JPPIPA 12(8) (2026)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

The Role of Artificial Intelligence in Formulating and Balancing Chemical Equations for Organic Compounds

Nasratullah Mahboob[1*], Tamann Quraishi[2]

[1] Department of Chemistry, Education Faculty, Samangan University, Samangan, Afghanistan
[2] Faculty of Computer Science, University of the People, United State of America (USA).

Received: June 21, 2026
Revised: August 20, 2026
Accepted: August 29, 2026
Published: August 31, 2026

Corresponding Author:
Nasratullah Mahboob
nmahboob091@gmail.com

DOI: 10.29303/jppipa.v12i8.16269

Abstract: Balancing chemical equations for organic compounds is a routine yet often tedious component of chemistry education and computational chemistry, and it becomes markedly harder as molecular size and functional-group complexity grow. This paper reviews how artificial intelligence (AI) techniques, ranging from classical algebraic and matrix-based solvers to modern machine-learning models such as graph neural networks and sequence-to-sequence transformers, are being used to formulate, balance, and validate chemical equations for organic reactions. We synthesise findings from the computational chemistry, cheminformatics, and chemistry-education literature to compare rule-based, algebraic, and learned approaches, and we illustrate the underlying logic with a worked example of ethanol combustion solved through the matrix null-space method. The review indicates that deterministic algebraic methods remain the most reliable choice for routine stoichiometric balancing, while learned models add clear value for predicting reaction outcomes and retrosynthetic routes in more complex organic transformations, albeit with lower top-1 accuracy and weaker interpretability. We further discuss the growing use of generative AI tools in chemistry classrooms, including in low-resource settings such as Afghan universities, and conclude that a hybrid architecture combining deterministic solvers with learned chemical-plausibility models is the most promising direction for AI-assisted equation balancing.

Keywords: Artificial intelligence; Chemical equation balancing; Chemistry education; Machine learning; Organic chemistry.

Introduction

Balancing a chemical equation is one of the first quantitative skills that chemistry students must master, and it remains a practical necessity for every chemist who needs to predict yields, plan a synthesis, or verify that a proposed reaction is even physically possible. The rule that governs this task is simple to state: the number of atoms of every element must be the same on both sides of the equation, because mass can neither be created nor destroyed. In practice, however, applying that rule to organic compounds is far from simple. Organic molecules routinely contain dozens of carbon, hydrogen, oxygen, nitrogen, and heteroatom centres, and a single combustion, esterification, or substitution reaction can involve several polyatomic species at once. Balancing such equations by inspection, the method most students are taught first, quickly becomes unreliable once the molecule grows beyond a handful of atoms (Blakley, 1982; Kumar, 2001).

Artificial intelligence has begun to change this picture in two related ways. On one hand, classical computer-assisted methods, most notably the representation of a chemical equation as a system of linear equations and the search for the null space of its coefficient matrix, turn balancing into a deterministic computation rather than a guessing game (Thorne, 2010; Zabadi & Assaf, 2017; Zhang et al., 2024). On the other hand, contemporary machine-learning systems, including rule-mining algorithms (Mohialden et al., 2023) and large neural architectures trained on millions of literature reactions (Coley et al., 2017; Schwaller et al.,

How to Cite:
E Mahboob, N., & Quraishi, T. (2026). The Role of Artificial Intelligence in Formulating and Balancing Chemical Equations for Organic Compounds. *Jurnal Penelitian Pendidikan IPA*, *12*(8), 252–258. https://doi.org/10.29303/jppipa.v12i8.16269

2019), no longer merely balance a given equation; they can propose what the products of a reaction should be in the first place, or work backwards from a target organic molecule to plausible starting materials. These two threads, exact algebraic balancing and learned chemical reasoning, are increasingly combined in practice, and understanding how they relate to each other is important for anyone building or teaching with AI-based chemistry tools.

