EDITORIAL

Artificial intelligence in chemistry and drug design

Nathan Brown[1] · Peter Ertl[2] · Richard Lewis[2] · Torsten Luksch[3] · Daniel Reker[4,5] · Nadine Schneider[2]

Introduction

The discovery of molecular structures with desired properties for applications in drug discovery, crop protection, or chemical biology is among the most impactful scientific challenges. However, given the complexity of biological systems and the associated cost for experiments and trials, molecular design is also scientifically very challenging, prone to failure, inherently expensive and time consuming [1, 2]. To improve our odds and the timelines in this process, and to identify good starting points, unbiased incorporation of knowledge through continuous analysis of literature and patents from different scientific fields is required [3]. The number of yearly publications is increasing, and a good collaboration between scientific experts across disciplines is required to fully evaluate the potential of a hypothesis. The theoretical space of chemistry, even when limited by molecular size, is huge [4] and dramatically exceeds what we can assess experimentally and even computationally. How to navigate through it efficiently and select molecules that satisfy the multiple parameters that need to be optimized and that are synthetically accessible [5]? The number of existing data points at the beginning of a project are low. How can we enrich projects in short time frames with informative molecules and data that are subsequently used to drive the design?

With these questions in mind, it comes as no surprise that data mining and statistics have been integrated into molecular discovery and design pipelines to provide computational support in the prioritization of molecular hypotheses [6, 7]. Machine learning algorithms have been part of the routine toolbox of computational and medicinal chemists for decades. The recent increase in applications and coverage of these methodologies has been attributed to advances in computational power, the growing amount of digitized research data, and an increasing theoretical understanding of the algorithms and their shortcomings. However, given the gradual character of these evolutions, it might be counterintuitive to expect a dramatic revolution of molecular design. Nevertheless, extravagant claims have been made for the ability of Artificial Intelligence (AI) to accelerate the design process [8, 9]; how well founded are these claims? While there is unquestionably a lot of potential in novel computational tools, it is important to scrutinize them and compare their performance to already existing methods, to objectively distinguish real progress from promotion. Only such careful evaluations will enable us to shed light on whether novel artificial intelligence methods contribute to an evolution or a revolution of the established scientific discipline of computer-assisted molecular design [10].

The historical context of machine learning in molecular design

Machine learning and AI are not new to researchers in computer-assisted molecular design. The pioneering work of Hansch and Fujita [6], as well as Free and Wilson [7], established the field of quantitative structure–activity relationship (QSAR) modelling. In their groundbreaking work, they used focused datasets as small as a series of a dozen chemical derivatives to fit equations that would anticipate fairly

✉ Richard Lewis
Richard.lewis@novartis.com

✉ Torsten Luksch
torsten.luksch@syngenta.com

✉ Daniel Reker
reker@mit.edu

[1] BenevolentAI, 4-8 Maple Street, London W1T 5HD, UK

[2] Novartis Institutes for BioMedical Research, 4056 Basel, Switzerland

[3] Syngenta Crop Protection AG, 4332 Stein, Switzerland

[4] Koch Institute for Integrative Cancer Research and MIT-IBM Watson AI Lab, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

[5] Division of Gastroenterology, Hepatology and Endoscopy, Department of Medicine, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, 02115, USA

complex phenotypic effects such as toxicity [11]. Spurred by this success, a large research area has emerged that focuses specifically on (a) identifying approaches to describe chemical structures in more detail, to capture the characteristics that govern their properties such as pharmacophores and three dimensional structure but also autonomously learned representations [12, 13], and (b) derive increasingly complex mathematical relationships that aim at describing the causal relationship between these chemical characteristics and the biological properties of interest for predictive purposes [14, 15]. Through an increasing amount of structural information [16], as well as data generation through combinatorial libraries and high-throughput screening, first applications of more complex machine learning models became feasible. However, the excitement and promise was shortly after followed with disenchantment. The growing field of QSAR learnt hard lessons in the 1990s about model validation, control experiments and other pitfalls [17]. Specifically, the overly broad application of computational models as hard filters for data sets that had not been covered in the training data led to an increasing disappointment in this technology.

