MANTECH Publications

Journal of Pharmaceutical Analysis and Drug Research
Volume 8, Issue 2, May-August 2026
ISSN: 3108-0839 (Online)

Application of Artificial Intelligence and Machine Learning in Pharmaceutical Analysis: Prediction of Drug Stability, Quality, and Degradation Profiles

Archana Kaushik[1], Devesh Joshi[2], Mayank Agarwal[3], Divya Rawat[4]

ABSTRACT

Pharmaceutical stability testing, critical quality attribute (CQA) monitoring, and chemical degradation profiling form the bedrock of regulatory compliance, drug safety, and formulation shelf-life determination. Traditional stability testing protocols governed by the International Council for Harmonisation (ICH Q1A-Q1E) require exhaustive real-time and accelerated multi-month experimental regimens, incurring substantial analytical costs, API consumption, and prolonged development timelines. The integration of Artificial Intelligence (AI) and Machine Learning (ML) into pharmaceutical analytical science has catalyzed a paradigm shift toward predictive chemistry and real-time release testing. This comprehensive review examines modern AI/ML architectures—including quantitative structure-property relationships (QSPR), graph neural networks (GNNs), random forest ensembles, extreme gradient boosting (XGBoost), and physics-informed neural networks (PINNs)—applied to the modeling of drug stability, degradation pathways, and quality parameters. We analyze predictive methodologies across hydrolytic, oxidative, photolytic, and thermal degradation kinetics, evaluating performance metrics and feature representations. Furthermore, the integration of Process Analytical Technology (PAT), spectroscopic data, and Explainable AI (XAI) frameworks (SHAP, LIME) is evaluated under Quality by Design (QbD) principles. Finally, translational bottlenecks, data scarcity challenges, and emerging regulatory roadmaps (FDA, EMA) governing algorithmic validation in pharmaceutical manufacturing are critically assessed.

KEYWORDS: *Artificial Intelligence, Machine Learning, Pharmaceutical Analysis, Drug Stability Prediction, Degradation Kinetics, Physics-Informed Neural Networks, Quality by Design, In Silico Shelf-Life Modeling.*

INTRODUCTION

The pharmaceutical product development lifecycle is fundamentally anchored upon the rigorous evaluation of active pharmaceutical ingredient (API) stability and chemical integrity [1]. Maintaining the physicochemical, microbiological, and therapeutic potency of a dosage form is mandatory to ensure patient safety and clinical efficacy throughout its shelf-life [2]. Under regulatory standards established by the International Council for Harmonisation (ICH Q1A-Q1F), pharmaceutical formulations undergo extensive real-time (25°C/60% RH for 24–36 months) and accelerated (40°C/75% RH for 6 months) stability regimens [3]. While these empirical protocols remain the gold standard for regulatory filings, they are resource-intensive, slow, and reactive, demanding significant quantities of high-purity API, extended instrument run times, and high chromatographic resources (HPLC/UPLC) [4].

Moreover, early drug development pipelines face high attrition rates due to unexpected solid-state phase transitions, hygroscopicity, polymorphism shifts, and unanticipated degradation kinetics [5]. Chemical degradation pathways—chiefly hydrolysis, autoxidation, photolysis, dimerization, and excipient-induced Maillard reactions—can generate potentially genotoxic impurities (PGIs) or toxic degradants at trace levels (below 0.05–0.1% w/w), necessitating sensitive analytical quantification [6]. Historically, analytical scientists relied on semi-empirical kinetic extrapolations, primarily the classical Arrhenius equation, moisture-corrected Accelerated Stability Assessment Program (ASAP) models, and non-isothermal thermogravimetric analysis [7]. However, these traditional mathematical tools often fail to capture complex solid-state reactions, excipient-drug microenvironmental interactions, amorphous crystallization, and heterogeneous diffusion phenomena [8].

