STEM Education International, *e-ISSN: 3124-0755*
Vol. 2, No. 2, December 2026, Pages 119-141

STEM-EI
STEM Education International

From Predictive Accuracy to Experimental Validation: A Systematic Review of Artificial Intelligence in Detergent Formulation and Quality Assessment

Ahmadi[1,2], Aliefman Hakim[1*], Jamaluddin[3], Iwan Sumarlan[4], Aris Doyan[1], AA Sukarso[1]

[1]Doctoral Program in Science Education, University of Mataram, Mataram, Indonesia.
[2]Department of Chemistry Education, Mandalika University of Education Mataram, Mataram, Indonesia
[3]Department of Biology Education, Faculty of Teacher Training and Education, University of Mataram, Mataram, Indonesia.
[4]Department of Chemistry, Faculty of Teacher Training and Education, University of Mataram, Mataram, Indonesia.

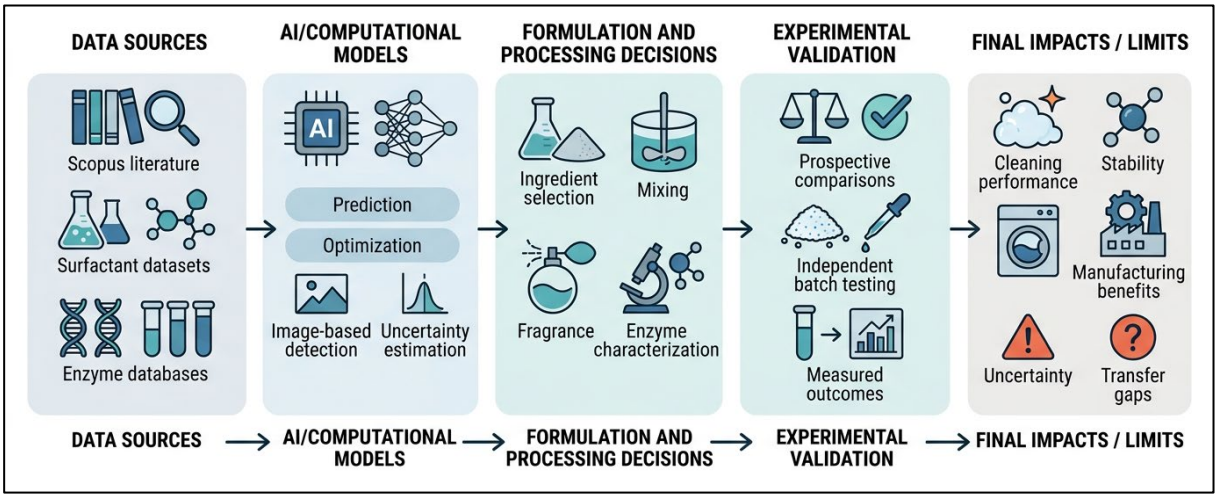
**Corresponding author: aliefman@unram.ac.id*

Abstract: Artificial intelligence supports detergent formulation and quality assessment, but predictive accuracy does not establish experimentally verified product improvement. This review examines how computational predictions inform ingredient selection, formulation, processing, and quality decisions. A Scopus-based systematic search and staged selection, documented in the author-provided PRISMA flow, yielded 37 records for evidence mapping and critical narrative synthesis, distinguishing direct detergent applications, enabling ingredient research, operational and educational contexts, and peripheral methods. Extraction considered computational approaches, data provenance, evaluation designs, experimental outcomes, uncertainty, and applicability; available abstracts constrained appraisal and eligibility judgments. Surfactant studies reported promising property predictions, while enzyme databases supplied supporting infrastructure. Direct applications addressed enzyme characterisation, fragrance modelling, powder mixing, and image-based instability detection. However, characterisation experiments did not consistently constitute prospective validation of model-selected formulations, and agreement with simulations differed from independent physical confirmation. Uncertainty undercoverage showed that favourable prediction metrics can coexist with unreliable confidence estimates. Regression-based optimisation supplied experimental comparisons without necessarily qualifying as AI. Operational savings and teaching-factory applications offered implementation context rather than evidence of improved detergent chemistry. Heterogeneous targets, sampling, and validation

designs prevented meaningful pooling of accuracy measures. Transfer to new formulations, batches, and operating conditions remained incompletely demonstrated, and single-database coverage limited search breadth. The proposed framework connects data, predictions, decisions, and measured outcomes. Progress requires prospective formulation comparisons, independent batch evaluation, calibrated uncertainty, and transparent reporting of reproducible benefits in cleaning, stability, and manufacturing.

Keywords: Artificial intelligence; detergent formulation; quality assessment; experimental validation; surfactant property prediction; enzymatic detergents; uncertainty quantification.

Graphical Abstract



INTRODUCTION

Detergent performance depends on ingredient selection, molecular interactions, processing, and use conditions. Surfactants, enzymes, stabilisers, and fragrances must function together: favourable isolated properties may not ensure successful mixing, storage, or application. Studies of pre-formulations reveal competing cleaning and foaming objectives, while liquid-formulation research identifies physical instability as a commercialisation barrier (Jangir et al., 2023; De Micco et al., 2022). Development therefore requires coordinated assessment of cleaning, stability, compatibility, and manufacturing consistency.

These challenges span molecular and process scales. Surfactant models incorporate temperature and chemical descriptors, whereas enzyme engineering must navigate extensive sequence-modification possibilities (Ali et al., 2026; Qin et al., 2024). Polymer–fragrance interactions affect diffusion, and powder-component proportions influence mixing uniformity (Qi et al., 2025; Cañamero et al., 2023). Ingredient behaviour must consequently be connected to formulation and manufacturing outcomes.

Experiments remain essential, but exploring many compositions and operating conditions consumes resources. Computational methods can prioritise experiments and characterise complex relationships, but the savings must be demonstrated. Rogers et al. (2025) combined neural-network attribution, symbolic regression, and experimental design to reduce model-identification effort in an in silico study.

Resource-driven machine learning in biomass research provides an analogy, not evidence of equivalent detergent benefits (Kovačić et al., 2024).

Here, AI includes learning-based and knowledge-based prediction, classification, and decision support, including neural networks and fuzzy inference. Surfactant applications use wavelet networks, neuro-fuzzy methods, and graph architectures (Tabaraki et al., 2021; Theis Marchan et al., 2026). Bayesian networks support fragrance-related optimisation, while regression studies examine competing formulation objectives (Qi et al., 2025; Jangir et al., 2023). Conventional regression, mechanistic simulation, and AI nevertheless have distinct assumptions and evidentiary roles.

Applications also address manufacturing and quality. Cañamero et al. (2023) generated mixing-prediction data through experimentally checked discrete element simulations. De Micco et al. (2022) classified liquid instability from images, and Silva et al. (2025) combined ANN-assisted amylase characterisation with detergent-compatibility and stain-removal experiments. These endpoints differ. Optimisation selects compositions or conditions against objectives; quality assessment evaluates attributes; active control additionally changes operating conditions in response to information.

The central question is whether accurate predictions yield experimentally verified improvements. A high coefficient of determination or low error indicates agreement for a specified target and dataset, not necessarily better cleaning, shelf life, or manufacturing. Simulator validation and agreement between a surrogate and that simulator are separate steps (Cañamero et al., 2023). Likewise, surfactant-property prediction alone cannot establish performance in unfamiliar multicomponent formulations.

Uncertainty and transferability further limit accuracy-based interpretation. Ali et al. (2026) reported mean test R^2 = 0.848 and scaffold-split R^2 = 0.870, but ensemble intervals achieved 75.5% empirical coverage at a nominal 95%. Calibration therefore requires separate evaluation. Machining research similarly illustrates operating-domain limits, although it supplies only an analogy for detergents (Aljinović et al., 2021). Independent samples, controls, application boundaries, and reproducibility are essential when predictions guide practical decisions.

