



Data-Driven Analytics to Identify School Absenteeism Associated Risk and Protective Factors for Secondary School Students



Knoo Lee¹, Barbara J. McMorris², Chih-Lin Chi^{2, 3}, Wendy S. Looman², Connie W. Delaney²
¹ Postdoctoral Fellow, University of Missouri Sinclair School of Nursing at Columbia, MO, 65201
² School of Nursing, University of Minnesota-Twin Cities, MN, 55455, ³Institute of Health Informatics, University of Minnesota-Twin Cities, MN, 55455

INTRODUCTION

Chronic absenteeism (CA), defined as missing at least 15 school days/year, is recognized as a national problem in the U.S. with devastating long-term impacts for students. Previous studies have been guided by a mixture of diverse CA definitions and measurements which could potentially harm the applicability of findings. Despite the number of CA-associated factors identified, studies utilizing a unified theoretical system to a wide range of risk and protective factors has been scarce.

OBJECTIVE AND AIMS

Objective: Applying machine learning methods with “The Kids and Teens at School” theoretical framework to student-level data, this study identified risk and protective factors that are associated with school absences.

METHODS

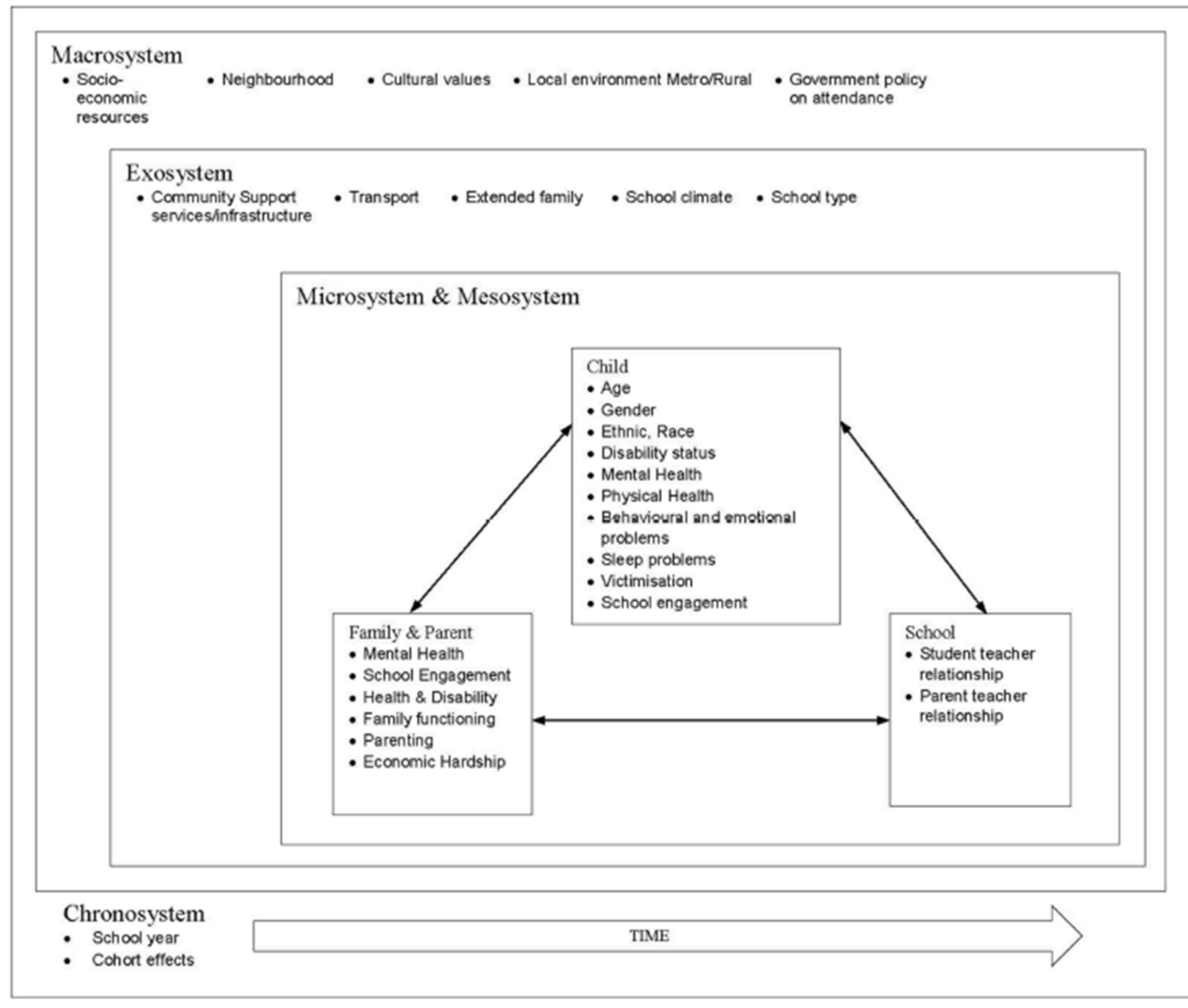
Using feature selection and prediction models performance comparison on de-identified student-level data (n = 121,005) from the Minnesota Student Survey 2016, factors associated with school absences were identified.

