştırmalarını doğrudan teşvik etmiştir [60]. Çok sayıda gizli katmana sahip DNN'ler kurulmuş ve 2015 yılında gelişmeleriyle çığır açan DL ile ilgili makale yayınlanmıştır [54]. Yine DeepTox (DNN'lere dayalı bir hesaplamalı toksiklik modeli), diğer ML yaklaşımına dayalı modelleri geride bırakmıştır [62]. DNN modelleri yüksek başarı oranı ile büyük veri kümelerini modellemek ve anlamlı özellikleri seçmek için DL yaklaşımları avantajlarından dolayı ilaç tasarımıyla uygulanmaya ve geliştirilmeye devam etmiştir.

Daha sonra, Pande ve ark., 2018 yılında moleküler makine öğrenimi için kapsamlı bir ölçüt olan MoleculeNet'i tanıtmışlardır [63]. MoleculeNet, bir AI yazılım aracı olan DeepChem'e bir verileri bölme mekanizmaları kitaplığı sunmakta ve tüm algoritmaları, çeşitli veri bölme seçenekleriyle değerlendirmektedir. Aynı zamanda MoleculeNet, veri seti koleksiyonlarındaki çeşitli özelliklerden yararlanarak makine öğrenimi algoritmalarının kıyaslanmasını yapmaktadır [41].

Son zamanlarda, Allen ve ark., ATOM Modelleme Boru Hattını (AMPL) geliştirerek ilerlemelere büyük katkıda bulunmuşlardır [64]. AMPL, silico ilaç tasarımıyla ilerlemede, modeller oluşturma ve paylaşmada açık kaynaklı, esnek ve genişletilebilir bir yazılım boru hattıdır. Model eğitimi ve değerlendirmesinde, veri setlerini bölmede çeşitli seçenekleri desteklemektedir. Ancak bölme, iç içe küme çapraz doğrulamaya benzer bir süreci takip etmektedir [65].

Veri oluşturma; eğitim, model seçimi (yani doğrulama) ve performans değerlendirmesi (yani test) olmak üzere veri setleri birbiriyle örtüşmeyen üç bölmeye bölünmüştür. AMPL ise, eğitim verilerinden alışılmışın dışında kimyasal uzaya kadar genelleşen modeller oluşturma önündeki engellere çeşitli yaklaşımlar sunan çok sayıda veri seti bölme algoritması sağlamaktadır [41]. Rastgele bölmeler, Bemis-Murcko yapı iskelesi ayırma, Butina kümeleme ve parmak izi farklılığı tabanlı algoritma gibi DeepChem'de yer alan çeşitli yöntemlere yardımcı olmaktadır [63].

Artan veri boyutu ve bilgisayar gücüyle yapay zekanın ilaç tasarımındaki tarihsel ilerlemesi Şekil 4'te verilmektedir.



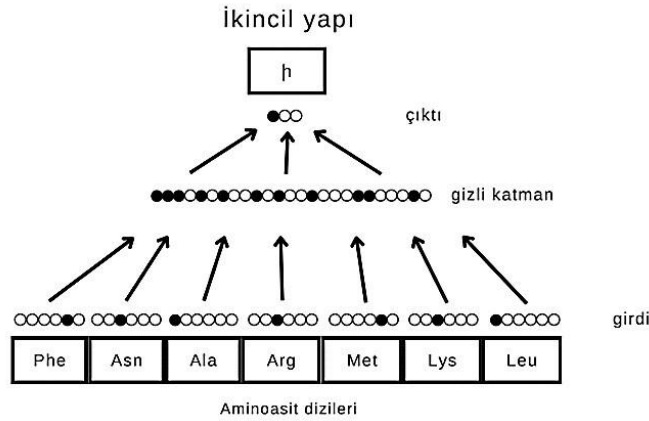
Şekil 4. Artan veri boyutu ve bilgisayar gücüyle yapay zekanın ilaç tasarımındaki tarihsel ilerlemesi [59]

Protein Katlanmasının Tahmininde Yapay Zeka Uygulamaları

Hastalıkların çoğu, protein fonksiyonlarının düzensizliği ile ilişkilidir. Protein yapısı temelinde, protein hedefleri üzerinde bulunan küçük aktif bölgelerin keşfedilmesi ve bu bölgelerle etkileşim yapabilen aktif moleküllerin geliştirilmesi için çeşitli ilaç tasarımı yaklaşımları kullanılabilir. Ancak,

proteinlerin üç boyutlu (3D) yapısının deneysel olarak aydınlatılması çok maliyetli ve zaman alıcı bir işlemdir [41]. Bu nedenle bilgisayar destekli teknikler, proteinlerin 3 boyutlu yapısının değerlendirilmesinde önemli rol oynamakla birlikte, 3 boyutlu yapısının netliği tartışmalıdır [66]. Bu durumun proteinlerin çok sayıda uzaysal konumunun olmasından kaynaklandığı öne sürülmüştür. Bu nedenle, protein yapısı tahmin edilirken genellikle, tek boyutlu (1D) yapısal niteliklerin dizileri, protein ikincil (2°) yapısı, solventle erişilebilen yüzey alanı, iskelet bükülme açısı ve benzeri gibi bileşene ayrılarak incelenmektedir [67].

Büyük protein dizilerinde veri setlerinin bulunması nedeniyle, AI tabanlı teknoloji, proteinlerin yapısal özelliklerini öngörmek için kapsamlı bir şekilde kullanılmıştır. İlk olarak, Qian ve ark. (1988) proteinlerin ikincil yapısını öngörmek için doğrusal olmayan sinir ağı algoritmasını kullanmışlardır. Kullanılan algoritmanın diyagramı Şekil 5'te verilmektedir [68].



Şekil 5. Ağ algoritmasının bir diyagramı [68]; Standart ağ 13 bitişik amino asitlik bir diziyi temsil eden 21 birim/grup ile 13 girdi grubuna sahiptir (yalnızca 7 girdi grubu ve 7 birim/grup gösterilmektedir). Giriş katmanından gelen bilgi, merkezi amino asit için ikincil yapı öngörüsünü temsil eden 3 çıkış biriminde bir aktivite paterni üretmek için "gizli katman" birimlerin bir ara katmanı tarafından dönüştürülmektedir

Kullanılan algoritma %64.3 başarı oranı ile diğer algoritmalarından daha başarılı bir sonuç vermiştir. Proteinlerin 3 boyutlu yapısını tam olarak öngörmek hala uzak bir durum olsa da DL algoritması, bu alanın ilerlemesinde önemli bir gelişim göstermiştir. Son zamanlarda DL yaklaşımları, proteinlerin 2° yapısını, solventle erişilebilen yüzey alanlarını, α -karbon atomları dihedral açısını ve iskelet bükülme açılarını öngörmek için kullanılmıştır [68-71].

Qi ve ark. DNN'leri çok görevli bölgesel protein yapısı tahmini geliştirmek için bir farklılaştırıcı olarak kullanmış ve protein çözücü erişilebilirliğini ve 2° yapılarını öngörmek ve evrimsel özellikleri kullanmak için tek bir sinir ağını yönetmişlerdir. Wang ve ark. koşullu rastgele feld (CRF) ile birlikte kolay sinir ağını birleştirmiş ve proteinlerin 2° yapısını öngörmek için bir DeepCNF tekniği geliştirmişlerdir. Bu teknikle, tahmin hassasiyeti %84'e yükselterek çözücü erişimi, temas numarası ve düzensiz alanlar dahil olmak üzere proteinlerin yapısal özelliklerini öngörmek için kullanılabilir hale gelmiştir. Ayrıca, Jo ve ark. protein yapı tanımlama performansını oldukça artıran bir derin öğrenme ağı tekniği (DN katlama) oluşturmuştur [69,71,72]. Bu teknik, belirli bir protein çiftinin karşılıklı gelen kısımlarının aynı yapısal katla ilişkili olup olmadığını tam olarak öngörebilmektedir. Protein yapı tahmini, bir protein ana yapısının fonksiyonlarını belirlediğinden, amino asit dizisinden bir proteinin 3D şeklini belirlemek için kullanılabilir [72-74].

2018 yılında DeepMind, protein yapısını tahmin etmek için AlphaFold'u kullanmıştır. AlphaFold, bilinmeyen kıvrımları doğru bir şekilde tahmin etme yeteneğine sahiptir. AlphaFold'un daha yüksek doğruluk oranı, proteinlerin yapıları arasındaki mesafe tahminlerinin doğruluğundan kaynaklanmaktadır [75,76]. 2020 yılına gelindiğinde yayımlanan bir diğer çalışma olan AlphaFold 2 ise, önceki çalışmadan

farklı olarak daha gelişmiş bir yapay zekâ algoritması kullanmıştır. AlphaFold 2, proteinlerin üç boyutlu yapılarını daha doğru bir şekilde tahmin edebilmekte ve önceki çalışmada karşılaşılan bazı zorlukları aşabilmektedir. AlphaFold 2'nin başarısı, daha önce elde edilen diğer yapısal verilerle karşılaştırıldığında %25 oranında daha doğru sonuçlar vermesiyle ölçülmüştür. Bu teknoloji, ilaç tasarımı alanında büyük bir potansiyel sunmaktadır ve biyolojik sistemlerdeki protein yapısının anlaşılmasına önemli katkılar sağlamaktadır [77].

Protein-Protein Etkileşimlerinin Tahmininde Yapay Zeka Uygulaması

Protein-protein etkileşimleri (PPIs) biyofiziksel kimyanın çekirdeğidir. Biyokimyasal süreçleri içeren biyolojik fonksiyonların çoğu, PPIs tarafından yakından kontrol edilmektedir [78-81]. PPIs pek çok biyolojik aktivitede önemli olmakla birlikte, aynı zamanda çok sayıda hastalıkla da doğrudan ilişkilidir [82]. Konvansiyonel hedeflerle (örneğin, nükleer reseptörler, G-protein eşli reseptörler (GPCR'ler), iyon kanalları, kinaz vb.), hedef alanı genişletmek ve küçük moleküllü ilaçların geliştirilmesine yardımcı olmak için PPIs'den yararlanılmaktadır. Bu nedenle, PPIs arayüz bölgesinin ayrıntılı bir şekilde anlaşılması, protein fonksiyonunun açıklanmasına katkıda bulunmasının yanı sıra, protein-protein kompleks yapılarına ve ilgili hastalık tedavisine dayalı ilaç tasarımı için de kritik önem taşımaktadır [83]. Bununla birlikte, PPIs için mevcut deneysel ölçüm yöntemlerinin yüksek maliyet, uzun zaman gereksinimleri, veri setinde yüksek yanlış pozitif ve negatif oranlar gibi dezavantajları nedeniyle PPIs hakkında bilgi çok sınırlı olmaktadır [84]. Bu nedenle, PPI arayüz tahmini için birçok hesaplama yöntemi oluşturulmuştur [85,86].

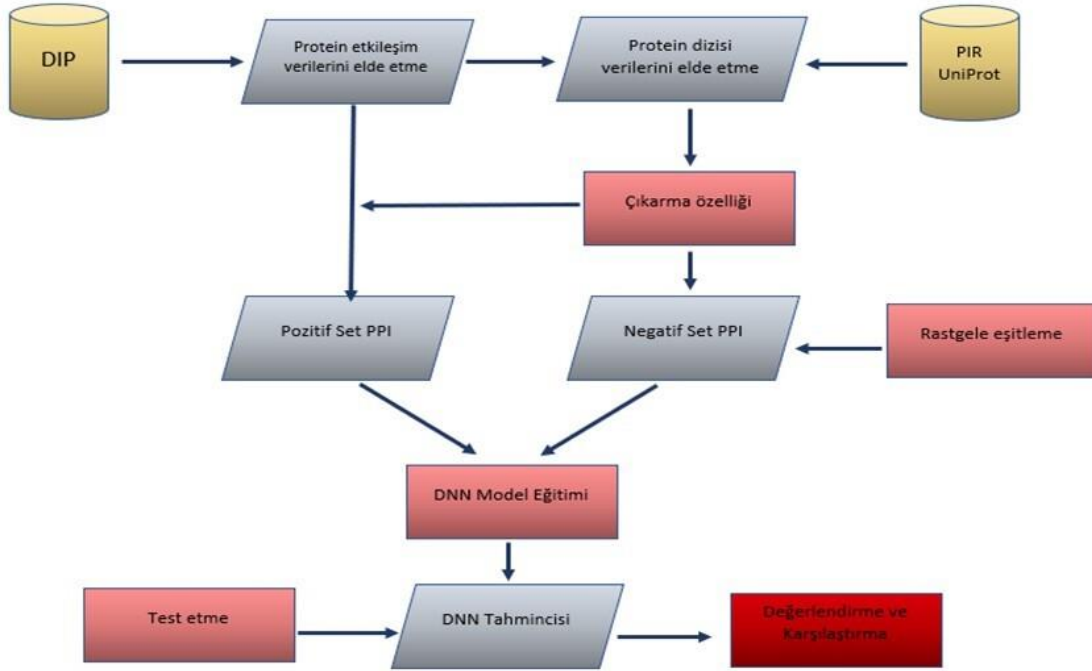
Mevcut PPI tahmini, temel olarak yapıya ve dizilere dayalı iki kategori içermektedir. Bunların arasında, çoğu PPI arabirimi korunaklı olduğundan, protein şablon yapısına dayalı yöntem daha basit ve daha güvenilir sonuç vermektedir. Örneğin, Maheshwari ve ark., zayıf homolog şablon yapısından PPI kalıntılarını tanımlamak için kullanılabilen, şablon yapısına dayalı eFindSitePPI tahmin yöntemini geliştirmişlerdir [87]. Bu yöntem hem deneysel protein yapısında hem de silico ile üretilen protein yapısında yüksek tahmin doğruluğuna sahiptir. Etkileşen iki proteinin 3D yapıları bilindiğinde, PPI arayüzü, tamamlayıcı prensibe dayalı protein-protein kenetleme yöntemleri kullanılarak tahmin edilebilmektedir [41,88].

Yapı tabanlı tahmin algoritmaları dizi tabanlı yöntemlerden daha iyi performans gösterse de bu yaklaşım bilinen protein yapılarının sayısı ve kalitesiyle sınırlıdır. Örneğin, şu anda bilinen bakteri, maya veya insandaki PPI'lerin %80'i hakkında çok az yapısal bilgi bulunmaktadır [89]. Protein dizisi verilerindeki logaritmik büyüme ile AI, dizi tabanlı yöntemler kullanarak PPI'leri tahmin etmede önemli ilerleme kaydetmiştir. 2016 yılında Du ve ark., protein dizilerinden Fisher fraksiyonu özelliklerini çıkarmak için etkileşimli profil gizli Markov modellerini (ipHMM'ler) kullanmışlardır. Bununla birlikte, ilk kez proteinlerin kalıntı-kalıntı temas matrisini tahmin etmek amacıyla bir DNN modeli oluşturmak için, seyrek otomatik kodlayıcı (SAE) kullanılmıştır. DL modelinin genel tahmin doğruluğu, geleneksel ML modelinden %15 daha yüksek olarak %80,82 olarak bulunmuştur [85]. Bu gelişmelerin ardından, Du ve ark. PPI'leri tahmin etmek için DeepPPI adlı bir DNN yöntemi tasarlamışlardır. Bu yöntemin işleyişi Şekil 6'da verilmektedir. Bu yöntemle katman katman soyut yöntemle protein dizilerinden yararlı özellikleri otomatik olarak belirleyerek, protein dizilerinin özgüllüğü öğrenilir hale gelmiştir. Sonuç olarak bu yöntemle, test veri setindeki doğruluk %92,50'ye kadar çıkmıştır [90].

Daha sonra Zeng ve ark. onaylanmış protein komplekslerini öngörmek için dizi yöntemini kullanan 'Complex Contact' web sunucusunu (<http://raptorx2.uchicago.edu/ComplexContact/>) geliştirmiştir. Sunucu önce proteinler arasındaki dizi homolojisini araştırmakta ve ardından iki adede kadar çoklu dizi hizalaması (MSA) oluşturmaktadır. Bundan sonra, proteinler arası teması tahmin etmek için derin evrişimli sinir ağı (ResNet) ve bununla birlikte evrimsel analiz yaklaşımları kullanılmaktadır [91]. Bu yöntem, protein dizisi homolojisi gereksinimini azaltmakta ve tahmin doğruluğunu önemli ölçüde artırmaktadır.

Son zamanlarda, Xie ve ark., PPI bölgesini tahmin etmek için bir CNN önermiş ve pozitif örnekleri iyileştirmek için kalıntı bağlama eğilimi kullanmıştır. Bu yöntem ile iyileştirilmiş veri setinde elde edilen sonuçlar %91.2 başarı oranını göstermiştir. Ayrıca bu yöntem, büyük bağlanma eğilimi olan numunelerde çok daha iyi sonuçlar vermiştir. Bu yöntem aynı zamanda, kalıntı atomları arasındaki

mesafe ile belirlenen pozitif örneklerde, oldukça büyük bir oranda yanlış pozitif PPI bölgelerinin varlığını ortaya koymuştur [92].



Şekil 6. Protein-protein etkileşimleri tahmini için derin sinir ağlarının çerçevesi [90]

Yapay Zeka ile Sanal Tarama Uygulamaları

Sanal tarama (VS), ilaç hedeflerine (genellikle proteinler) bağlanan aktif küçük molekülleri tanımlamak amacıyla yapılan hesaplamalı ilaç tasarımının ana yöntemlerinden biridir [93]. Bu amaçla, VS kaçınılmaz olarak bilgisayar programları kullanarak bileşiklerin geniş kapsamlı veri tabanlarını değerlendirmektedir [94]. İlaç geliştirmenin ilk aşamalarında uygun olmayan iskeletler içeren bileşikleri filtrelemek ve yeni hedefler bulmak için etkili bir yöntem olarak kullanılabilirdiğinden, yüksek maliyetli ve düşük başarı oranı gibi problemleri olan yüksek verimli taramaya (HTS) yardımcı olmak için önemli bir araç haline gelmiştir [49].

Sanal taramada ligand bazlı sanal tarama (LBVS) ve yapı bazlı sanal tarama (SBVS) olmak üzere başlıca iki yöntem kullanılmaktadır. LBVS, yüksek biyoaktivitelere sahip diğer ligandları tahmin etmek ve tanımlamak için aktif ligandlar arasındaki kimyasal ve uzaysal benzerlikleri ve fizikokimyasal analizi kullanan aktif ve inaktif ligandların deneysel verilerine dayanmaktadır. LBVS, 3D protein yapısal bilgisine dayanmadığından, bu yöntem öncelikle hedef yapı eksik olduğunda veya elde edilen yapısal doğruluk düşük olduğunda aktif ligandların tahmini için kullanılmaktadır [95].

Geleneksel makine öğrenimi yöntemlerinden olan RF, SVM, DT, KNN ve NB gibi ML yöntemleri, LBVS'de yaygın olarak kullanılmaktadır. Bu yöntemler, öngörülen hedef oranını başarılı bir şekilde iyileştirmekte ve yanlış hedef oranını azaltmaktadır [96]. Bununla birlikte, kimyasal alandaki aktif bileşiklerin seyrekliği (10^{60} teorik bileşik) ve sınırlı eğitim seti, geleneksel ML'lerin LBVS'de istenen başarıya ulaşma noktasında bazı kısıtlamalara sahip olmasına neden olmaktadır. Otomatik çıkarma ve katman katman öğrenme özelliklerine sahip DL'nin ortaya çıkışı, LBVS'nin daha da geliştirilmesine olanak sağlamıştır [97].

Xiao ve ark. açık kaynaklı TensorFlow programı ile büyük verilerin bir DNN prototipini oluşturmuş ve bunu kapsamlı bileşik kitaplıklarını taramak için bir LBVS cihazı olarak kullanmışlardır. 2015'ten önce PubChem'de 95 milyon antikanser hedef inhibitörü tarayan model, 2015'ten sonra inhibitörlerin %50'sini yalnızca %0.01-%0.09'luk bir yanlış pozitif oranıyla tanımlamış ve bu da DNN'lerin LBVS'deki uygulama potansiyelini tam olarak göstermiştir [98].

