5 dysfunctions of a team summary pdf free pdf templates

I'm not robot!

```
In computer science, Artificial intelligence (AI) additionally attributed as machine intelligence because machine are trained or customized to perform activities like a human brain (Poole et al. 1998; Vinod and Anand 2021; Gopal 2018). Artificial Intelligence (AI) can be categorized here as the field is dealing with a wide range of utilization and layouts
of numerous algorithms for interpreting and attaining knowledge from data. And the AI concept is firmly related to many fields like pattern recognition, probability theory, statistics, machine learning, and numerous procedures like fuzzy models, neural networks which are collectively known as "Computational Intelligence" Vinod and Anand (2021),
Engelbrecht (2007), Konar (2006), Duda et al. (2012), Webb (2003), Friedman et al. (2012), Webb (201
be characterized along with parameters. So, machines can be gain proficiency in the model with accessible parameters through the utilization of trained data. Furthermore, the model can predict the data in the future for recovering information from data (Alpaydin 2020). In this review, we are primarily focusing on qualities of AI approaches that are
 appropriate for drug development and discovery (Duch et al. 2007). Recently various factors were developed due to greater enthusiasm for utilizing machine learning approaches in the pharmaceutical industry. Figure 1 shows that the various fields of Drug Discovery and advancements utilized through machine learning. Every phase was performed
like a pipeline to represent therapeutic concepts. The respective phases represent unique iterations in time and cost expenditure. Here each phase is carried out to prove the effectiveness of the remedial treatment. The medical information was being mined and estimated accurately by using some 'omics' and 'smart automation tools'. Enlarging these
techniques into the biological field gives more opportunities as well as challenges in the pharmaceutical industry. Since numerous pharmaceutical industry is to distinguish the persuasive clinical hypothesis. With the obtained results, practitioners or clinicians can develop the medications. For establishing any type of drug in
pharmaceutical industries, the usage of machine learning approaches has checked out the performance. At this point, if included with unlimited storage, improvement appeared in datasets like size, types can provide premises to machine learning. In this way, it can access enormous data from pharmaceutical industries. Data types can have different
configurations like textual data, images, assay information, biometrics, and furthermore high dimensional omics data (Mamoshina et al. 2018). Thus, the AI field has developed from theoretical knowledge to real-world data. Information was widely improved for utilizing in PC hardware, for example, Graphical Processing Units (GPU), which makes
faster in processing (i.e., in computational techniques). Recently, the deep learning model is one of the machine learning algorithms (LeCun et al. 2015), it develops the models for making more accomplishment in broad daylight challenges (Chen et al. 2018; Hinton 2018). For the past 2 years, the usage of ML algorithms has a great extension within
pharmaceutical enterprises. In the clinical field, developing a new drug for persistent disease primarily relied on new medications. As of late, various drugs are improvised for recognizing dynamic components from traditional treatments such as penicillin. In chemical laboratories, it consists of natural substances, small molecules that aid in
 therapeutic medicine to detect substances such as cells or intact organisms. This procedure is called old-style pharmacology. High throughput screening with multiple libraries has normally expanded because of the human genome has permitted cloning strategies and furthermore improving refining of proteins in huge quantities. Screening activity for
large compounds through biological targets can be used to achieve a change in a disease called reverse pharmacology. Multiple hits can be generated from screening activity to provide cells and furthermore tests have been conducted in creatures for adequacy. In modern days, drug discovery has engaged with the performance of identifications on
screening hits, optimization techniques can build the drug effectiveness, affinity, stability of metabolic. If all requirements are satisfied by the compound, a particular drug will be developed in clinical trials if the drug is successful. In process of drug development and discovery, it requires lead optimization, target identification and validation, hit
discovery, clinical trials (Vohora and Singh 2018). In novel drug development, the cost expenditure can approximately 2.558 billion USD (DiMasi et al. 2016) and it is a tedious procedure in light of the fact that about 10–15 years have taken for selling in the market (Turner 2010). To accomplish a small number of molecules in drug development, many
investors are putting a lot of cash in developing exact progress in clinical trials. And still, 13% precision rate is lagging with disappointment. So as to conquer this issue, clinicians have utilized the Computer-assisted Drug Design CADD technique (Hassan Baig et al. 2016). By utilizing this strategy in drug discovery, the artificial techniques not just
provide the molecular properties (i.e., selectivity, distribution, absorption, bioactivity, metabolism, side effects, and excretion in the theoretical levels) but also provides the lead compounds such as ideal attributes in silico. Also, attrition cost in the preclinical state can be decreased through the utilization of multi-objective optimization techniques. In
drug discovery, computational intelligence provides various techniques for analyzing, learning and furthermore clarifies how such pharmaceutical was identified with AI for finding numerous medications in a programmed and integrated format (Duch et al. 2007). Therefore, many pharmaceutical industries have shown greater enthusiasm for
contributing to technologies, resources for retrieving accurate results in drug discovery. At last, this survey proposes AI techniques in the drug discovery area for targeting multiple applications in drug discovery and development by utilizing deep learning techniques. Along these lines, the AI field provides expected outcomes in concern of
computational intelligence in drug development and discovery (Table 1). List of Major Abbreviations ADMETAbsorption, Distribution, Metabolism And ExcretionToxicologyAEAutoEncoderAIArtificial IntellingenceANNArtificial Neural NetworksAUCArea under the ROC CurveCNNConvolution Neural NetworksCTComputed TomographyDLDeep
LearningDNNDeep Neural NetworksDPIDrug Protein InteractionGPCRG-Protein coupled receptorsGPUGraphical Processing UnitHARFHeterogeneity Aware Random ForestHTVSHigh-Throughput Virtual ScreeningLDALinear Discriminant AnalysisLRLogistic RegressionMAEMean Absolute ErrorMLMachine LearningMLPMulti Layer
 PerceptronMLRMultiple Linear RegressionMRIMagnetic Resonance ImagingNBCNaive Bayesian ClassificationNCENew Chemical EntitiesPNNProbabilistic Neural NetworksPPIProtein to Protein InteractionQSARQuantitative Structure-Activity RelationshipRBNRadial Basis function NetworksPRandom ForestRMSERoot Mean Square
 ErrorRNARibonucleic AcidRNNRecurrent Neural NetworksROCReceiver Operating Characteristic curveSARS-CoV-2Severe Acute Respiratory Syndrome CoronaVirus 2SMILESSimplified Molecular Input Line Entry SpecificationsSVMSupport Vector MachinesVAEVariational AutoEncodersThe rest of the article is arranged in the following way: Sect. 2
describes the application of AI in Drug design. Then, the various machine learning methods towards Drug discovery are discussed in Sect. 3. Various Drug design problems have discussed. Finally, Sect. 6 presents the research challenges with few possible suggestions in Drug
 discovery using Machine learning, and Sect. 7 concludes the article and provided some future directions. This section discusses a few application in drug design. Many impurities have appeared in the human body due to protein dysfunctions. Structural
drug design strategies are used to differentiate small molecules in protein targets. Protein targets. Protein structure in 3D format requires more exactness over de-novo prediction in 3D structure. By using deep learning and feature extraction tools, it is mandatory to
predict the secondary structure (Spencer et al. 2014) and residing the protein contacts (Li et al. 2017). It precisely gains the information on the connection among structure by utilizing deep learning techniques for improving the accuracy. To retrieve
 information from drug design of protein-protein computer structure, then it is mandatory to conduct investigations on PPI interface (Xue et al. 2015). Artificial Intelligence has been used in various applications like a prediction on drug-protein interactions, the discovery of drug efficacy, ensuring the safety biomarkers. The detailed discussion is given
as followsThe crucial step of drug development in silico is consisting of multiple biological sources for predictions, which relied on the countless unknown interactions. Therefore, semi-supervised training techniques should be used to address these unlabelled and labeled
date complications. Usually, only labeled data will produce better results. In addition, the semi-supervised technology integrates chemical structure, drug-protein interactions of various data sets such as ions, enzymes, and nuclear receptors provided well
predictable results (Xia et al. 2010). Drugs have an important priority in therapeutic activity, which is regulated by protein interactions. The drug-protein interactions. The drug-protein interaction database (DPI) focuses primarily on therapeutic protein targets, while knowledge gap to non-targets have an important priority in therapeutic activity, which is regulated by protein interactions. The drug-protein interactions are interactions.
for predicting protein targets for distributed drug molecules. In that study, the pool of 35 predictors had a major impact on the similarity between protein and drug profile are three types of similarity developed from the results of 35 predictors. Finally, the significant content, relationships, and
implications between database sources are of great importance for therapeutic activity (Wang and Kurgan 2020). In drug repurposing, while drug side effects are unavoidable and about 1,000 human proteins can cause critical
side effects. The proteomic scale method was used to predict side effects and protein goals. FINDSITEcomb is used to predict drug-protein interactions. The estimates showed greater disruption with a mean of 329 human targets for each drug (Zhou et al. 2015). Usually, a drug effect assessment looks at its biochemical activity. The effectiveness of
the therapeutic activity has posed a challenge to be properly coupled with the biochemical activity. The collection of a large amount of data on the effects of cellular drugs was undertaken to fill a gap that has been explored in the extensive content of cellular estimations and while this estimation is classified as a psychotropic drug. Here, the
microarray data can be analyzed by applying random trees to the forest and classifying them, providing a profile for the efficiency of biomarker gene expression. Accuracy of 88.9% of the random forest model used this efficacy profile for a drug treatment analysis. Therefore, at the cellular assessment level, general
genomic data are acceptable to reconcile the effects of new physiological drugs with clinical applications. Finally, in vitro signatures of gene expression data can identify the effectiveness of therapeutic activities that can help validating new
drugs requires predicting effectiveness and identifying targets. The proximity of medical illnesses helps to reduce the effectiveness of the treatment and also releases drugs to demonstrate the drug's effectiveness in therapeutic activity, as well as problems with
gene efficacy and various disorders. Here the network-based system is used to develop a drug-disease proximate measure that assesses the interactions between the disease and the drug diseases, offering a wide range of possibilities for
 conflict detection and drug repurposing (Guney et al. 2016). In drug development, the use of biomarkers supports the provision of safety measures that critically determine the biological and analytical indicators of a particular purpose and
 whether the desired standards are being met. For shareholders in the implementation of evaluating the experiment agreement, a stakeholder evaluation process is needed to adjust the unique characteristics of the biomarkers, as well as to determine how these innovations are analyzed, integrated, and interpreted, and how improved biomarkers and
conventional comparators are measured (Sistare et al. 2010). In the survey, we found that modern medicines are no safer than older drugs, even though with longer medical trial programs. These trails are placed on the market and impractical inspections are carried out which are not sufficient to be carried out systematically to ensure safety. Previous
drug-related signals can help in improving drug safety as well as identify underlying biomarkers, making them more toxic. However, the safety markers can be different for different for different target systems. However, the safety markers can be different for different target systems. However, the safety markers can be different for different for different target systems.
assessment by communicating with the PPI interface and show fantastic results when contrasted with the SVM technique (Du et al. 2016). Thus, the PPI's became more complex to utilize in biological techniques (Falchi et al. 2014; Scott et al. 2016). Each
PPI can be a mixture of various residues (Cukuroglu et al. 2014). New PPI can act as a modern class for pharmaceutical targets where disparate for different targets i.e., ion channels, GPCRs (G-Protein coupled receptors), kinases (Higueruelo et al. 2013; Santos et al. 2017). iFitDock is a docking tool used for investigating a few hotspots in PPIs.
Further, AI techniques have been utilized for distinguishing structures and hotspots in PPI interface (Fig. 2).AI innovation has a high priority in drug design through the enhancement of ML approaches and the collection of pharmacological data. AI does not rely upon any hypothetical improvements, but it has more essence in transforming medical
 information into studies like reusable methods. In general, there are different approaches such as Random Forest, Naive Bayesian Classification (NBC), Multiple Linear Regression (LR), Linear Discriminant Analysis (LDA), Probabilistic Neural Networks (PNN), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM),
 etc are considered in the context of ML (Lavecchia and Di Giovanni 2013). In order to gain capability in feature extraction and feature generalization, Al advancements are specifically used as a deep learning technique towards drug design. Also, Fig. 3 shows respective applications which illustrate an outline of AI procedures utilized to respond to
drug discovery queries in the review. A scope of classifier and regression strategies i.e. supervised learning techniques utilized to respond addresses desire expectations in continuous or categorical data factors, also unsupervised methods utilized in creating a model which empowers the clustering data. Many designed features in traditional ML
models are performed manually, but deep learning approaches will accelerate various features through available initialized data automatically because multi-layer features into complex features. One advantage of using deep learning approaches was, presence of low quantity
generalization blunders, so it recovers more exact results. CNN, RNN, Auto Encoder, DNN, and RBN are considered as different deep learning techniques. Summary of deep learning techniques which
are available in Deep Learning literature (Goodfellow et al. 2016). In drug discovery and development, many AI calculations are associated to analyse and predict the data. Here, few popular models like SVM, RF, and MLP discuss their effective use in drug discovery. SVM model is a supervised learning algorithm basically utilized in predicting the class
labelled data i.e., binary data. In SVM, x is considered as feature vector i.e., output for svm. Y \in {-1,1}. Here, Binary values are considered as classification task. Parameters in SVM u and b have considered for learning data in training set. In dataset,
(x(i),Y(i)) are considered as ith sample. Y can be represented as follows:Y(i)=if-1 and uTx+b=1 and uTx+b=1 and uTx+b=1 should be
maximized. In this way, Distance=2||U||maxU. Finally, it have to solve maxU2||U||minU2||U||Complete x(i) samples need to classify effectively in the SVM i.e., Y(i))(UTX(i)+b)\geqslant 1V\in 1,2,3,...N. The above equation was a hard-margin SVM, and we can
avoid this problem through applying linearly separable method. Using the slack variable \varepsilon(i) as constraints. In training data, each sample has its own slack variable \tau(i) as constraints. In training data, each sample has its own slack variable \tau(i) as constraints. In training data, each sample has its own slack variable \tau(i) as constraints. In training data, each sample has its own slack variable \tau(i) as constraints. In training data, each sample has its own slack variable \tau(i) as constraints. In training data, each sample has its own slack variable \tau(i) as constraints. In training data, each sample has its own slack variable.
function \phi to allow more flexibility in mapping. So, it maps multiple features like original space to high dimensional space (Noble 2006). Then, the quadratic optimization problem updates Eq. 1 as the following: Therefore, Y(i)(UT\phiX(i)+b)\geqslant1-\varepsilon(i), Vi\in1,2,3,...,NThe SVM widely used in drug discovery using its various kernels (Smola and Schölkopf 2004).