Use-phase and operational outcomes also need distinct interpretation. Muthusamy et al. (2022) predicted microplastic-fibre release from detergent choice and washing conditions; this environmental endpoint is not cleaning effectiveness. Hospital-laundry cost reductions likewise do not demonstrate improved formulation chemistry (Keswani, 2026). LICEDB addresses fragmented enzyme information and data limitations, but supporting data infrastructure does not itself validate a product (Gong et al., 2025).

Adjacent reviews provide relevant foundations. Qin et al. (2024) examined reductions in enzyme-engineering search spaces, while Kadam et al. (2025) reviewed AI integration in microbial biosurfactant production. These domains overlap with detergent development but differ from tracing predictions through formulation decisions to measured product outcomes. The present synthesis therefore centres on

the strength and application boundaries of that connection rather than algorithm performance alone.

The review evaluates AI support for detergent development and quality decisions through a Scopus-based search and staged selection documented in the author-provided PRISMA flow, followed by critical mapping of 37 retained records. Internal validation uses the originating dataset; external validation uses independently obtained application-relevant data; prospective validation tests predictions in subsequent experiments. Practical applicability means demonstrated usefulness under specified conditions. Direct detergent evidence is distinguished from ingredient resources, implementation contexts, and peripheral methodological parallels.

Four questions guide the review. RQ1: How does AI predict detergent-relevant ingredient properties, and how transferable are the predictions? RQ2: How does AI-assisted formulation and process optimisation produce experimentally verified improvements? RQ3: How reliably does AI assess detergent quality, stability, and performance under relevant conditions? RQ4: Which validation practices and evidence limitations determine practical readiness? The synthesis follows the theoretical framework and methods, distinguishes four evidence themes, and identifies priorities for reproducible experimental testing.

Theoretical Framework and Background

Physicochemical Foundations of Detergent Performance

Detergent performance emerges from interactions among surfactants, enzymes, polymers, builders, and processing conditions. Molecular properties must translate into cleaning, stability, compatibility, and manufacturing consistency. Mesostructure and formulation conditions provide a physicochemical basis for understanding why ingredient-level predictions require confirmation in complete products (Gradzielski et al., 2021; Salager et al., 2022; Phaodee & Weston, 2023).

The critical micelle concentration (CMC) marks the onset of substantial micellar aggregation. Accurate CMC prediction therefore cannot establish detergent effectiveness on its own. Dynamic interfacial behaviour matters during agitation, flow, and rapid interface formation. CMC is therefore treated here as an ingredient-level endpoint rather than a surrogate for cleaning performance.

Enzymatic detergents require catalytic activity under demanding chemical conditions. Isolated enzyme activity cannot replace compatibility and cleaning tests in complete formulations. Beyond molecular interactions, rheology and physical stability determine whether active ingredients remain uniformly distributed and are delivered effectively.

Computational Approaches to Formulation Research

Computational formulation research combines descriptor-based prediction, machine learning, physicochemical simulation, and hybrid models to explore large design spaces and potentially reduce experimental effort. Chemical descriptors capture characteristics such as size, polarity, hydrophobicity, and structure. CMC modelling also

requires attention to chemical coverage and applicability domains (Moriarty et al., 2023).

Machine learning can represent nonlinear relationships without complete mechanisms. Formulation studies still require documented conditions, interactions, and applicability limits. Molecular dynamics and coarse-grained simulations complement prediction by examining aggregation, molecular interactions, compatibility, and interfacial organisation. Hybrid methods integrate physical knowledge with data-driven prediction.

Formulation Optimisation and Quality Engineering

Formulation optimisation is multi-objective because improving one attribute may compromise another. DoE evaluates several variables and their interactions, while response surfaces identify favourable regions rather than isolated settings.

Quality engineering also considers stability, manufacturability, cost, and sustainability. Optimisation must resolve trade-offs among viscosity, phase behaviour, performance, and usability (Alves et al., 2022). Electrolyte-driven changes in aggregation and viscosity may alter foam performance (Seweryn et al., 2021). Practical optimisation consequently requires defined operating windows.

Predictive Validity, Experimental Validity, and Transferability

Industrial usefulness depends on the relevance and independence of validation data, not just numerical accuracy. Splitting and cross-validation estimate internal predictive performance but cannot guarantee transfer to unfamiliar formulations, batches, or conditions.

Uncertainty estimates and applicability-domain analysis can flag predictions that require confirmation. Reliable confidence intervals nevertheless require measured calibration (Ali et al., 2026). Simulation-trained predictions need independent physical checks. AI-assisted development should connect experiments, prediction, formulation decisions, and independent confirmation in an iterative process.

This review's framework distinguishes four linked elements: the data used to build a model, the predicted endpoint, the formulation or process decision supported by that endpoint, and independently measured outcomes. Three checkpoints concern data independence, uncertainty calibration, and experimental confirmation. This interpretive structure is proposed by the review; it is not a validated causal model. Figure 2, presented with the synthesis, illustrates these proposed validation checkpoints.

METHOD

Review Design and Documentation

The review combines systematic searching and staged selection with structured evidence mapping and critical narrative synthesis. The recorded search and selection pathway is reproduced in Figure 1, using the completed pathway documented in the

review records. The final bibliographic registry is the author-supplied SLR_MR_MACK.csv, containing 37 records. The reporting structure draws on PRISMA 2020 (Page et al., 2021). Because the retained registry spans chemical, operational, educational, and peripheral domains, the 37 records are not treated as 37 direct demonstrations of AI detergent effectiveness.

The four research questions guide interpretation of the retained material. Direct and closely related applications, enabling ingredient research, operational and educational contexts, and peripheral methodological parallels are coded separately. Background resources and peripheral records are retained for evidence mapping and scope analysis, but excluded from claims about direct chemical effectiveness. Targeted checks of publisher or bibliographic records support the revised interpretation. No prospective protocol registration, independent dual screening, or formal study-level risk-of-bias appraisal is claimed, because these procedures are not documented in the available review records.

Information Source and Search Strategy

Scopus was the bibliographic source recorded in the original selection diagram. The Boolean strategy combined AI terms with detergent and teaching-factory or learning-factory terms. The displayed expression consolidates the search and filtering constraints; the intermediate counts in Figure 1 describe the author-recorded selection pathway rather than separate searches using the already fully restricted expression:

TITLE-ABS-KEY(("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network*" OR "expert system*" OR "fuzzy logic" OR "computer vision" OR "generative AI" OR "large language model*") AND (detergent* OR "teaching factory" OR "teaching factories" OR "learning factory" OR "learning factories")) AND SRCTYPE(j) AND DOCTYPE(ar) AND LANGUAGE(english) AND (LIMIT-TO (SUBJAREA,"CHEM") OR LIMIT-TO (SUBJAREA,"ENGI") OR LIMIT-TO (SUBJAREA,"CENG") OR LIMIT-TO (SUBJAREA,"SOCI") OR LIMIT-TO (SUBJAREA,"BUSI") OR LIMIT-TO (SUBJAREA,"PHAR") OR LIMIT-TO (SUBJAREA,"COMP")) AND PUBYEAR > 2019 AND PUBYEAR < 2027.

The consolidated constraints concern publication years 2020–2026, English-language articles, journal sources, and the listed subject areas. The teaching-factory terms intentionally broadened retrieval to implementation environments; their presence does not make those records direct chemical evidence. Surfactant and enzyme applications were interpreted according to their explicit relevance to detergent-related tasks. The absence of dedicated ingredient terms remains a search-sensitivity limitation. No geographical restriction is specified. The exact search execution date and counts by individual Boolean clause are not recorded in the supplied diagram or identifier registry and are therefore not reconstructed.

