



UPTAKE OF SEASONAL FLU VACCINE

PRESNTATION BY: FLORENCE NGUUNI



CONTENT

01

INTRODUCTION

02

PROBLEM STATEMENT

03

BUSINESS OBJECTIVE

04

METRICS OF SUCCESS

05

MODELLING

06

CONCLUSION

07

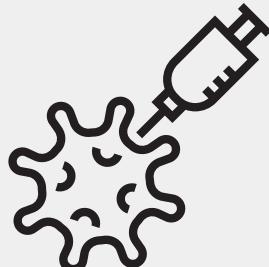
RECOMMENDATION



INTRODUCTION



Influenza, commonly known as the flu, is an illness that affects the respiratory system and can lead to severe complications, hospitalization, and even death, especially for vulnerable populations such as individuals over the age of 65, young children, and those with pre-existing conditions.



A vaccine for H1N1 became available to the public in October 2009, and the World Health Organization declared the end of the pandemic in August 2010. However, the H1N1 virus continues to circulate as a seasonal flu strain.



PROJECT OBJECTIVES



Main Objective

Main Objective: Develop a predictive model that can estimate the seasonal flu vaccine uptake by utilizing individual characteristics, background information, and behavioral patterns.



Objective n° 2

Additionally, the project seeks to uncover valuable patterns and associations between various factors and vaccine uptake, which can contribute to better understanding and improving public health strategies.

METRICS OF SUCCESS



The project's success criterion is to identify a model with the highest accuracy. The desired accuracy level for success is set at 80%.