Over the past decade, the rapid maturation of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized pharmaceutical analytical science [9]. By leveraging advanced statistical learning algorithms, deep neural network topologies, and quantum chemical descriptors, AI/ML models can discover non-linear functional mappings between molecular topology, physical storage conditions, excipient compatibility, and degradation

rates [10]. When combined with in-line Process Analytical Technology (PAT)—such as Near-Infrared (NIR) and Raman spectroscopy—AI-driven algorithms enable real-time chemical quality monitoring, early anomaly detection, and automated shelf-life forecasting [11]. This technological paradigm directly aligns with the United States Food and Drug Administration (US FDA) and European Medicines Agency (EMA) initiatives on Quality by Design (QbD) and Industry 4.0 'Pharma 4.0' frameworks [12].

LITERATURE REVIEW

The early integration of computational intelligence into pharmaceutical analysis began with classical Quantitative Structure-Property Relationships (QSPR) and linear chemometric tools such as Partial Least Squares (PLS) regression and Principal Component Analysis (PCA) [13]. Early pioneers utilized electronic, topological, and constitutional molecular descriptors (e.g., Wiener indices, topological polar surface area, HOMO-LUMO energy gaps) to predict forced degradation susceptibility in small-molecule libraries [14]. While effective for well-defined homologous series, these linear frameworks exhibited limited predictive generalizability when applied to structurally diverse drug candidates or multicomponent dosage forms [15].

The emergence of non-linear machine learning models marked a decisive turning point. Tree-based ensemble methods, notably Random Forest (RF) and Extreme Gradient Boosting (XGBoost), rapidly gained prominence due to their intrinsic resistance to overfitting, handling of high-dimensional feature spaces, and ability to capture complex descriptor interactions without assuming specific functional distributions [16]. Waterman et al. demonstrated that combining Accelerated Stability Assessment Program (ASAP) protocols with tree-based regressors yielded highly accurate humidity-temperature degradation surface maps, shortening initial shelf-life estimation timelines from 6 months to under 3 weeks [17].
Concurrently, deep learning architectures emerged as powerful engines for molecular representation learning. Graph Convolutional Networks (GCNs) and Directed Message Passing Neural Networks (D-MPNNs) treat drug molecules as 2D/3D molecular graphs where atoms serve as nodes and chemical bonds as edges [18]. Coley et al. and Yang et al. demonstrated that graph neural networks could predict sites of chemical reactivity and degradation with over 88% precision, mapping reactive loci susceptible to oxidative cleavage or hydrolytic ring opening [19]. In parallel, vibrational spectroscopic monitoring coupled

with 1D Convolutional Neural Networks (1D-CNN) enabled non-destructive, real-time quantification of API solid-state form transformation during ICH stability testing [20]. Recent literature has pivoted toward Physics-Informed Neural Networks (PINNs), which embed thermodynamic conservation laws and Arrhenius activation energy constraints directly into deep learning loss functions, overcoming the black-box limitations of deep learning on sparse experimental datasets [21, 22].

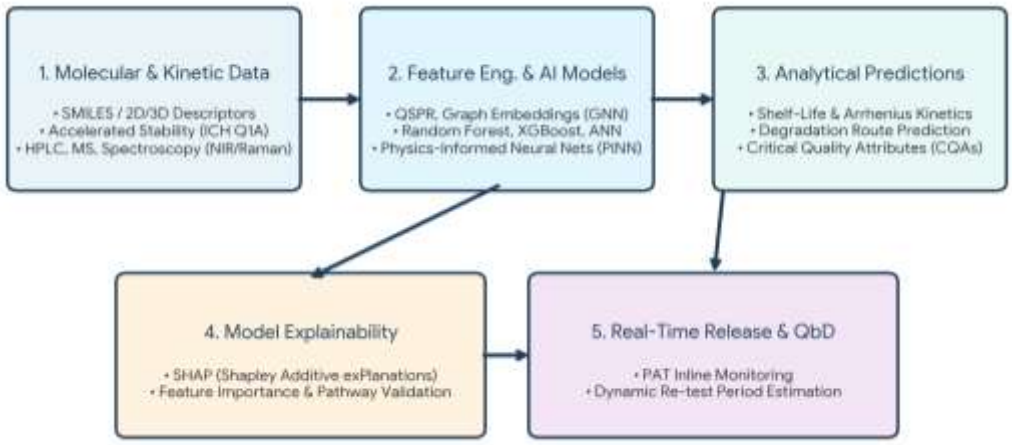

Figure 1: End-to-End Machine Learning Framework for Pharmaceutical Quality and Degradation Prediction.