SBVS genel olarak, hedefin üç boyutlu yapısının hesaplamalı veya deneysel yöntemlerle aydınlatılmasından sonra kullanılır. Bu yöntem esas olarak olası aktif ligandlar ve bağlanma yeri kalıntıları arasındaki etkileşimleri araştırmak için kullanılır ve genellikle LBVS yöntemlerinden daha iyi tahmin performansı gösterir [99]. Bununla birlikte, SBVS tabanlı yöntem, protein yapı sayısında üstel büyüme ve son derece karmaşık protein konformasyonları sorunuyla karşı karşıyadır. Problemi çözmenin anahtarı, ligandlar ve hedefler arasındaki ilişkiyi doğru bir şekilde tanımlamaktır. Protein-ligand bağlanmasındaki ve yapısal verilerdeki artışla birlikte, SBVS'nin daha da geliştirilmesi için yeni araçlar sağlayan AI teknolojisi kullanılarak protein-ligand etkileşimlerini tanımlamak mümkündür [100]. SVM, Boosting ve RF gibi geleneksel makine öğrenimi yöntemleri, ligandlar ve hedefler arasındaki moleküler etkileşimlerin doğrusal olmayan etkileşimlerini açıklayabilmiş ve SBVS'nin sonuçlarını iyileştirmek için başarıyla kullanılmıştır. Bir dizi protein-ligand bağlanma afinite sınıflandırıcısı da oluşturulmuştur [101].

Bu özelliklerin yanı sıra, geleneksel makine öğrenimi yöntemleri, büyük ölçekli uygulamaların elde edilmesini zorlaştıran ve hatta özellik çıkarma sürecinde ilgili bilgilerin kaybolmasına yol açan oldukça zahmetli manuel tanıma ve özellik çıkarma sorununa sahiptir. DL yöntemlerinin ortaya çıkmasıyla bu sorun başarılı bir şekilde çözülmüştür [102]

Pereira ve ark. ilk olarak 2016 yılında SBVS'nin performansını iyileştirmek için derin öğrenmeyi kullanmış ve DeepVS adı verilen geliştirilmiş bir SBVS yöntemi oluşturmak için derin evrişimli sinir ağları (DCNN) modelini kullanmışlardır. Bu yöntem, moleküler yerleştirmenin sonucunu DCNN'nin girdisi olarak almakta ve temel verilerden bileşik atom tipi, atomik kısmi yük ve atomik mesafe gibi ilgili özellikleri otomatik olarak öğrenebilmekte ve çıkarabilmektedir [103].

Ardından Ferrero ve ekibi, Open Targets platformundan elde edilen gen-hastalık ilişkisi verilerine dayanarak terapötik hedeflerin belirlenmesi üzerine araştırma yapmıştır. Bu çalışmada, belirli bir genin bir ilaç hedefi olma ihtimalini tahmin etmek için makine öğrenimi modelleri, RF, SVM, NN ve Gradyan Artırma Makinesi gibi yöntemler kullanılmıştır [104].

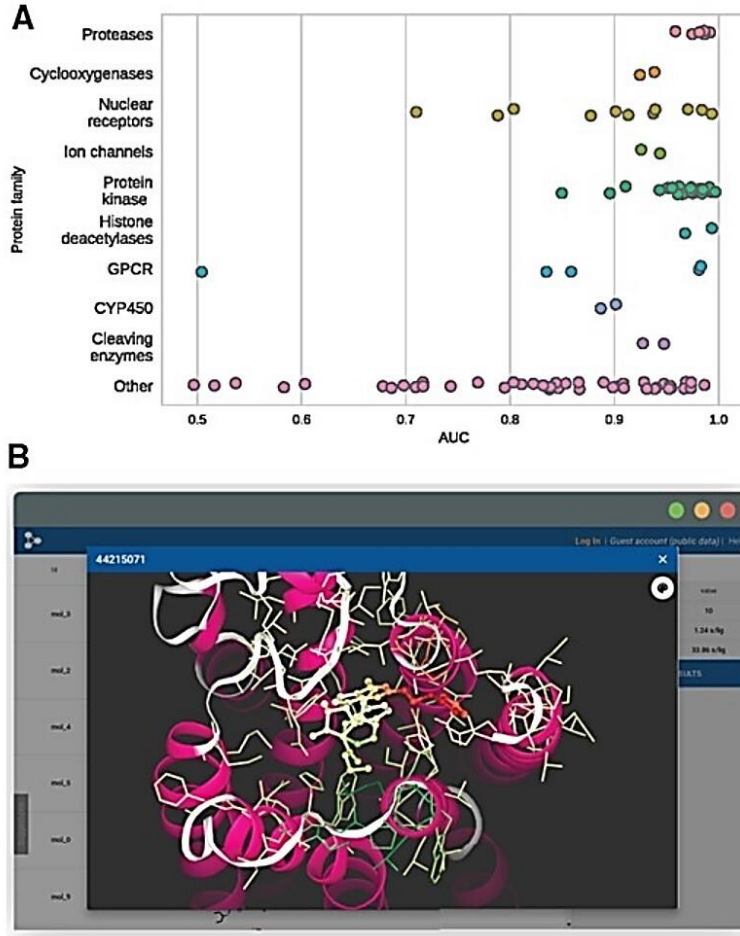
Donner ve ekibi ise, gen ekspresyon verilerinden oluşan moleküllerdeki fonksiyonel grupların yapısını taşıyan ve bu moleküllere en benzer bileşikler tespit etmek için bir yöntem geliştirmişlerdir. Bu yöntem, derin öğrenme ağlarının avantajlarını kullanarak ve verilerin bütünlüğünü koruyarak bileşiklerin karşılaştırılmasını sağlamaktadır. Bu yöntem, bileşiklerin yapısal olarak farklı olmasına rağmen, paylaşılan terapötik ve biyolojik hedefleri olan ilaçları belirleyebilmekte ve böylece bileşikler arasında daha önce bildirilmemiş fonksiyonel ilişkileri ortaya çıkarabilmektedir. DeepCodex adlı bir çevrimiçi platformda kullanıma sunulmuş olan bu yöntem, farklı yapısal özelliklere sahip bileşikler arasında bile paylaşılan biyolojik hedefleri tespit ederek, bileşikler arasında önceden bilinmeyen fonksiyonel ilişkileri ortaya çıkarabilmektedir [105,106].

Entegre Ağ Tabanlı Hücrel İmzalar Kitaplığı (LINCS, Library of Integrated Network-Based Cellular Signatures), gen ekspresyonu, protein aktivasyonu ve kimyasal bileşiklerin hücreler üzerindeki etkileri gibi biyolojik süreçlerin geniş bir yelpazesini kapsayan büyük ölçekli bir veritabanıdır. Bu veritabanı, ilaç tasarımı ve hastalık mekanizmalarının anlaşılması gibi birçok farklı araştırma alanında kullanılmaktadır [107]. Xie ve ark., LINCS'nin L1000 veritabanındaki transkriptom verilerini kullanarak ilaç-hedef etkileşimlerini öngörmek için bir derin öğrenme yöntemi ile bir çalışma gerçekleştirmişlerdir. Elde edilen sonuçlar, yöntemin doğruluğunun %98 oranında olduğunu bildirmişlerdir. Ayrıca bu çalışma, önceden toplanan verilerin derin öğrenme yaklaşımlarıyla yeniden kullanılmasının ilaçların tasarlanmasındaki önemini ortaya koymuştur [108,109].

Ayrıca, Skalic ve ark. büyük ölçekli aktif ve inaktif bileşikler sınıflandırabilen ve GPU hızlandırmasını gerçekleştirebilen BindScope adlı bir web uygulaması oluşturmak için DCNN modelini kullanmaktadır. Program, kullanıcıların aynı anda yüzlerce bileşiği görüntülemesine ve sonuçları etkileşimli olarak görselleştirmesine olanak tanımaktadır. BindScop uygulamasıyla ilgili örnek bir çalışma Şekil 7'de verilmektedir [110].

Son zamanlarda, Mendolia ve ark. IC₅₀ değerlerini kullanarak Cyclin-Dependent Kinaz 1 (CDK1) protein hedefine karşı aday bileşiklerin biyolojik aktivitesini tahmin etmek için moleküler parmak izleri üzerinde eğitilmiş yeni bir CNN yöntemini sunmuş ve bu çalışma için teknik geliştirmişlerdir. Bunlar; tek parmak izleri sınıflandırıcıları, önemli birleştirmeler için 1D ve 2D CNN sınıflandırıcıları, 1D CNN

sınıflandırıcılarına dayalı bir sınıflandırma ve VS'deki her farklı parmak izi tipinin çıktısını derin sinir ağına dayalı olarak sınıflandıran tekniklerdir [111].

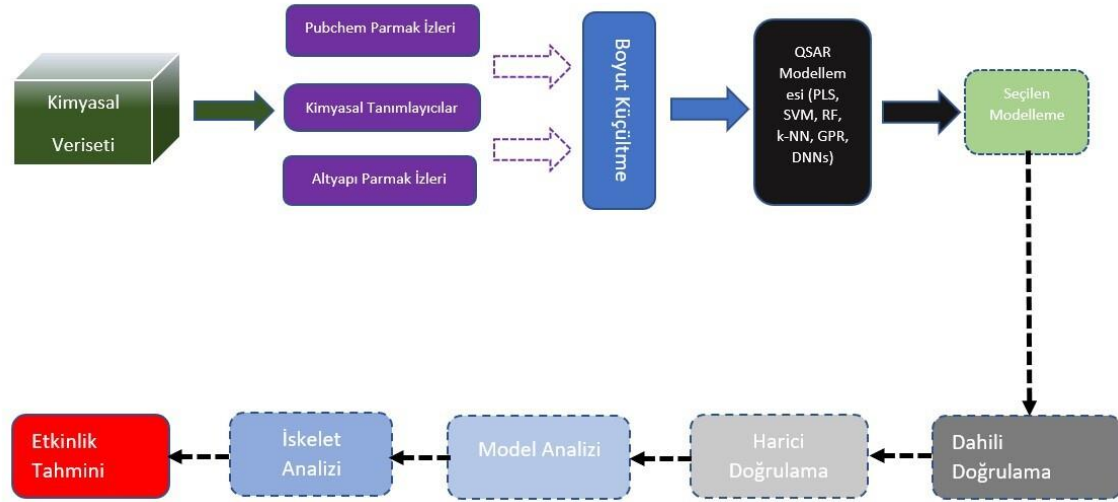


Şekil 7. BindScope web uygulamasındaki örnek çalışma [110]; A. Reseptör başına AUC değeri 5 kat çapraz validasyon kullanılarak değerlendirilmesi. B. Kenetlenmiş pozun, hangi ligand alanlarının olumlu katkıda bulunduğunu gösteren karşılık gelen ayırıcı lokalizasyon haritası ile görselleştirilmesi

Kantitatif Yapı-Aktivite İlişkisinde Yapay Zeka Uygulaması

Kantitatif yapı-aktivite ilişkisi (QSAR), fizikokimyasal özellikler, kimyasal yapı ve biyolojik aktiviteler arasındaki kantitatif haritalama ilişkisini oluşturmak için matematiksel yöntemler kullanılmaktadır [112-114]. Bu ilişkiler kurulduktan sonra, yapısal olarak çeşitli moleküler veri tabanını otomatik olarak taramak ve ardından laboratuvarında sentez ve analiz için en umut verici bileşikleri seçmek kolaylaşmaktadır. Böylece, deneysel kaynaklardan büyük ölçüde tasarruf edilebilmekte, deneyin başarısızlık oranı düşürülebilmekte, istenen özelliklerde yeni moleküllerin geliştirme süreci de hızlandırılabilir. QSAR yöntemi temel olarak veri toplama ve ön işleme, moleküler tanımlayıcıların üretimi ve seçimi, matematiksel bir modelin oluşturulması, model değerlendirme ve yorumlama ile model uygulamasını içermektedir. Bu uygulamanın iş akışı Şekil 8'de verilmektedir.

Yapay zekâ, kimyasal yapı ile biyolojik aktivite arasındaki ilişkinin sağlam bir modelini etkili bir şekilde oluşturabildiğinden, QSAR araştırmasının önemli bir parçası haline gelmiştir. 1990 gibi erken bir tarihte, Aoyama ve ark. sinir ağlarını (NN'ler) QSAR analizinde kullanılmıştır [115]. Daha sonra, RF, Boosting, GP, KNN, DL, Cubist ve SVM gibi çeşitli geleneksel ML yöntemleri de QSAR modellerini oluşturmak için yaygın olarak kullanılmıştır [116,117].



Şekil 8. QSAR modellemesinin iş akışı

Veri kümelerindeki sürekli artışla birlikte, QSAR modeli giderek daha karmaşık hale gelmekte ve geleneksel makine öğreniminde kullanılan sığ sinir ağı yönteminin büyük veri kümelerinin ihtiyaçlarını karşılaması zorlaşmıştır. DL'nin ortaya çıkışı, büyük veri kümeleri sorununu çözmek için yeni bir yol sağlamıştır. 2012'de Dahl'ın ekibi, çok görevli DNN'ler, Gauss süreci regresyonu (GPR) ve gradyan destekli makinelerden (GBM'ler) oluşan entegre bir model kullanarak 2012 Merck Kaggle Molecular Activity Challenge'ı kazanmıştır [61]. Bu çalışmayla, bileşik aktiviteyi tahmin etmede yeni bir dönem başlatan büyük bir veri setinde QSAR problemini çözmek için DL ilk kez kullanılmıştır. 2014 yılında Dahl ve ark., bir bileşiğin biyolojik ve kimyasal özelliklerini moleküler yapısından doğrudan tahmin edebilen çok görevli bir DNN geliştirmiştir [118]. Bu çalışmalara kadar Merck, DNN'leri ve RF'yi karşılaştırmak için birçok QSAR veri seti kullanmış ve 15 adet veri setinin 11 veri seti, DNN'lerde rastgele ormana göre daha iyi performans gösterirken, ikinci değerlendirmede, zamana dayalı bölme stratejisiyle 15 veri setinin 13'ü DNN'lerde rastgele ormana göre daha iyi performans göstermiştir [61].

Çok görevli QSAR modelleri, bileşikleri biyolojik etkilerinin yanı sıra deneysel bilgilere dayalı olarak kategorize edebilmektedir. Tenorio-Borroto ve ark., YSA temelinde çok görevli bir QSAR modeli geliştirmişlerdir. Bu modelleme, çoğullama deneylerinden elde edilen veri setini %92'lik bir hassasiyetle sınıflandırmıştır [119].

İlerleyen süreçte, kimyasalların immünotoksitesini öngörmek amacıyla, yeni çok görevli QSAR-ANN modellerini oluşturmak için birkaç tanımlayıcının kullanılmasıyla ilgili çalışmalar yürütülmüştür [120]. Ayrıca bu modelleme yöntemiyle, TMC-207, Sudoterb, Delamanid gibi güvenli antibakteriyel ilaçların bulunmasına yönelik ilk çalışma Speck-Planche ve ark. tarafından yapılmıştır. Bu modelleme ile elde edilen ilaçlar, *Mycobacterium spp.* üzerindeki antibakteriyel aktiviteleri umut vericidir [121]. QSAR-ANN modeli hem eğitim hem de öngörü setleri için kesinliği yüksek öngörü sonuçları elde etmiştir. Aynı program ve kavramsal süreç takip edilerek, değişik elverişli çoklu görev QSAR modelleri geliştirilmiştir [119,121,122].

Ramsundar ve ark. çoklu görev sınıflandırma ikilemindeki görev sayısının etkisine odaklanmıştır. Yaptıkları çalışmada, 200'den fazla hedef (ağırlıklı olarak proteinler) ve 1.6 milyon bileşik için genişletilmiş bağlantı parmak izleri (ECFP) olarak tanımlanan 37.8 milyon deneysel veri noktası bir araya getirilmiştir [123]. Böylelikle, çoklu görev ağı, RF, lojistik regresyon ve tek görevli sinir ağlarını geride bırakmıştır.

Daha sonra Xu ve ark. çok görevli bir DNN'ye aktarılan bir görevin, eğitim süreci sırasında ek QSAR görevlerinden bilgileri "ödünç alabileceğini" gözlemlemiştir. Çeşitli çalışmalar, yüksek yorumlamalı QSAR modelleri geliştirmede DNN'lerin geleneksel makine öğrenimi tekniklerine göre geliştirilmiş performansını göstermiştir [124].

Yakın zamanda ise Zhao ve ark., modelleri ayrı ayrı oluşturmak yerine, çok sayıda benzer biyolojik hedefi bir arada düşünerek QSAR modellerini oluşturmuştur. Ayrıca (çok görevli öğrenme) MTL tabanlı modelin (tek görevli öğrenme) STL tabanlı modellerden daha yüksek performans gösterdiği sonucuna ulaşmışlardır. Birden çok görev üzerinde paylaşılan bilgilerin kullanılmasına bağlı olarak, MTL modelleri, temel modellerin düşük doğruluğa sahip olduğu durumlarda daha da belirgin başarı performansına sahip olmuştur. MTL modellerinin üstünlüğü, %5'lik anlamlılık düzeyi ile Student's T testi tarafından da desteklenmiştir [125].

ADME/T Farmakokinetik Özelliklerin Tahmininde Yapay Zeka Uygulaması

İlaç tasarlama ve geliştirme sürecinde, hedef veya öncü bileşikler elde edildikten sonra, bu bileşiklerin farmakokinetik özellikleri (absorpsiyon, dağılım, metabolizma ve atılım) ve toksisitesi (ADME/T) üzerinde bir dizi test ve değerlendirmeler yapılmaktadır [126].

Son yıllarda, milyonlarca aktif bileşik keşfedilmiş olmasına rağmen, FDA tarafından onaylanan yeni ilaç molekülü sayısı çok kısıtlı kalmaktadır. Bunun başlıca nedeni, bu aktif bileşiklerin ADME/T özelliklerinin ilaç olmak için gereken standartları karşılamamasıdır [127,128]. Raporlara göre, yan etkilerin ve düşük etkinin yanı sıra; bazı ticari nedenlerle birlikte farmakokinetik özelliklerin yetersizliği (%39) ve klinik öncesi toksisite (%11) ilaç geliştirme başarısızlığının ana nedenlerindedir [129]. Bu nedenle, öncü bileşiklerin ADME/T özelliklerinin erken değerlendirilmesi ve optimizasyonu, üretim verimliliğini ve ilaç geliştirme başarı oranını artırmada önemli bir rol oynamaktadır. Bununla birlikte, ilaçların neden olduğu yan etkileri inceleyen *in vivo* toksikolojik deneyler en önemli standartları oluşturmasına rağmen, uzun zaman alması, birçok hayvan kullanılması ve yüksek maliyet gibi nedenlerden dolayı bilgisayar destekli ADME/T tahmini, zamanla ilaç tasarımı için tercih edilen bir yöntem olmuştur [130]. Süreç içerisinde, ADME/T özelliklerini tahmin etmek için çok sayıda model tanımlanmıştır [131]. Bu modellerle birlikte bilgisayar tabanlı araçların büyük ölçekli kullanımının ilaç geliştirme maliyetini %50 azaltabileceği öne sürülmüştür [132].

ADME/T için 1930'lu yıllardan 1990 ve hatta 2000'li yılların başlarına kadar aktif olarak basit çoklu doğrusal regrasyon yöntemleri, LFER (linear free energy relationship), QSPR (Quantitative structure property relationship), HTS, GSE (general solubility equation) yöntem ve hesaplamaları kullanılırken günümüzde daha kesin sonuçlar veren programların gelişmesiyle yerini daha gelişmiş yöntemlere bırakmıştır [127].

