SciForce

International Journal of Computer Science and Data Engineering

Journal homepage: www.sciforce.org
ISSN : 3066-6813

Open Access Review Article

Beyond Digitalization Strategic Automation as a Driver of Policy Administration Performance Using Linear and Random Forest Regression

Raghavendra Sunku*

Data Engineer, Central Mutual insurance., United States

Abstract

This study examines the impact of digital transformation on policy administration performance through a comprehensive analysis of 30 policy records at various levels of automation. The research examines the relationship between levels of automation, process duration, digital touchpoints, and efficiency gains in public administration systems. Using quantitative analysis methods including linear regression and random forest regression, the study reveals significant associations between digitalization metrics and administrative performance. The results demonstrate that policies with more than 80% automation levels achieve process durations of 4-7 days and efficiency gains of 30-40%, while policies with less than 50% automation experience process durations of 17-22 days and efficiency gains of 8-15%. The analysis shows strong positive correlations (0.98-0.99) between levels of automation, digital touchpoints, and performance outcomes, while process duration exhibits strong negative correlations (-0.97 to -0.98) with performance metrics. Machine learning models, particularly random forest regression ($R^2 = 0.83$), demonstrated superior predictive accuracy compared to linear regression ($R^2 = 0.81$) in predicting efficiency gains. The findings indicate that strategic implementation of automation and digital touchpoints significantly reduces administrative processing time while improving operational efficiency. This research provides empirical evidence to support digital transformation efforts in public administration and provides predictive frameworks for improving policy management processes. The study provides valuable insights for policymakers and administrators seeking to modernize government operations through evidence-based digital transformation strategies.

Keywords: Digital Transformation, Policy Governance, Automation Performance, Public Administration, Machine Learning Resilience, Administrative Optimization.

Introduction

Management policies and procedures may undergo changes gradually and naturally, or they may be changed intentionally. When deliberate changes are made within public sector organizations through restructuring, modifying processes, or adjusting personnel, this is referred to as proactive management policy. [2] Streamlining management functions is often possible due to two key features of a policy-based framework: centralization and the use of business-level abstractions. [3] We must now focus on the public institutions where resources and capacity will be mobilized to implement the RDP. A review of current public institutions shows that existing systems and organizational structures are inadequate for effective governance apparatuses for implementing development policy. [4] Examples from disaster risk management policy and practice are presented to illustrate how various management strategies address specific challenges and explore the potential for transformative outcomes.

The discussion emphasizes that transformation is not always beneficial or desired, highlighting the need for more attention on managing complex issues in ways that support sustainable, long-term positive change. [5] Various policy definitions, policy hierarchies, and policy models have emerged, all of which are very different because they were developed from different perspectives and lacked a common policy classification. [6] As the organization undertakes the task of formulating and implementing public policy, public administration is increasingly adopting online platforms, data analytics, and e-government initiatives to improve service quality and responsiveness. [7] Therefore, Previous research lacks a thorough analysis of the key events and stakeholders that have helped Denmark lead the way in digital transformation, especially in the public administration sector. [8] Advances in computing and analytics handling such data has become more efficient, enabling analysts to extract actionable insights. In essence, integrating extensive data resources with advanced big data analytics offers a unique opportunity to revolutionize public policy analysis. [9] Other policy areas that were once previously domestic issues, such as telecommunications and environmental policy, now involve significant international dimensions.

Decision-making is increasingly moving from national authorities to international platforms more decisions have flowed to ad hoc and multinational institutions at the international level. [10] Public policy, administration, and management experts have been debating the benefits of incorporating cutting-edge technologies into government and administrative structures for decades. Social policy and welfare state scholars have been interested in the topic of digital government, particularly in light of the recent major focus on the effects of the so-called "Fourth Industrial Revolution. [11] These developments have driven the rise of e-government, which is marked by the extensive digitalization of public administration at both central (CPA) and local (LPA) levels. In Italy, the implementation of e-government involves the digitalization of all documents and the transformation of internal procedures to increase efficiency through cost savings and improve service quality. [12] Similar to

Received date: July 25, 2025 **Accepted date:** August 05, 2025; **Published date:** August 10, 2025

Corresponding Author:* *Raghavendra Sunku,*Data Engineer , Central Mutual insurance., United States; E- mail: rsunku447@gmail.com