Various problems like Screen radiation protection and Gene Interaction using SVM-RBF(Radial Basis Function) (Matsumoto et al. 2011), Identify drug target interaction by Biased SVM (Wang et al. 2017), Predicting drug sensitivity prediction by Ensemble SVM,
and the Linear SVM used in Identify novel drug targets (Volkamer et al. 2012), Anti/non-anticancer molecule classification (Kapoorb et al. 2021). The SVM approach (Huang et al. 2021). The SVM approach (Huang et al. 2018) was used to quantify anti-cancer drugs based on cancer cell properties. To understand the relationship between cancer
cell properties and drug resistance, 24 drugs were tested on cancer cell lines (Gupta et al. 2016). In the treatment of oral cancer, the SVM-RBF (Radial Basis Function) approach has been used to find therapeutic compounds from a large collection of public databases (Bundela et al. 2015), the RBF is the popular kernel function used in various learning
algorithms. The RBF kernel takes two samples S1 and S2, represented as feature vectors in some input space K(S1,S2) = \exp(||S1-S2||22\sigma 2) where ||S1-S2||22\sigma 2| where |S1-S2||22\sigma 2|
values. In general, radiation therapy techniques help to protect against cancer. Therefore, the SVM method is used in virtual screening (Matsumoto et al. 2014). In this study, we found that the SVM approach worked better than
other techniques. When the target protein is known, we can find a suitable compound for the target protein. However, the SVM technique is mainly used to predict the outcome of targeted drugs. SVM has used sites to link global descriptors, taking into account various properties such as compactness and size. These descriptors can determine drug
scores for novel targets (Volkamer et al. 2012; Li et al. 2011). In therapeutic activities, the use of SVM helps to find different lead series
in the active compound to improve them in parallel in therapeutic activity (Warmuth et al. 2003). In contrast to other artificial neural networks, SVM demonstrated the ability to test drug similarity predictions of a wide variety of compounds. Because of this set of descriptors, the SVM outperformed the task and also reported that the SVM model
predicted better enzyme inhibitor quality for conventional QSAR (Zernov et al. 2003). Right now, the SVM model is the best methodology for prediction on and turned out to be more famous in drug discovery applications like a prediction on
                               lassification (Maltarollo et al. 2019). In designing new structures, the SVM approach was utilized for retrieving higher predicted results where depend on ligands (Hartenfeller and Schneider 2010). In the Activity process, to improve scoring capacity execution, the SVM approach was utilized for retrieving non-linear
relationships of energy terms from eHiTS and binding data which shows a lot of improvement in scoring power (Kinnings et al. 2007). SVM model was frequently utilized in virtual screening (Leelananda and Lindert 2016; Liew et al. 2009) and demonstrated best results (in the predicted
ratio called hits) and furthermore false-hit rates are decreased concurrently (counterfeit hit rates in the predicted hits) (Ma et al. 2019). The Random Forest algorithm was a
supervised algorithm. The name itself says, "This is a way of creating a forest from various perspectives to make it random". The significant advantage of the Random Forest algorithm was, it can relevant for both regression and classification issues. In the procedure of regression and classification tasks, overfitting can happen normally, so the
training data for fitting random forest tree. Alternate samples with n trained examples from X, Y then Xa,Ya.Classification tree fb must be training the data, invisible samples x' need to be predicted by averaging all individual trees on x'.In classification trees, majority voting can be considered. Finally,
random forest model produces better results due to the absence of increment in bias, it reduces variance in the model. The equation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression tree on x' can be represented in standard deviation for individual regression for individual r
Representation of scientific models and chemical structures are the fundamental issues in QSAR model (Dudek et al. 2006). At that point when descriptors are chosen, it is necessary to establish the best mathematical model (A Dobchev et al. 2006).
2014; Ning and Karypis 2011), a random forest algorithm was utilized (Fig. 4). The selection of molecular descriptors is seen as an important step in virtual screening to identify bioactive molecular descriptors shows predictions with lower accuracy. Hence, the random forest technique was
used to improve prediction and then select naturally trained molecular descriptors for kinase ligands, hormone receptors, enzymes, etc. (Cano et al. 2017). In the pharmaceutical industry, when developing drugs, the question that arises naturally is whether a prediction model trained with heterogeneous data is implemented as a similar prediction
model. Then the heterogeneity data were compiled for forecasting and model training. In this study, heterogeneity was treated as a problem with the latent distributed by means of an ensemble leaf node model. In general, an ensemble-based random forest model has
incorporated Heterogeneity Aware Random Forest (HARF) and assign specific weights to tree-based categories. Of course, the technique proposed by HARF gives better results than classical random forest, whereas drug feedback with the cancer disease types is something peculiar (Rahman et al. 2017). Immune network technology is to determine
new compounds from drug molecules. Using examples of sulfonamide properties, sulfonamides are divided into various prognostic effects over a period of time. Using a random forest approach, we selected molecular descriptors to achieve better accuracy than the simulation results for compounds designed for the drug (Samigulina and Zarina
2017). The Multilayer perception model is also known as a feed-forward neural network. MLP provides an outcome based on a set of input sources. For training any sort of information, the backpropagation approach is utilized. This model is similar to a directed graph because of the essence of multiple layers as input nodes and output nodes are
 associated with some weights (Pal et al. 1992). After processing the data, the perceptron can fluctuate each connected weight in the network. In this way, the presence of error in last data point i.e., nth ej(n)=aj(n)-Yj(n) Where a-targetvalue, Y-
the variable developed from the perception. Based on some corrections, weights in each node can be adjusted through decreasing error in the output i.e., \Delta W_i(n) = -\eta \delta E(n) \partial V_i(n) V_i(n)
oscillations. Yi \rightarrow previous neuron result. Depending on `Vj´ field, derivative can be calculated. Then, simplified derivative in output node can be-\partial \epsilon(n) \partial V_j(n) = \varphi(n) \partial V_j(n) = \varphi(n)
represented as the last node in the output layer. In case, changing any weights in a hidden layer, the activation function can be varied the weights in the output layer. Figure 5 performs specific computations to distinguish few features will be increased with available
weights to decide specific neuron was terminated or not. In this way, Multilayer perceptron uses backpropagation strategy with the activation function (Rosenblatt 1961). In this review, a multi-layer perceptron was utilized for predicting action between the drugs. This model has one advantage i.e., it does not require any structural information on
compounds because of the fact that it uses experimental data for predicting the accuracy (Stokes et al. 2020). Additionally, MLP was utilized to generate different compounds automatically with some advanced properties (Gómez-Bombarelli et al. 2018). Multilayer Perception
ArchitectureIn general, MLP can be used very easily and very quickly, but fulfilling its duties in training is very difficult, and MLP also does not offer any quarantee of global minimum performance (Gertrudes et al. 2012). The secondary structure of proteins offers a greater advantage in determining protein function, drug design. In that study, the MLF also does not offer any quarantee of global minimum performance (Gertrudes et al. 2012). The secondary structure of proteins offers a greater advantage in determining protein function, drug design. In that study, the MLF also does not offer any quarantee of global minimum performance (Gertrudes et al. 2012).
approach showed greater interest in classification success. However, in the experimental area, determining the secondary structure is more difficult and expensive. Finally, the results from the trained data were reported as a positive success compared to the classification (Yavuz et al. 2018). Deep learning is a part of machine learning, having the
capability to extract a greater level of features through utilization of multiple layers from input data (Deng and Dong 2014). Deep learning is an immense field that is creating massive premiums nowadays. Recently, deep learning is an immense field that is creating massive premiums nowadays. Recently, deep learning is an immense field that is creating massive premiums nowadays.
exactly is deep learning? In general, deep learning is the same neural network architecture that consists of several layers, and data can be transformed between these layers. It's still a significant popular expression, but the innovation behind it is genuine and very refined. So, models in deep learning can be developed through a strategy called greedy
layer-by-layer (Bengio et al. 2007). Figure 6 contrasts the powerful deep learning approaches with pooling layers and figure outs the critical issues and devise the most appropriate solution even problem was in a complex situation. In this review, deep learning algorithms have presented numerous models like DNN, CNN, RNN, Autoencoder in drug
discovery areas. The pooling layer is another structure that hinders the neural networks. The capacity of the pooling layer is to reduce boundary measurement and system computations and work independently on each feature map (channel). The motivation behind why max-pooling layers work so well in
various networks is that it enables the system to recognize the features very effectively after down-testing an input structure and it reduces the over-fitting. Deep Learning Architectures between input and output nodes (Bengio 2009). The
DNN architecture traces the outcomes in a mathematical model expected as a layer, also multiple layers were available in complex DNN, so that network is named as 'deep'. Deep learning models are introduced in QSAR modeling to retrieve feature extractions and
capabilities in chemical characters automatically. Dahl et.al had inspired by Kaggle's results of multi-task DNN. The results of multi-task DNN have demonstrated incredible execution in learning general features of sharing parameters (Dahl et al. 2014). Development of candidate drugs plays major desirable property in
oral delivery. Molecules in intestinal permeability can be assessed by computational technology through affording rapid and reasonable ways. Multiple studies focused on intestinal neural networks have been adopted for predicting the intestinal
permeabilities of peptides. The intestinal permeable of peptides consists of positive controlled data obtained through the peroral phage technique and random sequence data can be prepared through negative controlled data. Multiple statistical indicators like specificity, sensitivity, ROC score, enrichment curves, etc., are validated to produce
appropriate predictions. And the statistical results declared that models have good quality and can segregate in between random sequences and permeable with great levels of confidence. Finally, the ANN models demonstrated greater prediction than unpredictable one. So, this model can applicable for intestinal permeable peptide selection to
generate peptidomimetics (Jung et al. 2007). Multi-task neural networks integrated into a platform called 'DeepChem', it helps the multi-task neural networks have assessed performance in the multi-task deep networks was robust. Finally, the performance of
deep learning algorithms in QSAR models upgraded the prediction performance. Also, DNN played out a significant role in further research of hit-to-hit lead optimization. CNN is a subclass of DNN, ordinarily utilized for analyzing the visual images (Valueva et al. 2020). CNN also called shift-invariant ANN because frequently rely on weights. CNN is
regularized version of a multi-layer perceptron. The concept of multi-layer perceptron characterizes fully connected algorithm, a network can conquer the overfitting problem. The CNN algorithm examines the clinical field so that, every
neuron in a human cell appears like the visual cortex (Venkatesan and Li 2017). In ligand-protein interaction, many researchers utilized CNN model for prediction indicated the best correlation in the dataset (Jiménez et al. 2018). In protein-ligand
interaction, the CNN algorithm predicted binding affinities which can further increase scoring function but predictive capabilities must upgrade simultaneously. RNN algorithm predicted binding affinities which can be created in the network, connections can occur between the input node and the output node. In this way, a directed graph can be created in the network, connections can occur between the input node and the output node. In this way, a directed graph can be created in the network.
along with a temporal sequence. Likewise, the RNN network utilizes the internal memory to perform grouping in input variables (Dupond 2019). It also exhibits dynamic performance Miljanovic (2012) because the RNN algorithm struggled for two networks at a time with the general structure. Each network may contain various impulses i.e., finite
and infinite impulses. Determining the functionality of protein structure will play a vital role in secondary and tertiary structures. Previously, numerous algorithms relates to folding prediction have improved to encode in the protein structures. So, Visibelli has found α-helixes signals on a large dataset. To locate
specific occurrences in amino acids to characterize the specifications in secondary structure for deciding the helical moieties boundaries. The α-helixes occurrences are predicted through various ML models for validating the hypothesis equipped with an attention mechanism. This mechanism can interpret the weights of each input, model's decision
for prediction. At last, the similar subsequences show the experimental outcomes, where input code-driven in secondary structure information (Visibelli et al. 2020). Day by day, it has been turning out to be a challenge in improving affordable and effective treatments to humans without any prescience in drug target information. The deeDTnet is one of
the deep learning techniques that were embedded with 15 variations of phenotypic, chemicals, cellular profiles, genomics utilized to accelerate drug repurposing and target identification. Due to the presence of high accuracy, deepDTnet has been approved by U.S. Food and Drug Administration with the identification of novel targets to familiar drugs
Through experimental results, topotecan was an approved inhibitor that can directly be utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors to diminish transitional void in drug development (Zeng et al. 2020). In virtual screening, RNN utilized for human retinoic-acid receptors are acid retinoic-acid receptors.