Bu çalışmalardan biri olan ALOGPS, ilaç tasarımı (ADME/T ve HTS) ve çevre kimyası çalışmalarında kullanılan bir programdır ve bileşiklerin logP, suda çözünürlük ve pKa değerleri hakkında öngörüler sağlamaktadır. ALOGPS 2.1 sürümü, özellikle ilaç tasarımı çalışmalarında kullanılmak üzere tasarlanmıştır ve 1-oktanol/su içindeki partiyon katsayıları, logP ve sulu çözünürlük gibi özellikleri tahmin etmektedir. ALOGPS'un öne çıkan önemli özelliği, kullanıcı tarafından sağlanan yeni verileri birleştirme yeteneğine sahip olmasıdır. Bu özellik, ANN'nin kendi kendine öğrenme özelliği sayesinde mümkün olmaktadır. Bu nedenle, ALOGPS, ilaç tasarımı ve çevre kimyası çalışmalarında önemli bir araçtır ve yeni verilerin eklenmesiyle daha doğru öngörüler sağlayabilmektedir [133].

ADME/T tahmin modelini oluşturmak için SVM, GP, RF ve NB dahil olmak üzere bazı geleneksel ML algoritmaları kullanılmıştır [134-137] Ardından, güçlü genelleme ve özellik çıkarma yeteneği nedeniyle DL, ADME/T tahmin modelleri oluşturmak için kullanılmıştır. 2013 yılında, Lusci ve ark. ilk önce moleküler yapıyı kodlamak için yeni bir yönlendirilmemiş grafik tekrarlayan sinir ağı (UG-RNN) yöntemini kullanmış ve bu yöntemle kimyasal yapılar grafiğe dönüştürülmüştür. Bu yöntem, otomatik öğrenmeyi gerçekleştiren DL ağ yapısına dayalı bileşiklerin suda çözünürlüğünü etkili bir şekilde tahmin etmek için kullanılmışlardır [138].

Amerika Birleşik Devletleri federal kurumları (Çevre Koruma Ajansı (EPA), FDA ve Ulusal Sağlık Enstitüleri (NIH)) tarafından başlatılan veri yarışmasında, 21. Yüzyılda Toksisite Testi (Tox21) girişimi, en büyük çalışma olarak kabul edilmiştir. Tox21 girişiminde, kimyasalların insan sağlığı üzerindeki toksisite etkisini tahmin etmek, *in vivo* ve *in vitro* deneylerin sayısını azaltmak için daha etkili ve zaman kazandıran yöntemler geliştirilmesi amaçlanmıştır [139]. 2014 yılında yapılan Tox21 yarışmasında, Mayr ve ark., DL teknolojisini bileşiklerin toksisite tahminine uygulamış ve DeepTox

adlı, diğer yarışmacılardan daha başarılı bulunan, çok görevli bir DNN modeli geliştirmiştir. Sonuç, DL'nin toksisite tahmini açısından geleneksel yöntemlerden üstün olduğunu göstermiştir [62].

2015 yılına gelindiğinde ise Clark ve ark., ML tarafından oluşturulan, ADME/T tahmin modellerinin kullanılabilirliğini geliştirmek için NB'yi kullanan bir yazılım modülü oluşturulmuştur. Modül, Kimyasal Geliştirme Kiti (CDK) projesinde açık kaynaklı bir bileşen olarak piyasaya sürülmüş ve ADME/T, *in vivo* ve *in vitro* biyolojik aktiviteler, fizikokimyasal özellikler ve diğer amaçlar için bir dizi NB modeli oluşturabilen CDD Vault (Costumer Due Diligence) ve çeşitli mobil uygulamalarda uygulanmıştır [140].

Yine bu yılda en yeni derin öğrenme çerçevesi olan Yinelemeli İyileştirme Uzun Kısa Süreli Bellek (IterRefLSTM), grafik evrişimli bir sinir ağıyla entegre edilmiştir ve bu prototipleri, ilişkili ancak ayrı görevler arasında bilgi aktararak eğitmek için geliştirilmiştir. Bu çalışma, Tox21 ve veri tarama ve değerlendirme için kullanılan SIDER koleksiyonlarının aralıklı alt kümelerinde eğitilen prototiplerin performansını önemli ölçüde artırmış ve sonuç olarak; genellikle daha az girdi verisiyle kaybedilen bilgileri geri kazandırmıştır [141].

Ardından, Hughes ve ark. 702 adet epoksidasyon reaksiyon verisi ve DCNN'leri kullanarak epoksidasyon bölgesini doğru bir şekilde tahmin etmek için bir model oluşturmuştur. Model, yüzlerce epoksidasyon reaksiyonunu kantitatif olarak birleştirmiş ve epokside edilmiş ve epokside olmayan moleküller için sırasıyla %94.9'a kadar bir AUC ve %79.3'e kadar bir AUC ile atomik ve moleküler seviyelerde epoksidasyon bölgesi (SOE) tahminlerine ulaşılmıştır. Epokside metabolitler, genellikle ilaç toksisitesinin önde gelen nedeni olduğundan, epoksidasyon bölgesinin (SOE) doğru tahmini, daha güvenli ilaçlar elde etmek için metabolit oluşumu riskini etkili bir şekilde azaltabilmektedir [142].

Ayrıca, Xu ve ark. 475 adet ilaç bazında DL, DILI (Drug Induced Liver Injury) olarak adlandırılan, ilaca bağlı bir karaciğer hasarı tahmini modeli oluşturmak için UG-RNN moleküler kodlama yöntemini dikotomi ile birleştirmiştir. Sonuçlar, UG-RNN moleküler kodlamayı dikotomi ile birleştirmenin, ilaçların DILI riskini tahmin etmek için daha doğru sonuçlar verdiğini göstermiştir. Bu çalışma, ilaçların DILI riskini tahmin etmek için yeni bir yöntem sunmakla birlikte ilaç geliştirme sürecinde önemli bir rol oynamaktadır [143].

Iorio ve ark. tarafından yapılan çalışmada da, kanser hücrelerindeki gen mutasyon profilleri ve ekspresyon seviyeleri gibi faktörler göz önünde bulundurularak, ilaç moleküllerinin IC₅₀'sini tahmin etmek için elastik ağ modelleri kullanılmıştır. Bu çalışmada, elde edilen verilerin pratikteki sonuçlarla doğrulandığı bildirilmiştir [144]. Corte's-Ciriano ve ark. da aynı veri setini RF modellemelerini, ilaç moleküllerinin IC₅₀'sini tahmin etmek kullanmışlardır. Bu çalışmada, her bir tahminde istatistiksel güven ölçüsüne bağlı olarak öngörü performansının iyileştirilebileceği gösterilmiştir [145]. Her iki çalışma da ilaçların etkililiğini tahmin etmek için makine öğrenmesi yöntemlerini kullanmıştır. Elastik ağ modelleri ve RF modeli, ilaç geliştirme sürecinde önemli bir rol oynayabilmekte ve ilaçların etkililiğini tahmin etmek için yeni bir yaklaşım sunmaktadır.

Yine NB modelinin bir örneği olarak, PASS Online, farmakolojik etkiler, etki mekanizmaları, toksik ve yan etkiler dahil olmak üzere 4000'den fazla biyolojik aktivite türünü tahmin etmek için kullanılan bir programdır [146]. Bu program, bileşiğin yapısal formülüne dayanarak, farmakolojik etkiler, etki mekanizmaları, toksik ve yan etkiler gibi birçok biyolojik aktiviteyi öngörebilmektedir. DRABAL adlı etiket sınıflandırma yöntemi ise Bayes ağının yapı öğrenmesini kullanarak, 400.000'den fazla bileşiğin 1.4 milyondan fazla etkileşimini analiz edebilmekte ve PubChem BioAssay veri tabanından beş büyük HTS testi arasındaki mevcut ilişkileri inceleyebilmektedir. Bu yöntemler, bilgisayarda tasarlanan bileşiklerin biyolojik aktivitelerinin öngörülmesinde kullanılmakta ve ilaç tasarımı sürecinde önemli bir rol oynamaktadır [147].

Uygun moleküler tanımlayıcıların DL'nin ADME/T veri kümelerinde diğer geleneksel makine öğrenimi yöntemlerine göre herhangi bir iyileştirme sağlayıp sağlamadığını değerlendirmek için Korotcov ve ark., FCFP6 (FingerCode Print 6) parmak izlerini kullanarak farklı hesaplama yöntemlerini karşılaştırmak için farmasötik araştırmayla ilgili karmaşık uç noktalara sahip sekiz farklı ikili sınıflandırma veri seti seçmiştir. Dikkat çekici bir şekilde, farklı ölçümler veya veri kümeleri için sıralanmış puanlara dayalı olarak, DNN en iyi performansı sergilemiştir ve bu da DL'nin karmaşık veri kümeleriyle başa çıkma konusundaki güçlü yeteneğini göstermektedir [148].

Pande ekibi ise, dört farmasötik veri grubu (Kinase, Kaggle, Factors ve UV veri seti koleksiyonları) üzerinde, RF teknikleri üzerindeki çok görevli derin ağlarının üstünlüğünü ve dayanıklılığını da bildirmiştir. Böylelikle toksisiteyi tahmin etmek için derin öğrenmenin başarısı ve diğer ilaç tasarımı ve geliştirme girişimi alanlarındaki başarı, büyük ölçüde girdi verilerinin kalitesine ve miktarına bağlanmıştır [149].

Çok fazla veri noktasından anlamlı kimyasal bilgiler çıkarmak için Altae-Tran ve ark., aralıklı veri içeren görevler için iyileştirilmiş tahmin gücü sağlamak için "tek adımda öğrenme" yöntemini geliştirmiştir [150].

Li ve ark. ise insan sitokrom P450 (CYP450) izoformlarının eşzamanlı inhibisyonu için çok görevli bir model geliştirmiş ve çoklu görev modelinin, tek görev modellerinden, daha önce açıklanan sınıflandırıcılardan ve ortalama beş tasarımı görevinde geleneksel makine öğrenimi yöntemlerinden daha iyi öngörü sonuçları sunduğunu göstermişlerdir. Onların çoklu görev DNN modeli, on katlı çapraz doğrulama için %86,4 ve harici test veri setleri için %88.7 ortalama öngörü doğruluğu sunmuştur [151].

İlerleyen zamanlarda, Wenzel ve ark., DNN modellerini kullanarak, metabolik kararsızlık gösteren indol-3-karboksamid veya azaindol bağlı bir seri insan renin inhibitörü bileşiğinin deneysel ve tahmin edilen verileri ilişkilendirmiştir. Deneysel ve tahmin edilen veriler arasında iyi bir korelasyon bildirilmiştir. Ayrıca, çoklu görev modeli tek görev modelinden daha iyi sonuçlar vermiştir [152].

Günümüze kadar geliştirilmiş AI destekli bu yöntemlerle hesaplama güçleri geliştirilerek daha kısa sürede sonuç verecek şekilde çalışmalara devam edilmektedir.

İlaçların Yeniden Hedeflendirilmesinde Yapay Zeka Uygulaması

İlacın yeniden konumlandırılması olarak da bilinen ilacın yeniden hedeflendirilmesi, onaylanmış ilaçlardan yeni endikasyonların keşfedilmesi süreci olarak tanımlanmaktadır. Onaylanmış ilaçlar için mevcut bulunuyor oluşu ve bilinen güvenlik nedeniyle yeni endikasyonların incelenmesi, ilacın yeniden kullanımıyla birlikte, ilaç güvenliği ile ilgili sorunların riskini de etkili bir şekilde azaltabilmektedir [153].

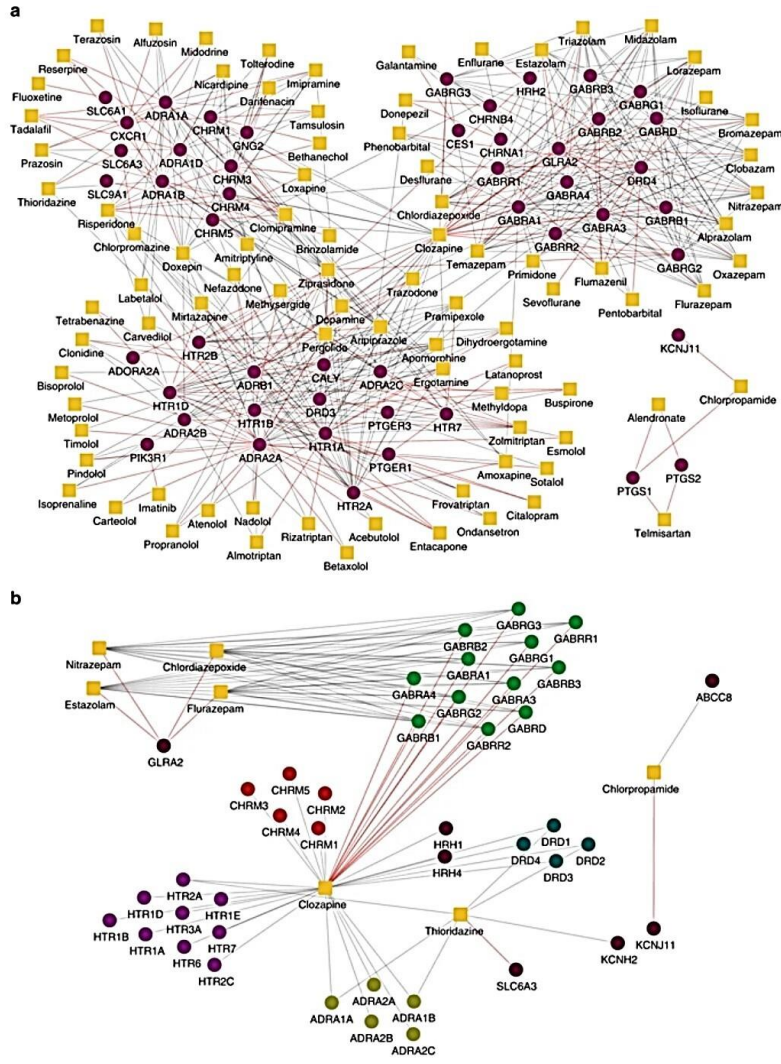
İlaç-hedef etkileşimi (DTI) tanımlaması, ilaç tasarımı ve ilacın yeniden hedeflendirilmesinde önemli bir rol almaktadır. Bununla birlikte, ilacın yeniden kullanım analizine ilişkin çok az veri vardır ve çeşitli biyoanaliz teknikleri, büyük ölçekli DTI'lerin gereksinimlerini karşılayamamaktadır. Böylece, DTI'leri hesaplamalı yöntemlerle tahmin etmek önemli bir araştırma konusu haline gelmiştir. Ligand ve yapı bazlı yöntemler, en yaygın kullanılan iki geleneksel hesaplamalı DTI tahmin yöntemidir [154].

Yapısal olarak benzer moleküllerin benzer biyolojik aktivitelere sahip olduğu varsayımına dayanan ligand bazlı yöntemler, hedef moleküllerin biyolojik aktivitelerini tahmin etmek için QSAR'ı uygulamaktadır [155]. Yapı bazlı yöntemler, hedefin kristal yapısına göre küçük molekülleri taramak için temel olarak moleküller yerleştirmeyi kullanmaktadır [156].

Sınırlı sayıda bilinen hedef aktif molekül ve hedef proteinlerin üç boyutlu yapısı nedeniyle, bu iki geleneksel hesaplama yöntemi büyük ölçüde yetersiz kalmaktadır. Son yıllarda, sürekli deneysel veri birikimi ve AI'nın heterojen verileri işlemedeki üstün performansı yardımıyla, DTI'leri tahmin etmek için birçok ML yöntemi uygulanmıştır [154,157-159]. En sık kullanılan makine öğrenimi modelleri, RF, SVM ve ANN gibi ikili sınıflandırıcılardır. DL yöntemleri de yüksek performansları ve çok soyut veri temsilini öğrenme yetenekleri nedeniyle çok daha fazla dikkat çekmiştir [160-162].

Mayr ve ark., ChEMBL gibi büyük ölçekli bir ilaç tasarımı veri seti üzerinde birkaç DL yönteminin performansını değerlendirmiş ve sonuçları, ilaç hedefleri tahmini için diğer geleneksel ML yöntemleri ile karşılaştırmıştır. İleri yönlü DNNs (FNNs) tüm rakip yöntemlerden önemli ölçüde daha iyi performans gösterdiğini ve belirli bir hedefi tahmin etme performansının *in vitro* analizlerle kıyaslanabilen, hatta daha iyi olduğu bulunmuşlardır [65].

Wen ve ark., ilk kez 2017 yılında DL'yi DTI tahminine uygulayan DeepDTIs adlı bir yöntem geliştirmiştir. Hedefleri sınıflandırmadan FDA onaylı ilaçlar ve hedefler arasındaki yeni DTI'leri doğru bir şekilde tahmin etmek için DBN'yi (deep-belief network) kullanmışlardır. Ayrıca DrugBank veri tabanından çıkarılan harici deneysel DTI verilerini test etmiş, uzaydaki tüm olası DTI'leri başarıyla tahmin etmiş ve en olası 10 adet DTI literatürde doğrulanmıştır [163].



Şekil 9. DTINet tarafından öngörülen ilaç-hedef etkileşimlerinin ağ görselleştirilmesi (en iyi 150 tahmini içeren genel ilaç-hedef etkileşim ağının görselleştirilmesi) [164]; a. Hedef ve ilaçlar sırasıyla mor daireler ve sarı kutular halinde gösterilmektedir. b. Literatürde bilinen deneysel veya klinik kanıtlarla desteklenebilecek yeni DTI tahminlerinin çeşitli örneklerinin ağ görselleştirilmesidir. İlaçlar sarı kutularda gösterilmekte, etkileşime giren hedeflerinin farklı aileleri farklı renklerde daireler halinde gösterilmektedir. İkisinde de, bilinen ilaç-hedef etkileşimleri gri, yeni tahmin edilen etkileşimler ise kırmızı çizgilerle gösterilmektedir

Heterojen veri kaynaklarının (ilaçlar ve hastalıklar arasındaki ilişki gibi) entegre edilmesi de DTI tahmininin doğruluğunun geliştirilmesine katkıda bulunabilmektedir. Bu ilişkiye dayanarak, Luo ve ark., heterojen ağlardan yeni DTI'ları tahmin etmek için DTINet adlı bir hesaplama süreci oluşturmuştur. DTINet tarafından öngörülen ilaç-hedef etkileşimlerinin ağ görselleştirilmesi Şekil 9'da verilmektedir. DTINet tarafından tahmin edilen üç ilaç ile siklo-oksijenaz (COX) proteini arasındaki yeni etkileşim, DTI'lerin nispeten güçlü bir tahmin performansı gösteren deneylerle de doğrulanmıştır [164].

CRISPR/Cas9, gen inaktivasyonu, aktivasyonu, mutasyonu ve kromozom yeniden düzenlemeleri gibi birçok genomik mühendislik işlemi için kullanılabilen çok yönlü bir araçtır [165]. Biyoteknoloji uzmanı Christof Fellmann, CRISPR-Cas'ın hedef molekülleri tanımlamaya yardımcı olmasının, ilaç tasarımında büyük bir etkiye sahip olabileceğini belirtmiştir. Bu teknoloji, hastalıklara neden olan veya engelleyen genleri ve proteinleri tanımlayarak, potansiyel ilaçlar için hedeflerin belirlenmesine yardımcı

olabilmektedir. Ayrıca, CRISPR-Cas, hastalıkları taklit eden hücrel ve hayvan modellerinin oluşturulmasını kolaylaştırarak, ilaçların güvenliği ve etkinliği hakkında daha doğru sonuçlar elde edilmesine olanak sağlamaktadır. Bu sayede, klinik çalışmalarda daha iyi öngörücüler oluşturulabilmekte ve ilaçların geliştirilmesi süreci hızlandırılabilir [166].

Molekül Transformörü-İlaç Hedef Etkileşimi (MT-DTI) adlı derin öğrenme tabanlı bir ilaç-hedef etkileşim modeli kullanılarak, SARS-CoV-2'nin viral proteinleri üzerinde etkili olabilecek ticari olarak kullanılan mevcut ilaçlar araştırılmıştır. Sonuçlar, HIV tedavisinde kullanılan antiretroviral ilaç olan Atazanavirin'in SARS-CoV-2 3C benzeri proteinaz karşı en etkili kimyasal bileşik olduğunu göstermiştir. Bununla birlikte Atazanavirin'in, Remdesivir, Efavirenz, Ritonavir ve Dolutegravir gibi diğer ilaçlardan daha yüksek bir inhibitör potensine sahip olduğu bildirilmiştir [167]. Bu çalışma, farklı bir viral enfeksiyonun tedavisinde kullanılan antiretroviralin başka bir viral enfeksiyondaki etkinliğini ortaya koyması açısından önemlidir ve MT-DTI modelinin ilaç tasarımı sürecinde kullanımının faydalı olabileceğini göstermektedir.

