

# Predictive Analytics for SSO Performance: Improving Authentication Response Times in Oracle Enterprise Environments Using Linear Regression, Random Forest Regression

Tirumala Rao Gundala\*

Consulting Technical Manager & Performance Architect, Oracle, United States

## Abstract

This research examines the implementation and optimization of single sign-on (SSO) systems within Oracle Insurance Policy Management environments, focusing on key performance metrics and their correlations. The study analyzes authentication response times, daily successful logins, integration complexity scores, and SSO success rates across 200 observations to identify patterns that affect system performance. Through extensive statistical analysis, including descriptive statistics, correlation analysis, and predictive modeling using linear regression and random forest regression techniques, this research reveals significant relationships between system performance indicators. The findings demonstrate a strong negative correlation (-0.79) between authentication response time and SSO success rate, indicating that slow authentication processes significantly reduce user authentication success. In addition, integration complexity shows a moderate negative correlation (-0.43) with SSO success rates, indicating that simpler system architectures yield better authentication outcomes. The Random Forest regression model achieved better predictive performance on the test data with an  $R^2$  of 0.805 compared to the  $R^2$  of 0.719 for linear regression, demonstrating the value of ensemble methods in complex authentication system analysis. Authentication response times varied significantly from 50.21ms to 1990.87ms, while SSO success rates ranged from 38.36% to 92.60%, highlighting the considerable variation in system performance across different implementations. This research provides practical insights for optimizing SSO applications in enterprise environments, particularly emphasizing the importance of reducing response times and reducing integration complexity to achieve reliable authentication outcomes. These findings provide valuable guidance for organizations looking to improve user experience and system reliability in large-scale Oracle-based insurance policy management systems.

**Key Words:** Single Sign-On (SSO), Authentication Performance, Oracle E-Business Suite, Integration Complexity, Random Forest Resilience, Insurance Policy Management.

## Introduction

These issues are addressed by single sign-on (SSO) technology, which enables users to authenticate only once and access multiple systems without having to log in repeatedly. In addition to streamlining access control, SSO also greatly reduces IT costs and password-related risks. However, there are significant scalability, interoperability, security, and compliance issues when adopting SSO in large-scale environments, especially when using Oracle E-Business Suite (EBS) R12.2. [2] The need for shared workstation support in healthcare settings and the high help desk costs associated with passwords are key drivers for ESSO adoption. However, whether explicitly stated or not, improved user experience is usually a more fundamental need. Despite attempts to address this issue with other reduced login methods and strategies, organizations are deploying ESSO systems when users are faced with an unmanageable volume of user IDs and passwords for at least the next two years. [3] This framework includes an application programming interface (API) for building single sign-on implementations

and a service provider interface (SPI) for building authentication plug-in modules that interface with existing authentication standards. It includes two implementations of the single sign-on SPI. Both implementations are instances of the password vault, an encrypted storage for stored usernames and passwords. [4] A popular authentication technique called single sign-on (SSO) allows users to access multiple web services with a single set of login credentials instead of multiple logins.

In a standard SSO process, a user authenticates with a previously registered identity provider in an attempt to access a trusted party (RP) service portal. [5] There are still many high-level insecure SSO implementations due to the complexity of the chosen protocol. While security testing for real-world SSO applications has recently received a lot of attention, most of the work done so far has focused on websites or relied on manual detection of specific, well-known vulnerabilities. [6] The M2X economy thus leverages ideas such as wireless sensor networks, the Internet of Things, and cyber-physical systems. For complex trust-establishment requirements involving multiple systems, devices, organizations, and human collaboration, single sign-on IA is not sufficient. [7] One of the main use cases of SAML is to facilitate authentication through single sign-on (SSO) functionality. Using a single set of credentials from a central authority or identity provider, single sign-on enables a user to authenticate to multiple service providers. [8] A unique feature of Oracle Cloud HCM is its adaptability. Unlike legacy systems that are constrained by rigid structures and manual processes, Oracle Cloud is designed to address the changing needs of global organizations. Its cloud-native architecture ensures seamless integration with enterprise systems, real-time data updates, and compliance with various regulatory requirements. This makes it well-suited for multinational organizations operating across

**Received date:** September 13, 2025 **Accepted date:** September 28, 2025; **Published date:** October 06, 2025

\*Corresponding Author: Gundala, Tirumala Rao, Consulting Technical Manager & Performance Architect, Oracle, United States ; E- mail: [Tirumalagundala7@gmail.com](mailto:Tirumalagundala7@gmail.com)

**Copyright:** © 2025 Gundala, Tirumala Rao. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

diverse legal and cultural landscapes. [9] Building and implementing test oracles that accurately evaluate software behavior against expected results is critical in any complex system, especially within insurance platforms, where accurate validation of rules and workflows, such as insurance and claims processing, is essential. Here are several essential methods and methodologies for building and managing these test oracles.