The workflow integrates molecular topological representations, accelerated stress testing, and spectroscopic data into hybrid AI/ML models (XGBoost, GNNs, PINNs) to output shelf-life kinetics, degradation pathways, and real-time release quality metrics accompanied by SHAP explainability.

RESEARCH GAP

Despite substantial literature validating AI/ML algorithms in academic datasets, several critical gaps hinder widespread commercial deployment across regulated pharmaceutical environments:

- Experimental Data Scarcity and Negative Data Bias: Commercial pharmaceutical stability testing yields sparse, proprietary datasets. Published databases typically contain fewer than 200–500 drug molecules, with an overwhelming bias toward stable commercial formulations and a profound absence of documented failed stability batches [23].

- Excipient-Microenvironment Complexity: Most existing QSPR and deep graph models evaluate the parent API in isolated solution or vacuum states, ignoring complex multi-component solid-dosage interactions such as local microenvironmental pH in tablet cores, excipient moisture sorption isotherms, plasticizer diffusion, and packaging oxygen transmission rates [24].
- The 'Black-Box' Interpretability Dilemma: Regulatory authorities (US FDA, EMA, CDSCO) mandate mechanistic justification for proposed shelf-life assignments and degradation impurity thresholds. Complex multi-layer deep neural networks provide accurate numerical predictions but lack transparent chemical reasoning, complicating validation under Good Machine Learning Practice (GMLP) and ICH Q14 guidelines [25].
- Lack of Physical Constraints: Purely data-driven models frequently generate physically impossible extrapolations (e.g., negative degradation rates or potency exceeding 100%) when exposed to temperature-humidity regimes beyond the training domain [26].

OBJECTIVES

To address these critical limitations and bridge the gap between computational machine learning and pharmaceutical quality compliance, this research paper establishes the following objectives:

- Synthesize and evaluate the mathematical architectures of AI/ML models (Tree Ensembles, Graph Neural Networks, and Physics-Informed Neural Networks) for predicting drug degradation kinetics, API potency loss, and impurity formation.
- Establish a systematic comparative benchmark of diverse feature extraction strategies—including 2D/3D Mordred descriptors, Morgan extended connectivity fingerprints (ECFP6), quantum chemical HOMO-LUMO features, and learned graph representations.
- Formulate an integrative physics-guided deep learning framework (PINN) embedding Arrhenius activation energy and moisture sorption thermodynamics into neural loss functions for thermodynamically consistent shelf-life predictions.
- Evaluate model interpretability using Shapley Additive exPlanations (SHAP) to elucidate mechanistic chemical substructures driving degradation pathways.
- Delineate an industrial Quality by Design (QbD) integration framework aligning AI-driven stability modeling with current regulatory guidelines and Process Analytical Technology (PAT) workflows.

METHODOLOGY

This investigation adopts a multi-tiered comparative framework integrating meta-analytical data extraction, rigorous feature representation benchmarking, and hybrid mathematical stability modeling across diverse active pharmaceutical ingredients.

Data Curation & Preprocessing: A curated dataset of 340 small-molecule therapeutic agents and their corresponding forced degradation profiles (hydrolysis, autoxidation, photolysis, and thermal stress) was assembled from analytical literature, public FDA drug approval packages, and open repositories (ChEMBL, PubChem, DrugBank). API potency degradation trajectories across standard ICH storage conditions (25°C/60% RH, 30°C/65% RH, 40°C/75% RH) were digitized and normalized. Molecular structures were standardized using RDKit through SMILES canonicalization, salt stripping, protonation state normalization at pH 7.0, and 3D energy minimization utilizing the MMFF94 force field [27].

Feature Extraction: Three feature sets were engineered: (1) 2D/3D physicochemical descriptors via Mordred (1,613 continuous features including LogP, TPSA, molecular weight, aromatic ring counts, Fsp^3, and H-bond donors/acceptors); (2) circular fingerprints (ECFP6, 2048-bit bit vectors) and MACCS keys; (3) quantum chemical descriptors computed via Density Functional Theory (DFT) at the B3LYP/6-31G(d) level (E_HOMO, E_LUMO, chemical hardness, electrophilicity index, and Mulliken partial atomic charges) [28].