(Kadurin et al. 2017). In producing the de novo drug design, the prediction must be conducted on biological performance. In this way, the RNN algorithm was utilized for generating molecules must be trained by the RNN algorithm
through conditional probability. Various classifiers performed data sampling however RNN with reinforcement learning has given 95% accuracy in scoring function (Mnih et al. 2015).'Deep Interact' was an integrative domain-based approach is utilized to predict PPI's through Deep Neural Network. Assortment of multiple PPIs is extended out from
(KUPS) Kansas University Proteomics Service and (DIP) Database of Interacting Proteins. It's highly fundamental to discover and analyze the cellular components in the specificity of interactions and explicit molecular protein complexes. The significant goal is to develop enormous scope high-throughput experiments through silico approach to
approaches in PPI prediction (Patel et al. 2017). Autoencoder is a class of artificial neural network, it retrieves information through unsupervised learning (Kramer 1991). Autoencoder objective is to represent the encoding data format in dimensionality reduction for maintaining a strategic distance from the 'noise' signal in the network. Along with
this, the autoencoder must explore input data and then copied to the output layer. Autoencoder has two areas i.e., Encoder and Decoder; and one hidden layer is considered as code. Encoder transfers input data to the hidden layer. The decoder can retrieve information for reproducing the signal output. Autoencoders was most
appropriate in dimensionality reduction and learning the data from generative models (Kingma and Welling 2013; Larsen et al. 2015). Considering encoder as \psi, such that \phi: Y \to E, \psi: E \to Y\Phi, \psi: 
code, `W' as weight matrix, `b' as bias vector, \sigma acts as activation function. Basically, biases and weights are randomly utilized and updated through backpropagation technique. Then, decoder maps `h' to `y' with same structure of `y':, Y' = \sigma'(W'h + b')Decoder consists \sigma', W', b' coefficients may vary in encoder i.e., \sigma, W, b coefficients. Mainly autoencoders
were trained to decrease reconstruction errors (loss).L(y,y')=|y-y'|/2=|y-\sigma'(W'(\sigma(Wy+b))+b')|/2Here, feature space `Y'. Also \phi(y) is a compressed format for input space `Y'. Also \phi(y) is a compressed format for input space `E' consists of less dimensionality than input space `E' consists of less dimensionality than input space `Y'.
identity function, finally, it was useless. In Autoencoders, test results despite everything to learn numerous valuable features from training set (Kingma and Welling 2019). In drug discovery, autoencoders utilized as unique architecture to deliver molecules through conducting experiments right into vermin (Zhavoronkov et al. 2019). In designing of
de-novo drug design, deep learning model i.e., autoencoder have utilized for generating the molecules. So, the autoencoder approach was employed with various classifiers like multilayer perceptron for generating new compounds automatically with appropriate properties (Gómez-Bombarelli et al. 2018). In many situations, the drug produces invalid
SMILES syntax, so as to defeat this issue, grammar variational autoencoder was utilized for developing SMILES syntax with more effectiveness (Pu et al. 2017) (Fig. 7). The review of drug discovery, hit to lead, lead optimization
techniques are discussed out. The drug design techniques rely on the databases which are inturn developed based on the different ML algorithms. The precise training, validation, and applicated error-prone protocols. The ML techniques are
introduced in most of the drug design processes to reduce the time as well as manual interference. The best example is QSAR, in which the huge data collection and training of datasets are considered as rate-limiting steps in defining the ligand-based virtual screening protocols and are now replaced by Denovo design techniques. The relationship
between drug discovery steps and algorithms is presented in Fig 8. The folding of secondary structure like β-sheets and α-helices, which is formed by the interaction of side-chain amino acid residues are very critical to regulating the smooth functioning of three-dimensional proteins. An accurate protein folding along with its prerogative active ligand
site can be experimentally obtained by X-ray crystallography, NMR-spectroscopy, and Cryogenic electron microscopic techniques (Cryo-EM). Information about the primary amino acid sequences of proteins/enzymes/receptors, both dissolved / insoluble, is stored on the UNIPROT server along with their targets and cellular functions. Based on
medicinal chemistry or pharmacological or biochemical studies, the main role of proteins is identified, and this information is also the basic unit for developing the protein folding predictions in the provided aminoacid (UNIPROT) sequence were compared with its
 experimentally derived PDB homologues which became a hopeful technique to refine the new protein models computationally and is also termed as "homology modeling". The homology modeling or comparative modeling is analyzed by the several algorithms which need to be implemented in either software modules (PRIME) or web servers (EXPASY
SWISS-MODEL) will definitely make a decision to predict the secondary structure folding with high accuracy within provided templates. However, the fine-tuning for the obtained homology models or template-based models are again scrutinized by Ramachandran analysis which can be sorted out by commercial modules (PRIME) or web servers
(QMEAN, PROCHEK). For further understanding, the homology of CHIKV nsP2 protease is described here (Fig. 9) which is obtained based on experimentally predicted VEEV nsP2 protease template by using insilico techniques. The inisilico techniques the computational databases to dig the information about the homology templates and provides the
best closest match as considering for more practical bioinformatics and medicinal chemistry applications. Figure 8 depicted the alignment of secondary structures such as α-helices, β-pleated sheets, and loop representations present in tertiary complexes. The surface view also useful for recognizing the hotspots present on the protein to bind with
incoming ligands/substrates. The sequence alignment mode also shows the mutations or differences in their primary sequences, it can be employed in different chemo-informatics approaches to identify the mutations or differences in their primary sequences, it can be employed in differences in their primary sequences, it can be employed in differences in their primary sequences, it can be employed in differences in their primary sequences, it can be employed in differences in their primary sequences, it can be employed in differences in their primary sequences.
as an emerging tool in the current SARS-COV2 pandemic towards the identification of new drug-like molecules (Fig. 10). In addition, selecting the best homologous model obtained from the above process is another major task that can be performed with SVOMA (Support-vector-machine Protein single-model Quality Assessment) servers or ProO3 or
ERRAT, which are operated by the Deep-learning methods. After going through the above steps, the best 3D protein template can be used for any basic drug chemistry study to identify hits that are part of a structure-based virtual screening protocol. To provide insight for homology modeling, the {"type":"entrez-protein","attrs"
{"text":"Q5XXP4","term id":"82020140","term id":"82020140","term text":"Q5XXP4"}} Q5XXP4 fasta sequence belongs to CHIKV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest VEEV nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by overlapping its closest veel nsP2 protease domain has been employed as a template by
understanding the above-specified concepts. Further, the active site residue position analysis of the finest developed model has been done and is found to have similar to VEEV nsP2 protease residues as shown in Fig. 10. The SWISS-MODEL also provides the information about percentage similarity along with structure alignment, the Fig. 10 shown
the overlap of similar active site residues consists of catalytic site (catalytic diad Cys and His). It also represents the conformational changes present in the new template which also considered as an essential parameter for drug interaction studies an essential parameter for drug interaction studies.
as enzymes, ion channels, proteins, or receptors. The binding site recognition for the NCE's is another key task for computational/bioinformatics experiments where more than one active site has existed in the protein. In the above cases, the predefined most popular web servers (FTMap), as well as specific modules such as "Sitemap" developed with
the help of algorithms, can define the preferential binding site to speed up the drug discovery process. A few other online programs like GHECOM, POCASA, Pocketome, SURFNET, ConCavity, LIGSITE, Q-SiteFinder, Fpocket, and PASS predicts the feasible binding sites located within the provided protein templates. Whereas, the metaPocket 2.0
program utilizes the above platforms to afford the most reliable ligand binding sites present on templates. Further, AI models like FD/DCA can also predict the druggable sites in the provided biological macromolecules. Recently, the DeepDTnet as a new target identifier in drug repurposing has been tested. The DeepDTnet strategy is developed by
amalgamating the multi-disease cellular targets, pathogenic genes (genomics), and drugs (chemical spaces) being utilized for their treatment. Prediction of protein folding Patients who experienced illnesses can be recognized through a structure-based drug design approach. Time and
cost consumption should be required for 3D structure of proteins. Because of the large amount of proteins. Because of the large amount of proteins are used to predict the 3D structure of proteins. Because of the large amount of proteins are used to predict the 3D structure of proteins.
extraction capabilities, deep learning approaches must apply prediction in backbone torsion angle (Li et al. 2017), secondary structure (Spencer et al. 2014), and protein residue contacts (Wang et al. 2017). At long last, the goal was to predict the 3D protein structure.
structure. Prediction of protein-protein structure PPI's are essential for biological processes and infections (Falchi et al. 2014; Scott et al. 2016). PPI can be characterized as 'it performs similar to networks for mathematical representation of physical contacts between binding regions in proteins have specific
biological importance. Also, it obtains the experimental and bioinformatics strategies from PPI's database (Li and Lai 2007; Szklarczyk et al. 2014). In this way, it turns into a new class for drug targets that are different from mainstream pharmaceutical targets
like ion channels, coupled receptors, G-protein, etc (Higueruelo et al. 2017). When contrasted with traditional drug targets, target PPI's reduces harmful impacts because of increment in biological selectivity of regulatory
impacts (Valkov et al. 2011). It is mandatory to learn fundamental ideas of the PPI interface on the protein-protein structure. Because of the less accessibility of PPI's data, it contributes many computational techniques for PPI interface
protection (Zhang et al. 2010). For example, a website name "eFindSite" (Maheshwari and Brylinski 2016) utilized for predicting PPI interfaces which consist of templates, residues, and sequence-related features for improving SVM, NBC techniques. If the chance of two interactive protein structures is vacant then it makes it easy for predicting the
 PPI interface (Vakser 2014) where it mainly relies on complementarity rules of protein docking (Chen et al. 2003) and SymmDock strategies (Schneidman-Duhovny et al. 2005). When two unbound proteins are integrated and converged as one protein, then a difficulty emerges for predicting the conformational change. When an equivalent
 accent sequence needs to be derived, deep learning models are used to predict PPI and better improvement is achieved compared to machine learning models such as SVM (Du et al. 2016) was mandatory. Considering druggable sites as
hotspots because of providing an enormous amount of binding free energy to convince the medical chemists (Cukuroglu et al. 2014). Bai et al. (2016), utilized two techniques i.e., fragment docking analysis for detecting druggable PPI sites. Fragment docking named "iFitDock", utilized for checking druggable hot spots(problem
areas) in the PPI interface. Further improvement for candidate binding locales needs to integrate similar small hot spots. At last, based on the evolutionary conservative level, the scoring function must be located to provide the finest protein binding spots. The PPI interface objective was to improve computational methodologies for locating the
best hot spots and significant structure of small modulator targets in the PPI interface. The Protein-Protein Interactions (PPI) is one of the major biological phenomena through which the basic units of the body (cell) transports the signals, ions, substrates, and energy production components that need to improve the pharmacological responses needed
by the body. In another circumstance, the PPI plays a critical role in the pathogenesis of the disease such as various types of cancer, especially colorectal carcinoma in humans is purely dependent on lifestyle as well as hereditary means. However, the pathogenesis of the colorectal carcinoma is linked with the
formation of malignant Adenomatous Polyposis Coli (APC) and its migration in the entire colorectal portion in the body is majorly occurs due to the interaction of APC protein with TCF4 component peptide are located in the pathogenic carcinoma cells. The example APC-Asef, β-catenin
TCF4 PPI has been illustrated in Fig. 12. In recent years, the PPI-based drug discovery programs are experimentally produced a hopeful pharmacological substance, in terms of cancer pathogenesis, APC-Asef PPI inhibitors are the best example which are delivered the basic peptides as an initial point to switch on the medicinal chemistry oriented drug
design projects. The importance of PPIs in understanding host pathogenic protein interactions is another extreme task that excites most vaccination programs. Battling against SARS-CoV2 that preferentially binds to the human
angiotensin converting enzyme-2 (hACE2) to enter into the alveoli mainstream of lungs and cause severe obstruction in respiratory syndrome. However, the time and cost for experimental prediction of PPI are considered as rate limiting barriers. In this regard, the different databases hosted the web servers (few are publicly available) framed by
targeting PPI which are prevailing as preliminary PPI identification tools to accelerate the medicinal chemistry research. The Hit discovery process is advanced in success which has been taken in drug discovery. In this procedure, small molecules are considered as hits for target binding to identify the best-altered functions. The detection of hit by
diverse algorithms is currently prevailing as a robust technique in the current drug discovery paradigm. An application of multivariate parameters (K-nearest neighbors (K-NN) and support vector machine (SVM)) on high-content screening (HCS) analysis in one such method produced a variety of hits against neurological complications. Drug
repurposing DeepDTnet's training parameters outperform other existing target identification techniques and rely on a minimum quantity of FDA-approved drugs (732 drugs) to produce beneficial therapeutic effects (human retinoic acid receptor gamma t-ROR-yt) of the existing topoisomerase inhibitor Topotecan (TPT). The
deepDTnet strategy also transfers several FDA drugs with different chemical scaffolds against GPCR with new targeted pharmacological actions. (See in Figs. 13, 14, 15). The deepDTnet algorithm is considered to be much more advantageous than NetLapRLS and KBMF2K methods as well as Naive Bayes, SVM, KNN, and Random Forest
algorithms. "Repurpose" refers "reprocess/reused/recycle". Drug Repurposing is characterized as 'locating new indications for drugs (Ashburn and Thor 2004). A significant reason
for utilizing the drug repurposing concept in drug discovery, because it exceptionally supportive to have multiple targets (Susan et al. 2017) in each drug which corresponds to various impacts. In this way, it provides high diversity in drug-disease relationships. Example: Few drugs extend its life expectancy such as "Metformin" which is an approved
medicine to deal with diseases like "type 2 Diabetes". In repurpose, essential elements are "drugs and disease genes are other elements utilized in drug repurposing. In order to show the interactions that have occurred in element (Lotfi
Shahreza et al. 2018), this can be performed through the networks, drug-disease networks, d
disease networks, drug-target networks, drug-target networks (Lotfi Shahreza et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguishable drugs have similar targets/effects (Yamanishi et al. 2018). In general, the network's model principle was, indistinguish and the network of the networ
(heterogeneous) networks. At last, consolidate the drug repurposing with drug target prediction to generate new targets and indications, then utilize the network diffusion algorithm and dimensionality reduction approach (Luo et al. 2017). Virtual
screening It is an AI strategy utilized in the drug discovery process for locating small molecules to distinguish bind structures for a drug target. In drug development, virtual screening also utilized software as well as algorithms to recognize hits from private chemical collections for retrieving unique hits inefficient way. After identification of new hits,
a further step needs to purify compounds with unfavorable scaffolds (framework) (Lavecchia and Di Giovanni 2013). And furthermore incorporates hardly includes few strategies like docking-based, similarity searching (Willett 2006), and machine learning methods (Leelananda and Lindert 2016). Based on the above
techniques, classification has taken two strategies i.e., structure-based and ligand-based virtual screening. When 3D-protein structure was accessible then molecular docking process can be widely utilized (Chen 2015). Many applications related to docking-based virtual screening have built (Talele et al. 2010) effectively without any impacts. May some
obstacles are present in this strategy such as the scoring function. A scoring function cannot estimate binding affinities (bond/relationship) with accuracy because insufficient arrangements and entropy impacts (Huang and Zou 2010) have taken protein flexibility which makes it more complicated (Chen 2015). Finally, many docking models considered
binding affinities and refuses remained like docking score, distance-time (Copeland 2010; Xing et al. 2017). When compared to docking-based virtual screening, the ligand-based virtual screening cannot confide to the 3D-protein structure. Its goal is to design bioactivity domains from molecular features (Lavecchia and Di Giovanni 2013). In this
concept, the aim is to persistently improve yields and to decrease false hit rates (Leelananda and Lindert 2016; Liew et al. 2009). DL strategies have been applied to retrieve great classification capacity, low generalization
error (LeCun et al. 2015; Thomas et al. 2014) and powerful feature extraction ability. Example: In virtual screening, sparse distribution method wastes a lot of time in searching process (Ma et al. 2014) among
the Simplified Molecular Input Line Entry Specifications (SMILES) and natural language relies on long short-term memory network architecture. ML techniques like DNN and gradient boosting trees provided the molecular libraries by RNN. Adversarial autoencoder models the molecular fingerprints to locate potential anti-cancer agents (Kadurin
et al. 2017). Routine techniques used after target identification are high through virtual screening (HTVS) and molecular docking techniques embedded in free energy perturbations, sampling, and scoring algorithms. The knowledge of active site for the protein/receptor where ligand would bind to mimic/antagonize the physiological role which is an energy perturbation are high through virtual screening (HTVS) and molecular docking techniques embedded in free energy perturbations, sampling, and scoring algorithms.