Sample Characteristic & Theoretical Framework & Data Analysis

Sample Characteristics

Study (N = 121,005)	Mean (SD) or %
Grade	
8th	35%
9th	36%
11th	29%
Gender: Male	50%
Race	
American Indian or Alaskan Native	2%
Asian	6%
Black	7%
Native Hawaiian, or Pacific Islander	1%
White	76%
Multiple Races	9%
Hispanic or Latino	7%
Primary Variables	
School Absences - High	6%
School Absences - Low	94%
In-school Suspension (any in the past month)	1.04 (0.26)
Out-of-school suspension (any in the past month)	1.02 (0.19)
School Engagement	3.15 (0.47)
Stayed home-sick (1 "non" to 5 "10+ times" in past month)	1.55 (0.75)
Teacher-student Relationship	2.04 (0.59)
Social Competency	3.07 (0.60)

KiTeS Framework

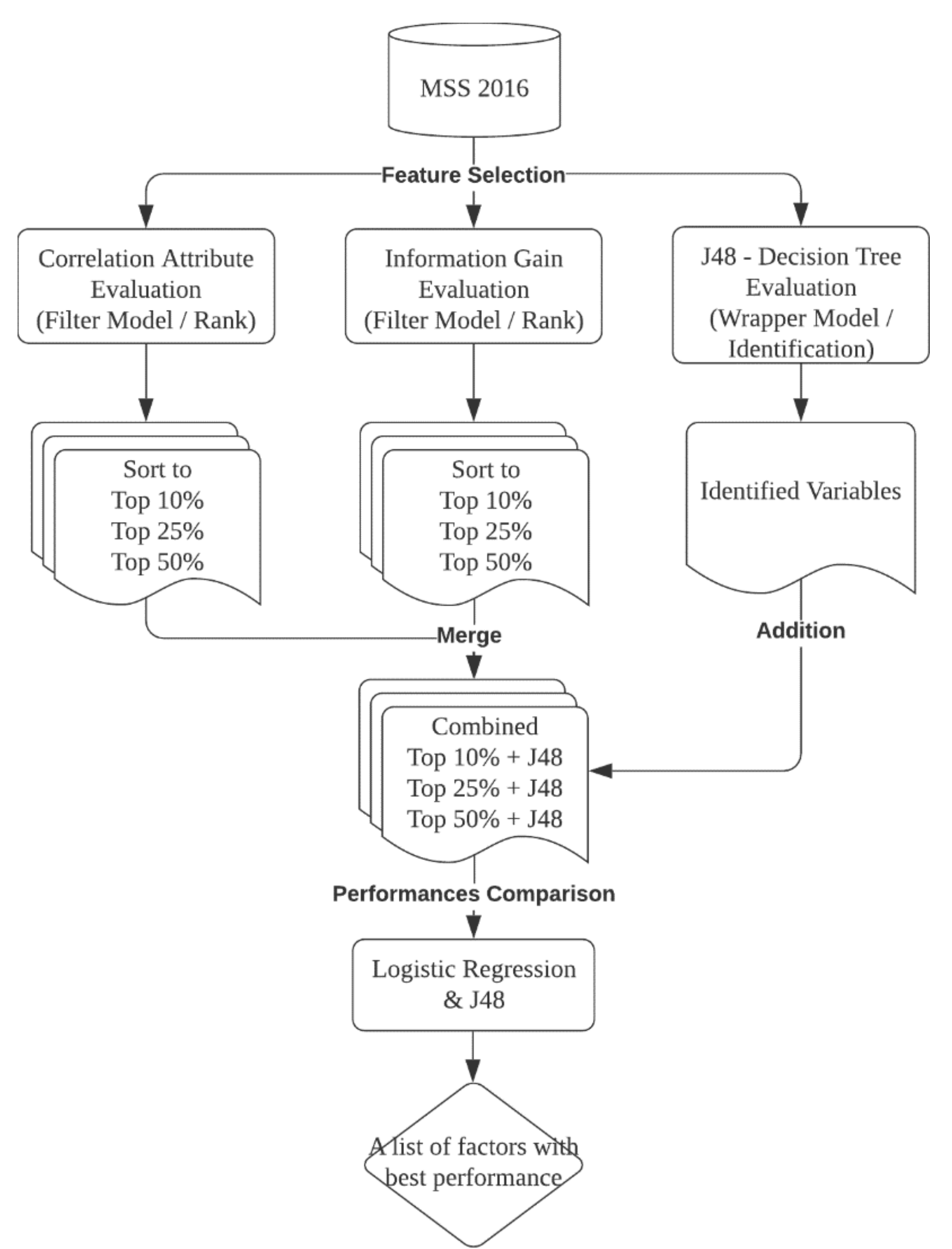


Results

By comparing three subsets of risk and protective factors acquired from feature selection procedure utilizing prediction models comparison, a total of 18 risk and protective factors associating with school absences were identified. All 18 factors are identified to be within either micro- or mesosystem such as substances usage, physical check-up, school engagement, and teacher-student relationship.

Data Analysis & Identified Factors

Analysis Pipeline



Identified Risk and Protective Factors

