The Indonesian Journal of Computer Science
www.ijcs.net
Volume 14, Issue 5, October 2025
https://doi.org/10.33022/ijcs.v14i5.5000

Models for Studying The Impact of Statistics Characteristics of Computing Network Gridlock on The Effectiveness Of Predictions Using Machine Learning

Adel Al-Janabi[1]
adelh.aljanabi@uokufa.edu.iq[1]
[1]University of kufa, Iraq

Article Information

Submitted : 10 Aug 2025
Reviewed : 29 Sep 2025
Accepted : 7 Oct 2025

Keywords

Network Traffic Analysis, Machine Learning Prediction, Traffic Modeling, Prediction Efficiency, Probabilistic Models

Abstract

The article endeavors to organize and classify a vast array of papers about contemporary approaches, strategies, and methods of data forecast across many domains, specifically with their usefulness in traffic forecast in networks of computers. The defined ordering is carried out within the context of the suggested concept model of forecast algorithms. This concept model emphasizes the qualities of both computational network activity models and traffic monitoring approaches that may be employed openly or implicitly in contemporary forecast software applications. It is demonstrated that the investigation of probabilistic characteristics for data definition, such as presence of considerable nonstationarity, certain nonlinear impacts in models of data, and the uniqueness of dissemination of data laws, could impact effectiveness in learning predictions.

A. Introduction

In spite of the fact that significant advancements in the creation of forecasting systems have been accomplished in the field of machine learning (ML), mostly based on neural networks, the challenges associated with their application have not been eliminated. Among the many reasons for this, one other thing that may be mentioned is that contemporary machine learning techniques do not taken into consideration certain mathematical aspects of the data that are random being predicted. This is due to the fact that in ML systems, the creation of a value prediction based on training data samples is often conducted without the use of mathematics in an effort to find inherent patterns and correlations probabilistic models of the data. This also applies to methods that use explicit mathematical models of the relationship of the predicted value in the future, such as linear, polynomial or exponential regression methods [1], but disregard mathematical characteristics of the input data, such as autocorrelation. These conditions are frequently complimented by the transparency of the change of input data into feature space, for instance, in deep learning algorithm [1], this makes it hard to be a priori judge the influence of sure properties of the contribution data on result. The article is a study of research on current forecasting methodologies the behavior of computer networks based on the use of software forecasting tools (IP) of the machine learning, in which the main attention is paid to methods of taken into consideration the influence of the features of the forecasted data on the effectiveness of traffic forecasting, taking into account the mathematical properties of the forecast models. In order to structure the huge array publication current approaches, procedures, and models for forecasting, on which this analysis was performed, a classification of forecasting models used in modern software ML is proposed, using the extended depiction of the concept model the forecasting problem proposed in [2], widely used in the practice of big data and Software engineering. Models of concept are a set of viewpoints about the elements and the model system's objectives, as well as their interactions, are represented in the form of one or another model [3]. The article provides a classification of approaches to the prediction problem [2], identifies its main elements, provides a general description of the problem of predicting the behavior of the computer system based on ML tools, and notes the connection between the difficulty of prediction future values of time sequence and general problem of ML [4].

The criteria for mathematics models of forecasting are defined. The impact of randomized time sequence' probabilistic features on its effectiveness prediction is analyzed, in particular possibility of prediction without exact considering the probabilistic distribution of the predicted data. It is shown that a number of characteristics reflect a bidirectional link between past events and the future values of time series modeling traffic; therefore, we consider nonlinear algorithms for predictors.

