Uncovering Hidden Trends in Absenteeism: A Machine Learning Perspective

Francis Bowen, Ph.D.
Butler University
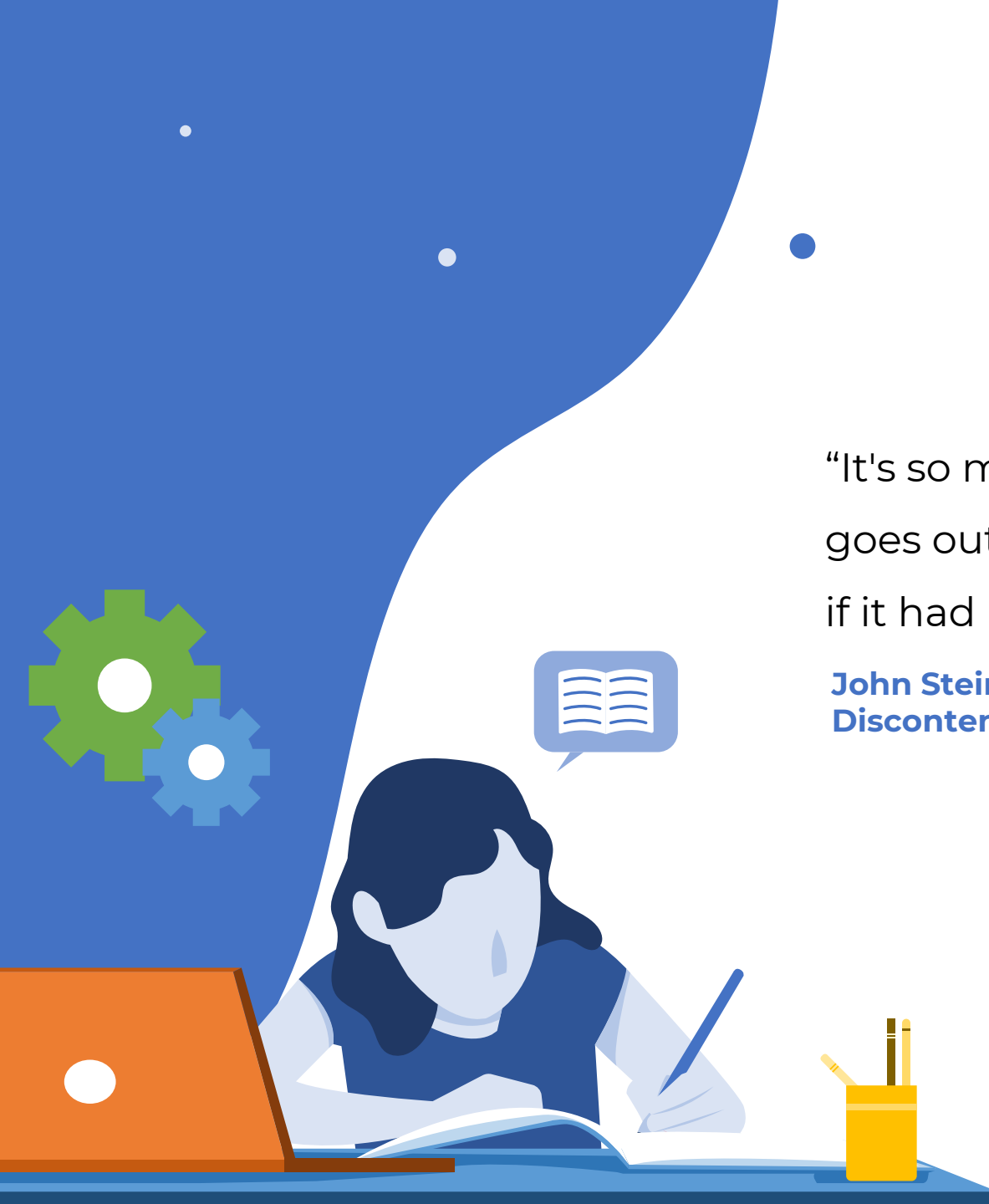
Carolyn Gentle-Genitty, Ph.D.
Indiana University

Marlin Jackson
Fight for Life Foundation

"It's so much darker when a light goes out than it would have been if it had never shone."

