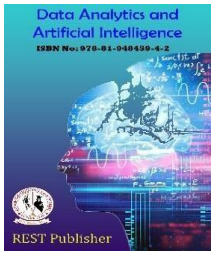
Data Analytics and Artificial Intelligence

Vol: 6(1), 2026

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/

DOI: https://doi.org/10.46632/daai/6/1/6



Machine Learning Analysis of Structural Mathematics: Predicting Matching Scores Using Random Forest Regression

Saandeep Sreerambatla

Lead ML Engineer, NobleAI

Corresponding Author Email: saandeep241@gmail.com

Abstract: *Creative mathematics represents a fundamental approach that differs from classical mathematics by rejecting the rule of the excluded middle and demanding explicit constructions rather than evidence. This study investigates the quantitative properties of creative mathematical structures through machine learning analysis, specifically examining four key parameters: axiom richness, evidence length, creative complexity, and creative fit score. Using a dataset of 100 observations, random forest regression was used to model and predict the creative fit score based on other parameters. Descriptive statistical analysis revealed significant variation across all measures, with axiom richness ranging from 20.44 to 98.95 (mean = 57.61), evidence length ranging from 33.27 to 493.25 steps (mean = 263.98), and creative complexity ranging from 10.20 to 49.60 (mean = 30.70). Correlation analysis found a strong positive relationship between theoretical richness and constructive fit score (r = 0.63), while evidence length showed a moderate negative relationship with fit (r = -0.64), indicating that longer evidence may reduce fit and performance. The random forest regression model showed exceptional performance on the training data, with an R^2 of 0.989 and minimal error measures (RMSE = 1.52, MAE = 1.23), while maintaining strong generalization on the test data (R^2 = 0.944, RMSE = 4.67). The findings indicate that systems with rich theoretical foundations and moderate complexity tend to achieve higher constructive fit, while excessive evidence length may compromise performance. This research contributes to the computational understanding of constructive mathematics by quantifying abstract theoretical properties and demonstrating the applicability of machine learning techniques in mathematical foundation analysis.*

Key Words: *Constructive Mathematics, Random Forest Regression, Theoretical Richness, Evidence Complexity, Machine Learning.*

1. INTRODUCTION

On a reflective day, a mathematician might find themselves pondering the nature of constructive mathematics. Although they may have encountered arguments supporting constructivist approaches, such perspectives might not resonate deeply with them especially if they are less concerned with philosophical underpinnings. [1] Modern constructive mathematics is shaped by two central concepts typically linked to constructive type theory and Dobbs' theory. While both are individually consistent with classical mathematics, they are mutually incompatible due to specific theoretical distinctions, which will be briefly discussed. [2] Constructive mathematics diverges from classical logic by adhering to intuitionist principles. Since Brower first introduced intuitionism, various interpretations and frameworks grounded in intuitionist logic have emerged. In this context, the term "structural mathematics" refers to Bishop-style mathematics, originally formulated by Egret Bishop in "Foundations of Constructive Analysis." This form of mathematics relies on constructive logic to ensure that proofs correspond to actual constructions. [3] In earlier work with Sabin, the need to unify foundational systems within constructive mathematics was emphasized seeking common ground across different frameworks such as axiomatic set theories. [4] A central goal of constructive inverse