PRISMA-S emphasizes reporting sources, platforms, dates, expressions, restrictions, and retrieval counts (Rethlefsen et al., 2021). Here, the original PRISMA flow documents aggregate search and selection counts, and the CSV supports the identities of the final retained records. These sources document the aggregate selection pathway and final bibliographic identities. Query-execution details, individual exclusion reasons, and

independent screening decisions are not separately available in the retained documentation.

Relevance Criteria and Evidence Categories

For direct chemical interpretation, relevance requires an identifiable computational task concerning detergent ingredients, formulation, processing, or product-quality assessment. Ingredient studies require an explicit detergent-related property or design problem. Conventional regression and mechanistic simulation are classified by their actual methods and retained as closely related comparisons where appropriate; a machine-learning label in a title does not by itself establish an incremental AI benefit. Database articles and reviews provide enabling infrastructure or background rather than primary demonstrations of improved detergent performance.

Demand forecasting, facility planning, education, environmental fibre release, and general manufacturing are classified as contextual evidence. Biomedical detergent-like mechanisms and analytical detergent-fibre fractions are peripheral to cleaning-product formulation. The final 37-record registry includes these domains, so inclusion in the map is explicitly separated from eligibility for direct detergent-effectiveness conclusions. The four thematic groups contain 6, 6, 13, and 12 records, respectively. They are relevance categories, not a hierarchy of study quality or a set of equally informative product experiments.

Screening Flow and Registry Reconciliation

The original flow records 207 Scopus records, with zero duplicate removals, zero automation exclusions, and zero other removals before screening. Of 207 screened records, 50 were excluded by the publication-year step, leaving 157 reports sought for retrieval. The diagram records zero unretrieved reports and 73 reports assessed for eligibility after the intervening article, language, and subject-area filtering pathway. The final included count is 37. The reduction from 157 to 73 is 84 by subtraction, but the original record does not separate that reduction into independently documented exclusion reasons.

The original diagram lists 37 out-of-scope exclusions at the 73-report eligibility stage, which is inconsistent with its final included count of 37. The revised Figure 1 preserves the assessed and included totals and shows 36 nonincluded records as an arithmetic reconciliation (73 − 37). This is not a newly verified count by exclusion reason. The CSV independently corroborates 37 retained record identities: 36 unique nonempty DOIs and 37 unique Scopus identifiers, with one DOI field blank. The blank DOI belongs to Dheerawat et al. (2023), identified by Scopus record 85187202082.

Extraction and Narrative Synthesis

Extraction summarises each record's application, computational method or evidence type, reported outcome, validation basis, and interpretation boundary. Bibliographic identities were reconciled using the 36 non-empty DOI values and the remaining Scopus identifier. Quantitative findings in the direct and ingredient tables were checked against accessible publisher records where available. NR-A denotes information not reported in the available abstract-derived material; it does not

establish that a procedure was absent from the original study. Metrics retain their target, validation context, and interpretation limits. The identifier CSV is not treated as an abstract or full-text dataset.

Narrative synthesis compares decision targets and validation designs rather than pooling heterogeneous accuracy values. Ingredient-property prediction, surrogate agreement with simulation, formulation experiments, stability classification, and operational or educational outcomes answer different questions. Section 5 interprets these differences through the Section 2 framework. It identifies what the mapped findings support, what remains uncertain, and what prospective comparisons would establish useful product improvements. Formal risk-of-bias ratings and a completed extraction appendix are not claimed.

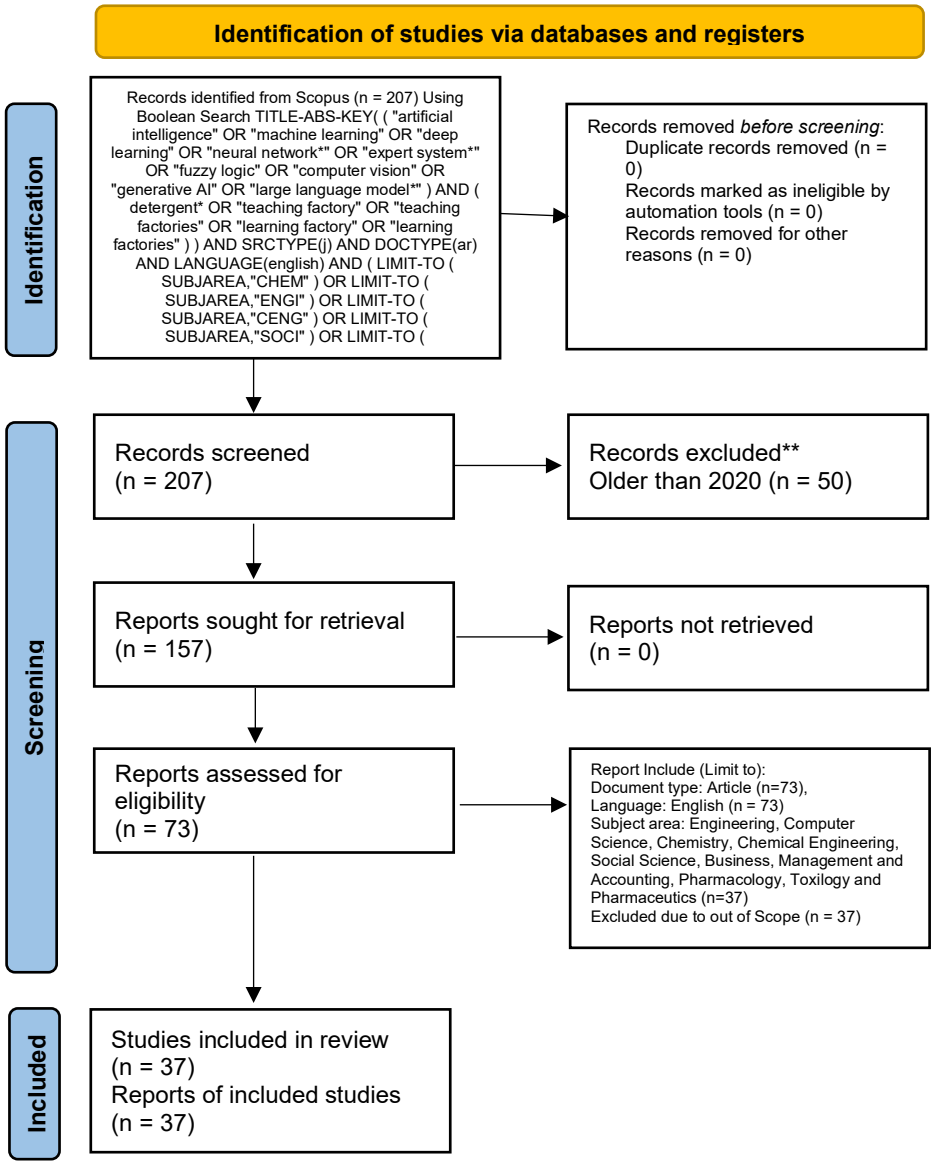


Figure 1. PRISMA-based search and selection flow reconstructed from the author-confirmed original diagram.

REVIEW OF THEMES AND FINDINGS

The final registry comprises 37 distinct Scopus records, all matched to the four evidence tables. Six records concern direct or closely related detergent tasks, six concern ingredient prediction or infrastructure, thirteen concern implementation contexts, and twelve concern peripheral methodological parallels. These counts describe the retained corpus and do not estimate prevalence across the entire field. Reviews, databases, conventional optimisation, and cross-domain applications are explicitly distinguished from direct AI product-effectiveness evidence.

Direct and Closely Related Detergent Applications

Table 1 distinguishes six direct or closely related applications. The studies address ingredient activity, formulation trade-offs, fragrance behaviour, powder mixing, visual instability, and computational model identification. These targets differ in their proximity to measured finished-product outcomes and should not be pooled as equivalent demonstrations of AI effectiveness.