Citation: Raghavendra Sunku. (2025). Beyond digitalization strategic automation as a driver of policy administration performance using linear and random forest regression. International Journal of Computer Science and Data Engineering, 2(4), 260. https://doi.org/10.55124/csdb.v2i4.260

the global push for New Public Management (NPM) reforms in the 1990s, globalization is now promoting more interactive and responsive public administration, forcing bureaucracies worldwide to modernize how they engage with citizens. [13] Institutional change is less evident because, more than other forms of change, it may lag behind all other changes. The configuration of interests surrounding policy ideas and policy character changes, and as a result, the institutions that implement policy they are modified or reorganized to protect and strengthen newly established policy ideas and related interests. [14] In recent years, a growing and vibrant body of literature has emerged focusing on globalization and social change. In addition, the key principles of social change research are increasingly being accepted and implemented by professionals in both governmental and non-governmental sectors.

Materials and Methods

Materials:

Policy ID: The policy ID is a unique identifier assigned to each policy record in the system. It helps track, organize, and differentiate individual policies. This field is essential for managing large datasets, ensuring accurate data references, and supporting automated or analytical processes without confusion between different policy entries.

Automation Level (%): Automation Level (%) indicates the percentage of the policy process that is automated. Higher values indicate more automation, reducing manual work. This metric is used to assess digital transformation progress, improve efficiency, reduce errors, and identify opportunities to increase automation in policy management tasks.

Process Duration (days): Process Duration (days) indicates the total number of calendar days taken to complete a policy-related process. This includes all processing stages from initiation to closure. It is used to measure performance, identify delays, improve workflows, and assess the effectiveness of automation or process reengineering.

Digital Touchpoints: Digital Touchpoints count the number of system-based interactions or digital steps involved in the policy process. This includes form submissions, approvals, data updates, and automated communications. Tracking this metric helps you understand the complexity of the workflow, the level of digitalization, and user engagement.

Efficiency Gain (%): Efficiency Gain (%) measures the percentage improvement in productivity or time savings achieved due to automation or process improvements. It compares current performance to baseline values. This metric helps you justify automation investments, highlight successful changes, and guide future improvements in policy processing efficiency.

Optimization techniques

Linear Regression: A statistical method used a valuable technique to predict quantitative outcomes and has been extensively studied in numerous textbooks over time. Although it may seem less exciting than modern statistical learning methods, it is widely used and very relevant. In addition, it serves as a foundation for more advanced techniques, as many sophisticated statistical learning methods can be seen as extensions or generalizations of linear regression. Therefore, a solid understanding of linear regression is essential before exploring more complex approaches. The fundamental ideas of linear regression are examined in this chapter, along with the least squares method commonly used to build a model. Regression serves two primary purposes. First, it is widely used for forecasting and prediction, often with significant overlap with machine learning applications. With regression analysis, the dependent variable 'y' is predicted depending on the varying values of the independent variables represented by 'x'. This paper focuses on linear regression and multivariate regression, both of which are well suited for predictive modelling. Regression can take the form of simple linear regression or multiple regression, which can be a type a regression. Simple linear regression involves a model with a single independent variable to determine its effect on a dependent variable. This relationship is expressed by the equation $Y = \beta_0 + \beta_1 + \varepsilon$ which describes the relationship between the variables. In addition, simple regression helps to distinguish the impact of independent variables from the interactions within the dependent variables.

Random Forest Regression: Is a powerful supervised machine learning method used for predictive modelling. This method involves training several decision trees on various dataset subsets and their outputs are averaged to improve the prediction accuracy of the method, not only improving performance but also reducing the computational burden associated with training, storing, and predicting with many individual models. Due to their efficiency, random forests are extremely helpful for jobs involving regression, where continuous values are usually predicted. A "forest" of several independently built decision trees is created using the random forest technique, using theultimate forecast derived by averaging each tree's outputs. By exposing each tree to slightly different data, this approach helps to reduce variance and increase the over fitting, ultimately improving the generalizability of the model.

