ORIGINAL RESEARCH **Open access**

Improving Predictive Accuracy and Model Interpretability through Data-Information Integrated Neural Networks

Priya Nambiar[1*], Suresh Menon[1], Kavita Pillai[2]

Abstract

Formulating a machine learning framework that simultaneously achieves strong predictive power and heightened transparency constitutes a core obstacle hindering the broader deployment of artificial intelligence across diverse practical settings. This paper proposes a Data Information Integrated Neural Network (DINN) methodology that weaves the correlational structure resident in the dataset directly into the model-building process. The forecasting accuracy of DINN is measured alongside that of a baseline artificial neural network (ANN). The proposed DINN approach is evaluated on two energy-domain case studies: the energy-efficiency cooling (ENC) and energy-efficiency heating (ENH) attributes of buildings, alongside the power output from an industrial gas turbine rated at 365 MW. With respect to ENC, DINN registers a reduced mean RMSE on the test partitions (RMSE_test = 1.23%) compared with the ANN counterpart (RMSE_test = 1.41%). In parallel, DINN formulations consistently delivered enhanced predictive fidelity when capturing the response variables across both case investigations. An input perturbation examination utilizing Gaussian-distributed noise generation reveals that the variable importance ordering produced by DINN aligns more coherently with the operational domain knowledge governing gas turbine power production. This investigation showcases the prospective merit of harnessing latent data structure for constructing high-performing models paired with strengthened interpretive capacity, thus paving the way for sector-wide uptake and further machine learning utilization opportunities.

Keywords Explainable AI, Model interpretation, Scientific machine learning, Artificial neural network, Loss function

*Correspondence:
Priya Nambiar
priya.nambiar@iisc.ac.in
[1] Department of Intelligent Energy Engineering, Faculty of Engineering, Indian Institute of Science, Bengaluru, India
[2] Department of AI for Smart Grid Optimization, Faculty of Engineering, Anna University, Chennai, India

Introduction

The rapid progress in information and communication technologies (ICT) has fundamentally altered the manner in which data is produced, preserved, and retrieved [1]. Escalating computational resources together with refined hardware functionalities have permitted the extensive utilization of accessible data for empirically driven system representation across myriad application spheres, giving rise to what is recognized as the fourth investigative paradigm, commonly labelled Data-Driven Science, succeeding Computational Science (through 2010), Theoretical Science (through 1950), and Empirical Science (through 1600). Technologies rooted in artificial intelligence, alongside machine learning (ML) techniques, stand as the central catalysts propelling developments within Data-Driven Science [2]. Expansive language models falling within the Generative Artificial Intelligence umbrella, along with their contemporary instantiation known as ChatGPT, carry significant implications for reshaping how information retrieval and task-oriented consumption occur, underscoring the practical relevance of language-based models in everyday contexts [3].

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Sophisticated ML modelling procedures can craft potent functional mappings involving high-dimensional input domains, yielding predictive precisions that markedly outstrip those attainable via low-dimensional regression schemes when confronting nonlinear, intricate function landscapes [4]. The inner predictive logic of ML architectures remains fundamentally opaque, and deciphering the rationale behind model outputs represents a pivotal impediment to large-scale rollouts of ML systems throughout industrial arenas encompassing chemical processing, petroleum and gas, and electricity generation sectors. Ordinarily, mechanistic understanding is either disregarded or proves exceptionally challenging to embed during ML model construction [5]; as a consequence, stakeholders regard these models with suspicion concerning their suitability for industrial settings where even negligible oversights could trigger disastrous outcomes [5].

The academic sphere has formulated strategies for elucidating model predictions, categorizing them under Ante-hoc and Post-hoc taxonomies [6]. ML constructs belonging to the Ante-hoc category are inherently transparent by virtue of their design yet suffer from constrained modelling fidelity [7]. In contrast, post-hoc explanatory methodologies are deployed after training opaque models to discern the relative weight or influence of predictor variables on response-variable forecasts [8]. Given that opaque models furnish elevated modelling fidelity relative to intrinsically transparent and low-order regression alternatives, the research community remains invested in fusing accessible knowledge streams throughout the model-building lifecycle to advance the predictive capability and interpretability of ML frameworks simultaneously.

Physics-informed neural networks (PINNs) undergo training under an altered objective function that has been supplemented with governing physical relationships and fundamental laws [9]. Consequently, throughout PINN formulation, parameter adjustments proceed only after conformity with stipulated constraints has been ensured [10]. Nonetheless, across numerous real-world scenarios where first-principles descriptions remain absent or constructing a mathematical characterization proves excessively burdensome owing to system magnitude and intricacy, parameter tuning may alternatively be steered by exploiting the informational content inherently embedded within the datasets harvested from the target system. The structural regularities contained in such datasets are distinctive to the system's operational states and boundary conditions, aspects which frequently defy straightforward mathematical articulation. Hence, assimilating data-inherent information through suitable statistical quantifiers may represent a viable route for guiding parameter evolution throughout ML model development, potentially manifesting as heightened interpretability when assessed through post-hoc lenses.

Numerous statistical measures exist for capturing informational presence and character within datasets through varied perspectives. Amongst these, the Pearson correlation coefficient serves as a conventional gauge for probing the linear interdependency linking two variables based on the observational data tied to each. The empirically quantified information conveyed by the Pearson correlation coefficient captures system behaviour across diverse operational regimes, making it conducive to integration when constructing accurate internal functional representations within ML models. Moreover, the pairwise variable interactions encapsulated by the Pearson correlation coefficient can fortify both the modelling efficacy and interpretive clarity of ML frameworks, particularly for highly nonlinear function spaces that resist straightforward approximation in practical applications. Across select recent investigations centred on object identification, image manipulation, and allied domains [11-16], correlational intelligence has been leveraged to boost predictive performance and interpretability in ML constructs. That said, the practice of embedding correlational intelligence during artificial neural network (ANN) formulation targeting regression-centric tasks remains undocumented within scholarly discourse. Likewise, scrutiny of the repercussions of assimilating empirically derived correlational intelligence on both modelling fidelity and interpretive performance for regression-oriented challenges in energy infrastructure is conspicuously absent from the literature, warranting examination to furnish a prospective remedy for the opaque disposition of ANN models.

In this contribution, we advance the Data Information integrated Neural Network (DINN) paradigm, which seeks to assimilate data intelligence via the correlation metric to augment the twin objectives of predictive performance and interpretive soundness. Bearing in mind the documented advantages of assimilating correlational intelligence to bolstered neural network modelling outcomes [13, 15, 16], it is postulated that accessible correlational intelligence, which

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

encapsulates empirically derived inter-variable associations, may help elevate both the modelling fidelity and interpretive robustness of ML algorithmic approaches when applied to energy systems. Drawing conceptual inspiration from the operational principles underpinning PINN, the objective function governing DINN is tailored to absorb extant data intelligence expressed through the Pearson correlation coefficient, which then participates in shaping the mechanism for updating learnable parameters (weights and biases) throughout the training phase. In addition, a stipulation is imposed upon the termination criterion for DINN training, seeking to curtail the absolute discrepancy between the correlation computed linking an input variable with the actual output variable and that computed with model-generated output variable values. Accordingly, it is anticipated that DINN, conditioned upon this bespoke cost function and halting condition, may deliver superior interpretive performance on regression-centric problems sourced from energy systems when scrutinized through established post-hoc analytical instruments.

The introduced DINN methodology is examined across two energy-system-focused case investigations – the energy efficiency cooling and energy efficiency heating performance characteristics of buildings, sourced from the University of California, Irvine publicly accessible repository [17, 18], along with a supplementary case concerning the modelling of power output from a 395 MW rated gas turbine employing authentic industrial operational records. The predictive competence of DINN is juxtaposed with that of a conventional artificial neural network (ANN) to probe the relative modelling strengths of the DINN formulation. Additionally, an input perturbation-centric strategy, rooted in Gaussian-distribution-based sample generation, is operationalized to juxtapose the interpretive outputs arising from DINN and ANN qualitatively. The qualitative interpretive framework draws upon system-specific domain expertise to substantiate the model-derived interpretations and finds traction among investigators possessing intimate familiarity with the functioning principles of their respective systems [19]. The distinctive contribution of this body of work lies in assimilating data intelligence, manifested through correlation, to steer the parameter refinement trajectory during ML model development. This tactic may prove instrumental in achieving concurrent gains in modelling fidelity and interpretive clarity for energy system applications via ML constructs, thereby advancing the frontiers of machine learning scholarship. This work does not introduce a novel methodology or network architecture exclusively for interpretive analysis; rather, it weaves data-driven interpretability considerations directly into the architectural conception of DINN, which, upon completion of training, is subsequently subjected to interpretability scrutiny via conventional analytical avenues. This embedding of data-centric interpretability considerations, nonetheless, customarily translates into elevated model efficacy and, by extension, refined interpretability.