John Steinbeck, The Winter of Our Discontent

Outline

01.
Introduction

02.
Previous Literature

03.
Machine Learning Methodologies

04.
Results and Discussion

05.
Conclusion

Social Emotional Learning (SEL)
(Casel, 2012)

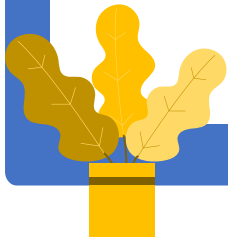Self-awareness

Self-management

Social awareness

Relationship skills

Responsible decision making

The Fight for Life Foundation

Individualized Interventions

Data-driven approach to offer support. behavior trend analysis, student and classroom analysis, automated student and class behavioral grouping, and office referral analysis

Targeted, Small Group Instruction

SEL video and lessons, implemented in small groups, as well as self-assessments that can trigger notifications to administration about concerns

Universal Instruction

Social Emotional Learning Lessons for PreK-12 - implemented with all students to develop self & social awareness, relationship skills, and responsible decision-making **Positive Behavioral Interventions & Supports system that acts as a decision-making game that leverages football terms** and analogies to reinforce positive behaviors
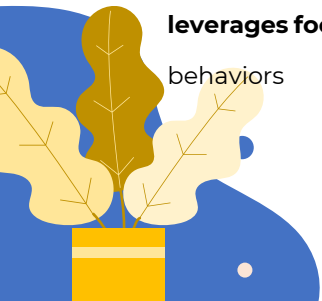
Core Values	
Description	**Code**
Enthusiastic in class	CV1
Focused within class	CV2
Meet or exceed expectations on assignments	CV3
Demonstrates initiative	CV4
Follow directions	CV5
Respect other's space	CV6
Respect for physical settings	CV7
Demonstrate accountability	CV8
Respectful communication	CV9
Positive relationships	CV10

First Down

Positively recognized core value

Extra Point

Positively recognized individual behavior

Sack

Negatively recognized core value

Flag

Negatively recognized individual behavior

Red Zone

Sack that results in an office referral

Using data collected in the Building Dreams Platform, what are the underlying factors of absenteeism?

Literature

Durlak et al., (2011), Taylor et al., (2017)

Two large scale studies of SEL, demonstrating the effectiveness of SEL in terms of drop-out rates and academic performance

Rastrollo-Guerrero J et al. (2020) and Albreiki B et al. (2021)

Extensive surveys of using **machine learning on academic behavior datasets.**

E. S. Bhutto et al (2020)

Demonstrated the effectiveness of **Support Vector Machines when applied to data centered on academic performance.**

- SEL is an effective tool that gives students the social and emotional foundations for success
- Machine learning can be used successfully on behavioral data from academic environments
- Very little research conducted at the K-12 levels
- No literature found applying machine learning directly to SEL-based data

Machine Learning

What is it?

"The field of study that gives computers the ability to learn without explicitly being programmed." – Arthur Samuel, AI Pioneer

Machine learning attempts to determine the relationships and patterns in data, improving in the presence of **more** and/or **better** data

Machine Learning Categories

Supervised Learning

Models developed by relating a set of inputs to a known set of results. Like humans, machines learn best through example.