Attribute	System in the KiTeS
Social competency Scale (SCS)	Microsystem
Tobacco Product Use (TBP)	Microsystem
School engagement (SE)	Mesosystem
Friends approval of substance use (FAS)	Mesosystem
Sent to office for disciplinary issue	Mesosystem
Marijuana use past year	Microsystem
Marijuana use frequency	Microsystem
Staying home due to sickness	Microsystem
Teacher-student relationship (TSR)	Mesosystem
Substance use – 1	Microsystem
Substance use – 2	Microsystem
ACEs	Mesosystem
In-school suspension	Mesosystem
Binge drinking – 2 (5 or more drinks in a row)	Microsystem
Out-of-school suspension	Mesosystem
Race & Ethnicity: American Indian Non-Hispanic	Microsystem
Race: Native Hawaiian or Pacific Islander only	Microsystem
Physical Checkup	Microsystem

DISCUSSION

The proximity of school absences to micro- and mesosystem emphasizes the area that requires attention to address CA. These data suggest focusing on factors within both systems (e.g., child’s health, substance usage, school engagement) could be an ideal cost-saving intervention for students missing school. Future research needs to gather perspectives on CA from key stakeholders such as school nurses.

ACKNOWLEDGEMENTS

Sophia Fund from the School of Nursing – University of Minnesota – Twin Cities