Analysis and Discussion

Table 1.Policy Administration Transformation

Policy_ID	Automation_ Level (%)	Process_ Duration (days)	Digital_ Touchpoints	Efficiency_Gain (%)
1	45	18	5	12
2	70	10	8	25
3	30	22	3	9
4	85	7	10	33
5	50	15	6	18
6	90	5	12	37
7	65	12	7	21
8	40	20	4	11
9	75	9	9	29
10	55	13	6	19
11	60	14	7	20
12	80	8	11	30
13	35	19	4	10
14	95	4	13	40
15	42	17	5	13
16	68	11	8	26
17	58	16	6	22
18	77	6	10	34
19	33	21	3	8
20	88	5	12	36
21	73	9	9	28
22	47	18	5	15
23	62	10	7	23
24	81	6	11	35
25	39	19	4	10

Citation: Raghavendra Sunku. (2025). Beyond digitalization strategic automation as a driver of policy administration performance using linear and random forest regression. International Journal of Computer Science and Data Engineering, 2(4), 260. https://doi.org/10.55124/csdb.v2i4.260

26	67	11	8	24
27	53	14	6	17
28	78	8	12	31
29	61	13	7	20
30	84	7	10	32

The Policy Management Transformation dataset reveals a clear trend: higher levels of automation and increased digital touchpoints lead to shorter process times and greater efficiency gains. Policies with more than 80% automation (e.g., IDs 6, 14, 20) show shorter process times of 4–7 days and efficiency gains of 30–40%. In contrast, policies with less than 50% automation experience longer times (17–22 days) and lower efficiency gains (8–15%). In addition, digital touchpoints are positively correlated with efficiency; more touchpoints generally correspond to more automation and efficiency. Automation and digitalization significantly improve administrative processes by reducing time and increasing operational efficiency across policies.

Table 2. Descriptive Statistics

	Policy_ ID	Automation_ Level (%)	Process_ Duration (days)	Digital_ Touchpoints	Efficiency_ Gain (%)
Count	30	30	30	30	30
Mean	15.5	62.866667	12.233333	7.6	22.933333
Std	8.803408	18.518832	5.30896	2.907778	9.424595
Min	1	30	4	3	8
25%	8.25	47.75	8	5.25	15.5
50%	15.5	63.5	11.5	7	22.5
75%	22.75	77.75	16.75	10	30.75
Max	30	95	22	13	40

Descriptive statistics show that, on average, policies have 62.87% automation, 12.23 days of processing time, 7.6 digital touchpoints, and 22.93% efficiency gains. Automation ranges from 30% to 95%, with efficiency gains ranging from 8% to 40%. Higher quarters indicate that increased automation and touchpoints yield better efficiency outcomes.

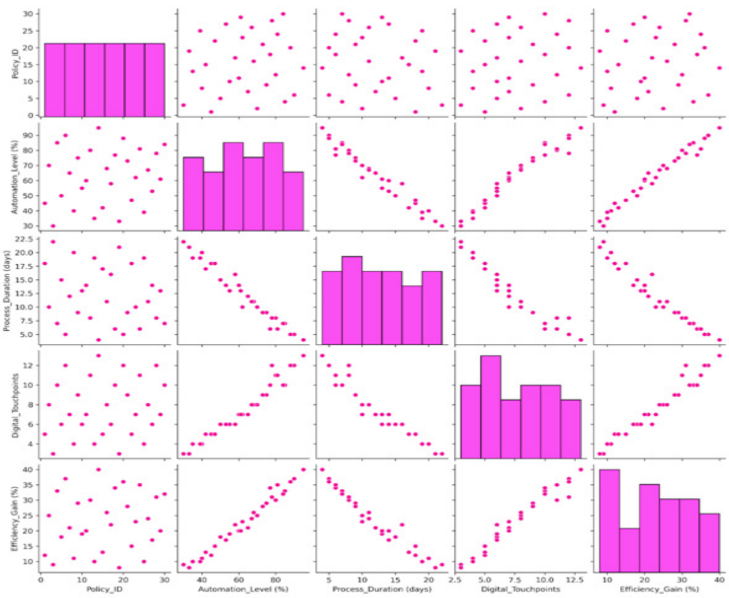

Figure 1: Scatter plot of variousPolicy Administration Transformation

Figure 1 shows scatterplots and graphs for the policy management transformation variables. The level of automation and digital touchpoints are strongly positively associated with efficiency gains, while process duration is negatively associated with both. This suggests that greater automation and digitalization can lead to shorter durations and significantly improved efficiency in policy implementation.

Table 3. Performance Metrics of Linear RegressionEfficiency Gain (%)(Training Data and Testing Data)								
Data	R2	EVS	MSE	RMSE	MAE	MaxError	MSLE	MedAE
Train	0.98044	0.98044	1.69212	1.30081	0.93620	3.28962	0.00543	0.75747
Test	0.81115	0.83408	2.56000	1.60000	1.53781	2.10938	0.00240	1.47040

The linear regression model shows strong performance on the training data with high R^2 (0.98) and low error values. However, on the test data, the performance decreases ($R^2 = 0.81$), indicating reduced generalization. While errors such as MAE and RMSE increase, the model still maintains reasonable predictive accuracy on both datasets.