essential task to initiate the HTVS protocol. Similarly, the ligand-based virtual scoring function is a fundamental component in molecular docking for assessing binding affinities towards
target (Huang and Zou 2010). In machine learning, mapping ability features can yield great accomplishment to extract physical, geometric, and chemical features (Khamis et al. (2015)) to retrieve scores. Based on scores, data-driven black box models which are considered to predict interactions in binding affinities and furthermore avoiding few
concepts in docking like physical function are very hard to study (Ain et al. 2015). Random Forest and SVM concepts identified with AI utilization for better performance in the scoring function. For instance, an SVM model can be utilized instead of a linear additive method related to the energy terms concept. Since an SVM can characterize the
relationship between experimental binding affinities and own energy terms i.e., can be extracted from docking program eHiTS. Thus, data gives better execution in scoring power (Kinnings et al. 2017). Numerous researchers initiated in utilizing the CNN model in image processing (LeCun et al. 2015) field
because CNN demonstrated better performance and protein-ligand interactions providing numerous features to CNN for predicting protein-ligand affinities. In the estimation of protein-ligand affinities (Jiménez et al. 2018) which have indicated better
correlation behavior in data sets. And essentially, deep learning represents its genuine intensity to increase abstract features from primitive features, since it's necessary to represent fundamental features for a compound-protein structure like molecule types, particle separation (LeCun et al. 2015) etc. A structure Deep VS, reliant on CNN model, got
familiar with abstract features from fundamental features to provide docking programs like GLIDE SP (Friesner et al. 2004) and ICM (Abagyan et al. 1994). Thus, the point in activity scoring was, choosing few features among protein-ligand interaction for predicting binding affinities with help of the CNN model, so it increases information scoring
function but it upgrades the predictive capabilities. It is also referred to as lead generation in the beginning phases of drug discovery. It locates small molecules referred to as hits from the High Throughput Screen (HTS) through deficient optimization approach
integrated with chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm "design layer"/Random Forest regression applied to create new biologically active chemical synthesis as well as mapping algorithm applied to create new biologically active chemical synthesis as well as mapping algorithm and the properties of the complex of the comple
from the hit analogs with the prediction of bioactivity analogs (Esposito et al. 2004). And primarily utilized in mathematical concepts to study quantitative mapping with physicochemical or structural objects and biological activities. QSAR analysis taken apart in foundation of mathematical models, selection and making the progression of molecular
descriptions, evaluation and interpretation methods, utilization techniques (Myint and Xie 2010). Here, mathematical models are considered issues in QSAR demonstration. When descriptors are chosen, then locating mathematical models is necessary to fit relationships in the structure-activity technique. In the
year 1964, Hansch equation was suggested by Hansch et al. For clarifying the 2D structure-activity relationship, utilize the parameters like physicochemical descriptors and linear regression models for presenting QSAR study as another section (Hansch et al. For clarifying the 2D structure-activity relationship, utilize the parameters like physicochemical descriptors and linear regression models for presenting QSAR study as another section (Hansch et al. For clarifying the 2D structure-activity relationship, utilize the parameters like physicochemical descriptors and linear regression models for presenting QSAR study as another section (Hansch et al. For clarifying the 2D structure-activity relationship, utilize the parameters like physicochemical descriptors and linear regression models for presenting QSAR study as another section (Hansch et al. For clarifying the 2D structure-activity relationship).
bioactivity description and chemical structure relationships have hypothesis concept to contribute substituent in compound activities (Free and Wilson 1964). Contrasted with the Hansch method, the Free-Wilson method can encode the chemical structure since it predicts legitimately from the chemical structure without any physiochemical
parameters. Random Forest and SVM are machine learning procedures, used in mathematical models (A Dobchev et al. 2014; Dudek et
was held in 2012 and a team called George Dahl's won the challenge in ensemble methods like gaussian progress regression, multi-task DNN. Along with this, Dahl et al. proceeded to work on the multi-task DNN concept and shown excellent performance
in single-task neural systems. Due to multi-task strategy, neural networks learn features from different parameters however tasks can be similar (Dahl et al. 2014). Ramsundar et al. (2017) utilized multi-task neural structures in drug development to assess the performance and finally, excellent results appeared in the random forests algorithm. Since
multi-task neural structures consolidated towards platform called Deeepchem. Subramanian utilized canvas descriptors for employing DNN. Prediction in binding affinities needs to reinforce the regression and classification model to gain results in human β-secretase-1 inhibitors (Subramanian et al. 2016). Usage of DNN model gives great results in
validation set i.e., classification capability gives 0.82 accuracy, it exhibits regression ability R2 with 0.74, MAE (Mean Absolute Error) is 0.52. DNN model utilizes the 2D descriptors and indicated better results when compared with force-field-based strategies because of the utilization of partial capability models in deep learning. At last, QSAR models
rely upon deep learning techniques which allots the better results in the future prediction role of hit-to-lead optimization research. De novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction role of hit-to-lead optimization research. The novo drug architecture prediction research ro
using a popular De novo model called the fragment-based approach. If at this point there are impracticalities and complexities in the molecular structure (Schneider et al. 2017), the risk arises in the development of the structure and becomes difficult in the assessment of bioactivity. Deep learning models utilized powerful knowledge and generative
```

capabilities to introduce a new structure with appropriate properties (Mullard 2017). In the De novo drug design process, the deep learning models acts as autoencoder to generate an appropriate format for new chemical entities (NCE's). Therefore, an embedment of autoencoder with multilayer perceptron classifier is also a value-added technique in

the generation of NCE's with predefined physicochemical properties. The syntax of the drug/chemical structure is produced in SMILES format which might be difficult to understand in many circumstances and grammar variational autoencoder (VAE) overcomes this problem to accelerate the process (Fig. 17). Smiles/SLN notation of antiviral compoundDeep reinforcement learning technique extended by Olivecrona et al. for predicting biological activities to develop new molecules by adjusting RNN model (Olivecrona et al. 2017). To obtain SMILES syntax, RNN model to be trained; where molecules can collect from chemBL. In reinforcement learning, agents act through actions in

```
activities under certain conditions. At this point, if the agent gets a positive reward, the actions made by the agent's trend can be renewed (Mnih et al. 2015). To acquire a high reward for activity scoring, then utilize the SVM technique to enhance few approaches relying upon ligands concept in the training set. Generate few molecules against
dopamine receptor 2-type for employing deep reinforcement learning model with RNN model. Along with this, it observed predictions have taken over 95% for structures in the bioactive region through the auto-encoders technique. To
generate new molecules automatically with appropriate properties then, Gomez-Bombarelli et al. (2018) integrated multilayer perceptron (MLP) and variational autoencoder (VAE). In PPI prediction, numerous tackles have taken placed due to (i) spending low expenditure in protein information, (ii) lack of known PPI to learn about the explicit virus, (iii)
 inefficient strategies due to sequence dissimilarity in viral families. The de-novo methodology motivation is to predict innovative PPI virus with its host. De-novo was a sequence-based negative examining framework that learns the diverse viruses in PPI to predict the innovative one, where the shared host proteins can exploit. For assessing
generalization, de-novo has endeavored to test the PPI's with various domains. At last, the De novo approach retrieved 81% accuracy in reducing the noisy negative associations and 86% accuracy in the viral protein prediction that utilized in the training period respectively. De-novo strategy accomplished more comparable in intra-species and single
virus-host prediction cases. In this way, it turns to be difficult to predict the PPI for a contaminated person and optimal accuracy is obtained when carrying out tests for the human-bacteria interactions (Eid et al. 2016). To develop biological and chemical prospects, multi-objective optimization technique and AI has given promising outcomes through
entrusting an automated De-novo compound structure like a human-creative mechanism. In this study, innovative perception pair, which relies on multi-objective technology, is to apply the RNN algorithm to automate unique molecules with a de-novo structure build on common properties found among constant physicochem properties for leading
trade-offs. In this view, multiple chemical libraries related to de-novo structure targeting acetylcholinesterase and neuraminidase. For assessing chemical feasibility, validity, drug-likeness, and diversity content were employed through numerous quality metrics. In the de-novo generative molecules, molecular docking has taken place for the evaluation
of posing and scoring through X-ray cognate ligands with similar molecular counterparts. At last, multi-objective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimization and AI are provided to use easily for customizable design techniques which especially effective optimizable effective effective
decoder, and predictor. Encoder plays a significant role in changing strings called discrete SMILES into latent (inactive) space, where vectors are considered as consider
the molecules. For retrieving a high prediction ratio in constant vectors, then utilize the gradient-based techniques i.e., Bayesian inference and gradient-based approach. By using both approaches, a significant advantage was delivering a high predictive ratio
consequently, where humans can comprehend the chemical structure. It does not correlate to chemical structure when SMILES syntax is invalid. To maintain a strategic distance from such difficulties, make the result source more constrained; Pu et al. 2017). For creating
molecular fingerprints, Kadurin et al. have utilized the AAE model, were later referred as druGAN. While using the AAE technique, it demonstrated excellent performance in the VAE model in areas of generation ability, error in reconstruction area, further extraction ability, (Kadurin et al. 2017). Coley et al. (2018) suggested locating whether the
generated molecule was synthetically accessed or not. Depending upon the reaction database, the neural network was trained because of the availability of excellent approximation capabilities for retrieving synthetic complexities i.e., the score in
product complexity must be greater than reactant (Andras 2017). Coley strived numerous attempts to build scoring function through encoding chemicals response into product pair and reactant pair for clarifying correlation inequalities between product and reactant complexities. To become familiar with any scoring capacity at that point, neural
 networks need to be trained where Coley utilized reactant and product pairs in a scope of 22 million. Along with this, the outcome determined with huge complexities in inverse synthetic planning yet additionally discloses synthetic complexities due to disposing of 22 million. Along with this, the outcome determined with huge complexities in the synthetic planning yet additionally discloses synthetic complexities due to disposing of 22 million.
the non-realistic molecules. The lead optimization is an essential step of the drug discovery process in which the best medicinally active fragment hits are considered leads to extend the medicinal chemistry projects. The main aim of the lead optimization is to eliminate the side effects/notorious effects of the existing active analogues by a minimal
structural modification to yield a better and safer scaffold. One such example is to increase the potency by tailor-made approaches to provide better active analogue. Here, the various
properties of ADME/T like Chemical and physical properties, Absorption, distribution, metabolism and excretion, Toxicity, and the ADME/T multi-task neural networks are discussed in the following sections. Chemical and physical properties have been utilized to reduce significant
 failures. At that point, deep learning models are utilized lead optimization techniques to improve unique methodologies (Lusci et al. (2013). Duvenaud et al. (2015) extracted data from molecular graph directly by adopting the CNN-ANN concept to perform prediction i.e., (MAE is 0.53+0.07) due to relied upon interpretability concept. Coley et al.
inspired Duvenaud's work and begun working for better results in molecular agueous concepts. And furthermore used the tensor-based convolutional techniques need to integrate features like a bond, atom levels. For
predicting molecular aqueous solution, Coley's employed an enormous number of atom level information compared to Duvenaud's model (Coley et al. 2017). Establishing a great correlation between Caco-2 permeability coefficients and oral drug absorption (P app) for predicting the candidate drug (P app) (Artursson and Karlsson 1991; Hubatsch et al. 2017).
2007) in the estimation of pharmacokinetic properties. To fabricate prediction templates with 30 descriptors (Wang et al. 2016) at that point, Wang et al. 2016) at that point, Wang et al. 2016) at that point, Wang et al. 2016 at that point, Wang et al. 2
great expectation capability. It follows QSAR principles from OECD (Organization for Economic Co-operation and Development). So as to persuade reliability and rationality, then follow the sequence of OECD standards. Absorption, distribution, metabolism and excretion Entering medicines or drugs into veins of the human body under some activity
site known as drug absorption. For examining the degree of absorption properties with a prediction of bioavailability molecules (Tian et al. 2011). In the usage of the MLR model, Tian et al. employed 1,014 molecules for bioavailability prediction
through molecular assets and structural fingerprints. By utilizing the genetic function technique, excellent results appeared in predictive performance as RMSE = 0.2355 and correlation coefficient is 0.71 respectively. Conveying drugs or medicine into the human body i.e., intracellular and interstitial fluids along with few drug absorption (Sim 2015)
properties called as drug distribution. Drug distribution at steady state (VDss) is a proportion of dosage from vivo stage into plasma reaction. The steady phase in drug distribution at steady state (VDss) is a proportion of dosage from vivo stage into plasma reaction. The steady phase in drug distribution at steady phase in drug distribution at steady phase in drug distribution process. Thus, VDss must be predicted; Lombardo and Jing have created PLS and Random Forest techniques
along with 1,096 molecules (Lombardo and Jing 2016). Here, board members are not satisfied with prediction results because 50% of molecules are accessible in twofold error. VDss may influence by the presence of obscure factors. To defeat this issue, intently taken as a challenge for VDss value in molecular structural data. If a drug or drug enters
the human body under the conditions applied, the drug itself tries to produce the current toxic metabolism. To ensure the metabolism to make predictions with high accuracy. Many AI strategies adopted a huge
amount of drug metabolism information to predict unique metabolic enzymes like UDP-glucuronosyltransferases (UGT's), cytochrome P450s, etc. Furthermore, neural networks trained in UGT metabolism (Dang et al. 2016). Eliminating dosage from
drugs and also metabolites from the human body referred to as drug excretion. Drug metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites from the human body referred to as drug excretion. Drug metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either with the usage of water (i.e., some drugs can be soluble in water) or it directly eliminated through the absence of metabolites are wiped out from the human body either water (i.e., some drugs can be soluble in water) and the absence of metabolites are wiped out from the absence of metabolites are wiped out from the human body either water (i.e., some drugs can be solved in water) and it is a solub eliminated through the absence of metabolites are wiped out from the abse
PCA technique with an expectation pace of 84% (Lombardo et al. 2014) accuracy. Toxicity and the ADME/T multi-task neural networks In clinical and preclinical damage accomplishment was reduced the adequacy of about 33% of significant molecules in drug localization, optimizing the significant molecules reducing risk hazards by predicting toxicity.