Table 1. Direct and closely related applications and their validation boundaries

Study	Target	Method or evidence	Mapped finding	Interpretation boundary
Silva et al. (2025)	Amylase characterisation and detergent additive testing	ANN-assisted characterisation; fermentation and detergent experiments	Amylase activity was 83.90 U/gds after 3 days; ANN characterised an optimum pH of 9 and 20 °C. Detergent compatibility and starch-stain tests were reported.	Physical application testing is relevant; prospective ANN-selected formulation improvement is not established by the available abstract.
Jangir et al. (2023)	Detergent pre-formulation optimisation	Linear, interaction and quadratic regression	Regression optima: RMI30 72.32 mL and cleaning time 17.67 s; rank-1 response pairs 88.9 mL/20 s and 81.2 mL/14 s. The source reports experimental validation.	Regression-based optimisation is a comparator; the reported optimum alone does not isolate an AI advantage.
Qi et al. (2025)	Fragranced detergent design	Molecular dynamics and Bayesian neural networks	Polymer–fragrance interactions and diffusion inform formulation design.	Mechanistic and computational guidance; independent finished-product improvement requires physical confirmation.
Cañamero et al. (2023)	Powder-detergent mixing	Experimentally checked DEM followed by ML prediction	25 simulated compositions at 45 rpm reached mixing indices 0.98–0.99 at 60 s. ML was evaluated against DEM outputs after experimental checking of DEM.	Physical simulator validation and ML-to-simulator agreement are separate; neither alone establishes cleaning improvement.
De Micco et al. (2022)	Liquid-formulation instability detection	Deep-learning image classification	Expert-labelled vial images support detection of visible instability.	Supports quality assessment; shelf-life prediction and batch transfer need separate evaluation.
Rogers et al. (2025)	Model identification and strategic sampling	ANN attribution, symbolic regression and experimental design	In-silico recovery of governing equations; a threefold reduction in required experiments was reported relative to unguided symbolic regression.	Computational efficiency evidence; physical detergent benefits are not established.

Ingredient Prediction and Supporting Infrastructure

Table 2 separates predictive ingredient studies from enzyme databases, reviews, and fermentation research. Ingredient relevance alone does not establish an AI contribution or product-level benefit. NR-A throughout the tables refers to information unavailable in the abstract-derived material, not a confirmed absence of a method.

Table 2. Ingredient research and infrastructure supporting detergent development.

Study	Target	Method or evidence	Mapped finding	Interpretation boundary
Tabaraki et al. (2021)	CMC of anionic gemini surfactants	MLR, PLS, wavelet networks and ANFIS	Descriptor-based modelling of 85 surfactants.	Ingredient screening; independent full-formulation performance is not established.
Theis Marchan et al. (2026)	Surfactant CMC prediction	Graph architectures, clustering and SHAP analysis	PharmHGT LogCMC R^2 = 0.943–0.947; molecular dynamics provided thermodynamic corroboration.	Chemical coverage, splitting and prospective measurements require full-text appraisal; product performance remains a separate endpoint.
Ali et al. (2026)	Six surfactant properties	Multi-task graph neural network and ensemble uncertainty	Mean held-out R^2 = 0.848; scaffold split 1,299/325, mean R^2 = 0.870; ensemble coverage 75.5% at nominal 95%.	Structural transfer was tested with zero scaffold overlap. Interval undercoverage limits confidence reliability; prospective product improvement remains a separate question.
Gong et al. (2025)	LICEDB enzyme resource	Database development and standardized retrieval	Infrastructure for enzyme information and design.	A resource, not an experimental demonstration of AI-selected detergent improvement.
Qin et al. (2024)	Enzyme engineering	Review of hotspot and library-size reduction	Strategies for narrowing enzyme-engineering searches.	Background synthesis rather than a new detergent-validation experiment.
Cardoso et al. (2020)	Biosurfactant production	Fermentation experiments	Ingredient-production evidence.	Supports experimental background; no identifiable AI contribution or finished-detergent validation is established here.

Operational and Educational Contexts

Table 3 maps implementation contexts. Operational, environmental, and educational outcomes may support future experimental systems, but they do not estimate detergent-formulation effectiveness. Their inclusion is explanatory, not evidence that these records satisfy primary chemical eligibility.

Table 3. Implementation contexts and limits of relevance to detergent chemistry

Study	Target	Method or evidence	Mapped finding	Interpretation boundary
Muthusamy et al. (2022)	Microplastic-fibre release during washing	ANFIS and washing experiments	Prediction of an environmental use-phase endpoint.	Fibre release is distinct from cleaning efficacy and formulation optimisation.
Dheerawat et al. (2023)	Washing-machine operation	Mamdani fuzzy control	Expert-derived control rules; physical-trial details NR-A.	Control-system context; experimentally verified cleaning or savings cannot be inferred.
Keswani (2026)	Production planning and hospital laundry	Fuzzy inference and ML decision support	Operational resource and cost outcomes reported for a hospital-laundry case.	Operational savings do not establish changed detergent chemistry or maintained cleaning quality.
Ghazouani et al. (2024)	Detergent demand forecasting	CNN–LSTM	Demand forecasting for inventory and production.	Forecast performance is not detergent-quality evidence.
Yudin et al. (2024)	Waste recognition	Detection and weakly supervised segmentation	Industrial image recognition.	Discarded-product recognition is distinct from formulation quality.
Schweitzer et al. (2023)	Automated vehicle scheduling	ML selection of scheduling strategies	Computational evaluation and a field experiment.	Transfer lesson concerns logistics, not detergent efficacy.
Aljinović et al. (2021)	Machining in a learning factory	ANN and response surfaces	Surface-roughness and power prediction.	Illustrates operating-domain constraints in another process.
Orozco et al. (2024)	Desktop learning factory	Image classification and Arduino hardware	Accessible instructional prototype.	No detergent-specific sensing or product validation demonstrated here.
Ramadhani Fajri et al. (2025)	Vocational electronics education	AI, extended reality and simulations	Educational outcomes in a combined intervention.	Learning outcomes do not isolate AI effects on chemistry.
Wahjusaputri et al. (2024)	Vocational teaching factories	Qualitative development and validity appraisal	Organisational and instructional development.	No independently evaluated AI model for detergent is specified in the available material.
Maksum et al. (2026)	Automotive vocational education	Case-based learning with AI, IoT and immersive technology	Development and quasi-experimental educational evaluation.	Combined instructional intervention; detergent product outcomes are outside scope.
Kangru et al. (2024)	Manufacturing laboratory redesign	Case study and automation simulation	Laboratory and learning-factory integration.	Infrastructure context rather than detergent-validation evidence.
Angrisani et al. (2020)	Instrumentation education	Project-based learning and IoT	Measurement education in a learning-enterprise setting.	Supports measurement literacy; educational evaluation is not chemical validation.

Peripheral Records and Scope Boundaries

Table 4 makes retrieval boundaries explicit. Keyword overlap with detergents can arise from analytical fibre fractions, biological membrane disruption, or unrelated applications. These records are not counted as direct detergent evidence; their methodological parallels remain bounded to the original research domains.