With increasing understanding of the algorithmic principles and their statistical interpretation, the concept of domains of applicability was introduced [18–20]. Such predictive confidence estimates enabled computational drug hunters to increase the transparency of the capabilities of their tools as well as adjust expectations. This led to an increasing number of successful applications of machine learning to drug discovery and design across academia and industry in the 2000s, which slowly rebuilt the trust of the community and led to a sustained growth of their use. By 2015, computational advances such as the broad inclusion of GPUs in modern computing frameworks and the increasing amount of available RAM, the training of larger and deeper neural nets became feasible. At the famous Kaggle challenge, a team from Toronto used a Deep Neural Net [21] to win a SAR challenge set by Merck. This competition is commonly perceived as a turning point in which a complex deep learning AI method had outperformed other machine learning approaches and therefore arrived as a useful tool for computational molecular design. Deep Learning can trace its roots back to the 1960s, in its theoretical form at least, with the work of Ivakhnenko and Lapa [22]. AI can trace its roots even further back to a workshop that was run at Dartmouth College in 1956. Even given AI's long history, and typically longer than many imagine, the field has had a number of 'winters' with expectations not matching reality. This has led to a number of setbacks for the field and it has taken time to recover from these. While now multiple promising applications of AI exist to derive molecular descriptors and understand their relationship to biological properties, these methods are inherently linked to big data. These algorithms are typically very data hungry before they can provide useful

solutions; as a bonus, they provide unprecedented opportunities to navigate large datasets.

Big data and navigation in chemical space

Analysis of very big chemical datasets is a major research area that can profit from the application of modern machine learning and AI-based methods. For many years the only larger public chemical data set available was the "NCI Open Database" [23], released in 1999 containing about 250,000 molecules. This database was used as a test case for validation of numerous "classical" cheminformatics methods and virtual screening techniques. Advent of PubChem [24] and later ChEMBL [25] databases considerably increased the amount of publicly available chemical data for model training and validation. PubChem currently contains more than 100 million unique compounds. ChEMBL, in its current 26th release, holds information on nearly 2 million compounds, 13 thousand targets, and 16 million relationships between these compounds and targets. Another useful source of public chemical data is the ZINC database [26] providing information about more than 230 million commercially available compounds. All these three data sources offer user friendly web interfaces, but since the data may be downloaded and processed locally, they also were used for development of several novel analysis and visualization tools [27, 28]. Recently, two new experimental developments have increased the amount of available data by several orders of magnitude. One of these technologies is DNA-based library synthesis [29], where a single library can contain tens or even hundreds of millions of molecules. Introduction of so called "readily available" virtual libraries offered currently by several compound vendors became another important factor in increasing the resolution of possible molecular solutions: the virtual molecules in these libraries are enumerated using exclusively validated synthetic protocols and available building blocks, thereby enabling the vendor to guarantee delivery of picked molecules in a relatively short time. The number of molecules in these libraries is reaching billions [30]. With these developments in mind, the community is expecting further increases in available chemical matter, so that in the next decades we are likely to witness datasets with several billion compound structures. This is an exponential growth, comparable with the Moore's law describing the increase in computer processing power, that will push the number of synthetically accessible molecules towards the size of the virtual chemistry database GDB-17 with 166 billion structures [4] and thereby enable the fine-tuned selection of molecular prototypes—if the amount of data can be appropriately handled.

Classical cheminformatics methods are often struggling with such very big data sets, although some recent

developments are promising [30–32]. Novel machine learning and AI-based approaches can help by adaptively navigating vast chemical spaces and autonomously focusing on the most promising regions. In this special issue, several such approaches are described: in the study by Varnek and colleagues, [33] Generative Topographic Mapping, a sophisticated dimensionality reduction method, was used to compare molecules in the company archive of a large pharmaceutical company with over 8 million commercially available samples. The method was enhanced by an AutoZoom function that focuses on the heavily populated areas of chemical space and automatically extracts substructures well representing these dense regions. The methodology was used to identify sets of commercial molecules maximally enhancing the chemical space covered by molecules already available in the investigated company archive. Such approaches enable the adaptive enrichment of compound sets.