Copyright@ REST Publisher

mathematics is to categorize mathematical theorems based on the logical or functional existence principles needed to establish them, spanning frameworks from intuitionist and constructive recursive systems to classical mathematics. This process focuses on determining which principles are both necessary and sufficient for deriving theorems within a foundational system. [5] Despite a general agreement among constructivist schools on the validity of countable choice, this remains a deeply debated and unresolved issue within the field arguably one of its greatest philosophical and logical challenges. [6] Finally, the discussion extends to educational research grounded in the investigative classroom approach, which emphasizes the development of teaching strategies that effectively support student learning in mathematics. This work is informed by direct classroom observation and analysis of teachers' cognitive processes during instruction. [7] It provides a level of abstraction that helps reveal underlying algebraic structures often obscured in traditional representations. This paper exemplifies how type theory serves as an effective language for articulating essential features that any foundational system in constructive mathematics should encompass. [8] The origins of modern constructive mathematics can be traced back to 1907, with the publication of L.E.J. Brower's doctoral thesis Over de Grondslagen deer Wiskunde, which introduced his intuitionist philosophy a broader worldview that extends beyond mathematics alone. [9] Constructive mathematics departs from traditional mathematics by rejecting the law of the excluded middle and, correspondingly, the principle of proof by contradiction (reduction ad absurdum). [10] While some classical theorems may be difficult or even impossible to prove constructively, constructive proofs, when available, tend to be highly informative and reveal deeper computational or procedural content. [11] One might expect that over the past thirty years, constructive mathematicians would have linear logic as an alternative to intuitionist logic. However, this has largely not occurred. It is arguable that aspects of linear logic are already inherently present in the foundations of constructive mathematics, dating back to Brower's early ideas. [12] My approach begins with a historical interpretation of constructive mathematics, distinguishing it from set-theoretic (or centurion) mathematics. Its philosophical and mathematical roots can be traced to Kronecker, with ancient influences from thinkers such as Proclus and Speusippus, and further shaped in the 20th century by Brower and Egret Bishop. Within this framework, a constructive function refers to a function whose existence is established through direct, explicit construction. [13] This dissertation focuses on the foundational aspects of constructive mathematics. Given that different branches of mathematics model aspects of reality in varied ways, there are intersections between the foundations of mathematics and foundations of physics. For instance, some of our explorations lead to hypothetical physical experiments that may provide statistical insights into whether the physical universe is fundamentally deterministic. [14] Constructive mathematics also encompasses the study of explicitly definable numbers, constructive topologies, and constructive metric spaces. It is defined by its commitment to concrete, step-by-step reasoning and the rejection of non-constructive proof techniques, instead requiring verifiable constructions as the basis for all mathematical claims.

2. MATERIALS AND METHOD

Materials:

Axiom Richness:

- Indicates the amount and diversity of basic axioms available within a constructive mathematical framework.
- A higher value indicates that the system is built on a broad and explicit set of axioms, allowing for the representation of more complex mathematical structures.
- In practical terms, it measures the theoretical depth and expressiveness of the constructive system.

Proof Length:

- Indicates the number of logical steps or derivations required to complete a constructive proof.
- A shorter proof length indicates greater efficiency in obtaining results, while a longer one may indicate greater complexity or redundancy.
- It captures the computational effort or logical complexity of the constructive reasoning process.

Constructive Complexity:

- Indicates the degree of difficulty in generating and verifying constructive proofs.
- It accounts for factors such as algorithmic complexity, resource requirements, and the complexity of constructive constraints (e.g., avoiding constructive principles).
- A higher value indicates that the proof requires more sophisticated constructive reasoning or computational resources.

Constructive Suitability Score:

- A composite measure that assesses how well a mathematical structure, proof, or approach conforms to constructive principles (i.e., ensuring that existence proofs provide explicit constructions, avoiding the excluded middle if they are invalid, etc.).
- It combines elements such as axiom richness, proof length, and constructive complexity into a performance-based score.
- Higher scores indicate better adherence and practical utility of the constructive approach.

Random Forest Regression: An ensemble approach is used for regression tasks in the supervised learning technique known as random forest regression. Instead of using a boost technique, it employs a packing strategy. Trees in a random forest are constructed concurrently and function independently of one another throughout the training process. By choosing the most common class for classification or by averaging predictions for regression, a random forest generates a number of decision trees during training, after which their results are combined to provide results. The hyper parameter, which is a specific percentage of the total, limits the number of features that can be separated at each node. This restriction encourages the balanced use of all predictor variables by preventing the model from depending too much on any one characteristic. Furthermore, when generating splits, each tree takes a random sample of data from the original dataset, adding variation that lessens over fitting.

3. ANALYSIS AND DISCUSSION

The dataset in Table 1 provides detailed values for axiom richness, proof length, construction complexity, and construction fit score across various cases. The distribution of these parameters highlights both the diversity and underlying trends within constructive mathematical structures. Starting with axiom richness, the values range from 20.44 to almost 99, showing a wide range of systems—from minimal axiomatic foundations to highly enriched ones. Some cases, such as 97.59 and 98.95, indicate very strong foundation foundations, while others range around 21–25. This variation indicates different levels of expressiveness among the observed systems. The proof length values show a significant spread from 33 steps to 493. Short proofs, such as those near 33–70, indicate efficient derivations, while longer ones, especially beyond 450, indicate more complex or redundant reasoning paths. This diversity reveals that constructivist frameworks can provide both concise and highly extended evidence, depending on their axial depth and problem structure. For constructivist complexity, the dataset reflects values between 10.20 and nearly 49.60, with some evidence being relatively straightforward, while others indicate that advanced constructivist reasoning is required. Higher complexity, often combined with moderate to high evidence lengths, suggests challenging proof tasks. Interestingly, some cases show lower complexity but longer proofs, emphasizing structural rather than methodological difficulty. The constructivist fit score varies from about 21.74 to 88.60. Most scores are in the moderate-high range (40–70), indicating that most systems show acceptable alignment with constructivist principles. Notably, scores above 80 indicate strong constructivist fit, where both performance and constructivist rigor are in balance. Together, the dataset captures a broad landscape of creative mathematical behavior, illustrating the trade-offs between print resourcefulness, resourcefulness, complexity, and relevance.