Materials and Method

Materials:

**Authentication Response Time (ms):** Authentication response time, measured in milliseconds, is the amount of time it takes for an authentication system to validate user credentials and return an access approval or denial. It reflects login speed and performance. Lower response times indicate faster, smoother authentication performance, improved user experience, scalability, and overall system responsiveness.

**Successful Logins Per Day:** Successful logins per day represent the number of user authentication attempts that successfully complete within 24 hours. This metric helps measure site usability, reliability, and system availability. High values generally indicate strong user engagement, while sudden drops may indicate technical issues, access barriers, or authentication process failures.

**Integration Complexity Score:** The Integration Complexity Score measures the relative difficulty of integrating an authentication or single sign-on system with existing applications. It considers factors such as configuration steps, required coding, compatibility, and maintenance. Higher scores indicate more effort, cost, and risk during implementation, while lower scores indicate simpler, smoother integration with fewer challenges.

**SSO Success Rate:** The single sign-on (SSO) success rate measures the percentage of login attempts that are successfully completed through SSO authentication. It indicates the reliability, efficiency, and user satisfaction of centralized login processes. A higher success rate reflects stable integrations and minimal login errors, while lower rates may highlight configuration issues, system incompatibilities, or failures.

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 based on different values of the independent variables. The variable '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. It is represented by the equation  $y = \beta_0 + \beta_1x + \epsilon$  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:** A useful supervised machine learning technique is random forest regression, which is used 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 the ultimate 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.

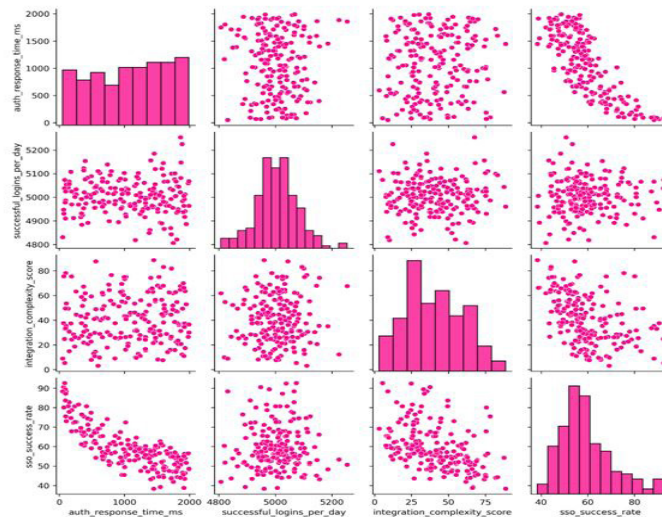
Analysis and Discussion

The dataset highlights four authentication performance metrics: response time, daily logins, integration complexity, and SSO success rate. Auth response time varies widely, from very low values near 50 ms to peaks near 2000 ms, indicating inconsistent system performance with some sessions occurring almost instantly while others suffer from latency. Despite this variability, successful logins per day remain relatively stable, typically ranging from 4900 to 5100, reflecting consistent site engagement regardless of speed differences. Integration complexity scores show significant fluctuations, ranging from 3 to nearly 90. Higher scores are often associated with lower SSO success rates, indicating that more complex integrations will reduce authentication reliability. Conversely, simpler systems often correspond to higher SSO rates above 70%, demonstrating the value of streamlined system design. SSO success rates range from less than 40% to more than 90%, revealing both operational risks and opportunities. Cases with best response times under 500 ms often achieve strong SSO outcomes, especially when the integration complexity is moderate. However, increased response times above 1500 ms are often associated with weaker success rates, reinforcing the link between speed, usability, and system reliability.