B. Description on the forecasting problems

Statistical mathematical models for forecasting randomized periods, series, and procedures have an existence of over eighty years [5], concentrated investigation in this area continues to this day. This is due to the huge variety of problems encountered when processing large data sets. These problems may require various accurate forecasting requirements, models of data, and time (efficiency) needs. As an example, in many cases of forecasting the values of financial indices, the condition the minimum mean squared error (MSE) of the foreseen data relative to the correct values is used (the basis of which the criteria and its several linear alterations are founded). However, this criterion is clearly inappropriate when predicting the sign of an increment. In some issues, it is quite fair to utilize certain numbers averaged across large previous samples as a criterion for the effectiveness of prediction, while some require specific estimates. There are prediction algorithms that employ numeric measurement data devoid making any presumptions about its statistical or probability qualities (for example, neural networks [1]), those that obliquely consider data as a sample unlike a regular sedentary process [1] or sequence [6]. In [2], a classification of approaches to the prediction problem was substantiated, which we will use further:

– A mathematical problem predicting a mystery precondition probability circulations (probability-theoretic);

– A mathematics problem for optimum decision making.

In the probability-theoretic approach [7], the prediction task is shadows: agreement at time t, according to the implementation of the process $x_1, x_2, \dots, x_t$, considered (here and below) as a time series [8], an estimate of the conditional probabilities $\gamma(x_{t+1}|x_1, x_2, \dots, x_t)$ on time t + 1 we shall examine how value we predicted is accomplished. The additional this estimation, the better prediction. The mathematical model of the $\gamma(x_{t+1}|x_1, x_2, \dots, x_t)$, estimate, like the restrictive prospect itself, is named a forecaster in literature [7]. Lengthways γ( ), forecasters are named software produces a loss function $(b, x), b \in B, x \in X,$, select the then assessment $^\wedge b_t$ as the predicted one, given the past states of $x_1, x_2, \dots, x_{t-1}$, so that the total losses entire $x = x_1, x_2, \dots, x_{t-1}$, sequence, i.e., the predicted $^\wedge b_t$ would minimize probable a posteriori fatalities. The policy is a meaning that revenues the output trajectory y for every sequence x in accordance with the provisional circulation purpose of $P(y|) or\, \gamma( )$ [9]. The third formulation forecasting problematic, which we will request algorithmic forecasting prototypical, consists by various conceptual models for attaining a forecast that do not obviously comprehend several of the elementary, in individual the limitations of the purposes called since firm machine learning lending library [10]. For example, this could be a search tree in the well-known predictor , if user theoretical classical deliberates the procedure of constructing forecast a certain procedure by way of a pursuit for adequate (according to designated efficiency principle) forecast amongst usual its probable values in motionless information domain [11]. It applies various methods of predicting future values of time series to the training set. This can be viewed as a sequential search for prediction models that improve on past results. These prototypes be able to be established overhead curriculums of exact simulations, but the involvement linguistic algorithmic prototypical does not contain of parameters and variables of these models. Since prediction is made on a finite sample, a natural performance measure for all three types of models is some statistical estimate that depends on the statistical properties of the data that must be reflected in estimate models.

C. Result and Discussion

Theoretical model of the forecast problematic

In [2] the following generalized form of presentation of theoretical prototypical of the forecast problematic is proposed:

$$FM = \{M_D^\theta(s), M_L, F_{ML}\}. \qquad (1)$$

Here $M_D^\theta(s)$, is a forecast model demarcated by data category D (binary, real), by a possibility amount θ *on* D and a construction S arranged D, which is understood as a way of structuring facts of category D as contribution variables of the predictor (e.g. dividing the data into a training and test sample, scalar or vector data, etc.); M_L is prototypical of losses (penalties) on or after the accepted estimate by a forfeiture purpose $L(X_S, Y_S)$ *where* $X_S, Y_S \in D$ are, respectively, the observed and predicted data with structure $S; F_{ML} = \{f_1, \ldots, f_m\}$ is established forecasters founded on $M_D^\theta(s)$ *and* M_L models, i.e. ways of calculating the estimate of the forthcoming value by the restricted chance delivery of the foretold value $\gamma(\ )$.

It is assumed that the predictor $f_p, p = 1, \ldots, m$, continuously resembles to an allowable forecast model, i.e. very its participation data and restrictions container be exclusively indomitable within the background of the prototypical currently under deliberation. It consents one to denote periods of models bestowing to the types of forecast data, the possibility processes used, and the forfeiture purposes after making verdicts around the suitability of the obtained prediction. In case, the need to take hooked on explanation the probabilistic individualities of the data follows from the above remark about the statistical nature of the assessment of the quality of the prediction.