De novo İlaç Tasarımında Yapay Zeka Uygulaması

De novo ilaç tasarımı, reseptörün 3 boyutlu yapısını kullanarak ilgilenilen hedef için beklenen aktiviteye sahip yeni kimyasal yapılar elde etmek için moleküler tasarım ve değerlendirme amacıyla bilgisayarlı algoritmalara dayalı bir yöntemdir. Yeni kimyasal yapıların biyolojik aktivitesini, ilaç metabolizmasını ve farmakokinetik (DMPK) özelliklerini, sentez için uygun koşulları ve sentezin uygulanabilirliğini öngörmesi beklenmektedir [168].

İlk *de novo* ilaç tasarımında, bağlama ceplerini hedeflemek için sterik ve elektronik olarak uygun olan ligandları geliştirmek amacıyla yapı bazlı bir yöntem kullanılmıştır. Bu yöntemle tasarlanan bileşikler genellikle zayıf DMPK özelliklerine sahiptir ve sentezlenmesi zordur [169]. Buna karşılık, ligand bazlı yöntem, kimyasal yapıların büyük bir veri tabanını oluşturmak için kullanılmaktadır. Molekülün kimyasal alanını araştırmak için DMPK özelliklerini, sentez fizibilitesini, biyolojik aktiviteyi ve hedef yapısı benzerliğini dikkate alan bir skor fonksiyonu kullanılmaktadır. Böylece sentetik olarak uygulanabilir birçok molekül elde edilebilmektedir [170].

İkinci yöntemdeki amaç ise, araştırmacıların mesleki bilgilerine dayalı olarak hedef yapısı analogları tasarlamak olmuştur. Yeni moleküller, sentez veya reaksiyon kuralları kullanılarak güvenilir ve verimli bir şekilde tasarlanabilmesine rağmen, bu tasarımlar genellikle laboratuvar ortamındaki uygulanabilirlik ve reaksiyonların zorlu doğası nedeniyle sınırlıdır [171].

Üçüncü yöntemde de amaç, uygun öngörde aktivite bölgesini, karşılık gelen moleküler yapıya eşlemeyi amaçlayan ters QSAR fikrini benimsemektir. Bu yöntem nispeten zordur, çünkü seçilen moleküler tanımlayıcıların yalnızca bir ileri QSAR tahmin modelinin oluşturulması için değil, aynı zamanda bir moleküler yapıya dönüşüm için de uygulanabilir olmasını gerektirmektedir [172].

Yeni ilaç tasarımının mevcut zorluklarının iyileştirilmesinde, güçlü genelleme ve öğrenme yetenekleriyle DL yaklaşımı, bazı beklenen özelliklere sahip yeni kimyasal yapıları otomatik olarak oluşturmak için kullanılmıştır. Olivecrona ve ark. *de novo* ilaç tasarımı için dizi bazlı bir optimizasyon yöntemi kullanmış ve öngörülen aktiviteye sahip bileşikler oluşturmak üzere ChEMBL veri tabanında önceden eğitilmiş RNN'de ince ayar yapmak için pekiştirmeli öğrenmeyi (RL) benimsemiştir. Model, dopamin tip 2 reseptörü için aktif bileşiklerin oluşumunu tahmin etmek için kullanıldığında, model tarafından tahmin edilen bileşiklerin %95'inden fazlasının aktif olduğu bildirilmiştir [173].

Kadurin ve ark. druGAN adı verilen, istenen moleküler özelliklere sahip yeni moleküllerin *de novo* tasarım için gelişmiş bir üretken karşıt otokodlayıcı (AAE) modeli önermiştir. Bu yöntem, moleküler parmak izlerinin ayarlanabilirliği açısından varyasyonel otomatik kodlayıcı VAE'den açıkça üstündür ve derin üretken modeli kullanarak spesifik anti kanser özelliklerine sahip yeni moleküller geliştirme yeteneğini ve verimliliğini önemli ölçüde artıran büyük moleküler veri setlerini işleme yeteneğine sahiptir [174].

Gomez-Bombarelli ve ark. ise moleküllerin farklı tasarımını çok boyutlu sürekli bir tasarıma dönüştürmek için bir teknik bildirmiştir. Amaçlanan özelliklere sahip yeni bileşikler otomatik olarak oluşturmak için çok katmanlı bir algılayıcı (MLP) ve varyasyonel otomatik kodlayıcı (VAE) entegre etmişlerdir. Bu DNN, bir kod çözücü, bir kodlayıcı ve bir tahmin ediciden oluşmaktadır. Kodlayıcı, farklı basitleştirilmiş moleküler giriş hattı giriş sistemi olan SMILES (Simplified Molecular Input Line

Entry System) dizisini gizli uzayda sürekli bir vektöre dönüştürmekte ve kod çözücü bu vektörleri tekrardan farklı SMILES dizisine dönüştürmektedir [175].

Zhavoronkov ve ark. geliştirdiği Generative Tensorial Reinforcement Learning (GENTRL) modeli de novo küçük molekül tasarımı için kullanılan bir derin öğrenme modelidir. Bu model, sentetik fizibilite, yenilik ve biyolojik aktiviteyi optimize ederek kullanılmaktadır. Bu modeli kullanarak araştırmacılar, fibroz ve diğer hastalıklarda rol oynayan Diskoidin Alan Receptor 1 (DDR1) adlı bir kinaz hedefi için güçlü inhibitörler keşfetmiştir. Keşfedilen bileşiklerin biyokimyasal özelliklerini doğrulamak için bileşikler önce sentezlemiş ve daha sonra hücre bazlı testlerle ölçümlerini yapmışlardır. Bu çalışmada, farelerde farmakokinetiği kanıtlanmış bir öncü aday bileşiği elde etmişlerdir. İki aydan daha kısa sürede sonuçlanan bu araştırma, derin öğrenme yöntemlerinin hızlı ve etkili moleküler tasarım için potansiyelini göstermektedir [176].

Ardından Skalic ve ark. tarafından, yeni moleküllerin üretilmesinde kullanılabilecek, zayıf yapısal çeşitlilik sorununu çözmek için çekirdek bileşiklerinin 3 boyutlu şekilden ve farmakodinamik özelliklerinden yararlanan bir yöntem önerilmiştir. Bu aynı zamanda öncü bileşik benzeri yeni bir tasarım yürütülmesi için uygulanan ilk yöntemdir. Bu yöntemde, çekirdek bileşiklerinin 3 boyutlu temsilini bozmak için önce bir VAE kullanılmış ve ardından CNN'ler ve RNN'lerden oluşan ağ sistemi tarafından SMILES dizi sembolü üretilmiştir. Son olarak SMILES analiz edilerek yeni moleküller elde edilmiştir. Bu yöntemle tasarlanan yeni iskeleler ve fonksiyonel gruplar, kimyasal alanda henüz keşfedilmemiş ancak öncü bileşik benzeri özelliklere sahip olabilecek alanları kapsayabilmektedir [177].

ADAMs (a Disintegrin and Metalloproteinase) hücre yapışması, hücresel göç, hücresel sinyalleme ve proteoliz gibi fizyolojik işlemlerde önemli transmembran proteinler ailesidir [178,179]. Bu protein ailesinden olan ADAM10, alzheimer, prion hastalığı, alerjik reaksiyonlar, glioblastoma ve pankreas kanseri gibi hastalıkların tedavisi için umut verici bir hedefdir [180-183]. ADAM10 inhibitörlerinin moleküler generatif modelini oluşturmak için, transfer öğrenmesi ile birlikte geçişli tekrarlayan birim (GRU) tabanlı bir derin sinir ağı kullanılmıştır. Bu çalışmada, GRU tabanlı generatif modelin, moleküllerin SMILES dizilerini doğru bir şekilde öğrenebildiği ve yeni potansiyel ADAM10 inhibitörleri üretebildiği gösterilmiştir. Geleneksel ligand bazlı yöntemlerle kıyaslandığında, GRU tabanlı model, kimyasal ligandların yalnızca SMILES bilgilerini gerektirmekte ve etkili geniş bir seri potansiyel yeni bileşikler oluşturabilmektedir [184].

Green ve ekibi ise 2020 yılında, kimyasal yapı oluşturma, deneysel tasarım, aktif öğrenme ve kemoinformatik araçlarını entegre eden otomatik moleküler tasarım için Biyolojik Tepki Analizi ve Tasarım Sistemi (BRADSHAW) adlı bir sistem geliştirmişlerdir. Bu sistem, tasarla-yap-test-et-analiz-et (DMTA) döngüsünde otomatik moleküler tasarım yapmayı mümkün kılmaktadır. Kolay kullanılan arayüzü sayesinde, büyük ölçekli otomatik tasarım erişimi kolaylaştırılmıştır ve yazılım geliştirme süreci en aza indirgenmiştir. Sistem geleneksel kemoinformatiklerin ve modern makine öğrenimi algoritmalarının kullanımını içermektedir. Geriye dönük bir durum çalışmalarında da, BRADSHAW'ın MMP12 inhibitörlerinin tasarımı için öncü bileşik optimizasyonunda başarıyla kullanılabildiği gösterilmiştir [185].

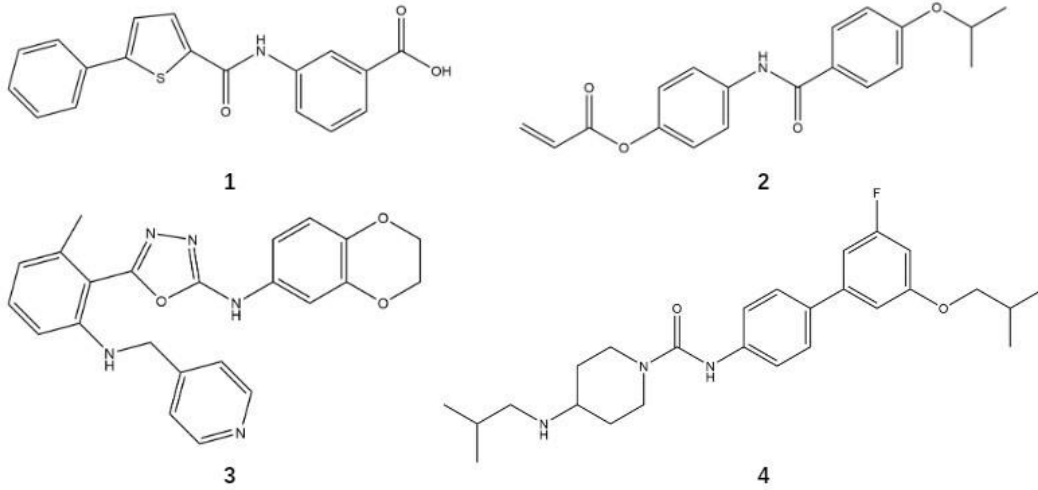
Yapay Zekâ Teknolojisiyle Geliştirilen İlaç Molekülleri ve Son Gelişmeler

Günümüzde her geçen gün yeni ilaç moleküllerine duyulan ihtiyaç artmaktadır. Yapay zekâ teknolojilerindeki gelişmeler de göz önüne alındığında yapay zekanın ilaç tasarımında kullanılmasıyla geliştirilen moleküller daha çok gündeme gelmektedir.

AI sistemleri *de novo* tasarım modelleriyle de uyum içinde aktif olarak kullanılabilmektedir [186-188]. Bu sistemler, denetimli ileri beslemeli ağların, denetimli ve yarı denetimli varyasyonel otomatik kodlayıcıların, tekrarlayan sinir ağlarının ve üretken çekişmeli ağların derin geri dönüşümünü (tersine çevirme) içermektedir [189-191]. Son zamanlarda, generatif AI'ya dayanan yenilikçi bir *de novo* moleküler tasarım kavramı önerilmiştir. Bu durum, bilinen biyoaktif bileşiklerden öğrenmenin bir yolu olarak umut vaat etmekte ve kalıtsal biyoaktivite ve sentezlenebilirliğe sahip yeni bileşikler kendi kendine tasarlanmaktadır [192].

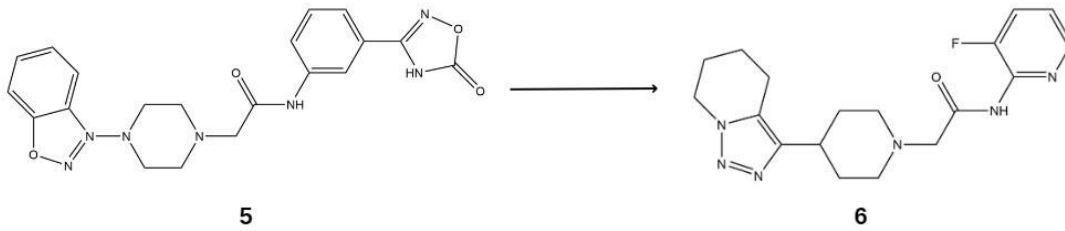
De novo kaynaklı bileşiklerin ilk başarılı dört sentezi (Şekil 10), generatif moleküler tasarımın ilaç tasarımına uygulanabilirliğini doğrulamaktadır. Derin sinir ağlarıyla üretilen RXR γ agonistleri

bileşik 1 ve 2 öncü prospektif tasarımlar olmuştur. Bu derin sinir ağları, ChEMBL'den SMILES biyoaktif bileşik dizileriyle önceden eğitilmiş ve transfer öğrenimi kullanılarak nükleer hormon reseptör hedefleri üzerinde ince ayar yapılmıştır. VEGFR2 kinaz inhibitörü olarak kabul gören bileşik 3, bilgisayar tarafından üretilen moleküllerin biyoaktivitesinin ligand-reseptör yerleştirme kullanılarak tahmin edildiği, ilgili bir SMILES tabanlı yaklaşımla oluşturulmuştur. Kısmi 5-HT_{2B} antagonisti olarak bildirilen bileşik 4 ise sanal ileri sentez için bir nöral ağ ile de novo yapı tasarımıyla üretilmiştir [190,192-194].



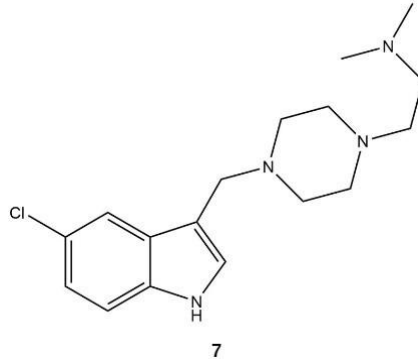
Şekil 10. De novo yöntemiyle elde edilen ilk başarılı bileşik örnekleri [190,192-194]

Bunun yanı sıra bir kimyasal serinin çok amaçlı optimizasyonu, potansiyel bir prelinik ilaç adayının arzu edilen tüm özelliklerini aynı anda karşılayan bileşikleri tanımlamayı amaçlamaktadır. Bu amaçla son zamanlarda, IKTOS ve Servier'deki bilim adamları, 11 adet hedef üzerinde çalışarak (fenotipik aktivite, 5-HT_{2A}, 5-HT_{2B}, α 1 ve D1 reseptörlerine karşı seçicilik, NaV 1.2 ve hERG iyon kanalları, sıçan ve insanda karaciğer mikrozomal stabilitesi ve Caco-2 Fab'lar ve Efflux) açıklanmayan başka bir hedef için ileri aşama bir öncü optimizasyon bağlamında bileşiklerin ligand tabanlı optimizasyonu için derin öğrenmenin kullanımını açıklamışlardır. 880 molekül içeren ilk veri setinde hiçbir molekül, bu 11 amaçtan herhangi birini karşılayamamıştır. Proje veri setinden QSAR modelleri geliştirilmiş ve bu modeller derin öğrenme uzun kısa süreli bellek (LSTM) modellerine dayanan tescilli bir SMILES tarafından oluşturulan yeni sanal yapıları puanlamak için kullanılmıştır. Ardından 11 hedef için oluşturulan yapılar üzerinde QSAR puanlaması kullanılmıştır. Son olarak, tüm hedefleri karşılayan 150 sanal yapı tahmini yapılmıştır. Çok büyük çoğunluğu ilk veri setinde bulunmayan bu sanal yapılardan 20 tane bileşik sentez için seçilmiş ve 11'i başarıyla sentezlenmiştir. Detaylı çalışmaların ardından 3 bileşik, belirtilmiş 11 hedef kriterinin hepsini karşılamıştır. 11 hedeften 9'unu karşılayan Şekil 11'de verilen bileşik 5, 11 hedefin tümünü karşılayan bileşik 6'nın tasarımına, sentezine ve tahliline yol açan QSAR eğitim setinin bir parçası olmuştur [195].



Şekil 11. Sentezi yapılan etkin bileşik 5 ve bileşik 6 molekülleri [195]

Daha önceki bir çalışmada, ETH Zurich ve Novartis'ten araştırmacılar, moleküler yapı için uyarlanabilir bir algoritma (MAntA olarak bilinen Moleküler Karınca Algoritması) kullanmışlardır [196]. Bu algoritma diğer 5-HT reseptörü alt tiplerine veya hERG potasyum kanalı da dahil olmak üzere 18 başka protein hedefinden oluşan bir panele bağlanmadan 5-HT_{2B} reseptörünün seçici antagonistleri olan küçük, sentetik olarak erişilebilir de novo tasarımları otomatik olarak üretmek için kullanılmıştır [197]. Bu yöntem sonucunda istenen seçiciliği sergileyen ve MAntA tarafından önerilen sentetik yolla uyumlu olarak, bir indirgeyici aminasyon adımında ticari olarak mevcut başlangıç malzemelerinden sentezlenen bileşiğin tasarımı gerçekleştirilmiştir. Sentezlenen bu bileşik 7, Şekil 12'de verilmektedir.



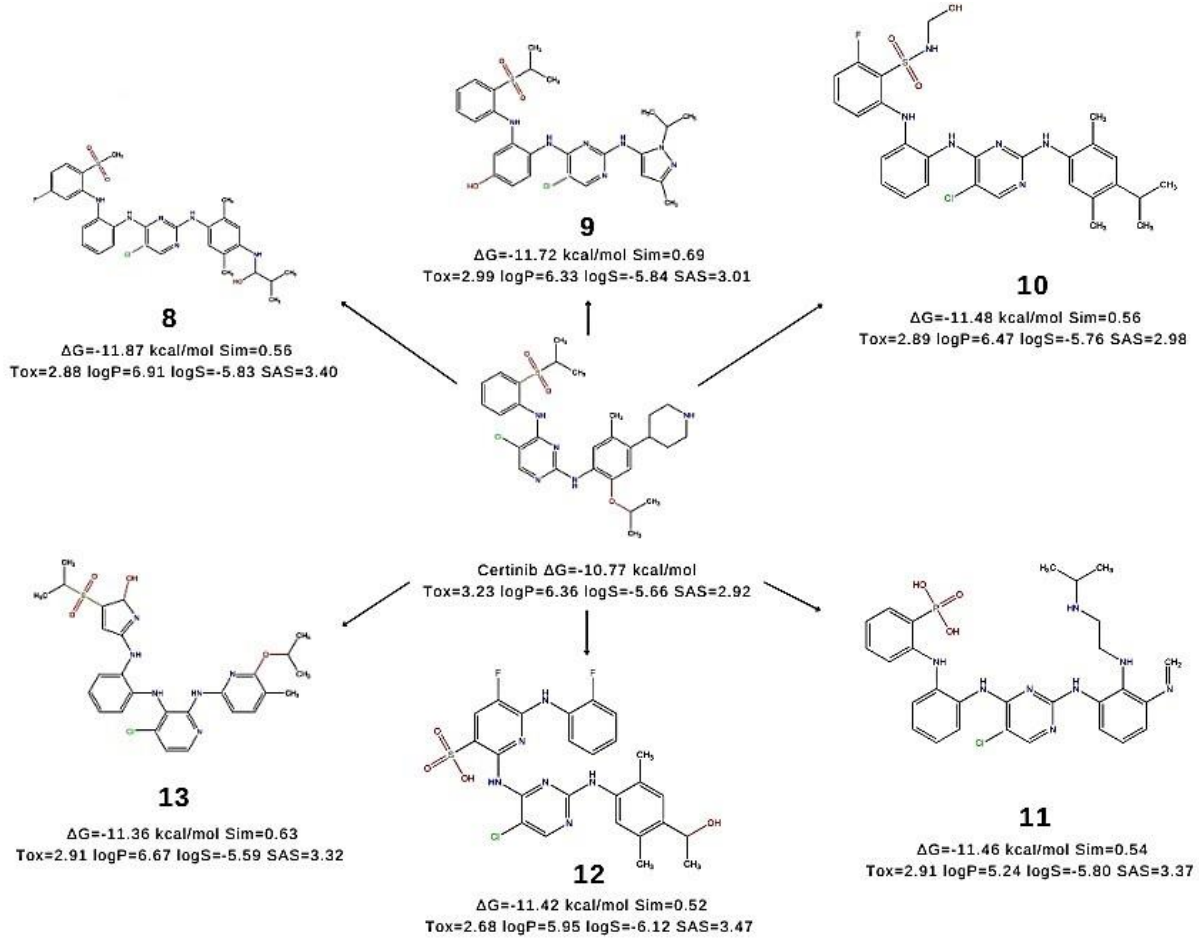
Şekil 12. MAntA ile tasarımı yapılan bileşik [197]

Yukarıda belirtilen molekül tasarımlarının yanı sıra, piyasadaki mevcut moleküllerin optimizasyonu ile daha etkili, güvenli ve daha ucuz bir şekilde elde edilebilen molekül tasarımları yapılabilmektedir.