Table 1: Descriptive statistics for the four main authentication metrics across 200 observations				
	Auth response time ms	Successful Logins Per day	Integration Complexity score	Sso Success rate
Count	200	200	200	200
Mean	1077.4886	5005.07	40.586695	59.241275
Std	570.54109	73.210134	19.28505	11.100403
Min	50.21	4807	3.096	38.356
25%	564.7075	4959.75	26.1765	51.8855
50%	1123.22	5007.5	38.444	57.056
75%	1571.205	5048.25	54.88125	65.28225
Max	1990.87	5254	88.654	92.6

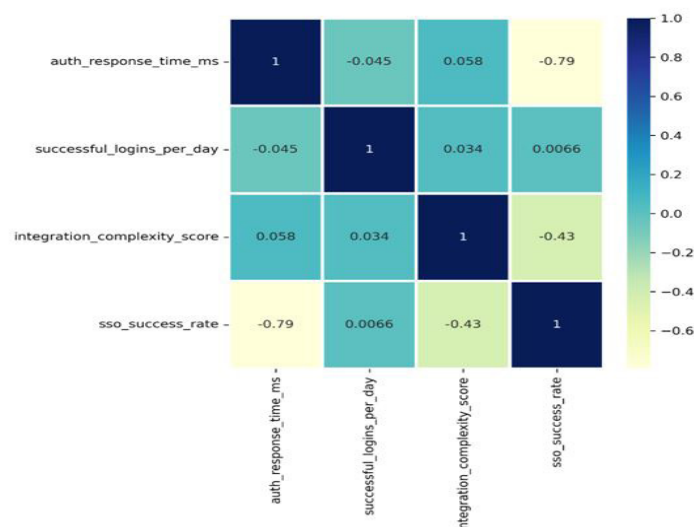
Table 1 provides descriptive statistics for the four main authentication metrics across 200 observations. The auth response time averages 1077.49 ms with a standard deviation of 570.54 ms, highlighting the considerable variability in login speed. The observed response times range from a minimum of 50.21 ms to a maximum of 1990.87 ms, indicating that while some users experience almost instant authentication, others experience significant delays. The intermediate range (564.71–1571.21 ms) also

indicates that most logins are moderately fast, but there are extreme values. The average number of successful logins per day is 5005.07, with a relatively low standard deviation of 73.21, indicating that daily site usage is stable. The values range from 4807 to 5254, reflecting consistent user engagement despite fluctuations in response time. The integration complexity score averages 40.59, with a wide spread (standard deviation 19.29), from very simple integrations at 3.10 to very complex systems at 88.65. High complexity appears less frequently, concentrated above the 75th percentile. Finally, the SSO success rate averages 59.24%, with values ranging from 38.36% to 92.60%, indicating varying reliability. The 25th–75th percentile range (51.89–65.28%) shows that most systems achieve moderate success, while exceptional or poor rates are exceptional.



**Figure 1:** Scatter plot of variousSingle Sign-On implementation for Oracle Insurance Policy Administration

Figure 1 shows pairwise scatter plots of four metrics: auth response time (ms), successful logins per day, integration complexity score, and SSO success rate. Each diagonal plot presents the distribution of a single variable through histograms, while diagonal plots show the relationships between pairs of variables. auth response time vs SSO success rate: There is a clear negative correlation; as response time increases, SSO success rate decreases. This indicates that slower authentication often leads to lower success rates. auth response time vs integration complexity: No strong linear pattern is evident, indicating that response time is not directly dependent on integration complexity. Successful logins per day: Most scatter plots involving this metric show no significant correlation, meaning that the number of user logins remains fairly constant regardless of response time, complexity, or SSO success rate. Integration complexity vs SSO success rate: There is a slight negative trend, indicating that more complex integrations may reduce SSO reliability, although the variability is high. Distributions: The histograms indicate that the authentication response time is very widespread, logins per day are around 5000, the integration complexity has a moderate spread, the SSO success rate shows a right-skewed distribution, and most systems achieve 50–70% success.



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

The heat map in Figure 2 illustrates the relationship between key process parameters and response metrics within an authentication system. Each cell represents the Pearson correlation coefficient between the two variables, ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation). The most significant relationship is between auth response time ms and sso success rate, which shows a strong negative correlation of -0.79. As authentication response time increases, the SSO success rate decreases significantly, indicating that performance issues can greatly impact user success. In contrast, successful logins per day shows almost no correlation with any other variable, indicating that it can be affected by external factors not captured here. Integration complexity score has a modest negative correlation (-0.43) with sso success rate, indicating that more complex integrations are associated with reduced SSO success. However, it has a much weaker positive correlation with the other variables.

