conferenceseries.com

10th International Conference on DATA SCIENCE AND MACHINE LEARNING APPLICATIONS

August 08, 2022 | Webinar

Generalized Robust Feature Selection

Bradford L Lott
Air Force Institute of Technology

Feature selection may be summarized as identifying salient data features to a given response. Understanding which features affect the response enables, in the future, limiting data collection to only consequential data; hence, the feature selection algorithm may lead to saving effort spent collecting and storing data, as well as reducing computational resources for making predictions. We propose a generalized robust approach to select the salient features of data sets. Our approach may also be applied to unsupervised datasets to understand which data streams provide unique information. Our approach identifies salient features robust to the subsequent predictive model applied. The proposed algorithm considers all provided variables, square variables, and two-way interactions as an extended data set. The algorithm implements a forward selection approach, based on correlation with the response, while fitting deep neural networks to the selected variables. These deep neural networks maintain an adaptive architecture which mirrors a full factorial experimental design. These networks assess both numeric and categorical values for both features and responses. Implementing this approach in ensemble with Recursive Feature Elimination we plot retained features versus predictive accuracy to create a new Pareto Frontier of feature selection methodologies, for the Wisconsin Breast Cancer problem instance. This Pareto Frontier highlights our ensemble approach as the top performing method in both feature reduction and predictive accuracy: compared to existing methods including the family of Boruta techniques, our method maintains fewer features and improves predictive accuracy.

Biography

Bradford Lott is an Operations Research Analyst whose research focuses in explainable artificial intelligence. Before working at the United States Air Force Research Laboratory's Explainable Artificial Intelligence Lab, Bradford received his Bachelors of Science in biomathematics from Florida State University. In 2022, Bradford completed his Masters of Science in operations research with a focus in data science at the Air Force Institute of Technology. Bradford's previous research domains include: disease modeling, agent based modeling, explainable artificial intelligence, deep learning, battle space acoustics, and design of experiments. Bradford's future research will focus on developing explainable artificial intelligence based tools to benefit the medical and public health communities..