Regression models: Relating a set of numbers to other numbers

Classification models: Relating data to categorical variables (labels)

Unsupervised Learning

Models developed to learn the final outcome from the patterns in the data

Clustering models: Finding the natural grouping of data

Dimensionality reduction: Finding comparable and smaller representations of data

Reinforcement Learning

Models developed that learn through a rewards-based system
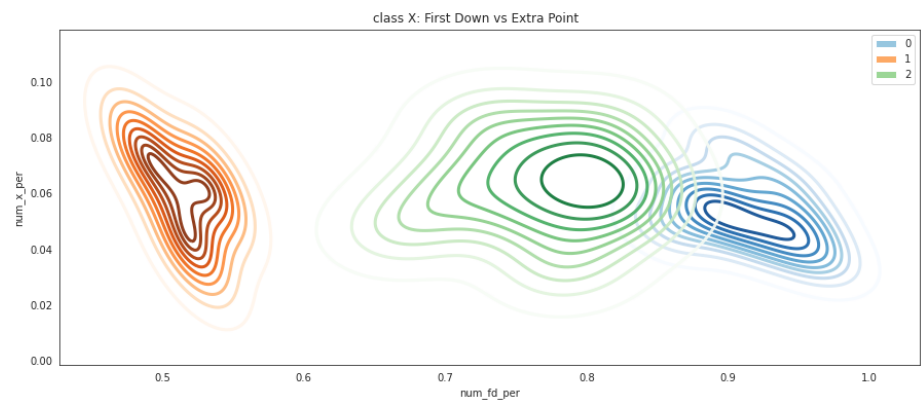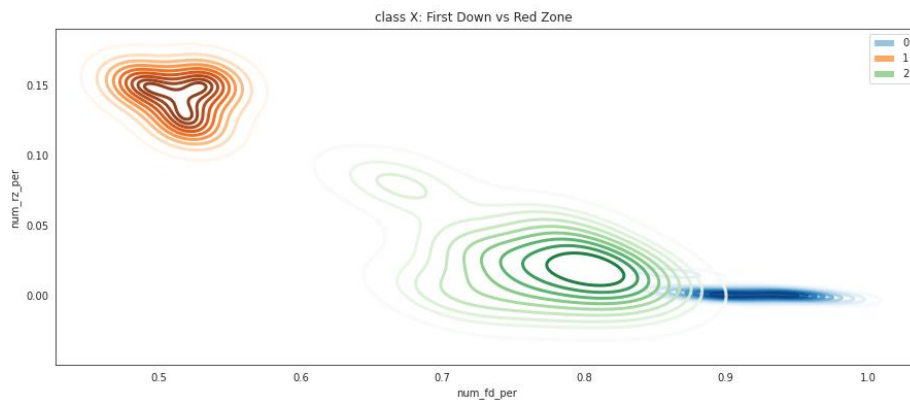
How FFLF Uses Unsupervised Learning

Dimensionality Reduction

In our proposed methodology, we perform dimensionality reduction **to represent each student with 5 dimensions of data**. This process aggregates each student's reports into a single data point with 5 dimensions which is used in other models

Clustering

Using the reduced representations for each students, **we create 3 groups; low-, medium-, and high-risk students per classroom**. Our results show that the data collected through Building Dreams are excellent differentiators.

Note: The following only shows 2 of the 8 clustering visualizations

How FFLF Uses Supervised Learning

Model Training

Training a classification model involves providing a list of aggregate reports for a student and an associated label

Classification

We use multiple classification models to associate a student's aggregate reports to one of three labels; low-, medium-, or high-risk. The data used to **train the model** to detect the **three labels** comes directly from **the clustering model**
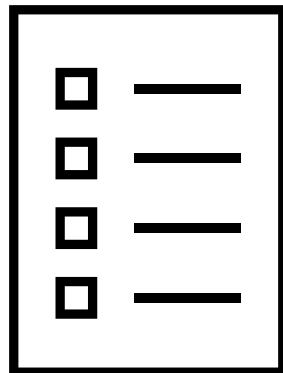
Aggregate report for Student X