Example 1.

Let $\theta = Prob\,(x_i = 1)$ possibility quantity $D = \{0,1\}^n$ usual classifications measurement n. After the topic of interpretation of (1), n is a stricture of the assembly S for the classification $X_t = \{x_1, x_2, \ldots, x_i, \ldots, x_{t-1}, x_t\} \in D$.

Let us consider a arrangement zeros and ones as a arrangement of sovereign trials with a loss function 0/1 loss a extent of the forecaster user's losses since guess mistakes equivalent to the amount of improper forecasts on a arrangement of interval n [6], i.e.

$$e_\Theta = E_\Theta\left(\sum_i^n \frac{I(b_t \neq x_t}{n}\right), \qquad (2)$$

Where $\Theta = Prob\,(x_t = 1)$ is the parameter of the distribution; I( ) is the indicator function; b_t *and*x_t are the predicted and true values, respectively; E_Θ means the mathematical expectation for the distribution .

Note that formula (2) is universal in the sense that it does not depend on the distribution on D. In the literature, criteria similar to (2) are called experiential risk minimization criteria [9].

We can deliberate the subsequent forecaster $b_t = f(x_{t-1})$ *where* $f \in F_{ML}$ then the defeat function is demarcated as the optimum prediction instruction for minimum loss principle with respect to e_n [6]:

$$b_t = \begin{cases} 1, & if\ Prob(1) > \frac{1}{2}; \\ 0, & if\ Prob(1) < \frac{1}{2}, \end{cases} \qquad (3)$$

In the case of $Prob\ (0)\ =\ Prob\ (1)\ =\ \frac{1}{2}$ the forecast is not fulfilled. Trendy circumstance, as is tranquil realize, the typical loss is equivalent to the fault possibility $1 - \Theta\ if\ \Theta > \frac{1}{2}$, $and\ \Theta$, if $\Theta \frac{1}{2}$ which can be conveyed

$$L_p = \min(\Theta, 1 - \Theta).$$

Meanwhile we are talking about using ML for prediction, let us consider whether there is a need to transform the formulated conceptual model of prediction to take into account the specifics of ML. Conceptual model of machine learning. One of the well-known and used formalizations of ML, so called because it estimates the probability of achieving a given classification accuracy after training [4]. In prediction is considered as accepting a hypothesis $h\ \in\ H,$ is a set of hypotheses about the predicted value of a random variable on a data set X with an unknown distribution, for which the objective function L computes the quality characteristics of the predicted values Y (labels [4]), by which a predictor $h_S\ :\ X\ \rightarrow\ Y$ should be selected, where S is a finite sample from X corresponding to the structure S (which may, for example, determine which elements from X should be included in the sample—see the example below). In the previous notation, accepting a hypothesis h means choosing a predictor f that realizes the prediction of some value from X. For example, in an ML tool such as linear regression, the learning problem may be defined on a subset of real numbers R for some number of variables (regression coefficients) d, with a predictor f( ) computed on a set of known observed values Y (labels in terminology) in the training set. The prediction hypothesis then corresponds to finding a linear function h that best approximates the relationships between variables X and Y (e.g., traffic volume versus time and attack intensity).

Thus, assessing the quality of the learning model requires using the same mathematical tools (probability measures, loss functions) on data sets of the same type as the prediction problem. At the same time, the structuring of S can divide the set of observed data into subsets of predictor training and its evaluation (testing). Therefore, the conceptual model (1) also covers the learning aspect. Below is a more complex example of setting and analyzing the influence of the structure of S.

Example 2.