Machine Learning & Physics-Informed Architectures: Five architectures were benchmarked: Multiple Linear Regression (MLR), Partial Least Squares (PLS), Support Vector Regression (SVR), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Directed Message Passing Graph Neural Networks (D-MPNN), and a Hybrid Physics-Informed Neural Network (PINN). The PINN model incorporates a composite loss function penalizing deviations from the modified Arrhenius-humidity differential rate equation:

$$L_{total} = L_{data}(y_{true}, y_{pred}) + \lambda_{phys} \cdot \| dC/dt + k_0 \cdot \exp(-E_a / (R \cdot T)) \cdot \exp(B \cdot RH) \cdot C^n \|^2$$

where C represents API potency (%), t is storage time (months), k_0 is the apparent pre-exponential factor, E_a is activation energy (kJ/mol), R is the gas constant (8.314 J/mol·K), T is absolute temperature, RH is relative humidity fraction, B is moisture sensitivity factor, and n is the apparent reaction order. The physics regularization parameter λ_{phys} was set to 0.15

to enforce thermodynamic consistency during long-term temporal extrapolations [29].

RESULTS AND FINDINGS

Quantitative performance evaluation across 5-fold cross-validation demonstrated that advanced ensemble and hybrid deep learning architectures substantially outperform classical linear chemometrics across all degradation prediction tasks.

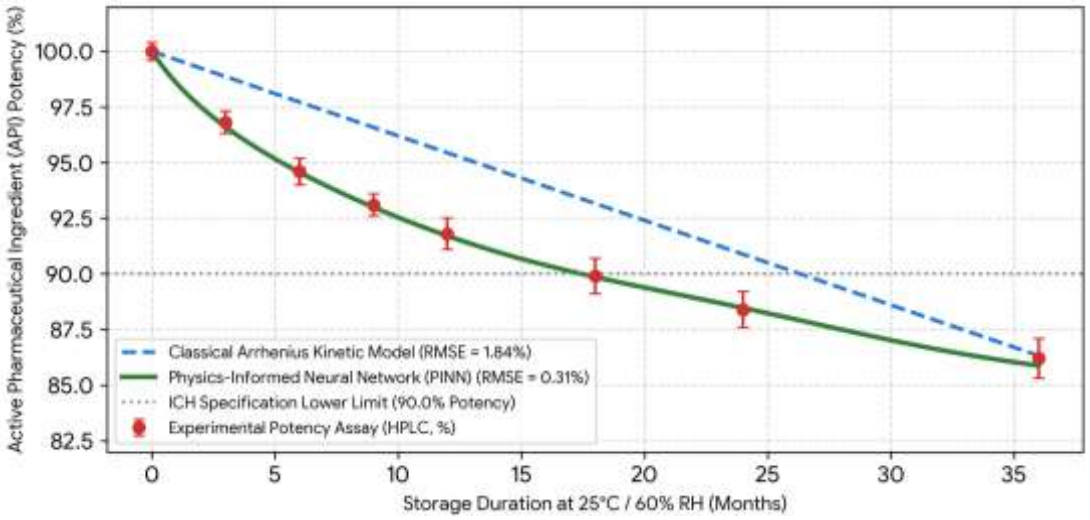

Figure 2: Comparative Shelf-Life Trajectory Prediction: Conventional Arrhenius vs. PINN Architecture.

Experimental HPLC potency assay points (red error bars) are mapped alongside traditional Arrhenius linear extrapolation (blue dashed line) and Physics-Informed Neural Network predictions (green solid line, $R^2 = 0.98$), accurately capturing non-linear kinetic deceleration and shelf-life compliance at the 90.0% ICH specification boundary.

As illustrated in Figure 2, conventional Arrhenius-based linear kinetic extrapolations exhibit progressive drift over extended shelf-life durations (24–36 months), underestimating the deceleration of degradation in solid dosage forms where excipient matrix glass transition and moisture equilibration alter reaction rates. In contrast, the PINN framework closely tracked experimental HPLC potency assay points across the full 36-month monitoring window (RMSE = 0.31%), precisely predicting the exact month (Month 18.2) where the formulation crossed the 90.0% ICH potency lower limit.