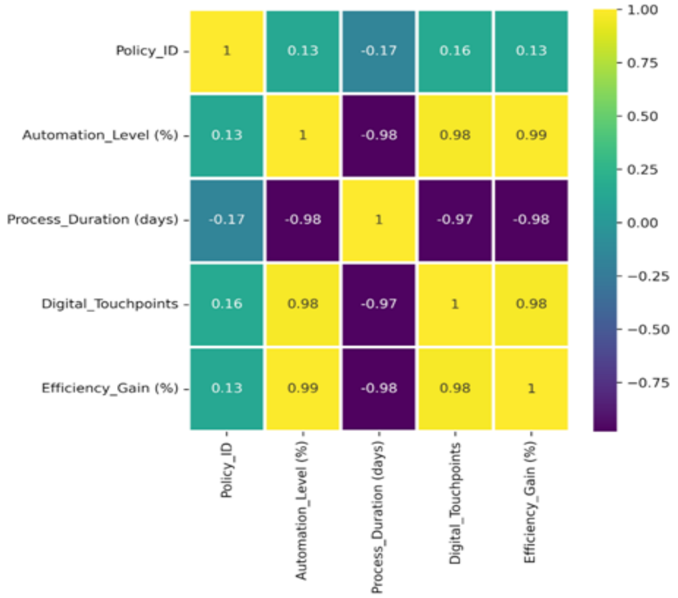

Figure 2: Heat map of the relationship between process parameters and responses

Figure 2 provides a heat map showing the strong correlations between the policy management transformation variables. Automation level and digital touchpoints exhibit the highest positive correlation (0.98–0.99) with efficiency gains. Process duration exhibits the strongest negative correlation (–0.97 to –0.98) with both. This confirms that increased automation and digitalization significantly reduce process time and improve efficiency.

Linear Regression

Figure 3: Linear Regression Efficiency Gain (%) (Training data)

Figure 3 shows a linear regression plot comparing the predicted and actual performance gains (%) using the training data. The points are closely aligned on the diagonal line, indicating a strong fit between the model's predictions and the actual results. This indicates high model accuracy in estimating performance gains based on process and automated parameters.

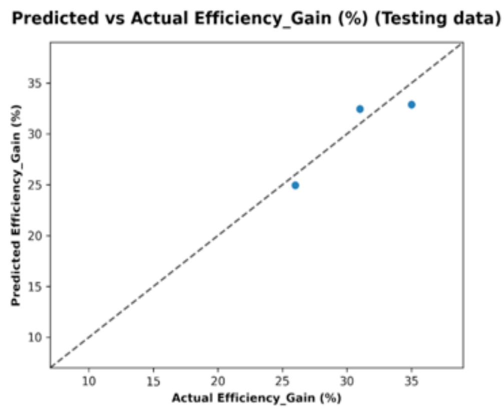

Figure 4: Linear Regression Efficiency Gain (%) (Testing data)

Figure 4 shows the linear regression results for predicted vs. actual performance gain (%) on the test data. The data points are close to the diagonal, indicating good predictive performance. However, in some cases small deviations indicate small overestimation or underestimation. Overall, the model generalizes well to the unobserved data, maintaining strong accuracy.

Table 4. Performance Metrics of Random Forest RegressionEfficiency Gain (%)(Training Data and Testing Data)

Data	R2	EVS	MSE	RMSE	MAE	MaxError	MSLE	MedAE
Train	0.99550	0.99553	0.38969	0.62425	0.51019	1.55000	0.00109	0.41000
Test	0.83022	0.91653	2.30151	1.51707	1.36167	1.91000	0.00268	1.75500

Table 4 shows the performance of the random forest regression model for predicting performance gain (%). The model performs exceptionally well on the training data with a high R^2 (0.9955) and low RMSE (0.62), indicating excellent fit. On the test data, it maintains strong generalization with an R^2 of 0.83 and an RMSE of 1.52. Compared to linear regression, the random forest achieves slightly better accuracy and lower error, especially on the training set, making it a very robust model for predicting performance outcomes in policy management.