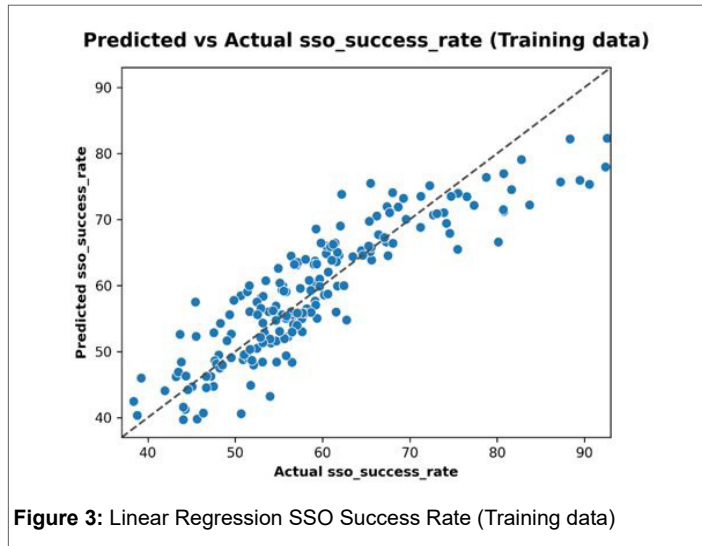


Figure 3 shows a scatterplot comparing predicted and actual SSO success rate values for the training data using a linear regression model. Each point on the plot represents an observation. The dashed diagonal line represents the best-case scenario where the predicted values exactly match the actual values (i.e., correct predictions). The cluster of points surrounding this line indicates that the linear regression model performs reasonably well in capturing the trend of the data. However, there is some significant scatter, particularly at the lower and upper limits of the actual SSO success rate, indicating that the model tends to underestimate at high success rates and overestimate at low ones. This pattern indicates regression to the mean, which is a common limitation in linear models. Despite this, the general alignment indicates that the model has captured meaningful relationships between features and the response variable. Although there is room for improvement in prediction accuracy, especially at extremes, the figure confirms the utility of the model in predicting SSO success rates. This insight could prompt further refinement of the model, perhaps by incorporating nonlinear techniques or additional features to better capture the complexity of the system.

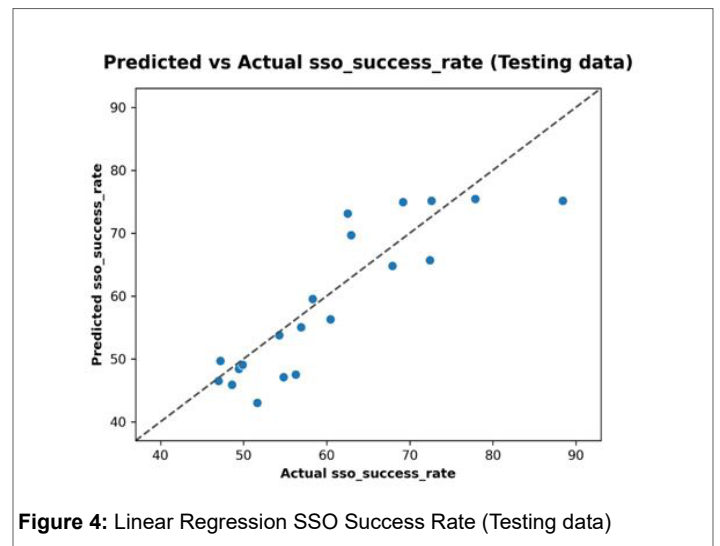


Figure 4 presents the predicted and actual SSO success rates for the test data, providing a visual assessment of how well the linear regression model generalizes to the unobserved data. Each point represents a test case, and the dashed line represents the best fit where the predictions are equal to the actual values. Compared to the training data in Figure 3, the scatter in this plot is more noticeable and less tightly clustered around the ideal line, especially for low and medium-level success rates. This indicates a modest decline in model performance when applied to new data, which is expected due to potential overfitting or the model's limited ability to capture complex, nonlinear relationships in the data. Some predictions fall significantly above or below the diagonal, indicating areas where the model overestimates or underestimates the actual success rate. Despite this, the overall trend is still positively aligned, indicating that the model retains some predictive power. However, the spread suggests that performance can be improved by improving feature selection, using regularization, or exploring more sophisticated models such as decision trees or ensemble methods. In summary, while the linear model shows potential, its generalization performance on the test set reveals limitations that warrant further optimization.

