

Comparison of machine learning models to predict US adolescent mental health trends

William Youkang Zhou, Walt Whitman High School, Bethesda, MD

Abstract

Objective

This study aimed to estimate recent trends in the frequency of adolescents feeling depressed or anxious using nationally representative data 2019-2023.

Method

The study utilized the 2019–2023 National Health Interview Survey’s Sample Child Interview, focusing on adolescents aged 12 to 17. The outcome measures were the frequency of adolescents seeming sad or depressed or anxious. Logistic regression models were applied to estimate the trends. Additionally, logistic regression, k-nearest neighbor (KNN) and decision tree machine learning models were trained to determine which approach best classified adolescents experiencing frequent depression or anxiety.

Results

The final sample included 14,167 adolescents. The percentage of adolescents reporting daily, weekly, or monthly feelings of anxiety increased from 22% in 2019 to 29% in 2023, with a peak of 36% in 2022. A similar pattern was observed for feelings of depression. Logistic regression results indicated that year indicators, White and African American adolescents (compared to other racial/ethnic groups), female gender, older age, U.S. citizenship, worse self-reported health, higher household income, higher parental education, and private insurance were significantly associated with a higher frequency of anxiety. Adolescents with married parents were less likely to report frequent anxiety. Patterns of the frequency of depression were similar. Machine learning models yielded similar results, with logistic regression demonstrating the highest accuracy in classifying whether an adolescent frequently felt depressed or anxious.

Conclusion

ML models produced similar results, with logistic regression demonstrating the highest accuracy in classifying adolescents who frequently experienced depression or anxiety. Further research is needed to explore the factors contributing to these trends and the utility of machine learning models in mental health surveillance.