Following an orthogonal approach, Tetko and colleagues [34] describe a focused library generator that is able to generate molecules with a higher chance to exhibit desired properties. The generator is based on the long short-term memory (LSTM) recurrent deep neural network with results directed by the reinforcement learning process to a specific target. As a proof of concept, Mdmx inhibitors were chosen as the objective for the presented study. The generated molecules were further refined by pharmacophore screening and molecular dynamics simulations. Additionally (and something that fortunately has become more commonplace in computational molecular design research), the source code of the generator is available at GitHub, which will allow other researchers to adapt it and use it in their own projects. Taken together, such adaptive approaches will improve the ability of research teams to navigate billions of possible structures to find molecular solutions that are sufficiently optimal for practical applications if the predictive algorithms are powerful enough and sufficiently validated.

Practical considerations for AI-based molecular design

The field of machine learning and AI has moved from theoretical studies to real-world applications. The field of chemoinformatics and especially QSAR have always been early adopters of statistical methods and machine learning, but in the past few years the development of novel algorithms in this area has drastically increased. Besides more conventional models like Random Forest, Gradient Boosted Trees, or Gaussian Processes, which have been applied very successfully in the past [35], novel techniques like deep neural nets (DNNs), convolutional neural nets (CNNs) or recurrent neural nets (RNNs) have been increasingly recognized as valuable additions to the toolbox of chemoinformaticians

[14, 15, 21, 36–38]. CNNs are especially attractive in this regard as they offer a different, data-driven way to extract molecular features [39, 40]. The promise of these novel techniques originates not only from slightly higher performance metrics in retrospective evaluations but even more importantly in an inherent ability to process unstructured data as well as navigating and manipulating the "latent" space. This has led to a series of specialized AI tools that can perform tasks that are not possible with "traditional" machine learning algorithms (see for example References [9, 41, 42]). Another series of publications has shown the ability of deep neural nets to use matrices of experimental observations (multitasking) rather than vectors to improve predictive accuracy [43, 44]—this is especially useful for noisy and smaller data sets, for which data collection experiments are time-consuming and expensive, for example in ADMET predictions [45–49]. Directly tackling this challenge is also possible with one shot learning [50] which enables learning from a low amount of data that is potentially better curated compared to high-throughput data. Conversely, to further combat low data limits and autonomously enable data generation, a new direction is the automation of experiments and "closing the loop" in the design-make-test-analysis (DMTA) cycle typically used in drug discovery programs [51]. Active learning [52] is being applied with increasing popularity to the analysis part of the DMTA cycle. This technique assists in selecting the most "interesting" compounds (most commonly the compounds that will help to improve the model) to test in the next cycle. The new results are then fed back into the system to improve model prediction quality and to rapidly increase the applicability domain of the model [53]. The design part of the DMTA cycle has received more attention, with generative chemistry methods well to the fore. Multiple new de novo design models based on RNNs [54–56], variational autoencoder (VAE) architectures [57–59] or generative adversarial networks (GAN) [60, 61] have been developed recently (see also Ref [62].). Most of these models are trained on molecule structures from large public compound collections like ChEMBL [25] or PubChem [24] (to ensure "druglikeness") and are able to generate completely novel molecules according to an objective function, for example, similarity to a given input structure or fitting to constraints in certain properties like logP or activity against a protein target. For the "make part" of the DMTA cycle retro-synthesis, reaction condition or reactivity prediction has been in the focus of the new DNN-based models [41, 63–66]. Here, substantial progress has been made in all areas given both access to more experimental data [67, 68] but also to the sophisticated techniques like Monte Carlo Tree Search (MCTS) which helps to identify the most likely synthetic routes in retro-synthesis planning using deep neural networks and symbolic AI [41]. In this special issue, Ghiandoni and colleagues present a novel reaction-based de novo

design algorithm [69] adapting previously published work on reaction vectors [70, 71] to optimise molecular structures that are likely to be more synthetically tractable. Using a recommender system, the authors demonstrate that their new methodology successfully prioritises the most relevant reaction vectors; this reduces the possibility of combinatorial explosion in the number of solutions while simultaneously ensuring that the probability of successful synthesis is high.

