IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2026

DOI: 10.17148/IARJSET.2026.131043

Artificial Intelligence and Machine Learning for Predicting Concrete Performance: A Critical Review of Modified Mixes, Durability, Structural Response and Sustainable Material Systems

Srushti Vijay Doke[1], Anand B. Kudoli*[2]

Department of Civil Engineering, Pimpri Chinchwad College of Engineering and Research

(affiliated to Savitribai Phule Pune University), Ravet, Pune 412101, Maharashtra, India[1,2]

*Corresponding Author

Abstract: The current research field of machine learning (ML) uses prediction models for concrete property assessment but these models show different levels of accuracy in their results. The review analyzes ML performance for concrete through four progressive levels of difficulty which include compressive-strength prediction and modified mixture prediction for fibres and polymers and waste materials and bacterial agents and freeze–thaw prediction and durability assessment and blast-structural response surrogate modeling. The review evaluates each task by assessing available data sets together with needed input characteristics and chosen algorithms and methods which assess model performance. The field of strength prediction stands as the most developed area because artificial neural networks and tree ensembles produce excellent results on benchmark datasets through their ability to perform interpolation on known mixtures. The available datasets do not include material descriptors which should contain information about fibre geometry and polymer chemistry and waste fineness and composition for creating sustainable modified mixtures. The prediction of durability faces obstacles because exposure data remains limited while showing different patterns and structural surrogate models draw their assumptions from the simulations which trained them. The research shows three main weaknesses which include limited dataset size and tests which split related specimens randomly and scientists who depend only on determination coefficient values. The study presents a ten-step process together with a reporting checklist which focuses on validation grouping and uncertainty reporting and verification through physical experiments. ML operates best as a screening system which supports decision-making processes through its ability to enhance engineering judgment and testing methods without replacing them.

Keywords: Machine learning; Artificial neural networks; Concrete strength prediction; Durability prediction; Explainable AI; Surrogate modelling; Sustainable concrete

1. INTRODUCTION

The performance of concrete depends on various elements which include binder composition and water–binder ratio and aggregate grading and admixtures and curing and age and exposure environment for durability assessment. The empirical equations show basic connections between these elements yet they fail to handle the complex nonlinear relationships which emerge in current concrete mixtures that include supplementary materials and fibers and chemical admixtures. The machine-learning methods directly extract these relationships from data which enables them to generate quick performance estimates that decrease the amount of trial batches needed for target achievement [1, 2]. Optimization algorithms work together with these systems to find material combinations which achieve the right balance between strength requirements and financial expenses and carbon footprint reduction [3].

The number of ML studies on concrete has grown rapidly, and reviews now catalogue dozens of algorithms applied to strength, workability and durability [2, 4]. Yet high reported accuracy does not always mean a model will perform well on a new mixture or in another laboratory. Young et al. demonstrated, using a large industrial dataset, that the information contained in mixture proportions alone limits how accurately strength can be predicted, regardless of the algorithm [5]. The fundamental understanding of these restrictions enables proper ML system handling because the technology functions effectively under suitable conditions.

The review follows a structure based on prediction tasks instead of grouping content according to specific algorithms. Section 2 presents the approach which serves as the foundation for this study. Section 3 presents the ML workflow

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2026

DOI: 10.17148/IARJSET.2026.131043

together with the primary algorithm families which form the basis of this study. The review analyzes four different sections between 4 and 7 which study strength prediction and sustainable mixture modifications and freeze-thaw prediction and structural-response surrogate models. The research paper explains interpretability in Section 8 while Section 9 identifies typical research flaws and offers a workflow for best practices and Section 10 establishes the main research directions.

2. REVIEW APPROACH

The research team conducted a critical narrative review to collect their necessary information. The research team accessed Google Scholar and Scopus databases to find journal articles and reviews which they retrieved through their search for machine learning and artificial neural network and random forest and gradient boosting and support vector and SHAP and compressive strength and durability and freeze–thaw and chloride and blast and surrogate model related documents. The research team selected studies which presented their data and validation procedures and research restrictions through proper documentation. The research included sources which contained complete bibliographic information together with their digital object identifiers. Studies report different accuracy levels because their data sets and evaluation metrics vary between each other.

3. THE ML WORKFLOW AND ALGORITHM FAMILIES

3.1 Workflow

A machine learning research project requires scientists to establish prediction goals while they collect their data which they then process for cleanliness before they select input variables and train their selected algorithms and validate their trained models and perform experimental validation of their predictions. The workflow summary in Figure 1 shows that most published research fails at two essential stages which involve data validation methods and experimental verification of results.