The motivation for this review is also pedagogical. Generative AI systems such as ChatGPT are now used widely by chemistry students and instructors (Kasneci et al., 2023; Rasul et al., 2023), and they are already being tested as tutoring aids in universities that face severe resource constraints, including Afghan higher-education institutions (Fazil et al., 2024). Whether such tools can be trusted to balance an organic equation correctly, or whether they should instead be paired with a deterministic solver behind the scenes, is a question with direct classroom relevance. This paper therefore has three aims: first, to summarise the main families of AI and computational methods used to formulate and balance organic chemical equations; second, to illustrate the mechanics of the most transparent of these methods, the matrix null-space approach, with a worked example; and third, to discuss what the chemistry-education literature tells us about the opportunities and risks of deploying these tools with students, including in low-resource settings.

Method

This paper follows a narrative review design combined with an illustrative computational example, an approach well suited to a topic that spans mathematics, cheminformatics, and education research. Literature was identified through Google Scholar, ACS Publications, Nature Portfolio journals, ScienceDirect, and arXiv, using combinations of the terms "artificial intelligence", "machine learning", "chemical equation balancing", "organic reaction prediction", "retrosynthesis", and "chemistry education". In keeping with common practice for this journal, priority was given to sources published within roughly the last ten years, while a small number of older, foundational papers on algebraic balancing were retained because they remain the methodological basis for later work. Sources were included if they described a computational or AI-based method for equation balancing, reaction-outcome prediction, or retrosynthesis, or if they empirically examined the use of AI tools in chemistry or STEM teaching.

The retrieved literature was organised into three method families that recur throughout the results: (i) rule-based and algebraic solvers, which treat balancing as a linear-algebra problem; (ii) machine-learning models for reaction-outcome and retrosynthesis prediction, which learn chemical transformations from large reaction corpora; and (iii) generative-AI tools used directly by students and instructors in the classroom. For the worked example, an unbalanced combustion reaction of ethanol was represented as an element-composition matrix, with one row per element and one column per chemical species, following the null-space formulation described by Thorne (2010) and Zabadi and Assaf (2017). The matrix was solved for its null space and the resulting solution vector was scaled to the smallest set of positive integers, reproducing the procedure that a Gauss-Jordan-based or NumPy-based AI balancing tool would perform internally. This example is used purely to make the abstract algebra concrete; it is not itself an experimental validation of any single software package.

Result and Discussion

Algebraic and matrix-based balancing of organic equations

The most transparent and, for simple stoichiometry, the most trustworthy computational method for balancing an organic equation is the algebraic or matrix approach. Every chemical species in a reaction can be written as a vector of its elemental composition, and the requirement that atoms balance becomes a homogeneous system of linear equations, $A \cdot x = 0$, where A is the element-composition matrix and x is the vector of unknown coefficients (Blakley, 1982; Krishna et al., 2016). Balancing then reduces to finding a basis vector of the null space of A and scaling it to the smallest integers, a computation that Gauss-Jordan elimination or a numerical linear-algebra library performs in a fraction of a second (Kumar, 2001; Zabadi & Assaf, 2017). Thorne (2010) showed that this null-space search can even be carried out with the matrix-inversion routine already built into an ordinary scientific calculator, which is part of why the method has become the standard "AI-adjacent" technique taught alongside manual balancing.

For most organic combustion, substitution, and esterification reactions, this method returns a single, unique, all-positive solution. Its main limitation appears with larger reaction networks or systems containing multiple independent reaction pathways, where the null space has a dimension greater than one and the direct output of Gaussian elimination may include invalid negative coefficients or fail to correspond to a chemically meaningful basis (Zhang et al., 2024). Zhang and colleagues (2024) address this by reformulating the balancing problem in terms of a positive affine monoid rather than a linear vector space, using the Hilbert basis

of that monoid to guarantee a canonical set of valid, non-negative elementary reactions. This distinction matters for organic chemistry specifically, because branched or cyclic reaction networks, such as those found in combustion mechanisms or metabolic pathways, are exactly the cases where a naive null-space calculation is most likely to misbehave.

