RESEARCH PAPER

Comparative Analysis of Machine Learning, Deep Learning and Liquid Neural Networks for Drug Delivery Prediction

[1]Sanika Satish Lad, [2]Anant Manish Singh, [3]Shifa Siraj Khan, [4]Afzal Siraj Khan and [5]Kush Pulindra Patel

[1]Master of Data Science, the University of Western Australia, Perth, Australia
[2]Department of Computer Engineering, Thakur College of Engineering and Technology, Mumbai, India
[3]Department of Information Technology, Thakur College of Engineering and Technology, Mumbai, India
[4]Department of Artificial Intelligence & Machine Learning, Thakur College of Engineering and Technology, Mumbai, India
[5]Master of Data Analytics Engineering, Northeastern University, Boston, MA, US
[1]ladsanika01@gmail.com, [2]anantsingh1302@gmail.com, [3]shifakhan.work@gmail.com, [4]afzalk280306@gmail.com and [5]patelkush9819@gmail.com

Received: 28th Feb, 2026; Revised: 6th March 2026; Accepted: 7th April, 2026; Available Online: 20th April, 2026

ABSTRACT
Drug delivery prediction has become an important problem in pharmaceutical informatics because formulation success depends on several tightly connected properties, including solubility, permeability, absorption, distribution behavior and toxicity. These properties rarely follow simple linear patterns and their behavior often changes across structurally different compounds which makes robust prediction difficult under real-world generalization settings. Artificial intelligence has emerged as a practical response to this difficulty because benchmark resources such as MoleculeNet and the Therapeutics Data Commons now provide curated molecular datasets, standardized evaluation tasks and scaffold-aware splitting schemes for realistic comparison. At the same time, the growing popularity of deep learning and newer liquid neural models has created a need for careful comparative work that separates genuine scientific benefit from architectural novelty. This paper presents a comparative research framework for machine learning, deep learning and Liquid Neural Networks in drug-delivery-relevant prediction tasks using real public datasets. The study is centered on benchmark datasets including ESOL, Lipophilicity, Caco2, HIA, BBB and AqSol which together cover regression and classification endpoints directly related to pharmaceutical transport and ADMET behavior. The proposed methodology is designed to stand out in four ways. First, it prioritizes scaffold-split generalization rather than relying only on random splits. Second, it requires repeated-run statistical testing to avoid reporting unstable gains. Third, it evaluates calibration and error behavior across chemical classes instead of reporting only a single mean score. Fourth, it includes rigorous feature-set ablations and a transparent discussion of when liquid architectures are scientifically justified. The main conclusion of the paper is straightforward. A strong drug-delivery prediction study should not ask only which model scores highest on average; it should ask which model generalizes best to unseen scaffolds, remains calibrated and is scientifically appropriate for the underlying prediction problem.

Keywords: *Drug delivery prediction, ADMET modeling, Scaffold-split generalization, Machine learning benchmarking, Deep learning for drug discovery, Liquid Neural Networks, Molecular property prediction, Pharmaceutical AI*
How to cite this article: Lad SS, Singh AM, Khan SS, Khan AS, Patel KP, Comparative Analysis of Machine Learning, Deep Learning and Liquid Neural Networks for Drug Delivery Prediction. Int J Drug Deliv Technol. 2026;16(5): 119-126. DOI: 10.25258/ijddt.16.5.12
Source of support: Nil.
Conflict of interest: None

1. INTRODUCTION

1.1 Background

Drug delivery systems are central to modern pharmaceutical development because the therapeutic value of a compound depends not only on biological activity but also on whether the compound reaches the right site, at the right concentration, for the right duration. For this reason, molecular properties such as aqueous solubility, lipophilicity, membrane permeability, intestinal absorption, blood-brain barrier penetration and toxicity are widely used as early indicators of delivery feasibility and downstream formulation success. [1][2]

In practical drug development, these properties are difficult to estimate experimentally for every new candidate at scale. Laboratory evaluation is time-consuming, costly and often constrained by assay availability which has made computational prediction an essential part of early screening pipelines. [1] Public benchmark initiatives have strengthened this area by organizing shared datasets and clearer evaluation

**Author for Correspondence: ladsanika01@gmail.com*

standards, allowing researchers to compare methods under more reproducible settings.