Table 4. Peripheral records and bounded methodological parallels

Study	Target	Method or evidence	Mapped finding	Interpretation boundary
Wang et al. (2024)	Chromium-contaminated soil washing	Multiple ML models and additional experiments	Prediction followed by additional experimental evaluation.	Useful validation analogy; soil remediation is outside cleaning-product formulation.
Hodyna et al. (2025)	Antifungal lysosomotropic compounds	QSAR, synthesis and biological assays	Computational analysis linked with physical and biological tests.	Biomedical detergent terminology does not indicate evidence of household formulation.
Li et al. (2025)	Antimicrobial peptide–membrane interactions	Gradient boosting and Raman spectroscopy	Complementary analytical evidence for membrane disruption.	Focus on biological mechanisms rather than detergent product development.
Azimikia et al. (2021)	CO2 absorption by nanofluids	Experiments and neural networks	Response modelling of gas absorption.	Cross-domain experimental-design analogy only.
Kimothi et al. (2026)	Lipophilic drug delivery	Review of formulation and performance	Pharmaceutical formulation background.	No AI detergent-validation experiment established.
Cano-Suñén et al. (2023)	Building performance	IoT and neural-network proof of concept	Real-world sensing and monitoring infrastructure.	Building outcomes cannot establish detergent manufacturing performance.
Wahyudi et al. (2026)	Lignocellulosic feedstock biodegradation	NIR spectroscopy and ML	Endpoint-dependent validation outcomes.	Detergent fibre denotes analytical fractions, not cleaning formulations.
Chinmayi et al. (2025)	Legume nutrition	Nearest-neighbour, SVM and random forest	Sensitivity to training and testing splits.	Agricultural prediction; methodological parallel only.
Fernández-Habas et al. (2022)	Pasture quality	PLS, random forest and feature selection	Spectroscopic nutrient prediction.	Analytical detergent-fibre measurements are outside product scope.
Zhang et al. (2026)	Grass–clover composition	Stacking and spectral preprocessing	Endpoint-dependent agricultural prediction.	No cleaning-product outcome.
Barnetson et al. (2020)	Pasture biomass and quality	UAV sensing and modelling	Field-scale predictive evaluation.	Transfer and heterogeneity analogy only.
Kovačić et al. (2024)	Agricultural-residue pretreatment	Ensemble ML	Process-variable modelling of lignin-related endpoints.	Detergent lignin is an analytical biomass endpoint, not detergent efficacy.

DISCUSSION

From Ingredient Prediction to Formulation Performance

The central synthesis is that evidence strength depends on how a prediction is used and independently checked, not simply on the sophistication of its algorithm. The

records in Tables 1 and 2 frequently target properties or process proxies rather than cleaning, storage stability, or product consistency. A model may be useful at this intermediate stage without demonstrating finished-product improvement. Accordingly, the review answers RQ1 conditionally: AI can support ingredient screening within the evaluated domain, but translation to a multicomponent detergent remains an additional empirical question.

Tabaraki et al. (2021), Theis Marchan et al. (2026), and Ali et al. (2026) illustrate different representations of surfactant properties. Their reported metrics cannot establish a ranking across studies unless targets, transformations, chemical coverage, and evaluation samples are comparable. Moriarty et al. (2023) provide a relevant previous-study comparison: graph-based prediction depends on balanced training coverage, and applicability-domain analysis helps identify unreliable molecular predictions. This supports interpreting newer architectures in relation to data coverage and a suitable baseline. Theis Marchan et al. (2026) report PharmHGT R^2 = 0.943–0.947 with molecular-dynamics corroboration. That evidence complements, but does not replace, prospective measurement of a model-selected formulation.

The physicochemical framework explains this boundary. Mesostructure and formulation conditions affect surfactant systems (Gradzielski et al., 2021; Salager et al., 2022), while HLD-based formulation provides an established chemistry-oriented perspective on composition choices (Phaodee & Weston, 2023). Electrolytes and aggregation also influence viscosity and functionality (Alves et al., 2022; Seweryn et al., 2021). These studies support treating a predicted molecular property as one input to formulation design. The review's inference is that validation should progressively add the interactions omitted from an ingredient model, rather than treating an accurate CMC prediction as proof of detergent efficacy.

What Experimental Evidence Establishes About Optimisation

The studies in Table 1 support several stages of development but do not establish the same causal claim. Silva et al. (2025) connect ANN-assisted enzyme characterisation to detergent-compatibility and starch-stain experiments. This is stronger application evidence than an enzyme-property prediction alone. However, the available abstract does not demonstrate a prospective comparison in which ANN-selected formulations outperform an otherwise equivalent conventional selection strategy. The appropriate interpretation is experimentally supported additive potential, with the incremental benefit of AI still requiring a design that isolates the selection method.

Jangir et al. (2023) report regression-based optimisation of competing pre-formulation responses. This provides a useful comparison because an experimentally useful optimisation procedure need not require a neural network. The study's terminology should not decide its evidentiary classification: the underlying method and validation determine whether it represents conventional statistical modelling, an AI approach, or a comparator. For RQ2, the stronger test is whether model-guided candidate selection improves a prespecified outcome relative to an appropriate baseline under equal experimental budgets.

Cañamero et al. (2023) link experimentally checked discrete-element modelling to ML prediction of detergent mixing. Their 25 simulated compositions reached mixing indices of 0.98–0.99 after 60 s at 45 rpm. This creates two distinct validation questions: whether the simulator represents physical mixing, and whether its surrogate reproduces simulated responses. A further physical test is needed to establish performance of an ML-selected operating condition on new batches. Qi et al. (2025) similarly contribute molecular guidance for fragranced formulations, while Rogers et al. (2025) report a threefold reduction in required in-silico experiments relative to unguided symbolic regression. Taken together, these studies suggest a productive role for computational prioritisation, but resource savings and product gains should be measured rather than inferred from computational speed or simulator agreement.

An informative prospective design would compare an AI-selected formulation with a conventional DoE or expert-selected comparator, holding ingredient constraints and testing budgets constant. Cleaning, viscosity, stability, and compatibility should be assessed on independently prepared batches, with assay conditions and uncertainty reported. This is a recommendation from the synthesis, not an experiment completed by the present review. It would distinguish useful prediction from useful selection and identify whether a gain in one attribute is purchased at the expense of another.

Quality Assessment Requires Endpoint-Specific Validation

De Micco et al. (2022) demonstrate a direct quality-assessment task using expert-labelled images of formulated liquids. Detecting visible instability can inform screening or monitoring, but it is not automatically a prediction of future shelf life. A present-state label and a time-to-failure endpoint require different reference measurements. The review therefore interprets this study as evidence for image-based assessment within its tested setting, while leaving long-term prognosis and production transfer as separate questions.

For RQ3, the relevant validation unit is often the independently prepared formulation or batch rather than the image. Multiple photographs from one vial can share formulation, lighting, and ageing information. An evaluation that partitions images without separating the underlying samples could overestimate transfer; this is a methodological risk to check in full texts, not an accusation that the mapped study used an invalid split. Batch-level testing, representative storage conditions, sensitivity to instability, and false-alarm behaviour would make a classifier's practical role clearer.

Mixing uniformity and visible instability are also distinct quality dimensions. Cañamero et al. (2023) concern process homogeneity, whereas De Micco et al. (2022) concern liquid stability assessment. Neither endpoint alone establishes cleaning efficacy, enzyme survival, or fragrance delivery. A quality system may combine these measurements, but the validity of each component and of the resulting release decision should be evaluated separately. This endpoint-specific interpretation prevents a successful classifier or surrogate from being generalised into an unsupported claim of comprehensive product quality.

Uncertainty Transferability and Decision Risk

Ali et al. (2026) provide an especially informative contrast between prediction, structural transfer, and uncertainty. They report mean held-out R^2 = 0.848 and mean scaffold-split R^2 = 0.870 using 1,299 training and 325 test molecules with zero scaffold overlap. Structural transfer is therefore supported within that evaluation. Yet their 60-model ensemble attained only 75.5% empirical coverage at a nominal 95%, demonstrating overconfidence in the reported intervals. These results support a qualified conclusion: promising accuracy and structural generalisation can coexist with inadequate calibration, and neither establishes finished-detergent improvement.

Moriarty et al. (2023) offer a complementary perspective by connecting molecular representations with uncertainty and applicability-domain analysis. Together, these sources motivate checking where a model is reliable as well as how accurate it is on average. A formulation decision near a viscosity or stability constraint can be sensitive to modest errors, even when aggregate accuracy appears favourable. Conversely, a less complex model may be sufficient for a screening decision with a broad acceptable range. The proposed implication is that tolerance should be tied to the actual decision and independently measured assay variability.