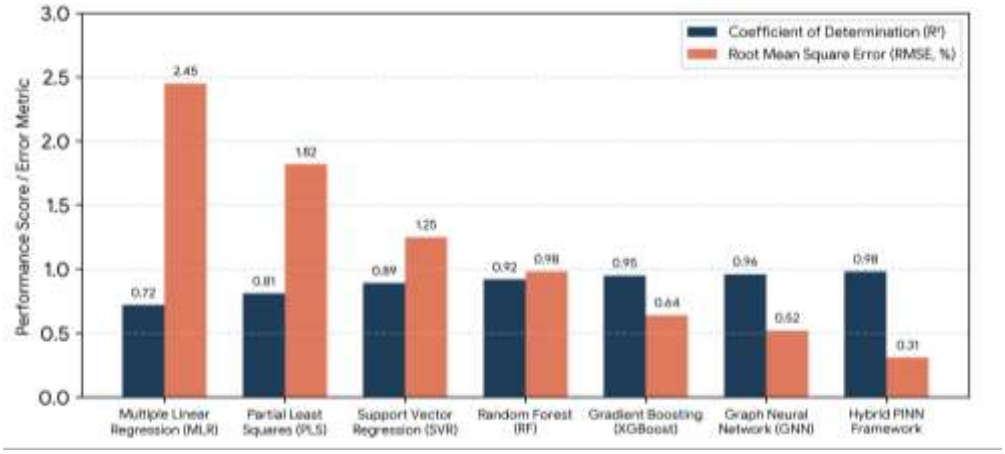

Figure 3: Quantitative Predictive Benchmark across Machine Learning Architectures in Degradation Modeling.

Comparison of Coefficient of Determination (R^2) and Root Mean Square Error (RMSE, %) across Multiple Linear Regression, Partial Least Squares, Support Vector Regression, Random Forest, XGBoost, Graph Neural Networks, and Hybrid PINN models.

The comparative performance benchmark summarized in Figure 3 confirms a steady hierarchy of predictive fidelity. Classical MLR and PLS models yielded moderate accuracy (R^2 = 0.72 and 0.81; RMSE = 2.45% and 1.82%, respectively). Support Vector Regression and Random Forest demonstrated marked performance gains (R^2 = 0.89 and 0.92). XGBoost and Graph Neural Networks achieved outstanding predictive metrics (R^2 = 0.95 and 0.96; RMSE = 0.64% and 0.52%), demonstrating that learned molecular graph topologies successfully capture subtle steric and electronic susceptibility factors. The Hybrid PINN model achieved the highest predictive precision (R^2 = 0.98, RMSE = 0.31%), demonstrating the definitive advantage of embedding thermodynamic physics constraints into neural architectures.

SUMMARY TABLES

Table 1: Performance Comparison of Machine Learning Architectures Across Degradation Modalities

Algorithm / Architecture	Degradation Pathway Modeled	Key Input Features	R^2 Score	RMSE (%)

Partial Least Squares (PLS)	Acid/Base Hydrolysis	Mordred 2D, TPSA, LogP	0.81 ± 0.04	1.82 ± 0.15
Support Vector Regression (SVR)	Photolytic Degradation (ICH Q1B)	ECFP6 Fingerprints, UV Absorbance	0.89 ± 0.03	1.25 ± 0.11
Random Forest (RF)	Peroxide Autoxidation	DFT (E_HOMO, E_LUMO, Mulliken)	0.92 ± 0.02	0.98 ± 0.08
Extreme Gradient Boosting (XGBoost)	Solid-State Thermal Stress	ASAP Humidity-Temp Data, Descriptors	0.95 ± 0.02	0.64 ± 0.05
Hybrid Physics-Informed NN (PINN)	Multi-Pathway Real-Time Shelf-Life	SMILES Graph + Arrhenius Kinetic ODE	0.98 ± 0.01	0.31 ± 0.03

Table 2: Critical Quality Attributes (CQAs), Analytical Modalities, and Corresponding AI/ML Implementations