Materials and Methods

Mathematical formulation of DINN

Consider X to be the array of predictor variables, $X = [X_1, X_2,..., X_m]$, with dimensionality m by N, where 'N' denotes the overall number of data points associated with X. This collection X is utilized to build the functional relationship with the response variable D (dimensioned 1 by N) via the DINN framework. The count of units within the input layer of DINN corresponds to the components of X. W1 constitutes a matrix capturing the synaptic weights bridging the input layer to the hidden layer (containing 'h' nodes) of DINN, with W1 having the shape m by h. Conversely, the synaptic weights extending from the hidden layer to the output layer of DINN are encapsulated in W2, which possesses the shape 1 by h. The overarching pattern of signal propagation and transformation occurring across the DINN structure is illustrated schematically in **Figure 1**. The precise mathematical operations unfolding within DINN are, nonetheless, specified below:

$$p_1 = \sum W_1 \odot X^T + b_1 \quad (1)$$

$$y_1 = f_1(p_1) \quad (2)$$

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

In this expression, b1 stands for the vector of bias terms positioned at the hidden layer nodes of DINN, shaped as m by 1; $W_1 \odot X$ represents the Hadamard (element-wise) product of X with the corresponding synaptic coefficients in W1, and the aggregate ($p_1 = \sum W_1 \odot X + b_1$) is evaluated at each hidden layer node; f1 signifies the nonlinear transfer function imposed upon the hidden layer, projecting p1 onto the output range of the activation function. Within DINN, we have adopted the hyperbolic tangent activation function at the hidden layer stage, which compresses p1 in a nonlinear fashion into the interval spanning -1 to 1 and assigns it to y1. The processed signal (y1) then advances from the hidden layer toward the output layer of DINN for continued computation, as detailed below:

$$p_2 = \sum W_2 \odot y_1 + b_2 \quad (3)$$

$$Z = f_2(p_2) \quad (4)$$

In this context, b2 refers to a bias scalar initialized at the output layer, sized 1 by 1; $W_2 \odot y_1$ captures the element-wise multiplication between y1 and W2 (the synaptic coefficients running from the hidden to the output layer of DINN). The resulting sum ($p_2 = \sum W_2 \odot y_1 + b_2$) is determined at the output layer of DINN and is subsequently acted upon by the activation function f2 to yield the network-generated estimate (Z) from DINN. Throughout this investigation, a linear activation function (f2) has been applied at the output layer of DINN.

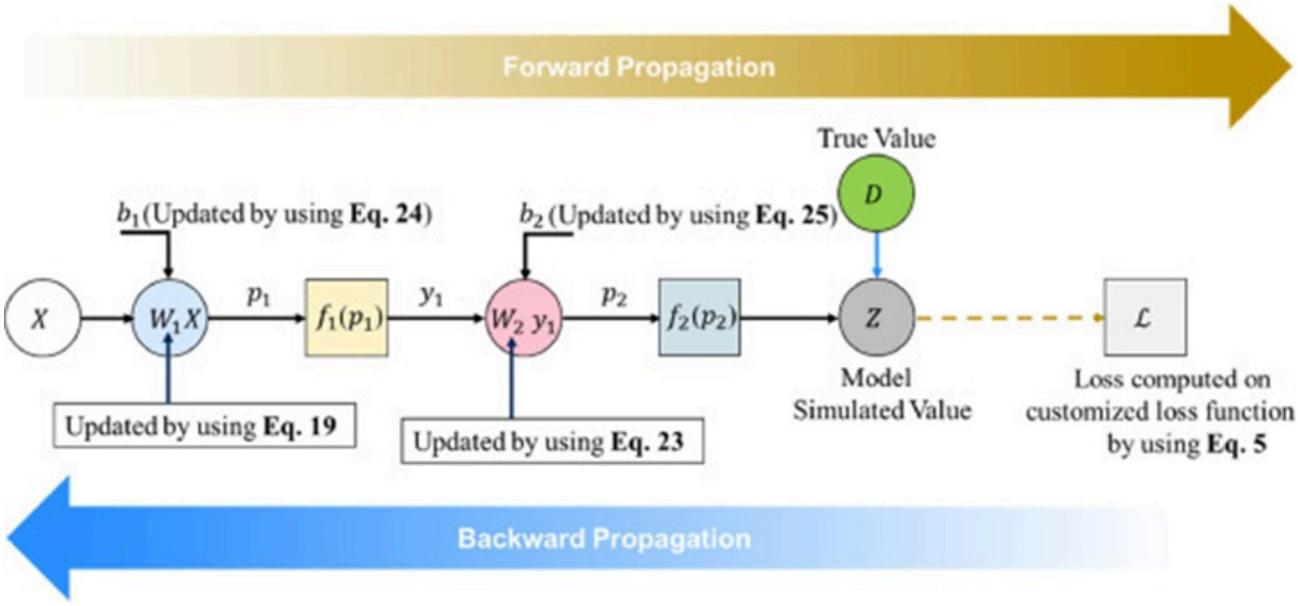

Figure 1. The generic schematic of the information flow and processing taking place in DINN.

The objective function ($\mathcal{L}$) devised within this study comprises the mean squared discrepancy (between the observed output variable (D) and the network-generated estimate (Z)), in addition to the mean squared divergence between the Pearson correlation evaluated for the components of Xi relative to D ($r_{X_i|D}$) and Z ($r_{X_i|Z}$). The Pearson correlation coefficient gauges the degree of linear association between a pair of variables, and its inclusion within the objective function brings in a supplementary term aimed at narrowing the gap between the correlation measure derived from the actual and network-generated responses. The rationale behind incorporating a correlation coefficient term carries a dual significance: (i) it directs the optimization of model parameters in a manner that bolsters the trained model's forecasting capability by mining and then assimilating the latent informational structure residing within the dataset, and (ii) the practitioner already harbours some awareness regarding the mutual dependence among input-output variables even before commencing network training, and that awareness gets encoded into the training procedure, culminating in a DINN

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

construct that delivers, upon appropriate examination, more meaningful interpretation relative to a standard ANN. As a result, the network-generated estimates progressively strengthen the alignment with Xi, mirroring the association originally embedded in the dataset, throughout the iterative training cycles. The bespoke objective function can pave the way toward building an efficient functional correspondence between the input-output variables for the fully trained DINN algorithm, which may subsequently furnish enhanced interpretive capability when juxtaposed with that of a conventional ANN model devoid of correlation intelligence within its objective function.

The Pearson correlation coefficient ranges from +1 (signifying a strong positive correlation) to -1 (signifying a strong negative correlation). Meanwhile, a Pearson correlation coefficient reading of zero points to the complete absence of any linear association between the pair of variables under consideration. It warrants emphasis here that the predictor variables deemed pertinent for modelling the response variable ought to be selected with care, drawing upon domain expertise or process insight, so that the correlation intelligence available for the variable pairings is meaningful and can aid in constructing a well-performing model. The tailored objective function for DINN differs from that of a traditional ANN model (ANN), which customarily features a single error expression. Nonetheless, weaving the correlation expression into the tailored objective function feeds into the parameter refinement process during the iterative training of DINN, which may equip the network to attain elevated performance in terms of both modelling fidelity and interpretative clarity. The two constituents within the proposed $\mathcal{L}$ are moderated by λ and are formulated as follows.

$$\mathcal{L} = \left(\frac{\lambda}{1+\lambda}\right).\frac{\sum (D - Z)^2}{N} + \left(\frac{1}{1+\lambda}\right).\frac{\sum_{i=1}^{m}\left(r_{X_i|D} - r_{X_i|Z}\right)^2}{N} \quad (5)$$

λ serves as a hyperparameter that regulates the relative influence of $\frac{\sum (D-Z)^2}{N}$ and $\frac{\sum_{i=1}^{m} \left(r_{X_i|D} - r_{X_i|Z}\right)^2}{N}$ in the evaluation of the objective function. The two error constituents within $\mathcal{L}$ are balanced through λ, and the practitioner may specify the weighted contributions assigned to each error constituent through the chosen value of λ. The two expressions bundled within $\mathcal{L}$ collectively act to reduce the divergence between the observed and network-generated responses and operate in synergy to drive down $\mathcal{L}$, thereby yielding a thoroughly trained DINN model. The gradient descent algorithm augmented with a momentum term is harnessed to arrive at the optimal parameter settings, i.e., weights and biases, given that this algorithm delivers rapid and stable error convergence relative to the behaviour of the plain gradient descent algorithm [20]. The partial derivative of $\mathcal{L}$ with respect to each parameter (W1, W2, b1, b2) is derived, the gradient information is relayed backward, and subsequently fused with the momentum-enhanced gradient descent scheme for iterative parameter adjustment until the minimum of $\mathcal{L}$ is attained.