Transferability also has several dimensions: new molecular structures, mixtures, ingredient suppliers, preparation batches, temperatures, and equipment. Success in one dimension does not guarantee success in another. For RQ4, grouped chemical splits can test structural transfer, whereas independent preparation batches and prospective application tests address different sources of variation. Calibration should be evaluated in the intended domain, and failed transfers should be reported. The map does not support a numerical estimate of how often such failures occur, because the extraction is incomplete and the validation designs are heterogeneous.

Implementation and Sustainability Without Scope Inflation

Table 3 shows that useful implementation outcomes extend beyond detergent chemistry. Muthusamy et al. (2022) address fibre release during washing, and Ghazouani et al. (2024) address demand forecasting. These tasks may matter to environmental management or production planning, but neither demonstrates that an AI-selected formula cleans better. Operational claims in Keswani (2026) likewise require interpretation against their own cost and resource endpoints. A reduction in operational expenditure cannot be translated into chemical effectiveness without evidence that cleaning and hygiene requirements were maintained.

Learning-factory studies provide another implementation perspective. Orozco et al. (2024), Wahjusaputri et al. (2024), and Kangru et al. (2024) describe instructional or organisational environments that could support experimentation. Angrisani et al. (2020) add a measurement-education perspective. The review infers that such environments may help develop the skills needed to collect reproducible data and evaluate model-guided decisions. The mapped educational outcomes, however, do not establish detergent benefits or validate the chemical models discussed in Tables 1 and 2.

Sustainability should similarly be evaluated through explicit trade-offs. Ingredient production, environmental release, washing conditions, and manufacturing resources belong to related but distinct parts of a product system. An improvement in a single proxy does not establish a net environmental benefit. Future studies could report cleaning performance alongside ingredient dose, water and energy use, and relevant environmental endpoints under comparable conditions. This extends the review's decision framework without treating contextual or peripheral records as primary evidence of a sustainable AI formulation.

Contribution of the Review and Research Priorities

The review contributes a decision-oriented synthesis of the supplied records. Its four relevance categories distinguish direct applications from enabling resources, implementation contexts, and peripheral analogies. Within each category, the data–prediction–decision–outcome framework separates the object predicted from the benefit claimed. This is the synthesis added by the review: diverse computational successes can be interpreted consistently by asking whether the evaluated evidence reaches the decision and product endpoint at issue. It is an organizing contribution rather than a new experimentally verified mechanism or a claim of field-wide prevalence.

The framework identifies three verification checkpoints. First, data and evaluation units should be sufficiently independent and representative of the intended application. Second, uncertainty should be assessed against observed outcomes where it informs selection or acceptance. Third, model-supported choices should be tested through independently measured outcomes and relevant comparators. These checkpoints complement ingredient and process studies by specifying the additional evidence needed to support stronger claims. They do not replace domain expertise, conventional DoE, or physicochemical interpretation.

The most useful research priority is a prospective formulation comparison that documents the initial data, candidate-selection rule, constraints, baseline strategy, and independently measured product outcomes. A second priority is evaluation across preparation batches and realistic operating or storage conditions. A third is transparent reporting of uncertain, unsuccessful, or non-transferable predictions. Enzyme resources such as LICEDB can enable future work (Gong et al., 2025), and enzyme-engineering and biosurfactant reviews identify related design opportunities (Qin et al., 2024; Kadam et al., 2025). Their contribution remains enabling until a specific design is tested.

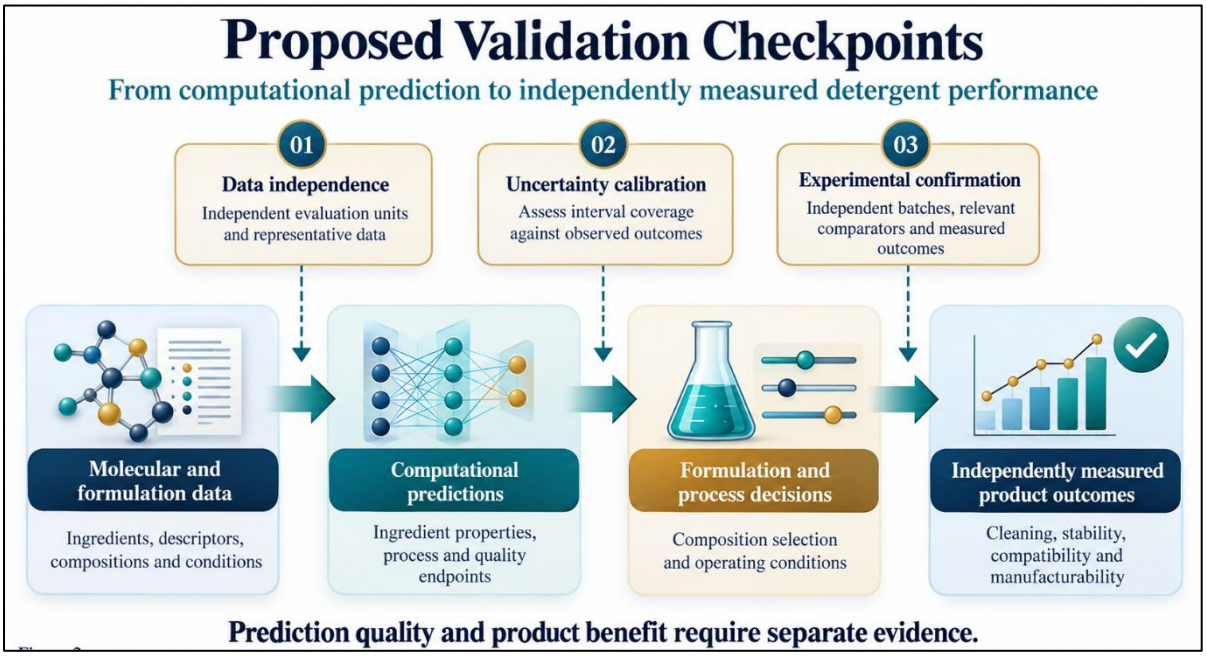


Figure 2. Conceptual framework connecting data, predictions, decisions, and measured outcomes. The checkpoints are interpretive appraisal questions proposed by this review.

Limitations and Strengths of the Conclusions

The PRISMA flow documents the search and selection pathway, and the final registry corroborates 37 distinct Scopus identifiers. This provides traceability of the final evidence corpus. Nevertheless, the search uses one database and mixes detergent chemistry with teaching-factory contexts, while dedicated surfactant and enzyme terms are absent. The exact search date, the 84-record intermediate reduction by specific reason, and individual full-text exclusion decisions are not available in the supplied documentation. PRISMA 2020 and PRISMA-S guide reporting, but these remaining omissions prevent a claim of complete reporting compliance (Page et al., 2021; Rethlefsen et al., 2021).

The CSV corroborates bibliographic identities but contains only DOI and Scopus-link fields. It cannot independently establish sample sizes, assay conditions, or validation designs. Abstract-based extraction and targeted source checks therefore limit assessment of data independence, baseline comparisons, and prospective experiments. NR-A marks unavailable information rather than confirmed methodological weakness. Full-text extraction and formal study-level appraisal remain incomplete in the available materials. Heterogeneous targets and datasets preclude meaningful pooling of accuracy, and the 37 retained records are not 37 direct AI detergent-effectiveness studies.

The conclusions are bounded to the documented 37-record corpus: several predictive and analytical tasks are supported, but independently confirmed benefits from AI-selected finished formulations are not consistently demonstrated in the available evidence. This is a judgment about the retained material, not proof that AI is ineffective

or that prospective validation is absent from the field. A broader chemistry-focused search, complete full-text extraction, and independent appraisal could strengthen or change the assessment. The framework remains an explicit contribution for interpreting heterogeneous evidence and identifying the next validation step.