1.2 Problem Statement

Traditional statistical models remain useful for interpretation but they often lack the flexibility needed to represent nonlinear and high-dimensional chemical-biological relationships. This limitation becomes more visible when data contain scaffold diversity, class imbalance, assay noise and hidden interactions among structural features.[3]

A second difficulty lies in evaluation design. Many predictive studies still rely heavily on random train-test partitions, even though scaffold-based separation provides a stricter and more realistic estimate of generalization to new chemical matter. As a result, models may appear stronger than they actually are when deployed on previously unseen compounds.[6]

1.3 Research Objectives

The first objective of this study is to build predictive models for drug-delivery-relevant outcomes using a consistent benchmarking pipeline over public molecular datasets. The second objective is to compare classical machine learning, deep learning and liquid neural models under the same preprocessing rules, feature sets and scaffold-aware evaluation strategy.[4]

The third objective is more analytical than competitive. It is to determine not only which model family performs best but also why it performs best, under which chemical settings and whether its architecture matches the scientific structure of the task. [4]

1.4 Research Questions

This study is guided by the following research questions:

1. Which model family provides the strongest predictive performance for drug-delivery-relevant tasks under scaffold-split evaluation?
2. Can Liquid Neural Networks outperform conventional machine learning and deep learning models in this setting?
3. Which model family generalizes best across structurally different drug compounds and chemical classes?
4. Under what conditions is a liquid architecture scientifically justified for pharmaceutical prediction rather than simply computationally fashionable?

1.5 Contributions

The contribution of this paper is a unified and rigorous framework for comparing ML, DL and LNN methods on real public drug-delivery-related benchmarks. The work also contributes a transparent evaluation design built around scaffold-split generalization, repeated-run statistical testing, calibration analysis, feature ablation and class-wise error interpretation.[6]

A further contribution is conceptual. Rather than assuming that a newer architecture is automatically better, the paper clearly separates tasks where liquid dynamics may be scientifically meaningful from tasks where they may not offer principled advantages.

2. LITERATURE REVIEW

2.1 Machine Learning in Drug Discovery and Delivery

Machine learning has played a long-standing role in molecular property prediction because traditional chemical descriptors and fingerprints already encode substantial structure-property information. Benchmark work in MoleculeNet includes methods such as logistic regression, support vector machines, random forests and other classical learners, showing that these approaches remain strong baselines, especially in modest-data regimes.[1]

In pharmaceutical prediction, these models are attractive because they are comparatively data-efficient, easier to tune and often easier to interpret at the descriptor level. For endpoints such as permeability or solubility, carefully designed tabular representations can still produce useful prediction quality, particularly when the dataset is not large enough to fully reward deep representation learning.[2][1]

2.2 Deep Learning in Pharmaceutical Prediction

Deep learning expands the modeling capacity of pharmaceutical AI by learning hierarchical representations rather than relying entirely on hand-crafted inputs. In molecular prediction, this includes dense neural networks on descriptor vectors as well as graph-based neural architectures that operate more directly on molecular topology.[7]

The main advantage of deep learning is that it can capture high-order dependencies that are difficult to specify manually. However, benchmark evidence also shows that deep models do not uniformly dominate classical methods; their gains depend on dataset size, target complexity, class balance and the evaluation split used.

2.3 Liquid Neural Networks

Liquid Neural Networks, particularly Liquid Time-constant Networks, model hidden states through continuous-time dynamics rather than a fixed discrete feedforward transformation. Their defining idea is that the internal time constants adapt according to the input and evolving hidden state which gives them flexibility in settings where temporal or dynamical behavior matters.[4]

This property has clear appeal for dynamic biological systems. If the task involves release curves, adaptive formulation states, dosing trajectories or other time-evolving signals, a liquid model may offer a more natural representation than a purely static architecture. If the task is only single-shot prediction from a fixed molecular structure, the justification becomes less obvious.[8]

2.4 Research Gap

The present literature offers many studies comparing conventional ML and DL in molecular prediction but there is still limited direct evaluation of liquid architectures in a benchmarked drug-delivery context. More importantly, few comparative studies combine scaffold-split testing, repeated-run significance analysis, calibration diagnostics, chemical-class error stratification and feature ablation within one coherent framework.[6]

This gap matters because model rankings can change once the evaluation becomes more realistic. A method that performs well under random split may lose its advantage under scaffold split or may show unstable confidence behavior across distinct chemical families.[3][6][1]

3. DATASET DESCRIPTION AND DATA SCIENCE PIPELINE

3.1 Dataset Overview

This study uses real benchmark datasets from MoleculeNet and the Therapeutics Data Commons because these resources are already recognized as standard references for molecular machine learning and ADMET prediction. The selected datasets are chosen to reflect several dimensions of drug delivery, from solubility and lipophilicity to permeability and absorption-related endpoints.[2][1]

The dataset subset is defined in Table 1.