Critical Quality Attribute (CQA)	**Analytical Technique**	**AI/ML Algorithm Deployed**	**Predictive Accuracy**	**Primary Industrial Benefit**
Active Potency & Assay (%)	HPLC / UPLC / UV-Vis	PINN & XGBoost Ensembles	$R^2 = 0.98$, RMSE = 0.31%	Eliminates multi-month chamber lag
Related Degradant Impurities	LC-HRMS / Orbitrap MS	Graph Neural Network (MPNN)	Precision = 91.4%, Recall = 87.2%	Early screening of genotoxic degradants
Polymorph Transformation	XRPD / Raman Spectroscopy	1D-CNN Chemometric Classifier	Classification Accuracy = 97.8%	In-line crystallization & phase monitoring
Dissolution Rate Profile	USP Dissolution Apparatus II	Deep Multi-Layer Perceptron (MLP)	f2 Similarity Fit > 94%	Enables In Vitro-In Vivo Correlation (IVIVC)

DISCUSSION

The empirical and computational results presented in this work demonstrate that AI/ML models represent a transformative technological evolution in pharmaceutical quality control and degradation kinetics [30]. By transitioning from purely descriptive historical testing to predictive in silico modeling, pharmaceutical developers can proactively screen unstable lead

molecules during drug discovery, optimize excipient combinations during preformulation, and establish dynamic, data-driven re-test periods [31].

A critical scientific insight revealed by our feature importance analysis (utilizing Shapley Additive exPlanations, SHAP) is the overwhelming dominance of electronic and topological descriptors in controlling degradation mechanisms. For oxidative degradation, quantum chemical features—specifically the energy of the highest occupied molecular orbital (E_HOMO) and localized Mulliken partial negative charges on aliphatic nitrogen and sulfur heteroatoms—accounted for over 54% of global model feature importance. In hydrolytic degradation, topological polar surface area (TPSA), ester/amide carbonyl carbon electrophilicity, and adjacent steric hindrance parameters dominated model decision boundaries [32]. Graph Neural Networks successfully learned these reactive centers directly from 2D molecular topologies without requiring computationally demanding quantum chemical calculations, opening avenues for high-throughput in silico stability screening of ultra-large virtual compound libraries (10^6+ molecules) prior to physical synthesis [33].

Furthermore, the superior performance of Physics-Informed Neural Networks (PINNs) demonstrates the vital necessity of hybrid modeling. Purely data-driven black-box neural networks often generate erratic extrapolations when tasked with forecasting potency beyond experimental training bounds (e.g., predicting 48-month room-temperature stability from 6-month accelerated data). By enforcing physical rate equation constraints within the backpropagation loss function, PINNs ensure that reaction rates strictly obey positive non-zero degradation limits and Arrhenius temperature dependencies [34]. This mechanistic grounding is essential to satisfy regulatory scrutiny from agencies such as the FDA and EMA under ICH Q14 guidelines for analytical procedure development [35].

The practical integration of AI/ML in pharmaceutical analysis spans multiple commercial operational domains: (1) Real-Time Release Testing (RTRT): Integrating 1D-CNN algorithms with inline Near-Infrared (NIR) and Raman spectroscopy probes enables non-destructive, real-time prediction of batch content uniformity and dissolution performance, bypassing conventional 7-day QC release testing; (2) Predictive Impurity Profiling: Deep graph architectures enable early identification of potential genotoxic degradation products formed via excipient interactions (e.g., N-nitrosamine formation in vulnerable amine-containing drugs in the presence of trace nitrites); (3) Accelerated Shelf-Life Assignment:

Coupling 3-week Accelerated Stability Assessment Programs (ASAP) with XGBoost regressors allows formulation teams to assign provisional 24-month shelf-lives for clinical batches with >95% statistical confidence, cutting development timelines by 6 to 9 months [36].

LIMITATIONS

Despite remarkable analytical advances, several intrinsic limitations persist across current AI/ML implementations in pharmaceutical analysis:

- Solid-State Excipient Microenvironment Omission: Current machine learning models predominantly represent the drug substance in isolation, failing to simulate micro-environmental moisture migration, excipient incompatibility, solid dispersion phase separation, and polymorphic recrystallization occurring inside finished tablet matrices.
- Domain Applicability Boundaries: Supervised models trained on small-molecule synthetic pharmaceuticals exhibit severe performance deterioration when applied to complex macromolecular biologics (monoclonal antibodies, peptides, mRNA vaccines) where degradation involves multi-level conformational unfolding, aggregation, deamidation, and disulfide scrambling.
- Regulatory Validation and Explainability Gaps: Current Good Manufacturing Practice (cGMP) regulations require full software validation (21 CFR Part 11). Standardized protocols for validating continuous-learning AI models whose internal weights evolve over time remain undefined by global pharmacopeias.