QSAR modelling has also concentrated on interpretability to assist the design part of DMTA; this assumes that the design is being carried out or supervised by skilled human experts. AI models are rather complex, in terms of their representations of molecules. For that reason they are often treated as black boxes and interpretation or understanding of what exactly is learned remains difficult [72]. The paper in this special issue from Webel et al. demonstrates the impact of deep learning to the area of identifying cytotoxic substructures in a large corpus of data [73]. Here, the authors use Deep Taylor Decomposition to identify these toxicophores in the training set so that one can more easily diagnose the structural drivers of toxicity. Such interpretability will enable to increase the credence into novel methodological developments and facilitate the implementation of such methods into established molecular design pipelines.

In an industrial setting, an important aspect is making all these novel machine-learning models and technologies operational: this includes deployment, access, reproducibility, monitoring and maintenance. In addition, these new machine-learning systems bring novel technical challenges in industrial settings which often are not directly obvious [74]. Green and colleagues [75] discuss how these novel methods can be made accessible to a broad range of scientists in GSK and how a smart design of the system can help with maintenance and deployment. Their system called BRADSHAW integrates methods for chemical structure generation, experimental design, active learning and cheminformatics tools to allow automated molecular design in the DMTA cycle. Due to a very modular design of their system they can incorporate many of these novel methods and models. In a retrospective case study they show how the system can be used successfully in lead optimization for the design of MMP12 inhibitors.

Control experiments—is AI really doing better?

In recent years there has been a resurgence of interest and demonstrated impact of Artificial Intelligence in a number of domains [9, 76, 77]. The biggest impact in recent years has been the advent of publicly available Deep Learning algorithms for processing image data and pattern recognition through the ImageNet [78] competition, leading to a

victory for Deep Learning in 2012. The recent advances, especially in Deep Learning, have led to a huge quantity of research conducted in this area and published online in preprints and peer-reviewed articles. Of particular interest here, is the great quantity of research directly at challenges in chemistry and, specifically, drug discovery and materials chemistry. Given the increasing importance of these new machine-learning methods in a plethora of fields, researchers are trying to better understand how these models work [79, 80]. As might be expected, these models have a high risk to learn something different than what was intended [81, 82]. Much work has still to be done to make these methods resilient to noise (brittleness) or overfitting [83]. Latter, i.e. memorization of training data by these models, can lead to a reduced performance on prospective data in the best case but also to security issues in the worst case [84, 85]. Due to these reasons, the establishment of a strong tool kit for validation of these models is crucial (see for example [86–88]). In this special issue, Lee and coworkers [89] have investigated a recent study on large scale comparison of deep learning models with more traditional methods on bioactivity prediction tasks [43]. They show how critical it is to choose the right metrics for benchmarking regarding data distribution and data biases to enable a fair comparison of the methods. Furthermore they suggest using precision and recall statistics in conjunction with the common area under the receiver-operator curve (AUC–ROC). Finally they report challenges in interpreting scaffold-splitting cross-validation results. They conclude that more research needs to be done in proper validation procedures for these models used in the field of chemoinformatics.

Conclusions

As is evident from the information covered in this perspective and by the plethora of scientific and media outlets, many opportunities exist now for the development of novel computational methods, data-driven workflows and algorithmic tools that lead to a higher degree of automation and improve the efficacy of certain components in the drug design process [37]. A particular focus lies on assisting the selection of which experiment to carry out next [52]. The tight integration of artificial intelligence into pharmaceutical, chemical, and crop protection research is inevitable and has the potential to significantly improve the efficiency and efficacy in molecular discovery.