Table 1. Comparison of computational and AI-based approaches to formulating and balancing organic chemical equations

Method	Underlying principle	Typical reliability	Key strength	Key limitation
Manual trial-and-error	Iterative adjustment of coefficients by inspection	High for short equations, error-prone for long ones	Requires no software	Slow and skill-dependent (Blakley, 1982)
Algebraic / matrix null-space	Solve a homogeneous linear system built from element counts	Deterministic, always correct if a unique solution exists	Exact, generalises to any equation (Kumar, 2001; Thorne, 2010)	Non-unique or negative solutions for degenerate systems (Zhang et al., 2024)
Rule-mining (e.g., Apriori)	Learns frequent balancing patterns from a corpus of solved reactions	High on equations similar to training examples	Fast on familiar reaction families (Mohialden et al., 2023)	Weak generalisation to novel organic scaffolds
ML reaction-outcome prediction	Sequence-to-sequence or graph models trained on large reaction databases	About 90% top-1 on USPTO patent reactions	Captures selectivity and real reaction context (Schwaller et al., 2019)	Black-box; sensitive to dataset bias (Coley et al., 2017)
Graph-based retrosynthesis	Graph neural networks predict bond/atom edits to decompose a target molecule	50-70% top-1 depending on model and benchmark	Handles complex, multi-step organic synthesis (Zhong et al., 2023)	Computationally heavy; limited interpretability (Somnath et al., 2021)

Rule-mining and hybrid AI approaches

Between purely manual balancing and full machine-learning reaction prediction sits a family of methods that mine patterns from previously solved equations rather than deriving coefficients from first principles. Mohialden et al. (2023) demonstrated this with an Apriori-based algorithm that treats a corpus of already-balanced reactions as a transaction database, extracts frequent itemsets of coefficient patterns, and uses them to reconstruct the coefficients of a new, unbalanced equation. This kind of approach can be fast and effective when the target equation is structurally similar to reactions already seen by the system, which makes it attractive for classroom software that only needs to cover a defined syllabus of reaction types. Its weakness is generalisation: an Apriori-style balancer offers no guarantee of correctness on a genuinely novel organic scaffold, unlike the deterministic matrix method, so in practice the two approaches are complementary rather than substitutes.

Machine learning for organic reaction outcome prediction

A different and, in many ways, more ambitious use of AI in organic chemistry is not balancing a given equation but predicting what the products of a reaction will be in the first place, essentially formulating the equation from scratch. Coley et al. (2017) were among the first to show that a neural model, in their case built on the Weisfeiler-Lehman graph representation of molecules, could learn to predict the outcome of an organic reaction directly from the structures of the reactants, outperforming earlier template-based expert systems. Fooshee et al. (2018) extended this line of work with deep neural networks trained on annotated reaction data, further improving the practical accuracy of outcome prediction for everyday organic transformations.

The most influential development in this space has been the Molecular Transformer of Schwaller et al. (2019), which reframes reaction prediction as a machine-translation problem: the SMILES text string of the reactants is "translated" into the SMILES string of the products using the same transformer architecture originally developed for human-language translation. Trained on roughly 480,000 reactions extracted from United States patents, the Molecular Transformer reaches about 90% top-1 accuracy and, notably, also produces a calibrated confidence score for each prediction, which is valuable for flagging cases where the model is unsure. The same architecture has since been shown to unify forward reaction prediction and retrosynthesis within a single model (Schwaller et al., 2019), and to transfer well to specialised domains such as carbohydrate and enzymatic chemistry once fine-tuned on smaller, domain-specific datasets. Because these models are trained end-to-end on real literature

reactions, their output equations are balanced almost by construction, but their correctness is only as good as the training data, and later analyses have shown that patent-derived datasets carry systematic biases that a large model can inadvertently learn (Fooshee et al., 2018).

Graph-based retrosynthesis for complex organic molecules

Formulating an equation is only half the challenge in organic synthesis; chemists frequently need to work backwards from a target molecule to a set of plausible starting materials, a task known as retrosynthesis. Because retrosynthesis is fundamentally about editing a molecular graph, graph neural networks (GNNs) have become the dominant AI architecture for this task. Somnath et al. (2021) framed retrosynthesis as a two-stage graph-editing problem, first breaking the product into synthons and then completing those synthons into full reactants, an approach subsequently refined by Chen et al. (2023) with G2Retro and by Gao et al. (2022) with the SemiRetro framework. Zhong et al. (2023) unified the two stages into a single end-to-end model, Graph2Edits, inspired by the arrow-pushing notation organic chemists already use to describe reaction mechanisms, which improved both accuracy and interpretability on complex, multi-step organic reactions. Complementary template-free approaches, such as the graph-to-graph attention model of Lin et al. (2023) and the permutation-invariant graph-to-sequence model of Tu and Coley (2022), together with transformer-based systems such as RetroPrime (Wang et al., 2021), report top-1 accuracies in the range of roughly 50-70% on the widely used USPTO-50K benchmark, figures that are respectable but still well below the near-certain correctness of a deterministic matrix solver for simple stoichiometry.