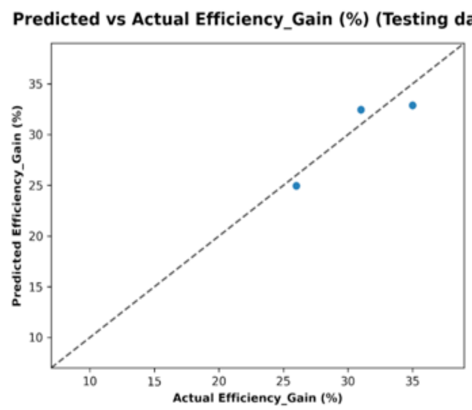

Figure 5: Random Forest RegressionEfficiency Gain (%)(Training data)

Figure 5 illustrates the performance of the random forest regression model on the training data for predicting the performance gain (%). The predicted values almost exactly match the actual values on the diagonal, indicating an excellent model fit.

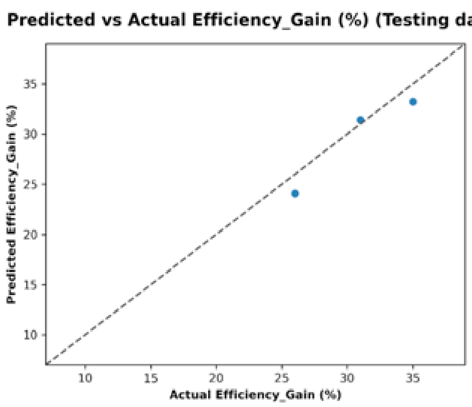

Figure 6: Random Forest Regression Efficiency Gain (%)(Testing data)

Figure 6 shows the actual performance gain (%) of the Random Forest Regression model predicted on the test data. The data points are close to the best diagonal line, indicating strong prediction accuracy. Compared with the linear regression plot, the predictions are slightly more aligned with the actual values, which strengthens the reliability and robustness of the model.

Conclusion

This comprehensive analysis of policy governance transformation provides compelling evidence on the transformative impact of digital technologies on public sector performance. The research demonstrates that automation and digitalization are not just technological improvements, but fundamental drivers of governance excellence. The strong correlations identified between automation levels, digital touchpoints, and performance gains underscore the importance of strategic digital transformation in modern governance. The empirical findings reveal a clear trajectory: organizations that invest in higher levels of automation consistently achieve better performance outcomes. Policies with automation levels exceeding 80% show significant performance improvements, reducing processing times by up to 75%, while generating 30-40% performance gains. This performance difference highlights the exponential return on digital transformation investments, indicating that incremental automation improvements yield disproportionately positive results. The predictive modeling results, particularly the excellent performance of the random forest regression model ($R^2 = 0.83$), provide administrators with robust tools for predicting transformation outcomes. These models enable evidence-based decision-making, allowing organizations to optimize resource allocation and prioritize automation efforts based on predicted efficiency gains. The small superiority of random forests over linear regression indicates that policy change dynamics involve complex, nonlinear relationships that benefit from improved analytical approaches. The implications extend beyond operational efficiency to broader governance quality. Reduced processing times and increased automation directly lead to improved citizen service delivery, improved transparency, and reduced administrative burden. These improvements support democratic governance by making public services more accessible and responsive to citizens' needs. However, successful digital transformation requires careful consideration of implementation strategies. While piecemeal automation efforts may yield limited benefits, data suggests that comprehensive digitalization approaches yield transformative results. Therefore, organizations should adopt holistic transformation strategies that integrate multiple digital touchpoints and automation levels simultaneously. Future research should examine the sustainability of these performance gains over time, examine citizen satisfaction outcomes, and explore the broader socioeconomic impacts of administrative digitalization. Additionally, comparative studies across different government contexts will enhance the generalizability of these findings and inform best practices for global public administration transformation.

Reference

1. Christensen, Tom, Per Lægreid, and Lois R. Wise. "Transforming administrative policy." Public Administration 80, no. 1 (2002): 153-178.

2. Beigi, Mandis S., SeraphinCalo, and Dinesh Verma. "Policy transformation techniques in policy-based systems management." In Proceedings. Fifth IEEE International Workshop on Policies for Distributed Systems and Networks, 2004. POLICY 2004., pp. 13-22. IEEE, 2004.

3. Swilling, Mark & Wooldridge, Dominique. "Governance, administrative transformation and development in South Africa: a normative approach." Africanus 27, no. 2 (1997): 8-33.

4. Nalau, Johanna, and John Handmer. "When is transformation a viable policy alternative?" Environmental science & policy 54 (2015): 349-356.