The adjustment rule for the synaptic weight matrix W1 under the momentum-enhanced gradient descent algorithm is expressed as:

$$W_1^{new} = W_1 - \eta V_{w_1} \quad (6)$$

Within this expression, η corresponds to the learning rate hyperparameter, and Vw1 denotes the velocity matrix characterized as [21]:

$$V_{w_1} = \beta V_{w_1} + (1-\beta)\frac{\partial \mathcal{L}}{\partial W_1} \quad (7)$$

In this representation, β stands for the momentum coefficient, and V_{w_1} begins as a null matrix with the same dimensions as W1. The gradient term $\frac{\partial \mathcal{L}}{\partial W_1}$ is evaluated through successive application of the chain rule:

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

$$\frac{\partial \mathscr{L}}{\partial W_1} = \frac{\partial \mathscr{L}}{\partial Z}\frac{\partial Z}{\partial p_2}\frac{\partial p_2}{\partial y_1}\frac{\partial y_1}{\partial p_1}\frac{\partial p_1}{\partial W_1} \quad (8)$$

$$\frac{\partial \mathscr{L}}{\partial Z} = \frac{\partial}{\partial z}\left(\frac{\sum (D-Z)^2}{N} + \frac{\sum_{i=1}^{m}\sum \left(r_{X_i|D} - r_{X_i|Z}\right)^2}{N}\right) = \frac{-2(D-Z)}{N} - r_{X_i|Z}.\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right) \quad (9)$$

Where,

$$X_i^{\mu} = X_i - \overline{X}_i \quad (10)$$

$$Z^{\mu} = Z - \overline{Z} \quad (11)$$

$$\mathscr{B}_i = \sum_{i}^{m}(X_i - \overline{X}_i).(Z - \overline{Z}) \quad (12)$$

$$\mathscr{M} = \sum (Z - \overline{Z})^2 \quad (13)$$

$$\frac{\partial Z}{\partial p_2} = \frac{\partial p_2}{\partial p_2} = 1 \quad (14)$$

$$\frac{\partial p_2}{\partial y_1} = \frac{\partial}{\partial y_1}(W_2 \odot y_1 + b_2) = W_2 \quad (15)$$

$$\frac{\partial y_1}{\partial p_1} = \frac{\partial}{\partial p_1}\left(\frac{e^{p_1} + e^{-p_1}}{e^{p_1} - e^{-p_1}}\right) = 1 - y_1^2 \quad (16)$$

$$\frac{\partial p_1}{\partial W_1} = \frac{\partial}{\partial W_1}(W_1 \odot X + b_1) = X \quad (17)$$

Incorporating **Eq. 9-17** back into **Eq. 8**:

$$\frac{\partial \mathscr{L}}{\partial W_1} = -\left(\frac{2(D-Z)}{N} + r_{X_i|Z}.\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right)\right) W_2^T (1 - y_1^2) X^T \quad (18)$$

Drawing upon **Eqs. 6, 7, 18** may be restated as:

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

$$W_1^{new} = W_1 + \eta\left(\beta V_{W_1} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}\cdot\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right)\right)W_2^T\left(1-y_1^2\right)X^T;\right) \quad (19)$$

In a parallel fashion, the adjustment rule governing the synaptic weights extending from the hidden layer nodes to the output layer node of DINN takes the form [21]:

$$W_2^{new} = W_2 - \eta\, V_{W_2} \quad (20)$$

$$V_{W_2} = \beta V_{W_2} + (1-\beta)\frac{\partial \mathscr{L}}{\partial W_2} \quad (21)$$

In this case, VW_2 mirrors the dimensionality of W_2 and is set to a null matrix initially. The gradient can be unpacked by invoking the chain rule as outlined below:

$$\frac{\partial \mathscr{L}}{\partial W_2} = \frac{\partial \mathscr{L}}{\partial Z}\frac{\partial Z}{\partial p_2}\frac{\partial p_2}{\partial W_2} = -\left(\frac{2(D-Z)}{N} + r_{X_i|Z}\cdot\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right)\right)y_1 \quad (22)$$

Leveraging **Eq. 21** together with **Eq. 22**, **Eq. 20** may be reformulated as:

$$W_2^{new} = W_2 + \eta\left(\beta V_{W_2} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}\cdot\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right)\right)y_1\right) \quad (23)$$

Employing an analogous derivation, the update rules for the bias terms residing at the hidden layer (b1) and the output layer (b2) are rendered as:

$$b_1^{new} = b_1 + \eta\left(\beta V_{b_1} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}\cdot\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right)W_2^T\left(1-y_1^2\right)\right)\right) \quad (24)$$

$$b_2^{new} = b_2 + \eta\left(\beta V_{b_2} + (1-\beta)\left(\frac{2(D-Z)}{N} + r_{X_i|Z}\cdot\left(\frac{X_i^{\mu}}{\mathscr{B}_i} - \frac{Z^{\mu}}{\mathscr{M}}\right)\right)\right) \quad (25)$$

The parameter refinement process persists until any one of the halting criteria is met, namely: the slope descends beneath 0.00000001, the objective function value evaluated on the testing partition falls short of the prescribed goal value, and the absolute divergence between $r_{X_i|D}$ and $r_{X_i|Z}$, expressed as $\left|r_{X_i|D}\right|$ - $\left|r_{X_i|Z}\right|$, is driven to zero. Reducing the absolute divergence between $r_{X_i|D}$ and $r_{X_i|Z}$ guides the DINN algorithm to closely replicate the distribution profile exhibited by the true observations of the response variable when plotted against the collection of predictor variables. This process refines the precision of the functional correspondence linking the input and output variables. The halting criterion $\left|r_{X_i|D}\right|$ - $\left|r_{X_i|Z}\right|$ is poised to bolster both the modelling accuracy and interpretive performance of the DINN construct, given that the correlation intelligence embedded in the dataset is overtly capitalized upon throughout model development, in stark contrast to the ANN paradigm, which imposes no additional requirement to be fulfilled before the training procedure concludes.

Evaluation criteria

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Two statistical indicators, the coefficient of determination (R^2) alongside the root-mean-squared-error (RMSE), are applied to appraise the forecasting fidelity attained by the DINN framework. The computational forms for R^2 and RMSE are detailed below:

$$R^2 = 1 - \frac{\sum_i^N (Z_i - D_i)^2}{\sum_i^N (D_i - \overline{D}_i)^2} \quad (26)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_i - D_i)^2} \quad (27)$$

In these expressions, Zi and Di stand for the model-inferred outputs and the observed values of the target variables, respectively, iterated over i = 1, 2, 3,…, N, where N equals the aggregate number of recorded instances. R^2 reflects the proportion of variability in the target that the model succeeds in explaining, extending from a floor of zero (denoting virtually no explanatory power) up to a ceiling of one (signifying flawless replication of the observed responses). RMSE, for its part, conveys the typical magnitude of offset separating the observed readings from the model-inferred outputs and is driven downward as aggressively as achievable to realize a dependable functional correspondence bridging the predictor and target quantities under the DINN paradigm.

Results and Discussion

Case investigations rooted in energy system domains

The DINN procedure is exercised across a pair of case investigations – the first draws upon a publicly recognized benchmark corpus concerning energy efficiency cooling (ENC) together with energy efficiency heating (ENH). In contrast, the complementary case investigation focuses on reproducing the electrical output delivered by an operational industrial gas turbine installation boasting a 395 MW nameplate capacity. Through this dual setup, DINN has been applied to benchmark datasets within the research community, while its modelling merits have additionally been scrutinized in a real-world industrial scenario. It deserves explicit emphasis that the modelling merits of DINN are measured against those delivered by an ANN construct whose training relies upon a conventional objective function comprising exclusively mean-squared-error constituents. This side-by-side arrangement creates a pathway for diagnosing the degree to which reconfiguring the objective function influences the modelling merits exhibited by DINN in the examined case investigations. Both the DINN and the ANN constructs are realized within the framework of MATLAB 2019 b. The head-to-head performance metrics for DINN and ANN are set out in the segments that follow.

