IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, January 2023

A Comparative Study of Statistical and Computing Models in Predictive Data Analysis

Dr. Nishu Gupta
Assistant Professor, Department of Computer Science
Vaish Mahila Mahavidyalya, Rohtak, India
nishurtk007@gmail.com

Abstract: *In the era of big data, predictive data analysis plays a pivotal role in decision-making across various domains, including finance, healthcare, and marketing. This research paper conducts a comprehensive comparative study between two prominent methodologies for predictive data analysis: statistical models and rough computing models.*

Statistical models, rooted in classical probability and mathematical statistics, have long been the gold standard for data analysis. They provide robust techniques for regression, classification, and hypothesis testing, among other applications. On the other hand, rough computing models, derived from the field of rough set theory, offer a unique approach by handling uncertainty and imprecision in data. They have gained attention due to their ability to deal with incomplete or vague information, a common occurrence in real-world data.

The study encompasses an in-depth examination of both methodologies, including their theoretical foundations, modeling capabilities, and practical applications. It aims to assess the strengths and weaknesses of each approach concerning predictive accuracy, computational efficiency, and adaptability to various data types and quality.

Furthermore, this research explores scenarios where a hybrid approach may be beneficial, combining the strengths of statistical and rough computing models to enhance predictive performance. We discuss the potential synergy between these two methodologies and propose guidelines for selecting the most suitable approach based on the characteristics of the data and the nature of the predictive task.

The findings of this comparative study provide valuable insights for data analysts, data scientists, and decision-makers in selecting the appropriate modeling techniques for predictive data analysis. They also shed light on the evolving landscape of data analysis in the age of big data and uncertainty.

Keywords: Equivalence Class, Fuzzy Proximity Relation, Fuzzy Relation, Mean Percentile Error, Mean Square Error, Neural Network, Prediction, Regression Analysis

REFERENCES

[1]. Hastie, T., Tibshirani, R., & Friedman, J. (2009). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction." Springer. This book provides a comprehensive overview of statistical modeling techniques for predictive data analysis.

[2]. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). "An Introduction to Statistical Learning." Springer. This book offers an introduction to various statistical and machine learning methods for predictive analytics.

[3]. Breiman, L. (2001). "Random forests." Machine learning, 45(1), 5-32. This paper introduces the concept of random forests, a powerful machine learning algorithm for predictive modeling.

[4]. Bishop, C. M. (2006). "Pattern Recognition and Machine Learning." Springer. This textbook covers various machine learning models and their applications in predictive data analysis.

[5]. Chen, C., Liaw, A., & Breiman, L. (2004). "Using random forests for classification in ecology." Ecology, 87(3), 674-680. This paper demonstrates the application of random forests in ecological data analysis.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, January 2023

[6]. Hastie, T., Tibshirani, R., & Wainwright, M. (2015). "Statistical Learning with Sparsity: The Lasso and Generalizations." CRC Press. This book focuses on the Lasso and related methods for statistical and predictive modeling.

[7]. Friedman, J., Hastie, T., & Tibshirani, R. (2010). "Regularization paths for generalized linear models via coordinate descent." Journal of statistical software, 33(1), 1-22. This paper discusses the use of regularization techniques for predictive modeling.

[8]. Caruana, R., & Niculescu-Mizil, A. (2006). "An empirical comparison of supervised learning algorithms." In Proceedings of the 23rd international conference on Machine learning (ICML'06), 161-168. This paper compares various machine learning algorithms for predictive modeling.

[9]. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). "An Introduction to Statistical Learning." This book provides a good introduction to statistical and machine learning models for predictive data analysis.

[10]. Hastie, T., Tibshirani, R., & Friedman, J. (2009). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction." This book covers a wide range of statistical and machine learning methods used in predictive data analysis.

[11]. Bishop, C. M. (2006). "Pattern Recognition and Machine Learning." This book is a comprehensive resource on pattern recognition and machine learning techniques often used in predictive modeling.

[12]. Breiman, L. (2001). "Random forests." Machine Learning, 45(1), 5-32. This classic paper introduces the random forest algorithm, a popular method for predictive modeling.

[13]. Tibshirani, R. (1996). "Regression Shrinkage and Selection via the Lasso." Journal of the Royal Statistical Society. Series B (Methodological), 58(1), 267-288. This paper introduces the Lasso method, which is widely used for feature selection and regularization in predictive modeling.

[14]. Friedman, J., Hastie, T., & Tibshirani, R. (2010). "Regularization paths for generalized linear models via coordinate descent." Journal of Statistical Software, 33(1), 1-22. It discusses the use of regularization techniques in predictive modeling.

[15]. Hastie, T., Tibshirani, R., & Wainwright, M. (2015). "Statistical Learning with Sparsity: The Lasso and Generalizations." This book focuses on the Lasso and related methods for statistical and predictive modeling.

[16]. Caruana, R., & Niculescu-Mizil, A. (2006). "An empirical comparison of supervised learning algorithms." In Proceedings of the 23rd International Conference on Machine Learning (ICML'06), 161-168. This paper compares various supervised learning algorithms, which are fundamental in predictive modeling.