Dataset	Relevance to drug delivery	Samples	Task type	Benchmark usage
ESOL	Aqueous solubility	1,128 [1]	Regression	MoleculeNet solubility benchmark [1]
Lipophilicity	Membrane partition tendency	4,200 [1]	Regression	MoleculeNet physicochemical prediction [1]
Caco2	Intestinal permeability proxy	906 [2][9]	Regression	TDC ADMET benchmark [2]
HIA	Human intestinal absorption	578 [2]	Classification	TDC ADMET benchmark [2]
BBB	Blood-brain barrier penetration	1,975 [2]	Classification	TDC ADMET benchmark [2]
AqSol	Large-scale aqueous solubility	9,982 [2]	Regression	TDC ADMET benchmark [2]

These datasets are suitable because they cover both regression and classification settings while remaining connected to practical pharmaceutical screening. Together they create a balanced test bed for comparing low-capacity and high-capacity model families.

3.2 Data Preprocessing

The preprocessing pipeline begins with molecule validation and canonical representation of SMILES strings, followed by duplicate removal and exclusion of records with missing target labels. This step is necessary because inconsistencies in molecular encoding can distort both descriptor generation and scaffold assignment.[2]

After cleaning, the input space is organized into four feature groups: physicochemical descriptors, structural fingerprints, graph-based atomic representations and combined multimodal representations. Numerical features are standardized using training-fold statistics only and any encoding or normalization step is fit within each split to avoid information leakage.[1][2]

3.3 Exploratory Data Analysis

Exploratory analysis should summarize target distributions, descriptor ranges, missingness patterns, scaffold frequency and inter-feature correlation. This matters because pharmaceutical datasets are rarely uniform; some endpoints show skewed distributions, while others contain sparse classes or chemically clustered regions that can bias model evaluation.

A scaffold-level profile is especially important. Since scaffold split intentionally separates structural cores between training and test compounds, the analysis should quantify how much structural novelty is introduced at test time and why this makes the prediction task more difficult than random splitting.[3]

3.4 Feature Engineering

Feature engineering should remain hypothesis-driven rather than purely expansive. Descriptor sets should include interpretable physicochemical properties such as molecular weight, logP, hydrogen-bond donor count, hydrogen-bond acceptor count, topological polar surface area and rotatable bonds because these features are mechanistically linked to solubility, permeability and absorption behavior.

Fingerprint representations provide a stronger structural encoding for classical models, while graph-based inputs allow deep models to learn relational structure directly from atoms and bonds. Dimensionality reduction can be considered as an optional ablation but it should not be assumed beneficial without testing because some predictive information may be lost when the chemical space is compressed too aggressively.[7]

4. METHODOLOGY

4.1 Experimental Framework

The central experimental design uses scaffold split as the primary evaluation protocol because both MoleculeNet and TDC identify scaffold-aware partitioning as a more realistic test of molecular generalization than random train-test splitting. Random split may still be reported as a secondary comparison but it should not be used as the main basis for scientific claims.[6][2]

To avoid unstable conclusions, each model should be trained across repeated runs with different initialization seeds. Repeated-run analysis allows the study to report mean performance, standard deviation, confidence intervals and significance testing instead of presenting a single lucky score as the final result.[10]

4.2 Machine Learning Models

The machine learning block includes Linear Regression or Elastic Net for regression tasks, Logistic Regression for classification, Support Vector Regression or Support Vector Classification, Random Forest and Gradient Boosting or XGBoost-style ensembles. These models form a strong baseline because they work well with descriptor and fingerprint representations and often remain competitive in medium-sized molecular datasets.

4.3 Deep Learning Models

The deep learning block includes a fully connected neural network on tabular descriptors or fingerprints and a graph-based model for molecular structure learning. This combination is important because it separates two questions: whether neural nonlinearities improve tabular prediction and whether topology-aware feature learning improves over hand-crafted features.[11]

Hyperparameter tuning should be performed under a common budget across models to maintain fairness. Search spaces should include depth, hidden dimension, dropout, learning rate, batch size, activation choice and regularization strength depending on the architecture.