Bu amaçla, otomatik olarak yeni ilaç benzeri moleküller üretmek için çoklu özellik optimizasyonuna dayalı bir üretken ağ kompleksi (GNC) geliştirilmiştir. GNC temelde, SMILES dizilerini, DNN tabanlı molekül üreticileri, önceden eğitilmiş kodlayıcıları ve kod çözücülerini kullanmaktadır [198]. Geliştirilen GNC ile alternatif ilaç adaylarını tasarlamak amacıyla son yıllardaki kritik hastalıklar ve çeşitli kanserleri tedavi etmek için FDA onaylı, tek ve çok hedefli ilaçlar üzerinde çalışmalar yapılmıştır.

Tek hedefli ilaç tasarımı 'Ceritinib' üzerinde çalışılmıştır. Novartis tarafından geliştirilmiş ve akciğer kanseri türlerini tedavi etmek için Nisan 2014'te FDA tarafından onaylanmıştır. Ceritinib son derece pahalıdır ve ABD'de Ceritinib bazlı tedavinin aylık maliyeti günümüzde yaklaşık 11.428 dolardır. Ceritinib, ALK (anaplastik lenfoma kinaz) tirozin kinaz reseptörünü inhibe etmektedir. Hedef ALK'ya potansiyel inhibitörler tasarlamak için eğitilen GNC, ceritinib molekülünü referans molekül olarak kullanarak alternatif moleküller tasarlamıştır. Tasarlanan moleküllerin enerjileri (ΔG), ilaca benzerlik puanları (Sim), hesaplanan toksisite (Tox), log P, log S değerleri ve sentetik erişilebilirlik puanları (SAS'lar) referans molekülle kıyaslandığında eşdeğer ve hatta daha başarılı bağlanma afinitesi gösterebilen 6 başarılı moleküle (bileşik 8-13) Şekil 13'te yer verilmiştir [198].

Çift hedefli ilaç tasarlamada ise, 'Ribociclib' üzerinde çalışılmıştır. Ribociclib (ChEMBL ID: CHEMBL3545110) Novartis ve Astex Pharmaceuticals tarafından geliştirilmiş ve bazı meme kanseri türlerini tedavi etmek için 2017 yılında FDA tarafından onaylanmıştır. ABD'de Ribociclib tedavisinin aylık maliyeti yaklaşık 10.950 dolardır. Ribociclib, CDK4 ve CDK6 olmak üzere iki farklı hedefi inhibe etmektedir. Alternatif ribociclib molekülü tasarımlarında CDK4 ve CDK6'ya bağlanma afiniteleri öncelikli olarak göz önünde bulundurulup en başarılı bulunan 6 molekül (bileşik 14-19) Şekil 14'te gösterilmiştir. Tasarlanan moleküllerin çoğu, referans ilaç Ribociclib'den daha iyi bağlanma afinitelerine sahiptir. Örneğin CDK4'e birinci ve dördüncü bileşiklerin afinitelerinin Ribociclib'inkinden daha yüksek olduğu ve diğer üçünün Ribociclib ile benzer afinitelere sahip olduğu saptanmış, CDK6'ya olan bağlanma afiniteleri ise tüm bileşiklerde daha iyi olduğu bildirilmiştir [198].



Şekil 13. Ceritinib molekülü ve alternatif olarak üretilen ilk altı molekül (hedef ALK için öngörülen ΔG 'leri, ilaca benzerlik puanları (Sim), hesaplanan toksisite (Tox), log P, log S değerleri ve sentetik erişilebilirlik puanları (SAS'lar)) [198]

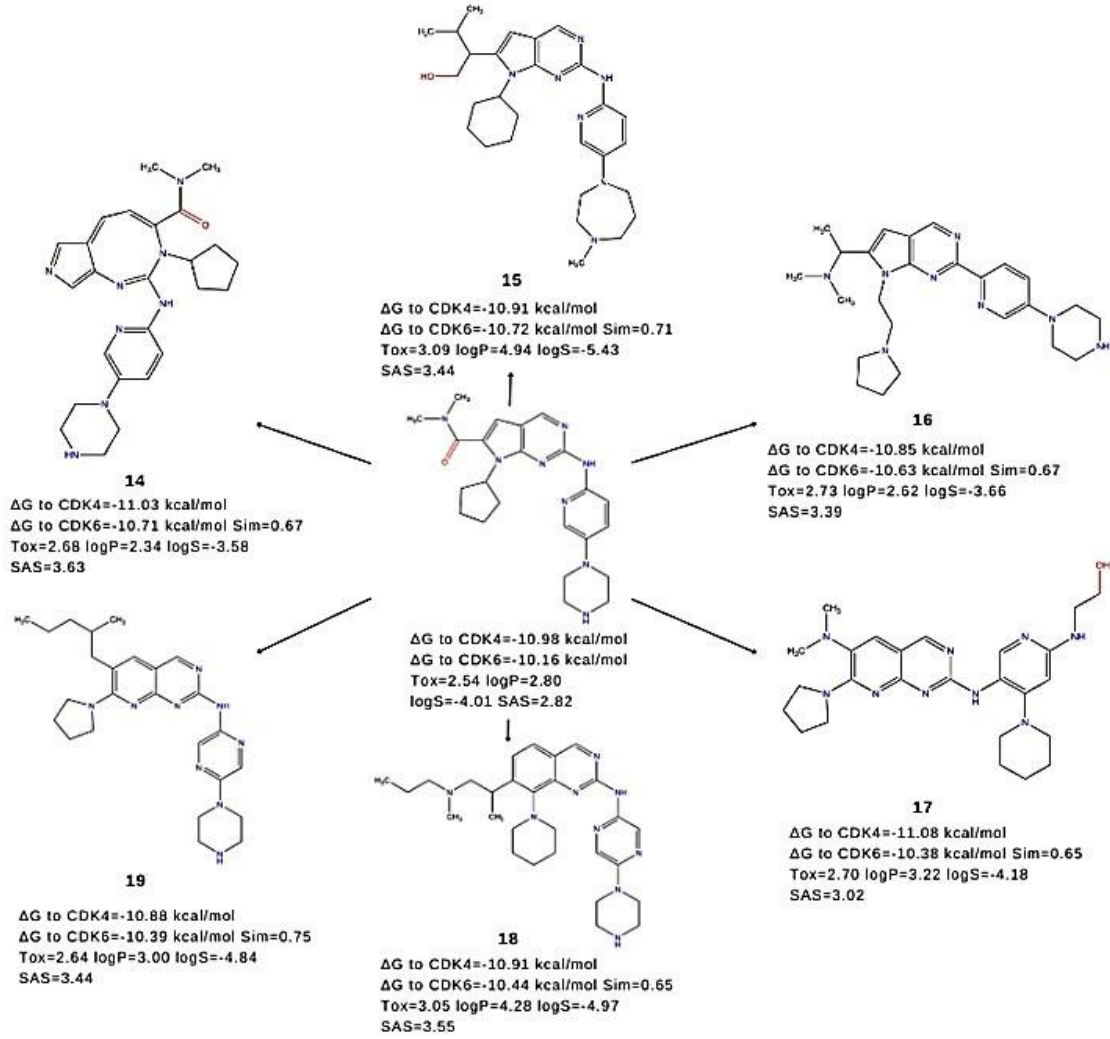
Geliştirilen bu yöntemle Ceritinib ve Ribociclib dışında Acalabrutinib, İdelasidib, Dabrafenib gibi piyasada bulunan birkaç ilaç molekülüne alternatif ilaç adayı oluşturmak için çalışmalar devam etmektedir.

Bu çalışmaların yanı sıra, entegre DMTA platformlarının uygulamaları, ilaç tasarımı ve optimizasyonuna uygulanabilmektedir [199]. Bu amaçla, AbbVie'deki araştırmacılar, küçük bileşik kitaplıkların otomatik sentezi için, esas olarak ticari olarak temin edilebilen bileşenlerden oluşturulan entegre bir robotik platform geliştirmişlerdir [200]. Ardından seçici hepsin inhibitörlerinin üzerinde çalışılmasıyla birlikte, otomatik DMTA döngülerinde bilinen hepsin inhibitör bileşiğinden Şekil 15'te verilen bileşik 20'den yeni bir bileşik olan bileşik 21 tasarlanmıştır. Tasarım sonucu, hepsin'e karşı yarı maksimum inhibitör konsantrasyon (IC_{50}) değeri $1\mu M$ 'den 33 nM'ye yükseltilmiş ve ürokinaz tipi plazminojen aktivatörüne göre seçicilik 30 kattan 100 kata çıkarılmıştır. Tüm bu çalışmaların yaklaşık 1.5 saat gibi kısa bir sürede bittiği gözlemlenmiştir.

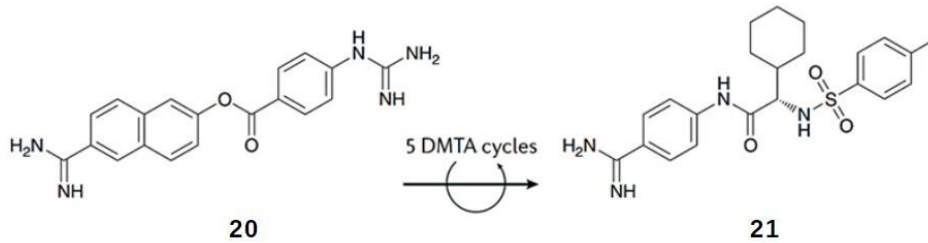
Birçok yeni ilaç molekülüne ihtiyaç duyulduğu gibi antibiyotiğe dirençli bakterilerin hızla ortaya çıkması nedeniyle, yeni bir antibiyotik tasarlama ihtiyacı da artmaktadır. Ne yazık ki, yeni antibiyotiklerin tasarımı giderek zorlaşmaktadır, çünkü doğal ürün keşfi, artık aynı moleküllerin tekrar tekrar keşfedildiği dereplikasyon sorunuyla boğuşmaktadır [201].

Makine öğrenimindeki son gelişmeler göz önüne alındığında, artık yeni yapısal antibiyotik sınıflarını tanımlamak amacıyla moleküler özellik tahmininde algoritmik çözümler başarıyla uygulanmaktadır. Erken dönem ilaç tasarımının büyük ölçüde in silico gerçekleştirilmesine izin veren

metodolojilerin benimsenmesi, mevcut deneysel yaklaşımların ulaşamayacağı geniş kimyasal alanların keşfedilmesini sağlamaktadır [202].



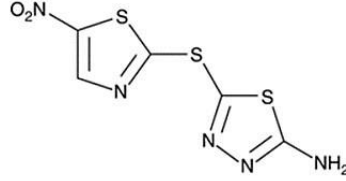
Şekil 14. Ribociclib molekülü ve alternatif olarak üretilen ilk altı molekül (CDK4 ve CDK6 hedeflerine yönelik öngörülen ΔG 'leri, ilaca benzerlik puanları (Sim), hesaplanan toksisite (Tox), log P, log S değerleri ve sentetik erişilebilirlik puanları (SAS'lar)) [198]



Şekil 15. Hepsin inhibitörü (20) ve tasarlanan yeni bileşik (21) [200]

Sinir ağı modelleri, parmak izi vektörleri, otomatik tanımlayıcılar gibi yaklaşımların kullanılmasıyla yeni antibakteriyel aktiviteye sahip molekülleri tasarlayabilen derin bir sinir ağı eğitilmiştir. Birden fazla kimyasal veri setleri üzerinde tahminler yapılmış ve Drug Repurposing

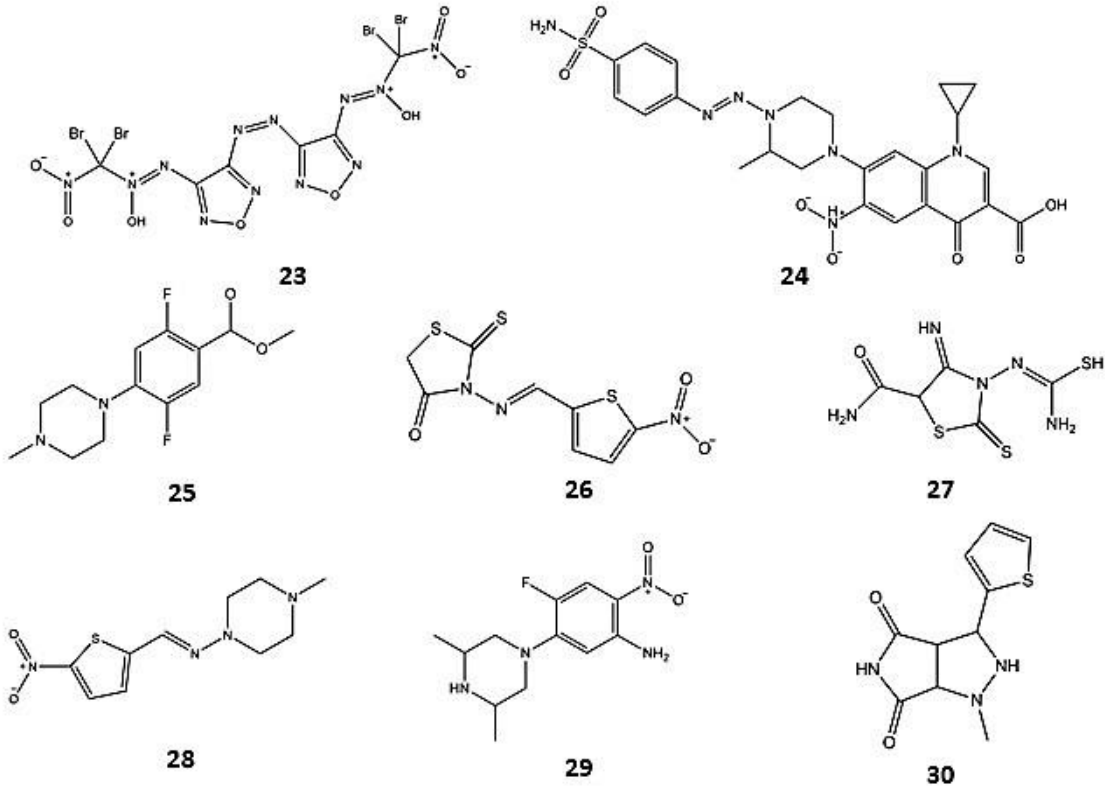
Hub'dan, geleneksel antibiyotiklerden yapısal olarak farklı olan ve *Mycobacterium tuberculosis* ve karbapenem dirençli *Enterobacteriaceae* dahil olmak üzere geniş bir filogenetik patojen spektrumuna karşı bakterisidal aktivite gösteren bir molekül olan c-Jun N-terminal kinaz inhibitörü SU3327 (Halicin) tasarlanmıştır. Ayrıca Şekil 18'de gösterilen bileşik 22, Halicin molekülü, fare türlerinde *Clostridioides difficile* ve pan ilaç rezistanlı *Acinetobacter baumannii* enfeksiyonlarını etkili bir şekilde tedavi etmiştir [203,204].



22

Şekil 16. Halicin molekülü

Halicin'e ek olarak, ZINC15 veritabanında bulunan 107 milyon molekülden deneysel olarak test edilmiş 23 farklı molekül seti tasarlanmış ve oluşan bu setten, bilinen antibiyotiklerden yapısal olarak uzak Şekil 17'de verilen 8 ek antibakteriyel bileşik keşfedilmiştir. Bu çalışmada, bu moleküllerden bileşik 23 ve bileşik 24, dikkate değer bir şekilde güçlü geniş spektrumlu aktivite sergilemiş ve *E. coli*'deki bir dizi antibiyotik direnci belirleyicisinin üstesinden geldiği bildirilmiştir. Bu çalışma, makine öğreniminin, öncü bileşik tanımlamasının doğruluk oranını aynı anda artırarak ve tarama çalışmalarının maliyetini azaltarak antibiyotik tasarlama çabaları üzerinde sahip olabileceği önemli etkiyi vurgulamaktadır [204].



Şekil 17. Keşfedilen antibakteriyel etkili bileşikler [204]

Tüm bu gelişmelerle birlikte, yakın zaman önce Al ile ilaç tasarımı kilometre taşı sayılabilecek gelişmeler, 30 Ocak 2020'de Exscientia'nın, Faz 1 Klinik denemesine girmek için Al tarafından tasarlanan ilk molekülü DSP-1181'i duyurmasıyla devam etmiştir. DSP-1181 bileşiği ile ilgili çalışmalar, işbirlikçi Sumitomo Dainippon Pharma tarafından uzun etkili ve güçlü bir serotonin 5-HT1A reseptör agonisti olarak, obsesif-kompulsif bozukluğunu (OKB) tedavi etmek için faz I klinik çalışmalarına kadar ilerletilmiştir. Normalde on yıla kadar sürebilen bu süreç 12 ayda sonuçlanmıştır. Yapısal analizler sonucunda, iddia edilen molekülün, FDA'nın 1967'de onayladığı sık kullanılan birinci nesil bir antipsikotik ajan olan Haloperidol ile benzerliği ortaya konmuştur [205-207].

Ardından Aralık 2020'de, Exscientia'nın en gelişmiş, adozin A2a reseptör antagonisti kabul edilen, öncü ilaç adayı EXS21546, Birleşik Krallık'ta çeşitli tümör türleri için immüno-onkoloji tedavisi amacıyla Faz 1 klinik çalışmasına başlamıştır. EXS21546 üzerindeki çalışmalar, yapay zeka firması Exscientia ve ilaç firması Evotec arasındaki işbirliğinin bir parçası olarak devam etmektedir [207].

DSP-0038 ise Exscientia ve Sumitomo Dainippon Pharma arasındaki iş birliğinin bir parçası olarak keşfedilen çift hedefli bir 5-HT1a reseptör agonisti ve 5-HT2a reseptör antagonisti olarak kabul edilmektedir. Mayıs 2021'e gelindiğinde, Amerika Birleşik Devletleri'nde, DSP-0038'in klinik çalışmasının başlatıldığı açıklanmıştır. Şu anda Alzheimer hastalığı psikoza tedavisi için araştırılmaktadır [207].

SONUÇ VE TARTIŞMA

Yapay zekâ ve makinelerde öğrenme kavramı, bir fikir olarak ortaya çıkmasından günümüze kadar sürekli gelişerek hayatımızda çığır açmaya devam etmektedir. Yaşanan gelişmeler matematik, istatistik, mühendislik ve tıp gibi alanlara yön vermeye devam ederken, ilaç tasarımı da yapay zekâ ve bilgisayar destekli teknolojilerden faydalanılmaktadır.