Table 2. Performance Metrics of Linear Regression SSO Success Rate(Training, Testing Data)									
Data	Symbol	R2	EVS	MSE	RMSE	MAE	MaxError	MSLE	MedAE
Train	LR	0.78663	0.78663	26.14639	5.11335	4.00695	15.24765	0.00685	3.09775
Test	LR	0.71942	0.74155	34.12245	5.84144	4.57210	13.24868	0.00870	2.91240

Table 2 provides performance metrics for the linear regression model for predicting SSO success rate on both the training and test datasets. The model shows reasonably strong performance on several key regression evaluation metrics, with some notable insights. For the training data, both the  $R^2$  (coefficient of determination) and the explained variance score (EVS) are approximately 0.786, indicating that the model explains about 78.6% of the variance in the SSO success rate. The root mean square error (RMSE) is 5.11, and the mean absolute error (MAE) is 4.01, reflecting relatively low prediction errors on average. The maximum error is 15.25, showing a large deviation in the prediction. The mean square log error (MSLE) and mean absolute error (MedAE) are low, indicating stable prediction performance. For the test data, this model shows a slight performance drop -  $R^2$  drops to 0.719, and RMSE increases to 5.84, with a high MAE of 4.57. Although this drop is expected due to generalization to unobserved data, the measurements still indicate good predictive ability.



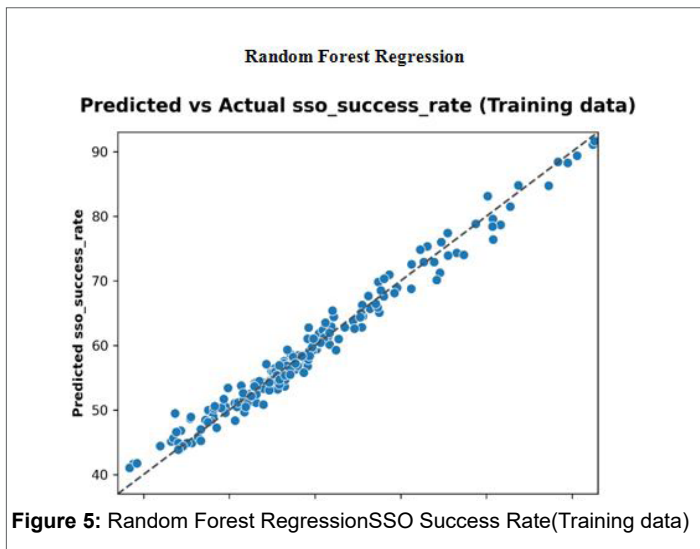


Figure 5 illustrates the performance of the random forest regression model in predicting the SSO (single sign-on) success rate using the training dataset. The scatter plot compares the actual values (x-axis) with the predicted values (y-axis) of sso success rate. Each point on the plot represents a data event. The presence of a dotted 45-degree reference line (where predicted equals actual) helps us visually assess the prediction accuracy. The closer the points are to this diagonal line, the better the model performed. From the plot, we observe a strong alignment between the predicted and actual values, indicating that the model has effectively captured the underlying patterns in the training data. The points are tightly clustered around the diagonal with minimal scatter, indicating low prediction error. This strong correlation reflects a high level of model accuracy and a good fit to the training data. However, while this performance is promising, it is very important to evaluate the model on unobserved (test) data to ensure that it generalizes well and avoids overfitting.