Classification Models
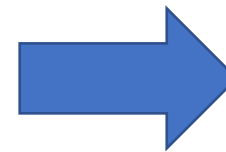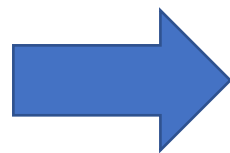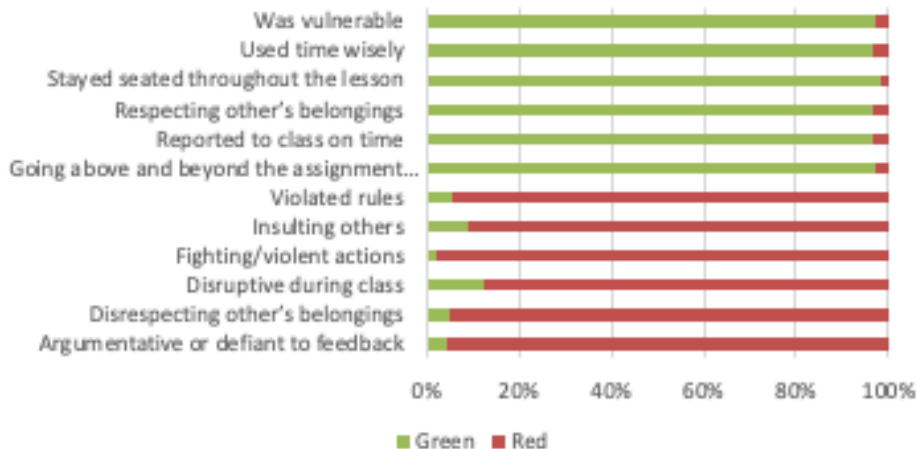
Low, medium, or high-risk

99.1%

Validation accuracy
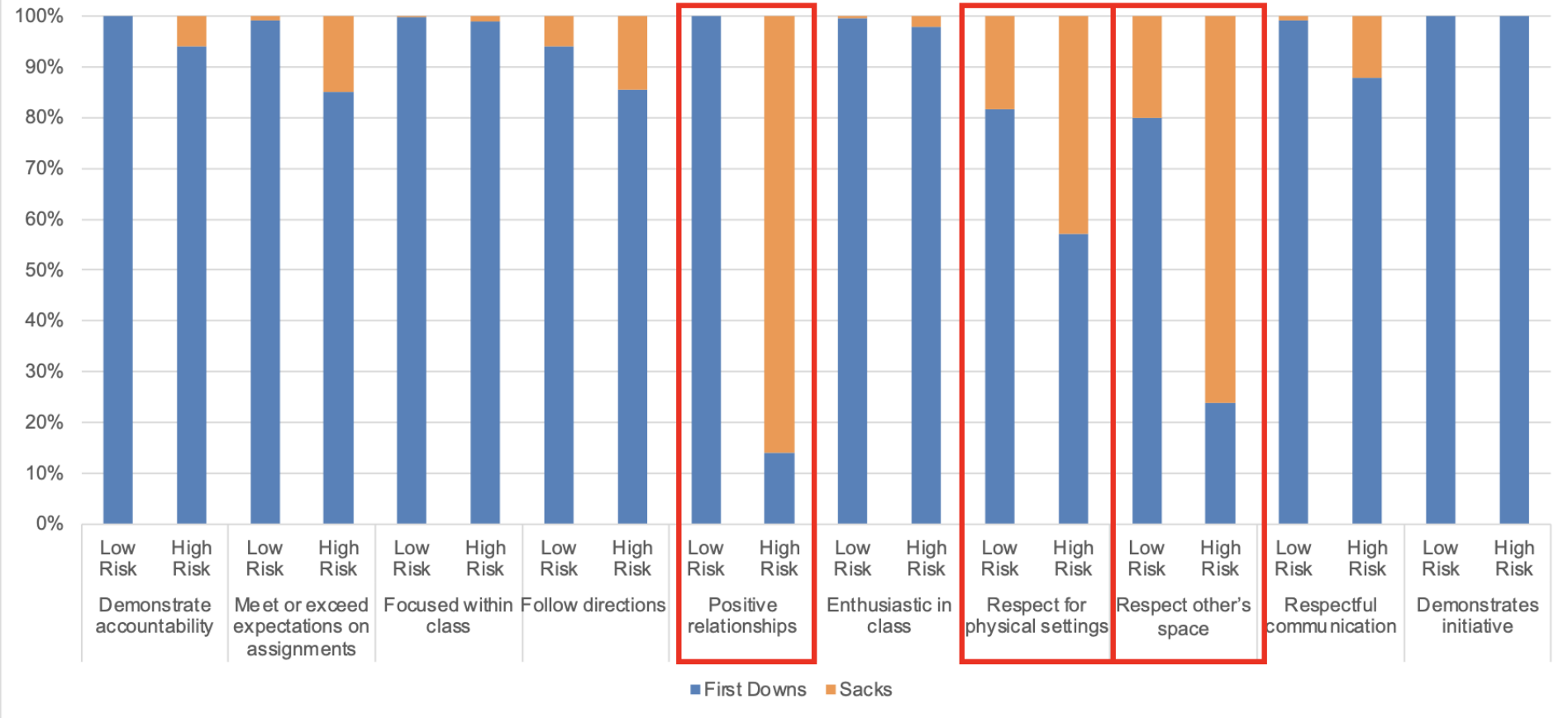
Using the Final Classifier