Case investigation–1: energy efficiency cooling & energy efficiency heating

The energy efficiency cooling (ENC) together with the energy efficiency heating (ENH) data collections, both tied to the thermal behaviour of built structures, encompass a suite of eight predictor quantities: relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution. These assembled records span a total of 768 distinct observational instances. A share representing 80% of these observational instances is channelled into the training stage, leaving the residual 20% as an untouched evaluation portion throughout the construction of models targeting ENC and ENH. The complement of processing nodes stationed within the hidden layer is held constant at 10, a choice that facilitates a meaningful comparison of the modelling merits delivered by DINN and ANN alongside those reported for a feed forward neural network (FFNN) within the broader scholarly conversation [22]; every model configuration is subjected to 20 separate rounds of training to neutralize the variability in measured performance arising from distinct parameter initialization seeds. In parallel, a collection of candidate settings for the hyperparameter λ is surveyed, sweeping across the band extending from 0.1 through 0.2, during the conditioning of the DINN models.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Figure 2 juxtaposes the modelling merits realized by the DINN and ANN formulations when tasked with reconstructing ENC and ENH profiles. The assessment indices are determined over both the training corpus and the held-out evaluation corpus for each of the two target variables. The dispersion of RMSE values for DINN and ANN across the evaluation corpus for ENC and ENH is shown graphically as box plots in **Figure 2a**. It becomes apparent that the average RMSE attained by DINN on the evaluation corpus for ENC and ENH registers at 1.23% and 0.56%, readings that sit notably beneath the corresponding figures posted by ANN, specifically 1.41% and 0.65%. Furthermore, for the collection of trained ANN and DINN candidates, the assessment indices computed between the observed and model-inferred outputs for the comparably stronger-performing realizations for ENC and ENH are presented in **Figure 2b**. With reference to ENC, the R^2 statistic stays broadly aligned across both the training corpus and the evaluation corpus when contrasting the DINN and ANN models. Despite this similarity, the RMSE estimates delivered by DINN for the training corpus and the evaluation corpus stand at 1.04% and 1.14%, both of which undercut those emanating from the ANN counterpart, namely RMSE_train = 1.08% and RMSE_test = 1.24%. A matching pattern emerges when inspecting the modelling merits of DINN and ANN in the ENH context, where the R^2 value calculated on the training corpus remains largely equivalent between the two models. Nonetheless, the DINN-generated estimates across both the training corpus and the evaluation corpus yield consistently lower RMSE values than those sourced from ANN. The modelling merits of the two formulations with respect to ENH can be succinctly encoded as: (RMSE_train = 0.46%) DINN < (RMSE_train = 0.51%) ANN and (RMSE_test = 0.43%) DINN < (RMSE_test = 0.52%) ANN. Beyond this, the modelling merits of DINN are additionally placed alongside those of an FFNN configuration for ENC and ENH as documented in prior scholarly work [22]. Those investigators performed 20 repeated training cycles of the FFNN under a structurally identical configuration. They disclosed the averaged findings from the repetition runs as: RMSE_test = 1.63% and RMSE_test = 0.63% for ENC and ENH, respectively [22], numerals that appreciably surpass those achieved by DINN as delineated earlier. Balancing the predictive merits evidenced by DINN against those of ANN and FFNN reveals unmistakably that DINN commands superior modelling strength for the ENC and ENH tasks. This phenomenon may plausibly be traced back to the assimilation of data-borne intelligence via the correlation structure connecting the predictor and target variables. This feature likely assisted the parameter refinement machinery in securing the uplifted predictive capacity of DINN.

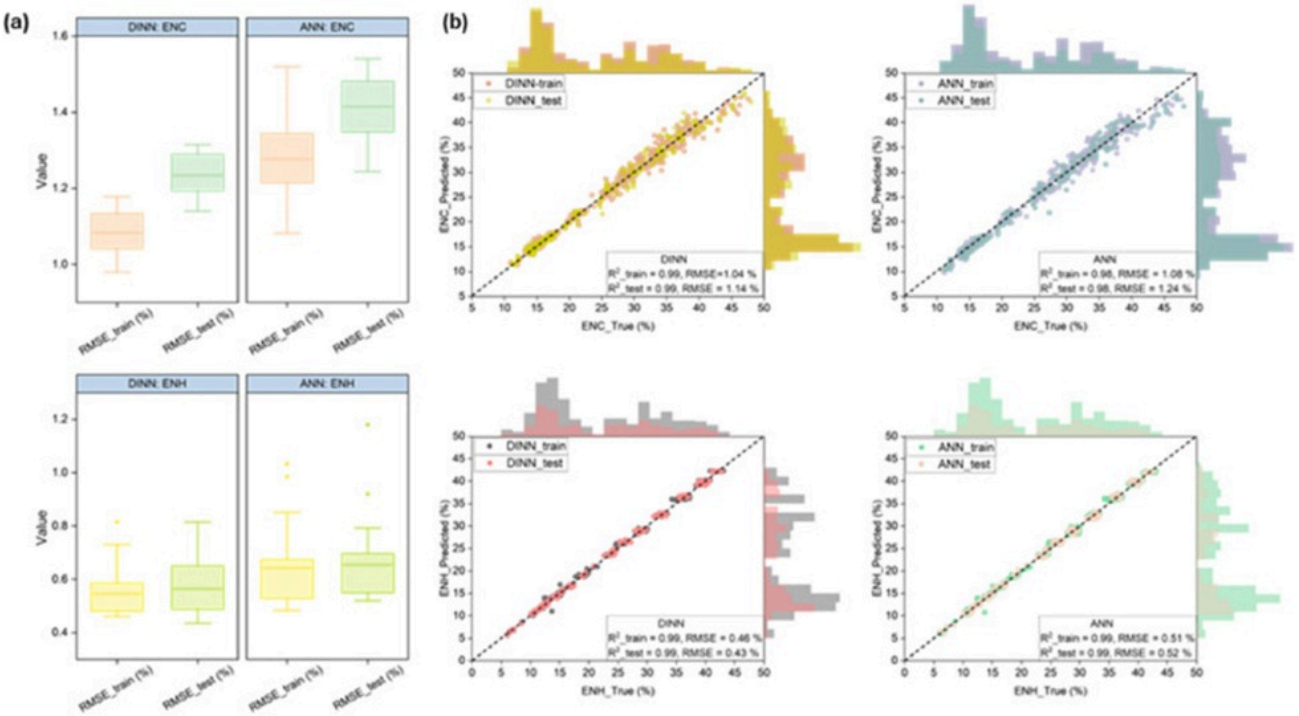

Figure 2. Comparison of modelling performance of the DINN with the ANN model for ENC and ENH: (a) The graphical visualization of variability in the RMSE on the test dataset is presented for 20 training runs of the ANN and DINN models under the same initial conditions for ENC and ENH, and (b) True vs. predicted responses of DINN and ANN, which have

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

comparatively better performance metrics among the 20 trained models. The performance metrics computed on the testing dataset for DINN appear to be better than those of ANN for ENC and ENH, thereby showing improved modelling performance of DINN. The bars along the edges of the graph represent the data-distribution density profiles for the training and testing dataset with respect to true and model-predicted responses.

The predictor quantities feeding into the ENC and ENH formulations are characterized by mutual interlinkages, as encoded within the observational records through the operative ranges of the respective quantities. Notably, however, a couple of the predictor quantities – orientation together with glazing area – are distinguished by relatively restricted operative ranges and exhibit weaker coupling with the remainder of the predictor quantities across the evaluation corpus. This circumstance provides an opportunity to examine the behavioural response to fluctuations in these two nominated predictor quantities across the evaluation corpus for both the DINN and ANN models applied to ENC and ENH, and to check the model-inferred outputs against the observed target readings. Such a parametric probing exercise navigates confined subregions of the evaluation corpus with the intent of diagnosing both the predictive trustworthiness of the models and the degree to which they succeed in generalizing the inter-variable linkages. The DINN- and ANN-inferred outputs for ENC and ENH under systematic variation of orientation and glazing area are depicted in **Figure 3**. The DINN-inferred outputs are traced in red. At the same time, the ANN-derived equivalents appear in blue, with the observed target readings shown as solid black lines. The outputs furnished by both the DINN and ANN configurations in response to the manipulation of the two predictor quantities remain in reasonable congruence with the observed target readings, an indication that both formulations possess adequate modelling competence across the evaluation corpus. When one undertakes a finer-grained assessment of predictive competence between the two formulations, it becomes evident that DINN consistently produces prediction traces that are superior to those associated with ANN. This observation reinforces the conclusion that DINN possesses heightened predictive precision and greater generalization capacity relative to the conventional ANN formulation and highlights the competitive advantage secured by embedding correlation intelligence within the DINN architecture for refined predictive analytics. Furthermore, the improved predictive competence documented for DINN under parametric sweeps of the two quantities also supports efforts to decipher the model's interpretability and gauge the correctness of the model's projections by aligning them against the observed target readings.

