doi:10.2533/chimia.2023.22 Chimia 77 (2023) 22–30 © K. Jorner

Putting Chemical Knowledge to Work in Machine Learning for Reactivity

Kjell Jorner*

Abstract: Machine learning has been used to study chemical reactivity for a long time in fields such as physical organic chemistry, chemometrics and cheminformatics. Recent advances in computer science have resulted in deep neural networks that can learn directly from the molecular structure. Neural networks are a good choice when large amounts of data are available. However, many datasets in chemistry are small, and models utilizing chemical knowledge are required for good performance. Adding chemical knowledge can be achieved either by adding more information about the molecules or by adjusting the model architecture itself. The current method of choice for adding more information is descriptors based on computed quantum-chemical properties. Exciting new research directions show that it is possible to augment deep learning with such descriptors for better performance in the low-data regime. To modify the models, differentiable programming enables seamless merging of neural networks with mathematical models from chemistry and physics. The resulting methods are also more data-efficient and make better predictions for molecules that are different from the initial dataset on which they were trained. Application of these chemistry-informed machine learning methods promise to accelerate research in fields such as drug design, materials design, catalysis and reactivity.

Keywords: Digital chemistry · Machine learning · Reactivity

Kjell Jorner received his PhD (2018) in computational physical organic chemistry from Uppsala University, Sweden, under the supervision of Henrik Ottosson. After postdoctoral work at AstraZeneca UK (2018–2020) on reaction prediction models in process chemistry, he received an International Postdoc grant from the Swedish Research Council to work with Alán Aspuru-Guzik at the University of Toronto (2021–2022). His research there focused on computer-assisted design of functional molecules and catalysts using machine learning and artificial intelligence. Since January 2023, Kjell is Assistant Professor of Digital Chemistry at ETH Zurich, where he focuses his research on chemistry-informed machine learning for catalysis.

Artificial intelligence (AI) is a broad research field that aims to use computers to accomplish tasks that were previously only achievable with active input from intelligent humans. Some of the most famous tools in the AI toolbox is *machine learning* (ML), which refers to algorithms that learn from data, and *deep learning* (DL), which refers to such algorithms based on neural networks (NNs) with many layers. During the last 5–10 years, interest in applying ML in chemistry has exploded, with countless research articles, reviews,[1,2] perspectives[3,4] and books.[5,6] The large interest in these methods stems from their potential to accelerate discovery and development of chemical solutions to important societal challenges and to bring these to the market faster. Sustainable energy production,[7] chemical production,[8,9] drug design,[10] and the computer-aided synthesis of (drug-like) molecules[11] are just some of the challenges where the application of AI in chemistry can make a difference.

The new wave of AI methods in chemistry follows significant advances in computer science.[12] One application where DL excels is image recognition (Fig. 1a), with models trained on the large datasets of the internet era.[13] These powerful models represent a shift from the previous hand-crafted, rule-based expert AI systems, to NNs that learn the rules implicitly from the data. The rationale behind this shift is that ever more flexible NNs can come up with more and sometimes better rules than experts, provided that they are given sufficient data to learn from. One example of this transition is when Google Translate went from an expert system consisting of 500,000 lines of code to a new and better DL system with only 500 lines of code.[14] Another recent triumph for DL is image generation, where algorithms such as DALL-E[15,16] and Stable Diffusion[17] are mature enough to co-create scientific journal cover art[18] and illustrations.[19] Deep learning-based AI algorithms are now the champions not only of chess, but of more complex games such as Go[20] and multi-player poker.[21]

The advances in methodology in computer science have also spilled over to the physical sciences, where algorithms such as graph neural networks (GNNs) are applied to chemical problems (Fig. 1b). Triumphs include the AlphaFold algorithm for protein structure prediction (Fig. 1c),[22] that beats previous expert-devised systems and has the potential to boost research in areas such as structure-based drug design and biocatalysis.[23–25] In quantum chemistry, promising advances include machine-learned approximations to the universal density functional,[26,27] described as one of the holy grails that would enable computer-aided catalysis and reaction design.[28]

Generalization, Chemical Space and Applicability Domain

Despite these promising advances, successful application of ML in chemistry has so far been limited. The root of the problem has to do with the amount and type of available data for training the models. Generally speaking, flexible DL methods need large data sets for optimal performance. For comparison, DALL-E was

Correspondence: Prof. K. Jorner, E-Mail: kjell.jorner@chem.ethz.ch
ETH Zürich, Institute of Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, Vladimir-Prelog-Weg 1, HCI E 137, CH-8093 Zürich, Switzerland

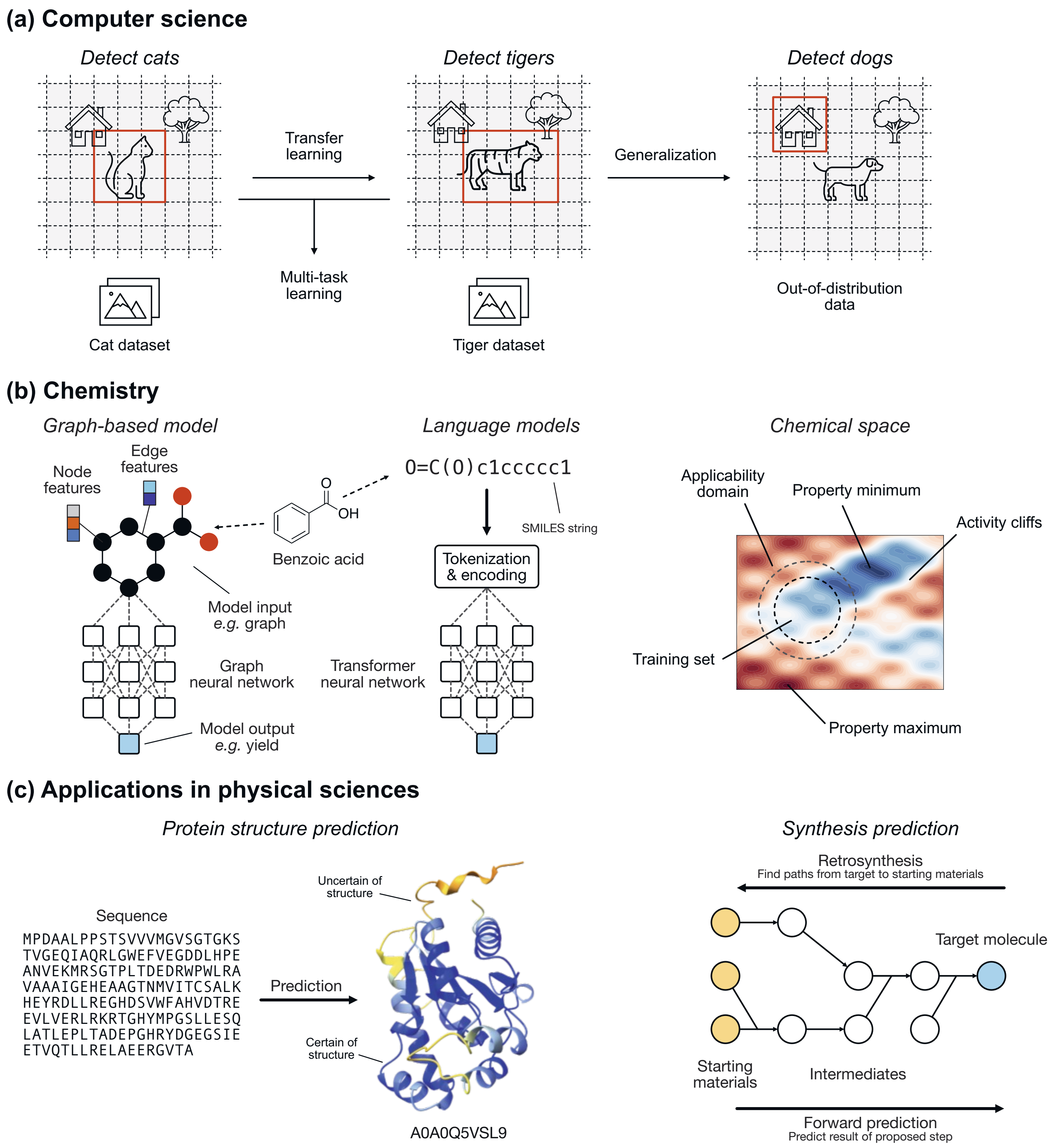

Fig. 1. (a) Application of deep learning originating in computer science. Computer vision is a perfect area for using transfer learning and multi-task learning on multiple related image recognition tasks and datasets. Models struggle with out-of-distribution data that is different from what they were trained on. (b) Application of deep learning in chemistry mainly involves converting the molecular structure into graph-based or text-based representations and using the corresponding architectures. Models have an applicability domain in chemical space that depends on the training set. Generalization is difficult, not least due to activity cliffs. (c) Successful applications of deep learning include tasks with much available data, such as protein structure prediction and synthesis prediction.

trained on 250 million text-image pairs[15] and Stable Diffusion on 2.3 billion images.[29] In chemistry, large datasets exist only for some applications. AlphaFold was trained[22] on 170,000 protein structures from the Protein Data Bank (PDB).[30] Models for retrosynthesis planning (Fig. 1c) are most often trained[31] on a dataset of *ca.* 3.3 million reactions text-mined from the US Patent Office,[32] or on data from commercial databases such as Reaxys, also with millions of reactions.[33] In contrast, the size of reactivity datasets modeled using traditional approaches can have as little as 11 data points and often not larger than a hundred.[34] Even the larger datasets such as the PDB and reaction databases are considerably smaller than the datasets used for image and language models.

Translation models for languages with much less available data often perform poorly, and in an infamous example, image recognition models trained on an insufficient sample of photos led to critical failures that humiliated minorities and sparked global discussions about the bias inherited by these technologies.[35] In the same way, DL models trained for chemistry face problems when making predictions outside the *chemical space* on which they were trained (Fig. 1b). In statistical terminology, the models do not generalize well. A central concept with a long history in chemistry is the *applicability domain,*[36] which constitutes a description of the types of molecules and property ranges for which a model is reliable. Any ML model used for regulatory purposes in chemistry is actually recommended to have an applicability domain according to OECD guidelines.[37] Generalization outside the training space is particularly important in chemistry, where different compound classes occupy distinct parts of 'chemical space' and where properties can change unpredictably with structure in terms of so-called *activity cliffs.*[38,39] While a lesser extent of generalization is required for library searches of similar compounds, more unconstrained inverse design[40] of molecules with new functionalities places much larger requirements on models with very large applicability domains.[41]

Utilizing Existing Data better or Gathering more Informative New Data

The problem of generalization has typically been tackled by approaches such as *transfer learning,* or *multi-task learn*ing (Fig. 1a). In transfer learning, a NN model is first trained on a (larger) dataset for one task in a stage called *pre-training.* The pre-trained model is then further trained on a second (smaller) dataset for a related task in a stage called *fine-tuning.* The idea is that the model learns general trends from the first task that are also valid for the second task. In multi-task learning, the model learns on both datasets and tasks at the same time instead of the sequential training used in transfer learning. Both multi-task and transfer learning effectively increase the size of the training data. There are successful examples of transfer learning in chemistry where tasks are similar, for example, learning computed atomization energies using density functional theory (DFT) vs. coupled cluster theory,[42] or improving reaction activation energies computed at a lower level of theory.[43,44] However, applications to, in particular, experimental data has been limited, with successful examples including the work by Reymond on carbohydrate reactivity.[45] In another recent example, researchers from the Janssen pharmaceutical company demonstrated improved predictions on in-house chemistry when pre-training on both patent data and academic data.[46]

Another approach to handle the problem of limited data is active learning, where an algorithm suggests which data points should be collected next to maximally improve the predictive accuracy of the ML model.[47] However, due to time constraints it is often not practical to gather new data as it comes from costly simulations or time- and labor-intensive experiments. Active learning should rather be seen as an approach for guiding the collection of new data in such a way that it is maximally informative for ML, than as a method for improving the models using existing data. Although gathering new data, and making it findable, accessible, interoperable and reusable (FAIR) is a very important task,[48] we will here focus on another approach that can make a difference with the data that we already have.

Adding more Information about the Data or the Task

As an alternative to utilizing the data better with transfer- or multi-task learning, or gathering new data with active learning, we might consider using additional information about the data or the prediction task itself. The success of DL over expert systems is rooted in the unbiased approach to learning. But when faced with too little data, we have to re-evaluate. Why should we throw away most of what we as chemists know about our problem or any additional information that we might have on the molecules or materials in question? While the old expert systems relied only on our own intuition to craft expert rules, we have now surrendered that intuition almost completely. The answer to our problems might be to put just enough chemical knowledge into the models to help them on the way, but not bias them too much. Machine learning practitioners prefer to talk about *inductive bias* or *strong priors.*[49] As expressed by Goodfellow and co-workers:

> *"In order to generalize well, machine learning algorithms need to be guided by prior beliefs about what kind of function they should learn."*[12]

In the simplest case, this can just mean choosing the right model architecture for the modality of the data (architecture prior).[50] For example, molecules are a natural fit for GNNs. But we can go further. It has recently become apparent that accommodation of molecular symmetries such as translation and rotation in three-dimensional GNNs leads to better accuracy and data efficiency.[51] This is a natural prior for molecular data that has no apparent drawbacks. Similarly, information such as bond order or hybridization state is routinely added to GNNs.[52] In the following two sections, we will explore two recent trends in these directions: (1) adding more chemical information through expert descriptors, and (2) incorporating physical models into the ML frameworks themselves through differentiable programming. The main focus will be on applications related to reactivity prediction due to our own interest in this area. Reactivity is addressed in a broad sense, ranging from activation energies of individual reactions to predictions of selectivity and retrosynthesis pathways.

Combining Descriptors with Deep Learning for Reaction Prediction

Machine learning has a long history in chemistry, going from linear free energy relationships from the first half of the last century to chemometrics and cheminformatics in the second half.[53] One of the machine-readable *molecular representations*[54] that have dominated is *descriptors* (Fig. 2a).[55] The translation of a molecule into its representation is sometimes called *featurization,* or in DL lingo, *embedding.*[56] The descriptor representation is most often combined with traditional ML algorithms such as multivariate linear regression or non-linear methods such as Random Forest.[54] In contrast, modern DL architectures for molecules are based either on graph representation of molecules or natural language processing of string representations of molecules such as SMILES[57] or SELFIES (Fig. 1b).[58] Given sufficient data, these models learn their own optimal representation of the molecules rather than rely on expert encoding schemes.[59] While DL approaches generally perform better in the big-data regime and traditional descriptors better in the low-data regime,[56] recent advances have shown the advantages of combining both approaches.[60]

The more traditional descriptor-based approaches are sometimes called quantitative structure–activity relationships (QSAR), especially when referring to biological activity, and sometimes quantitative structure–property (QSPR) relationships more generally. For reactivity, the concept QSRR is often used, or QSSR when referring specifically to selectivity prediction. QSRR has a long history under the name of linear free energy relationships,[53] with perhaps the first study by Brønstedt and Pedersen in 1924.[61] One of the first modern QSSR models for catalysis was a study from 1997 by Norrby and co-workers, who related the regio- and stereoselectivity of palladium-catalyzed allylation reactions to steric and strain descriptors.[62] This type of descriptors are often described as 'expert crafted',[56] as the right choice of descriptors requires some domain knowledge. The procedure is related to the concept of *feature engineering* in ML. In more recent times, descriptors from quantum mechanics (QM) have been used together with more sophisticated models by, among others, the groups of Sigman.[63,64] and Doyle.[65,66] A related approach uses *compara-*

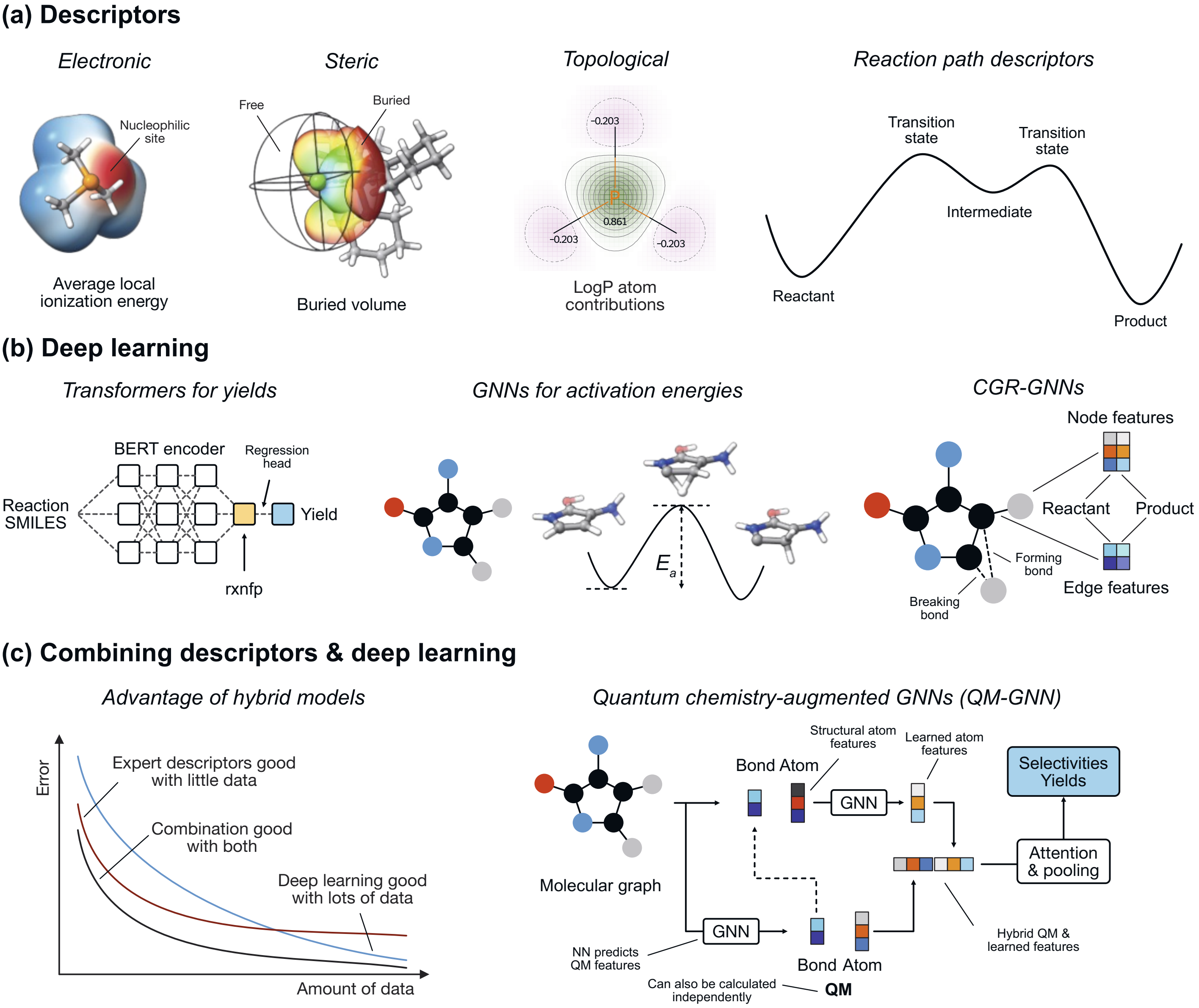

Fig. 2. (a) Traditional descriptors used for reactivity are often steric or electronic and computed from quantum mechanics and the 3D structure of the molecule. 'Topological' descriptors are calculated from just the 2D structure. Descriptors are mostly calculated for reactants, but additional information can be gained by incorporating descriptors calculated for products, intermediates and transition states along the reaction path. (b) Deep learning approaches based on graphs and SMILES have been used to predict, e.g., activation energies and reaction yields. The CGR-GNN combines features of both the reactant and product graphs into a reaction graph representation. (c) Initial work combining descriptors with deep learning have shown promise in the low-to-medium data regime. The QM-GNN combines QM descriptors with learned features from a GNN.

tive molecular field analysis (CoMFA), an approach originating in drug design in which molecules are aligned in a box and steric and electronic properties are calculated on a grid.[67,68] These descriptors depend on the three-dimensional structure of the molecules, but so-called *topological* or two-dimensional descriptors that only depend on the molecular connectivity are also available (Fig. 2a).[69] Descriptors are normally calculated for reactants, but product descriptors can also be calculated.[68,70,71]

Descriptor approaches have also been combined with mechanistic QM calculations of the reaction pathway (Fig. 2a).[72] Perhaps the first study in this direction was by Sigman and Toste, who included transition state descriptors to model the oxidative amination of tetrahydroisoquinolines under chiral-anion phase-transfer catalysis.[73] High-energy intermediates can also be informative, as shown, for example, by calculation of regioselectivities in electrophilic aromatic substitution reactions by Norrby and co-workers by including the energies of the σ complexes (Wheland intermediates).[74] Buttar and co-workers used both ground state and transition state descriptors to predict activation energies and selectivities for the nucleophilic aromatic substitution reaction,[75] and Zhao and co-workers showed that introduction of transition state features increased the predictive performance when modelling activation energies also for enzymatic reactions.[76] While reaction path calculations can be time-consuming, Grzybowski and co-workers recently showed that information from approximate transition states could be used to predict the facial selectivity of Michael additions and Diels–Alder cycloadditions.[77] The calculations can also be accelerated by using semi-empirical methods, as shown by Hartwig, Norrby and co-workers for iridium-catalyzed borylation of C–H bonds reactions, leading to run-times of only a few minutes per reaction.[78] Their resulting hybrid ML model with mechanistic information achieved 100% accuracy on site-selectivity prediction, outclassing highly skilled synthetic chemists.

The downsides of descriptor approaches include requiring an expert to choose appropriate descriptors that are then also specific to a certain reaction. QM descriptors are also computationally expensive and manually time-consuming to calculate, a problem that can partly be alleviated by automatic workflows,[79] descriptor databases,[80] or by calculating them with ML approaches.[81,82] An alternative way to incorporate quantum-chemical information in a more general way is the so-called *quantum machine learning* (QML) approach,[83,84] where molecular representations are created based on two-, three-body and higher order interaction terms inspired from quantum mechanics.[85] Recently, Corminboeuf and co-workers have developed QML reaction representations that do not require adaptation for different reaction types.[71,86] We also note that there is a rich literature on the use of descriptors for heterogeneous catalysis, including *d-band theory,*[87] *linear scaling relations* and *volcano plots.*[88] We refer the interested reader to some of the excellent reviews for further details.[89]

The methods discussed so far make use of more traditional ML approaches in combination with hand-picked descriptors. We will now examine more modern DL methods where the methods of choice are GNNs based on the molecular connectivity graph, or natural language models such as Transformers, based on the SMILES string representation of molecules and reactions (Fig. 1b).[90] Whether working on molecular graphs or SMILES, recent investigations show close connections between GNNs and transformers under the unifying umbrella of *geometric deep learning.*[91,92] For periodic materials, architectures such as crystal graph convolutional neural networks[93] or periodic graph transformers[94] have been developed. There is also a long history of applying NNs in chemistry. A prominent example from 1990 is the prediction of electrophilic aromatic substitution selectivity by scientists from Upjohn using connection table representations and shallow feed-forward NNs.[95,96] Modern deep NN models have proven proficient at both retrosynthesis and forward reaction prediction of different types.[2] In a similar application as the previous example from Upjohn, Jensen and co-workers developed a multitask model for regioselectivity in C-H functionalization,[97] that outperformed previous approaches based on quantum-chemical calculations. In an elegant approach, Heid and Green developed a GNN based on the condensed graph of reaction (CGR),[98,99] which can be seen as a superposition of the reactant and products graphs (Fig. 2b).[100] The model showed very good performance on a range of quantitative reactivity tasks. Reymond, Laino and co-workers used transformer models to predict reaction yields (Fig. 2b), observing good performance on literature HTE datasets but poorer performance on patent data, likely due to different data homogeneity and quality.[101] Activation energies is another target that was tackled by Green and co-workers for gas-phase isomerization reactions of small molecules (Fig. 2b).[102]

While pure DL models typically require many hundred or thousands of data points, descriptor models have been used with as little as 11 data points.[34] Although it is statistically questionable whether conclusions based on models with so few data points can really be trusted,[103] it seems clear also from other studies that models based on well-chosen descriptors perform better in the low-data regime than those having access to only molecular structural information.[75] One natural approach is to marry both approaches and get the best of both worlds (Fig. 2c). This is what Green, Jensen and co-workers[81] did by developing quantum mechanics-augmented graph neural networks (QM-GNNs). By incorporating traditional reactivity descriptors such as atomic charges, Fukui indices and NMR chemical shifts into a GNN model, the performance increased significantly in the low-data regime and the models generalized better to unseen regions of chemical space as demonstrated by a scaffold split (Fig. 2c). To circumvent the high cost of generating the descriptors with QM calculations, the authors trained NNs to predict them directly from the chemical structure. Coley and Stuyver later applied the same architecture to S_N2 nucleophilic substitution and E2 elimination reactions.[104] They showed that the models utilized descriptor information according to ‘hard’ (*e.g.,* atomic charges) and ‘soft’ (*e.g.,* Fukui indices) interactions from Pearson’s theory of hard and soft acids and bases.[105] Studies by Balcells and co-workers[106] and Gomes and co-workers[107] have explored adding QM information from the natural bond orbital (NBO) theory[108] into GNNs. Recently, Schneider and co-workers, including scientists from Roche, utilized 2D and 3D GNNs for the prediction of binary reaction outcome, yield and regioselectivity of late-stage borylation of druglike molecules.[109] Incorporating QM atomic charges into the models failed to produce any significant improvement, although the authors acknowledged that the iridium-catalyzed borylation reaction that they modelled is mainly sensitive to steric rather than electronic effects. These handful of studies indicate that there can be a substantial value of including descriptor information in DL models, although it remains to be seen which descriptors are most informative and which predictions tasks benefit the most.

Building Auto-differentiable Physical Models

All modern DL frameworks are built on an *automatic differentiation* (AD) engine that trains the NNs.[110] In practice, this is usually done by automatic application of the familiar *chain rule* for expressing the derivatives of composite functions. Neural network models built on large datasets have achieved great accuracy for predicting the energies of molecular systems. These models are sometimes called *neural network potentials* (NNPs) or *machine learning interatomic potentials.* Based on the pioneering work of Behler and Parinello,[111] numerous models have been proposed.[112] Arguably the most successful in molecular chemistry has been the ANI family of models.[113] While the initial version[113] covered only the chemical elements H, C, N and O, the later ANI-2x[114] included also S, F, and Cl, and Schrödinger-ANI[115] expanded the set to also include P. In the materials modelling field, 3D GNNs such as GemNet[116] have shown promising performance on the OC20 and OC22 datasets[117,118] for predicting absorption energies and geometries of small molecules on metal and oxide surfaces. Although this task is not directly related to catalytic activity, absorption energies can later be correlated with energy barriers through linear scaling relationships (Fig. 3a).[119]

Conventional NNPs do not work for modelling reactivity as they are not trained on off-equilibrium data. Custom-made reactive potentials can be constructed, but they are reaction specific.[120] Very recently, Lubbers, Messerly, Smith and co-workers expanded the ANI framework with an active learning scheme based on a computational ‘nanoreactor’ (Fig. 3a).[121] They showed that the resulting dataset covered a large part of reactive chemical space and that the NNP, dubbed ANI-nr, could be used for applications such as methane combustion or an *in silico* Miller-Urey experiment. Development of further datasets including reactive geometries will be key in expanding the utility of reactive NNPs.[122] In spite of these promising advances, the current generation of NNPs are limited by the use of pairwise potentials. Expanding the potentials such as ANI-nr to additional elements beyond H, C, N and O therefore entails enormous amounts of additional training data for parametrization, a problem well-known from for example density functional tight binding.[123] There is therefore a large need of developing models which can circumvent these problems and generalize well with less training data.

The same software tools that enable automatic differentiation of NNs can also be used to make physical models differentiable (Fig. 3b).[124] Initial work in this direction in chemistry has focused on the automatic calculation of higher-order derivatives for quantum-chemical methods,[125–127] but recent developments have significantly widened the scope. Once a physical model is coded in an auto-differentiable way, it can be combined with NN components to create hybrid models (Fig. 3b). Integrating traditional DL with strong priors using differentiable physical models will likely lead to better performance, data efficiency and generalization. In a famous example, a team led by researchers from DeepMind

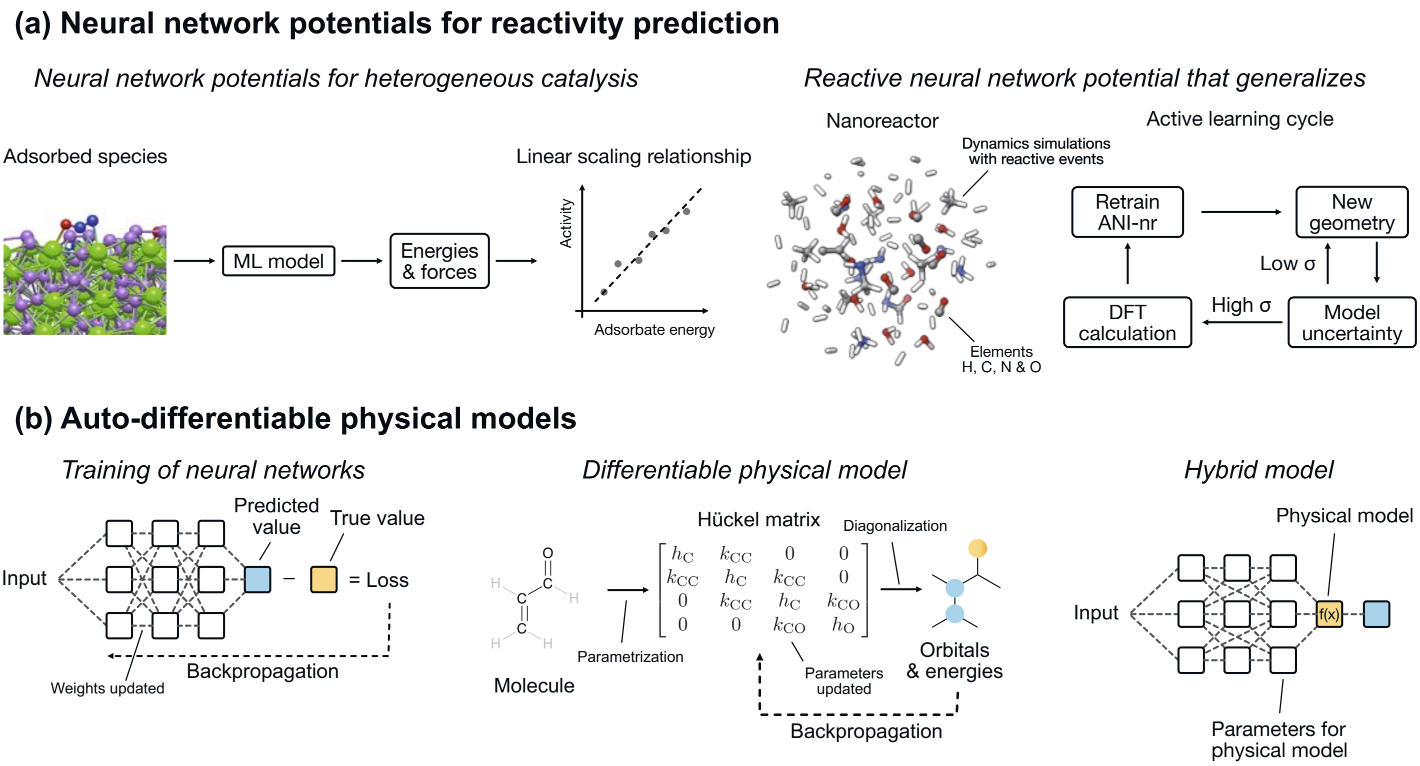

Fig. 3. (a) Neural network potentials have seen limited used in reactivity prediction. Generalizable reactive NNPs, such as ANI-nr, are under development. Non-reactive NNPs can also be used for reactivity prediction by utilizing linear scaling relationships between adsorbate energies and catalyst activity. (b) Neural networks are trained through backpropagation of loss. Differentiable physical models can be 'trained' in a similar way to optimize parameters based on data. Combination of NNs and differentiable physical models into hybrid models is a promising venue for creating models that generalize well.

showed that a machine learned density functional approximation could extrapolate well beyond the training set and for example beat standard functionals for reaction barriers heights.[27] In a similar vein, Vinko and co-workers demonstrated that a NN approximation of the density functional could be trained with experimental data from only eight diatomic molecules and generalize to larger molecular systems.[26] Finding better approximations to the universal density functional is indeed one of the holy grails of computational chemistry, that could allow more accurate simulations of chemical reactivity.[28]

Despite these promising advances, a typical DFT calculation still scales on the order of $\mathcal{O}(N_{bf}^{3}) - \mathcal{O}(N_{bf}^{4})$, where N_{bf} is the number of basis functions used to describe the system. Hybrid approaches with much better scaling are needed to be competitive with pure ML, that in favorable cases scales as $\mathcal{O}(N_{a}^{1})$ where N_a is the number of atoms in the system. This requirement is especially important for high-throughput virtual screening, reaction network exploration[128] and generative models, where thousands or millions of molecules need to be evaluated. The OrbNet model from Manby, Miller and co-workers[129,130] uses atomic orbital features from semi-empirical quantum chemistry to reduce the scaling to $\mathcal{O}(N_{bf}^{3})$.[131] Similarly, Tretiak and co-workers combined semi-empirical methods with the hierarchically interacting particle neural network (HIP-NN) model and showed significantly improved generalization compared to pure NN approaches.[132,133] Yaron and co-workers implemented an auto-differentiable DFTB model, learning the parameters of more classical functions rather than using a NN.[134] Further speed-ups need to use simpler models such as physically informed force fields with $\mathcal{O}(N_{a}^{2})$ scaling. An example is the ReaxFF reactive force field[135] that was recently rewritten with differentiable programming to allow much faster parametrization and easier introduction of new functional forms into the force field.[136] Hybrid force field and quantum-chemical approaches that incorporate some electronic effects have a scaling somewhere between $\mathcal{O}(N_{a}^{2})$ and $\mathcal{O}(N_{bf}^{3})$.[137,138] Even though this is far from the $\mathcal{O}(N_{a}^{1})$ possible with pure NNPs, it may be good enough to generate hundreds of thousands of datapoints for many applications.

As examples of further applications, AD has been used to tune the parameters of QML representations,[139] and it can also be utilized for inverse design. Aspuru-Guzik and co-workers recently used an auto-differentiable Hückel molecular orbital code to optimize model parameters and perform inverse design by alchemical gradient- based optimization of atom identity.[140] Ongoing work with introducing automatic differentiation at the compiler level[141] will open possibilities for exploiting the many large legacy codebases built up in science over the last decades.

Conclusions and Outlook

Models based on deep learning have demonstrated outstanding performance for chemical problems given that sufficient training data on the order of thousands or millions of data points is available. These amounts of data are accessible for a handful of applications, such as protein structure prediction, but generally datasets in chemistry are small, in the order of tens or hundreds of data points. Expert-chosen descriptors or physical models are traditionally used to model these datasets, but the resulting models do not scale as well with increasing dataset size compared to deep learning. Recent developments show promising directions or merging deep learning with descriptors and physical models. Descriptors have been combined with graph neural networks for reactivity prediction, and auto-differentiable implementations of physical models have been merged with deep learning. The resulting models are generally more data efficient than pure deep learning and generalize better outside the chemical space of the training data. Data efficiency is important when modelling the small datasets found in chemistry, while good generalization is crucial for the application of the models in inverse design and high-throughput virtual screening. We believe that these emerging methods are the first of a new generation of ML models in chemistry, where prior chemical knowledge is incorporated to a larger extent. The subsequent application of these more robust and generalizable models would enable breakthroughs to create chemical solutions to our societal problems.

Acknowledgements

K.J. thanks Dr. Eno Paenurk, Dr. Sergio Pablo-García, Mr. Stefan P. Schmid, as well as the anonymous reviewers for critically reading the manuscript and providing helpful feedback.

Received: November 14, 2022

[1] A. C. Mater, M. L. Coote, *J. Chem. Inf. Model*. **2019**, *59*, 2545, https://doi.org/10.1021/acs.jcim.9b00266.
[2] P. Schwaller, A. C. Vaucher, R. Laplaza, C. Bunne, A. Krause, C. Corminboeuf, T. Laino, *WIREs Comput. Mol. Sci*. **2022**, *12*, https://doi.org/10.1002/wcms.1604.
[3] A. Aspuru-Guzik, R. Lindh, M. Reiher, *ACS Cent. Sci*. **2018**, *4*, 144, https://doi.org/10.1021/acscentsci.7b00550.
[4] A. Tkatchenko, *Nat. Commun*. **2020**, *11*, 4125, https://doi.org/10.1038/s41467-020-17844-8.
[5] J. P. Janet, H. J. Kulik, 'Machine Learning in chemistry', American Chemical Society, Washington, DC, USA, **2020**, https://doi.org/10.1021/acs.infocus.7e4001.
[6] H. M. Cartwright, Ed., 'Machine learning in chemistry: The impact of artificial intelligence', Royal Society Of Chemistry, Cambridge, **2020**, https://doi.org/10.1039/9781839160233.
[7] D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwaraknath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, A. Aspuru-Guzik, *Nat. Rev. Mater*. **2018**, *3*, 5, https://doi.org/10.1038/s41578-018-0005-z.
[8] P. Anastas, N. Eghbali, *Chem. Soc. Rev*. **2010**, *39*, 301, https://doi.org/10.1039/B918763B.
[9] R. A. Sheldon, I. W. C. E. Arends, U. Hanefeld, 'Green chemistry and catalysis', Wiley, **2007**, https://doi.org/10.1002/9783527611003.
[10] J. Jiménez-Luna, F. Grisoni, N. Weskamp, G. Schneider, *Expert Opin. Drug Discov*. **2021**, *16*, 949, https://doi.org/10.1080/17460441.2021.1909567.
[11] A. Thakkar, S. Johansson, K. Jorner, D. Buttar, J.-L. Reymond, O. Engkvist, *React. Chem. Eng*. **2021**, *6*, 27, https://doi.org/10.1039/D0RE00340A.
[12] I. Goodfellow, Y. Bengio, A. Courville, 'Deep learning', MIT Press, **2016**.
[13] A. Krizhevsky, I. Sutskever, G. E. Hinton, *Commun. ACM* **2017**, *60*, 84, https://doi.org/10.1145/3065386.
[14] Blue lines: @Google's old Translate program, 500 k lines of stats-focused code. Green: Now, 500 lines of @tensorflow. https://t.co/MwNPjYd4KJ, https://twitter.com/DynamicWebPaige/status/915326707107844097, accessed October 23, 2022.
[15] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, I. Sutskever, **2021**, https://doi.org/10.48550/arXiv.2102.12092.
[16] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, M. Chen, **2022**, https://doi.org/10.48550/arXiv.2204.06125.
[17] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, **2022**, https://doi.org/10.48550/arXiv.2112.10752.
[18] Mostly created in @midjourney, starting with several attempted ideas and prompts, with the winners evolving *via* lots of 'variation' requests https://t.co/yreiUw5gR9, https://twitter.com/richardhwest/status/1584755290649415681 , accessed November 13, 2022.
[19] Editorial, *Nat. Biomed. Eng*. **2022**, *6*, 1087, https://doi.org/10.1038/s41551-022-00957-4.
[20] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, *Nature* **2016**, *529*, 484, https://doi.org/10.1038/nature16961.
[21] N. Brown, T. Sandholm, *Science* **2019**, *365*, 885, https://doi.org/10.1126/science.aay2400.
[22] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583, https://doi.org/10.1038/s41586-021-03819-2.
[23] F. Ren, X. Ding, M. Zheng, M. Korzinkin, X. Cai, W. Zhu, A. Mantsyzov, A. Aliper, V. Aladinskiy, Z. Cao, S. Kong, X. Long, B. H. M. Liu, Y. Liu, V. Naumov, A. Shneyderman, I. V. Ozerov, J. Wang, F. W. Pun, A. Aspuru-Guzik, M. Levitt, A. Zhavoronkov, **2022**, https://doi.org/10.48550/arXiv.2201.09647.
[24] F. Wong, A. Krishnan, E. J. Zheng, H. Stärk, A. L. Manson, A. M. Earl, T. Jaakkola, J. J. Collins, *Mol. Syst. Biol*. **2022**, *18*, https://doi.org/10.15252/msb.202211081.
[25] D. Yi, T. Bayer, C. P. S. Badenhorst, S. Wu, M. Doerr, M. Höhne, U. T. Bornscheuer, *Chem. Soc. Rev*. **2021**, *50*, 8003, https://doi.org/10.1039/D0CS01575J.
[26] M. F. Kasim, S. M. Vinko, *Phys. Rev. Lett*. **2021**, *127*, 126403, https://doi.org/10.1103/PhysRevLett.127.126403.
[27] J. Kirkpatrick, B. McMorrow, D. H. P. Turban, A. L. Gaunt, J. S. Spencer, A. G. D. G. Matthews, A. Obika, L. Thiry, M. Fortunato, D. Pfau, L. R. Castellanos, S. Petersen, A. W. R. Nelson, P. Kohli, P. Mori-Sánchez, D. Hassabis, A. J. Cohen, *Science* **2021**, *374*, 1385, https://doi.org/10.1126/science.abj6511.
[28] K. N. Houk, F. Liu, *Acc. Chem. Res*. **2017**, *50*, 539, https://doi.org/10.1021/acs.accounts.6b00532.
[29] CompVis/stable-diffusion: A latent text-to-image diffusion model, https://github.com/CompVis/stable- diffusion, accessed October 29, 2022.
[30] H. M. Berman, *Nucleic Acids Res*. **2000**, *28*, 235, https://doi.org/10.1093/nar/28.1.235.
[31] P. Schwaller, R. Petraglia, V. Zullo, V. H. Nair, R. A. Haeuselmann, R. Pisoni, C. Bekas, A. Iuliano, T. Laino, *Chem. Sci*. **2020**, *11*, 3316, https://doi.org/10.1039/C9SC05704H.
[32] D. M. Lowe, **2012**, https://doi.org/10.17863/CAM.16293.
[33] M. H. S. Segler, M. Preuss, M. P. Waller, *Nature* **2018**, *555*, 604, https://doi.org/10.1038/nature25978.
[34] D. M. Lustosa, A. Milo, *ACS Catal*. **2022**, 7886, https://doi.org/10.1021/acscatal.2c01741.
[35] M. Weiss, P. Tonella, in '2021 14th IEEE Conference on Software Testing, Verification and Validation (ICST)', IEEE, Porto de Galinhas, Brazil, **2021**, 24, https://doi.org/10.1109/ICST49551.2021.00015.
[36] T. Hanser, C. Barber, J. F. Marchaland, S. Werner, *SAR QSAR Environ. Res*. **2016**, *27*, 865, https://doi.org/10.1080/1062936X.2016.1250229.
[37] OECD, 'Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models', **2014**.
[38] J. L. Medina-Franco, A. L. Chávez-Hernández, E. López-López, F. I. Saldívar-González, *Mol. Inform*. **2022**, 2200116, https://doi.org/10.1002/minf.202200116.
[39] M. Aldeghi, D. E. Graff, N. Frey, J. A. Morrone, E. O. Pyzer-Knapp, K. E. Jordan, C. W. Coley, *J. Chem. Inf. Model*. **2022**, *62*, 4660, https://doi.org/10.1021/acs.jcim.2c00903.
[40] B. Sanchez-Lengeling, A. Aspuru-Guzik, *Science* **2018**, *361*, 360, https://doi.org/10.1126/science.aat2663.
[41] A. Nigam, R. Pollice, G. Tom, K. Jorner, L. A. Thiede, A. Kundaje, A. Aspuru-Guzik, **2022**, https://doi.org/10.48550/arXiv.2209.12487.
[42] J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev, A. E. Roitberg, *Nat. Commun*. **2019**, *10*, 2903, https://doi.org/10.1038/s41467-019-10827-4.
[43] S. Heinen, G. F. von Rudorff, O. A. von Lilienfeld, *J. Chem. Phys*. **2021**, *155*, 064105, https://doi.org/10.1063/5.0059742.
[44] E. H. E. Farrar, M. N. Grayson, *Chem. Sci*. **2022**, *13*, 7594, https://doi.org/10.1039/D2SC02925A.
[45] G. Pesciullesi, P. Schwaller, T. Laino, J.-L. Reymond, *Nat. Commun*. **2020**, 11, 4874, https://doi.org/10.1038/s41467-020-18671-7.
[46] P. Neves, K. McClure, J. Verhoeven, N. Dyubankova, R. Nugmanov, A. Gedich, S. Menon, z. Shi, J. Wegner, **2022**, https://doi.org/10.26434/chemrxiv-2022-5775s-v2.
[47] E. Shim, J. A. Kammeraad, Z. Xu, A. Tewari, T. Cernak, P. M. Zimmerman, *Chem. Sci*. **2022**, *13*, 6655, https://doi.org/10.1039/D1SC06932B.
[48] C. Bo, F. Maseras, N. López, *Nat. Catal*. **2018**, *1*, 809, https://doi.org/10.1038/s41929-018-0176-4.
[49] K. P. Murphy, 'Machine learning: A probabilistic perspective', MIT Press, Cambridge, MA, **2012**.
[50] F. Chollet, 'Deep learning with Python', Manning Publications, Shelter Island, **2021**.
[51] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, B. Kozinsky, *Nat. Commun*. **2022**, *13*, 2453, https://doi.org/10.1038/s41467-022-29939-5.
[52] C. Merkwirth, T. Lengauer, *J. Chem. Inf. Model*. **2005**, *45*, 1159, https://doi.org/10.1021/ci049613b.
[53] W. L. Williams, L. Zeng, T. Gensch, M. S. Sigman, A. G. Doyle, E. V. Anslyn, *ACS Cent. Sci*. **2021**, *7*, 1622, https://doi.org/10.1021/acscentsci.1c00535.
[54] L. Pattanaik, C. W. Coley, *Chem* **2020**, *6*, 1204, https://doi.org/10.1016/j.chempr.2020.05.002.
[55] M. Karelson, V. S. Lobanov, A. R. Katritzky, *Chem. Rev*. **1996**, 96, 1027, https://doi.org/10.1021/cr950202r.

[56] L. C. Gallegos, G. Luchini, P. C. St. John, S. Kim, R. S. Paton, *Acc. Chem. Res*. **2021**, *54*, 827, https://doi.org/10.1021/acs.accounts.0c00745.

[57] D. Weininger, in 'Handbook of Chemoinformatics', Ed. J. Gasteiger, Wiley-VCH Verlag GmbH, Weinheim, Germany, **2003**, 80, https://doi.org/10.1002/9783527618279.ch5.

[58] M. Krenn, F. Häse, A. Nigam, P. Friederich, A. Aspuru-Guzik, *Mach. Learn. Sci. Technol*. **2020**, *1*, 045024, https://doi.org/10.1088/2632-2153/aba947.

[59] K. Yang, K. Swanson, W. Jin, C. Coley, P. Eiden, H. Gao, A. Guzman-Perez, T. Hopper, B. Kelley, M. Mathea, A. Palmer, V. Settels, T. Jaakkola, K. Jensen, R. Barzilay, *J. Chem. Inf. Model*. **2019**, *59*, 3370, https://doi.org/10.1021/acs.jcim.9b00237.

[60] S.-Q. Zhang, L.-C. Xu, S.-W. Li, J. C. A. Oliveira, X. Li, L. Ackermann, X. Hong, *Chem. - Eur. J.* **2022**, chem.202202834, https://doi.org/10.1002/chem.202202834.

[61] J. N. Brönsted, K. Pedersen, *Z. Phys. Chem*. **1924**, *108U*, 185, https://doi.org/10.1515/zpch-1924-10814.

[62] J. D. Oslob, B. Åkermark, P. Helquist, P.-O. Norrby, *Organometallics* **1997**, *16*, 3015, https://doi.org/10.1021/om9700371.

[63] J. P. Reid, M. S. Sigman, *Nat. Rev. Chem*. **2018**, *2*, 290, https://doi.org/10.1038/s41570-018-0040-8.

[64] J. P. Reid, M. S. Sigman, *Nature* **2019**, *571*, 343, https://doi.org/10.1038/s41586-019-1384-z.

[65] D. T. Ahneman, J. G. Estrada, S. Lin, S. D. Dreher, A. G. Doyle, *Science* **2018**, *360*, 186, https://doi.org/10.1126/science.aar5169.

[66] A. M. Żurański, J. I. Martinez Alvarado, B. J. Shields, A. G. Doyle, *Acc. Chem. Res*. **2021**, *54*, 1856, https://doi.org/10.1021/acs.accounts.0c00770.

[67] K. B. Lipkowitz, M. Pradhan, *J. Org. Chem*. **2003**, *68*, 4648, https://doi.org/10.1021/jo0267697.

[68] A. F. Zahrt, J. J. Henle, B. T. Rose, Y. Wang, W. T. Darrow, S. E. Denmark, *Science* **2019**, *363*, eaau5631, https://doi.org/10.1126/science.aau5631.

[69] T. T. Metsänen, K. W. Lexa, C. B. Santiago, C. K. Chung, Y. Xu, Z. Liu, G. R. Humphrey, R. T. Ruck, E. C. Sherer, M. S. Sigman, *Chem. Sci*. **2018**, *9*, 6922, https://doi.org/10.1039/C8SC02089B.

[70] H. Clements, A. Flynn, B. Nicholls, D. Grosheva, T. Hyster, M. Sigman, **2021**, https://doi.org/10.26434/chemrxiv-2021-9gd5m.

[71] P. van Gerwen, A. Fabrizio, M. D. Wodrich, C. Corminboeuf, *Mach. Learn. Sci. Technol.* **2022**, *3*, 045005, https://doi.org/10.1088/2632-2153/ac8f1a.

[72] K. Jorner, A. Tomberg, C. Bauer, C. Sköld, P.-O. Norrby, *Nat. Rev. Chem*. **2021**, *5*, 240, https://doi.org/10.1038/s41570-021-00260-x.

[73] M. Orlandi, F. D. Toste, M. S. Sigman, *Angew. Chem. Int. Ed*. **2017**, *56*, 14080, https://doi.org/10.1002/anie.201707644.

[74] A. Tomberg, M. J. Johansson, P.-O. Norrby, *J. Org. Chem*. **2019**, *84*, 4695, https://doi.org/10.1021/acs.joc.8b02270.

[75] K. Jorner, T. Brinck, P.-O. Norrby, D. Buttar, *Chem. Sci*. **2021**, *12*, 1163, https://doi.org/10.1039/D0SC04896H.

[76] S. Luo, L. Liu, C.-J. Lyu, B. Sim, Y. Liu, H. Gong, Y. Nie, Y.-L. Zhao, *Cell Rep. Phys. Sci*. **2022**, 101128, https://doi.org/10.1016/j.xcrp.2022.101128.

[77] M. Moskal, W. Beker, S. Szymkuć, B. A. Grzybowski, *Angew. Chem. Int. Ed*. **2021**, *60*, 15230, https://doi.org/10.1002/anie.202101986.

[78] E. Caldeweyher, M. Elkin, G. Gheibi, M. Johansson, C. Sköld, P.-O. Norrby, J. Hartwig, **2022**, https://doi.org/10.26434/chemrxiv-2022-7qw68.

[79] A. M. Żurański, J. Y. Wang, B. J. Shields, A. G. Doyle, *React. Chem. Eng*. **2022**, *7*, 1276, https://doi.org/10.1039/D2RE00030J.

[80] T. Gensch, G. dos Passos Gomes, P. Friederich, E. Peters, T. Gaudin, R. Pollice, K. Jorner, A. Nigam, M. Lindner-D'Addario, M. S. Sigman, A. Aspuru-Guzik, *J. Am. Chem. Soc*. **2022**, *144*, 1205, https://doi.org/10.1021/jacs.1c09718.

[81] Y. Guan, C. W. Coley, H. Wu, D. Ranasinghe, E. Heid, T. J. Struble, L. Pattanaik, W. H. Green, K. F. Jensen, *Chem. Sci*. **2021**, *12*, 2198, https://doi.org/10.1039/D0SC04823B.

[82] S. V. Shree Sowndarya, J. N. Law, C. E. Tripp, D. Duplyakin, E. Skordilis, D. Biagioni, R. S. Paton, P. C. St. John, *Nat. Mach. Intell*. **2022**, *4*, 720, https://doi.org/10.1038/s42256-022-00506-3.

[83] O. A. von Lilienfeld, *Angew. Chem. Int. Ed*. **2018**, *57*, 4164, https://doi.org/10.1002/anie.201709686.

[84] B. Huang, O. A. von Lilienfeld, *Chem. Rev*. **2021**, *121*, 10001, https://doi.org/10.1021/acs.chemrev.0c01303.

[85] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, M. Ceriotti, *Chem. Rev.* **2021**, *121*, 9759, https://doi.org/10.1021/acs.chemrev.1c00021.

[86] S. Gallarati, R. Fabregat, R. Laplaza, S. Bhattacharjee, M. D. Wodrich, C. Corminboeuf, *Chem. Sci*. **2021**, *12*, 6879, https://doi.org/10.1039/D1SC00482D.

[87] B. Hammer, J. K. Nørskov, *Nature* **1995**, *376*, 238, https://doi.org/10.1038/376238a0.

[88] J. K. Nørskov, T. Bligaard, J. Rossmeisl, C. H. Christensen, *Nat. Chem*. **2009**, *1*, 37, https://doi.org/10.1038/nchem.121.

[89] Z.-J. Zhao, S. Liu, S. Zha, D. Cheng, F. Studt, G. Henkelman, J. Gong, *Nat. Rev. Mater*. **2019**, *4*, 792, https://doi.org/10.1038/s41578-019-0152-x.

[90] P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, C. A. Hunter, C. Bekas, A. A. Lee, *ACS Cent. Sci*. **2019**, *5*, 1572, https://doi.org/10.1021/acscentsci.9b00576.

[91] M. M. Bronstein, J. Bruna, T. Cohen, P. Veličković, **2021**, https://doi.org/10.48550/arXiv.2104.13478.

[92] K. Atz, F. Grisoni, G. Schneider, *Nat. Mach. Intell*. **2021**, *3*, 1023, https://doi.org/10.1038/s42256-021-00418-8.

[93] T. Xie, J. C. Grossman, *Phys. Rev. Lett*. **2018**, *120*, 145301, https://doi.org/10.1103/PhysRevLett.120.145301.

[94] K. Yan, Y. Liu, Y. Lin, S. Ji, **2022**, https://doi.org/10.48550/arXiv.2209.11807.

[95] D. W. Elrod, G. M. Maggiora, R. G. Trenary, *J. Chem. Inf. Comput. Sci*. **1990**, *30*, 477, https://doi.org/10.1021/ci00068a020.

[96] D. W. Elrod, G. M. Maggiora, R. G. Trenary, *Tetrahedron Comput. Methodol*. **1990**, *3*, 163, https://doi.org/10.1016/0898-5529(90)90050-I.

[97] T. J. Struble, C. W. Coley, K. F. Jensen, *React. Chem. Eng*. **2020**, *5*, 896, https://doi.org/10.1039/D0RE00071J.

[98] S. Fujita, *J. Chem. Inf. Comput. Sci*. **1986**, *26*, 205, https://doi.org/10.1021/ci00052a009.

[99] A. Varnek, D. Fourches, F. Hoonakker, V. P. Solov'ev, *J. Comput. Aided Mol. Des*. **2005**, *19*, 693, https://doi.org/10.1007/s10822-005-9008-0.

[100] E. Heid, W. H. Green, *J. Chem. Inf. Model*. **2022**, *62*, 2101, https://doi.org/10.1021/acs.jcim.1c00975.

[101] P. Schwaller, A. C. Vaucher, T. Laino, J.-L. Reymond, *Mach. Learn. Sci. Technol*. **2021**, *2*, 015016, https://doi.org/10.1088/2632-2153/abc81d.

[102] C. A. Grambow, L. Pattanaik, W. H. Green, *J. Phys. Chem. Lett*. **2020**, *11*, 2992, https://doi.org/10.1021/acs.jpclett.0c00500.

[103] F. E. Harrell, 'Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis', Springer, New York, **2015**.

[104] T. Stuyver, C. W. Coley, *J. Chem. Phys*. **2022**, *156*, 084104, https://doi.org/10.1063/5.0079574.

[105] R. G. Pearson, *J. Am. Chem. Soc*. **1963**, *85*, 3533, https://doi.org/10.1021/ja00905a001.

[106] H. Kneiding, R. Lukin, L. Lang, S. Reine, T. B. Pedersen, R. De Bin, D. Balcells, **2022**, https://doi.org/10.26434/chemrxiv-2022-fd43k-v2.

[107] D. Boiko, T. Reschützegger, B. Sanchez-Lengeling, S. Blau, G. Gomes, **2022**, https://doi.org/10.26434/chemrxiv- 2022-nz4pc.

[108] E. D. Glendening, C. R. Landis, F. Weinhold, *WIREs Comput. Mol. Sci*. **2012**, *2*, 1, https://doi.org/10.1002/wcms.51.

[109] D. F. Nippa, K. Atz, R. Hohler, A. T. Müller, A. Marx, C. Bartelmus, G. Wuitschik, I. Marzuoli, V. Jost, J. Wolfard, M. Binder, A. F. Stepan, D. B. Konrad, U. Grether, R. E. Martin, G. Schneider, **2022**, https://doi.org/10.26434/chemrxiv-2022-gkxm6.

[110] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, J. M. Siskind, *J. Mach. Learn. Res*. **2018**, *18*, 1.

[111] J. Behler, M. Parrinello, *Phys. Rev. Lett*. **2007**, *98*, 146401, https://doi.org/10.1103/PhysRevLett.98.146401.

[112] E. Kocer, T. W. Ko, J. Behler, *Annu. Rev. Phys. Chem*. **2022**, *73*, 163, https://doi.org/10.1146/annurev- physchem-082720-034254.

[113] J. S. Smith, O. Isayev, A. E. Roitberg, *Chem. Sci*. **2017**, *8*, 3192, https://doi.org/10.1039/C6SC05720A.

[114] C. Devereux, J. S. Smith, K. K. Davis, K. Barros, R. Zubatyuk, O. Isayev, A. E. Roitberg, *J. Chem. Theory Comput*. **2020**, *16*, 4192, https://doi.org/10.1021/acs.jctc.0c00121.

[115] J. Stevenson, L. D. Jacobson, Y. Zhao, C. Wu, J. Maple, K. Leswing, E. Harder, R. Abel, **2019**, https://doi.org/10.26434/chemrxiv.11319860.v1.

[116] J. Gasteiger, F. Becker, S. Günnemann, in 'Conference on Neural Information Processing Systems (NeurIPS)', **2021**.

[117] L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W. Hu, A. Palizhati, A. Sriram, B. Wood, J. Yoon, D. Parikh, C. L. Zitnick, Z. Ulissi, *ACS Catal*. **2021**, *11*, 6059, https://doi.org/10.1021/acscatal.0c04525.

[118] R. Tran, J. Lan, M. Shuaibi, S. Goyal, B. M. Wood, A. Das, J. Heras-Domingo, A. Kolluru, A. Rizvi, N. Shoghi, A. Sriram, Z. Ulissi, C. L. Zitnick, **2022**, https://doi.org/10.48550/arXiv.2206.08917.

[119] J. Pérez-Ramírez, N. López, *Nat. Catal*. **2019**, *2*, 971, https://doi.org/10.1038/s41929-019-0376-6.

[120] T. A. Young, T. Johnston-Wood, H. Zhang, F. Duarte, *Phys. Chem. Chem. Phys*. **2022**, *24*, 20820, https://doi.org/10.1039/D2CP02978B.

[121] S. Zhang, M. Makoś, R. Jadrich, E. Kraka, K. Barros, B. Nebgen, S. Tretiak, O. Isayev, N. Lubbers, R. Messerly, J. Smith, **2022**, https://doi.org/10.26434/chemrxiv-2022-15ct6.

[122] M. Schreiner, A. Bhowmik, T. Vegge, J. Busk, O. Winther, **2022**, https://doi.org/10.48550/arXiv.2207.12858.

[123] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput*. **2017**, *13*, 1989, https://doi.org/10.1021/acs.jctc.7b00118.

[124] S. S. Schoenholz, E. D. Cubuk, in 'Advances in Neural Information Processing Systems', 33, Curran Associates, Inc., **2020**.

[125] T. Tamayo-Mendoza, C. Kreisbeck, R. Lindh, A. Aspuru-Guzik, *ACS Cent. Sci*. **2018**, *4*, 559, https://doi.org/10.1021/acscentsci.7b00586.

[126] A. S. Abbott, B. Z. Abbott, J. M. Turney, H. F. Schaefer, *J. Phys. Chem. Lett*. **2021**, *12*, 3232, https://doi.org/10.1021/acs.jpclett.1c00607.

[127] M. F. Kasim, S. Lehtola, S. M. Vinko, *J. Chem. Phys*. **2022**, *156*, 084801, https://doi.org/10.1063/5.0076202.

[128] A. L. Dewyer, A. J. Argüelles, P. M. Zimmerman, *Wiley Interdiscip. Rev. Comput. Mol. Sci*. **2018**, *8*, e1354, https://doi.org/10.1002/wcms.1354.

[129] Z. Qiao, M. Welborn, A. Anandkumar, F. R. Manby, T. F. Miller, *J. Chem. Phys*. **2020**, *153*, 124111, https://doi.org/10.1063/5.0021955.

[130] Z. Qiao, A. S. Christensen, M. Welborn, F. R. Manby, A. Anandkumar, T. F. Miller, *Proc. Natl. Acad. Sci. U.S.A*. **2022**, *119*, e2205221119, https://doi.org/10.1073/pnas.2205221119.

[131] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *WIREs Comput. Mol. Sci*. **2021**, *11*, https://doi.org/10.1002/wcms.1493.

[132] T. Zubatiuk, B. Nebgen, N. Lubbers, J. S. Smith, R. Zubatyuk, G. Zhou, C. Koh, K. Barros, O. Isayev, S. Tretiak, *J. Chem. Phys*. **2021**, *154*, 244108, https://doi.org/10.1063/5.0052857.

[133] G. Zhou, N. Lubbers, K. Barros, S. Tretiak, B. Nebgen, *Proc. Natl. Acad. Sci. U.S.A*. **2022**, *119*, e2120333119, https://doi.org/10.1073/pnas.2120333119.

[134] F. Hu, F. He, D. J. Yaron, **2022**, https://doi.org/10.48550/arXiv.2210.11682.

[135] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel- Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, A. C. T. van Duin, *npj Comput. Mater*. **2016**, *2*, 15011, https://doi.org/10.1038/npjcompumats.2015.11.

[136] M. C. Kaymak, A. Rahnamoun, K. A. O'Hearn, A. C. T. van Duin, K. M. Merz, H. M. Aktulga, *J. Chem. Theory Comput*. **2022**, *18*, 5181, https://doi.org/10.1021/acs.jctc.2c00363.

[137] N. L. Allinger, F. Li, L. Yan, J. C. Tai, *J. Comput. Chem*. **1990**, *11*, 868, https://doi.org/10.1002/jcc.540110709.

[138] S. Spicher, S. Grimme, *Angew. Chem. Int. Ed.* **2020**, *59*, 15665, https://doi.org/10.1002/anie.202004239.

[139] H. Gao, J. Wang, J. Sun, *J. Chem. Phys*. **2019**, *150*, 244110, https://doi.org/10.1063/1.5097293.

[140] R. A. Vargas-Hernández, K. Jorner, R. Pollice, A. Aspuru-Guzik, **2022**, https://doi.org/10.48550/arXiv.2211.16763.

[141] W. Moses, V. Churavy, in 'Advances in Neural Information Processing Systems', 33, Eds. H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, H. Lin, Curran Associates, Inc., **2020**, 12472

License and Terms