Although slight increases in retrospective accuracy are unlikely to qualitatively change the ability of machine learning to support the drug discovery and development pipeline [10], we anticipate an enthusiasm for this technology, coupled to technological and algorithmic advances, to significantly further the field and increase the contribution of

computational tools in the chemical sciences. A possible inflection point for the field will be the concurrent progress initiated by the convergence of multiple AI branches, such as natural language processing, computer vision, and robotics. This might very well amplify the increase in available information, change our ability to automate and increase reproducibility of experiments, as well as accelerate our understanding of the inner-workings AI. We are still a very long way from a completely in silico discovery process; the need to perform experiments is still vital.

With these advantages in mind, novel challenges will occur. First and foremost, similar to the emergence of applicability domains, a consensus among the community needs to be reached about what appropriate controls are to validate and assess novel AI tools [90]. Specifically relevant will be the proper implementation of adversarial controls to reduce the risk of overfitting, brittleness, and other classical machine learning challenges [84, 91], which are easily overlooked with increasing model complexity. Another important challenge that arises with increasingly complex models will be the potential for attacks or simply unrobust predictive behavior [85, 92]. This is a recurrent hot topic in deep learning research and its implications for novel computational tools in molecular design will need to be carefully considered.

In this special issue, we have carefully picked a selection of classical challenges in computer-assisted molecular design and have invited some of the leading scientists in their respective disciplines to contribute studies that propose avant-garde computational approaches to address these challenges and evaluate and contextualize their potential to accelerate drug discovery. We expect that this special issue will provide an overview of the possibilities that these novel tools hold, but also provide important examples on proper quality control, validation, and domain of applicability assessments. We hope that this will serve as a compendium to stir further discussions and guide the future development of novel AI-tools to guide molecular design.

References

1. Mullard A (2014) New drugs cost US$2.6 billion to develop. Nat Rev Drug Discov 13:877–877
2. Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 3:711–715
3. Searls DB (2005) Data integration: challenges for drug discovery. Nat Rev Drug Discov 4:45–58
4. Ruddigkeit L, van Deursen R, Blum LC, Reymond J-L (2012) Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. J Chem Inf Model 52:2864–2875
5. Lipinski C, Hopkins A (2004) Navigating chemical space for biology and medicine. Nature 432:855–861
6. Hansch C, Fujita T (1964) p-σ-π Analysis. A method for the correlation of biological activity and chemical structure. J Am Chem Soc 86:1616–1626
7. Free SM Jr, Wilson JW (1964) A mathematical contribution to structure-activity studies. J Med Chem 7:395–399
8. Zhavoronkov A, Ivanenkov YA, Aliper A et al (2019) Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol 37:1038–1040
9. Stokes JM, Yang K, Swanson K et al (2020) A deep learning approach to antibiotic discovery. Cell 180:688–702.e13
10. Morrison C (2019) AI developers tout revolution, drugmakers talk evolution. Nat Biotechnol. https://doi.org/10.1038/d41587-019-00033-4
11. Holzgrabe U (1994) QSAR: Hansch analysis and related approaches, H. Kubiny, VCH, Weinheim 1993. 232 Seiten, 60 Abb. und 32 Tab. 158,– DM. ISBN 3-527-30035-X. Pharm Unserer Zeit 23:192–193
12. Todeschini R, Consonni V (2000) Methods and principles in medicinal chemistry. Handbook of molecular descriptors. Wiley-VCH, Weinheim
13. Yang K, Swanson K, Jin W et al (2019) Are learned molecular representations ready for prime time?. Massachusetts Institute of Technology, Cambridge
14. Vamathevan J, Clark D, Czodrowski P et al (2019) Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18:463–477
15. Chen H, Engkvist O, Wang Y et al (2018) The rise of deep learning in drug discovery. Drug Discov Today 23:1241–1250
16. Lewis RA (2005) A general method for exploiting QSAR models in lead optimization. J Med Chem 48:1638–1648
17. Dearden JC, Cronin MTD, Kaiser KLE (2009) How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). SAR QSAR Environ Res 20:241–266
18. Varnek A, Baskin I (2012) Machine learning methods for property prediction in chemoinformatics: Quo Vadis? J Chem Inf Model 52:1413–1437
19. Fechner N, Jahn A, Hinselmann G, Zell A (2010) Estimation of the applicability domain of kernel-based machine learning models for virtual screening. J Cheminform 2:2
20. Sheridan RP, Feuston BP, Maiorov VN, Kearsley SK (2004) Similarity to molecules in the training set is a good discriminator for prediction accuracy in QSAR. J Chem Inf Comput Sci 44:1912–1928
21. Ma J, Sheridan RP, Liaw A et al (2015) Deep neural nets as a method for quantitative structure-activity relationships. J Chem Inf Model 55:263–274
22. Ivakhnenko AG, Lapa VG (1967) Cybernetics and forecasting techniques. American Elsevier Pub. Co., New York
23. Voigt JH, Bienfait B, Wang S, Nicklaus MC (2001) Comparison of the NCI open database with seven large chemical structural databases. J Chem Inf Comput Sci 41:702–712
24. Kim S, Chen J, Cheng T et al (2019) PubChem 2019 update: improved access to chemical data. Nucleic Acids Res 47:D1102–D1109
25. Mendez D, Gaulton A, Bento AP et al (2019) ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res 47:D930–D940
26. Sterling T, Irwin JJ (2015) ZINC 15—ligand discovery for everyone. J Chem Inf Model 55:2324–2337
27. Reymond J-L (2015) The chemical space project. Acc Chem Res 48:722–730