(Guengerich 2010). Prediction can perform through techniques called structural alerts and rule-based expert knowledge for toxicity prediction. Along these, Xu et al. created a prediction model named acute-oral toxicity, for predicting results on
molecular graph encoding CNN (MGE-CNN). Predicted outcomes indicated as better when compared with SVM model (Youjun et al. 2017). Therefore, the MGE-CNN model succussed because of feature extraction, model development, molecular encoding is similar in training for neural networks. The advantage was, the issue can alter through
molecular fingerprints because of accessibility of flexibility in the MGE-CNN model. For acquiring great fragments relates to structural alerts, Xu et al. utilized toxic features for fingerprints which characterizes TOX Alerts (Sushko et al. 2012). If parameters were comparative, then it's necessary to correlate with trained multi-task neural networks and
performance demonstrated better results contrasted with single task neural networks (Mayr et al. 2016) because of sharing parameters and more supportive towards multiple tasks for retrieving similar features. At last, some information is provided to the human body when drug absorption, distribution, metabolism, and excretion has handled and
prediction improved through performing multi-tasking neural networks. Here, single-task and multi was tasks contrasted by Kearnes et al. 2016). The AI and ML algorithms prevailed as the main computational scoring functions for
evaluation when a predicted value was added as a parameter, which is involved in the basic drug discovery paradigm (Stork et al. 2020), it illustrated in 18. The detailed applications of the ML algorithms specified in the e-resource are described in the following sections (Fig. 18).ML in Pan-assay interference screening (PAINS) The precise information
about hits can be obtained from primary or secondary biological screening assays of purchasable/commercially available databases which were the most important parameters before starting medicinal chemistry projects. Thus, elimination of the compounds has been exhibited its presence in different cellular biological assays considered as pan-assay
derived hits could reduce the cost and time of the medicinal chemists. The pan assay information can be accessed from the PAINS database on request. Therefore, the Hit Dexter 2.0 web server has been launched compiled from Pubchem library and screening assays. The Hit Dexter 2.0 web server has been launched compiled from Pubchem library and screening assays.
newly designed compound and thus anyone can easily eliminate the pan-assay interfering compound at the initial stage itself (Stork et al. 2019).ML in drug metabolic site prediction The identification of metabolic site prediction and thus anyone can easily eliminate the pan-assay interfering compound at the initial stage itself (Stork et al. 2019).ML in drug metabolic site prediction The identification of metabolic site for any kind of drug or new chemical entity is very essential before its administration into the human body. The
prediction of drug metabolism can be done by animal models (preclinical studies) which was a rate-limiting step as well as costly and it is mandatory to retrieve therapeutic approval of new chemical entities. The site of metabolism can be predicted by several modules among "ADMET Predictor" of SimulationsPlus tools have gained attention and is
pure works on the models compiled by the artificial intelligence algorithms. The FAME3 is one of the online servers which predicts the region for the given drug/compound which undergoes metabolism validated databases gathering phase-1/phase-2 metabolic parameters associated with several databases validated by comparing with Matthews
correlation coefficient (MCC) (Stork et al. 2020). It is also important to have an overview of the chemical modification of drugs/NCE's which are undergone the metabolism and thus can be used in calculating dosage regimen, dosage frequency, toxicity, and other beneficial side effects. The online services such as GLORY/GLORYx provides the precise
information about the possibilities of new metabolite and their relevant formation data with respect to mitochondrial cytochromeP450 enzyme and conjugations (de Bruyn Kops et al. 2019).ML in skin sensitivity parameter prediction of skin sensitivity is one of the essential criteria for assessing safety parameters of the new
drugs/compounds and it is patient to patient to patient to patient to patient to patient specifications. In this regard, the AI models such as Random Forest based MACCS (RF_MACCS) and support vector machine (SVM) based PaDEL (SVM_PaDEL) algorithms trained with approximately 1400 ligands linked with local lympho node assay (LLNA) information (Stork et al. 2020; Vranic et al.
2019).ML in natural product identification The ML trained with 265000 natural product isolates and synthetic libraries validated by MCC is being used as a basic predictive model NP Scout on the prediction of sources for the query molecules.
might provide information about their natural product sources and could become a part of natural product-based drug discovery, numerous clinicians and specialists confronted challenges towards target validation, computational pathology data, identification of prognostic biomarkers in clinical
preliminaries. By regulating the molecular target activity, drugs can be developed through the utilization of ultimate methodologies in drug discovery for altering the infection state. By inaugurating a program in drug development, target identification requires a therapeutic hypothesis for modulating target regulation in the outcome of the infection
state. When available evidence is identified for that target, it can be considered as target identification, outcomes can be retrieved through clinical preliminaries, yet it's necessary to concentrate on target validation efforts for
successful projects. The diseases incorporate metabolomic, transcriptomic, proteomic profiles that are available in-patient clinical material. With the clinical databases provides the primitive target identification and target validation. For predicting target identification, it requires appropriate
strategies for yielding legitimate statistical models.ML approaches are used in target identification because of the increment of data-driven target identification experiments. In target identification experimental). By
using ML approaches, prediction can be taken placed on known properties of targets, causalities, driven targets. ML techniques can apply from various perspectives in the target identification field. For predicting genes with dysphoria, a decision-tree classifier need to be trained on a protein-protein localization network (Costa et al. 2010). So,
distinguished few key parameters in decision-tree inspection i.e., extracellular path, transcription factors, metabolic paths. John et al. improved a classifier model called SVM with genomic details for classifying proteins towards non-drug and drug spots in ovarian and breast cancer (Jeon et al. 2014). mRNA expression, network topology, proteins
protein interaction, DNA copy numbers are the key segments in classification and recognized 122 cancer targets globally. Targets identified as 462, 266, and 355 related to pancreatic, breast, and ovarian tumors. Peptide inhibitors were validated through the prediction of two targets. Outcomes in the cell culture approach were identified as more
prominent anti-proliferative effects. Although, in pancreatic tumors, usage of inhibitors shown twice greater inhibition on cells. To distinguish transcription factor site with transcriptome information (Ament et al. 2018). By utilizing LASSO and regression
models in mouse striatum, a genome-scale has been created for 718 transcription factors. Transcriptional factor modules are recognized to provide treatment in the early phases of Huntington's disease. In tissue-related anti-aging treatment in the early phases of Huntington's disease. In tissue-related anti-aging treatment in the early phases of Huntington's disease. In tissue-related anti-aging treatment in the early phases of Huntington's disease. In tissue-related anti-aging treatment in the early phases of Huntington's disease.
new muscles. When contrasted with supervised machine learning models, SVM exposed feature selection and linear kernels are generally appropriate for identifying biomarkers. Predicted targets can be developed through ML i.e., blind drugs can furtherly be utilized for therapeutic assumptions. For identifying affiliations like gene-disease, drug-
disease, target-drugs, then apply NLP kernel strategies in Medline concept (Bravo et al. 2015). Many supervised learning techniques rely upon EU-ADR [European Union Adverse Drug Reaction] database for disease genes identification in the Medline concept. NLP technique is used in the extraction of biological entity events (Kim et al. 2017). For
identifying therapeutic treatment through novel targets, ML is the best extension for understanding biological aspects. The splicing signal model is utilized to predict alternate signal (Leung et al. 2014). Binding the integrative splicing signals (Jha et al. 2017) like RNA
sequencing data and CLIP-seq splicing data indicated knock-down results. To identify variations in Alzheimer disease (Vaquero-Garcia et al. 2016), then code models like complex variants and de-novo designs must integrate for prediction. ML can predict cancer-related drug impacts (Iorio et al. 2016). So that, ML investigated how DNA-methylation,
somatic mutation data, genome-wide data impacts the drug feedback. To identify molecular features, then utilize logical models, ANOVA, and machine learning models like random forests for predictive data types in cancer regions. Data utilized from RNAi
screens to locate molecular features from 501 cancer lines, so it predicts 769 genes from cancer cells (Tsherniak et al. 2017). 171 chemicals are necessary to locate in genetic affiliations because targetable vulnerabilities revealed as oncotypes don't influence cancer therapy (McMillan et al. 2018). The models used in predictive data types how therapy
in cancer-intrinsic medicine. Many queries emerge for developers i.e., how specific drugs are developed for the given targets in small molecules for delivering drugs. In this way, a random forest algorithm must train on genomic attributes like physicochemical and
cavities of 1,187 compounds in non-drug adhesive sites against 99 protein collection (Nayal and Honig 2006). Additionally, length and configuration are considered significant features in surface cavities. For predicting drug targets, distinctive physicochemical properties from protein sequences applied SVM's (Li and Lai 2007; Bakheet and Doig 2009).
DL model (Bakheet and Doig 2009). Proteins occupy explicit locations in PPI network to associate exceptionally (Jeon et al. 2015). ML algorithms utilized newly developed targets to predict blind drugs for reducing search space, but drug target requires more endorsements. Predicting the clinical trial success in
drug targets is a complicated goal for target validation and identification. Along ML approaches, omics information utilized 332 drug targets, so it can come up short or accomplishment in the third phase of clinical trials through multivariate compound selection (Rouillard et al. 2018). Gene-expression data is identified as successful prediction across
tissue layers with high variance and less RNA mean expression in clinical trials. In this way, the drug target was confirmed that specific disease expression can influence tissue region (Kumar et al. 2016). For predicting de-novo therapeutic drug targets, (Koscielny et al. 2017) ML classifiers should train from open platform (Ferrero et al. 2017).
Significant indications are key data types such as genetic data, gene expression for predicting therapeutic drug targets. In such cases, ML approaches constrained because of data absence and sparse data are fundamental purposes behind failure in drug development programs. Practically, to initiate any drug in the market, it considers the length of
time period due to more advancement in technology, new models like biologics (antibodies were included) can accessible and small molecular drug design may not same as today. Additional constraints are developed to predict medicine because it can fail or succeed with accessible metadata in public space. Using the ML approach, biomarker
discovery is used to improve clinical trial performance by differentiating drugs and understanding drugs are drugs and understanding drugs and understanding drugs are drugs and understanding drugs and understanding drugs are drugs are drugs and understanding drugs are drugs are drugs and understanding drugs a
stages of clinical trials. Usage of ML algorithms allows predicting translational biomarkers in preclinical data assortment. After data validation, corresponding biomarkers and models must investigate the patient indications and lastly propose the medication. In literature, several papers provided information relates to predictive models and
biomarkers, and last, few were utilized in clinical trials. Various factors like model rebuilding, designing, data accessing, data quality and software, model selection are necessary for a clinical trials. Various factors like model rebuilding, designing, data accessing, data quality and software, model selection are necessary for a clinical trials. Various factors like model rebuilding, designing, data accessing, data quality and software, model selection are necessary for a clinical trials.
(Food and Drug Administration) led (MAQC II) MicroArray Quality Control evaluated ML algorithms for predicting gene expression data (Shi et al. 2010) in the final stage of clinical trials. In this project, 6 microarray data collections were analyzed by 36 independent groups to develop predictive models for classifying in the end stage of clinical sites
For modelling appropriate approaches in a clinical trial, information incorporates data quality, skilled scientists, control processes. Multiple myeloma and gene expression
are continuous variables. By utilizing Cox regression models, it confirmed to predict (Zhan et al. 2006) patient risk factors through gene expression models (Shaughnessy et al. 2008; Mulligan et al. 2008; Mulligan et al. 2007) can be highlighted due to the absence of
predefined classes that can perform prediction in clinical trials. To evaluate regression models, NCI (National Cancer Institute) challenge is to build drug predictive models (Costello et al. 2014). Each group must utilize the best model with key parameters in training data collection (i.e., treating 35 breast tumor cells with 31 drugs) and models ought
to be verified through similar blind testing data collection (i.e., treating 18 breast tumor cells with similar 31 drugs). For generating more predictive techniques, six sorts of data profiles are considered i.e., RNA microarray, reverse protein phase array, SNP (Single Nucleotide Polymorphism) array, DNA methylation status, exome
sequencing for 44 groups are utilized for applying multiple regression models like sparse linear regression models. In differentiating, few groups maintained technical details like feature selection
quality control, data reduction, tuning ML parameters, splitting strategy, and biological data like gene expression data to improve the predictive model. Numerous drugs are convenient in the development of the predictive model when compared to other strategies. Challenge of NCI-DREAM needs to maintain a data collection and outcomes for
evaluating, improving group factor analyses in validation (Bunte et al. 2016), Random forest framework (Rahman et al. 2017) and other approaches (Huang et al. (2017); Hejase and Chan (2018); Hejase 
parameters in drug sensitivity sites across tumors in tissue cells. PD1 (Programmed cell Death 1) inhibitor endorsed by FDA in 2017, at that situation, genetic biomarkers utilized s pembrolizumab as inhibitors for tumors. It was the first endorsement made by FDA that relates to genetic biomarkers other than tumor type (Boyiadzis et al. 2018), which
can highlight the biomarker disclosure. Recently, predictive biomarkers indicated improvement in ML other than different oncology data types. For improving drug responses in patients, ML algorithms ought to apply multi-omics data (Tasaki et al. 2018). And gradient regression tree is utilized for improving polygenic risk scores in predicting clinical
trials (Paré et al. 2017). Tested outcomes in UK Biobank explained the presentation of SNP model is individuals such as cardiac arrests, breast cancers, inflammatory bowel cancers, at that point, genome-wide scored data must develop (Khera et al.