Table 2. Representative performance figures reported in the literature for different computational and AI approaches

Method	Task	Benchmark dataset	Reported top-1 accuracy	Source
Matrix null-space	Coefficient balancing	Deterministic (any solvable equation)	100% (when unique solution exists)	Thorne (2010)
Apriori rule-mining	Coefficient balancing	Corpus of solved reactions	~96%	Mohialden et al. (2023)
Molecular Transformer	Reaction outcome prediction	USPTO (~480,000 reactions)	~90%	Schwaller et al. (2019)
G2GT (graph-to-graph)	Single-step retrosynthesis	USPTO-50K	54%	Lin et al. (2023)

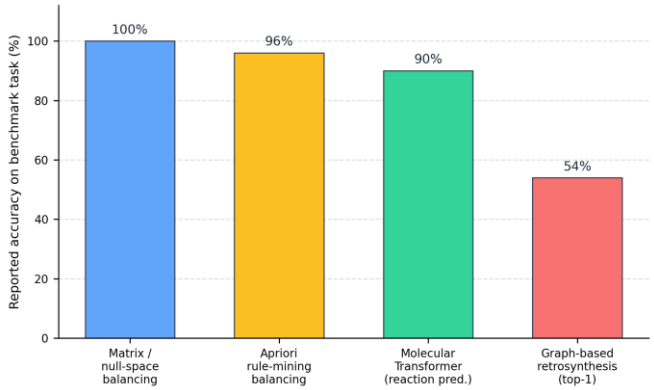

Figure 1. Illustrative comparison of reported top-1 accuracy across benchmark tasks for representative computational and AI methods discussed in this review

A worked example: an AI-assisted balancing pipeline

To make the abstract algebra concrete, Figure 2 sketches the pipeline that a typical AI-assisted balancing tool follows internally, whether it is implemented as a simple matrix solver or wrapped inside a chatbot interface: the input reaction is parsed into a machine-readable representation, an element-composition matrix is built, a solver (deterministic or learned) proposes coefficients, and those coefficients are validated and reduced to integers before being returned to the user

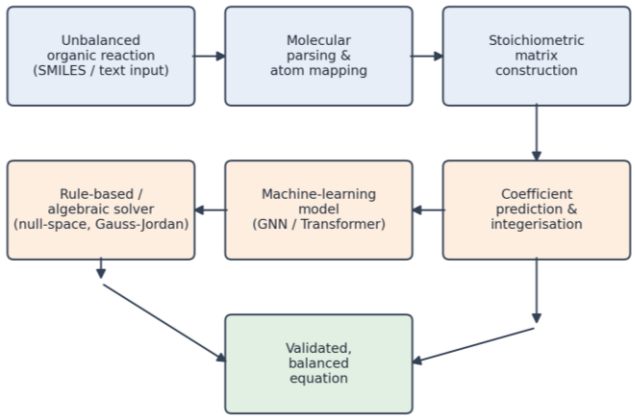

Figure 2. Conceptual pipeline of an AI-assisted system for formulating and balancing organic chemical equations, combining a deterministic algebraic solver with an optional learned model for chemical plausibility checking

Consider the unbalanced combustion of ethanol, $C_2H_6O + O_2 \rightarrow CO_2 + H_2O$. Following the null-space method, each species is written as a column vector of its

carbon, hydrogen, and oxygen content, giving the element-composition matrix in Table 3

Table 3. Element-composition (stoichiometric) matrix for the combustion of ethanol, with reactant columns given a positive sign and product columns a negative sign

Element	C2H6O (a)	O2 (b)	CO2 (c)	H2O (d)
Carbon (C)	2	0	-1	0
Hydrogen (H)	6	0	0	-2
Oxygen (O)	1	2	-2	-1