CONCLUSIONS

AI can support detergent-relevant ingredient prediction, formulation exploration, process assessment, and instability detection. The 37-record map does not establish that these predictive gains consistently produce independently verified improvements in finished detergents. Ingredient accuracy, simulator agreement, product testing, and implementation outcomes require separate interpretation. Relevant conventional modelling and experimental comparators remain essential for determining the incremental benefit of AI.

The review contributes four relevance categories and a framework connecting data, predictions, decisions, and measured outcomes. Its proposed checkpoints concern independent and representative evaluation, uncertainty calibration, and experimental confirmation of model-supported choices. Prospective comparisons across independently prepared batches are the priority for stronger claims about cleaning, stability, compatibility, and manufacturability. Conclusions remain limited by single-database coverage, incomplete extraction, heterogeneous endpoints, and gaps in the detailed search and screening documentation.

ACKNOWLEDGEMENT

Thanks to the University of Mataram and the researchers in the Research program for the intensive discussion

CONFLICT OF INTEREST

The authors declared that they have no competing interests.

DATA AVAILABILITY

The search and selection pathway is reported in Figure 1, and the final 37-record bibliographic registry is supplied as SLR_MR_MACK.csv. Appendix A links each retained record to its DOI or Scopus identifier and evidence category. The CSV contains identifiers rather than experimental datasets or a complete full-text extraction archive.

AUTHOR CONTRIBUTION

All authors contributed to the interpretation of results, reviewed the manuscript critically, approved the final version, and agreed to be accountable for all aspects of the work.

USE OF GENERATIVE ARTIFICIAL INTELLIGENCE

Generative AI assisted manuscript restructuring, language editing, and development of the interpretive synthesis. No new experimental results were generated. Search counts were transcribed from the authors' original PRISMA flow, with arithmetic reconciliation explicitly identified. Numerical values are attributed to the cited sources; the authors remain responsible for source verification, interpretation, and journal-policy compliance.

REFERENCES

Ali, A.-F. A. N., Huang, H., & AL-Bakhrani, A. A. (2026). Multi-task graph neural networks for comprehensive surfactant property prediction. Digital Discovery, 5(9), 3702–3715. https://doi.org/10.1039/d6dd00135a

Aljinović, A., Bilić, B., Gjeldum, N., & Mladineo, M. (2021). Prediction of surface roughness and power in turning process using response surface method and ANN. Tehnicki Vjesnik, 28(2), 456–464. https://doi.org/10.17559/TV-20190522104029

Alves, A. A., Sousa, F. J., & Sebastião, M. (2022). Influence of electrolytes on the structural and viscosity properties of mixed anionic–nonionic–zwitterionic surfactants in detergent formulations. Journal of Surfactants and Detergents, 25(6), 775–787. https://doi.org/10.1002/jsde.12613

Angrisani, L., Arpaia, P., Bonavolontá, F., Moccaldi, N., & Schiano Lo Moriello, R. (2020). A "learning small enterprise" networked with a FabLab: An academic course 4.0 in instrumentation and measurement. Measurement: Journal of the International Measurement Confederation, 150. https://doi.org/10.1016/j.measurement.2019.107063

Azimikia, R., Mazaheri, H., Joshaghani, A. H., & Joudaki, E. (2021). Presentation of an experimental model of CO2 gas absorption by spraying method by NiO nanofluids and carbon nanotube in the presence magnetic field. European Physical Journal Plus, 136(10). https://doi.org/10.1140/epjp/s13360-021-01972-1

Barnetson, J., Phinn, S., & Scarth, P. (2020). Estimating plant pasture biomass and quality from UAV imaging across Queensland's rangelands. AgriEngineering, 2(4), 523–543. https://doi.org/10.3390/agriengineering2040035

Cañamero, F. J., Doraisingam, A. R., & Álvarez-Leal, M. (2023). Mixing performance prediction of detergent mixing process based on the discrete element method and machine learning. Applied Sciences, 13(10). https://doi.org/10.3390/app13106094

Cano-Suñén, E., Martínez, I., Fernández, Á., Zalba, B., & Casas, R. (2023). Internet of things (IoT) in buildings: A learning factory. Sustainability, 15(16). https://doi.org/10.3390/su151612219

Cardoso, S. L., Dantas, R., Costa, C. S. D., Campos, E. S., & Tambourgi, E. B. (2020). Evaluation of the medium fermentative parameters for the production of biosurfactant using the Candida gloebosa. Chemical Engineering Transactions, 79, 475–480. https://doi.org/10.3303/CET2079080

Chinmayi, H. K., Flynn, K. C., Baath, G. S., Gowda, P., Northup, B., & Ashworth, A. (2025). Monitoring legume nutrition with machine learning: The impact of splits in training and testing data. Applied Soft Computing, 176. https://doi.org/10.1016/j.asoc.2025.113186

De Micco, M., Gragnaniello, D., Zonfrilli, F., Guida, V., Villone, M. M., Poggi, G., & Verdoliva, L. (2022). Stability assessment of liquid formulations: A deep learning approach. Chemical Engineering Science, 262. https://doi.org/10.1016/j.ces.2022.117991

Dheerawat, K., Pirzada, U. S. M., & Kataria, H. R. (2023). Design of economical fuzzy logic controller for washing machine. Journal of Multiple-Valued Logic and Soft Computing, 41(1–2), 209–222. https://www.scopus.com/pages/publications/85187202082?origin=resultslist

Fernández-Habas, J., Carriere Cañada, M., García Moreno, A. M., Leal-Murillo, J. R., González-Dugo, M. P., Abellanas Oar, B., Gómez-Giráldez, P. J., & Fernández-Rebollo, P. (2022). Estimating pasture quality of Mediterranean grasslands using hyperspectral narrow bands from field spectroscopy by random forest and PLS regressions. Computers and Electronics in Agriculture, 192. https://doi.org/10.1016/j.compag.2021.106614

Ghazouani, I., Masmoudi, I., Mejri, I., & Layeb, S. B. (2024). A CNN–LSTM hybrid deep learning model for detergent products demand forecasting: A case study. International Journal of Supply and Operations Management, 11(4), 417–429. https://doi.org/10.22034/ijsom.2024.109931.2752

Gong, L., Liu, F., Zhang, C., Ming, Y., Mou, Y., Yuan, Z., Jiang, H., Gao, B., Lu, F., & Zhang, L. (2025). LICEDB: Light industrial core enzyme database for industrial applications and AI enzyme design. Database, 2025. https://doi.org/10.1093/database/baaf001

Gradzielski, M., Duvail, M., Molina, P. M. d., Simon, M., Talmon, Y., & Zemb, T. (2021). Using microemulsions: Formulation based on knowledge of their mesostructure. Chemical Reviews, 121(10), 5671–5740. https://doi.org/10.1021/acs.chemrev.0c00812

Hodyna, D., Kovalishyn, V., Shulha, Y., Trokhimenko, O., Aksenovska, O., Rogalsky, S., & Metelytsia, L. (2025). Antifungal activity and cytotoxicity of imidazole-and morpholine-based lysosomotropic detergents. Innovative Biosystems and Bioengineering, 9(1), 26–44. https://doi.org/10.20535/IBB.2025.9.1.315523

Jangir, K., Gour, A., Suniya, N. K., Meena, S. K., & Parihar, K. (2023). Multi-objective optimization of detergent pre-formulations using machine learning techniques. Journal of the Indian Chemical Society, 100(1). https://doi.org/10.1016/j.jics.2022.100815

Kadam, V., Kusal, S., Patil, S., & Singh, P. (2025). Advances and challenges in the integration of artificial intelligence in microbial biosurfactant bioprocess. AMB Express, 15(1). https://doi.org/10.1186/s13568-025-01988-1

Kangru, T., Moor, M., Rudz, A., & Jänes, A. (2024). Reconstructing a manufacturing laboratory to a learning factory: A TTK UAS case study. Journal of Machine Engineering, 24(3), 133–140. https://doi.org/10.36897/jme/191243