İlaç tasarımı yapay zekâ çalışmaları ilkel algoritmalarla başlamış olup süreç içinde makine öğreniminin gelişmesiyle yerini ML tekniklerine ve makine öğreniminin de daha gelişmiş alt dalı olan DL'e bırakmıştır. Geliştirilen yapay zekâ uygulamaları, ilaç tasarımı önemli aşamaları olan proteinlerin katlanması tahmininde, protein-protein etkileşimlerinin tahmininde, sanal taramada, kantitatif yapı-aktivite ilişkilerinin incelenmesinde, ADME/T farmakokinetik özelliklerinin tahmininde, ilaçların yeniden hedeflendirilmesinde ve de novo ilaç tasarımı aktif olarak kullanılmakta ve çok önemli rol almaktadır. Özellikle derin sinir ağları ve tekrarlayan sinir ağları gibi yapay sinir ağları, ilaç tasarımı yapay zekanın kullanımına büyük ölçüde yön vermiştir.

Yapay zekânın kullanımının yaygınlaşmasıyla birlikte ilk çalışmalar QSAR çalışmaları ile başlamış olup başarılı sonuçlar elde edilmiştir. İlerleyen çalışmalar doğrultusunda DL algoritmalarının ML algoritmalarından daha başarılı sonuçlar verdiği gözlemlenmiştir.

İlaç tasarımı için büyük önem taşıyan protein-protein etkileşimleri ve protein yapısının deneysel olarak belirlenmesi zor, zaman alıcı ve maliyetli bir süreç olsa da yaşanan bu gelişmelerle birlikte PPI ve protein yapısının tahmini üzerinde DL'nin önemi açıkça görülmektedir.

Yine sanal tarama çalışmalarında uygulanan ML ve DL uygulamaları ilaç-hedef etkileşimlerini öngörmede ve terapötik hedefleri belirlemede başarılı bir şekilde kullanılmış ve aktif olarak kullanılmaya devam edilmektedir. ADME/T üzerinde AI uygulamalarının ardından öncü bileşiğin erken değerlendirilmesi ve optimizasyonu konularında zaman ve üretim verimliliği açısından ciddi bir başarı gözlenmiştir. Bu ilerlemelerinin yanı sıra, ilaçları yeniden hedeflendirmek amacıyla yapılan çalışmalarda biyoteknoloji alanındaki ilerleme büyük ölçüde hızlanmış ve yakın zaman önce yaşanan SARS-CoV-2 pandemisi için Atazanavir gibi alternatif ilaçlar tespit edilmiştir.

İlaç endüstrisinde uygulanan yapay zekâ ve ML algoritma teknolojilerindeki gelişmelere rağmen, bu teknolojilerin özel olarak ilaç tasarımı sürecine ve genel olarak ilaç endüstrisine uygulanması ve entegrasyonu ile ilgili birçok zorluk bulunmaktadır. Bu sorunlardan biri verimsiz veri entegrasyonudur. Bu sorun, ham veri kümelerini içerebilecek veriler, işlenmiş veriler, meta veriler veya aday veriler gibi veri kümeleri arasında var olan çeşitlilikten kaynaklanmaktadır. Verimli analiz için bu veri kümeleri toplanmalı ve harmanlanmalıdır, ancak şu anda bunu yapmanın pratik bir yöntemi yoktur. Bu, ilaç tasarımı süreci başlamadan önce yapılmalıdır, çünkü uygun şekilde biçimlendirilmiş veriler olmadan

makine öğrenimi algoritmalarının çıktıları hatalı olmaktadır. Bu nedenle, ilaç tasarım süreci başlamadan önce mevcut verileri veri bankalarına entegre etmek için daha verimli yöntemler gerekmektedir.

Diğer bir sorun ise mesleki beceri ve tecrübe yetersizliğidir. Şu anda ilaç endüstrisinde çalışan birçok kişi, yapay zekâ sistemlerini çalıştırmak için gerekli becerilere veya niteliklere sahip değildir. Aynı şekilde yapay zekâ veri bilimindeki birçok kişi de moleküler kimya ve biyolojide yetkin değildir, uygun algoritmalar oluşturmak için kimya ve biyoloji bilgisiyle birlikte yapay zekâ alanında da uzmanlık gereklidir. Ancak çok azı her ikisinde de uzmandır ve yapay zekâyı farmasötik bağlamda uygulamak için doğru beceri kombinasyonuna sahiptir.

Bunlarla birlikte 'kara kutu' fenomeni olarak bilinen, model mekanizmasının belirsiz olmasıyla şeffaflık ve yorumlanabilirlik çoğu modelde yetersiz kalmaktadır ve sonuçları açıklamak için sınırlı yöntemlere sahiptir. Ayrıca, ilaç şirketleri tarafından üretilen büyük hacimli biyomedikal veriler genellikle halktan gizlenmekte ve pahalı özel ticari veriler olarak görüldüğünden bazı büyük veri setleri kolayca elde edilememektedir. Bu sebeplerden dolayı üretilen sonuçlara güven eksikliği oluşmaktadır ve ilaç endüstrisinde makine öğrenimi ve yapay zekâyı şüpheyle yaklaşılmaktadır. Yapay zekâyı olan bu güvensizlik ilaç endüstrisinde yapay zekânın gelişimi için finansman eksikliğine sebep olmaktadır. Bu durum, potansiyeline kıyasla daha yavaş, daha az verimli araştırma ve geliştirmeye yol açarak ilaç endüstrisinde yapay zekâ ile ilgili ilerlemelerde azalmaya yol açabilmektedir. Bunlar, gerçek gelişimin önünde duran farklı engellerdir ve yapay zekânın ilaç tasarımına entegre edilmesi için aşılması gereken zorluklardır.

Tüm bu zorluklara rağmen, birçok molekül tasarımı ve sentez aşamalarını ön görmesi bakımından ilaç tasarımında AI'nın kullanımının ciddi bir yer kapladığı görülmektedir. Son zamanlarda, OKB tedavisi için DSP-1181'in ve hiçbir veri setinde bulunmayan tamamen özgün bir antibiyotik molekülü olan Halicin'in tasarımı buna en güzel örneklerdir. DSP-1181 ile ilgili çalışmalar faz aşamalarına kadar başarılı bir şekilde ilerlemiş ancak faz 1 aşamalarında yetersiz bulunmuştur. Yine de bu çalışma, normalde onlarca yılı bulabilecekken sadece 12 ayda sonuçlanması büyük bir başarı olarak kabul edilmektedir. Geleneksel antibiyotiklerden yapısal olarak farklı olan Halicin ise *Mycobacterium tuberculosis* ve karbapenem dirençli *Enterobacteriaceae* dahil olmak üzere geniş bir patojen spektrumuna karşı bakterisidal aktivite gösteren bir molekül olarak tasarlanmıştır. Halicin *Clostridioides difficile* ve pan ilaç rezistanlı *Acinetobacter baumannii* enfeksiyonlarını etkili bir şekilde tedavi etmesinin görülmesinden sonra henüz deneysel aşamada olsa da kullanılabilir en güçlü antibiyotik molekülü özelliği taşımaktadır.

Bu örneklerde görüldüğü gibi ilaç tasarımında yapay zekâ teknolojilerinin kullanımı hem zamandan hem maliyetten önemli ölçüde tasarruf sağlayarak bu süreci ciddi bir şekilde destekleyip hızlandırmıştır. Bununla birlikte ilaç şirketleri ile yapay zekâ şirketlerinin iş birliği değer kazanmış olup ortak yürüttükleri çalışmaların gün geçtikçe arttığı açıkça görülmektedir. Bu durum ilaç tasarımına yeni bir bakış açısı ve canlılık kazandırmıştır. Yakın gelecekte tıbbi verilerin daha fazla birikmesi ve yapay zekâ algoritmalarının geliştirilmesiyle birlikte, yapay zekâ teknolojisinin yeni ilaç tasarımı ve geliştirmenin tüm alanlarını kapsamaması ve ilaç tasarımında en çok kullanılan yöntem haline gelmesi beklenmektedir.

YAZAR KATKILARI

Kavram: Ö.T., N.A.; Tasarım: Ö.T., N.A.; Denetim: Ö.T.; Kaynaklar: - ; Malzemeler: - ; Veri Toplama ve/veya İşleme: - ; Analiz ve/veya Yorumlama: Ö.T., N.A.; Literatür Taraması: N.A.; Makalenin Yazılması: Ö.T., N.A.; Kritik İnceleme: Ö.T.; Diğer: -

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar bu makale için gerçek, potansiyel veya algılanan çıkar çatışması olmadığını beyan ederler.

KAYNAKLAR

1. Lo, Y.C., Ren, G., Honda, H.L., Davis, K. (2020). Artificial intelligence-based drug design and discovery.

- Cheminformatics and Its Applications. [\[CrossRef\]](#)
2. Mandal, S., Moudgil, M., Mandal, S.K. (2009). Rational drug design. *European Journal of Pharmacology*, 625(1-3), 90-100. [\[CrossRef\]](#)
 3. Zhong, F., Xing, J., Li, X., Liu, X., Fu, Z., Xiong, Z., Lu, D., Wu, X., Zhao, J., Tan, X., Li, F., Luo, X., Li, Z., Chen, K., Zheng, M., Jiang, H. (2018). Artificial intelligence in drug design. *Science China Life Sciences*, 61(10), 1191-1204. [\[CrossRef\]](#)
 4. Gertrudes, J.C., Maltarollo, V.G., Silva, R.A., Oliveira, P.R., Honorio, K.M., da Silva, A.B.F. (2012). Machine learning techniques and drug design. *Current Medicinal Chemistry*, 19(25), 4289-4297. [\[CrossRef\]](#)
 5. Hessler, G., Baringhaus, K.H. (2018). Artificial intelligence in drug design. *Molecules*, 23(10), 2520. [\[CrossRef\]](#)
 6. Michie, D., Spiegelhalter, D.J., Taylor, C.C., Campbell, J. (Eds.) (1994). *Machine Learning, Neural and Statistical Classification* ABD: Ellis Horwood.
 7. Kaul, V., Enslin, S., Gross, S.A. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, 92(4), 807-812. [\[CrossRef\]](#)
 8. Intelligent drug discovery powered by AI, A report from the Deloitte Centre for Health Solutions. Retrieved 14.07.2023, from: <https://www2.deloitte.com/content/dam/Deloitte/my/Documents/risk/my-risk-sdg3-intelligent-drug-discovery.pdf>.
 9. Greenhill, A.T., Edmunds, B.R. (2020). A primer of artificial intelligence in medicine. *Techniques and Innovations in Gastrointestinal Endoscopy*, 22(2), 85-89. [\[CrossRef\]](#)
 10. Hoogenboom, S.A., Bagci, U., Wallace, M.B. (2020). Artificial intelligence in gastroenterology. The current state of play and the potential. How will it affect our practice and when? *Techniques and Innovations in Gastrointestinal Endoscopy*, 22(2), 42-47. [\[CrossRef\]](#)
 11. Le Berre, C., Sandborn, W.J., Aridhi, S., Devignes, M.D., Fournier, L., Smail-Tabbone, M., Danese, S., Peyrin-Biroulet, L. (2020). Application of artificial intelligence to gastroenterology and hepatology. *Gastroenterology*, 158(1), 76-94.e2. [\[CrossRef\]](#)
 12. LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. [\[CrossRef\]](#)
 13. Sarkar, C., Das, B., Rawat, V S., Wahlang, J.B., Nongpiur, A., Tiewsoh, I., Lyngdoh, N.M., Das, D., Bidarolli, M., Sony, H.T. (2023). Artificial intelligence and machine learning technology driven modern drug discovery and development. *International Journal of Molecular Sciences*, 24(3), 2026. [\[CrossRef\]](#)
 14. D'Souza, S., Prema, K.V., Balaji, S. (2020). Machine learning models for drug-target interactions: current knowledge and future directions. *Drug Discovery Today*, 25(4), 748-756. [\[CrossRef\]](#)
 15. Ippolito, M., Ferguson, J., Jenson, F. (2021). Improving facies prediction by combining supervised and unsupervised learning methods. *Journal of Petroleum Science and Engineering*, 200, 108300. [\[CrossRef\]](#)
 16. Linton-Reid, K. (2020). Introduction: An overview of AI in oncology drug discovery and development. In *Artificial Intelligence in Oncology Drug Discovery and Development*, Intechopen, pp:1-14. [\[CrossRef\]](#)
 17. Bohr, H. (2020). Drug discovery and molecular modeling using artificial intelligence. In *Artificial Intelligence in Healthcare* (pp. 61-83). Elsevier. [\[CrossRef\]](#)
 18. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., Blaschke, T. (2018). The rise of deep learning in drug discovery. *Drug Discovery Today*, 23(6), 1241-1250. [\[CrossRef\]](#)
 19. Jing, Y., Bian, Y., Hu, Z., Wang, L., Xie, X.Q.S. (2018). Deep learning for drug design: An artificial intelligence paradigm for drug discovery in the big data era. *The AAPS Journal*, 20(3), 58. [\[CrossRef\]](#)
 20. Gunavathi, C., Sivasubramanian, K., Keerthika, P., Paramasivam, C. (2021). A review on convolutional neural network based deep learning methods in gene expression data for disease diagnosis. *Materials Today: Proceedings*, 45, 2282-2285. [\[CrossRef\]](#)
 21. Hubel, D.H., Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3), 574-591. [\[CrossRef\]](#)
 22. Hubel, D.H., Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106-154. [\[CrossRef\]](#)
 23. Ramesh, A., Kambhampati, C., Monson, J., Drew, P. (2004). Artificial intelligence in medicine. *Annals of The Royal College of Surgeons of England*, 86(5), 334-338. [\[CrossRef\]](#)
 24. Amisha, Malik, P., Pathania, M., Rathaur, V. (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, 8(7), 2328. [\[CrossRef\]](#)
 25. Hamet, P., Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism*, 69, S36-S40. [\[CrossRef\]](#)
 26. Moran, M.E. (2007). Evolution of robotic arms. *Journal of Robotic Surgery*, 1(2), 103-111. [\[CrossRef\]](#)
 27. Weizenbaum, J. (1966). Eliza-a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45. [\[CrossRef\]](#)
 28. Kuipers, B., Feigenbaum, E.A., Hart, P.E., Nilsson, N.J. (2017). Shakey: From conception to history. *AI*

- Magazine, 38(1), 88-103. [CrossRef]
29. Kulikowski, C.A. (2015). An opening chapter of the first generation of artificial intelligence in medicine: the first rutgers aim workshop, june 1975. *Yearbook of Medical Informatics*, 24(01), 227-233. [CrossRef]
 30. Ferrucci, D., Levas, A., Bagchi, S., Gondek, D., Mueller, E.T. (2013). Watson: beyond jeopardy! *Artificial Intelligence*, 199-200, 93-105. [CrossRef]
 31. Comendador, B.E.V., Francisco, B.M.B., Medenilla, J.S., Nacion, S.M.T., Serac, T.B.E. (2015). Pharmabot: a pediatric generic medicine consultant chatbot. *Journal of Automation and Control Engineering*, 3(2), 137-140. [CrossRef]
 32. Ni, L., Lu, C., Liu, N., Liu, J. (2017). MANDY: Towards a smart primary care chatbot application. In: Chen, J., Theeramunkong, T., Supnithi, T., Tang, X. (eds) *Knowledge and Systems Sciences. Communications in Computer and Information Science*, vol 780. Springer, Singapore. [CrossRef]
 33. Artificial intelligence: Google's AlphaGo beats Go master Lee Sedol. In: *Technology. BBC NEWS*. 12 March 2016. from: <http://www.bbc.com/news/technology-35785875#> Erişim Tarihi: 14.07.2023
 34. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587), 484-489. [CrossRef]
 35. Chang, A.C. (2020). History of artificial intelligence in medicine. *Intelligence-Based Medicine*, (pp. 29-42). Academic Press. [CrossRef]
 36. Precision Medicine World Conference (PMWC) 2018 Silicon Valley PMWC Precision Medicine World Conferen. Retrieved 01.06.2023, from: <https://past.pmwintl.com/fabien-beckers-2018sv/>. Erişim Tarihi: 14.07.2023.
 37. Chat GPT. Retrieved 01.06.2023, from: <https://chat-gpt.org/tr>. Erişim Tarihi: 14.07.2023.
 38. Chat GPT. Retrieved 01.06.2023, from: <https://openai.com/research/gpt-4>. Erişim Tarihi: 14.07.2023.
 39. Drug Statics. Retrieved 01.06.2023, from: <https://go.drugbank.com/stats>. Erişim Tarihi: 14.07.2023.
 40. Dalkara, S., Saraç S. 2016, s.148-187 *Farmasötik Kimya 1. 4. Baskı*, Ankara:Hacettepe Üniversitesi.
 41. Tripathi, N., Goshisht, M.K., Sahu, S.K., Arora, C. (2021). Applications of artificial intelligence to drug design and discovery in the big data era: A comprehensive review. *Molecular Diversity*, 25(3), 1643-1664. [CrossRef]
 42. Ashburn, T.T., Thor, K.B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nature Reviews Drug Discovery*, 3(8), 673-683. [CrossRef]
 43. DiMasi, J.A., Grabowski, H.G., Hansen, R.W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47, 20-33. [CrossRef]
 44. Domingos, P., Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine Learning*, 29(2/3), 103-130. [CrossRef]
 45. Hou, T., Wang, J., Li, Y. (2007). Adme evaluation in drug discovery. 8. the prediction of human intestinal absorption by a support vector machine. *Journal of Chemical Information and Modeling*, 47(6), 2408-2415. [CrossRef]
 46. Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P. (2003). Random forest: A classification and regression tool for compound classification and qsar modeling. *Journal of Chemical Information and Computer Sciences*, 43(6), 1947-1958. [CrossRef]
 47. Rayhan, F., Ahmed, S., Shatabda, S., Farid, D.M., Mousavian, Z., Dehzangi, A., Rahman, M.S. (2017). İdti-esboost: identification of drug target interaction using evolutionary and structural features with boosting. *Scientific Reports*, 7(1), 17731. [CrossRef]
 48. Cao, D.S., Xu, Q.S., Liang, Y.Z., Chen, X., Li, H.D. (2010). Automatic feature subset selection for decision tree-based ensemble methods in the prediction of bioactivity. *Chemometrics and Intelligent Laboratory Systems*, 103(2), 129-136. [CrossRef]
 49. Lavecchia, A., Giovanni, C. (2013). Virtual screening strategies in drug discovery: A critical review. *Current Medicinal Chemistry*, 20(23), 2839-2860. [CrossRef]
 50. Hansch, C., Fujita, T. (1964). ρ - σ - π analysis. A method for the correlation of biological activity and chemical structure. *Journal of the American Chemical Society*, 86(8), 1616-1626. [CrossRef]
 51. Zefirov, N.S., Palyulin, V.A. (2002). Fragmental approach in qsar. *Journal of Chemical Information and Computer Sciences*, 42(5), 1112-1122. [CrossRef]
 52. McGregor, M.J., Muskal, S.M. (1999). Pharmacophore fingerprinting. 1. application to qsar and focused library design. *Journal of Chemical Information and Computer Sciences*, 39(3), 569-574. [CrossRef]
 53. Gozalbes, R., Doucet, J., Derouin, F. (2002). Application of topological descriptors in qsar and drug design: history and new trends. *Current Drug Target -Infectious Disorders*, 2(1), 93-102. [CrossRef]
 54. Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P.,

