

(a) Data Collection and Preprocessing

CPCD Four-Wave Survey

Cluster Sampling Method, Five Primary and Middle Schools Selected;
Baseline survey (T1, n=8,825; response rate=98.41%)

Data Entry Quality Control

Randomly double-enter 5% of each group's questionnaires; consistency >95%.

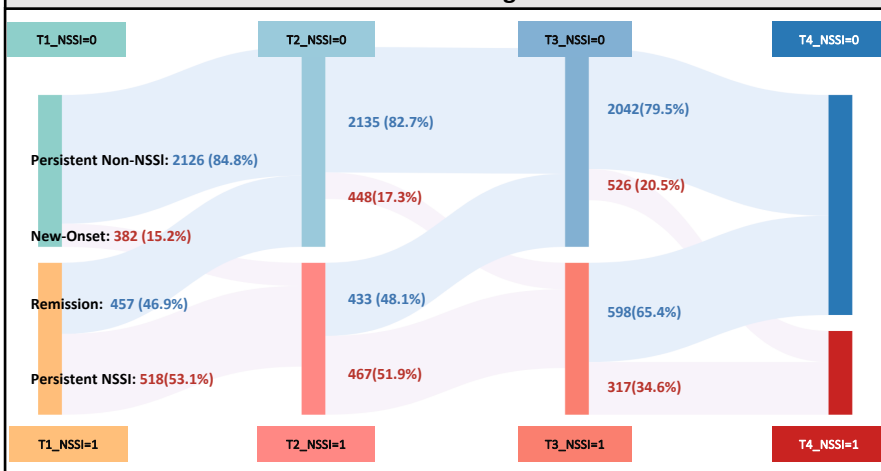
Exclusion Criteria

- Samples missing NSSI variables in any of the four survey waves;
- Samples with overall variable missing rate $\geq 50\%$;
- Samples where students were below 3rd grade at baseline survey

Final Analysis Sample

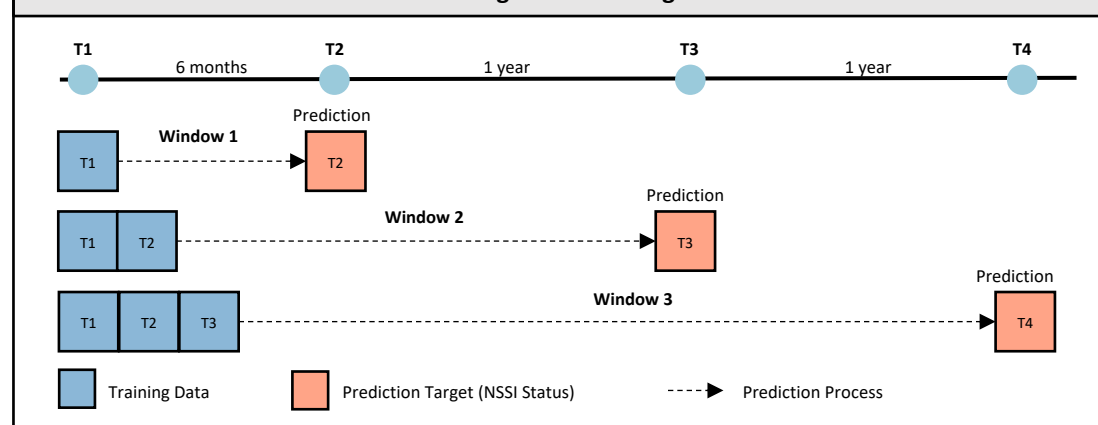
N=3483 (Male=1781, Female=1702), Wave=4

Transition of NSSI among Adolescents

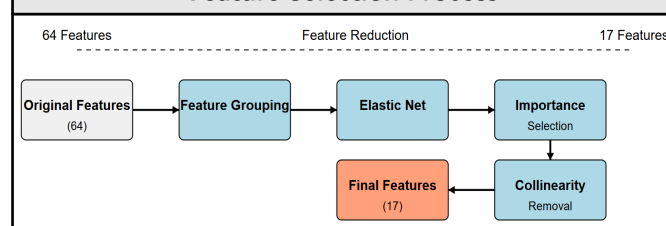


(b) Predictive Model Building

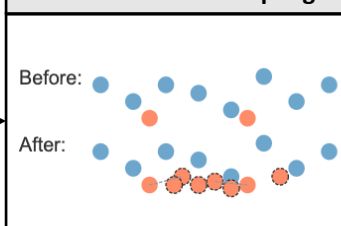
Sliding Window Design



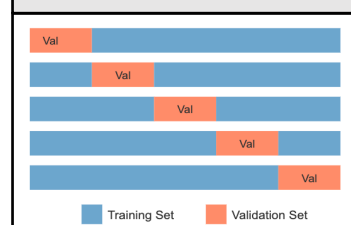
Feature Selection Process



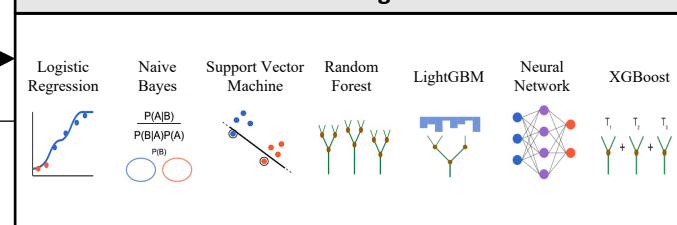
SMOTENC Oversampling



5-Fold Cross-Validation



Machine Learning Prediction

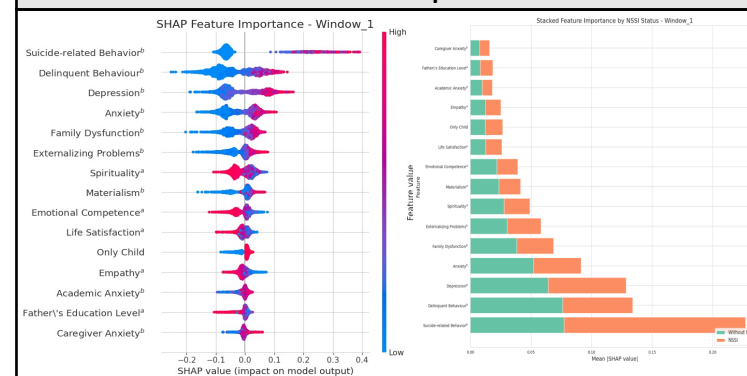


(c) Visualization Analysis

Model Performance

ROC-AUC, PRC-AUC, Accuracy, Precision, Recall, F1-Score

SHAP Feature Importance



SHAP Dependency Graph