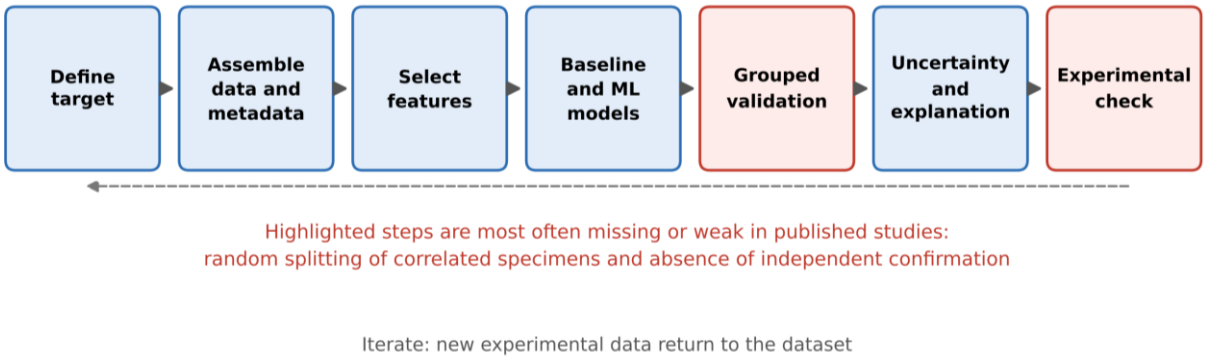


Figure 1. Machine-learning workflow for concrete performance prediction, with the steps most often neglected in published studies highlighted.

3.2 Algorithm families

Table 1 compares the algorithm families most often used for concrete prediction. No family is consistently best; performance depends on dataset size, feature types and the degree of nonlinearity [2, 4].

Table 1. Algorithm families used for concrete performance prediction.

Family	Strengths	Limitations	Typical use
Linear and polynomial regression	Transparent, few parameters	Misses complex interactions	Baselines, small datasets
Artificial neural networks	Capture nonlinear interactions	Need tuning and larger datasets; opaque	Strength of complex mixtures [1]
Support vector regression	Effective on small–medium datasets	Kernel and hyperparameter sensitive	Strength, workability

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2026

DOI: 10.17148/IARJSET.2026.131043

Family	Strengths	Limitations	Typical use
Decision-tree ensembles (random forest, gradient boosting)	Strong on tabular data; robust to scaling	Can overfit small datasets; extrapolate poorly	Strength and durability [6]
Gaussian process regression	Provides uncertainty estimates	Scales poorly to large datasets	Small experimental datasets
Hybrid and optimised models	Can improve accuracy	Risk of over-tuning to test data	Research comparisons

4. COMPRESSIVE-STRENGTH PREDICTION

Yeh's early work showed that artificial neural networks could predict the strength of high-performance concrete more accurately than regression, using mixture proportions and age as inputs [1]. The dataset he assembled became a benchmark, and many later studies have applied increasingly complex algorithms to it. Chou et al. compared several data-mining techniques on strength data and found that an ensemble of regression trees gave the best accuracy among the methods tested, ahead of single neural networks and linear regression [7]. Ben Chaabene et al. reviewed this literature and found that ensemble methods and neural networks generally outperform single models, but that many studies reuse the same public datasets and validate on random splits [2]. Nunez et al. reached similar conclusions in a systematic review of computational intelligence for modern concrete mixtures and called for larger, more diverse datasets [4]. A review focused on self-compacting concrete reported comparable trends, with artificial neural networks and hybrid optimisation methods the most common [8].

The review by Kudoli and Bala compared traditional empirical models with ML models for predicting the physical and mechanical properties of cement and cement mortar, and concluded that ML models can capture nonlinear relationships that empirical equations miss, provided that the training data adequately represent the materials and conditions of interest [9]. This caveat is critical. The industrial study by Young et al. showed that, even with thousands of mixtures, prediction errors remain substantial because batch-to-batch variation in materials and testing is not captured by mixture proportions [5]. High accuracy on a benchmark therefore indicates good interpolation, not universal predictive power.

5. MODIFIED AND SUSTAINABLE MIXTURES

5.1 The feature problem

The existing dataset lacks various elements which become apparent when researchers establish altered combination models. The model fails to learn about any variable which does not show up in the input data. Table 2 lists descriptors that are needed for each type of modification but are often missing.

Table 2. Input descriptors needed for modified and sustainable concrete.