4.4 Liquid Neural Network Model

The liquid model in this paper follows the Liquid Time-constant Network idea, where hidden states evolve according to continuous-time dynamics rather than a fixed one-pass transformation. A simplified formulation can be written as:[4]

$$\frac{dh(t)}{dt} = f(h(t), x(t), t; \theta)$$

where $h(t)$ is the hidden state, $x(t)$ is the input signal and θ denotes trainable parameters. The prediction is produced by a readout function:

$$\hat{y} = g(h(T))$$

This architecture is scientifically strongest when the input contains genuine temporal structure, such as drug-release trajectories, dissolution measurements over time, patient-specific dosing sequences or adaptive formulation signals. If the task uses only static molecular descriptors from single compounds, the LNN should be evaluated carefully and interpreted as an exploratory model unless a valid dynamic representation is explicitly constructed.[8]

4.5 Evaluation Metrics

Regression tasks should be evaluated using Mean Absolute Error and Root Mean Squared Error because they capture complementary aspects of prediction quality. The formulas are:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Classification tasks should report Accuracy, Precision, Recall, F1-score and ROC-AUC with probability-based evaluation extended to calibration metrics such as Brier score and expected calibration error when class probabilities are available. The standard formulas are:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

For probabilistic calibration, the Brier score is given by:

$$\text{Brier} = \frac{1}{n}\sum_{i=1}^{n} (p_i - y_i)^2$$

In addition to predictive metrics, computational efficiency should be reported using training time, inference time and approximate model complexity. This is especially important when comparing simpler tabular baselines with heavier deep or liquid architectures.

5. EXPERIMENTAL SETUP

The experimental setup should explicitly report hardware, software versions, random seeds, molecular processing tools and training frameworks. At minimum, the study should state CPU or GPU type, available RAM, Python version, RDKit version and the deep learning libraries used for model training.

To make the comparison reproducible, the same scaffold partitions and feature-generation rules should be used across all model families. Hyperparameter tuning must be performed only on the training data with a separate validation stage and no test information should influence model selection.

A practical tuning design is to use a fixed search budget across families. This prevents the comparison from favoring one family simply because it received more optimization effort.[1]

6. RESULTS AND COMPARATIVE ANALYSIS

6.1 Quantitative Results

The most important expectation from benchmark evidence is that scaffold-split performance will be lower than random-split performance for many tasks because structural overlap between train and test compounds is intentionally reduced [3][6]. This does not indicate model failure; it indicates that the evaluation is more realistic.

Within this setting, the likely pattern is task-dependent rather than universal. Classical ML may remain competitive for descriptor-rich and relatively small datasets, deep models may gain more when molecular structure learning is central and liquid models are most likely to help when the input contains meaningful dynamic information rather than static descriptors alone.

6.2 Model Performance Comparison

The comparison should not stop at average performance values. Each model family should be examined across predictive accuracy, generalization stability, calibration quality and robustness to chemical diversity.

A scientifically credible result section should therefore report the following for every task and model: mean score across repeated runs, standard deviation, 95% confidence interval and the difference between random-split and scaffold-split performance. This presentation gives a much clearer picture than a single leaderboard number.[6]

6.3 Scaffold-Split Generalization Analysis

Scaffold split is central to this paper because it directly tests the ability of a model to extrapolate beyond the structural frameworks seen during training. In pharmaceutical settings, this is closer to the real use case, where a model is often asked to evaluate new analog series rather than near-duplicates of molecules it has already seen.

A model that performs slightly worse on random split but degrades less sharply under scaffold split may be more useful in practice than a model with the highest random-split score. For this reason, the drop from random to scaffold performance should be treated as an explicit robustness indicator.

6.4 Repeated-Run Statistical Testing

Repeated-run analysis is necessary because neural models can vary meaningfully across random initializations and optimization trajectories. A fair comparison should therefore estimate the distribution of performance rather than a single point estimate.[10]

For each dataset, the study should compute mean and standard deviation over repeated runs and apply paired significance testing between competing models. When normality is plausible, a paired t-test is acceptable; otherwise, the Wilcoxon signed-rank test provides a more distribution-robust alternative. Statistical significance should be interpreted together with effect size, not in isolation.

