

## Insurance Cold Calling Optimization Using Random Forest Regression Compared with Logistic Regression for Improved Accuracy

### INTRODUCTION

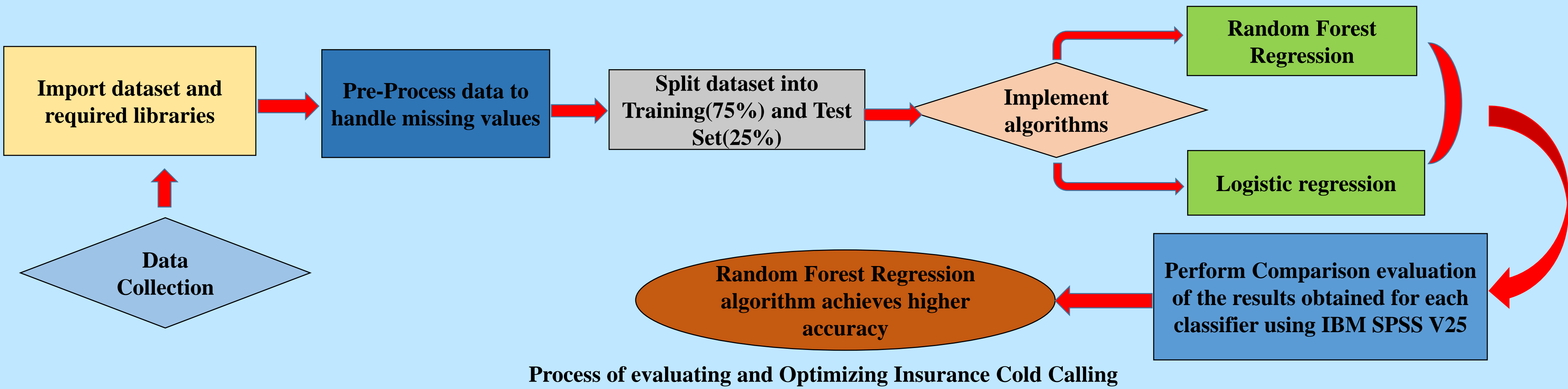
- The primary aim of this research is to optimize the cold-calling process for insurance companies, focusing on improving efficiency, reducing resource consumption, and enhancing overall productivity. By refining cold-calling strategies, insurance companies can more effectively attract new customers and maximize profits
- To identify factors like call timing, audience demographics, and communication skills that affect cold-calling success and to evaluate the effectiveness of various machine learning techniques, such as Random Forest Regression and Logistic regression, in streamlining marketing and sales initiatives for insurance companies
- Use machine learning algorithms, such as Random Forest Regression and Logistic regression, to analyze and optimize cold-calling strategies
- Enhanced sales outcomes for insurance companies, with higher conversion rates and increased customer satisfaction, resulting from more targeted and data-driven cold-calling methods