Modification	Essential descriptors	Commonly missing
Fibres	Type, volume fraction, length, aspect ratio, tensile strength	Surface treatment, dispersion quality
Polymers	Polymer type, polymer–cement ratio, solids content	Curing regime (dry/wet), product chemistry
Waste marble and fillers	Replacement basis (binder or aggregate), fineness, $CaCO_3$ content	Source identity, particle-size distribution
Recycled aggregate	Replacement level, water absorption, density	Adhered mortar content, pre-treatment
Bacterial agents	Species, spore dosage, nutrient, carrier	Viability, crack width at testing

5.2 Fibres, polymers and waste materials

Fibre effects are strongest in tensile, flexural and toughness properties, which are less often recorded than compressive strength [10]. A model trained only on compressive strength therefore captures the least important effect of fibres. Polymer-modified mixtures depend on chemistry and curing, so a single "polymer dosage" input may conflate products with different behaviour. Waste marble may replace cement or fine aggregate, and the two routes act through different mechanisms; Sathe et al. tested both routes and reported mechanical, durability and microstructural data that show why replacement basis should be encoded explicitly [11]. For sustainable mixtures in general, source characterisation data

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2026

DOI: 10.17148/IARJSET.2026.131043

such as chemical composition and fineness are needed to avoid models that learn dataset-specific correlations rather than material behaviour [12].

5.3 Bacterial self-healing

Bacterial concrete introduces biological variables—species, viability, nutrient and carrier—as well as crack width and moisture during healing. Guddad et al. examined the mechanical and chemical-durability performance of *Bacillus subtilis*-based concrete [13]. Datasets of this kind are small and heterogeneous, so simple, physically informed models and careful standardisation of data collection are more appropriate at present than high-capacity algorithms that would overfit.

6. DURABILITY AND FREEZE–THAW PREDICTION

The prediction of durability stands as a more complex task than strength prediction because results depend on both time progression and environmental exposure. Taffese and Sistonen conducted a review of ML applications for assessing reinforced concrete durability and service life but they found that limited data availability together with varying testing approaches presented major challenges [14]. The targets which include chloride diffusion coefficient and carbonation depth and mass loss and relative dynamic modulus represent separate measurements that require specific definitions before combining them together.Freeze–thaw prediction illustrates these challenges. Frost resistance depends on the degree of saturation, the air-void system, pore structure and the test protocol, as well as on mixture composition [15]. Vishwanath et al. combined experimental freeze–thaw testing of sustainable concrete with AI/ML-based prediction, showing how models can complement lengthy exposure tests [16]. For such models to generalise, inputs must include exposure variables—saturation procedure, temperature range, number of cycles—and validation must hold out whole mixtures, since repeated measurements of the same specimen at successive cycles are strongly correlated.

7. STRUCTURAL-RESPONSE SURROGATES

ML can also approximate structural response. The time-based structural response to blast and impact loads receives solution through explicit-dynamics finite-element analysis although the process demands extensive computational resources. The work of Gudadappanavar et al. involved using explicit dynamics to investigate how distance from target and explosive mass and blast angle affected the behavior of reinforced concrete frames during TNT explosions [17]. The process of parametric campaigns produces training data which surrogate models use to generate response predictions within seconds. The ML model which Almustafa and Nehdi developed predicts how RC slabs will react when exposed to blast loading [18].

Surrogates inherit every assumption of the simulations used to train them, including constitutive models, strain-rate effects and boundary conditions. They should be validated against independent simulations and, where possible, experiments, and they should not be used outside the range of parameters in their training data. A surrogate trained on frames of one geometry cannot be assumed to apply to another.

8. INTERPRETABILITY

Engineers face difficulties when model interpretation fails to provide the necessary understanding which they require. The research paper [6] presents tree-ensemble methods which produce feature-importance measures and SHAP uses Shapley-value methods to distribute prediction values across all input features by following a standardized procedure [19]. These tools help to check whether a model has learned physically sensible relationships—for example, that strength increases with age and decreases with water–binder ratio. The authors present their model but they fail to describe how causal relationships emerge because input variables which correlate with each other will generate different levels of significance that cannot be predicted. The quality of explanations depends on the same data distribution which generates their explanations. The research work by Rudin supports the use of interpretable models which operate independently from black-box models for vital decision processes because these models deliver similar performance results [20]. Physics-informed machine learning provides an additional solution which incorporates physical laws and restrictions into the model to generate acceptable predictions for areas which exist beyond the most populated training data points [21].