- Sainath, T., Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6), 82-97. [\[CrossRef\]](#)
55. Aoyama, T., Suzuki, Y., Ichikawa, H. (1989). Neural network applied to pharmaceutical problems. I. method and application to decision making. *Chemical and Pharmaceutical Bulletin*, 37(9), 2558-2560. [\[CrossRef\]](#)
56. Tetko, I.V., Villa, A.E.P., Aksenova, T.I., Zielinski, W.L., Brower, J., Collantes, E.R., Welsh, W.J. (1998). Application of a pruning algorithm to optimize artificial neural networks for pharmaceutical fingerprinting. *Journal of Chemical Information and Computer Sciences*, 38(4), 660-668. [\[CrossRef\]](#)
57. Tetko, I.V., Villa, A.E.P., Livingstone, D.J. (1996). Neural network studies. 2. variable selection. *Journal of Chemical Information and Computer Sciences*, 36(4), 794-803. [\[CrossRef\]](#)
58. Agatonovic-Kustrin, S., Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22(5), 717-727. [\[CrossRef\]](#)
59. Zhu, H. (2020). Big Data and artificial intelligence modeling for drug discovery. *Annual Review of Pharmacology and Toxicology*, 60(1), 573-589. [\[CrossRef\]](#)
60. Gawehn, E., Hiss, J.A., Schneider, G. (2016). Deep learning in drug discovery. *Molecular Informatics*, 35(1), 3-14. [\[CrossRef\]](#)
61. Ma, J., Sheridan, R.P., Liaw, A., Dahl, G.E., Svetnik, V. (2015). Deep neural nets as a method for quantitative structure-activity relationships. *Journal of Chemical Information and Modeling*, 55(2), 263-274. [\[CrossRef\]](#)
62. Mayr, A., Klambauer, G., Unterthiner, T., Hochreiter, S. (2016). Deeptox: Toxicity prediction using deep learning. *Frontiers in Environmental Science*, 3. [\[CrossRef\]](#)
63. Wu, Z., Ramsundar, B., Feinberg, E.N., Gomes, J., Geniesse, C., Pappu, A.S., Leswing, K., Pande, V. (2018). Moleculenet: A benchmark for molecular machine learning. *Chemical Science*, 9(2), 513-530. [\[CrossRef\]](#)
64. Minnich, A.J., McLoughlin, K., Tse, M., Deng, J., Weber, A., Murad, N., Madej, B.D., Ramsundar, B., Rush, T., Calad-Thomson, S., Brase, J., Allen, J.E. (2020). Ampl: A data-driven modeling pipeline for drug discovery. *Journal of Chemical Information and Modeling*, 60(4), 1955-1968. [\[CrossRef\]](#)
65. Mayr, A., Klambauer, G., Unterthiner, T., Steijaert, M., Wegner, J.K., Ceulemans, H., Clevert, D.-A., Hochreiter, S. (2018). Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. *Chemical Science*, 9(24), 5441-5451. [\[CrossRef\]](#)
66. Sheridan, R.P. (2013). Time-split cross-validation as a method for estimating the goodness of prospective prediction. *Journal of Chemical Information and Modeling*, 53(4), 783-790. [\[CrossRef\]](#)
67. Heffernan, R., Paliwal, K., Lyons, J., Dehzangi, A., Sharma, A., Wang, J., Sattar, A., Yang, Y., Zhou, Y. (2015). Improving prediction of secondary structure, local backbone angles and solvent accessible surface area of proteins by iterative deep learning. *Scientific Reports*, 5(1), 11476. [\[CrossRef\]](#)
68. Qian, N., Sejnowski, T.J. (1988). Predicting the secondary structure of globular proteins using neural network models. *Journal of Molecular Biology*, 202(4), 865-884. [\[CrossRef\]](#)
69. Qi, Y., Oja, M., Weston, J., Noble, W.S. (2012). A unified multitask architecture for predicting local protein properties. *PloS One*, 7(3), e32235. [\[CrossRef\]](#)
70. Spencer, M., Eickholt, J., Cheng, J. (2015). A deep learning network approach to ab initio protein secondary structure prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 12(1), 103-112. [\[CrossRef\]](#)
71. Wang, S., Peng, J., Ma, J., Xu, J. (2016). Protein secondary structure prediction using deep convolutional neural networks. *Scientific Reports*, 6(1), 18962. [\[CrossRef\]](#)
72. Jo, T., Hou, J., Eickholt, J., Cheng, J. (2015). Improving protein fold recognition by deep learning networks. *Scientific Reports*, 5(1), 17573. [\[CrossRef\]](#)
73. Dill, K.A., Ozkan, S.B., Shell, M.S., Weikl, T.R. (2008). The protein folding problem. *Annual Review of Biophysics*, 37(1), 289-316. [\[CrossRef\]](#)
74. Dill, K.A., MacCallum, J.L. (2012). The protein-folding problem, 50 years on. *Science*, 338(6110), 1042-1046. [\[CrossRef\]](#)
75. Wang, L., Ding, J., Pan, L., Cao, D., Jiang, H., Ding, X. (2019). Artificial intelligence facilitates drug design in the big data era. *Chemometrics and Intelligent Laboratory Systems*, 194, 103850. [\[CrossRef\]](#)
76. Senior, A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A.W., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792), 706-710. [\[CrossRef\]](#)
77. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Židek, A., Bridgland, A., Cowie, A.,

- Meyer, C., Laydon, A., Velankar, S., Kleywegt, G.J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S.A.A., Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, 596(7873), 590-596. [\[CrossRef\]](#)
78. Goshisht, M.K., Moudgil, L., Rani, M., Khullar, P., Singh, G., Kumar, H., Singh, N., Kaur, G., Bakshi, M. S. (2014). Lysozyme complexes for the synthesis of functionalized biomaterials to understand protein-protein interactions and their biological applications. *The Journal of Physical Chemistry C*, 118(48), 28207-28219. [\[CrossRef\]](#)
79. Goshisht, M.K., Moudgil, L., Khullar, P., Singh, G., Kaura, A., Kumar, H., Kaur, G., Bakshi, M.S. (2015). Surface adsorption and molecular modeling of biofunctional gold nanoparticles for systemic circulation and biological sustainability. *ACS Sustainable Chemistry & Engineering*, 3(12), 3175-3187. [\[CrossRef\]](#)
80. Khullar, P., Goshisht, M.K., Moudgil, L., Singh, G., Mandial, D., Kumar, H., Ahluwalia, G.K., Bakshi, M. S. (2017). Mode of protein complexes on gold nanoparticles surface: Synthesis and characterization of biomaterials for hemocompatibility and preferential dna complexation. *ACS Sustainable Chemistry & Engineering*, 5(1), 1082-1093. [\[CrossRef\]](#)
81. Mahal, A., Goshisht, M. K., Khullar, P., Kumar, H., Singh, N., Kaur, G., Bakshi, M.S. (2014). Protein mixtures of environmentally friendly zein to understand protein-protein interactions through biomaterials synthesis, hemolysis, and their antimicrobial activities. *Phys. Chem. Chem. Phys.*, 16(27), 14257-14270. [\[CrossRef\]](#)
82. Scott, D.E., Bayly, A.R., Abell, C., Skidmore, J. (2016). Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. *Nature Reviews Drug Discovery*, 15(8), 533-550. [\[CrossRef\]](#)
83. Azzarito, V., Long, K., Murphy, N.S., Wilson, A.J. (2013). Inhibition of α -helix-mediated protein-protein interactions using designed molecules. *Nature Chemistry*, 5(3), 161-173. [\[CrossRef\]](#)
84. Rao, V.S., Srinivas, K., Sujini, G.N., Kumar, G.N.S. (2014). Protein-protein interaction detection: methods and analysis. *International Journal of Proteomics*, 2014, 1-12. [\[CrossRef\]](#)
85. Du, T., Liao, L., Wu, C.H., Sun, B. (2016). Prediction of residue-residue contact matrix for protein-protein interaction with fisher score features and deep learning. *Methods*, 110, 97-105. [\[CrossRef\]](#)
86. Shin, W.H., Christoffer, C.W., Kihara, D. (2017). *In silico* structure-based approaches to discover protein-protein interaction-targeting drugs. *Methods*, 131, 22-32. [\[CrossRef\]](#)
87. Maheshwari, S., Brylinski, M. (2016). Template-based identification of protein-protein interfaces using eFindSitePPI. *Methods*, 93, 64-71. [\[CrossRef\]](#)
88. Vakser, I.A. (2014). Protein-protein docking: From interaction to interactome. *Biophysical Journal*, 107(8), 1785-1793. [\[CrossRef\]](#)
89. Mosca, R., Céol, A., Aloy, P. (2013). Interactome3D: Adding structural details to protein networks. *Nature Methods*, 10(1), 47-53. [\[CrossRef\]](#)
90. Du, X., Sun, S., Hu, C., Yao, Y., Yan, Y., Zhang, Y. (2017). Deepppi: Boosting prediction of protein-protein interactions with deep neural networks. *Journal of Chemical Information and Modeling*, 57(6), 1499-1510. [\[CrossRef\]](#)
91. Zeng, H., Wang, S., Zhou, T., Zhao, F., Li, X., Wu, Q., Xu, J. (2018). Complexcontact: A web server for inter-protein contact prediction using deep learning. *Nucleic Acids Research*, 46(W1), W432-W437. [\[CrossRef\]](#)
92. Xie, Z., Deng, X., Shu, K. (2020). Prediction of protein-protein interaction sites using convolutional neural network and improved data sets. *International Journal of Molecular Sciences*, 21(2), 467. [\[CrossRef\]](#)
93. Rester, U. (2008). From virtuality to reality-virtual screening in lead discovery and lead optimization: A medicinal chemistry perspective. *Current Opinion in Drug Discovery and Development*, 11(4), 559-568.
94. Walters, W.P., Stahl, M.T., Murcko, M.A. (1998). Virtual screening-an overview. *Drug Discovery Today*, 3(4), 160-178. [\[CrossRef\]](#)
95. Gonczarek, A., Tomczak, J.M., Zaręba, S., Kaczmar, J., Dąbrowski, P., Walczak, M.J. (2018). Interaction prediction in structure-based virtual screening using deep learning. *Computers in Biology and Medicine*, 100, 253-258. [\[CrossRef\]](#)
96. Plewczynski, D., Spieser, S., Koch, U. (2009). Performance of machine learning methods for ligand-based virtual screening. *Combinatorial Chemistry & High Throughput Screening*, 12(4), 358-368. [\[CrossRef\]](#)
97. Bohacek, R.S., McMartin, C., Guida, W.C. (1996). The art and practice of structure-based drug design: A molecular modeling perspective. *Medicinal Research Reviews*, 16(1), 3-50. [\[CrossRef\]](#)
98. Xiao, T., Qi, X., Chen, Y., Jiang, Y. (2018). Development of ligand-based big data deep neural network models for virtual screening of large compound libraries. *Molecular Informatics*, 37(11), 1800031. [\[CrossRef\]](#)
99. Ferreira, L., dos Santos, R., Oliva, G., Andricopulo, A. (2015). Molecular docking and structure-based drug design strategies. *Molecules*, 20(7), 13384-13421. [\[CrossRef\]](#)

100. Akbar, R., Jusoh, S.A., Amaro, R.E., Helms, V. (2017). Enri: A tool for selecting structure-based virtual screening target conformations. *Chemical Biology and Drug Design*, 89(5), 762-771. [\[CrossRef\]](#)
101. Cheng, T., Li, Q., Zhou, Z., Wang, Y., Bryant, S.H. (2012). Structure-based virtual screening for drug discovery: A problem-centric review. *The AAPS Journal*, 14(1), 133-141. [\[CrossRef\]](#)
102. Bengio, Y., Courville, A., Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828. [\[CrossRef\]](#)
103. Pereira, J.C., Caffarena, E.R., dos Santos, C.N. (2016). Boosting docking-based virtual screening with deep learning. *Journal of Chemical Information and Modeling*, 56(12), 2495-2506. [\[CrossRef\]](#)
104. Ferrero, E., Dunham, I., Sanseau, P. (2017). *In silico* prediction of novel therapeutic targets using gene-disease association data. *Journal of Translational Medicine*, 15(1), 182. [\[CrossRef\]](#)
105. DeepCodex: A deep code for gene expression data. Retrieved 01.06.2023, from: <http://deepcodex.org> Erişim Tarihi: 14.07.2023.
106. Donner, Y., Kazmierczak, S., Fortney, K. (2018). Drug repurposing using deep embeddings of gene expression profiles. *Molecular Pharmaceutics*, 15(10), 4314-4325. [\[CrossRef\]](#)
107. Duan, Q., Flynn, C., Niepel, M., Hafner, M., Muhlich, J.L., Fernandez, N.F., Rouillard, A.D., Tan, C.M., Chen, E.Y., Golub, T.R., Sorger, P.K., Subramanian, A., Ma'ayan, A. (2014). Lincs canvas browser: Interactive web app to query, browse and interrogate lincs 11000 gene expression signatures. *Nucleic Acids Research*, 42(W1), W449-W460. [\[CrossRef\]](#)
108. Xie, L., He, S., Song, X., Bo, X., Zhang, Z. (2018). Deep learning-based transcriptome data classification for drug-target interaction prediction. *BMC Genomics*, 19(S7), 667. [\[CrossRef\]](#)
109. Vanhaelen, Q., Mamoshina, P., Aliper, A.M., Artemov, A., Lezhnina, K., Ozerov, I., Labat, I., Zhavoronkov, A. (2017). Design of efficient computational workflows for *in silico* drug repurposing. *Drug Discovery Today*, 22(2), 210-222. [\[CrossRef\]](#)
110. Skalic, M., Martínez-Rosell, G., Jiménez, J., De Fabritiis, G. (2019). Playmolecule bindscope: Large scale cnn-based virtual screening on the web. *Bioinformatics*, 35(7), 1237-1238. [\[CrossRef\]](#)
111. Mendolia, I., Contino, S., Perricone, U., Ardizzone, E., Pirrone, R. (2020). Convolutional architectures for virtual screening. *BMC Bioinformatics*, 21(S8), 310. [\[CrossRef\]](#)
112. Esposito, E.X., Hopfinger, A.J., Madura, J.D. (2004). Methods for applying the quantitative structure-activity relationship paradigm (pp. 131-213). [\[CrossRef\]](#)
113. Myint, K.Z., Xie, X.Q. (2010). Recent advances in fragment-based qsar and multi-dimensional qsar methods. *International Journal of Molecular Sciences*, 11(10), 3846-3866. [\[CrossRef\]](#)
114. Lei, T., Li, Y., Song, Y., Li, D., Sun, H., Hou, T. (2016). Admet evaluation in drug discovery: 15. accurate prediction of rat oral acute toxicity using relevance vector machine and consensus modeling. *Journal of Cheminformatics*, 8(1), 6. [\[CrossRef\]](#)
115. Aoyama, T., Suzuki, Y., Ichikawa, H. (1990). Neural networks applied to pharmaceutical problems. III. Neural networks applied to quantitative structure-activity relationship (QSAR) analysis. *Journal of Medicinal Chemistry*, 33(9), 2583-2590. [\[CrossRef\]](#)
116. Dobchev, D., Pillai, G., Karelson, M. (2014). *In silico* machine learning methods in drug development. *Current Topics in Medicinal Chemistry*, 14(16), 1913-1922. [\[CrossRef\]](#)
117. Dong, J., Yao, Z.J., Zhu, M.F., Wang, N.N., Lu, B., Chen, A.F., Lu, A.P., Miao, H., Zeng, W.B., Cao, D.S. (2017). Chemsar: An online pipelining platform for molecular sar modeling. *Journal of Cheminformatics*, 9(1), 27. [\[CrossRef\]](#)
118. Dahl, G.E., Jaitly, N., Salakhutdinov, R. (2014). Multi-task Neural Networks for QSAR Predictions.
119. Tenorio-Borroto, E., Peñuelas Rivas, C.G., Vásquez Chagoyán, J.C., Castañedo, N., Prado-Prado, F.J., García-Mera, X., González-Díaz, H. (2012). Ann multiplexing model of drugs effect on macrophages; theoretical and flow cytometry study on the cytotoxicity of the anti-microbial drug gi in spleen. *Bioorganic Medicinal Chemistry*, 20(20), 6181-6194. [\[CrossRef\]](#)
120. Tenorio-Borroto, E., Peñuelas-Rivas, C.G., Vásquez-Chagoyán, J.C., Castañedo, N., Prado-Prado, F.J., García-Mera, X., González-Díaz, H. (2014). Model for high-throughput screening of drug immunotoxicity-study of the anti-microbial gi over peritoneal macrophages using flow cytometry. *European Journal of Medicinal Chemistry*, 72, 206-220. [\[CrossRef\]](#)
121. Speck-Planche, A., Cordeiro, M. (2013). Simultaneous modeling of antimycobacterial activities and admet profiles: a chemoinformatic approach to medicinal chemistry. *Current Topics in Medicinal Chemistry*, 13(14), 1656-1665. [\[CrossRef\]](#)
122. Speck-Planche, A., Dias Soeiro Cordeiro, M.N. (2017). Speeding up early drug discovery in antiviral research: a fragment-based *in silico* approach for the design of virtual anti-hepatitis C leads. *ACS Combinatorial Science*, 19(8), 501-512. [\[CrossRef\]](#)
123. Ramsundar, B., Kearnes, S., Riley, P., Webster, D., Konerding, D., Pande, V. (2015). Massively Multitask

- Networks for Drug Discovery.
124. Xu, Y., Ma, J., Liaw, A., Sheridan, R.P., Svetnik, V. (2017). Demystifying multitask deep neural networks for quantitative structure-activity relationships. *Journal of Chemical Information and Modeling*, 57(10), 2490-2504. [\[CrossRef\]](#)
 125. Zhao, Z., Qin, J., Gou, Z., Zhang, Y., Yang, Y. (2020). Multi-task learning models for predicting active compounds. *Journal of Biomedical Informatics*, 108, 103484. [\[CrossRef\]](#)
 126. Kharkar, P. (2010). Two-dimensional (2D) *in silico* models for absorption, distribution, metabolism, excretion and toxicity (ADME/T) in drug discovery. *Current Topics in Medicinal Chemistry*, 10(1), 116-126. [\[CrossRef\]](#)
 127. Wang, Y., Xing, J., Xu, Y., Zhou, N., Peng, J., Xiong, Z., Liu, X., Luo, X., Luo, C., Chen, K., Zheng, M., Jiang, H. (2015). *In silico* ADME/T modelling for rational drug design. *Quarterly Reviews of Biophysics*, 48(4), 488-515. [\[CrossRef\]](#)
 128. Xue, H., Li, J., Xie, H., Wang, Y. (2018). Review of drug repositioning approaches and resources. *International Journal of Biological Sciences*, 14(10), 1232-1244. [\[CrossRef\]](#)
 129. Kennedy, T. (1997). Managing the drug discovery/development interface. *Drug Discovery Today*, 2(10), 436-444. [\[CrossRef\]](#)
 130. Merlot, C. (2010). Computational toxicology-a tool for early safety evaluation. *Drug Discovery Today*, 15(1-2), 16-22. [\[CrossRef\]](#)
 131. Khanna, I. (2012). Drug discovery in pharmaceutical industry: productivity challenges and trends. *Drug Discovery Today*, 17(19-20), 1088-1102. [\[CrossRef\]](#)
 132. Tan, J.J., Cong, X.J., Hu, L.M., Wang, C.X., Jia, L., Liang, X.J. (2010). Therapeutic strategies underpinning the development of novel techniques for the treatment of HIV infection. *Drug Discovery Today*, 15(5-6), 186-197. [\[CrossRef\]](#)
 133. Tetko, I.V., Bruneau, P. (2004). Application of ALOGPS to predict 1-octanol/water distribution coefficients, logP, and logD, of AstraZeneca in-house database. *Journal of Pharmaceutical Sciences*, 93(12), 3103-3110. [\[CrossRef\]](#)
 134. Kortagere, S., Chekmarev, D., Welsh, W.J., Ekins, S. (2008). New predictive models for blood-brain barrier permeability of drug-like molecules. *Pharmaceutical Research*, 25(8), 1836-1845. [\[CrossRef\]](#)
 135. Obrezanova, O., Csányi, G., Gola, J.M.R., Segall, M.D. (2007). Gaussian processes: A method for automatic qsar modeling of adme properties. *Journal of Chemical Information and Modeling*, 47(5), 1847-1857. [\[CrossRef\]](#)
 136. Lombardo, F., Obach, R.S., DiCapua, F.M., Bakken, G.A., Lu, J., Potter, D.M., Gao, F., Miller, M.D., Zhang, Y. (2006). A hybrid mixture discriminant analysis-random forest computational model for the prediction of volume of distribution of drugs in human. *Journal of Medicinal Chemistry*, 49(7), 2262-2267. [\[CrossRef\]](#)
 137. Klon, A.E., Lowrie, J.F., Diller, D.J. (2006). Improved native bayesian modeling of numerical data for absorption, distribution, metabolism and excretion (ADME) property prediction. *Journal of Chemical Information and Modeling*, 46(5), 1945-1956. [\[CrossRef\]](#)
 138. Lusci, A., Pollastri, G., Baldi, P. (2013). Deep architectures and deep learning in chemoinformatics: The prediction of aqueous solubility for drug-like molecules. *Journal of Chemical Information and Modeling*, 53(7), 1563-1575. [\[CrossRef\]](#)
 139. Krewski, D., Acosta, D., Andersen, M., Anderson, H., Bailar, J.C., Boekelheide, K., Brent, R., Charnley, G., Cheung, V.G., Green, S., Kelsey, K.T., Kerkvliet, N.I., Li, A.A., McCray, L., Meyer, O., Patterson, R.D., Pennie, W., Scala, R.A., Solomon, G.M., Staff of Committee on Toxicity Test. (2010). Toxicity testing in the 21st century: A vision and a strategy. *Journal of Toxicology and Environmental Health, Part B*, 13(2-4), 51-138. [\[CrossRef\]](#)
 140. Clark, A.M., Dole, K., Coulon-Spektor, A., McNutt, A., Grass, G., Freundlich, J.S., Reynolds, R.C., Ekins, S. (2015). Open source bayesian models. 1. application to adme/tox and drug discovery datasets. *Journal of Chemical Information and Modeling*, 55(6), 1231-1245. [\[CrossRef\]](#)
 141. Wenlock, M.C., Carlsson, L.A. (2015). How experimental errors influence drug metabolism and pharmacokinetic qsar/qspr models. *Journal of Chemical Information and Modeling*, 55(1), 125-134. [\[CrossRef\]](#)
 142. Hughes, T.B., Miller, G.P., Swamidass, S.J. (2015). Modeling epoxidation of drug-like molecules with a deep machine learning network. *ACS Central Science*, 1(4), 168-180. [\[CrossRef\]](#)
 143. Xu, Y., Dai, Z., Chen, F., Gao, S., Pei, J., Lai, L. (2015). Deep learning for drug-induced liver injury. *Journal of Chemical Information and Modeling*, 55(10), 2085-2093. [\[CrossRef\]](#)
 144. Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., Cokelaer, T., Greninger, P., van Dyk, E., Chang, H., de Silva, H., Heyn,