Solving $A \cdot x = 0$ for the null space and scaling the resulting vector to the smallest positive integers yields a = 1, b = 3, c = 2, d = 3, so the balanced equation is C2H6O + 3O2 → 2CO2 + 3H2O, exactly the coefficients that a student would eventually reach by trial and error, but obtained here in a single deterministic step. Figure 3 walks through this calculation visually. For an equation this small, the algebraic method and a well-trained machine-learning model will usually agree; the practical advantage of the AI-assisted pipeline becomes clearer for larger organic molecules, redox reactions with several coupled half-equations, or reaction networks with multiple simultaneous products, where manual inspection is slow and error-prone but the underlying linear algebra scales cleanly (Krishna et al., 2016; Zhang et al., 2024).

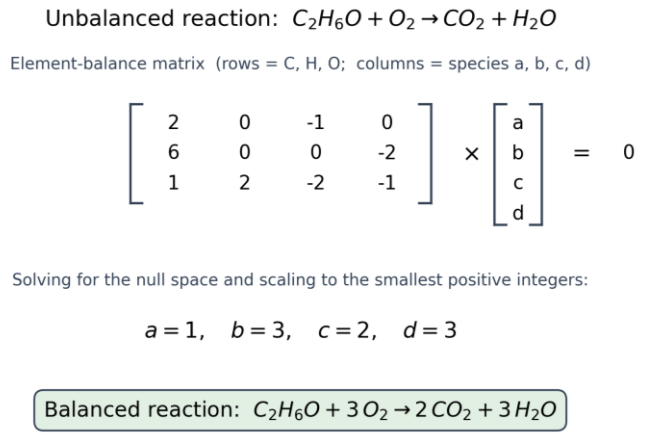

Figure 3. Worked example of the matrix null-space method applied to the combustion of ethanol, from the unbalanced reaction through the element-composition matrix to the balanced equation

Generative AI and chemistry education, including the Afghan context

Beyond dedicated balancing software, general-purpose generative AI systems such as ChatGPT are now widely used by chemistry students for homework help, concept explanation, and even equation balancing (Kasneci et al., 2023; Rasul et al., 2023). Ausat et al. (2023) and Firat (2023) both found that such tools can support self-directed learning without necessarily replacing the teacher, provided students still develop the underlying reasoning skills. At the same time, chemistry-education researchers have documented clear shortcomings: Tyson (2023) catalogued cases where ChatGPT produced chemically incorrect explanations with high apparent confidence, and Farrokhnia et al. (2024) note in their SWOT analysis that hallucinated or subtly unbalanced equations are a recurring risk precisely because large language models generate text token by token rather than verifying atom counts symbolically. Rahman and Watanobe (2023) similarly argue that generative AI is best positioned in education as a drafting and explanation aid that should be paired with a verification step, a conclusion that matches the hybrid architecture this review recommends: let a language model handle natural-language explanation, but let a deterministic algebraic solver, of the kind described above, perform and check the actual balancing.

This distinction is especially relevant in resource-constrained settings. Fazil et al. (2024) surveyed university students in Afghanistan and found that AI tools were already influencing engagement and academic performance despite considerable infrastructural constraints, echoing broader findings that AI-enabled e-learning can partly offset limited laboratory access, unreliable connectivity, and large class sizes in Afghan higher education. For a subject like organic chemistry, where practical laboratory work to observe real reactions is often limited, a reliable AI balancing and reaction-prediction tool, even a lightweight one built around the matrix method described in this paper, could meaningfully extend what students are able to practise outside the classroom, provided the tool is transparent about when it is reasoning symbolically and when it is only pattern-matching.