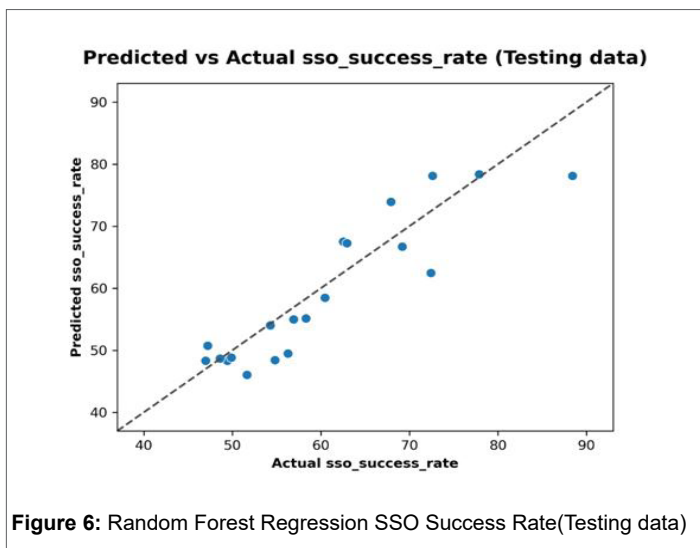


Table 3 summarizes the performance metrics of the random forest regression (RFR) model used to predict SSO success rate. On the training dataset, the model performs very well, with an  $R^2$  of 0.97695 and an explained variance score (EVS) of 0.97696, indicating that almost all of the variation in the success rate is explained by the model. The error values are very low, with  $MSE = 2.82$ ,  $RMSE = 1.68$ , and  $MAE = 1.32$ , indicating

that the predictions are close to the true values. The mean absolute error ( $MedAE = 1.11$ ) reinforces the accuracy of the model, while the maximum error (5.81) shows that only a few cases deviate significantly. On the test dataset, the performance decreases but remains robust. The  $R^2$  of 0.8054 and EVS of 0.8179 indicate that the model explains 80% of the variation in the unobserved data. The errors increase with  $MSE = 23.67$ ,  $RMSE = 4.86$ , and  $MAE = 3.86$ , showing larger prediction gaps compared to training. The maximum error rises to 10.26, reflecting occasional large deviations. However, the MSLE remains low (0.00574), confirming that the impact of logarithmic error is minimal.

## Conclusion

This detailed analysis of the single sign-on implementation in Oracle Insurance Policy Management systems reveals important insights that can significantly improve organizational authentication strategies. The research demonstrates that authentication response time emerges as the most influential factor affecting SSO success rates, with a strong negative correlation of -0.79. This finding underscores the critical importance of optimizing system performance to maintain user satisfaction and operational efficiency in large-scale enterprise environments. The study's comparative analysis between linear regression and random forest regression models provides valuable methodological insights for predictive analytics in authentication systems. While linear regression achieves respectable performance with an  $R^2$  of 0.719 on the test data, the superior performance of the random forest regression model ( $R^2 = 0.805$ ) demonstrates the effectiveness of ensemble methods in capturing the complex, nonlinear relationships inherent in authentication system dynamics. This suggests that organizations should consider advanced machine learning approaches when analyzing and optimizing their SSO implementations. Integration complexity analysis reveals an important trade-off between system sophistication and reliability. The research indicates that simple, well-designed architectures often outperform complex implementations, with complexity scores ranging from 3.10 to 88.65 and their modest negative correlation with success rates (-0.43).

This finding challenges the conventional assumption that more sophisticated systems necessarily produce better results, instead supporting streamlined approaches that prioritize user experience and system stability. Despite the varying response times and success rates, the consistency in daily successful logins (average 5005.07 with a low standard deviation) indicates that user behavior is relatively stable even when faced with authentication challenges. This resilience suggests that while users may tolerate occasional authentication issues, persistent issues can ultimately impact user engagement and system adoption. From a practical implementation perspective, the research provides actionable recommendations for organizations deploying SSO solutions in Oracle environments. Priority should be given to reducing authentication response times below 500ms, as this threshold appears to be associated with success rates greater than 70%. In addition, integration strategies should favor simplicity over complexity, focusing on robust, streamlined architectures that reduce potential points of failure. Future research directions should explore temporal aspects of authentication performance, examine the impact of external factors such as network infrastructure and user behavior patterns, and explore the measurable impacts of these findings across different organizational sizes and industry sectors. The development of real-time monitoring systems that can predict and prevent authentication failures based on identified performance indicators represents a promising avenue for improving SSO reliability and user experience in enterprise environments.