TABLE 2. Descriptive Statistics

	Axiom Richness	Proof Length	Constructive Complexity	Constructive Suitability Score
Count	100	100	100	100
Mean	57.614459	263.98091	30.704053	55.629361
Std	23.799153	137.762289	11.73705	15.029465
Min	20.441769	33.267501	10.202463	21.748464
25%	35.456061	143.742127	21.075195	44.356381
50%	57.131396	267.643683	32.502197	55.898147
75%	78.41625	390.106291	40.094678	66.680836
Max	98.950955	493.255713	49.602154	88.605211

The descriptive statistics in table 2 provide a quantitative overview of four key parameters: axiom richness, proof length, construct complexity, and construct fit score. Each parameter is measured on 100 observations, providing insight into central tendencies and variations within the dataset. For axiom richness, the mean value is approximately 57.61, with a standard deviation of 23.79. This indicates a moderate level of axiomatic diversity on average, but also considerable variation across observations. Values range from a minimum of 20.44 to a maximum of 98.95, suggesting that some systems are based on very limited axioms and others are very enriched. Proof length has the highest variability among the parameters, with a mean of 263.98 and a large standard deviation of 137.76. The range is also wide, from as short as 33.27 steps to as long as 493.25 steps, reflecting significant differences in the efficiency and complexity of the evidence across contexts. The mean of 267.64 indicates that many of the evidence is long, closely matching the mean. Constructive complexity shows a mean of 30.70, with values ranging between 10.20 and 49.60. The interquartile range (21.07 to 40.09) highlights that most of the evidence falls within the moderate complexity range, balancing feasibility with rigor. The constructive fit score averages 55.62, with relatively low variability (SD = 15.03). Scores range from 21.74 to 88.60, with an interquartile range of 44.35 to 66.68, indicating that most approaches are moderately to highly suitable for constructive mathematics.