- H., Deng, X., Egan, R.K., Liu, Q., Garnett, M.J. (2016). A landscape of pharmacogenomic interactions in cancer. *Cell*, 166(3), 740-754. [\[CrossRef\]](#)
145. Cortés-Ciriano, I., van Westen, G.J.P., Bouvier, G., Nilges, M., Overington, J.P., Bender, A., Malliavin, T. E. (2016). Improved large-scale prediction of growth inhibition patterns using the NCI60 cancer cell line panel. *Bioinformatics*, 32(1), 85-95. [\[CrossRef\]](#)
146. Lagunin, A., Zakharov, A., Filimonov, D., Poroikov, V. (2011). Qsar modelling of rat acute toxicity on the basis of pass prediction. *Molecular Informatics*, 30(2-3), 241-250. [\[CrossRef\]](#)
147. Soufan, O., Ba-Alawi, W., Afeef, M., Essack, M., Kalnis, P., Bajic, V.B. (2016). Drabal: Novel method to mine large high-throughput screening assays using bayesian active learning. *Journal of Cheminformatics*, 8(1), 64. [\[CrossRef\]](#)
148. Korotcov, A., Tkachenko, V., Russo, D.P., Ekins, S. (2017). Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. *Molecular Pharmaceutics*, 14(12), 4462-4475. [\[CrossRef\]](#)
149. Ramsundar, B., Liu, B., Wu, Z., Verras, A., Tudor, M., Sheridan, R.P., Pande, V. (2017). Is multitask deep learning practical for pharma? *Journal of Chemical Information and Modeling*, 57(8), 2068-2076. [\[CrossRef\]](#)
150. Altae-Tran, H., Ramsundar, B., Pappu, A.S., Pande, V. (2017). Low data drug discovery with one-shot learning. *ACS Central Science*, 3(4), 283-293. [\[CrossRef\]](#)
151. Li, X., Xu, Y., Lai, L., Pei, J. (2018). Prediction of human cytochrome P450 inhibition using a multitask deep autoencoder neural network. *Molecular Pharmaceutics*, 15(10), 4336-4345. [\[CrossRef\]](#)
152. Wenzel, J., Matter, H., Schmidt, F. (2019). Predictive multitask deep neural network models for adme-tox properties: Learning from large data sets. *Journal of Chemical Information and Modeling*, 59(3), 1253-1268. [\[CrossRef\]](#)
153. Novac, N. (2013). Challenges and opportunities of drug repositioning. *Trends in Pharmacological Sciences*, 34(5), 267-272. [\[CrossRef\]](#)
154. Chen, X., Yan, C.C., Zhang, X., Zhang, X., Dai, F., Yin, J., Zhang, Y. (2016). Drug-target interaction prediction: Databases, web servers and computational models. *Briefings in Bioinformatics*, 17(4), 696-712. [\[CrossRef\]](#)
155. Durán, F., Alonso, N., Caamaño, O., García-Mera, X., Yañez, M., Prado-Prado, F., González-Díaz, H. (2014). Prediction of multi-target networks of neuroprotective compounds with entropy indices and synthesis, assay, and theoretical study of new asymmetric 1,2-rasagiline carbamates. *International Journal of Molecular Sciences*, 15(9), 17035-17064. [\[CrossRef\]](#)
156. Kitchen, D.B., Decornez, H., Furr, J.R., Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews Drug Discovery*, 3(11), 935-949. [\[CrossRef\]](#)
157. Cao, D.S., Liu, S., Xu, Q.S., Lu, H.M., Huang, J.H., Hu, Q.N., Liang, Y.Z. (2012). Large-scale prediction of drug-target interactions using protein sequences and drug topological structures. *Analytica Chimica Acta*, 752, 1-10. [\[CrossRef\]](#)
158. Yao, Z.J., Dong, J., Che, Y.J., Zhu, M.F., Wen, M., Wang, N.N., Wang, S., Lu, A.P., Cao, D.S. (2016). Targetnet: A web service for predicting potential drug-target interaction profiling via multi-target sar models. *Journal of Computer-Aided Molecular Design*, 30(5), 413-424. [\[CrossRef\]](#)
159. Ding, H., Takigawa, I., Mamitsuka, H., Zhu, S. (2014). Similarity-based machine learning methods for predicting drug-target interactions: A brief review. *Briefings in Bioinformatics*, 15(5), 734-747. [\[CrossRef\]](#)
160. Cao, D.S., Zhang, L.X., Tan, G.S., Xiang, Z., Zeng, W.B., Xu, Q.S., Chen, A.F. (2014). Computational prediction of drug-target interactions using chemical, biological, and network features. *Molecular Informatics*, 33(10), 669-681. [\[CrossRef\]](#)
161. Byvatov, E., Fechner, U., Sadowski, J., Schneider, G. (2003). Comparison of support vector machine and artificial neural network systems for drug/non-drug classification. *Journal of Chemical Information and Computer Sciences*, 43(6), 1882-1889. [\[CrossRef\]](#)
162. Romero-Durán, F.J., Alonso, N., Yañez, M., Caamaño, O., García-Mera, X., González-Díaz, H. (2016). Brain-inspired cheminformatics of drug-target brain interactome, synthesis, and assay of TVP1022 derivatives. *Neuropharmacology*, 103, 270-278. [\[CrossRef\]](#)
163. Wen, M., Zhang, Z., Niu, S., Sha, H., Yang, R., Yun, Y., Lu, H. (2017). Deep-learning-based drug-target interaction prediction. *Journal of Proteome Research*, 16(4), 1401-1409. [\[CrossRef\]](#)
164. Luo, Y., Zhao, X., Zhou, J., Yang, J., Zhang, Y., Kuang, W., Peng, J., Chen, L., Zeng, J. (2017). A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nature Communications*, 8(1), 573. [\[CrossRef\]](#)
165. Luo, J. (2016). Crisp/Cas9: From genome engineering to cancer drug discovery. *Trends in Cancer*, 2(6), 313-324. [\[CrossRef\]](#)

166. Scott, A. (2018). How crispr is transforming drug discovery. *Nature*, 555(7695), S10-S11. [\[CrossRef\]](#)
167. Beck, B.R., Shin, B., Choi, Y., Park, S., Kang, K. (2020). Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Computational and Structural Biotechnology Journal*, 18, 784-790. [\[CrossRef\]](#)
168. Schneider, G., Fechner, U. (2005). Computer-based de novo design of drug-like molecules. *Nature Reviews Drug Discovery*, 4(8), 649-663. [\[CrossRef\]](#)
169. Bohm, H.J. (1992). The computer program ludi: A new method for the de novo design of enzyme inhibitors. *Journal of Computer-Aided Molecular Design*, 6(1), 61-78. [\[CrossRef\]](#)
170. Schneider, G., Geppert, T., Hartenfeller, M., Reisen, F., Klenner, A., Reutlinger, M., Hähnke, V., Hiss, J. A., Zettl, H., Keppner, S., Spänkuch, B., Schneider, P. (2011). Reaction-driven de novo design, synthesis and testing of potential type II kinase inhibitors. *Future Medicinal Chemistry*, 3(4), 415-424. [\[CrossRef\]](#)
171. Besnard, J., Ruda, G. F., Setola, V., Abecassis, K., Rodriguiz, R.M., Huang, X.P., Norval, S., Sassano, M. F., Shin, A.I., Webster, L.A., Simeons, F.R.C., Stojanovski, L., Prat, A., Seidah, N.G., Constam, D.B., Bickerton, G.R., Read, K.D., Wetsel, W.C., Gilbert, I.H., Hopkins, A.L. (2012). Automated design of ligands to polypharmacological profiles. *Nature*, 492(7428), 215-220. [\[CrossRef\]](#)
172. Miyao, T., Kaneko, H., Funatsu, K. (2016). Inverse qspr/qsar analysis for chemical structure generation (from y to x). *Journal of Chemical Information and Modeling*, 56(2), 286-299. [\[CrossRef\]](#)
173. Olivecrona, M., Blaschke, T., Engkvist, O., Chen, H. (2017). Molecular de-novo design through deep reinforcement learning. *Journal of Cheminformatics*, 9(1), 48. [\[CrossRef\]](#)
174. Kadurin, A., Nikolenko, S., Khrabrov, K., Aliper, A., Zhavoronkov, A. (2017). druGAN: An advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties *in silico*. *Molecular Pharmaceutics*, 14(9), 3098-3104. [\[CrossRef\]](#)
175. Gómez-Bombarelli, R., Wei, J.N., Duvenaud, D., Hernández-Lobato, J.M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T.D., Adams, R.P., Aspuru-Guzik, A. (2018). Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2), 268-276. [\[CrossRef\]](#)
176. Zhavoronkov, A., Ivanenkov, Y.A., Aliper, A., Veselov, M.S., Aladinskiy, V.A., Aladinskaya, A.V., Terentiev, V.A., Polykovskiy, D.A., Kuznetsov, M.D., Asadulaev, A., Volkov, Y., Zholus, A., Shayakhmetov, R.R., Zhebrak, A., Minaeva, L.I., Zagribelnyy, B.A., Lee, L. H., Soll, R., Madge, D., Aspuru-Guzik, A. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology*, 37(9), 1038-1040. [\[CrossRef\]](#)
177. Skalic, M., Jiménez, J., Sabbadin, D., De Fabritiis, G. (2019). Shape-based generative modeling for de novo drug design. *Journal of Chemical Information and Modeling*, 59(3), 1205-1214. [\[CrossRef\]](#)
178. Camodeca, C., Nuti, E., Tepshi, L., Boero, S., Tuccinardi, T., Stura, E.A., Poggi, A., Zocchi, M.R., Rossello, A. (2016). Discovery of a new selective inhibitor of a disintegrin and metalloprotease 10 (ADAM-10) able to reduce the shedding of NKG2D ligands in Hodgkin's lymphoma cell models. *European Journal of Medicinal Chemistry*, 111, 193-201. [\[CrossRef\]](#)
179. Healy, E.F., Romano, P., Mejia, M., Lindfors, G. (2010). Acetylenic inhibitors of ADAM10 and ADAM17: *in silico* analysis of potency and selectivity. *Journal of Molecular Graphics and Modelling*, 29(3), 436-442. [\[CrossRef\]](#)
180. Tippmann, F., Hundt, J., Schneider, A., Endres, K., Fahrenholz, F. (2009). Up-regulation of the α -secretase ADAM10 by retinoic acid receptors and acitretin. *The FASEB Journal*, 23(6), 1643-1654. [\[CrossRef\]](#)
181. Altmeppen, H.C., Prox, J., Krasemann, S., Puig, B., Kruszewski, K., Dohler, F., Bernreuther, C., Hoxha, A., Linsenmeier, L., Sikorska, B., Liberski, P.P., Bartsch, U., Saftig, P., Glatzel, M. (2015). The sheddase ADAM10 is a potent modulator of prion disease. *ELife*, 4. [\[CrossRef\]](#)
182. Kohutek, Z.A., diPierro, C.G., Redpath, G.T., Hussaini, I.M. (2009). ADAM-10-mediated n-cadherin cleavage is protein kinase c- α dependent and promotes glioblastoma cell migration. *The Journal of Neuroscience*, 29(14), 4605-4615. [\[CrossRef\]](#)
183. Woods, N., Trevino, J., Coppola, D., Chellappan, S., Yang, S., Padmanabhan, J. (2015). Fendiline inhibits proliferation and invasion of pancreatic cancer cells by interfering with ADAM10 activation and β -catenin signaling. *Oncotarget*, 6(34), 35931-35948. [\[CrossRef\]](#)
184. Shi, T., Huang, S., Chen, L., Heng, Y., Kuang, Z., Xu, L., Mei, H. (2020). A molecular generative model of ADAM10 inhibitors by using GRU-based deep neural network and transfer learning. *Chemometrics and Intelligent Laboratory Systems*, 205, 104122. [\[CrossRef\]](#)
185. Green, D.V.S., Pickett, S., Luscombe, C., Senger, S., Marcus, D., Meslamani, J., Brett, D., Powell, A., Masson, J. (2020). Bradshaw: A system for automated molecular design. *Journal of Computer-Aided Molecular Design*, 34(7), 747-765. [\[CrossRef\]](#)
186. Szymkuć, S., Gajewska, E.P., Klucznik, T., Molga, K., Dittwald, P., Startek, M., Bajczyk, M., Grzybowski,

- B.A. (2016). Computer-assisted synthetic planning: The end of the beginning. *Angewandte Chemie International Edition*, 55(20), 5904-5937. [CrossRef]
187. Segler, M.H.S., Preuss, M., Waller, M.P. (2018). Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698), 604-610. [CrossRef]
188. Coley, C.W., Green, W.H., Jensen, K.F. (2018). Machine learning in computer-aided synthesis planning. *Accounts of Chemical Research*, 51(5), 1281-1289. [CrossRef]
189. Button, A., Merk, D., Hiss, J.A., Schneider, G. (2019). Automated de novo molecular design by hybrid machine intelligence and rule-driven chemical synthesis. *Nature Machine Intelligence*, 1(7), 307-315. [CrossRef]
190. Yuan, W., Jiang, D., Nambiar, D.K., Liew, L.P., Hay, M.P., Bloomstein, J., Lu, P., Turner, B., Le, Q.T., Tibshirani, R., Khatri, P., Moloney, M.G., Koong, A.C. (2017). Chemical space mimicry for drug discovery. *Journal of Chemical Information and Modeling*, 57(4), 875-882. [CrossRef]
191. Segler, M.H.S., Kogej, T., Tyrchan, C., Waller, M.P. (2018). Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Central Science*, 4(1), 120-131. [CrossRef]
192. Merk, D., Friedrich, L., Grisoni, F., Schneider, G. (2018). De novo design of bioactive small molecules by artificial intelligence. *Molecular Informatics*, 37(1-2), 1700153. [CrossRef]
193. Papadatos, G., Gaulton, A., Hersey, A., Overington, J.P. (2015). Activity, assay and target data curation and quality in the ChEMBL database. *Journal of Computer-Aided Molecular Design*, 29(9), 885-896. [CrossRef]
194. Olley, D. (ed.) *Artificial intelligence: How knowledge is created, transferred, and used* (Elsevier, 2019).
195. Perron, Q. Deep learning for ligand-based de novo design in lead optimization: A real life case study. Presented at the XXV EFMC International Symposium on Medicinal Chemistry (2018).
196. Rodrigues, T., Hauser, N., Reker, D., Reutlinger, M., Wunderlin, T., Hamon, J., Koch, G., Schneider, G. (2015). Multidimensional de novo design reveals 5-HT_{2B} receptor-selective ligands. *Angewandte Chemie International Edition*, 54(5), 1551-1555. [CrossRef]
197. Reutlinger, M., Rodrigues, T., Schneider, P., Schneider, G. (2014). Multi-objective molecular de novo design by adaptive fragment prioritization. *Angewandte Chemie International Edition*, 53(16), 4244-4248. [CrossRef]
198. Gao, K., Nguyen, D.D., Tu, M., Wei, G.W. (2020). Generative network complex for the automated generation of drug-like molecules. *Journal of Chemical Information and Modeling*, 60(12), 5682-5698. [CrossRef]
199. Trobe, M., Burke, M.D. (2018). The molecular industrial revolution: Automated synthesis of small molecules. *Angewandte Chemie International Edition*, 57(16), 4192-4214. [CrossRef]
200. Baranczak, A., Tu, N.P., Marjanovic, J., Searle, P.A., Vasudevan, A., Djuric, S.W. (2017). Integrated platform for expedited synthesis-purification-testing of small molecule libraries. *ACS Medicinal Chemistry Letters*, 8(4), 461-465. [CrossRef]
201. Cox, G., Sieron, A., King, A.M., De Pascale, G., Pawlowski, A.C., Koteva, K., Wright, G.D. (2017). A common platform for antibiotic dereplication and adjuvant discovery. *Cell Chemical Biology*, 24(1), 98-109. [CrossRef]
202. Camacho, D.M., Collins, K.M., Powers, R.K., Costello, J.C., Collins, J.J. (2018). Next-generation machine learning for biological networks. *Cell*, 173(7), 1581-1592. [CrossRef]
203. De, S.K., Stebbins, J.L., Chen, L.H., Riel-Mehan, M., Machleidt, T., Dahl, R., Yuan, H., Emdadi, A., Barile, E., Chen, V., Murphy, R., Pellecchia, M. (2009). Design, synthesis, and structure-activity relationship of substrate competitive, selective, and in vivo active triazole and thiadiazole inhibitors of the c-jun n-terminal kinase. *Journal of Medicinal Chemistry*, 52(7), 1943-1952. [CrossRef]
204. Stokes, J.M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N.M., MacNair, C.R., French, S., Carfrae, L.A., Bloom-Ackermann, Z., Tran, V.M., Chiappino-Pepe, A., Badran, A.H., Andrews, I.W., Chory, E.J., Church, G.M., Brown, E.D., Jaakkola, T.S., Barzilay, R., Collins, J.J. (2020). A deep learning approach to antibiotic discovery. *Cell*, 180(4), 688-702.e13. [CrossRef]
205. Malandraki-Miller, S., Riley, P.R. (2021). Use of artificial intelligence to enhance phenotypic drug discovery. *Drug Discovery Today*, 26(4), 887-901. [CrossRef]
206. DSP-1181. Retrieved 14.07.2023, from: <https://www.exscientia.ai/dsp-1181>.
207. Wills T. AI drug discovery: Assessing the first AI-designed drug candidates to go into human clinical trials, CAS, 2022. Retrieved 14.07.2023, from: <https://www.cas.org/resources/cas-insights/drug-discovery/ai-designed-drug-candidates>.