6.5 Calibration and Error Stratification by Chemical Class

Calibration analysis is important because a model can be accurate on average and still produce unreliable confidence estimates. In drug-delivery prediction, poorly calibrated probabilities can distort screening priorities, especially when a model is used for candidate triage or risk-sensitive filtering.

Error stratification by chemical class adds another necessary layer. This makes the evaluation far more informative for real pharmaceutical decision-making.

6.6 Feature-Set Ablation

Ablation analysis should compare descriptor-only, fingerprint-only, graph-only and multimodal feature settings under the same scaffold partitions. This reveals whether a model's gain comes from better architecture, richer input representation or simply more information.

This part is especially important for LNN interpretation. If the liquid model only improves after being given a pseudo-sequential representation of static features, our study carefully examine whether the gain reflects a meaningful dynamic inductive bias or merely an alternative nonlinear transformation.

7. DISCUSSION

The broader interpretation of this comparative study is that architecture choice should follow problem structure rather than trend. For many benchmark endpoints related to drug delivery, the data are still static molecule-level observations and this means strong classical or graph-based models may be more scientifically natural than liquid temporal models.

That said, liquid models should not be dismissed. They may become genuinely valuable when the prediction target includes time-varying formulation behavior, release kinetics, longitudinal patient response or adaptive biological dynamics. In those settings, continuous-time hidden-state evolution can be justified by the science of the problem rather than by novelty alone.[8]

The comparison also highlights an important methodological lesson. The most useful pharmaceutical predictor is not necessarily the model with the highest headline score but the one that remains stable under scaffold shift, behaves sensibly across chemical classes and expresses uncertainty in a trustworthy way.

8. PRACTICAL IMPLICATIONS

The results of such a framework can directly support pharmaceutical screening, formulation optimization and candidate prioritization. A calibrated and scaffold-robust model can reduce wasted experimental effort by identifying which compounds are promising and which predictions are too uncertain to trust without laboratory validation.[5]

In personalized or adaptive medicine settings, the architectural discussion becomes even more important. Static predictors may remain sufficient for one-time property estimation, whereas dynamic liquid models may be more appropriate when treatment, release or response unfolds over time.[8]

9. LIMITATIONS

Public benchmark datasets are extremely useful but they do not fully represent the entire complexity of real drug delivery systems. Most benchmark tasks remain proxy endpoints at the molecule level rather than full formulation-level, patient-level or time-resolved delivery processes.

This limitation affects the interpretation of liquid architectures in particular. Since LNNs are designed for dynamic settings, a benchmark composed mostly of static molecular labels may underestimate their true scientific value in genuinely time-dependent pharmaceutical applications.[8]

A second limitation is that benchmark quality depends on assay consistency, label noise and dataset coverage. Even when the split protocol is rigorous, the biological and chemical diversity of public data still constrains how broadly conclusions can be generalized.

10. FUTURE WORK

Future research should move beyond static molecular endpoints toward datasets that contain release profiles, dissolution curves, pharmacokinetic trajectories, pharmacodynamic responses and clinically contextual delivery outcomes. Such data would allow a stronger and more scientifically grounded test of whether liquid neural models truly offer an advantage.

Another promising direction is hybrid modeling. Graph encoders could be used to learn molecular structure, while liquid modules could model temporal adaptation over release or response trajectories. This type of combination may prove more useful than treating liquid networks as a universal replacement for all molecular models.[8]

Future work should also incorporate explainable AI methods so that prediction differences across ML, DL and LNN families can be interpreted in chemically meaningful terms. This is particularly valuable when the goal is not only prediction but also formulation insight and candidate design support.[5]

11. CONCLUSION

This paper presents a clear and rigorous framework for comparing machine learning, deep learning and Liquid Neural Networks in drug-delivery-relevant prediction tasks using real public benchmark datasets. The central argument is that a meaningful comparison must go beyond average accuracy and include scaffold-split generalization, repeated-run statistics, calibration, chemical-class error analysis and feature ablation.

The evidence from benchmark literature supports scaffold-aware evaluation as the stronger estimate of real-world molecular generalization. It also supports a balanced scientific interpretation of liquid models: they are promising when the task is genuinely dynamic but they should not be assumed superior for static molecular-property prediction without a valid temporal rationale.

In this sense, the paper stands out not by making exaggerated claims but by making disciplined ones. The scientifically best model for drug delivery prediction is the one that matches the structure of the pharmaceutical problem, generalizes across unseen chemistry and remains reliable when its predictions are used to guide real decisions.