Challenges and limitations

Several limitations temper the promise of AI-assisted equation balancing for organic chemistry. First, purely learned models inherit the biases of their training corpora; because most large reaction datasets are extracted from patent literature, they under-represent reaction types that are common in teaching but rare in industrial patents (Coley et al., 2017; Fooshee et al., 2018). Second, deterministic algebraic solvers, while exact, can return non-unique or chemically invalid (negative-coefficient) solutions once a reaction network becomes large or contains coupled pathways, a problem only partially resolved by more recent monoid-based reformulations (Zhang et al., 2024). Third, generative language models can produce fluent but factually wrong chemistry, and their confidence is not a reliable signal of

correctness (Tyson, 2023), which makes human or symbolic verification essential before such output reaches students. Finally, most of the machine-learning literature reviewed here evaluates methods on curated organic-chemistry benchmarks such as USPTO-50K; how well these systems perform on the messier, less well-defined equations that arise in an actual classroom or research notebook remains comparatively under-studied.

Conclusion

Artificial intelligence has meaningfully changed how organic chemical equations are formulated and balanced, but not by replacing the underlying chemistry with a single black-box model. Deterministic algebraic and matrix null-space methods remain the most reliable tool for the core task of balancing a given equation, and they scale far better than manual inspection once molecules or reaction networks grow large. Machine-learning models, from rule-mining systems to graph neural networks and transformers, add genuine value on the harder, related problems of predicting reaction outcomes and planning retrosynthetic routes for complex organic molecules, though with lower accuracy and less interpretability than the algebraic approach. Generative AI chatbots, meanwhile, are already in students' hands and offer real pedagogical benefits, but the chemistry-education literature is consistent in warning that their fluency can outpace their correctness. The most promising direction, and the one implicitly adopted by the more recent hybrid systems reviewed here, is to combine a deterministic solver that guarantees a mass-balanced answer with a learned model that supplies chemical plausibility, explanation, and context. For resource-constrained institutions such as those in Afghanistan, where laboratory access and instructor time are often limited, a transparent tool of this kind could extend meaningful chemistry practice well beyond the classroom, provided it is built and taught with a clear understanding of what each component can and cannot be trusted to do.

Acknowledgments
The author thanks colleagues in the Faculty of Education at Samangan University for their comments on an earlier draft of this manuscript.

Author Contributions
Conceptualization, N.M.; methodology, N.M.; investigation, M.H.; writing-original draft preparation, N.M.; writing-review and editing, N.M.; visualization, T.Q. The author has read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
The author declares no conflict of interest.

References

Ausat, A. M. A., Massang, B., Efendi, M., Nofirman, & Riady, Y. (2023). Can ChatGPT replace the role of the teacher in the classroom: A fundamental analysis. *Journal on Education*, 5(4), 16100-16106. Retrieved from https://www.studocu.com/cl/document/universi...8004-072fc476e24a1788181670

Blakley, G. R. (1982). Chemical equation balancing: A general method which is quick, simple, and has unexpected applications. *Journal of Chemical Education,* 59(9), 728-734. https://doi.org/10.1021/ed059p728

Chen, Z., Ayinde, O. R., Fuchs, J. R., Sun, H., & Ning, X. (2023). G2Retro as a two-step graph generative model for retrosynthesis prediction. *Communications Chemistry*, 6, 102. https://doi.org/10.1038/s42004-023-00897-3

Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H., & Jensen, K. F. (2017). Prediction of organic reaction outcomes using machine learning. *ACS Central Science,* 3(5), 434-443. https://doi.org/10.1021/acscentsci.7b00064

Fazil, A. W., Hakimi, M., Shahidzay, A. K., & Hasas, A. (2024). Exploring the broad impact of AI technologies on student engagement and academic performance in university settings in Afghanistan. *RIGGS: Journal of Artificial Intelligence and Digital Business,* 2(2), 108-118. https://doi.org/10.31004/riggs.v2i2.268

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International,* 61(3), 460-474. https://doi.org/10.1080/14703297.2023.2195846

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching,* 6(1), 57-63. https://doi.org/10.37074/jalt.2023.6.1.22

Fooshee, D., Mood, A., Gutman, E., Tavakoli, M., Urban, G., Liu, F., Huynh, N., Van Vranken, D., & Baldi, P. (2018). Deep learning for chemical reaction prediction. *Molecular Systems Design & Engineering,* 3(3), 442-452. https://doi.org/10.1039/C7ME00107J

Gao, Z., Tan, C., Wu, L., & Li, S. Z. (2022). *SemiRetro: Semi-template framework boosts deep retrosynthesis prediction. arXiv preprint* arXiv:2202.08205.