## References

1. Gosangi, Sreenivasula Reddy. "Scalable Single Sign-On Architecture: Securing Access in Large Enterprise Systems." *International Journal of Technology, Management and Humanities* 10, no. 02 (2024): 27-33.
2. Kreizman, Gregg. "MarketScope for Enterprise Single Sign-On." *Gartner RAS Core Research Note G 170568* (2011).
3. Chepuru, Anitha, K. Venugopal Rao, and Amardeep Matta. "Web Applications Access Control Single Sign On."
4. Alom, Ifteher, SudipBhujel, and Yang Xiao. "VeriSSO: A Privacy-Preserving Legacy-Compatible Single Sign-On Protocol Using Verifiable Credentials." *Cryptology ePrint Archive* (2025).
5. Sridhar Kakulavaram. (2024). Artificial Intelligence-Driven Frameworks for Enhanced Risk Management in Life Insurance. *Journal of Computational Analysis and Applications (JoCAAA)*, 33(08), 4873–4897. Retrieved from <https://www.eudoxuspress.com/index.php/pub/article/view/2996>
6. Shi, Shangcheng, Xianbo Wang, and Wing Cheong Lau. "MoSSOT: An automated blackbox tester for single sign-on vulnerabilities in mobile applications." In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, pp. 269-282. 2019.
7. Norta, Alex, AlexandrKormiltyn, ChibuzorUdokwu, VimalDwivedi, Sunday Aroh, and IgnasNikolajev. "A blockchain implementation for configurable multi-factor challenge-set self-sovereign identity authentication." In *2022 IEEE International Conference on Blockchain (Blockchain)*, pp. 455-461. IEEE, 2022.
8. Peram, S. R. (2024). Automated Label Detection and Recommendation System Using Deep Convolution Neural Networks and SPSS-Based Evaluation. *International Journal of Computer Science and Data Engineering*, 1(2), 258. <https://doi.org/10.55124/csdb.v1i2.258>
9. Müller, Johannes, and Jan Oupický. "Post-quantum XML and SAML Single Sign-On." *Proceedings on Privacy Enhancing Technologies* 2024, no. 4 (2024): 525-543.
10. Pareek, Chandra Shekhar. "Advancing Test Oracle Methodologies for Operational Excellence in Life Insurance." *IJSAT-International Journal on Science and Technology* 16, no. 1 (2025).
11. Gonugunta, Krishna C., and TsakiridisSotirios. "Advanced Oracle Methodologies for Operational Excellence." *International Journal of Modern Computing* 3, no. 1 (2020): 11-25.
12. Raghavendra Sunku. (2024). AI-Powered Forecasting and Insights in Big Data Environments. *Journal of Business Intelligence and Data Analytics*, 1(2), 254. <https://doi.org/10.55124/jbid.v1i2.254>
13. Adabala, Sai Krishna. "Transforming HR Processes with Oracle Cloud: Best Practices for Implementation."
14. Radha, Vedala, and D. Hitha Reddy. "A survey on single sign-on techniques." *Procedia Technology* 4 (2012): 134-139.
15. De Clercq, Jan. "Single sign-on architectures." In *International Conference on Infrastructure Security*, pp. 40-58. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.
16. PK Kanumarlapudi. "Improving Data Market Implementation Using Gray Relational Analysis in Decentralized Environments" *Journal of Artificial Intelligence and Machine Learning.*, 2024, vol. 2, no. 1, pp. 1–7. doi: <https://dx.doi.org/10.55124/jaim.v2i1.271>
17. Pashalidis, Andreas, and Chris J. Mitchell. "A taxonomy of single sign-on systems." In *Australasian conference on information security and privacy*, pp. 249-264. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003.
18. Jannett, Louis, Maximilian Westers, Tobias Wich, Christian Mainka, Andreas Mayer, and VladislavMladenov. "Sok: Sso-monitor-the current state and future research directions in single sign-on security measurements." In *2024 IEEE 9th European Symposium on Security and Privacy (EuroS&P)*, pp. 173-192. IEEE, 2024.
19. Ghasemisharif, Mohammad, Amrutha Ramesh, Stephen Checkoway, Chris Kanich, and Jason Polakis. "O single {Sign-Off}, where art thou? an empirical analysis of single {Sign-On} account hijacking and session management on the web." In *27th USENIX security symposium (USENIX security 18)*, pp. 1475-1492. 2018.