28. Borrel A, Kleinstreuer NC, Fourches D (2018) Exploring drug space with ChemMaps.com. Bioinformatics 34:3773–3775

29. Goodnow RA, Dumelin CE, Keefe AD (2017) DNA-encoded chemistry: enabling the deeper sampling of chemical space. Nat Rev Drug Discov 16:131–147

30. Hoffmann T, Gastreich M (2019) The next level in chemical space navigation: going far beyond enumerable compound libraries. Drug Discov Today 24:1148–1156

31. NextMove Software|SmallWorld. Available at https://www.nextmovesoftware.com/smallworld.html. Accessed 24 May 2019

32. Walters WP (2019) Virtual chemical libraries. J Med Chem 62:1116–1124

33. Lin A, Beck B, Horvath D et al (2019) Diversifying chemical libraries with generative topographic mapping. J Comput Aided Mol Des. https://doi.org/10.1007/s10822-019-00215-x

34. Xia Z, Karpov P, Popowicz G, Tetko IV (2019) Focused library generator: case of Mdmx inhibitors. J Comput Aided Mol Des. https://doi.org/10.1007/s10822-019-00242-8

35. Sheridan RP, Wang WM, Liaw A et al (2016) Extreme gradient boosting as a method for quantitative structure-activity relationships. J Chem Inf Model 56:2353–2360

36. Sanchez-Lengeling B, Aspuru-Guzik A (2018) Inverse molecular design using machine learning: generative models for matter engineering. Science 361:360–365

37. Schneider P, Walters WP, Plowright AT et al (2019) Rethinking drug design in the artificial intelligence era. Nat Rev Drug Discov. https://doi.org/10.1038/s41573-019-0050-3

38. de Almeida AF, de Almeida AF, Moreira R, Rodrigues T (2019) Synthetic organic chemistry driven by artificial intelligence. Nat Rev Chem 3:589–604

39. Kearnes S, McCloskey K, Berndl M et al (2016) Molecular graph convolutions: moving beyond fingerprints. J Comput Aided Mol Des 30:595–608

40. Yang K, Swanson K, Jin W et al (2019) Analyzing Learned Molecular Representations for Property Prediction. J Chem Inf Model 59:3370–3388

41. Segler MHS, Preuss M, Waller MP (2018) Planning chemical syntheses with deep neural networks and symbolic AI. Nature 555:604–610

42. Méndez-Lucio O, Baillif B, Clevert D-A et al (2020) De novo generation of hit-like molecules from gene expression signatures using artificial intelligence. Nat Commun 11:10

43. Mayr A, Klambauer G, Unterthiner T et al (2018) Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. Chem Sci 9:5441–5451