Jin, W., Coley, C. W., Barzilay, R., & Jaakkola, T. (2017). Predicting organic reaction outcomes with

Weisfeiler-Lehman network. *Advances in Neural Information Processing Systems*, 30, 2607-2616.

Kasneci, E., Sessler, K., Kuchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Gunnemann, S., Hullermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences,* 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Krishna, Y. H., et al. (2016). Balancing chemical equations by using matrix algebra. *World Journal of Pharmacy and Pharmaceutical Sciences*, 6, 994-999.

Kumar, D. D. (2001). Computer applications in balancing chemical equations. *Journal of Science Education and Technology,* 10(4), 347-350. https://doi.org/10.1023/A:1012295119087

Lin, Z., Yin, S., Shi, L., Zhou, W., & Zhang, Y. (2023). G2GT: Retrosynthesis prediction with graph-to-graph attention neural network and self-training. *Journal of Chemical Information and Modeling*, 63(7), 1894-1905. https://doi.org/10.1021/acs.jcim.2c01302

Mohialden, Y. M., Hussien, N. M., & Al-Rada, W. A. A. (2023). Automated chemical equation balancing using the Apriori algorithm. *Journal La Multiapp*, 4(3), 92-97. https://doi.org/10.37899/journallamultiapp.v4i3.852

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13(9), 5783. https://doi.org/10.3390/app13095783

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching,* 6(1), 41-56. https://doi.org/10.37074/jalt.2023.6.1.29

Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., & Lee, A. A. (2019). Molecular Transformer: A model for uncertainty-calibrated chemical reaction prediction. *ACS Central Science*, 5(9), 1572-1583. https://doi.org/10.1021/acscentsci.9b00576

Schwaller, P., Petraglia, R., Zullo, V., Nair, V. H., Haeuselmann, R. A., Pisoni, R., Bekas, C., Iuliano, A., & Laino, T. (2019). Molecular Transformer unifies reaction prediction and retrosynthesis across pharma chemical space. *Chemical Communications,* 55(81), 12152-12155. https://doi.org/10.1039/C9CC05122H

Segler, M. H. S., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*, 555, 604-610. https://doi.org/10.1038/nature25978

Somnath, V. R., Bunne, C., Coley, C. W., Krause, A., & Barzilay, R. (2021). Learning graph models for retrosynthesis prediction. *Advances in Neural Information Processing Systems*, 34, 9405-9415.

Thorne, L. R. (2010). *An innovative approach to balancing chemical-reaction equations: A simplified matrix-inversion technique for determining the matrix null space*. The Chemical Educator, 15, 304-308.

Tu, Z., & Coley, C. W. (2022). Permutation invariant graph-to-sequence model for template-free retrosynthesis and reaction prediction. *Journal of Chemical Information and Modeling*, 62(15), 3503-3513. https://doi.org/10.1021/acs.jcim.2c00321

Tyson, J. (2023). Shortcomings of ChatGPT. *Journal of Chemical Education,* 100(8), 3098-3101. https://doi.org/10.1021/acs.jchemed.3c00361

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems,* 30, 5998-6008.

Wang, X., Li, Y., Qiu, J., Chen, G., Liu, H., Liao, B., Hsieh, C.-Y., & Yao, X. (2021). RetroPrime: A diverse, plausible and transformer-based method for single-step retrosynthesis predictions. *Chemical Engineering Journal,* 420, 129845. https://doi.org/10.1016/j.cej.2021.129845

Zabadi, A. M., & Assaf, R. (2017). From chemistry to linear algebra: Balancing chemical reaction equation using algebraic approach. *International Journal of Advanced Biotechnology and Research*, 8, 24-33.

Zhang, Z., Zhang, X., Zhao, Y. X., & Yang, S. A. (2024). *Balancing chemical equations from the perspective of Hilbert basis*. arXiv preprint arXiv:2410.06023.

Zhong, W., Yang, Z., & Chen, C. Y.-C. (2023). Retrosynthesis prediction using an end-to-end graph generative architecture for molecular graph editing. *Nature Communications*, 14, 3009. https://doi.org/10.1038/s41467-023-38851-5