Let the prediction in the binary sequence according to rule (3) be made taking into account that the k values preceding the predicted moment t are unit values. If the parameter Θ (i.e. the probability Prob (1)) is unknown, then it is estimated as the proportion of successes (units) Θ_n on some finite subsequence of length n and Θ is the mathematical expectation of this estimate (on the set of such n-subsequences). In [12] it is shown that the conditional mathematical expectation of the estimate Θ_n provided that the estimate is made on the subsequence following the specified k units, is less than the true value Θ. This means that within the framework of rule (3), the a priori estimate of the prediction error at the moment following the specified k successes, when using Θ_n as Prob(1), will be higher than in the absence of fixing the specified k events. Thus, the properties of the data structure S in model (1) (in this example, the presence of k selected observations) affect the efficiency of the forecast. At the same time, ignorance of the distribution of data D in (1) means the unfeasibility of an a priori estimate of the correctness forecast by this forecast algorithm. Therefore, a theoretical model is needed that allows one to rely not on exact awareness disseminations, but awareness of certain possessions of scatterings.

Inspiration of probabilistic possessions of arbitrary time sequence on forecasting competence

The conditional distribution predicted data $\gamma(\,)$ (predictor) is determined by the statistical dependence of the data in the past and future (linear or nonlinear prototypes, correspondence constructions) and stationarity the information in intelligence selected probability rules leading the alteration in data during surveillance remain unchanged, at least on the training sample for which prediction should through. Aimed at instance, extremely connected time sequence, linear regression models (for example, autoregressive integrated moving average)[13] give a better forecast than in the case of their weak correlation. In additional arguments, if, since the prototypical possessions of time sequences, we limit ourselves merely to the association possessions of arrangement and the above-considered provisional possibility of the forecast $\gamma(x_{t+1}|x_1, x_2, \dots, x_t)$ is obviously connected to connection (as is the situation, aimed at sample, for a Gaussian distribution, at what time correspondence is corresponding to unconventionality), then, based on the gradation of connection, one can imagine one or additional excellence of forecast by a lined reversion forecaster.

Let us consider which of the measured characteristics of processes can be used to assess the predictability of data by a particular predictor. According to the proposed intangible model of the forecast problematic (1), which integrates probabilistic models of both data and decision making (through the loss function), we consider the following classes of measured characteristics of samples of random time series associated with probabilistic laws of behavior:

- Properties of autocorrelation functions of events remote in time and the structural properties of the trajectories of the corresponding processes determined by them,
- Nonlinearity of the relationships between past and future values of time series,
- Possible influence of the properties of heavy tails of distributions,
- Estimates of the degree of stationarity of the time series under consideration.

Autocorrelation properties of data and their impact on predictability. The ability to predict using available data obviously depends on the degree of correlation between their values corresponding to remote events. A well-known model of such dependence, which, as shown by long-term studies, is observed in traffic for many real networks [14], is the so-called property of a random process. A time series $X_t, t \in Z$, its covariance function [15]

$$r(t) = E\big((X_0 - EX_0)(X_t - EX_t)\big) \sim c_\gamma |t|^{2-2H} \text{ as } t \to \infty .$$

Here $c_\gamma > 0$ is some constant; H is the Hurst exponent, which is calculated as the mathematical expectation:

$$H = \lim_{n\to\infty} E[\tfrac{R(n)}{S(n)}] ,$$

Where $R(n)$ is assortment (distance) of accrued nonconformities of the main morals from the callous value of sequence, $S(n)$ is typical unconventionality, n is value of time intermission (the amount of facts on a section of period sequence). The Hurst advocate H characterizes the period unravelling connected (for sample, more strongly than a convinced inception value) traffic samples from each other. The most well-known model of a process [16], which has the specified correlation function. It is ordinary to adopt that circulation forecasting with properties can be quite effective (accurate) from the point of view of the loss function used, meanwhile data approximately the preceding obviously influences forecast forthcoming. It is significant that the curves of arbitrary procedures with possessions devour comparison, i.e. repeatability of designs changed period rulers [3, 17]. It is known that this property is most pronounced for $H \in \left(\frac{1}{2}, 1\right)$ [7, 18]. It is imperative, as soon as the preliminary fractional Brownian procedure (modeling this traffic with properties) is non-stationary, its increases are definitely associated if $H \in \left(\frac{1}{2}, 1\right)$, uncorrelated if $H = ½$, and negatively correlated $if\ H \in \left(\frac{1}{2}, 1\right)$. This assets is significant for foreseeing the emblem of the increases. $if\ H \notin \left(\frac{1}{2}, 1\right)$ then loses its self-similarity properties and in this case it cannot be used for prediction. An example of a method that depends on the sign of the autocorrelation of the values of the time series is the method [19]. In [19, 20], the efficiency of forecasting the emblem of alteration $y_{i+1} - y_i$:

$$sign(y_{i+1} - y_i) = sign(y_i)\ , \quad (4)$$

Where $y_i = x_i - Ex_i, i = 1, \dots, t$ is a sequence of centered random variables with zero mean. An indication of opportunity of consuming (4) as a forecaster is lack of correlation (or very feeble association) of consecutive variances in values of the period sequence (procedure) under consideration by a propensity toward undesirable association of the $y_{i+1} - y_i\ and\ y_i$ meanwhile correspondence of two arbitrary variables revenues a propensity to alteration in the conflicting bearing [16].

Recall that for $H < 1/2$ the increments have a negative correlation and demonstrate a short-range dependence, for $H > 1/2$ the autocorrelation of the increments is positive. Thus, two consecutive increments, as a rule, have the same direction. A sample of practicality of forecasting the emblem of circulation increases is, meant for sample, a situation as soon as occurrences on network are approved in a method that cannot be noticed by antivirus software, and predictors consume to depend on an analysis of variations in circulation capacity or the trend of variations in the intensity of recording in the log file [14]. Stationarity most of the real processes under consideration in are formally not stationary, and stationarity can be spoken of with confidence individual to a convinced degree. To assess this grade, arithmetical assessments on behalf of stationarity [17], but it is challenging to comprise their consequences in the forecast prototypical, meanwhile nothing can be used in the model except for instinctive qualitative reasoning that motionless procedures are improved forecast than motionless ones. However, for processes, more substantiated heuristic rules for the relationship between stationarity and predictability can be obtained. In [19], it is exposed that comparatively little explanations can careful motionless, at slightest in terms of alteration, since with resemblance the variance of the model mean reductions additional gradually than the mutual of the model size, i.e. $var\left(X^{(m)}\right) \sim a_2 m^{-\beta}, m \rightarrow \infty, 0 < \beta < 1$,

Where a_2 is a confident continuous. It is important autocorrelations decline by way of hyperbolic purposes rather than exponentially. Consequently, by means of a convinced gradation of conformism, we can correspondingly express stationarity with respect to autocorrelation purposes. Hurst advocate H as a stationarity indicator. In [18], it is revealed Hurst advocate H can be used as a stationarity standard meant for a analogous procedure. Values of $H > 1$ indicate (taking into account the statistical difficulty of its estimation based on a finite sample in the vicinity of $H = 1$) that the predicted process cannot be considered stationary. On behalf of a motionless comparable procedure $H \in (0.5, 1)$ the earlier the cost Hurst advocate to 1, the leisurelier the variance deteriorations as the period ruler growths, and the road traffic becomes more pulsating and, accordingly, non-stationary.

The influence of properties and stationarity properties on the technique of solving forecasting problems. If the cause callous rectangular measure of forecast accuracy is used and the series can be considered stationary, then the square of the linear prediction error aimed at a still period sequence is proportional to $1 - \rho^2$ [8], where $\rho = ACF(1), ACF(1)$ resources the correlation purpose in interval 1. Taking into account that on behalf of procedure autocorrelation purpose depends exponentially on H [20], a monotonic association amid the source callous square fault of lined forecast and the Hurst advocate in the intermission $(1/2, 1)$ is obvious. This examination container valuable, for example, as soon as consuming such forecasting software utensils as the autoregressive affecting typical prototypical with a moving average besides autoregressive combined touching usual. On behalf of the forecast of unadulterated arbitrary tread, the prediction of the sign change (4) can be useful, as discussed above. The theory of linear prediction with the criterion, founded on the prototypical, undertakes that period sequence has a determinate mathematical expectation then alteration. Nevertheless, aimed at period sequence with possessions, this may not continuously be true, in particular if the data streams have a Pareto distribution [21]. The inspiration parameters of probability distributions of procedures with on forecast. Frequent revisions of sequence have exposed their deliveries, as distributions with heavy tails, are close to the Pareto distribution:

$$Pareto\ (X) = \frac{ab^a}{x^{a+1}},$$

Where $x \geq a$. In this case, the mean and variance:

$$\mu\_Pareto\ (x) = \frac{ab}{a-1}\ ;\ Var\_Pareto = \frac{ab^2}{(a-1)^2(a-2)}\ ,$$

Where $x \geq a > 0, b > 0$ obviously, μ_Pareto and Var_Pareto do not exist for $a = 1\ and\ 2$, and the practice of forecasters to infer the obviousness of sequence with respect is unacceptable. Accounting for nonlinear properties of traffic trajectories. The self-similarity of the traffic trajectory considered above reproduces nonlinear association among the previous and the forthcoming, where degree of nonlinearity depends on H [2, 16]. As noted, for $H = \frac{1}{2}$ the property is lost [20]. This means that the H parameter can designate whether forecasted period sequence consumes the possessions of an unadulterated arbitrary saunter, for which it is obviously difficult to effectively use ML, or whether it has some correlation structure, which would mean the potential for effective forecasting. For example, in [22] for the Internet of Things (IOT), the H constraint on behalf of network circulation is used as a priori information in training. Meanwhile the comparison possessions is glowing interpretable in expressions of how consultants describe traffic, its practice increases interpretability model, namely: the aptitude to comprehend in what way statistics possessions, in specific resemblance, affect the superiority of the prediction.

A number of prediction tools use self-similarity destruction as a predictor methods founded the reflection that attendance of a distributed denial of service (DDoS) occurrence decreases the gradation of comparison of ordinary circulation, meanwhile DDoS gears do not produce personality related circulation [14].

Nonlinear prototypes and neural processes

First of all, we note that an obvious example of the application of a nonlinear algorithm is the considered predictor (4). However, a nonlinear forecasting model in modern PIs is usually applied whichever in nonlinear reversion prototypes or in artificial neural complexes [23], on behalf of sample, deep learning models driven by data, in totaling to machine learning algorithms by way of sustenance trajectory deterioration [10, 21]. A neural network can consume difficulties owing to terminated fitting, at what time the prototypical works thriving lone with data from the exercise set, familiarizing to keep fit examples in its place of knowledge to categorize statistics that was not involved in exercise. It is laidback to understand such a nonlinear phenomenon as resemblance can donate to this. On the similar period, such a shared method to decrease over appropriate as failure rotating off certain neurons by a convinced prospect on certain intermission of data as of the exercise procedure, might not effort because the exercise set determination alike to the preceding one. Unexpected variations [5] in mist circulation can effortlessly disordered in a neural network through circulation irregularities, which also indications to keeping fit disorganization. The severity of such non-stationarity can be estimated by the values of H.

Support vector regression [23] is also used as a nonlinear approach to traffic forecasting, but the special of appropriate kernel purposes and optimum strictures is same complex. Samples of such a choice are briefly discussed in [6] (interesting results of studying the influence of data models on the use of the self-similarity property are given in [7]). On or after the presented examination of the requirement of the competence of consuming forecast prototypes on explicit possessions of period sequence, it shadows that a normal method of taking addicted to explanation the requirement of the possessions of predictors happening the conduct of the data (stationarity, nonlinearity of requirement) should be their connected choice in the procedure of resolving network management problematic. For example, in [2], artificial neural network and lined deterioration procedures were considered to predict the reply period and amount of cloud facilities at dissimilar phases of reserve usage.