2018).RNA sequencing for single-cell innovation is widely utilized in advanced biomarker discoveries and gene clustering. This technique is utilized to locate lineages of trace development, determining cell states, novel cell varieties. Here, reducing estimations in gene expression from thousand cells into the low-dimensional regions was the
develops 2D structure in the multi-dimensional regions for distinguishing cell data into the encoded feature of latent space, VAE's (Variational autoencoders) utilized for determining subpopulations in hidden tumour (Sabrina et al. 2019). Encoded features
assessed few relationships in gene cell subpopulations. This strategy contains a data pre-processing technique since it relies upon unsupervised learning. RNA sequencing of single-cell data utilized the VASC model for data visualization (Wang and Jin 2018). When testing was conducted on 20 informational sets, results indicated more superior to VASC
model other than SIMLR (Wang et al. 2017) and ZIFA (Pierson and Yau 2015) reduction models. By utilizing ML approaches, feature selection received huge advancements in biomarker discovery. For extracting appropriate structures in clusters (Tan et al. 2016), many specialists have claimed unsupervised deep learning methods. To locate explicit
structures in VAE encoded features, then the VAE technique must compete with TCGA (The Cancer Genome Atlas) data in RNA sequencing (Way and Greene 2017). To upgrade identifications in carcinoma disease, Beck et al. (2011) explained data integration techniques, image analysis with gene expression data to identify the squamous cells in lungs
And CNN model showed better execution in predicting the cardiac failures i.e., (AUC=0.97) from endomyocardial biopsy data other than (AUC=0.73 and 0.75) trained samples (Nirschl et al. 2018). From the above examples, the usage of ML approaches has shown success in biomarker discovery and still, numerous issues need to be rectified. A few
 issues considered as; one classifier must understandable by end-users for clinical adoptions. Another key issue was, every approach needs to validate the multi-institutional, multi-site data sets for determining the generalizability approach needs to validate the multi-institutional, multi-site data sets for determining the generalizability approach. Many community parties tended to key issues and providing a quick advancement like model extraction and
interpretations in biological sites (Finnegan and Song 2017), key optimization and training algorithms (Angermueller et al. 2016), model reproducib0ility (Hutson 2018). The word pathology refers to a realistic field, each pathologist clarified what can see from a glass slide through visual assessment. A lot of information is produced through glass slides
for example, which cell type is arranged in tissue layer and spatial context. In this way, it is generally imperative to examine relationships between immune cells and immune-oncology cancers. In clinical trials, before choosing a patient to test with thousands of compounds, pharmaceutical industries must realize how the particular drug can treat
 patient cells and tissues in the body. Because of rapid advancements in clinical trials, locating biomarkers became more significant for victims i.e., who can ready to react to the therapy. Fast improvement in digital pathology can discover new biomarkers with more reasonable, precise, and high-throughput behavior for reducing time in drug
development, and also victims can access therapy very fast. Prior to applying deep learning models, many algorithms related to image analysis propelled me to collaborate with pathology study is to capture to handcraft graphical features in computers. The objective of digital pathology study is to capture to handcraft graphical features in computer scientists are required to handcraft graphical features in computers. The objective of digital pathology study is to capture to handcraft graphical features in computers.
recognize etymological descriptors largely utilized in hematoxylin and easin (H&E) structures. Here, Nuclear morphometry is an implementation in the digital study for explaining relationships between prognosis (Veltri et al. 2000) and features created by PCs. From the spatial context, Beck et al. (2011) identified tissues in stroma cancer and stroma
survival features in breast cancer. Recently, the Nuclear orientation structure was explained by LU et al. (2017) for clarifying survival features in oral cancers and breast cancers (Cheng et al. 2018). In many conditions, antibodies utilized immunohistochemical stains for targeting image proteins. With the absence of deep learning tools, morphology
can detect tissues in sophisticated data. Investigation of immuno-oncology permits ML approaches for generating high throughput features to explain thousands of cells associated with a spatial context, and impossible tasks given for pathologists. Usage of DL methods shows improvement more precisely for tissue and cell detection in cancer
environments. Many different features are explained spatial context associations for cells and tissues through scale estimations. Understanding heterogeneity concept in breast-cancer population to utilize lymphocytes in biomarkers (Mani et al. 2016). The cell-cell relationship was examined and delivered outcomes through cell locations like CD8+
PD1+ and cell densities for distinguishing carcinoma Merkel cell to respond in pembrolizumab (Giraldo et al. 2017). For leading a trial, utilized the number of tissues for each stain. In this circumstance, ML models and feature selections must be incorporated to
predict the therapeutic response. The CNN model is well applicable for digital pathology works since a single biopsy was utilized to train feasible pixels. So, DL models automatically learn structured features from various classification tasks (Janowczyk and Madabhushi 2016). Here model was, M-CNN (Multi-scale CNN) considered as a supervised
 learning technique for phenotyping images with high-content cells (Godinez et al. 2017), where it restricts a few models with their customized steps. Converting image pixel values to phenotype images, then the M-CNN approach demonstrated more accuracy at classification levels. For creating objectives in image analysis, numerous DL methods
utilized in tubules (Romo-Bucheli et al. 2016), lymphocytes (Saltz et al. 2016), lymphocytes (Saltz et al. 2017; Bychkov et al. 2017; Bychkov et al. 2017; Bychkov et al. 2017; Bychkov et al. 2018; Cruz-Roa et al. 2017; Bychkov et al. 2018; Cruz-Roa et al. 2018; Cr
methodologies. Utilization of deep learning models can stimulate data acquisition (Cohen et al. 2018) of MRI (Magnetic Resonance Imaging) or it diminishes dosage for radiation in CT (Computed Tomography) image process (Chen et al. 2017). The quality of images improved a lot in noise signal ratio, spatial resolution; so, applications like victim
stratification, disease prediction, image qualification have correspondingly improved. The deep learning framework is another study (Coudray et al. 2018) which determines to predict the usage of mutated genes called lung cancers from hematoxylin & eosin (H & E)-stained images. In image analysis, numerous deep learning procedures are required to
perform explicit tasks; So, integration of image analysis and deep learning algorithms can be accommodated for problem-solving. In numerous issues, usage of DL techniques can outperform the results, however, it was not an image analysis tool because of lack of flexibility. Likewise, many scientific experts are accessible for any classification tasks.
However, it consumes a lot of money to generate. To defeat this challenge (Turkki et al. 2016) immunohistochemistry staining would utilize to mitigate this problem. Due to community tasks, it provides more data for pathologists to build annotations for many use-cases. The transparency issue is another challenge to digital pathology. Black-box is a
known methodology in deep learning strategies. In classification tasks, decision-making is unclear. For understanding numerous mechanisms in drug development, interpretable outcomes can be accommodating in locating potential biomarkers and drug targets for predictive response in therapy. Additionally, trust should be improved in generating
assembled features with interpretability. In clinical trials, the large sample size required to apply DL techniques legitimately for predictive response in therapy is a further challenge. The DL requires countless sample examples in clinical trials. Sometimes, integrating data in clinical trials can be possible however the existence of bias can make the
outcomes difficult for interpretation. Corredor et al. (2019) and Saltz et al. (2018) explained numerous models related to image analysis and DL models for predictive response in therapy, at that point CNN model used to identify features in sub-sequent graph and lymphocytes situated in H&E-stained cells. In the future, DL consists of more
capabilities to replace nuclear detection and traditional segmentation algorithms for providing spatial context features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML methods related to various tasks in Drug discoveryMethodElement/features (Table 2). Different ML method 2). Different ML method 2). Different ML method 2).
convolutional fingerprints Duvenaud et al. (2015), Coley et al. (2017)CNNSubsequent graph- based features HE-stained tissuesTo predict the disease response in lymphocytes Saltz et al. (2017)RNNMolecular GraphGenerating molecules with predicted
biological activity Olivecrona et al. (2017)RNNSMILESTo predict molecular properties Pu et al. (2017)RNNSMILESGenerating novel molecules Olivecrona et al. (2017)RNNSMILESTo predict molecular markers for predicting cancer
dependencies Tsherniak et al. (2017)DNNAtom pair descriptors and donor-acceptor pair descriptor and donor-acceptor pair descriptor and donor-acceptor pair descriptor and donor-acceptor pair descriptor and donor-acceptor pair descriptors and protein featuresDrug target interactions Wang et al. (2015)DNNMolecular descriptor and donor-acceptor pair descriptor and donor-acceptor and donor-acceptor pair descriptor and donor-acceptor acceptor and donor-acceptor and donor-acce
expression signaturesIt identifies the predicted high-risk subgroup of victims Zhan et al. (2006)Gradient boosted regression treesGenome-wide polygenic scoresTo identify high-risks of breast cancer, coronary artery type 2 diabetes diseases in patients. Khera et al. (2018)VAEsRNA sequencing datasetDifferentiating hidden cancer subpopulations without a coronary artery type 2 diabetes diseases in patients.
effectively Sabrina et al. (2019) Auto Encoder Fingerprints Virtual screening Kadurin et al. (2017) Abagyan R, Totrov M, Kuznetsov D. Icm-a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem. 1994;15(5):488-506. [Google Scholar] Ain QU, Aleksandrova
A, Roessler FD, Ballester PJ. Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. Wiley Interdisciplinary Reviews. Comput Molec Sci. 2015;5(6):405–424. [PMC free article] [PubMed] [Google Scholar] Alpaydin E. Introduction to machine learning. Cambridge: MIT press; 2020. [Google Scholar]
Scholar] Ament SA, Pearl JR, Cantle JP, Bragg RM, Skene PJ, Coffey SR, Bergey DE, Wheeler VC, MacDonald ME, Baliga NS, et al. Transcriptional regulatory networks underlying gene expression changes in huntington's disease. Mol Syst Biol. 2018;14(3):e7435. [PMC free article] [PubMed] [Google Scholar] An H, Li M, Gao J, Zhang Z, Ma S, Chen Y, 
 Incorporation of biomolecules in metal-organic frameworks for advanced applications. Coord Chem Rev. 2019;384:90–106. [Google Scholar]Andras P. High-dimensional function approximation with neural networks for large volumes of data. IEEE Trans Neural Netw Learn Syst. 2017;29(2):500–508. [PubMed] [Google Scholar]Angermueller C,
Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. Mol Syst Biol. 2016;12(7):878. [PMC free article] [PubMed] [Google Scholar]Artursson P, Karlsson J. Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (caco-2) cells. Biochem Biophys Res Commun
1991;175(3):880-885. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discovery. 2004;3(8):673-683. [PubMed] [Google Scholar]Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing and developing new uses for existing and developing new uses for existing and developing new us
Cheng RR, Jiang H, Onuchic JN. Elucidating the druggable interface of protein- protein interactions using fragment docking and coevolutionary analysis. Proc Natl Acad Sci. 2016;113(50):E8051-E8058. [PMC free article] [PubMed] [Google Scholar]Bakheet TM, Doig AJ. Properties and identification of human protein drug targets. Bioinformatics
2009;25(4):451-457. [PubMed] [Google Scholar]Beck AH, Sangoi AR, Leung S, Marinelli RJ, Nielsen TO, Van De Vijver MJ, West RB, Van De Rijn M, Koller D. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Trans Med. 2011;3(108):108ra113-108ra113. [PubMed] [Google Scholar]Bengio Y.
Learning deep architectures for AI. Norwell: Now Publishers Inc; 2009. [Google Scholar] Bengio Y, Lamblin P, Popovici D, Larochelle H (2007) Greedy layer-wise training of deep networks advances in neural information processing systems Boyiadzis MM, Kirkwood JM, Marshall JL, Pritchard CC, Azad NS, Gulley JL. Significance and implications of fda
approval of pembrolizumab for biomarker-defined disease. J Immunother Cancer. 2018;6(1):1-7. [PMC free article] [PubMed] [Google Scholar]Bravo Å, Piñero J, Queralt-Rosinach N, Rautschka M, Furlong LI. Extraction of relations between genes and diseases from text and large-scale data analysis: implications for translational research. BMC
 Bioinform. 2015;16(1):55. [PMC free article] [PubMed] [Google Scholar]Bukhman YV, Skolnick J. Biomolquest: integrated database-based retrieval of protein structural and functional informatics. 2001;17(5):468–478. [PubMed] [Google Scholar]Bundela S, Sharma A, Bisen PS. Potential compounds for oral cancer treatments.
resveratrol, nimbolide, lovastatin, bortezomib, vorinostat, berberine, pterostilbene, deguelin, andrographolide, and colchicine. PLoS ONE. 2015;10(11):e0141719. [PMC free article] [PubMed] [Google Scholar]Bunte K, Leppäaho E, Saarinen I, Kaski S. Sparse group factor analysis for biclustering of multiple data sources. Bioinformatics.
2016;32(16):2457-2463. [PubMed] [Google Scholar]Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep. 2018;8(1):1-11. [PMC free article] [PubMed] [Google Scholar]Cabreiro F, Au C, Leung K-Y, Vergara-
Irigaray N, Cochemé HM, Noori T, Weinkove D, Schuster E, Greene NDE, Gems D. Metformin retards aging in c. elegans by altering microbial folate and methionine metabolism. Cell. 2013;153(1):228–239. [PMC free article] [PubMed] [Google Scholar]Cano G, Garcia-Rodriguez J, Garcia-Garcia A, Perez-Sanchez H, Benediktsson JA, Thapa A, Barr A.
Automatic selection of molecular descriptors using random forest: Application to drug discovery. Expert Syst Appl. 2017;72:151–159. [Google Scholar]Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery
Drug Discovery Today. 2018;23(6):1241-1250. [PubMed] [Google Scholar]Chen R, Li L, Weng Z. Zdock: an initial-stage protein-docking algorithm. Proteins Struct Funct Bioinf. 2003;52(1):80-87. [PubMed] [Google Scholar]Chen R, Li L, Weng Z. Zdock: an initial-stage protein-docking algorithm. Proteins Struct Funct Bioinf. 2003;52(1):80-87. [PubMed] [Google Scholar]Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, Zhou J, Wang G. Low-dose ct with a residual encoder-decoder convolutional
neural network. IEEE Trans Med Imaging. 2017;36(12):2524-2535. [PMC free article] [PubMed] [Google Scholar]Cheng L, Lewis JS, Dupont WD, Janowczyk A, Madabhushi A. An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease.
specific survival. Mod Pathol. 2017;30(12):1655–1665. [PMC free article] [PubMed] [Google Scholar]Cheng L, Romo-Bucheli D, Wang X, Janowczyk A, Ganesan S, Gilmore H, Rimm D, Madabhushi A. Nuclear shape and orientation features from h&e images predict survival in early-stage estrogen receptor-positive breast cancers. Lab Invest.