REFERENCES

[1]. Wu, Z., Ramsundar, B., Feinberg, E. N., Gomes, J., Geniesse, C., Pappu, A. S., Leswing, K., & Pande, V. (2017). MoleculeNet: a benchmark for molecular machine learning. *Chemical science*, *9*(2), 513–530. https://doi.org/10.1039/c7sc02664a

[2]. Huang, K., Fu, T., Glass, L. M., Zitnik, M., Xiao, C., & Sun, J. (2022). **Therapeutics Data Commons: Machine learning datasets and tasks for drug discovery and development**. *Proceedings of Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS Datasets and Benchmarks)*.

[3]. Niu, Z., Xiao, X., Wu, W. *et al.* PharmaBench: Enhancing ADMET benchmarks with large language models. *Sci Data* **11**, 985 (2024). https://doi.org/10.1038/s41597-024-03793-0

[4]. Hasani, R., Lechner, M., Amini, A., Rus, D., & Grosu, R. (2020). *Liquid time-constant networks*. arXiv. https://doi.org/10.48550/arXiv.2006.04439

[5]. **Wu, Y., Wang, N., Xiong, P., Wang, R., Deng, J., & Ouyang, D. (2025).** Artificial intelligence for drug delivery: Yesterday, today and tomorrow. *Acta Pharmaceutica Sinica B*. Advance online publication. https://doi.org/10.1016/j.apsb.2025.09.022

[6]. Meng, J., Chen, P., Wahib, M. *et al.* Boosting the predictive performance with aqueous solubility dataset curation. *Sci Data* **9**, 71 (2022). https://doi.org/10.1038/s41597-022-01154-3

[7]. **Mohamed, A., Galal, N., Brooks, B. R., & Amin, M. (2026).** Graph neural networks model based on atomic hybridization for predicting drug targets. *Journal of Chemical Information and Modeling*. https://doi.org/10.1021/acs.jcim.5c02961

[8]. **Han, P., Wang, J., Liu, D., Liu, L., & Song, T. (2025).** Robust temporal knowledge inference via pathway snapshots with liquid neural network. *Methods, 241*, 24–32. https://doi.org/10.1016/j.ymeth.2025.05.003

[9]. Le, H., Ren, W., Kim, J. *et al.* CaliciBoost: Performance-driven evaluation of molecular representations for caco-2 permeability prediction. *J Cheminform* **17**, 184 (2025). https://doi.org/10.1186/s13321-025-01137-7

[10]. **Dong, T., You, L., & Chen, C. Y.-C. (2025).** Multi-objective drug design with a scaffold-aware variational autoencoder. *Chemical Science, 16*, 13352–13367. https://doi.org/10.1039/D4SC08736D

[11]. Wen, T., Cai, X., & Li, J. (2025). Graph Neural Networks vs. Traditional QSAR: A Comprehensive Comparison for Multi-Label Molecular Odor Prediction. *Molecules (Basel, Switzerland)*, *30*(23), 4605. https://doi.org/10.3390/molecules30234605

[12]. Suksaeree, J., Maneewattanapinyo, P., & Monton, C. (2026). AI-enabled, QbD-aligned Predictive, and Sustainable Design of Natural Polymer-based Drug Delivery Systems. *AAPS PharmSciTech*, *27*(2), 90.

[13]. Manigandan, M., Gayathri, K., Vithiya, G., Vishnuvarthan, N., & Saranraj, P. (2026). Machine Learning and AI in Nanotechnology with Ionic Liquids and Nanofluid and Bionanofluids. In *Ionic Liquids, Nanofluids and Nanotechnology: Innovations, Applications and Sustainability* (pp. 553-566). Cham: Springer Nature Switzerland.

[14]. Ashwini, A., & Preemi, G. (2026). Innovations of Machine Learning with AI-Driven Nanotechnology in Ionic Liquids, Nanofluids, and Bionanofluids. In *Ionic Liquids, Nanofluids and Nanotechnology: Innovations, Applications and Sustainability* (pp. 533-552). Cham: Springer Nature Switzerland.

