



**PREDICTING SUICIDAL IDEATION AMONG FILIPINO YOUTH: AN
APPLICATION OF MODELING WITH CLASS IMBALANCE DATA**

TEODOLFO F. BONITEZ

A Special Problem Submitted in Partial Fulfillment

Of the Requirements for the Degree of

MASTER OF STATISTICS

School of Statistics

University of the Philippines

Diliman, Quezon City

May 2017

ABSTRACT

Class imbalance data is a common problem for many field of researches, especially when the variable of interest measures a rare event, characteristic, or instance. Class imbalance occurs when the ratio of instances between the larger class (or majority group) and smaller class (or minority group) is very large. In this study, models for predicting suicidal ideation among Filipino youth are constructed. To address potential class imbalance data issues, this study utilizes data-level and algorithm-level approaches. Data-level approach consists of resampling methods (oversampling, undersampling, and hybrid techniques), while algorithm-level approaches considered in this study are Support Vector Machine (SVM), Fuzzy SVM, Adaptive Boosting, and Logistic Regression. The data is derived from the 2015 National Youth Assessment Study. Results suggest that relatively higher/improved sensitivity measures are observed using logistic regression and SVM models when using undersampling than when using oversampling. Also, fuzzy SVM and adaptive boosting models tend to have high measures in correctly classifying the majority group and have relatively improved measures (compared to data-level approaches) in correctly classifying the minority group. From the different models implemented in the actual survey data, several predictors of suicidal ideation are identified: youth who do harm to one's self, youth who have negative perception in life, youth with anorexic symptom, difference between urban and rural living, and average monthly income.

Keywords: *Class imbalance data; data-level approach, algorithm-level approach, suicide ideation, National Youth Assessment Study (2015 NYAS)*