- Applied final classifier to all students and examined underlying reasons for first downs, sacks, and red zones
- There is a **clear difference in absence-related reasons** between **low and high-risk students**

Reasons by Percentage (All Students)

Core Value Differences

First Down vs Sacks by Cluster

The core value differentiators between low and high-risk students are all related to peer relationships

Conclusion

Machine Learning Effectiveness

We have proposed an effective **data-driven** approach for analyzing SEL-based data. The multi-phased approach leverages both **unsupervised and supervised models** for identifying low, medium, and high-risk students
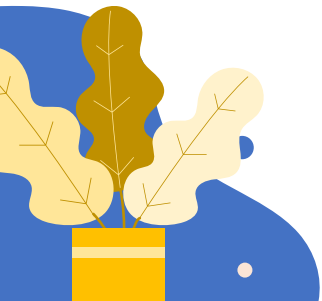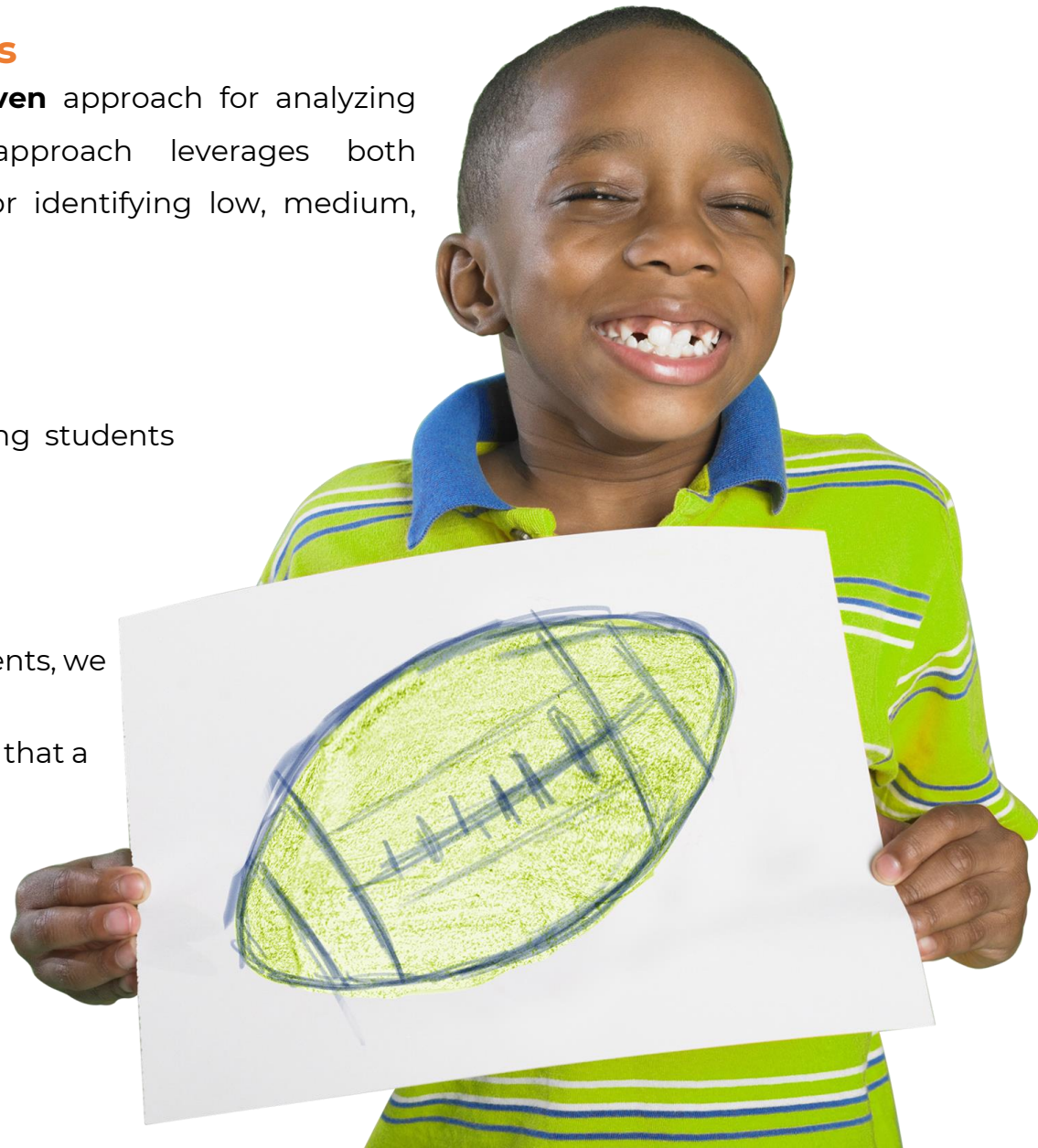
Absenteeism