44. Whitehead TM, Irwin BWJ, Hunt P et al (2019) Imputation of assay bioactivity data using deep learning. J Chem Inf Model 59:1197–1204

45. Montanari F, Kuhnke L, Ter Laak A, Clevert D-A (2020) Modeling physico-chemical ADMET endpoints with multitask graph convolutional networks. Molecules 25:44

46. Ramsundar B, Liu B, Wu Z et al (2017) Is multitask deep learning practical for pharma? J Chem Inf Model 57:2068–2076

47. Wenzel J, Matter H, Schmidt F (2019) Predictive multitask deep neural network models for ADME-Tox properties: learning from large data sets. J Chem Inf Model 59:1253–1268

48. Xu Y, Ma J, Liaw A et al (2017) Demystifying multitask deep neural networks for quantitative structure-activity relationships. J Chem Inf Model 57:2490–2504

49. Zhou Y, Cahya S, Combs SA et al (2019) Exploring tunable hyperparameters for deep neural networks with industrial ADME data sets. J Chem Inf Model 59:1005–1016

50. Altae-Tran H, Ramsundar B, Pappu AS, Pande V (2017) Low data drug discovery with one-shot learning. ACS Cent Sci 3:283–293

51. Schneider G (2018) Automating drug discovery. Nat Rev Drug Discov 17:97–113

52. Reker D, Schneider G (2015) Active-learning strategies in computer-assisted drug discovery. Drug Discov Today 20:458–465

53. Reker D, Schneider P, Schneider G (2016) Multi-objective active machine learning rapidly improves structure-activity models and reveals new protein-protein interaction inhibitors. Chem Sci 7:3919–3927

54. Segler MHS, Kogej T, Tyrchan C, Waller MP (2018) Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Cent Sci 4:120–131

55. Olivecrona M, Blaschke T, Engkvist O, Chen H (2017) Molecular de-novo design through deep reinforcement learning. J Cheminform 9:48

56. Ertl P, Lewis R, Martin E, Polyakov V (2017) In silico generation of novel, drug-like chemical matter using the LSTM neural network. arXiv preprint arXiv:171207449

57. Winter R, Montanari F, Noé F, Clevert D-A (2019) Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations. Chem Sci 10:1692–1701

58. Gómez-Bombarelli R, Wei JN, Duvenaud D et al (2018) automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci 4:268–276

59. Jin W, Barzilay R, Jaakkola T (2018) Junction tree variational autoencoder for molecular graph generation. arXiv preprint arXiv:180204364

60. Sanchez-Lengeling B, Outeiral C, Guimaraes GL, Aspuru-Guzik A (2017) Optimizing distributions over molecular space An objective-reinforced generative adversarial network for inverse-design chemistry (ORGANIC). ChemRxiv. https://doi.org/10.26434/chemrxiv.5309668.v2

61. Prykhodko O, Johansson S, Kotsias P-C et al (2019) A de novo molecular generation method using latent vector based generative adversarial network. J Cheminform 11:74

62. Elton DC, Boukouvalas Z, Fuge MD, Chung PW (2019) Deep learning for molecular design—a review of the state of the art. Mol Syst Design Eng 4:828–849

63. Coley CW, Green WH, Jensen KF (2018) Machine learning in computer-aided synthesis planning. Acc Chem Res 51:1281–1289

64. Engkvist O, Norrby P-O, Selmi N et al (2018) Computational prediction of chemical reactions: current status and outlook. Drug Discov Today 23:1203–1218

65. Gao H, Struble TJ, Coley CW et al (2018) Using machine learning to predict suitable conditions for organic reactions. ACS Cent Sci 4:1465–1476

66. Coley CW, Jin W, Rogers L et al (2019) A graph-convolutional neural network model for the prediction of chemical reactivity. Chem Sci 10:370–377

67. Lowe DM (2012) Extraction of chemical structures and reactions from the literature. PhD University of Cambridge, Cambridge