Keswani, M. (2026). An intelligent decision-making system integrating fuzzy logic and machine learning for pharmaceutical closed-loop supply chain and production planning in an uncertain environment. International Journal of Pharmaceutical and Healthcare Marketing, 1–21. https://doi.org/10.1108/IJPHM-05-2025-0078

Kimothi, S., Ramtirthkar, C. R., Garg, V., Ghewari, P., & Durgawale, P. (2026). Lipophilic drug delivery systems: Formulation and performance. International Journal of Drug Delivery Technology, 16, 119–127. https://doi.org/10.25258/ijddt.16.119-127

Kovačić, Đ., Radočaj, D., Samac, D., & Jurišić, M. (2024). Influence of thermal pretreatment on lignin destabilization in harvest residues: An ensemble machine learning approach. AgriEngineering, 6(1), 171–184. https://doi.org/10.3390/agriengineering6010011

Li, C., Zhou, Q.-Q., Yu, T., Zhang, X., Kang, Z., Zhou, M., Wang, Y., Zou, X., Wang, Z., Guo, J., Li, Q., Lu, G., Xie, H., & Guo, J. (2025). Molecular-level understanding of membrane disruption behaviors of antimicrobial peptides by gradient boosting algorithm-assisted raman spectroscopy. Journal of Chemical Information and Modeling, 65(16), 8603–8613. https://doi.org/10.1021/acs.jcim.5c01392

Maksum, H., Simatupang, W., Irfan, D., & Hasan, H. (2026). Bridging the industry 4.0 skills gap: Development of case-based learning model in AI-powered teaching factory for automotive vocational education. International Journal of Interactive Mobile Technologies, 20(10), 112–127. https://doi.org/10.3991/ijim.v20i10.61583

Moriarty, A., Kobayashi, T., Salvalaglio, M., Angeli, P., Striolo, A., & McRobbie, I. (2023). Analyzing the accuracy of critical micelle concentration predictions using deep learning. Journal of Chemical Theory and Computation, 19(20), 7371–7386. https://doi.org/10.1021/acs.jctc.3c00868

Muthusamy, L. P., Periyasamy, A. P., Militky, J., & Thangasamy, P. R. (2022). Adaptive neuro-fuzzy inference system to predict the release of microplastic fibers during domestic washing. Journal of Testing and Evaluation, 50(1). https://doi.org/10.1520/JTE20210175

Orozco, E., Cárdenas, P. C., López, J. A., & Rodriguez, C. K. (2024). Low-cost desktop learning factory to support the teaching of artificial intelligence. HardwareX, 18. https://doi.org/10.1016/j.ohx.2024.e00528

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D.,

Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, Article n71. https://doi.org/10.1136/bmj.n71

Phaodee, P., & Weston, J. S. (2023). Review: Implementing the hydrophilic–lipophilic deviation model when formulating detergents and other surfactant-related applications. Journal of Surfactants and Detergents, 26(3), 277–286. https://doi.org/10.1002/jsde.12660

Qi, Y., Zhao, L., Tang, H., Zhang, L., & Gani, R. (2025). Computer aided formulation design based on molecular dynamics simulation: Detergents with fragrance. Computers and Chemical Engineering, 192. https://doi.org/10.1016/j.compchemeng.2024.108919

Qin, Z., Yuan, B., Qu, G., & Sun, Z. (2024). Rational enzyme design by reducing the number of hotspots and library size. Chemical Communications, 60(76), 10451–10463. https://doi.org/10.1039/d4cc01394h

Ramadhani Fajri, B., Mohd Kamaruzaman, F., Omar, M., Lofandri, W., & Samala, A. D. (2025). Smart teaching factory: Integrating extended reality, artificial intelligence, and animated simulations for transformative vocational education. Salud, Ciencia y Tecnologia, 5. https://doi.org/10.56294/saludcyt20251769

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. Systematic Reviews, 10, Article 39. https://doi.org/10.1186/s13643-020-01542-z

Rogers, A. W., Lane, A., Mendoza, C., Watson, S., Kowalski, A., Martin, P., & Zhang, D. (2025). Integrating feature attribution and symbolic regression for automatic model structure identification and strategic sampling. Computers and Chemical Engineering, 197. https://doi.org/10.1016/j.compchemeng.2025.109036

Salager, J., Márquez, R., Bullón, J., & Forgiarini, A. (2022). Formulation in surfactant systems: From-Winsor-to-HLDN. Encyclopedia, 2(2), 778–839. https://doi.org/10.3390/encyclopedia2020054

Schweitzer, F., Bitsch, G., & Louw, L. (2023). Choosing solution strategies for scheduling automated guided vehicles in production using machine learning. Applied Sciences, 13(2). https://doi.org/10.3390/app13020806

Seweryn, A., Wasilewski, T., & Bocho–Janiszewska, A. (2021). Correlations between the type of aggregates in the bulk phase and the functionality and safety of all-purpose cleaners. International Journal of Molecular Sciences, 22(12), 6592. https://doi.org/10.3390/ijms22126592

Silva, A. S., Neves, K. M., Freitas, R. F., das Chagas, T. P., Salay, L. C., da Silva, E. G. P., Uetanabaro, A. P. T., & da Costa, A. M. (2025). Valorization of cocoa and peach-palm wastes for the production of amylases by Pleurotus pulmonarius CCB19 and its application as an additive in commercial detergents. Applied Biochemistry and Biotechnology, 197(5), 3271–3294. https://doi.org/10.1007/s12010-024-05147-3

Tabaraki, R., Khodabakhshi, M., & Fatahi, G. (2021). Quantitative structure-critical micelle concentration modeling of anionic gemini surfactants, comparison of MLR, PLS, WNN, and ANFIS models with eigenvalue and correlation ranking methods. Journal of the Iranian Chemical Society, 18(10), 2703–2711. https://doi.org/10.1007/s13738-021-02225-9

Theis Marchan, G., Balogun, T. O., Territo, K., Das, D., Olayiwola, T., Kumar, R., & Romagnoli, J. A. (2026). Toward harnessing AI for surfactant chemistry: Prediction of critical micelle concentration. Computational Materials Science, 265. https://doi.org/10.1016/j.commatsci.2026.114548

Wahjusaputri, S., Nastiti, T. I., & Sukmawati, W. (2024). Development of artificial intelligence-based teaching factory in vocational high schools in Central Java province. Journal of Education and Learning, 18(4), 1234–1245. https://doi.org/10.11591/edulearn.v18i4.21422

Wahyudi, I., Munawar, A. A., Kaloudis, E., Samsudin, A. A., & Wajizah, S. (2026). Machine learning-enhanced NIR spectroscopy for intelligent monitoring of lignocellulosic feedstuffs biodegradation by lignocellulolytic fungi. International Journal of Design and Nature and Ecodynamics, 21(7), 1957–1969. https://doi.org/10.18280/ijdne.210710

Wang, J., Zhang, H., Wang, X., Liu, X., & Deng, H. (2024). Modeling rapidly discriminative strategies of Cr contaminated soils through machine learning. Journal of Environmental Chemical Engineering, 12(3). https://doi.org/10.1016/j.jece.2024.112921

Yudin, D., Zakharenko, N., Smetanin, A., Filonov, R., Kichik, M., Kuznetsov, V., Larichev, D., Gudov, E., Budennyy, S., & Panov, A. (2024). Hierarchical waste detection with weakly supervised segmentation in images from recycling plants. Engineering Applications of Artificial Intelligence, 128. https://doi.org/10.1016/j.engappai.2023.107542

Zhang, S., Gnyp, M. L., Xin, P., Junklewitz, P., Jasper, J., Kalaitzidis, C., & Manevski, K. (2026). Estimating grass-clover mixtures quality using field spectroradiometry and stacking-based ensemble learning models. Computers and Electronics in Agriculture, 251. https://doi.org/10.1016/j.compag.2026.112081