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 10, October 2026

DOI: 10.17148/IARJSET.2026.131043

9. METHODOLOGICAL WEAKNESSES AND A BEST-PRACTICE WORKFLOW

The published concrete ML studies show multiple common weaknesses because their datasets remain small while their models show high complexity and their random train–test splitting method results in specimen and mixture duplicates appearing across both data sets and their hyperparameter tuning process targets the test data and their physical unit error measurements remain unknown because they depend solely on R^2 and their results lack experimental validation and uncertainty measurements [2, 4, 5]. The model performance metrics increase because of these methods but they fail to prove model effectiveness when used with different data sets. The issue extends past concrete because Roberts et al. demonstrated that random cross-validation produces inaccurate prediction error estimates for structured ecological data which contains grouped observations so they suggested blocked or grouped cross-validation approaches [22].

Table 3 proposes a minimum reporting checklist that addresses these issues.

Table 3. Minimum reporting checklist for concrete ML studies.

Item	What to report
1. Target	Property, test standard, specimen size, age
2. Data source	Number of mixtures and specimens; laboratories; public or private
3. Inputs	All features with units; material characterisation; exposure variables
4. Pre-processing	Missing data treatment, outlier removal, scaling
5. Baseline	Performance of a simple regression model
6. Splitting	Grouped by mixture, source or study; split proportions
7. Tuning	Procedure and data used (training only)
8. Metrics	MAE and RMSE in physical units, R^2, residual plots
9. Uncertainty	Prediction intervals; domain of applicability
10. Confirmation	Independent experimental or simulation check

The aim is not to find a universally superior algorithm but to build a dependable tool for a defined mixture family, exposure or structural problem. Model simplicity, interpretability and the cost of errors should influence algorithm choice as much as accuracy.

10. RESEARCH PRIORITIES AND CONCLUSIONS

The organization pursues four main priorities which include creating shared datasets that contain detailed material characterization and exposure information. The organization needs to establish validation procedures together with uncertainty reporting systems which will become their standard operating procedure. The models which contain physical knowledge must maintain the patterns which science has already discovered and the established boundaries which exist. The organization needs to verify the proposed mixtures through independent experimental testing which will confirm the model-based recommendations

The main conclusions are as follows.

1. Strength prediction is the most mature application, but high accuracy on benchmark data mainly reflects interpolation; mixture proportions alone limit achievable accuracy.
2. Modified and sustainable mixtures need material descriptors—fibre geometry, polymer chemistry, replacement basis and waste characterisation—that most datasets lack.
3. Durability and freeze–thaw prediction require consistent targets, exposure variables and validation that holds out whole mixtures.
4. Structural surrogates can accelerate blast-response screening but inherit simulation assumptions and must not be extrapolated.
5. Interpretability tools help check physical plausibility but do not establish causality.
6. With grouped validation, uncertainty reporting and experimental confirmation, ML can become a reliable decision-support tool for concrete design.

CRediT authorship contribution statement

Srushti Vijay Doke: Investigation, Data curation, Visualization, Writing – original draft. **Anand B. Kudoli:** Conceptualization, Methodology, Supervision, Validation, Writing – review and editing.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2026.131043

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

No new data were generated in this review. All information discussed is available in the cited literature.

Declaration of generative AI use

During the preparation of this work, the authors used ChatGPT solely for language editing and grammar checking. After using this tool, the authors reviewed and revised the content as necessary and take full responsibility for the content of the publication.

REFERENCES

[1]. I.-C. Yeh, Modeling of strength of high-performance concrete using artificial neural networks, Cem. Concr. Res. 28 (12) (1998) 1797–1808. https://doi.org/10.1016/S0008-8846(98)00165-3

[2]. W. Ben Chaabene, M. Flah, M.L. Nehdi, Machine learning prediction of mechanical properties of concrete: critical review, Constr. Build. Mater. 260 (2020) 119889. https://doi.org/10.1016/j.conbuildmat.2020.119889

[3]. M.A. DeRousseau, J.R. Kasprzyk, W.V. Srubar, Computational design optimization of concrete mixtures: a review, Cem. Concr. Res. 109 (2018) 42–53. https://doi.org/10.1016/j.cemconres.2018.04.007

[4]. I. Nunez, A. Marani, M. Flah, M.L. Nehdi, Estimating compressive strength of modern concrete mixtures using computational intelligence: a systematic review, Constr. Build. Mater. 310 (2021) 125279. https://doi.org/10.1016/j.conbuildmat.2021.125279

[5]. B.A. Young, A. Hall, L. Pilon, P. Gupta, G. Sant, Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods, Cem. Concr. Res. 115 (2019) 379–388. https://doi.org/10.1016/j.cemconres.2018.09.006

