

Compare & pull request

...



6 minutes ago

- _____
- _____
- _____

Training Set Features: Contains various features related to individuals' demographic information, behavioral responses, and health beliefs. Training Set Labels: Contains the vaccination status (H1N1 vaccine and seasonal flu vaccine) for the individuals in the training set features.

Data Summary

Training Set Features: Includes columns such as respondent_id, h1n1_concern, h1n1_knowledge, behavioral indicators (e.g., behavioral_antiviral_meds, behavioral_avoidance), and demographic information (e.g., age_group, education, race, sex). Training Set Labels: Contains the columns respondent_id, h1n1_vaccine, and seasonal_vaccine.

Modeling

Methodology

The modeling process involves several key steps:

Data Preprocessing: This includes handling missing values, encoding categorical variables, and scaling numerical features. Feature Engineering: Additional features may be created based on existing ones to improve model performance. Model Training: Various machine learning models such as Logistic Regression, Random Forest, and Gradient Boosting are trained using the preprocessed data. Model Selection: The best model is selected based on its performance on validation data.

Models Used

Logistic Regression Random Forest Gradient Boosting

Evaluation

Metrics

The models are evaluated using several performance metrics:

Accuracy: The proportion of correctly classified instances. Precision: The proportion of positive identifications that are actually correct. Recall: The proportion of actual positives that were identified correctly. AUC-ROC: The area under the receiver operating characteristic curve, which measures the model's ability to distinguish between classes.

Results

Random Forest: Showed good accuracy and recall, highlighting its ability to handle complex interactions between features. Gradient Boosting: Delivered the best overall performance with the highest accuracy and AUC-ROC, indicating strong predictive capability.

Conclusion

The H1N1 vaccination prediction project successfully identified key factors influencing vaccination uptake and developed a predictive model to aid public health efforts. The Gradient Boosting model emerged as the best-performing model, offering significant insights for targeted vaccination campaigns. These findings can help public health authorities and

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)