MODELING

01



DECISION TREES

02



RANDOM FOREST

03



ADABOOST

04

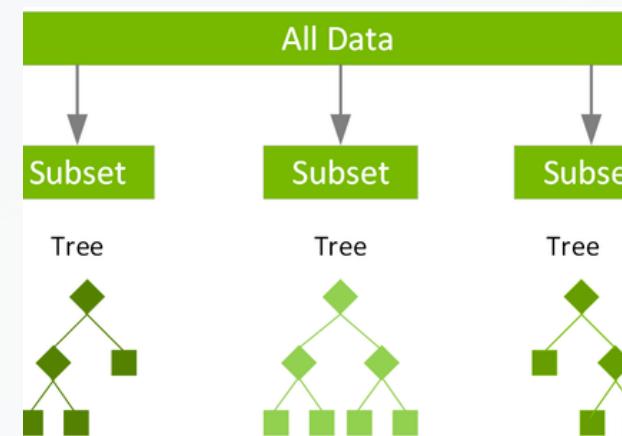
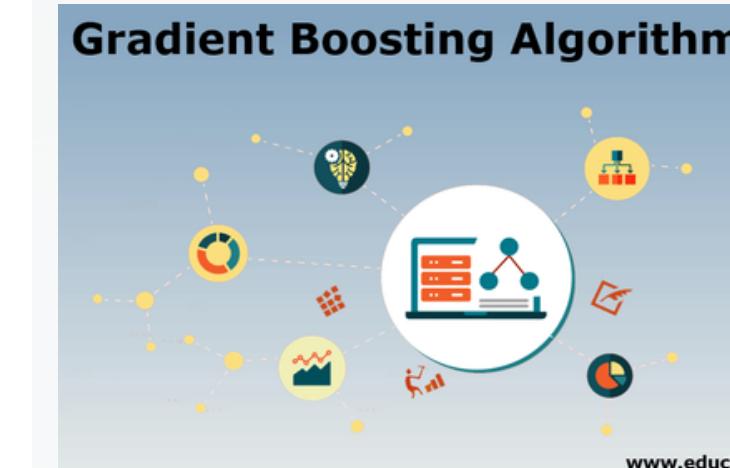
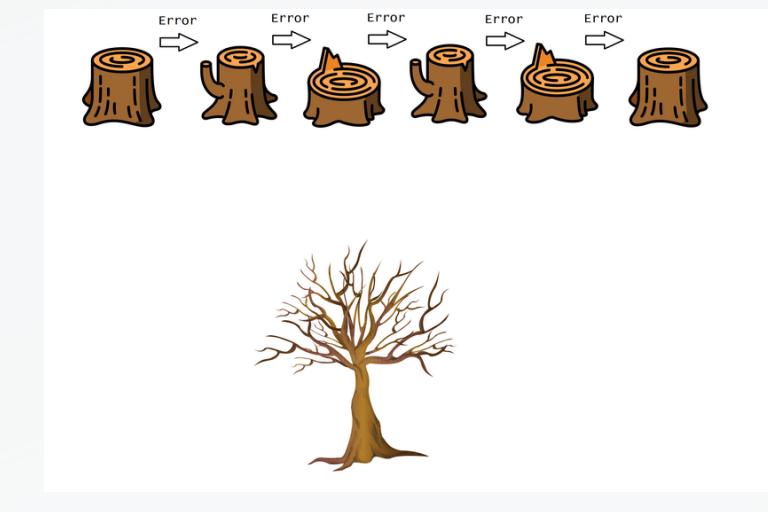
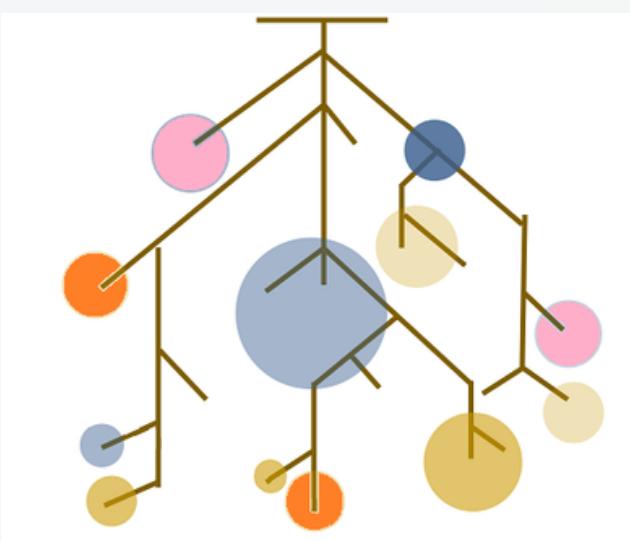
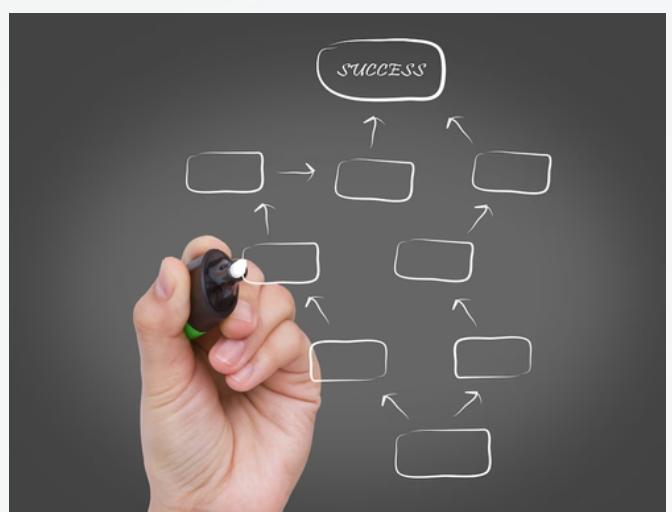