2018;98(11):1438-1448. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson Med. 2018;80(3):885-894. [PMC free article] [PubMed] [Google Scholar]Cohen O, Zhu B, Rosen MS. Mr fingerprinting deep reconstruction network (drone) Magn Reson MB. Mr fingerprinting deep reconstruction network (drone) Magn Reson MB. Mr fingerprinting deep reconstruction network (drone) Magn Reson MB. Mr fingerprinting deep reconstruction network (drone) Magn Reson MB. Mr fingerprinting deep reconstruction network (drone) Magn Reson MB. Mr fingerprinting deep reconstruction network (drone) Magn Reson MB. Mr fingerprinting deep reconstruction network (drone) MB. Mr fingerprinting deep reco
graphs for physical property prediction. J Chem Inf Model. 2017;57(8):1757–1772. [PubMed] [Google Scholar]Copeland RA. The dynamics of drug-target interactions: drug-target
residence time and its impact on efficacy and safety. Expert Opin Drug Discov. 2010;5(4):305–310. [PubMed] [Google Scholar]Cornell M, Paton NW, Wu S, Goble CA, Miller CJ, Kirby P, Eilbeck K, Brass A, Hayes A, Oliver SG (2001) Gims-a data warehouse for storage and analysis of genome sequence and functional data. In: Proceedings 2nd Annual
IEEE International Symposium on Bioinformatics and Bioengineering (BIBE 2001). IEEE, pp 15–22Corredor G, Xiangxue Wang Yu, Zhou CL, Pingfu F, Syrigos K, Rimm DL, Yang M, Romero E, Schalper KA, et al. Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell
lung cancer. Clin Cancer Res. 2019;25(5):1526-1534. [PMC free article] [PubMed] [Google Scholar]Costa PR, Acencio ML, Lemke N (2010) A machine learning approach for genome-wide prediction of morbid and druggable human genes based on systems-level data. In: BMC genomics, vol 11. Springer, Berlin, p S9 [PMC free article] [PubMed]Costello
JC, Heiser LM, Georgii E, Gönen M, Menden MP, Wang NJ, Bansal M, Hintsanen P, Khan SA, Mpindi J-P, et al. A community effort to assess and improve drug sensitivity prediction algorithms. Nat Biotechnol. 2014;32(12):1202–1212. [PMC free article] [PubMed] [Google Scholar]Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö
D, Moreira AL, Razavian N, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med. 2018;24(10):1559–1567. [PubMed] [Google Scholar]Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, Shih NNC, Tomaszewski J, González FA, Madabhushi A. Accurate and
reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent. Sci Rep. 2017;7:46450. [PMC free article] [PubMed] [Google Scholar]Cukuroglu E, Engin HB, Gursoy A, Keskin O. Hot spots in protein-protein interfaces: towards drug discovery. Prog Biophys Mol Biol. 2014;116(2-3):165-173
[PubMed] [Google Scholar]Dahl GE, Jaitly N, Salakhutdinov R (2014) Multi-task neural networks for gsar predicts ugt-mediated metabolism. Bioinformatics. 2016;32(20):3183-3189. [PMC free article] [PubMed] [Google Scholar]de Bruyn
KC, Stork C, Šícho M, Kochev N, Svozil D, Jeliazkova N, Kirchmair J. Glory: generator of the structures of likely cytochrome p450 metabolism. Front Chem. 2019;7:402. [PMC free article] [PubMed] [Google Scholar] De Haes W, Frooninckx L, Van Assche R, Smolders A, Depuydt G, Billen J, Braeckman BP, Schoofs
L, Temmerman L. Metformin promotes lifespan through mitohormesis via the peroxiredoxin prdx-2. Proc Natl Acad Sci. 2014;111(24):E2501-E2509. [PMC free article] [PubMed] [Google Scholar] Decaux O, Lodé L, Magrangeas F, Charbonnel C, Gouraud W, Jézéquel P, Attal M, Harousseau J-L, Moreau P, Bataille R, et al. Prediction of survival in
multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients: a study of the intergroupe francophone du myelome. J Clin Oncol. 2008;26(29):4798–4805. [PubMed] [Google Scholar]Deng L, Dong Yu. Deep learning: methods and
applications. Found Trends Sign Process. 2014;7(3-4):197-387. [Google Scholar]Desai B, Dixon K, Farrant E, Feng Q, Gibson KR, van Hoorn WP, Mills J, Morgan T, Parry DM, Ramjee MK, et al. Rapid discovery of a novel series of abl kinase inhibitors by application of an integrated microfluidic synthesis and screening platform. J Med Chem.
2013;56(7):3033-3047. [PubMed] [Google Scholar]DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: new estimates of r&d costs. J Health Econ. 2016;47:20-33. [PubMed] [Google Scholar]Ding J, Condon A, Shah SP. Interpretable dimensionality reduction of single cell transcriptome data with deep generative models
Nat Commun. 2018;9(1):1-13. [PMC free article] [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning methods in drug development. Curr Top Med Chem. 2014;14(16):1913-1922. [PubMed] [Google Scholar]Dobchev DA, G Pillai G, Karelson M. In silico machine learning met
chemical libraries based on artificial intelligence and pair based multi-objective optimization. J Chem Inform Model [PubMed] Du T, Liao L, Wu CH, Sun B. Prediction of residue-residue contact matrix for protein interaction with fisher score features and deep learning. Methods. 2016;110:97–105. [PubMed] [Google Scholar] Duch W,
Swaminathan K, Meller J. Artificial intelligence approaches for rational drug design and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. Curr Pharm Des. 2007;13(14):1497–1508. [PubMed] [Google Scholar]Dudek AZ, Arodz T, Gálvez J. Computational methods in developing and discovery. (In the develo
 quantitative structure-activity relationships (qsar): a review. Comb Chem High Throughput Screen. 2006;9(3):213-228. [PubMed] [Google Scholar]Dupond S. A thorough review on the current advance of neural network structures. Annu Rev Control. 2019;14:200-230. [Google Scholar]Dupond S. A thorough review on the current advance of neural network structures. Annu Rev Control. 2019;14:200-230. [Google Scholar]Dupond S. A thorough review on the current advance of neural network structures.
Aspuru-Guzik A, Adams RP (2015) Convolutional networks on graphs for learning molecular fingerprints. In: Advances in neural information processing systems, pp 2224–2232Eid F-E, ElHefnawi M, Heath LS. Denovo: virus-host sequence-based protein-protein interaction prediction. Bioinformatics. 2016;32(8):1144–1150. [PubMed] [Google
Scholar]Engelbrecht AP. Computational intelligence: an introduction. New Jersy: John Wiley & Sons; 2007. [Google Scholar]Esposito EX, Hopfinger AJ, Madura JD (2004) Methods for applying the quantitative structure-activity relationship paradigm. In: Chemoinformatics. Springer, pp 131–213 [PubMed]Etzold T, Ulyanov A, Argos P. [8] srs:
 information retrieval system for molecular biology data banks. Methods Enzymol. 1996;266:114-128. [PubMed] [Google Scholar]Falchi F, Caporuscio F, Recanatini M. Structure-based design of small-molecule protein interaction modulators: the story so far. Future Med Chem. 2014;6(3):343-357. [PubMed] [Google Scholar]Ferrero E, Dunhar
I, Sanseau P. In silico prediction of novel therapeutic targets using gene-disease association data. J Transl Med. 2017;13(10):e1005836. [PMC free article] [PubMed] [Google Scholar] Finnegan A, Song JS. Maximum entropy methods for extracting the learned features of deep neural networks. PLoS Comput Biol. 2017;13(10):e1005836. [PMC free article]
 [PubMed] [Google Scholar]Free SM, Wilson JW. A mathematical contribution to structure-activity studies. J Med Chem. 1964;7(4):395-399. [PubMed] [Google Scholar]Friedman J, Hastie T, Tibshirani R. The elements of statistical learning, volume 1. 1. New York: Springer series in statistics; 2001. [Google Scholar]Friesner RA, Banks JL, Murphy RB,
Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. J Med Chem. 2004;47(7):1739–1749. [PubMed] [Google Scholar]Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S,
Michalovich D, Al-Lazikani B, et al. Chembl: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012;40(D1):D1100-D1107. [PMC free article] [PubMed] [Google Scholar]Gertrudes JC, Maltarollo VG, Silva RA, Oliveira PR, Honorio KM, Da Silva ABF. Machine learning techniques and drug design. Curr Med Chem.
2012;19(25):4289-4297. [PubMed] [Google Scholar]Ghasemi F, Mehridehnavi A, Fassihi A, Pérez-Sánchez H. Deep neural network in gsar studies using deep belief network. Appl Soft Comput. 2018;62:251-258. [Google Scholar]Giraldo NA, Kaunitz GJ, Cottrell TR, Berry S, Sunshine JC, Nguyen P, Xu H, Orgutsova A, Church CD, Miller NJ et al. (2017)
The differential association of pd-1, pd-11, and cd8+ cells with response to pembrolizumab and presence of merkel cell polyomavirus (mcpyv) in patients with merkel cell carcinoma (mcc)Godinez WJ, Hossain I, Lazic SE, Davies JW, Zhang X. A multi-scale convolutional neural network for phenotyping high-content cellular images. Bioinformatics
2017;33(13):2010-2019. [PubMed] [Google Scholar]Goodfellow I, Bengio Y, Courville A, Bengio
             nent. Drug metabolism and pharmacokinetics, p 1010210090 [PMC free article] [PubMed]Guney E, Menche J, Vidal M, Barábasi A-L. Network-based in silico drug efficacy screening. Nat Commun. 2016;7(1):1–13. [PMC free article] [PubMed] [Google Scholar]Gunther EC, Stone DJ, Gerwien RW, Bento P, Heyes MP. Prediction of clinical drug
efficacy by classification of drug-induced genomic expression profiles in vitro. Proc Natl Acad Sci. 2003;100(16):9608-9613. [PMC free article] [PubMed] [Google Scholar]Guo Y, Lezheng Yu, Wen Z, Li M. Using support vector machine combined with auto covariance to predict protein-protein interactions from protein sequences. Nucleic Acids Res.
2008;36(9):3025-3030. [PMC free article] [PubMed] [Google Scholar]Gupta S, Chaudhary K, Kumar R, Gautam A, Nanda JS, Dhanda SK, Brahmachari SK, Raghava GPS. Prioritization of anticancer drugs against a cancer using genomic features of cancer cells: A step towards personalized medicine. Sci Rep. 2016;6(1):1-11. [PMC free article]
[PubMed] [Google Scholar]Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, Aguilera-Iparraguirre J, Hirzel TD, Adams RP, Aspuru-Guzik A. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci. 2018;4(2):268–276. [PMC free article] [PubMed] [Google
Scholar]Hansch C, Fujita T. Additions and corrections-analysis. a method for the correlation of biological activity and chemical structure. J Am Chem Soc. 1964;86(24):5710. [Google Scholar]Hartenfeller M, Schneider G (2010) De novo drug design. In: Chemoinformatics and computational chemical biology. Springer, Berlin, pp 299–323
[PubMed]Hassan Baig M, Ahmad K, Roy S, Mohammad Ashraf J, Adil M, Haris Siddiqui M, Khan S, Amjad Kamal M, Provazník I, Choi I. Computer aided drug design: success and limitations. Curr Pharm Des. 2016;22(5):572-581. [PubMed] [Google Scholar]Hejase HA, Chan C. Improving drug sensitivity prediction using different types of data. CPT:
Pharmacometrics Syst Pharmacol. 2015;4(2):98–105. [PMC free article] [PubMed] [Google Scholar]Hinton G. Deep learning—a technology with the potential to
transform health care. JAMA. 2018;320(11):1101–1102. [PubMed] [Google Scholar]Ho Tin K (1995) Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition, vol 1. IEEE, pp 278–282Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Wayne X. Applications of support vector machine (svm)
learning in cancer genomics. Cancer Genom Proteom. 2018;15(1):41-51. [PMC free article] [PubMed] [Google Scholar]Huang C, Mezencev R, McDonald JF, Vannberg F. Open source machine-learning algorithms for the prediction of optimal cancer drug therapies. PLoS ONE. 2017;12(10):e0186906e0186906. [PMC free article] [PubMed] [Google Scholar]Huang C, Mezencev R, McDonald JF, Vannberg F. Open source machine-learning algorithms for the prediction of optimal cancer drug therapies.
Scholar]Huang S-Y, Zou X. Inclusion of solvation and entropy in the knowledge-based scoring function for protein-ligand interactions. J Chem Inf Model. 2010;50(2):262–273. [PMC free article] [PubMed] [Google Scholar]Hubatsch I, Ragnarsson EGE, Artursson P. Determination of drug permeability and prediction of drug absorption in caco-2
monolayers. Nat Protoc. 2007;2(9):2111. [PubMed] [Google Scholar] Hutson M (2018) Artificial intelligence faces reproducibility crisis [PubMed] Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves E, Barthorpe S, Lightfoot H, et al. A landscape of pharmacogenomic interactions in cancer. Cell. 2016;166(3):740-
754. [PMC free article] [PubMed] [Google Scholar] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [Pathol Inform 7 [PMC free article] [PubMed] [Google Scholar] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [Pathol Inform 7 [PMC free article] [PubMed] [Google Scholar] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [Google Scholar] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [Google Scholar] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis; a comprehensive tutorial with selected use cases. [PubMed] [anowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis and cases. [PubMed] [anowczyk A, Madabhushi A, M
drug targets using machine learning, inhibitor design and high-throughput screening. Genome Med. 2014;6(7):1-18. [PMC free article] [PubMed] [Google Scholar][ha A, Gazzara MR, Barash Y. Integrative deep models for alternative splicing. Bioinformatics. 2017;33(14):i274-i282. [PMC free article] [PubMed] [Google Scholar][iménez J, Skalic M,
Martinez-Rosell G, De Fabritiis G. K deep: Protein-ligand absolute binding affinity prediction via 3d-convolutional neural networks. J Chem Inf Model. 2018;58(2):287–296. [PubMed] [Google Scholar]Jung E, Kim M, Jung DH, Rhee H, Shin J-M, Choi K, Kang S-K, Kim M-K, Yun C-H, et al. Artificial neural network models for prediction of intestinal
permeability of oligopeptides. BMC Bioinform. 2007;8(1):245. [PMC free article] [PubMed] [Google Scholar]Kadurin A, Aliper A, Kazennov A, Mamoshina P, Vanhaelen Q, Khrabrov K, Zhavoronkov A. The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology. Oncotarget. 2017;8(7):10883.
[PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: a mini-review. Front Physiol. 2015;6:366. [PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: a mini-review. Front Physiol. 2015;6:366. [PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: a mini-review. Front Physiol. 2015;6:366. [PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: a mini-review. Front Physiol. 2015;6:366. [PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: a mini-review. Front Physiol. 2015;6:366. [PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: a mini-review. Front Physiol. 2015;6:366. [PMC free article] [PubMed] [Google Scholar]Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems because the proteins using machine learning and systems because the proteins are proteins using machine learning and systems because the proteins are proteins as a protein because the proteins are proteins as a protein because the protein 
diagnostic tools and decision support analytics for clinical informatics. Artif Intell Prec Health From Conc Appl, p 31Kearnes S, Goldman B, Pande V (2016) Modeling industrial admet data with multitask networks. arXiv preprint arXiv:1606.08793Khamis MA, Gomaa W, Ahmed WF. Machine learning in computational docking. Artif Intell Med.
2015;63(3):135-152. [PubMed] [Google Scholar]Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018;50(9):1219-1224. [PMC free article] [PubMed]
[Google Scholar]Kim ES, Herbst RS, Wistuba II, Lee JJ, Blumenschein GR, Tsao A, Stewart DJ, Hicks ME, Erasmus J, Gupta S, et al. The battle trial: personalizing therapy for lung cancer. Cancer Discov. 2011;1(1):44-53. [PMC free article] [PubMed] [Google Scholar]Kim J, Kim J, Lee H. An analysis of disease-gene relationship from medline abstracts
by digsee. Sci Rep. 2017;7(1):1-13. [PMC free article] [PubMed] [Google Scholar]Kingma DP, Welling M (2013) Auto-encoding variational autoencoders. arXiv preprint arXiv:1312.6114Kingma DP, Welling M (2013) Auto-encoding variational bayes.
machine learning-based method to improve docking scoring functions and its application to drug repurposing. J Chem Inf Model. 2011;51(2):408-419. [PMC free article] [PubMed] [Google Scholar]Korbar B,
Olofson AM, Miraflor AP, Nicka CM, Suriawinata AA, Hassanpour S (2017) Deep learning for classification of colorectal polyps on whole-slide images. J Pathol Inform 8 [PMC free article] [PubMed]Koscielny G, An P, Carvalho-Silva D, Cham JA, Fumis L, Gasparyan R, Hasan S, Karamanis N, Maguire M, Papa E, et al. Open
targets: a platform for therapeutic target identification and validation. Nucleic Acids Res. 2017;45(D1):D985-D994. [PMC free article] [PubMed] [Google Scholar]Kraus VB. Biomarkers as drug development tools:
discovery, validation, qualification and use. Nat Rev Rheumatol. 2018;14(6):354–362. [PubMed] [Google Scholar]Kumar V, Sanseau P, Simola DF, Hurle MR, Agarwal P. Systematic analysis of drug targets confirms expression in disease-relevant tissues. Sci Rep. 2016;6:36205. [PMC free article] [PubMed] [Google Scholar]Larsen ABL, Sønderby SK
(2015) Generating faces with torch. URL . htmlLavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: a critical review. Curr Med Chem. 2013;20(23):2839–2860. [PubMed] [Google Scholar]Leelananda SP, Lindert S.
Computational methods in drug discovery. Beilstein J Org Chem. 2016;12(1):2694-2718. [PMC free article] [PubMed] [Google Scholar]Leung MKK, Xiong HY, Lee LJ, Frey BJ. Deep learning of the tissue-regulated splicing code. Bioinformatics. 2014;30(12):i121-i129. [PMC free article] [PubMed] [Google Scholar]Leung MKK, Xiong HY, Lee LJ, Frey BJ. Deep learning of the tissue-regulated splicing code.
Deep learning methods for protein torsion angle prediction. BMC Bioinform. 2017;18(1):417. [PMC free article] [PubMed] [Google Scholar]Li Q, Lai L. Prediction of potential drug targets based on simple sequence properties. BMC Bioinform. 2007;8(1):353. [PMC free article] [PubMed] [Google Scholar]Li B, Shin H, Gulbekyan G, Pustovalova O,
Nikolsky Y, Hope A, Bessarabova M, Schu M, Kolpakova-Hart E, Merberg D, et al. Development of a drug-response modeling framework to identify cell line derived translational biomarkers that can predict treatment outcome to erlotinib or sorafenib. PLoS ONE. 2015;10(6):e0130700e0130700. [PMC free article] [PubMed] [Google Scholar]Li L, Wang
B, Meroueh SO. Support vector regression scoring of receptor-ligand complexes for rank-ordering and virtual screening of chemical libraries. J Chem Inf Model. 2011;51(9):2132-2138. [PMC free article] [PubMed] [Google Scholar]Liew CY, Ma XH, Liu X, Yap CW. Sym model for virtual screening of lck inhibitors. J Chem Inf Model. 2009;49(4):877-
885. [PubMed] [Google Scholar]Lombardo F, Jing Y. In silico prediction of vol of distribution in humans. extensive data set and the exploration of linear and nonlinear methods coupled with molecular interaction fields descriptors. J Chem Inf Model. 2016;56(10):2042–2052. [PubMed] [Google Scholar]Lombardo F, Obach RS, Varma MV, Stringer R,
Berellini G. Clearance mechanism assignment and total clearance prediction in human based upon in silico models. J Med Chem. 2014;57(10):4397-4405. [PubMed] [Google Scholar]Lotfi Shahreza M, Ghadiri N, Mousavi SR, Varshosaz J, Green JR. A review of network-based approaches to drug repositioning. Brief Bioinform. 2018;19(5):878-892.
[PubMed] [Google Scholar]Luo Y, Zhao X, Zhou J, Yang J, Zhang Y, Kuang W, Peng J, Chen L, Zeng J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. Nat Commun. 2017;8(1):1–13. [PMC free article] [PubMed] [Google Scholar]Luo Y, Zhao X, Zhou J, Yang J, Zhang Y, Kuang W, Peng J, Chen L, Zeng J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. Nat Commun. 2017;8(1):1–13. [PMC free article] [PubMed] [Google Scholar]Luo Y, Zhao X, Zhou J, Yang J, Zhang Y, Kuang W, Peng J, Zhang Y, Kuang W, Peng J, Zhang Y, Kuang W, Peng J, Zhang Y, Zhao X, Zhou J, Yang J, Zhang Y, Zhao X, Zhou J, Yang J, Zhang Y, Zhao X, Zhou J, Zhang Y, Zhao X, Zhou J, Zhao X, Zhou
Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. J Chem Inf Model. 2013;53(7):1563-1575. [PMC free article] [PubMed] [Google Scholar]Ma XH, Jia J, Zhu F, Xue Y, Li ZR, Chen YZ. Comparative analysis of machine learning methods in ligand-based virtual screening of large
compound libraries. Comb Chem High Throughput Screen. 2009;12(4):344-357. [PubMed] [Google Scholar]Madeira F, Park YM, Lee J, Buso N, Gur T,
Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, et al. The embl-ebi search and sequence analysis tools apis in 2019. Nucleic Acids Res. 2019;47(W1):W636–W641. [PMC free article] [PubMed] [Google Scholar]Maheshwari S, Brylinski M. Template-based identification of protein-protein interfaces using efindsiteppi. Methods.
2016;93:64-71. [PubMed] [Google Scholar]Maltarollo VG, Kronenberger T, Espinoza GZ, Oliveira PR, Honorio KM. Advances with support vector machines for novel drug discovery. Expert Opin Drug Discov. 2019;14(1):23-33. [PubMed] [Google Scholar]Mamoshina P, Volosnikova M, Ozerov IV, Putin E, Skibina E, Cortese F, Zhavoronkov A. Machine
learning on human muscle transcriptomic data for biomarker discovery and tissue-specific drug target identification. Front Genet. 2018;9:242. [PMC free article] [PubMed] [Google Scholar]Mani NL, Schalper KA, Hatzis C, Saglam O, Tavassoli F, Butler M, Chagpar AB, Pusztai L, Rimm DL. Quantitative assessment of the spatial heterogeneity of
tumor-infiltrating lymphocytes in breast cancer. Breast Cancer Res. 2016;18(1):78. [PMC free article] [PubMed] [Google Scholar]Martin-Montalvo A, Mercken EM, Mitchell SJ, Palacios HH, Mote PL, Scheibye-Knudsen M, Gomes AP, Ward TM, Minor RK, Blouin M-J, et al. Metformin improves healthspan and lifespan in mice. Nat Commun. 2013;4(1):1-
9. [PMC free article] [PubMed] [Google Scholar] Matlock MK, Hughes TB, Swamidass SJ. Xenosite server: a web-available site of metabolism prediction tool. Bioinformatics. 2015;31(7):1136–1137. [PubMed] [Google Scholar] Matsumoto A, Aoki S, Ohwada H. Comparison of random forest and svm for raw data in drug discovery: prediction of radiation
protection and toxicity case study. Int J Mach Learn Comput. 2016;6(2):145. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter S. Deeptox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. [Google Scholar]Mayr A, Klambauer G, Unterthiner T, Hochreiter Sci. 2016;3:80. [Google Scholar]Mayr A, Hochreiter Sci. 2016;3:80. [Google Scholar]Mayr A, Hochreiter Sci. 2016;3:80. [Google Scholar]Mayr A, Hochreiter Sc
KR, Peyton M, et al. Chemistry-first approach for nomination of personalized treatment in lung cancer. Cell. 2018;173(4):864-878. [PubMed] [Google Scholar]Miljanovic M.
Comparative analysis of recurrent and finite impulse response neural networks in time series prediction. Indian J Comput Sci Eng. 2012;3(1):180-191. [Google Scholar]Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, et al. Human-level control through deep reinforcement
learning. Nature. 2015;518(7540):529-533. [PubMed] [Google Scholar] Morita A, Ariyasu S, Wang B, Asanuma T, Onoda T, Sawa A, Tanaka K, Takahashi I, Togami S, Nenoi M, et al. As-2, a novel inhibitor of p53-dependent apoptosis, prevents apoptotic mitochondrial dysfunction in a transcription-independent manner and protects mice from a lethal
dose of ionizing radiation. Biochem Biophys Res Commun. 2014;450(4):1498-1504. [PubMed] [Google Scholar] Mulligan G, Mitsiades C, Bryant B, Zhan F, Chng WJ, Roels S, Koenig E, Fergus A, Huang Y, Richardson P, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood.
2007;109(8):3177-3188. [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. Int J Mol Sci. 2010;11(10):3846-3866. [PMC free article] [PubMed] [Google Scholar] Myint KZ, Xie X-Q. Recent advances in fragment-based qsar and multi-dimensional qsar methods. [PMC free article] [PubMed] [Google Scholar] 
sites. Proteins Struct Funct Bioinf. 2006;63(4):892-906. [PubMed] [Google Scholar]Ning X, Karypis G. In silico structure-activity-relationship (sar) models from machine learning: a review. Drug Dev Res. 2011;72(2):138-146. [Google Scholar]Nirschl JJ, Janowczyk A, Peyster EG, Frank R, Margulies KB, Feldman MD, Madabhushi A. A deep-learning
classifier identifies patients with clinical heart failure using whole-slide images of h&e tissue. PLoS ONE. 2018;13(4):e0192726. [PMC free article] [PubMed] [Google Scholar] Olivecrona M, Blaschke T, Engkvist O, Chen H. Molecular de-
novo design through deep reinforcement learning. J Cheminform. 2017;9(1):48. [PMC free article] [PubMed] [Google Scholar]Pal SK, Mitra S (1992) Multilayer perceptron, fuzzy sets, classifiaction [PubMed]Paré G, Mao S, Deng WQ. A machine-learning heuristic to improve gene score prediction of polygenic traits. Sci Rep. 2017;7(1):1–11. [PMC free
article] [PubMed] [Google Scholar]Patel S, Tripathi R, Kumari V, Varadwaj P. Deepinteract: deep neural network based protein-protein interaction prediction tool. Curr Bioinform. 2017;12(6):551-557. [Google Scholar]Patel K, Jordan EJ, Park JH, Suresh K, Smith CM, Lemmon AA, Mossé Yaël P, Lemmon MA, Radhakrishnan R (2021) Computational
studies of anaplastic lymphoma kinase mutations reveal common mechanisms of oncogenic activation. Proc Natl Acad Sci 118(10) [PMC free article] [PubMed]Pierson E, Yau C. Zifa: Dimensionality reduction for zero-inflated single-cell gene expression analysis. Genome Biol. 2015;16(1):1–10. [PMC free article] [PubMed]Pierson E, Yau C. Zifa: Dimensionality reduction for zero-inflated single-cell gene expression analysis.
(2017) How the random forest algorithm works in machine learning. Retrieved December, 21Poole D, Mackworth A, Goebel R (1998) Computational intelligence Pu Y, Wang W, Henao R, Chen L, Gan Z, Li C, Carin L (2017) Adversarial symmetric variational autoencoder. In: Advances in neural information processing systems, pp 4330–4339Rahman R,
Matlock K, Ghosh S, Pal R. Heterogeneity aware random forest for drug sensitivity prediction. Sci Rep. 2017;7(1):1–11. [PMC free article] [PubMed] [Google Scholar]Rahman R, Otridge J, Pal R. Integratedmrf: random forest-based framework for integrating prediction from different data types. Bioinformatics. 2017;33(9):1407–1410. [PMC free article]
```

And the control of th	article] es of clinically samaras D, for computer- cticle] hbar Y, Discovery. D-131. [PMC Deep lidated gene common g distribution. le broach to ab ork C, Chen Y, s. s. Svenja P, chini A, Wyder d drug s. MSystems andreas M, fingerprints. cal. Defining a le] Hyvönen M tional neural crix F, Lynch KW, Cell Biochem. fareendran S. ceutical s B, Cao D-S. encoders with Kurgan L. Scholar]Wang rity learning. nt latent space le Scholar]Xia ach for Google ch clonal J. Xenosite: . gle c Hematol. Honig BH. ation of potent

