Title: StackIL10: A stacking ensemble model for the improved prediction of IL-10 inducing peptides.

Author Name: Izaz Ahmmed Tuhin, Md. Rajib Mia, Md. Monirul Islam, Imran Mahmud, Henry Fabian Gongora, Carlos Uc Rios, Imran Ashraf, Md. Abdus Samad

Abstract: Interleukin-10, a highly effective cytokine recognized for its anti-inflammatory properties, plays a critical role in the immune system. In addition to its well-documented capacity to mitigate inflammation, IL-10 can unexpectedly demonstrate pro-inflammatory characteristics under specific circumstances. The presence of both aspects emphasizes the vital need to identify the IL-10-induced peptide. To mitigate the drawbacks of manual identification, which include its high cost, this study introduces StackIL10, an ensemble learning model based on stacking, to identify IL-10-inducing peptides in a precise and efficient manner. Ten Amino-acid-composition-based Feature Extraction approaches are considered. The StackIL10, stacking ensemble, the model with five optimized Machine Learning Algorithm (specifically LGBM, RF, SVM, Decision Tree, KNN) as the base learners and a Logistic Regression as the meta learner was constructed, and the identification rate reached 91.7%, MCC of 0.833 with 0.9078 Specificity. Experiments were conducted to examine the impact of various enhancement techniques on the correctness of IL-10 Prediction. These experiments included comparisons between single models and various combinations of stacking-based ensemble models. It was demonstrated that the model proposed in this study was more effective than singular models and produced satisfactory results, thereby improving the identification of peptides that induce IL-10..

Keyword: Peptide Identification, StackIL10, Machine Learning, Ensemble Learning, Stacking Classifier, Amino Acid Composition, Feature Extraction, Logistic Regression.

DOI: https://doi.org/10.1371/journal.pone.0313835