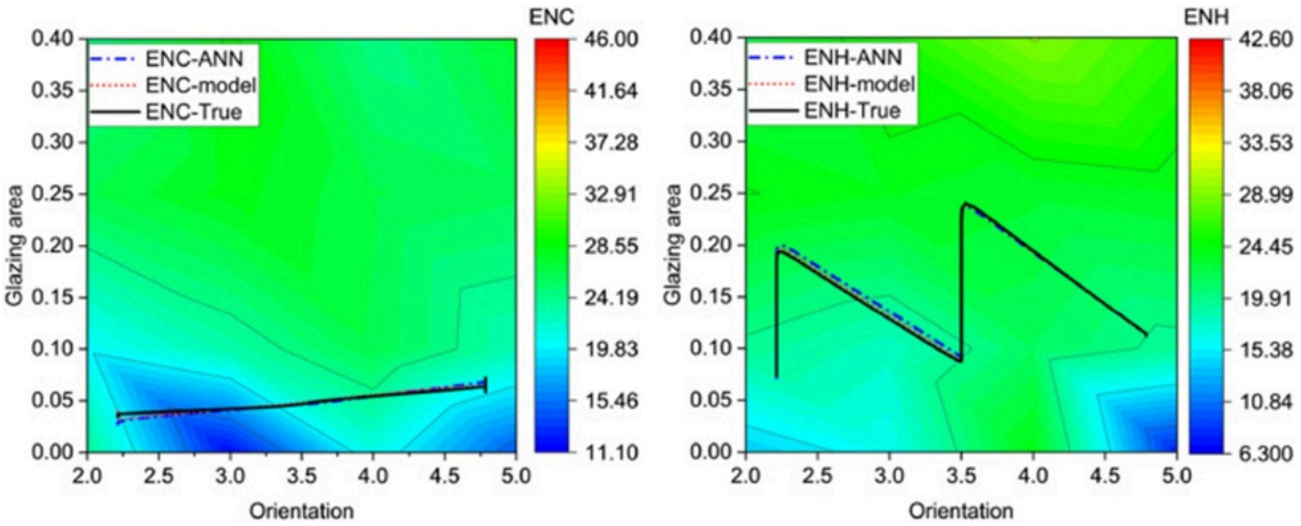

Figure 3. Effect of varying orientation and glazing area on the ENC and ENH performance as studied by the DINN and ANN models. The model predicted responses are compared with the true values to evaluate the predictive accuracy of the models.

Case investigation–2: electrical output from an operational industrial gas turbine

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

The electricity produced by a utility-grade gas turbine emerges from the coordinated functioning of numerous supporting subsystems woven into the generation infrastructure. The process quantities participate in elaborate, deeply nonlinear mutual dependencies that ultimately govern the power output behaviour. For this inquiry, the selection of predictor quantities has been guided by a combination of experiential familiarity with station operations and consultation of the surrounding scholarly discourse [23, 24]. The ensuing collection of variables has thus been marshalled to develop the empirically grounded process representation for electrical generation: compressor discharge air temperature (CDT), compressor discharge air pressure (CDP), fuel gas flow rate (GFR), performance heater gas outlet temperature (PHT), fuel gas temperature (FGT), ambient temperature (AT), ambient pressure (AP), and ambient humidity (AH).

The measurement archive, aggregating 578 logged entries, was obtained directly from the gas-firing power station, with the records split 80/20 between the parameter-fitting phase and the holdback-testing phase. The dimensionality of the hidden processing layer represents a decisive structural hyperparameter that delimits how much representational flexibility the ANN machinery can draw upon to render the target mapping with acceptable precision. Acting on this premise, a methodical traversal of the architectural configuration space has been performed, with the hidden complement being adjusted through a multiplicative window ranging from 1× to 2.5× relative to the count of input-layer predictor nodes [25]. The regularisation coefficient λ, conversely, remains pegged at 0.4 for the entirety of the DINN conditioning procedure. Performance yardsticks are logged during both the parameter-optimisation pass and the subsequent evaluation pass throughout the DINN build process. **Figure 4** portrays how the modelling fidelity of DINN and ANN shifts as a function of the hidden-layer complement, with the corresponding RMSE trajectories for the two paradigms displayed via spider-web charts. A detailed side-by-side reading of predictive fidelity across the optimisation and evaluation passes leaves little doubt that the DINN variant with 10 hidden computational nodes yields the most favourable RMSE on both the optimisation and evaluation subsets when contrasted with other DINN instantiations with differing hidden-layer complements. Mirroring this outcome, the ANN paradigm reaches its lowest RMSE when configured with 18 hidden computational nodes. The scatterplot relating observed readings to model-reconstructed estimates of gas turbine power (GT Power) corresponding to the optimally tuned hidden-layer arrangement in each paradigm is offered in **Figure 4.** The DINN and ANN formulations return broadly overlapping R^2 magnitudes across both the optimisation subset and the evaluation subset. Despite this similarity, the predictive exactitude realised by the two paradigms diverges considerably once RMSE is adopted as the evaluative lens on the optimisation and evaluation subsets. For DINN, the pair of readings RMSE_train = 1.20 MW and RMSE_test = 1.25 MW are recorded, both of which are markedly lower than the parallel figures produced by the ANN comparison case, namely RMSE_train = 2.0 MW and RMSE_test = 2.17 MW. Weighing the performance ledger of the DINN formulation against that of the ANN formulation substantiates that DINN secures appreciably keener predictive precision, spanning both the optimisation subset and the evaluation subset, in emulating the electricity generation signature of the gas turbine asset.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