68. Reaxys. In: Reaxys. Available at www.reaxys.com. Accessed 1 Jan 2020

69. Ghiandoni GM, Bodkin MJ, Chen B et al (2020) Enhancing reaction-based de novo design using a multi-label reaction class recommender. J Comput Aided Mol Des. https://doi.org/10.1007/s10822-020-00300-6

70. Patel H, Bodkin MJ, Chen B, Gillet VJ (2009) Knowledge-based approach to de novo design using reaction vectors. J Chem Inf Model 49:1163–1184

71. Hristozov D, Bodkin M, Chen B et al (2012) ChemInform abstract: validation of reaction vectors for de novo design. ChemInform 43:50

72. Sheridan RP (2019) Interpretation of QSAR models by coloring atoms according to changes in predicted activity: how robust is it? J Chem Inf Model 59:1324–1337

73. Webel HE, Kimber TB, Radetzki S et al (2020) Revealing cytotoxic substructures in molecules using deep learning. J Comput Aided Mol Des. https://doi.org/10.1007/s10822-020-00310-4

74. Sculley D, Holt G, Golovin D et al (2015) Hidden technical debt in machine learning systems. Adv Neural Inf Process Syst 2:2503–2511

75. Green DVS, Pickett S, Luscombe C et al (2019) BRADSHAW: a system for automated molecular design. J Comput Aided Mol Des. https://doi.org/10.1007/s10822-019-00243-7

76. Cui J, Zhang H, Han H et al (2018) Improving 2D Face Recognition via Discriminative Face Depth Estimation. 2018 International Conference on Biometrics (ICB)

77. Cha KH, Petrick N, Pezeshk A et al (2020) Evaluation of data augmentation via synthetic images for improved breast mass detection on mammograms using deep learning. J Med Imag (Bellingham) 7:012703

78. Fei-Fei L, Deng J, Li K (2010) ImageNet: constructing a large-scale image database. J Vision 9:1037–1037

79. Samek W, Müller K-R (2019) Towards explainable artificial intelligence explainable. AI: interpreting, Explaining and Visualizing Deep Learning. Springer, Cham, pp 5–22

80. Alber M, Lapuschkin S, Seegerer P et al (2019) iNNvestigate neural networks. J Mach Learn Res 20:1–8

81. Sieg J, Flachsenberg F, Rarey M (2019) In need of bias control: evaluating chemical data for machine learning in structure-based virtual screening. J Chem Inf Model 59:947–961

82. Lapuschkin S, Wäldchen S, Binder A et al (2019) Unmasking Clever Hans predictors and assessing what machines really learn. Nat Commun 10:1096

83. Heaven D (2019) Why deep-learning AIs are so easy to fool. Nature 574:163–166

84. Wallach I, Heifets A (2018) Most ligand-based classification benchmarks reward memorization rather than generalization. J Chem Inf Model 58:916–932

85. Carlini N, Liu C, Kos J, et al (2018) The secret sharer: measuring unintended neural network memorization & extracting secrets. arXiv preprint arXiv:180208232

86. Wu Z, Ramsundar B, Feinberg EN et al (2018) MoleculeNet: a benchmark for molecular machine learning. Chem Sci 9:513–530

87. Brown N, Fiscato M, Segler MHS, Vaucher AC (2019) GuacaMol: benchmarking models for de novo molecular design. J Chem Inf Model 59:1096–1108

88. Raschka S (2018) Model evaluation, model selection, and algorithm selection in machine learning. arXiv preprint arXiv:181112808

89. Robinson MC, Glen RC, Lee AA (2020) Validating the validation: reanalyzing a large-scale comparison of deep learning and machine learning models for bioactivity prediction. J Comput Aided Mol Des. https://doi.org/10.1007/s10822-019-00274-0

90. Walters WP, Murcko M (2020) Assessing the impact of generative AI on medicinal chemistry. Nat Biotechnol 38:143–145

91. Chuang KV, Keiser MJ (2018) adversarial controls for scientific machine learning. ACS Chem Biol 13:2819–2821

92. Eykholt K, Evtimov I, Fernandes E, et al (2018) Robust Physical-World Attacks on Deep Learning Visual Classification. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition