IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

AI Based Modelling of Chemical Plant

Mr. B. L. Pangarkar[1], Pratik Kale[2], Aniket Kadu[3], Ravindra Agale[4] Vaibhav Tathe[5]
Assistant Professor, Department of Chemical Engineering[1]
UG Students, Department of Chemical Engineering[2-5]
Pravara Rural Engineering College, Loni, Ahmednagar

Abstract: *This project focuses on the design and implementation of an AI-based modeling system for a distillery plant, aiming to enhance process efficiency, optimize resource utilization, and ensure consistent product quality. Traditional distillery operations often face challenges such as fluctuating feedstock quality, energy inefficiencies, and manual control limitations. To address these issues, this project integrates Artificial Intelligence (AI) techniques—including machine learning, predictive analytics, and process optimization—into the core operational workflow of the distillery.*

The AI model is trained using historical and real-time process data from key distillery units such as fermentation, distillation, and dehydration. It enables predictive control over critical parameters such as temperature, pH, flow rates, and alcohol concentration. Advanced algorithms like Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) are employed for process modeling, anomaly detection, and yield prediction. The system also incorporates a user-friendly interface for real-time monitoring and decision-making support.

By deploying this AI-based solution, the distillery can achieve significant improvements in operational reliability, product consistency, energy efficiency, and sustainability. This project demonstrates the transformative potential of AI in modernizing traditional industrial processes and sets a foundation for smart manufacturing in the beverage and biofuel industries..

Keywords: *Artificial Intelligence*

I. INTRODUCTION

The distillation process is at the heart of many industries, particularly in the production of alcoholic beverages and biofuels like ethanol. A distillery plant involves several interconnected processes such as fermentation, distillation, and dehydration, which are highly sensitive to variations in raw material quality, environmental conditions, and operational parameters. Traditionally, the control and optimization of these processes have relied on manual monitoring or basic automation systems, which often lead to inefficiencies, inconsistent product quality, and increased operational costs.

With the rapid advancement in Artificial Intelligence (AI) and machine learning technologies, there is a significant opportunity to revolutionize distillery plant operations. AI-based modeling offers powerful tools for process prediction, real-time optimization, fault detection, and decision support. By leveraging large volumes of process data, AI algorithms can learn complex patterns, forecast outcomes, and provide actionable insights that improve plant performance and reduce human error.

This project aims to develop and implement an AI-based modeling system tailored for a distillery plant. The system is designed to monitor key process variables, predict potential issues, and optimize operations in real-time. By integrating AI into the distillation workflow, the plant can achieve greater operational efficiency, enhanced product consistency, and better compliance with quality and environmental standards.

The scope of this project includes data collection and preprocessing, model development using algorithms such as Artificial Neural Networks (ANN) and Regression Analysis, validation of the models, and deployment in a simulated or real plant environment. The ultimate goal is to demonstrate how AI can be a transformative force in the distillery industry, paving the way for smart and sustainable manufacturing practices the description of engineering relevant data for P&IDs are combined (e.g. ISO 15926 (International Organization for Standardization, 2013), ISO 10628 (International Organization for Standardization, 2012a), IEC 62424 (International Electrotechnical Commission, 2016

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

), ISO 10209 (International Organization for Standardization, 2012b). In particular, these include plant breakout structures, instrumentation, properties of equipment and components, and piping topology. The DEXPI informa- tion model is already offered by some manufacturers and is exchange- able via a Proteus XML schema (Proteus XML, 2017). At the same time, DEXPI provides the possibility to be used as a platform for dig- ital plant data in process industry (Wiedau et al., 2019), which can significantly reduce the development time of chemical and biotechno- logical production plants. Additionally, interoperability increases due to the

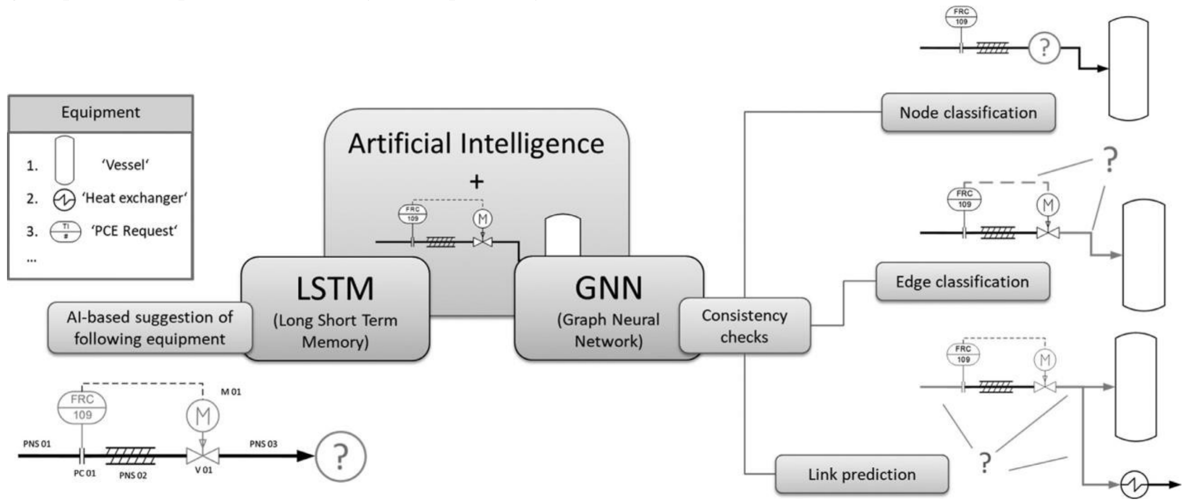

Fig. 1. Use cases of artificial intelligence to accelerate and improve the synthesis of P&IDs

continuous integration of DEXPI into existing engineering soft- ware (Fillinger et al., 2017). The uniform and machine-readable format as well as the increasing acceptance of the DEXPI format in the process industry improve the potential for the application in the field of data sci- ence and allow the application of artificial intelligence (Wiedau et al., 2021).

I- based processing of P&IDs

AI-assisted P&ID synthesis and their respective modeling approaches. In the first use case, a node prediction generates suggestions about subsequent compo- nents based on a recurrent neural network (RNN). These suggestions can support the user and decrease the time of the drawing process. The sec- ond approach uses graph neural networks (GNN), which are neural net- works especially developed for the modeling of graphs. They can

learn the topologies of process plants, which are stored in the form of a graph, and enable a consistency check during drawing by comparing the mod- els with drawn P&IDs. The use of a neural network offers the advantage that patterns and rules for generating P&IDs can be learned from exist- ing plant topologies. Therefore, no explicit heuristics need to be stored. Furthermore, the models can be adapted to the user's requirements and preferences by re-training them with data from the user. GNNs can be used to perform various classifications and predictions, which then can be used for consistency checking. The node classification allows for pre- dicting information, such as the equipment classes of individual nodes to check whether components within a P&ID are present at meaning- ful positions. In contrast, edge classification focuses on the prediction of edge information. In relation to the P&ID, these are for example the connection types (piping or signal lines) or the information, whether a pipeline is insulated or heated. In parallel, a link prediction can be used, which provides information about the probability of a possible link between two components. This enables the validation of connec- tions within a P&ID as well as the

suggestion of connections during the drawing process. In the following, the preprocessing for converting the P&IDs into graphs as well as the two modeling concepts is introduced in more detail. The results for a node prediction via RNNs as well as a node classification based on a GNN is presented.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

II. FUTURE SCOPE AND BENEFITS

The implementation of AI in distillery plant operations is just the beginning of a broader transformation in the process industry. As the technology matures and more real-time data becomes available, several enhancements can be made:

1. Real-Time AI Control Systems Integration with live sensor data and PLCs/SCADA systems can enable real- time process control and dynamic optimization without human intervention.
2. Integration with IoT (Internet of Things) Using IoT-enabled smart sensors can improve data collection, remote monitoring, and maintenance prediction.
3. Cloud-Based Data Analytics Leveraging cloud platforms can allow for centralized data processing, remote access, and advanced analytics at sca
4. Digital Twin TechnologyA digital replica of the entire distillery plant can be created for simulations, predictive maintenance, and scenario analysis.
5. Sustainability OptimizationFuture models can focus on reducing energy consumption, emissions, and waste through AI-driven environmental impact assessments.
6. Cross-Industry ApplicationsThe modeling framework developed can be adapted to other process industries such as breweries, pharmaceuticals, and chemical plants.

The adoption of AI-based modeling in a distillery plant brings numerous advantages, both operational and strategic:

1. Improved Process EfficiencyAI can optimize process parameters to reduce cycle time, minimize energy use, and enhance output yield.
2. Consistent Product Quality Predictive models ensure that product specifications (like alcohol content and purity) remain within desired limits despite feedstock variations.
3. Early Fault Detection and Maintenance Machine learning can detect equipment anomalies before they lead to failures, reducing downtime and maintenance costs.
4. Cost Reduction By minimizing waste, energy usage, and manual intervention, overall operational costs can be significantly reduced.
5. Enhanced Decision-Making Data-driven insights empower plant operators and management to make better, faster, and more informed decisions.
6. Scalability and Flexibility The AI system can be scaled for larger operations or adapted for different types of distillation setups (e.g., ethanol, spirits, essential oils).

Results - GNN node classification

To quantify how well GNNs are suited to classify individual pieces of equipment based on their location in the P&ID topology, the recursive GNN presented in chapter 3 is trained with the P&ID graphs presented

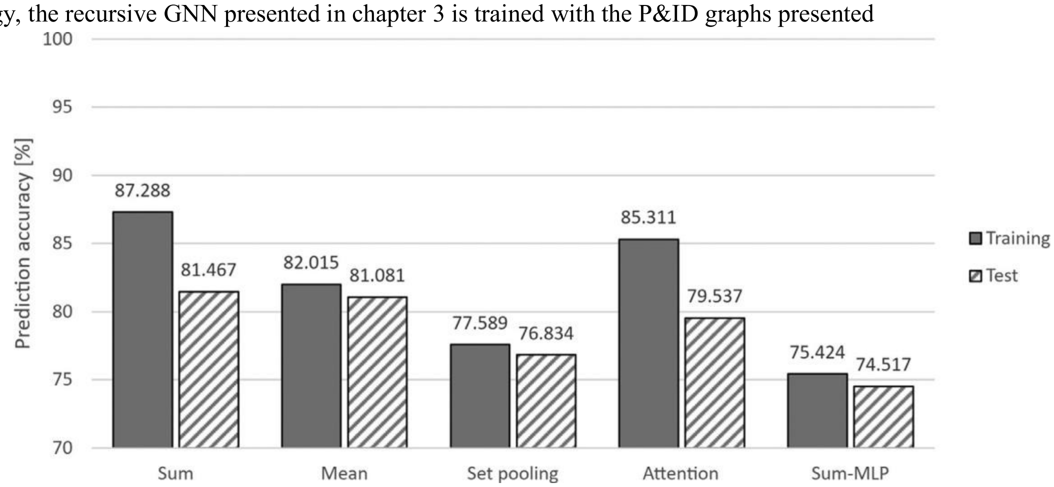

Fig. 9. Results of the node classification in a P&ID graph via recursive GNN grouped by the applied aggregation functions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

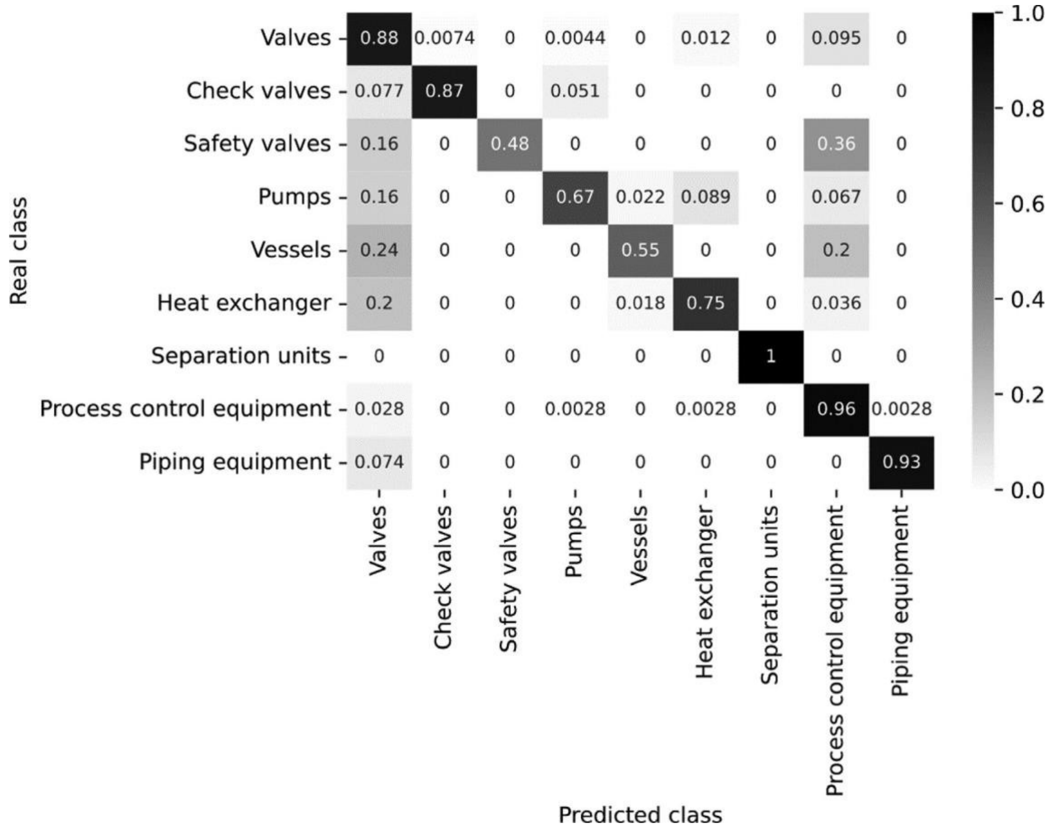

Fig. 10. Normalized confusion matrix of the recursive GNN with sum aggregation for the test data set.

before. For this purpose, the models were implemented and trained using the Deep Graph Library (version 0.7.2) with PyTorch frame- work in Python (version 3.7) . The calculations were performed on an Intel® Xeon® W-2155 (3.31 GHz) CPU in combination with 128 GB RAM. The aggregation functions listed previously are used in the GNN and compared against each other. The prediction accuracy is used as an evaluation metric, which was defined in chapter 3.2. For the examined models, the accuracy is shown below in Fig. 9 .

The results for all models show deviations among each other. The ac- curacy for the training for all models is between 75.4% (sum-MLP) and 87.3% (sum) while the accuracy for the test data varies between 74.5% (sum-MLP) and 81.5% (sum). It is striking that the gap between the test and training accuracies for the sum aggregation as well as the atten- tion aggregation is larger than for the remaining aggregation functions, at about 6 percentage points. Additionally, the results show that sim- pler aggregation algorithms such as sum and arithmetic mean achieve higher training accuracies than the more complex aggregations using attention, set pooling or sum- MLP. It is hypothesized that this is due to the fact that all neighborhood information is equally important in predicting the component class. For this reason, learning the individual P&ID components works particularly well when the neighborhood information is aggregated with the same weight, i.e., equally important. This is especially true for the sum or mean.

To check how well the classification can be performed for the dif- ferent classes, the confusion matrix of the model with sum aggregation using the test data set is also considered, see Fig. 10 . The columns in the matrix describe the predicted classes, while the rows represent the real classes. Consequently, the main diagonal displays the number of correctly classified components (TP).

The main diagonal shows that most components of each class are cor- rectly classified. This way, all separation units are correctly classified. Process control equipment (PCE, 96%) and piping components (93%) are also almost completely correctly assigned. With over 87% predic- tion accuracy, valves and check valves are also classified sufficiently well, although it is noticeable that there is confusion between valves and piping equipment and valves and

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, June 2025 Impact Factor: 7.67

safety valves. However, with less than 10%, this is still within tolerable limits. The classification of the remaining components is much more difficult for the GNN. Thus, 39% of safety valves are identified as piping equipment and 16% as standard valves. This is not surprising since both classes are usually found in sim- ilar positions of a P&ID. Classification of pumps (67%), vessels (55%) and heat exchangers (75%) is only reasonably satisfactory. It is notice- able that all three classes are mainly classified as valves or PCEs. Classes which are particularly strongly represented in the data set according to Table 1 . The GNN models should therefore be further optimized in the future. At this point, it is conceivable to integrate the comparatively un- derrepresented classes more strongly into the training by introducing weighting factors. Furthermore, it would be conceivable to use a larger k , which would aggregate more information. However, this results in a larger computational effort.

III. RESULTS AND DISCUSSIONS

Benefits to the Industry

a. Increased Efficiency and Reduced Time-to-Market

Faster Design and Engineering: By automating the conversion of PFDs into PIDs, AI significantly reduces the time engineers spend on manual diagram creation, leading to faster plant design and commissioning. This is particularly important in industries like distilling, where the plant design phase often faces tight deadlines.

Improved Iteration Speed: AI allows for rapid iterations of PFD-to-PID conversions, facilitating quicker modifications during the design process. This results in a more agile design phase and reduces the time needed to respond to changes in project scope or regulations.

b. Cost Reduction

Lower Engineering Costs: Reducing the amount of manual labor involved in creating PIDs from PFDs results in direct cost savings. Fewer resources are required for drawing, analyzing, and correcting PIDs, reducing labor costs and overheads in the design phase.

Reduced Errors and Rework: AI-driven models reduce human errors in converting PFDs to PIDs, leading to fewer mistakes in the final design. This results in less rework, which can be costly both in terms of time and materials.

Benefits to Society

a. Improved Workplace Safety

Reduced Human Error in Design: By automating the process and reducing human involvement in the creation of critical PIDs, AI minimizes the risks of human errors that could lead to safety issues. This is especially important in high-risk industries such as distillery plants, where improper PIDs can result in operational hazards, leaks, or system failures.

Compliance with Safety Standards: AI ensures that the generated PIDs adhere to safety standards such as ISO, IEC, and NFPA codes, ensuring that distillery plants are designed with built-in safety features (e.g., emergency shutoff valves, pressure relief systems) to mitigate risks to workers and the community.

b. Environmental Sustainability

Energy Efficiency: AI-based optimization in the design phase can lead to more energy-efficient distillery operations. By fine-tuning control systems and improving process flows, AI helps reduce energy consumption, resulting in lower carbon footprints and more sustainable production processes.

Waste Reduction: Optimizing the design for efficiency not only helps reduce energy consumption but also minimizes waste generation. This leads to fewer by-products, lower emissions, and a more sustainable production cycle, which benefits the environment.

IV. CONCLUSION

The AI-based modeling of distillery plants through the conversion of PFDs into PIDs offers a powerful tool for transforming how industrial plants are designed and operated. This approach provides significant improvements in speed, accuracy, cost-effectiveness, and sustainability, while also enhancing safety and efficiency. The scalability of the AI model across various industries opens up broader possibilities for the future of process design and optimization.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, June 2025 Impact Factor: 7.67

Ultimately, this technology is a key enabler of the next generation of smart manufacturing, digitalization, and sustainability in industrial operations. Through its implementation, the distillery industry, along with other sectors, can experience enhanced productivity, reduced environmental impact, and greater safety—creating a positive ripple effect across industries and society as a whole.

REFERENCES

1. Schwaller et al. [2018]
Philippe Schwaller, Théophile Gaudin, Dávid Lányi, Costas Bekas, and Teodoro Laino."found in translation": predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models.Chemical science, 9(28):6091–6098, 2018.doi:10.1039/c8sc02339e.
Schwaller et al. [2019]
2. Philippe Schwaller, Teodoro Laino, Théophile Gaudin, Peter Bolgar, Christopher A. Hunter, Costas Bekas, and Alpha A. Lee.Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction.ACS central science, 5(9):1572–1583, 2019.doi:10.1021/acscentsci.9b00576.
Schwaller et al. [2020]
3. Philippe Schwaller, Riccardo Petraglia, Valerio Zullo, Vishnu H. Nair, Rico Andreas Haeuselmann, Riccardo Pisoni, Costas Bekas, Anna Iuliano, and Teodoro Laino.Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy.Chemical science, 11(12):3316–3325, 2020. Towler and Sinnott [2008]
4. Gavin P. Towler and R. K. Sinnott.Chemical engineering design - Principles, practice and economics of plant and process design.Elsevier/Butterworth-Heinemann, Amsterdam and Boston, 2008.ISBN 9780750684231.
Toghraei [2019]
5. Moe Toghraei.Piping and instrumentation diagram development, volume First edition.John Wiley & Sons Inc, Hoboken, NJ, USA, 2019.ISBN 9781119329343.
Uzuner and Schembecker [2012]
6. Hülya Uzuner and Gerhard Schembecker.Wissensbasierte erstellung von r&i-fließbildern.Chemie Ingenieur Technik, 84(5):747–761, 2012.doi:10.1002/cite.201100230.
Blitz et al. [1994]
7. Helmut Blitz, Jochen Engelke, Raymund Sonnenschein, and Henner Schmidt-Traub.Rechnergestützte konfigurierung von ri- fließbildern am beispiel von pumpen.Chemie Ingenieur Technik, 66(4):470–475, 1994.doi:10.1002/cite.330660404.
Obst et al. [2013]
8. Michael Obst, Falk Doherr, and Leon Urbas.Wissensbasiertes assistenzsystem für modulares engineering.Automatisierungstechnik, 61(2):103–108, 2013.doi:10.1524/auto.2013.0011.
Fleischer-Trebes et al. [2017]
9. Christoph Fleischer-Trebes, Nicolai Krasberg, Christian Bramsiepe, and Norbert Kockmann.Planungsansatz für modulare anlagen in der chemischen industrie.Chemie Ingenieur Technik, 89(6):785–799, 2017.doi:10.1002/cite.201600083.
Hohmann et al. [2017]
10. Lukas Hohmann, Katrin Kössl, Norbert Kockmann, Gerhard Schembecker, and Christian Bramsiepe.Modules in process industry
11. a life cycle definition.Chemical Engineering and Processing: Process Intensification, 111:115–126, 2017.doi:10.1016/j.cep.2016.09.017.
Eilermann et al. [2018]
12. Martin Eilermann, Christian Post, Heiko Radatz, Christian Bramsiepe, and Gerhard Schembecker.A general approach to module-based plant design.Chemical Engineering Research and Design, 137:125–140, 2018.doi:10.1016/j.cherd.2018.06.039.
Venkatasubramanian [2019]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, June 2025
Impact Factor: 7.67

13. Venkat Venkatasubramanian.The promise of artificial intelligence in chemical engineering: Is it here, finally?AIChE Journal, 65(2):466–478, 2019.doi:10.1002/aic.16489.
Brown et al. [2020]
14. Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei.Language models are few- shot learners.arXiv, pages 1–75, 2020.doi:10.48550/arXiv.2005.14165.
Raffel et al. [2019]
15. Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J.
Liu.Exploring the limits of transfer learning with a unified text-to-text transformer.Journal of Machine Learning Research, 21(140):1– 67, 2019.doi:10.48550/arXiv.1910.10683.
Vaswani et al. [2017]
16. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin.Attention is all you need.arXiv, pages 1–15, 2017.doi:10.48550/arXiv.1706.03762.
Popel et al. [2020]
17. Martin Popel, Marketa Tomkova, Jakub Tomek, Łukasz Kaiser, Jakob Uszkoreit, Ondřej Bojar, and Zdeněk Žabokrtský.Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals.Nature
communications, 11(4381):1–15, 2020.doi:10.1038/s41467-020-18073-9.
Weininger [1988]
18. David Weininger.Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules.Journal of Chemical Information and Modeling, 28(1):31–36, 1988.doi:10.1021/ci00057a005.
Weininger et al. [1989]
19. David Weininger, Arthur Weininger, and Joseph L. Weininger.Smiles. 2. algorithm for generation of unique smiles notation.Journal of Chemical Information and Computer Sciences, 29(2):97–101, 1989.doi:10.1021/ci00062a008.
Schwaller et al. [2018]
20. Philippe Schwaller, Théophile Gaudin, Dávid Lányi, Costas Bekas, and Teodoro Laino."found in translation": predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models.Chemical science, 9(28):6091–6098, 2018.doi:10.1039/c8sc02339e.
Schwaller et al. [2019]