### General Cold Calling Statistics



Fig 1. General Cold Calling statistics

### MATERIALS AND METHODS



### RESULTS

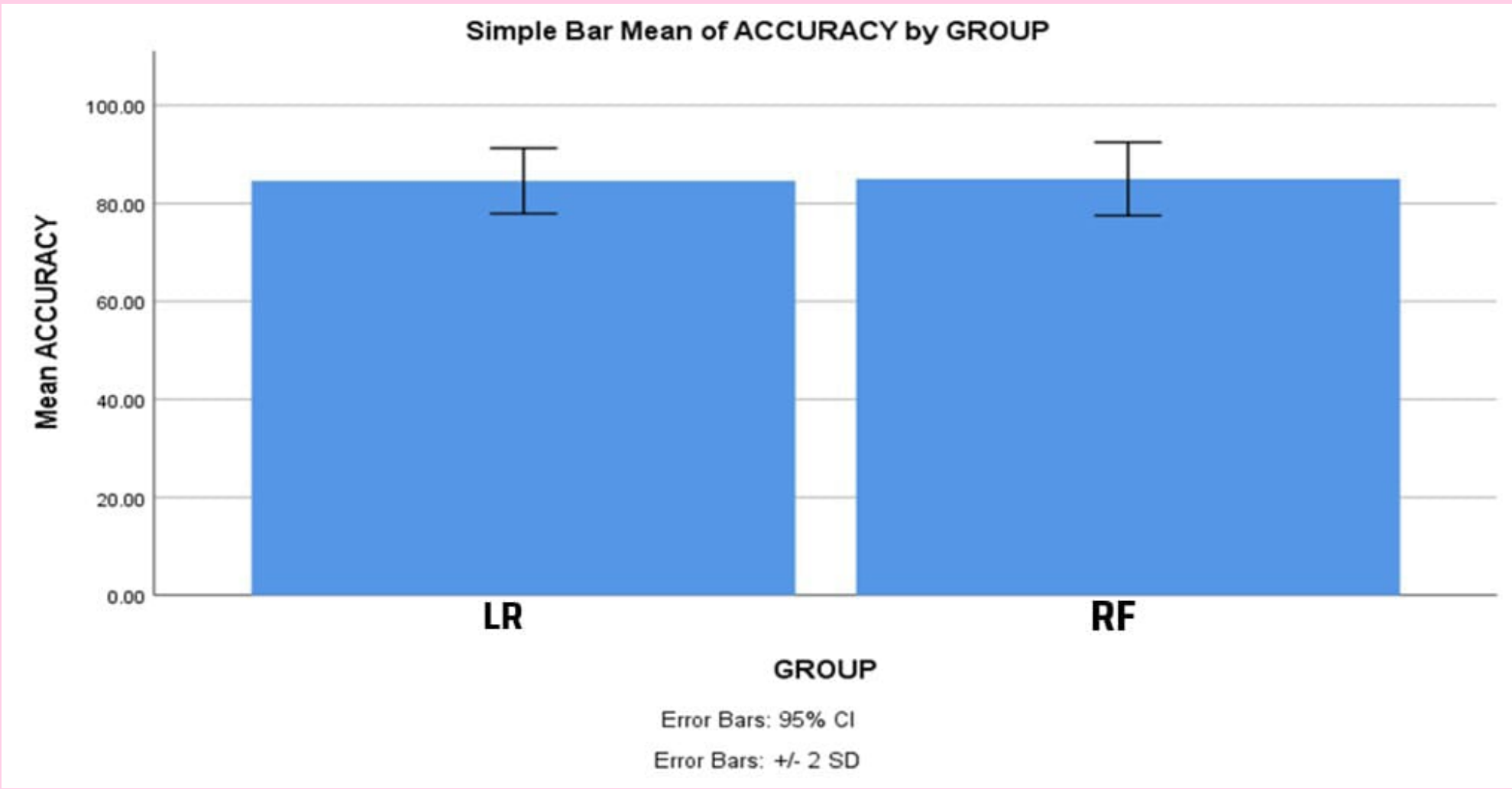


Fig 2. Mean Accuracy Graph

Table 1. The Mean Accuracy of the Random Forest Regression algorithm and Logistic regression

Groups		N	Mean	STD. Deviation	STD. Error Mean
Accuracy	Random Forest Regression	10	85.4000	3.74166	1.18322
	Logistic Regression	10	84.0000	3.06232	0.96839

- The figure indicates the mean Accuracy of the Random forest regression which is greater than the Logistic regression algorithm. X axis is Random forest regression vs Logistic regression, Y axis is Mean Accuracy. Error bar is +/- 2 SD

### DISCUSSION AND CONCLUSION

- Random Forest Regression outperformed Logistic Regression in predicting insurance cold-call outcomes, achieving a higher average accuracy of 85.4% compared to 84.0%
- Statistical analysis revealed differences in standard deviation and mean standard error between the two algorithms, indicating variability in predictions and model stability
- Independent variable statistical calculation showed a significance level of 0.273 in comparing the accuracy rates between Random Forest Regression and Logistic Regression
- The superiority of Random Forest Regression is attributed to its effectiveness in optimizing cold-calling procedures, aligning with the procedural nature of the cold calling domain
- Logistic Regression, while offering interpretability, may struggle with capturing complex temporal relationships in insurance cold calling data
- Acknowledgment of study limitations, including its focus on a specific dataset and potential variation in algorithm performance across different data sources or insurance firms

### BIBLIOGRAPHY

- Anderson, Raymond. 2022. The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation. Oxford University Press. DOI: 10.1093/9780199226405.001.000
- Brownlee, Jason. 2021. Optimization for Machine Learning. Machine Learning Mastery. DOI: 10.1017/CBO9781316402276
- Denuit, Michel, Donatien Hainaut, and Julien Trufin. 2022. Effective Statistical Learning Methods for Actuaries I: GLMs and Extensions. Springer Nature
- DOI: 10.1007/978-3-030-25827-6
- Géron, Aurélien. 2021. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. “O’Reilly Media, Inc.”
- Kumar, Sandeep, Anuj Sharma, Navneet Kaur, Lokesh Pawar, and Rohit Bajaj. 2024. Optimized Predictive Models in Health Care Using Machine Learning. John Wiley & Sons. 10.1002/9781394175376.ch11