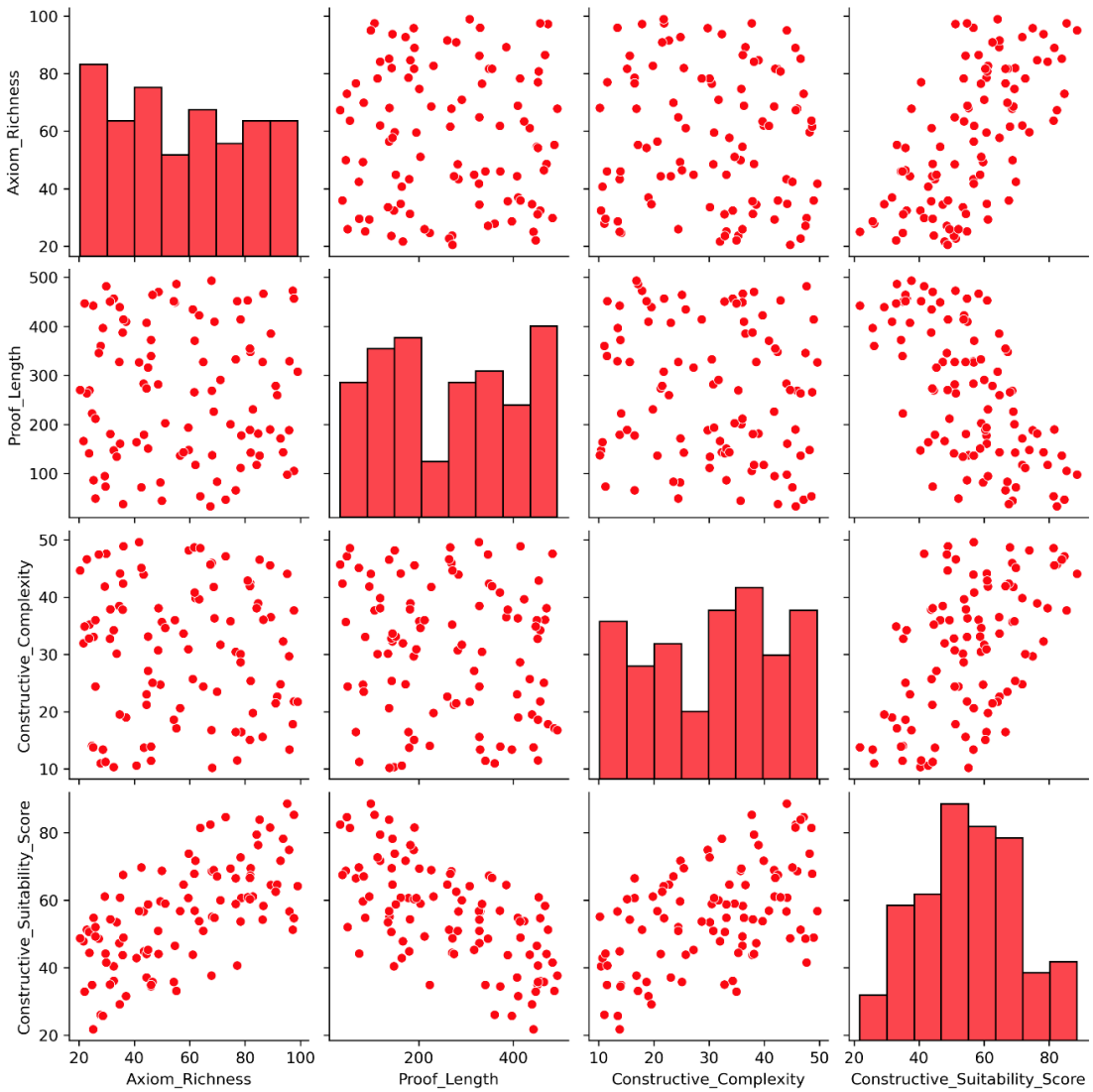


FIGURE 1. Scatter Plot of Various Constructive Mathematics

Figure 1 presents a pairwise diagram (scatterplot matrix) that visualizes the relationships between four key metrics in constructive mathematics: axiom richness, proof length, constructive complexity, and constructive relevance score. Each diagonal cell shows the distribution of a variable through a histogram, while the off-diagonal scatterplots show pairwise relationships. A strong positive correlation between constructive complexity and constructive relevance score is visually evident, suggesting that as the complexity of a construct increases, so does its relevance. In contrast, proof length does not show a clear linear relationship with the other variables, indicating that longer proofs do not necessarily correlate with richer theorems or greater complexity. Axiom richness appears to be uniformly distributed and does not show a distinct trend with any of the other metrics. The scatterplots use red dots for visual clarity, and the histograms are shaded in translucent red to highlight density.