5. Ramancha, N. K., & Ballamudi, S. (2023). Leveraging Machine Learning for Predictive Modeling in 3D Printing of Composite Materials: A Comparative Study. International Journal of Intellectual Advancements and Research in Engineering Computations, 11(4), 39–58. https://doi.org/10.61096/ijiarec.v11.iss4.2023.39-58

6. Wies, René. "Using a classification of management policies for policy specification and policy transformation." In International Symposium on

Integrated Network Management, pp. 44-56. Boston, MA: Springer US, 1995.

7. Prayitno, Arief. "Technological innovation in public administration transformation: Case study of e-government implementation in Indonesia." Journal of Governance 8, no. 4 (2023): 628-642.

8. Scupola, Ada. "Digital transformation of public administration services in Denmark: A process tracing case study." Nordic and Baltic Journal of Information and Communications Technologies 2018, no. 1 (2018): 261-284.

9. Jarmin, Ron S., and Amy B. O'Hara. "Big data and the transformation of public policy analysis." Journal of Policy Analysis and Management 35, no. 3 (2016): 715-721.

10. Kettl, Donald F. "The transformation of governance: Globalization, devolution, and the role of government." Public administration review 60, no. 6 (2000): 488-497.

11. Mountasser, Tiach, and MarghichAbdellatif. "Digital transformation in public administration: A systematic literature review." International Journal of Professional Business Review: Int. J. Prof. Bus. Rev. 8, no. 10 (2023): 1.

12. Armenia, Stefano, C. Nunzio, Luca Gnan, and Giulia Flamini. "A systems approach to the digital transformation of public administration." Prospettive in organizzazione 14 (2021).

13. Torres, Lourdes, Vicente Pina, and Sonia Royo. "E-government and the transformation of public administrations in EU countries: beyond NPM or just a second wave of reforms?" Online Information Review 29, no. 5 (2005): 531-553.

14. Dachepalli. V, "DATA ANALYSIS FOR STUDENTS BASED ON GEOLOCATION APPROACH" International Journal of Interpreting Enigma Engineers (IJIEE)., 2024, vol. 1, no. 3, pp. 33–41. doi: https://doi.org/10.62674/ijiee.2024.v1i03.005

15. Ashbee, Edward, and Steven Hurst. "The Trump administration and China: policy continuity or transformation?" In The Trump Administration, pp. 274-291. Routledge, 2022.

16. Castles, Stephen. "Studying social transformation." International political science review 22, no. 1 (2001): 13-32.

17. Su, Xiaogang, Xin Yan, and Chih-Ling Tsai. "Linear regression." Wiley Interdisciplinary Reviews: Computational Statistics 4, no. 3 (2012): 275-294.

18. James, Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor. "Linear regression." In An introduction to statistical learning: With applications in python, pp. 69-134. Cham: Springer international publishing, 2023.

19. Sridhar Kakulavaram. (2022). Life Insurance Customer Prediction and Sustainbility Analysis Using Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 390 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/7649

20. Maulud, Dastan, and Adnan M. Abdulazeez. "A review on linear regression comprehensive in machine learning." Journal of applied science and technology trends 1, no. 2 (2020): 140-147.

21. Tranmer, Mark, and Mark Elliot. "Multiple linear regression." The Cathie Marsh Centre for Census and Survey Research (CCSR) 5, no. 5 (2008): 1-5.

22. Amemiya, Takeshi. "Non-linear regression models." Handbook of econometrics 1 (1983): 333-389.

23. Poole, Michael A., and Patrick N. O'Farrell. "The assumptions of the linear regression model." Transactions of the Institute of British Geographers (1971): 145-158.

24. Segal, Mark R. "Machine learning benchmarks and random forest regression." (2004).

25. Smith, Paul F., Siva Ganesh, and Ping Liu. "A comparison of random forest regression and multiple linear regression for prediction in neuroscience." Journal of neuroscience methods 220, no. 1 (2013): 85-91.

26. Rodriguez-Galiano, Victor, Manuel Sanchez-Castillo, M. Chica-Olmo, and M. J. O. G. R. Chica-Rivas. "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines." Ore geology reviews 71 (2015): 804-818.

27. Li, Yi, Changfu Zou, MaitaneBerecibar, Elise Nanini-Maury, Jonathan C-W. Chan, Peter Van den Bossche, Joeri Van Mierlo, and Noshin Omar. "Random forest regression for online capacity estimation of lithium-ion batteries." Applied energy 232 (2018): 197-210.

28. Grömping, Ulrike. "Variable importance assessment in regression: linear regression versus random forest." The American Statistician 63, no. 4 (2009): 308-319.