The final model was effective at identifying students who **exhibit high levels of absenteeism**

Underlying Factors

Through analysis of low and high-risk students, we found that **peer relationships** are the clear differentiating factor. This finding suggests that a program rooted in SEL can be a means for improving absenteeism

Thank You

FIGHT FOR LIFE
F O U N D A T I O N

FFLF.org
information@FightforLifeFoundation.com

BUTLER

References

Albreiki B, Zaki N, Alashwal H. A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques. Education Sciences. 2021; 11(9):552. https://doi.org/10.3390/educsci11090552

A. S. Alblawi and A. A. Alhamed, "Big data and learning analytics in higher education: Demystifying variety, acquisition, storage, NLP and analytics," 2017 IEEE Conference on Big Data and Analytics (ICBDA), 2017, pp. 124-129, doi: 10.1109/ICBDAA.2017.8284118

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions. Educational Technology & Society, 25(1), 28–47. https://www.jstor.org/stable/48647028

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (cite arxiv:1810.04805Comment: 13 pages)

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. Child Develop- ment, 82(1): 405–432

E. S. Bhutto, I. F. Siddiqui, Q. A. Arain and M. Anwar, "Predicting Students' Academic Performance Through Supervised Machine Learning," 2020 International Conference on Information Science and Communication Technology (ICISCT), 2020, pp. 1-6, doi: 10.1109/ICISCT49550.2020.9080033.

Rastrollo-Guerrero J, Gómez-Pulido J and Durán-Domínguez A 2020 Analyzing and Predicting Students' Performance by Means of Machine Learning: A Review Applied Science 10 1042

Somers, R., Cunningham-Nelson, S., & Boles, W. (2021). Applying natural language processing to automatically assess student conceptual understanding from textual responses. Australasian Journal of Educational Technology, 37(5), 98–115. https://doi.org/10.14742/ajet.7121

Taylor RD, Oberle E, Durlak JA, Weissberg RP. Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects. Child Dev. 2017 Jul;88(4):1156-1171. doi: 10.1111/cdev.12864. PMID: 28685826