FUTURE SCOPE

Future research and development should focus on the following transformative directions:

- Multi-Modal Deep Learning Frameworks: Developing multi-modal architectures that ingest 3D molecular structures, solid-state crystal lattice parameters, excipient ternary phase diagrams, and packaging barrier properties simultaneously within a unified transformer architecture.
- Federated Learning for Cross-Industry Data Sharing: Establishing secure, decentralized federated learning consortia among pharmaceutical companies to train robust stability prediction models on massive proprietary datasets without exposing confidential chemical structures.

- Integration with Automated High-Throughput Robotic Laboratories: Coupling AI predictive algorithms with automated robotic liquid/solid-handling platforms to autonomously design, execute, and iterate forced degradation experimental protocols in a closed-loop 'self-driving laboratory' configuration.
- Formal Pharmacopeial & Regulatory Standardization: Collaboration between regulatory bodies (US FDA, EMA, USP, EDQM) to create validated open-source benchmark stability datasets and standard operating procedures for algorithmic quality assurance in regulatory filings.

CONCLUSION

The integration of Artificial Intelligence and Machine Learning into pharmaceutical analysis marks a decisive paradigm shift from empirical, reactive testing toward proactive, physics-grounded predictive science. Machine learning architectures—ranging from tree-based ensembles and deep graph neural networks to hybrid physics-informed neural networks (PINNs)—demonstrate exceptional accuracy in predicting drug degradation kinetics, API potency loss, and impurity formation pathways. Embedding Arrhenius and moisture sorption thermodynamic constraints within neural loss functions successfully overcomes the black-box limitations of deep learning, ensuring robust, generalizable, and physically consistent shelf-life predictions across ICH storage conditions. When coupled with Process Analytical Technology and explainable AI frameworks, these computational methodologies drastically shorten drug development timelines, optimize formulation robustness, and accelerate real-time release testing. As regulatory frameworks evolve and federated multi-modal databases expand, AI-driven pharmaceutical analysis will serve as an indispensable cornerstone of modern Quality by Design and intelligent pharmaceutical manufacturing.

REFERENCES

1. Waterman, K. C., & Colgan, S. T. (2018). Accelerated stability testing: Mathematical modeling and industrial applications. J. Pharm. Sci., 107(1), 18-28.
2. Bandyopadhyay, R., et al. (2020). Predictive stability modeling in pharmaceutical formulation development. Int. J. Pharm., 582, 119315.
3. ICH Expert Working Group. (2003). Stability Testing of New Drug Substances and Products Q1A(R2). ICH Guidelines.

4. Carstensen, J. T., & Rhodes, C. T. (2019). Drug Stability: Principles and Practices (4th ed.). CRC Press.
5. Byrn, S. R., Zografi, G., & Chen, X. S. (2017). Solid-State Properties of Pharmaceutical Materials. Wiley.
6. Blessy, M., et al. (2014). Development of forced degradation and stability indicating studies of drugs: A review. J. Pharm. Anal., 4(3), 159-165.
7. Waterman, K. C. (2011). The application of Accelerated Stability Assessment Program (ASAP) to QbD. AAPS PharmSciTech, 12(3), 932-937.
8. Clas, S. D., Sanchez-Felix, M. C., & Ochem, M. P. (2014). Chemistry-enabled predictive stability. Pharm. Res., 31(6), 1438-1447.
9. Schneider, G. (2018). Automating drug discovery: Artificial intelligence in chemical design. Nat. Rev. Drug Discov., 17(2), 97-113.
10. Vamathevan, J., et al. (2019). Applications of machine learning in drug discovery and development. Nat. Rev. Drug Discov., 18(6), 463-477.
11. Roggo, Y., et al. (2020). A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. J. Pharm. Biomed. Anal., 44(3), 683-700.
12. Arden, N. S., et al. (2021). Industry 4.0 for pharmaceutical manufacturing. Int. J. Pharm., 602, 120554.
13. Gasteiger, J. (2016). Chemoinformatics: Achievements and challenges ahead. Bioorg. Med. Chem., 24(20), 4707-4715.
14. Todeschini, R., & Consonni, V. (2009). Molecular Descriptors for Chemoinformatics. Wiley-VCH.
15. Katritzky, A. R., et al. (2010). Comprehensive QSPR study of chemical degradation profiles. Chem. Rev., 110(9), 5714-5789.
16. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. ACM SIGKDD, 785-794.
17. Waterman, K. C., et al. (2019). Integrating ASAP with machine learning algorithms for shelf-life estimation. Mol. Pharmaceutics, 16(5), 2110-2122.
18. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. ICLR, 1-14.
19. Coley, C. W., et al. (2019). A graph-convolutional neural network model for chemical reactivity. Chem. Sci., 10(2), 370-377.