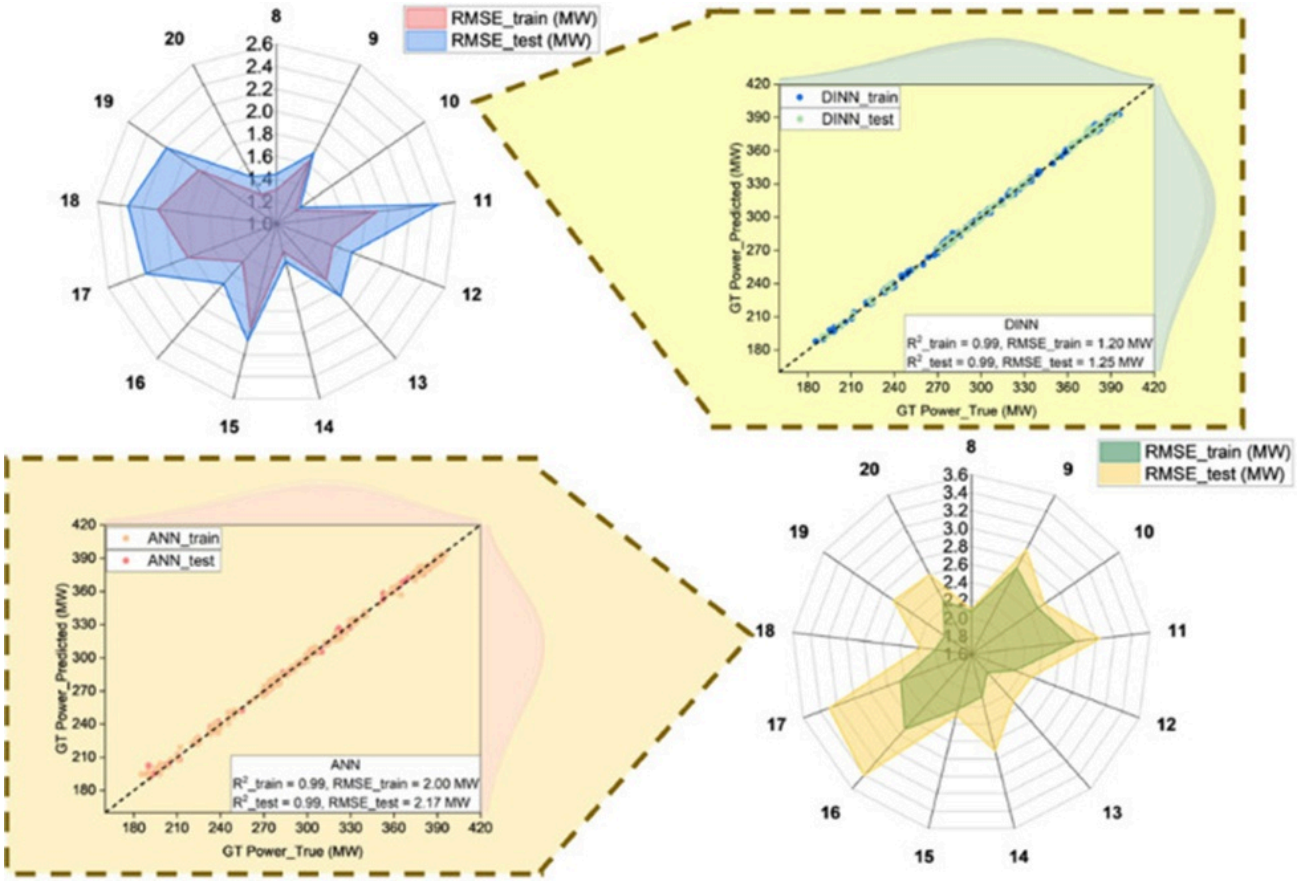

Figure 4. Predictive performance of DINN and ANN to model the power generation from an industrial-scale 395 MW capacity gas turbine. The performance metrics are measured for training and testing datasets to investigate the effectiveness of the functional map created between the input and output variables for DINN and ANN. The data-distribution density profiles for the training and testing datasets with respect to true and model-predicted responses are drawn along the edges of the graph.

Throughout the preceding passage, the modelling exactitude exhibited by the DINN and ANN formulations across the span of their construction has been set forth, with performance catalogued separately over the optimisation corpus and the evaluation corpus. In strands of more recent scholarship, investigators have turned attention toward corroborating machine learning artefacts by directing them to generate forecasts over an external verification pool – denoting a batch of data points purposely sequestered and hence essentially unseen over the training cycles, yet originating from the identical physical system being studied. An external verification undertaking of this character represents a notably rigorous crucible for appraising the extent to which a trained machine learning construct generalises [26, 27]. Aligned with this philosophy, an external verification pool comprising 144 stochastically selected operational states culled from the gas turbine generation log has been sourced from the station and submitted to the already-trained DINN and ANN artefacts for forward prediction. The forecasts emitted by the two artefacts are plotted together with the ground-truth readings, the assessment yardsticks R^2 and RMSE are calculated, and the residual mismatches are graphed in **Figure 5a**. Stacking the assessment yardsticks for DINN against those for ANN, it can be noted that the two artefacts achieve broadly equivalent R^2 scores over the external verification pool. The RMSE yardstick computed for DINN, by contrast, sits below the magnitude obtained for the ANN artefact (RMSE_DINN = 1.51 MW < RMSE_ANN = 2.75 MW). Beyond this observation, some 92.4% of the residual mass for DINN, versus 76.4% for ANN, is contained within the interval flanked by -2.5 MW and 2.5 MW, a finding that betrays the tendency of DINN to produce residuals of systematically smaller amplitude relative to those originating from ANN throughout this moderately wide residual interval.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

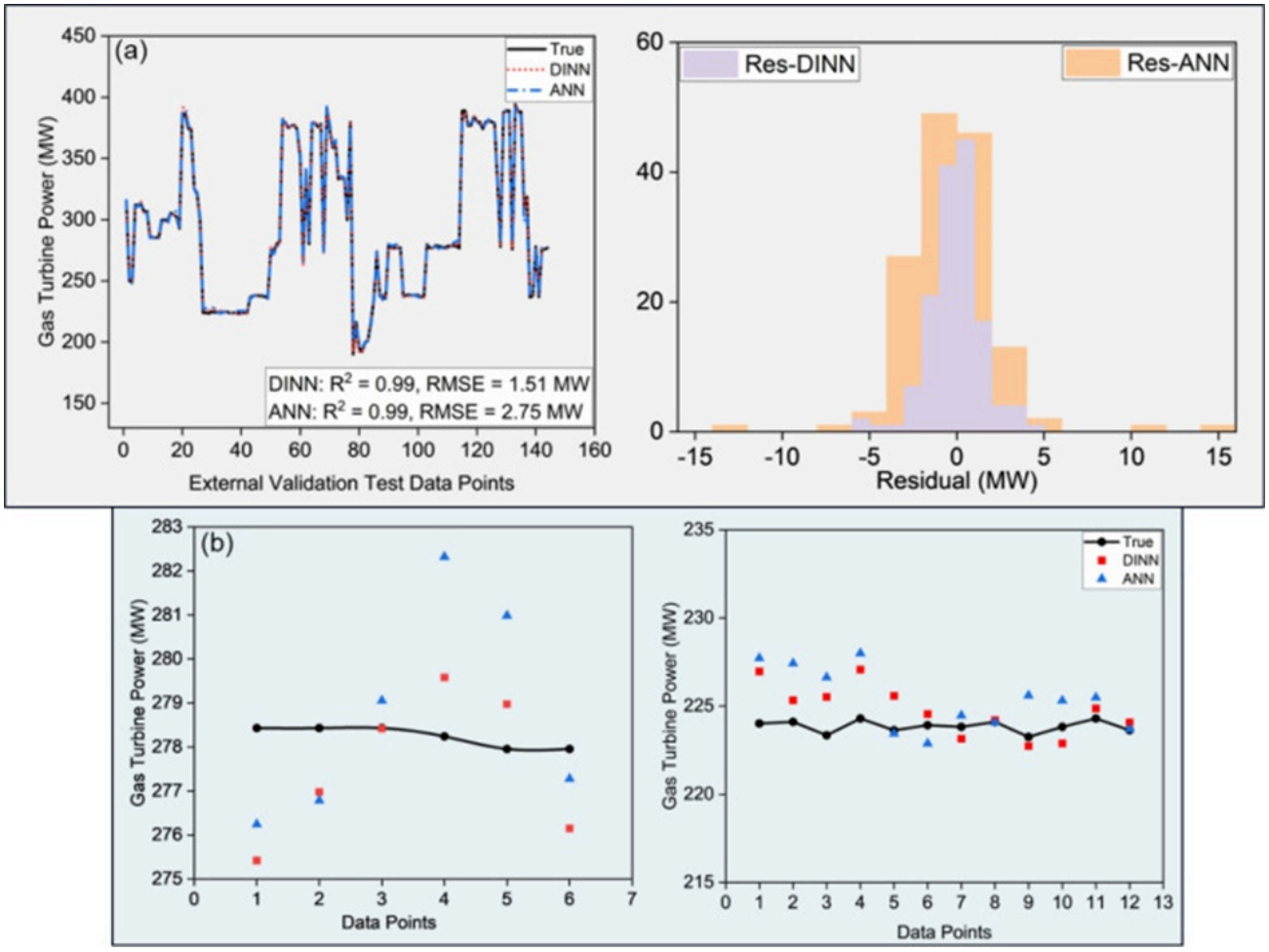

Figure 5. Predictive performance of DINN and ANN for the external validation dataset: (a) DINN- and ANN-model-based responses mapped against the true observations for the external validation dataset, along with the residual distributions, are shown, and (b) The predictive performance of DINN and ANN in the local space of gas turbine power built with the significant contribution of ambient temperature and ambient humidity is presented. The models' predicted responses are plotted against the true observations of the gas turbine power to show the predictive accuracy of the two models.

The capacity of both modelling paradigms to generalise is subjected to additional scrutiny within delimited pockets of the gas turbine power operational landscape, homing in specifically on fluctuations in ambient temperature and ambient humidity. At the same time, the remaining predictor quantities are kept near-constant in the vicinity of their customary operating values. Because ambient environmental conditions drift independently of the other predictor quantities throughout generation activities, the imprint that ambient temperature and ambient humidity leave on the live power output is intrinsically woven into the collected operational chronicle, thereby constituting an organic ground truth. This set of circumstances allows interrogating how the already-trained DINN and ANN artefacts behave across narrow sectors of the output variable domain, with the respective forecasts held up against the true readings to form a judgement on the fidelity of the trained artefacts. **Figure 5b** maps out the congruence between DINN- and ANN-generated estimates and the observed power output under two distinguishable gas turbine generation modes, scrutinised over a batch of input conditions extracted from the external verification pool, wherein ambient temperature and ambient humidity are permitted to vary. At the same time, the remaining predictor quantities are constrained to remain nearly constant. Scanning the dispersion of the DINN- and ANN-based estimates around the authentic gas turbine power trajectory across the two generation modes, the overarching impression gained is that DINN appears inclined to yield forecast traces that track the ground-truth readings with superior fidelity. There are occasional moments where ANN eclipses DINN in forecast accuracy. Taken as a whole, nonetheless, the DINN-based estimates seem to settle somewhat more tightly around the

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

ground-truth readings. An exhaustive vetting of DINN has accordingly been carried out through the external verification protocol, and its predictive exactitude has been juxtaposed with that of the ANN paradigm to gauge the modelling rigour of DINN when brought to bear on the problem of forecasting electrical generation from a gas turbine asset.