[15]. Joseph, J., Sahu, J., Baldota, J., Parab, V. A. V., Sharma, T. P., & Walsinge, S. (2026). Advancements in Nanotechnology: The Role of Machine Learning and AI in the Development of Ionic Liquids, Nanofluids, and Bionanofluids. In *Ionic Liquids, Nanofluids and Nanotechnology: Innovations, Applications and Sustainability* (pp. 435-478). Cham: Springer Nature Switzerland.

[16]. Zeng, M., Deng, P., Yang, Q., Hu, J., Li, J., Tang, Q., ... & Zhang, L. (2026). Machine learning-guided composite ionic liquid-based system for dual-drug delivery targeting redox homeostasis and STAT3–PI3K axis in psoriasis therapy. *Bioactive Materials*, *58*, 107-122.

[17]. Cheong, E., Radford, D. C., & Gormley, A. J. (2026). Automated active learning to optimize hydrogel drug release profiles. *Journal of Controlled Release*, 114602.

[18]. Kantesaria, R., & Panda, H. S. (2026). A Review on AI-Based Data-Driven Models for Optimization of Nanocarriers as Drug Delivery Systems. *ACS Biomaterials Science & Engineering*.

[19]. Yuan, X. Y., Hua, Y., Aubry, N., Zhussupbekov, M., Antaki, J. F., Zhou, Z. F., & Peng, J. Z. (2022). Real-time prediction of transarterial drug delivery based on a deep convolutional neural network. *Applied Sciences*, *12*(20), 10554.

[20]. Piroozmand, F., Mohammadipanah, F., & Sajedi, H. (2023). Artificial neural network (ANN) in drug delivery. In *A handbook of artificial intelligence in drug delivery* (pp. 97-122). Academic Press.

[21]. Islam, N., Akhtar, Y., Ahmad, S., Junjua, M. U. D., Hendy, A. S., Alballa, T., & Khalifa, H. A. E. W. (2024). Advancing drug delivery: Neural network perspectives on nanoparticle-mediated treatments for cancerous tissues. *Nanotechnology Reviews*, *13*(1), 20240129.

[22]. Salarpour, S., Salarpour, S., & Dogaheh, M. A. (2025). Advancing pharmaceutical science with artificial neural networks: a review on optimizing drug delivery systems formulation. *Current Pharmaceutical Design*, *31*(7), 507-520.

[23]. Wang, S., Di, J., Wang, D., Dai, X., Hua, Y., Gao, X., ... & Gao, J. (2022). State-of-the-art review of artificial neural networks to predict, characterize and optimize pharmaceutical formulation. *Pharmaceutics*, *14*(1), 183.

[24]. Panchpuri, M., Painuli, R., & Kumar, C. (2025). Artificial intelligence in smart drug delivery systems: A step toward personalized medicine. *RSC pharmaceutics*, *2*(5), 882-914.

[25]. Biswas, A. A., Dhondale, M. R., Agrawal, A. K., Serrano, D. R., Mishra, B., & Kumar, D. (2024). Advancements in microneedle fabrication techniques: Artificial intelligence assisted 3D-printing technology. *Drug Delivery and Translational Research*, *14*(6), 1458-1479.

[26]. Nagy, B., Galata, D. L., Farkas, A., & Nagy, Z. K. (2022). Application of artificial neural networks in the process analytical technology of pharmaceutical manufacturing—a review. *The AAPS journal*, *24*(4), 74.

[27]. Prajapati, J. B., Paliwal, H., Saikia, S., Prajapati, B. G., Prajapati, D. N., Philip, A. K., & Faiyazuddin, M. (2023). Impact of AI on drug delivery and pharmacokinetics: The present scenario and future prospects. In *A Handbook of Artificial Intelligence in Drug Delivery* (pp. 443-465). Academic Press.

[28]. Lin, Z., Chou, W. C., Cheng, Y. H., He, C., Monteiro-Riviere, N. A., & Riviere, J. E. (2022). Predicting nanoparticle delivery to tumors using machine learning and artificial intelligence approaches. *International journal of nanomedicine*, 1365-1379.

[29]. Bannigan, P., Aldeghi, M., Bao, Z., Häse, F., Aspuru-Guzik, A., & Allen, C. (2021). Machine learning directed drug formulation development. *Advanced Drug Delivery Reviews*, *175*, 113806.

[30]. He, S., Leanse, L. G., & Feng, Y. (2021). Artificial intelligence and machine learning assisted drug delivery for effective treatment of infectious diseases. *Advanced drug delivery reviews*, *178*, 113922.