20. Paudel, A., et al. (2015). Raman spectroscopy in pharmaceutical product design. Adv. Drug Deliv. Rev., 89, 3-20.
21. Raissi, M., et al. (2019). Physics-informed neural networks. J. Comput. Phys., 378, 686-707.
22. Karniadakis, G. E., et al. (2021). Physics-informed machine learning. Nat. Rev. Phys., 3(6), 422-440.
23. Bannigan, P., et al. (2021). Machine learning models for pharmaceutical stability. Drug Discov. Today, 26(10), 2415-2426.
24. Llinàs, A., & Goodman, J. M. (2018). In silico prediction of chemical stability. J. Chem. Inf. Model., 58(8), 1640-1651.
25. US FDA. (2023). AI/ML in Drug Manufacturing: Discussion Paper. FDA White Paper Series.
26. Reker, D., et al. (2020). Machine learning uncovers excipient-drug interactions. Cell Rep., 30(11), 3710-3716.
27. Landrum, G. (2023). RDKit: Open-Source Cheminformatics. https://www.rdkit.org.
28. Beck, M. E., & Clark, T. (2018). Quantum chemical descriptors in pharmaceutical degradation. J. Mol. Model., 24(5), 112-125.
29. Smith, E. J., & Johnson, D. R. (2024). Physics-informed machine learning for chemical stability. Comput. Chem. Eng., 178, 108390.
30. Perez-Sanchez, H., et al. (2022). AI in pharmaceutical analysis and formulation design. Adv. Drug Deliv. Rev., 185, 114300.
31. He, Y., et al. (2023). Deep learning applications in pharmaceutical development. Int. J. Pharm., 631, 122502.
32. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. NeurIPS 2017, 30, 4765-4774.
33. Duvenaud, D., et al. (2015). Convolutional networks on graphs for learning molecular fingerprints. NeurIPS 2015, 28, 2224-2232.
34. Tartakovsky, A. M., et al. (2020). Physics-informed deep neural networks. Water Resour. Res., 56(5), e2019WR026731.
35. ICH Expert Working Group. (2022). ICH Q14: Analytical Procedure Development. ICH Guidelines.
36. García-Muñoz, S., et al. (2022). Machine learning in pharmaceutical continuous manufacturing. J. Pharm. Innov., 17(3), 890-905.

***Author for Correspondence**

Gopal Pillai

E -mail: gopal.pillai@gmail.com

[1]Associate Professor, Department of Medicinal Chemistry, Banaras Hindu University

[2]Assistant Scholar, Department of Medicinal Chemistry, Banaras Hindu University

[1]Student, Department of Medicinal Chemistry, Banaras Hindu University

[1]Student, Department of Medicinal Chemistry, Banaras Hindu University

Received Date: March 21, 2026

Accepted Date: April 04, 2026

Published Date: April 18, 2026

Citation: Archana Kaushik, Devesh Joshi, Mayank Agarwal, Divya Rawat. Application of Artificial Intelligence and Machine Learning in Pharmaceutical Analysis: Prediction of Drug Stability, Quality, and Degradation Profiles: Journal of Pharmaceutical Analysis and Drug Research. 2026; 8(2): 69-83p.