As a terminal analytical move, unearthing the proportional contribution that each candidate predictor makes toward the reconstruction of the target quantity is indispensable, given the inherently inscrutable character of the ANN modelling tradition. The ranked ordering of predictor saliency communicates the relative leverage each quantity exercises upon the model's computed output. The soundness of the input-output functional mapping built by the data-driven ML model, when probed through variable saliency dissection, furnishes a conduit for deciphering the logic underpinning the model's forecasts [28]. Furthermore, the saliency hierarchy assigned across the predictor quantities can be squared against foundational process understanding, thereby functioning as a vehicle for authenticating the coherence of the model's interpretive conduct [29].

For this inquiry, the input perturbation framework was adopted as the vehicle for dissecting predictor saliency, complemented by a qualitative orientation toward evaluating the interpretive trustworthiness delivered by each modelling paradigm [19]. The qualitative pathway for probing a model's interpretability fuses the practitioner's experiential intuition with foundational domain insights to corroborate the interpretive picture emerging from the model. Within the input perturbation protocol, the predictor being scrutinised is driven through its complete operating envelope, spanning from the minimal to the maximal recorded setting. In contrast, all remaining predictors are pegged at their respective mean values [30]. One may of course pursue alternative experimental design philosophies within the broader input perturbation methodology. At the subsequent stage, the technique overlays stochastic disturbances onto the designated input configuration and averages across the resultant output variable realisations, thereby accommodating the natural dispersion in the operational states of the predictors. This yields a granular inquiry at each operating point, drawing upon a sizeable ensemble of noise-corrupted instances to isolate the contribution of the predictor under examination. The workflow is then cycled across the full parametric sweep of that predictor and propagated over every predictor in the set to yield a standardised ranking of importance. In this investigation, we have generated a pool of 10,000 noise draws obeying a Gaussian probability density, scaled in proportion to one percent of the floor value characterising each predictor, and these noise draws are mobilised within the input-perturbation-driven variable importance dissection to illuminate the model's internal logic when forming its forecasts. The choice of a Gaussian model for the noise realisations rests on the observation that transducer readings in gas turbine power generation settings ordinarily exhibit Gaussian-like distributional properties.

Figure 6 displays the ranking of predictor variables according to their influence upon electrical generation as revealed through the input perturbation procedure when executed by way of the DINN and ANN models respectively. Each model quantifies the percentage-based importance allocated across predictors to enable comparative scrutiny. CDP together with CDT surface are the two decidedly pre-eminent predictors governing the electrical output of the gas turbine, with percentage importance shares of 75.4% and 12.5% under DINN, and 57.7% and 18.3% under ANN. The predictor assuming the third ordinal position in the DINN-derived ranking is GFR, which carries a percentage importance of 5.1%. Under the ANN model, in contrast, FGT occupies the third slot within the importance ordering, registering a percentage importance of 7.1%. The leading three predictors identified by DINN—CDP, CDT, and GFR—together amass around 95% of the cumulative importance assigned to the reconstruction of gas turbine power. Within the ANN model, a broader assembly consisting of CDP, CDT, FGT, AT, and GFR must be aggregated to approach a comparable 95% cumulative importance threshold for the same forecasting target. CDP and CDT are measurements that capture the thermodynamic state of the airstream exiting the compressor stage. GFR, analogously, quantifies the rate at which fuel is admitted into the combustor.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

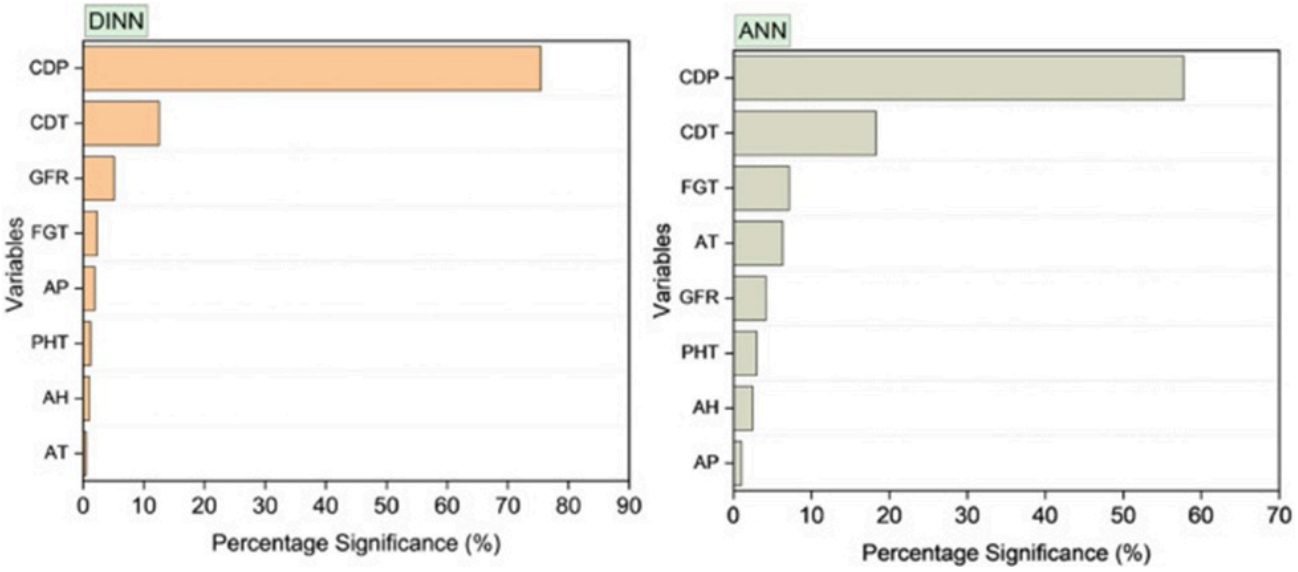

Figure 6. The significance of the input variables to predict the power generation from the gas turbine. The variables' significance analysis presents the importance of each variable in predicting the response variable, thereby contributing to model interpretation.

The chemical reaction between fuel and compressed oxidant liberates a combustion gas stream at elevated temperature, which is subsequently expanded across the turbine stages to extract mechanical work and, in turn, electrical power. The generation yield is thus acutely attuned to the conditioning of both the air and fuel streams that conjointly produce the propulsive combustion gases. This relationship stands as an elemental operational axiom underpinning gas-fired electricity production. The third-place ranking accorded to GFR within the DINN framework therefore dovetails convincingly with the operational logic and established domain understanding of gas turbine power generation when appraised through the variable importance lens, signalling that DINN supplies a more defensible interpretive account of the power generation phenomenon insofar as the ordering of predictor relevance is concerned. Under the ANN-generated importance hierarchy, by comparison, GFR is displaced to the fifth rank.

In contrast, FGT ascends to the third most influential predictor—an arrangement of priorities that accords less naturally with the accumulated process know-how and engineering judgement governing gas-fired power station operations. The input perturbation strategy deployed to excavate model interpretability spotlights the value of embedding correlational intelligence within the parameter refinement cycle during the DINN construction stage, a design choice whose payoff surfaces as sharpened interpretability of the model-based forecasts. Practitioners across the industrial landscape and operational engineers stand poised to extract benefit from the elevated interpretability that DINN provides, whether in formulating supervisory strategies for the efficient stewardship of their industrial installations, in executing informed operational decision-making, or in engineering training simulators geared toward operator certification and the familiarisation of junior engineers, thus enabling a richer appreciation of system dynamics under diverse operational regimes. Harnessing the intelligence latent within operational datasets inside the modelling algorithm can therefore serve as a powerful enabler toward simultaneously securing heightened predictive exactitude and interpretations that resonate with established domain principles. This tandem achievement propels the Industry-4.0 aspiration for cognitively aware management of industrial systems.