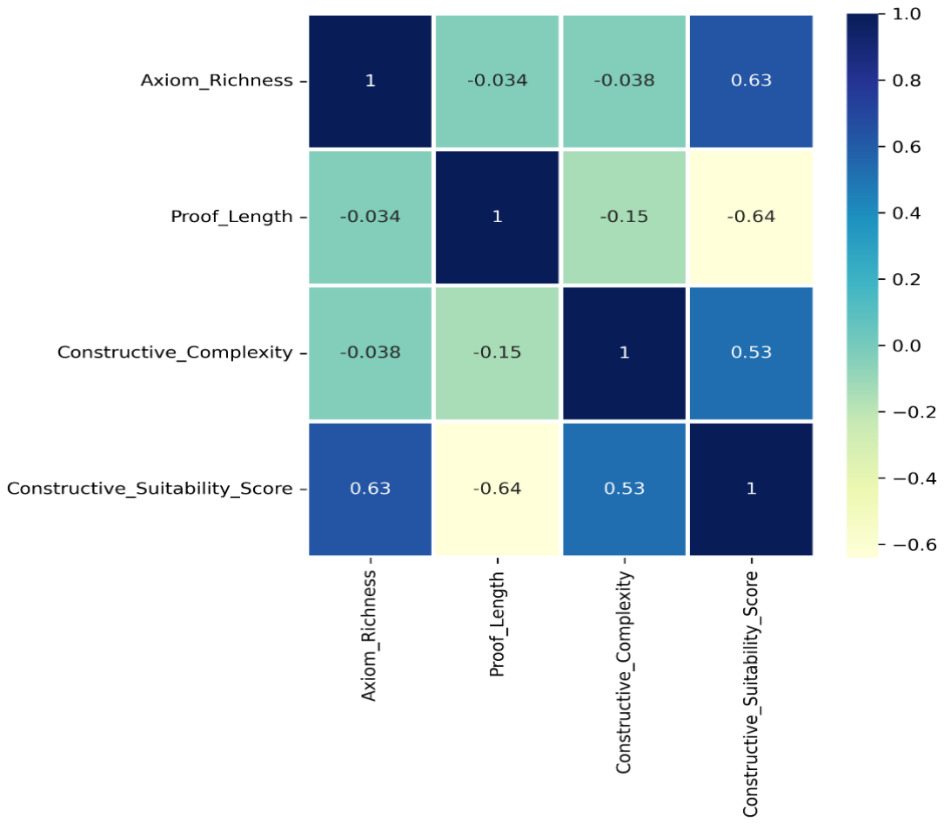


FIGURE 2. Heat Map of the Relationship Between Process Parameters and Responses

Figure 2 shows a correlation heat map measuring the relationships between four key variables in constructive mathematics: axiom richness, proof length, constructive complexity, and constructive complexity score. Each cell represents the Pearson correlation coefficient between a pair of variables, with values ranging from -1 (perfect negative complexity) to +1 (perfect positive complexity). The color gradient—from dark blue (strongly positive) to yellow-green (strongly negative) helps with visual interpretation. There is a significant positive complexity (0.63) between axiom richness and constructive complexity score, indicating that rich ampleximic systems tend to provide greater relevance in constructive systems. Constructive complexity is also associated with a moderate positive complexity (0.53) in the matching score, indicating that more complex constructive systems are generally more relevant. Interestingly, proof length shows a moderate negative correlation with both constructive complexity score (-0.64) and constructive complexity (-0.15), indicating that it may reduce both long complexity and complex complexity. The negligible correlations between axiom richness and proof length (-0.034) or constructive complexity (-0.038) indicate weak or non-linear relationships.

Copyright@ REST Publisher

Random Forest Regression

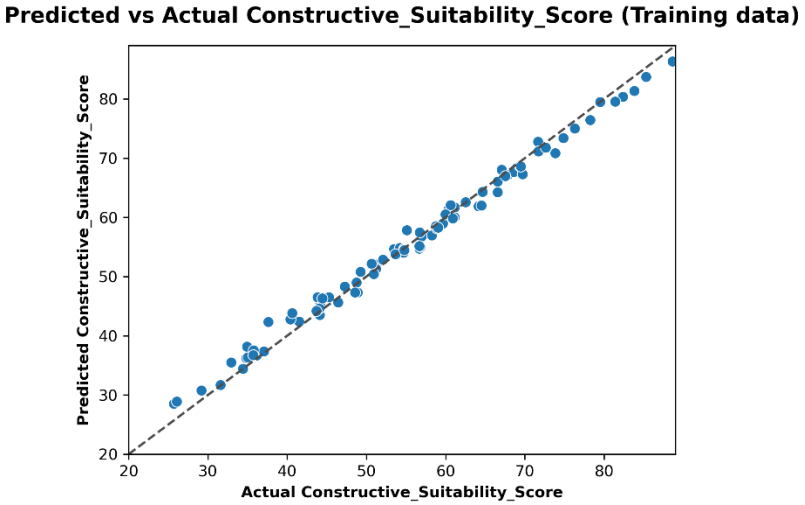


FIGURE 3. Random Forest Regression Constructive Suitability Score (Training data)

The figure 3 illustrates the relationship between the actual and predicted structural fit scores generated by the random forest regression model for the training dataset. The scatter points are closely aligned on the diagonal reference line, indicating perfect prediction accuracy. This strong alignment indicates that the model has effectively learned the underlying patterns in the training data, producing predictions that closely match the actual values. The tight clustering of points around the line further indicates low variance and minimal prediction error. In addition, the absence of significant deviations or outliers indicates that the model generalizes well within the training sample and captures nonlinear relationships between input features. The plot demonstrates that the random forest regression model provides reliable and highly accurate predictions during training, highlighting its robustness and suitability for estimating the structural fit score based on given numerical parameters.