[6]. L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32. https://doi.org/10.1023/A:1010933404324

[7]. J.-S. Chou, C.-K. Chiu, M. Farfoura, I. Al-Taharwa, Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques, J. Comput. Civ. Eng. 25 (3) (2011) 242–253. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000088

[8]. S. Terlumun, M.E. Onyia, F.O. Okafor, Predicting the compressive strength of self-compacting concrete using artificial intelligence techniques: a review, AI Civ. Eng. 3 (1) (2024) 14. https://doi.org/10.1007/s43503-024-00029-3

[9]. A.B. Kudoli, A. Bala, A comprehensive review of machine learning and traditional models for predicting the physical and mechanical properties of cement and cement mortar, J. Mines Met. Fuels 73 (2025) 2923–2931. https://doi.org/10.18311/jmmf/2025/49408

[10]. A.B. Kudoli, R.S. Patil, M.M. Yeole, Recent advances in polypropylene fibre-reinforced concrete: a systematic review on mechanical characteristics and durability performance, J. Mines Met. Fuels 73 (2025) 2887–2896. https://doi.org/10.18311/jmmf/2025/49411

[11]. S. Sathe, M. Sonawane, M. Yeole, S. Shekokar, P.D. Gunaware, Effects of waste marble as partial replacement of cement and fine aggregate on the mechanical, durability and microscopic properties of concrete, Mag. Concr. Res. 78 (17–18) (2026) 1064–1088. https://doi.org/10.1680/jmacr.25.00443

[12]. O. Turk, S. Yehia, A. Abdelfatah, M. Elchalakani, Sustainable concrete production: the potential of utilizing recycled waste materials, J. Build. Eng. 98 (2024) 111467. https://doi.org/10.1016/j.jobe.2024.111467

[13]. M.V. Guddad, M. Selvakumar, G.M. Basavanagowda, M.M. Yeole, D.B. Firake, Y. Dibdalli, P. Sunagar, Microbially induced calcite precipitation (MICP)-based bio-mediated self-repairing concrete incorporating Bacillus subtilis: mechanical performance and chemical durability enhancement, SSRG Int. J. Civ. Eng. 13 (8) (2026) 362–374. https://doi.org/10.14445/23488352/IJCE-V13I8P122

[14]. W.Z. Taffese, E. Sistonen, Machine learning for durability and service-life assessment of reinforced concrete structures: recent advances and future directions, Autom. Constr. 77 (2017) 1–14. https://doi.org/10.1016/j.autcon.2017.01.016

[15]. M.M. Abbas, R. Muntean, The effectiveness of different additives on concrete's freeze–thaw durability: a review, Materials 18 (5) (2025) 978. https://doi.org/10.3390/ma18050978

[16]. K.N. Vishwanath, M.M. Yeole, A. Ananthi, D. Parameswari, S.T. Ramya, M. Vijayakumar, Freeze–thaw resistance of sustainable concrete: experimental investigation, damage mechanisms, and AI/ML-based performance prediction, Int. J. Artif. Intell. Mach. Learn. 6 (13s) (2026) 218–230. https://doi.org/10.51483/IJAIML.6.13s.2026.218-230

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2026.131043

[17]. B. Gudadappanavar, M. Selvakumar, M.M. Yeole, V.A. Hosur, Y. Dibdalli, C. Morales-Verdejo, P. Sunagar, Explicit-dynamics assessment of reinforced concrete frame response under TNT blast: stand-off, charge weight and incidence angle effects, SSRG Int. J. Civ. Eng. 13 (3) (2026) 255–273. https://doi.org/10.14445/23488352/IJCE-V13I3P119

[18]. M.K. Almustafa, M.L. Nehdi, Machine learning model for predicting structural response of RC slabs exposed to blast loading, Eng. Struct. 221 (2020) 111109. https://doi.org/10.1016/j.engstruct.2020.111109

[19]. S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (1) (2020) 56–67. https://doi.org/10.1038/s42256-019-0138-9

[20]. C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nat. Mach. Intell. 1 (5) (2019) 206–215. https://doi.org/10.1038/s42256-019-0048-x

[21]. G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nat. Rev. Phys. 3 (6) (2021) 422–440. https://doi.org/10.1038/s42254-021-00314-5

[22]. D.R. Roberts, V. Bahn, S. Ciuti, M.S. Boyce, J. Elith, G. Guillera-Arroita, et al., Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography 40 (8) (2017) 913–929. https://doi.org/10.1111/ecog.02881

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License