D. Conclusion

By way of the literature examination demonstrations, numerous contemporary forecast tools founded the moralities of machine learning do not effort efficiently enough outstanding to the noticeable nonlinearity of circulation variations, their non-stationarity, also the essential on behalf of a great volume of preceding explanations. This article is an effort to somewhat organize besides categorize the enormous flow of publication off contemporary approaches, methods, and prototypes for predicting statistics of numerous natures from the opinion of interpretation of the opportunity of a priori valuation of effectiveness of predicting statistics standards by corresponding. For this purpose, the characteristics of probabilistic models explicitly or implicitly used in modern software prediction tools are highlighted, allowing one to measure the attendance of important non-stationarity trendy facts streams, some nonlinear effects, and the specifics of distribution laws that affect the effectiveness of predictor training.

E. References

[1] O. Aouedi, V. A. Le, K. Piamrat, and Y. Ji, "Deep learning on network traffic prediction: Recent advances, analysis, and future directions," ACM Computing Surveys, vol. 57, pp. 1-37, 2025, https://doi.org/10.1145/3703447.

[2] Y. Feng, J. Li, J. Mirkovic, C. Wu, C. Wang, H. Ren, et al., "Unmasking the Internet: A Survey of Fine-Grained Network Traffic Analysis," IEEE Communications Surveys & Tutorials, 2025, http://dx.doi.org/10.1109/COMST.2025.3545541.

[3] D. Jiang, Z. Wang, X. Liu, Q. Xu, T. Zou, R. Zhang, et al., "Toward Synthetic Network Traffic Generating in NTN-Enabled IoT: A Generative AI Approach," IEEE Internet of Things Journal, vol. 12, pp. 2174-2187, 2025, https://doi.org/10.1109/JIOT.2024.3468209.

[4] Z. Yang, Y. Jin, J. Liu, X. Xu, Y. Zhang, and S. Ji, "Research on cloud platform network traffic monitoring and anomaly detection system based on large language models," arXiv preprint arXiv:2504.17807, 2025, https://doi.org/10.48550/arXiv.2504.17807.

[5] J. Koumar, K. Hynek, T. Čejka, and P. Šiška, "CESNET-TimeSeries24: Time Series Dataset for Network Traffic Anomaly Detection and Forecasting," Scientific Data, vol. 12, p. 338, 2025, https://doi.org/10.1038/s41597-025-04603-x.

[6] C. Liang, Z. Zhang, Y. Li, Y. Wang, M. He, F. Xia, et al., "Simulation, prediction and optimization for synthesis and heavy metals adsorption of schwertmannite by machine learning," Environmental Research, vol. 265, p. 120471, 2025, https://doi.org/10.1016/j.envres.2024.120471.

[7] J. Liu and J. Chen, "Applications and trends of machine learning in building energy optimization: A bibliometric analysis," Buildings, vol. 15, p. 994, 2025, http://dx.doi.org/10.3390/buildings15070994.

[8] A. Equbal, A. Equbal, Z. A. Khan, and I. A. Badruddin, "Machine learning in additive manufacturing: A comprehensive insight," International Journal of Lightweight Materials & Manufacture, vol. 8, 2025, https://doi.org/10.1016/j.ijlmm.2024.10.002.

[9] F. Gyaase, O. S. Okunlola, A. M. Olusanya, A. Adedokun, O. I. Somoye, and A. T. Erhieyovwe, "Utilizing machine learning for predictive analysis of emission levels to ensure compliance in refinery operations," International Journal of