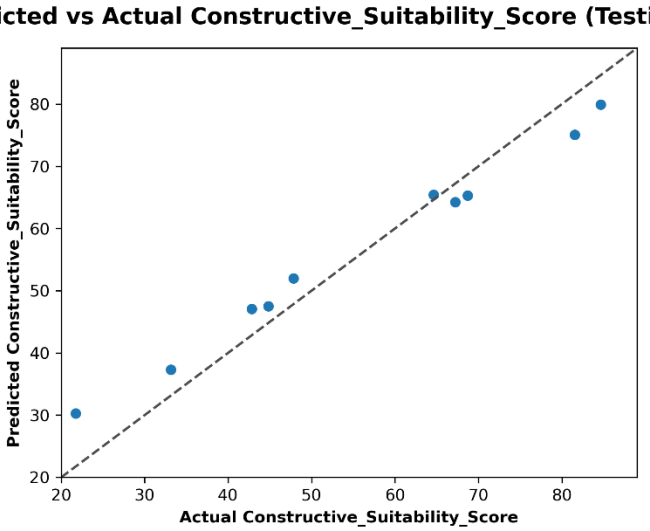


FIGURE 4. Random Forest Regression Constructive Suitability Score (Testing data)

The figure 4 presents a comparison between the actual and predicted constructive fit scores for the test dataset using the Random Forest regression model. Unlike the training data plot, the test data points are sparsely and widely

scattered, indicating moderate predictive performance. Although many points are close to the diagonal line indicating correct prediction some deviations are visible, indicating slight underestimation or overestimation by the model for specific cases. This variation reflects the model's ability to generalize to unseen data, while also revealing areas where prediction accuracy decreases due to limited sample heterogeneity or unobserved patterns in the test dataset. The Random Forest model maintains a reasonably strong relationship between the actual and predicted scores, but shows that some prediction errors are common when applied to new data. The visual trend still indicates a positive correlation, indicating that the model is effective but not flawless in predicting constructive fit scores.

TABLE 2. Random Forest Regression Constructive Suitability Score (Training, Testing data)

Data	Symbol	R2	EVS	MSE	RMSE	MAE	MaxError	MSLE	MedAE
Train	RFR	0.98869	0.98874	2.32078	1.52341	1.22639	4.71599	0.00109	1.03162
Test	RFR	0.94413	0.94541	21.76782	4.66560	4.21083	8.54299	0.01478	4.14381

Table 2 presents the performance metrics of the random forest regression (RFR) model for predicting the constructive fit score using both the training and testing datasets. The training results show exceptionally strong performance, with an R^2 value of 0.98869 and an explained variance score (EVS) of 0.98874, indicating that this model captures almost all of the variation in the training data. The low MSE (2.32), RMSE (1.52), and MAE (1.23) further confirm the high accuracy and minimal error levels of the model during training. The small MSLE and median absolute error also reflect stable predictions without significant deviation. For the testing data, an R^2 value of 0.94413 and an EVS of 0.94541 show that the model generalizes well to unobserved data, albeit with a slightly reduced accuracy compared to the training set. The high error values – MSE (21.76), RMSE (4.66), and MAE (4.21) – indicate moderate prediction deviations, which are expected due to the limited experimental samples.

4. CONCLUSION

This research successfully applied machine learning techniques to analyze and predict the constructive fit of mathematical structures, providing quantitative insights into the relationships between fundamental constructive parameters. The study examined 100 cases across four important dimensions—axiom richness, proof length, constructive complexity, and constructive fit score and revealed important patterns in how these factors interact within constructive mathematical systems. The correlation analysis found meaningful relationships that advance our understanding of constructive mathematics. The strong positive correlation between theoretical richness and constructive fit score demonstrates that systems built on more extensive axiom richness foundations show better alignment with constructive principles. Conversely, the negative correlation between proof length and fit suggests that performance is valued in derivation in constructive approaches, where excessively long proofs may indicate redundancy or structural inefficiency rather than completeness. The random forest regression model proved to be very useful for this application, achieving exceptional accuracy on training data and maintaining strong generalization to unseen test data. The robust performance of this model confirms the feasibility of using machine learning to estimate abstract mathematical properties, opening up new avenues for computational analysis of fundamental mathematical systems. The small performance degradation on test data is typical and reflects the inherent challenge of generalizing to novel cases with limited sample heterogeneity. These findings have important implications for both theoretical and applied constructive mathematics. Optimal constructive systems balance print expressiveness with resource efficiency, avoiding unnecessary complexity, while maintaining sufficient richness to address a variety of mathematical structures. Future research could extend this analysis to larger datasets, incorporate additional parameters such as algorithm performance or computational resource requirements, and explore other machine learning architectures. The methodological, quantitative, data-driven approaches established here demonstrate that they can meaningfully complement traditional philosophical and logical analyses in understanding the foundations of constructive mathematics, guide the development of more effective constructive frameworks, and support educational strategies in teaching constructive reasoning.