Conclusion

The intelligence woven into data repositories can be instrumental in unlocking gains in both the predictive potency and the interpretive transparency of machine learning artefacts. In this manuscript, we advance a modelling algorithm that absorbs the correlational architecture embedded in the variable dataset and leverages it to condition an effective Data

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Information integrated Neural Network (DINN). The cost function underpinning DINN is supplemented with a term capturing the mean squared departure between the Pearson correlation magnitudes computed for the observed outputs and the model-reconstructed outputs with respect to the predictor variables. This reconfigured cost function consequently directs the iterative adjustment of learnable parameters via the momentum-augmented gradient descent scheme.

The proposed DINN algorithm is applied to a duo of case investigations originating from the energy systems arena, namely the energy efficiency cooling (ENC) and heating (ENH) performance of built structures, together with the electrical generation from a 395 MW rated gas turbine facility. The modelling capability of DINN is measured alongside that of a baseline ANN formulation. For the ENC case, DINN exhibits a reduced mean RMSE over the evaluation splits (RMSE_test = 1.23%) relative to the ANN comparison (RMSE_test = 1.41%). Furthermore, the DINN-based predictive capability is found to fall below results previously reported in the literature for ENC ((RMSE_test = 1.23%) DINN < (RMSE_train = 1.63%) FFNN). Analogous findings attesting to the superior predictive capability of DINN are recorded across the remaining case investigations.

An external confirmation exercise is undertaken on both the DINN and ANN models in the gas turbine power context to examine the generalisation prowess of the two paradigms. DINN delivers a diminished forecast discrepancy (RMSE_DINN = 1.51 MW < RMSE_ANN = 2.75 MW) and a residual distribution concentrated within a tighter discrepancy band compared with that of ANN. Additionally, an input-perturbation-based variable importance dissection is performed to construct the ranking of predictor influence upon the electrical output of the gas turbine. The variable importance ordering derived from DINN aligns more coherently with the domain understanding characterising gas turbine power generation.

The present investigation delineates the merits of enriching the cost function to shepherd the parameter refinement journey toward the assembly of a highly predictive DINN model that concurrently manifests improved interpretability. The incorporation of data-borne intelligence into the operational mechanics of ML models thus holds the potential to simultaneously advance both the predictive and interpretive dimensions of model performance.

Acknowledgements

None

Conflict of interest

None

Financial support

Punjab Education Endowment Fund (PEEF) provided funding to Waqar Muhammad Ashraf for pursuing his PhD at University College London.

Ethics statement

None

Received: 07 Jul 2022 Revised: 12 Sep 2022 Accepted: 28 Sep 2022
Published online: 15 November 2023

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

1. Kang C, Liu Y, Zhang H, Wang L, Chen X, Li J, et al. Big data analytics in China's electric power industry: modern information, communication technologies, and millions of smart meters. IEEE Power Energy Mag. 2018;16(3):54-65.

2. Tsihrintzis GA, Sotiropoulos DN, Jain LC. Machine learning paradigms: advances in data analytics. Cham: Springer; 2019.

3. Hopkins AM, Logan JM, Kichenadasse G, Sorich MJ, Rowland A, Karapetis CS. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. JNCI Cancer Spectr. 2023;7(2).

4. Muhammad Ashraf W, Dua V, Ghafoor A, Moeen U, Farhan M, Raza A, et al. Optimization of a 660 MWe supercritical power plant performance—A case of Industry 4.0 in the data-driven operational management. Part 2. Power generation. Energies (Basel). 2020;13(21):5619.

5. Chen Z, Zhang Y, Li X, Wang H, Liu S, Zhao Q, et al. Interpretable machine learning for building energy management: a state-of-the-art review. Adv Appl Energy. 2023;100123.

6. Antoniadi AM, Du Y, Guendouz Y, Wei L, Mazin C, Becker S, et al. Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review. Appl Sci. 2021;11(11):5088.

7. Lisboa P, Saralajew S, Vellido A, Fernández-Navarro F, Biganzoli E, Gacto M, et al. The coming of age of interpretable and explainable machine learning models. Neurocomputing. 2023;535:25-39.

8. Kenny EM, Keane MT. Explaining Deep Learning using examples: optimal feature weighting methods for twin systems using post-hoc, explanation-by-example in XAI. Knowl Based Syst. 2021;233:107530.

9. Cai S, Wang Z, Zhang X, Li Y, Chen H, Liu J, et al. Physics-informed neural networks (PINNs) for fluid mechanics: a review. Acta Mech Sin. 2021;37(12):1727-38.

10. Almajid MM, Abu-Al-Saud MO. Prediction of porous media fluid flow using physics informed neural networks. J Pet Sci Eng. 2022;208:109205.

11. Li S, Deng M, Lee J, Sinha A, Barbastathis G, Kim Y, et al. Imaging through glass diffusers using densely connected convolutional networks. Optica. 2018;5(7):803-13.

12. Ou X, Liu L, Tu B, Zhang G, Xu Z, Chen Y, et al. A Hyperspectral Image Change Detection Framework With Self-Supervised Contrastive Learning Pretrained Model. IEEE J Sel Top Appl Earth Obs Remote Sens. 2022;15:7724-40.

13. Fan W, Chen X, Liu Y, Zhang H, Wang L, Li J, et al. 2D shape reconstruction of irregular particles with deep learning based on interferometric particle imaging. Appl Opt. 2022;61(32):9595-602.

14. Nagaraju TV, Prasad CD, Reddy KS, Kumar A, Singh B, Gupta R, et al. Predicting California Bearing Ratio of Lateritic Soils Using Hybrid Machine Learning Technique. Buildings. 2023;13(1):255.

15. Liu M, He P, Wang H, Li Y, Zhang X, Chen S, et al. BIT-MI Deep Learning-based Model to Non-intrusive Speech Quality Assessment Challenge in Online Conferencing Applications. In: Proc. Interspeech 2022. 2022. p. 3288-92.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press

16. Sun X, Zhang Y, Li Q, Wang H, Chen L, Zhao F, et al. Dual-task convolutional neural network based on the combination of the U-Net and a diffraction propagation model for phase hologram design with suppressed speckle noise. Opt Express. 2022;30(2):2646-58.

17. Tsanas A, Xifara A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy Build. 2012;49:560-7.

18. Oh SJ, Schiele B, Fritz M. Towards Reverse-Engineering Black-Box Neural Networks. In: Samek W, Montavon G, Vedaldi A, Hansen LK, Müller KR, et al., editors. Explainable AI: interpreting, explaining and visualizing deep learning. Cham: Springer International Publishing; 2019. p. 121-44.

19. Saleem R, Yuan B, Kurugollu F, Anjum A, Liu L, Wang J, et al. Explaining deep neural networks: a survey on the global interpretation methods. Neurocomputing. 2022;513:165-80.

20. Yu H, Wilamowski BM. Levenberg–marquardt training, in intelligent systems. Boca Raton: CRC Press; 2018. p. 12-1-12-16.

21. Ng A. Improving deep neural networks: hyperparameter tuning, regularization and optimization [Internet]. Deeplearning.ai on Coursera; 2017.

22. Arnaldo I, O'Reilly UM, Veeramachaneni K. Building predictive models via feature synthesis. In: Proceedings of the 2015 annual conference on genetic and evolutionary computation. 2015.

23. Qu Z, Zhang Y, Li X, Wang H, Liu S, Zhao Q, et al. Prediction of electricity generation from a combined cycle power plant based on a stacking ensemble and its hyperparameter optimization with a grid-search method. Energy. 2021;227:120309.

24. Hundi P, Shahsavari R. Comparative studies among machine learning models for performance estimation and health monitoring of thermal power plants. Appl Energy. 2020;265:114775.

25. Ashraf WM, Dua V. Machine learning based modelling and optimization of post-combustion carbon capture process using MEA supporting carbon neutrality. Digital Chem Eng. 2023;8:100115.

26. Ashraf WM, Dua V. Artificial intelligence driven smart operation of large industrial complexes supporting the net-zero goal: coal power plants. Digital Chem Eng. 2023;8:100119.

27. Zhang W, Li Y, Wang H, Chen X, Liu S, Zhao Q, et al. Machine learning based prediction and experimental validation of arsenite and arsenate sorption on biochars. Sci Total Environ. 2023;904:166678.

28. Kumar R, Singh AK. Chemical hardness-driven interpretable machine learning approach for rapid search of photocatalysts. NPJ Comput Mater. 2021;7(1):197.

29. Zhao S, Xu Z, Wang C. Interpretable machine learning for predicting and evaluating hydrogen production via supercritical water gasification of biomass. J Clean Prod. 2021;316:128244.

30. Nourani V, Fard MS. Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. Adv Eng Softw. 2012;47(1):127-46.

© 2023 The Author(s). | Published by Institute of Smart Energy Systems Press