Engineering & Technology, vol. 14, pp. 24-29, 2025, http://dx.doi.org/10.14419/6x85gr51.
[10] M. Abdellatief, L. S. Wong, N. M. Din, A. N. Ahmed, A. M. Hassan, Z. Ibrahim, et al., "Sustainable foam glass property prediction using machine learning: A comprehensive comparison of predictive methods and techniques," Results in Engineering, vol. 25, p. 104089, 2025, https://doi.org/10.1016/j.rineng.2025.104089.
[11] S. Uddin, S. Yan, and H. Lu, "Machine learning and deep learning in project analytics: methods, applications and research trends," Production Planning & Control, vol. 36, pp. 873-892, 2025, https://doi.org/10.1080/09537287.2024.2320790.
[12] K. Zhong, X. Tan, S. Liu, Z. Lu, X. Hou, and Q. Wang, "Prediction of slope failure probability based on machine learning with genetic-ADASYN algorithm," Engineering Geology, vol. 346, p. 107885, 2025, http://dx.doi.org/10.1016/j.enggeo.2024.107885.
[13] G. Zhou, X. Guo, Z. Liu, T. Li, Q. Li, and K. Xu, "Trafficformer: an efficient pre-trained model for traffic data," in 2025 IEEE Symposium on Security and Privacy (SP), 2025, pp. 1844-1860, https://doi.org/10.1109/SP61157.2025.00102.
[14] B.-l. Cen, Y. Xue, K. Zhang, L. Jia, and H.-d. He, "Study on traffic flows with connected vehicles and human-driven vehicles," Applied Mathematics and Computation, vol. 490, p. 129182, 2025, https://doi.org/10.1016/j.amc.2024.129182.
[15] S. Jung, A. Hardy, and M. J. Kochnederfer, "Inferring traffic models in terminal airspace from flight tracks and procedures," Journal of Aerospace Information Systems, pp. 1-11, 2025, https://doi.org/10.48550/arXiv.2303.09981.
[16] L. Huang, J. Wang, Y. Huang, M. Zhu, Y. Wen, and Y. Zhou, "Probabilistic prediction of ship destinations based on traffic pattern awareness in maritime networks," Ocean Engineering, vol. 316, p. 119933, 2025, https://doi.org/10.1016/j.oceaneng.2024.119933.
[17] J. Chen, L. Yang, Y. Yang, L. Peng, and X. Ge, "Spatio-temporal graph neural networks for missing data completion in traffic prediction," International Journal of Geographical Information Science, vol. 39, pp. 1057-1075, 2025, https://doi.org/10.1080/13658816.2024.2381221.
[18] S. Mtsweni, B. F. Bakare, and S. Rathilal, "Artificial Neural Network-Based Prediction of Clogging Duration to Support Backwashing Requirement in a Horizontal Roughing Filter: Enhancing Maintenance Efficiency," Water, vol. 17, p. 2319, 2025, https://doi.org/10.3390/w17152319.
[19] R. Buller, J. Damborsky, D. Hilvert, and U. T. Bornscheuer, "Structure prediction and computational protein design for efficient biocatalysts and bioactive proteins," Angewandte Chemie, vol. 137, p. e202421686, 2025, https://doi.org/10.1002/anie.202421686.
[20] L. Dai, W. Zhu, X. Quan, R. Meng, S. Chai, and Y. Wang, "Deep Probabilistic Modeling of User Behavior for Anomaly Detection via Mixture Density Networks," arXiv preprint arXiv:2505.08220, 2025, https://doi.org/10.48550/arXiv.2505.08220.

[21] P. Zhao, W. Zhang, X. Cao, and X. Li, "Denoising diffusion probabilistic model-enabled data augmentation method for intelligent machine fault diagnosis," Engineering Applications of Artificial Intelligence, vol. 139, p. 109520, 2025, https://doi.org/10.1016/j.engappai.2024.109520.
[22] S.-V. Oprea and A. Bâra, "Unveiling the nexus between energy storage and electricity markets in academic publications. A data-driven analysis of emerging trends and market dynamics using NLP, sentiment analysis and probabilistic modeling," Journal of Energy Storage, vol. 106, p. 114917, 2025, https://doi.org/10.1016/j.est.2024.114917.
[23] C. Surget, S. Dubreuil, J. Morio, C. Mattrand, J.-M. Bourinet, and N. Gayton, "A sensitivity analysis based trade-off between probabilistic model identification and statistical estimation," Reliability Engineering & System Safety, vol. 254, p. 110545, 2025, https://doi.org/10.1016/j.ress.2024.110545.