Copyright@ REST Publisher

REFERENCES

[1]. Martin-Lof, Per. "Constructive mathematics." Lecture at the Aarhus University. –1966 (1966).
[2]. Maietti, Maria Emilia, and Giovanni Sambin. "Toward a minimalist foundation for constructive mathematics." From Sets and Types to Topology and Analysis: Practicable Foundations for Constructive Mathematics 48 (2005): 91-114.
[3]. Crosilla, Laura. "Predicativity and constructive mathematics." In Objects, structures, and logics: FilMat studies in the philosophy of mathematics, pp. 287-309. Cham: Springer International Publishing, 2021.
[4]. Maietti, Maria Emilia. "A minimalist two-level foundation for constructive mathematics." Annals of pure and applied logic 160, no. 3 (2009): 319-354.
[5]. Ishihara, Hajime. "Reverse mathematics in Bishop's constructive mathematics." Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences CS 6 (2006): 43-59.
[6]. Bauer, Andrej. "Five stages of accepting constructive mathematics." Bulletin of the American Mathematical Society 54, no. 3 (2017): 481-498.
[7]. Nakahara, Tadao, and Masataka Koyama. "Study of the constructive approach in mathematics education." PUB TYPE EDRS PRICE (1998): 85.
[8]. Maietti, Maria Emilia, and Giuseppe Rosolini. "Quotient completion for the foundation of constructive mathematics." Logica Universalis 7 (2013): 371-402.
[9]. Bridges, Douglas, and Steeve Reeves. "Constructive mathematics in theory and programming practice." Philosophia Mathematica 7, no. 1 (1999): 65-104.
[10]. Lietz, Peter. "From constructive mathematics to computable analysis via the realizability interpretation." PhD diss., Technische Universität Darmstadt, 2004.
[11]. Shulman, Michael. "Affine logic for constructive mathematics." Bulletin of Symbolic Logic 28, no. 3 (2022): 327-386.
[12]. Lorenzen, Paul. "Constructive mathematics as a philosophical problem." Compositio Mathematica 20 (1968): 133-142.
[13]. Coquand, Thierry. "Recursive functions and constructive mathematics." Constructivity and computability in historical and philosophical perspective (2014): 159-167.
[14]. Waaldijk, Frank. "On the foundations of constructive mathematics." preprint, April (2001).
[15]. Miao, Keming. "PROOF THAT THERE IS NO ALGORITHM TO THE OPTIMAL WORK ASSIGNMENT IN CONSTRUCTIVE MATHEMATICS." European Journal of Economics and Management Sciences (2022): 16.
[16]. Tyurin, A. S., and P. V. Saraev. "Construction of quantile regression using natural gradient descent." Applied Mathematics and Control Sciences 2 (2023): 43-52.
[17]. Natekin, Alexey, and Alois Knoll. "Gradient boosting machines, a tutorial." Frontiers in neurorobotics 7 (2013): 21.
[18]. Aydin, Fatma Nur, Kubra Atalay Kabasakal, and Ismail Dilek. "Investigating the Predictive Performance of Process Data and Result Data in Complex Problem Solving Using the Conditional Gradient Boosting Algorithm." Journal of Intelligence 13, no. 3 (2025): 29.
[19]. Kaloop, Mosbeh R., Deepak Kumar, Pijush Samui, Jong Wan Hu, and Dongwook Kim. "Compressive strength prediction of high-performance concrete using gradient tree boosting machine." Construction and Building Materials 264 (2020): 120198.
[20]. Huang, Yifei, Yuhua Liu, Chenhui Li, and Changbo Wang. "GBRTVis: online analysis of gradient boosting regression tree." Journal of Visualization 22, no. 1 (2019): 125-140.

Copyright@ REST Publisher