GRADIENT
BOOST

05



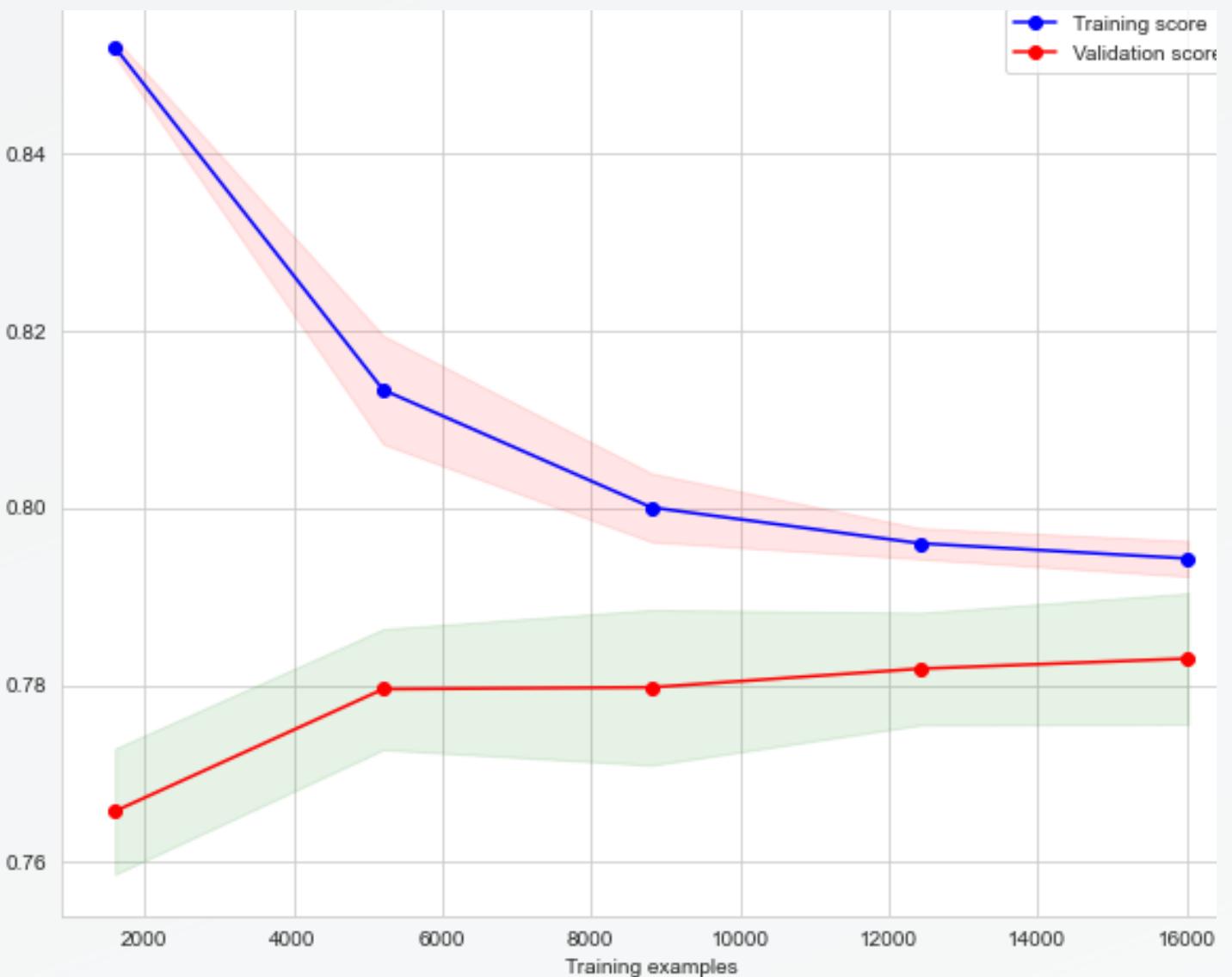
XGBOOST



BEST MODEL



GRADIENT BOOSTING



Model	Model_Train_Accuracy	Model_Test_Accuracy	CV_Accuracy	Testing_time_sec	log_loss
0 Decision_tree		1.000	0.686	0.686	5.20 10.853
1 Random_forest		0.773	0.771	0.768	35.21 7.920
2 Gradient_boosting		0.793	0.787	0.783	37.75 7.340
3 AdaBoosting	0.775	0.780	0.776	9.53	7.604
4 XGBoost	0.779	0.781	0.773	15.15	7.578

CONCLUSION

The dataset provides intriguing insights, its biases, particularly the overrepresentation of certain demographics, need to be considered.



The dataset used for analysis exhibits inherent bias, with 80% of respondents being white individuals and a high representation of college graduates, employed individuals, and a specific income range.



To obtain more comprehensive and representative insights, it is essential to strive for a more balanced dataset.



RECOMMENDATIONS

STRATEGY N°1

Improve on Data Collection



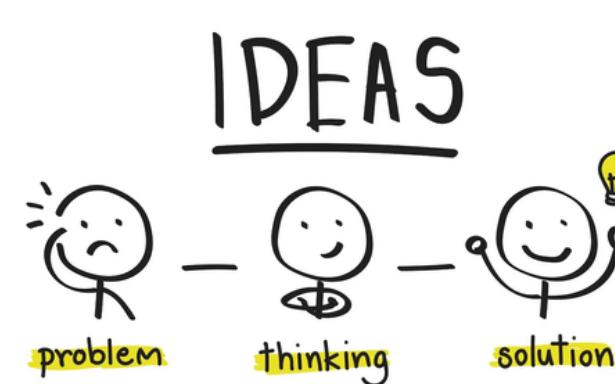
STRATEGY N°2

Mitigate Bias in model training



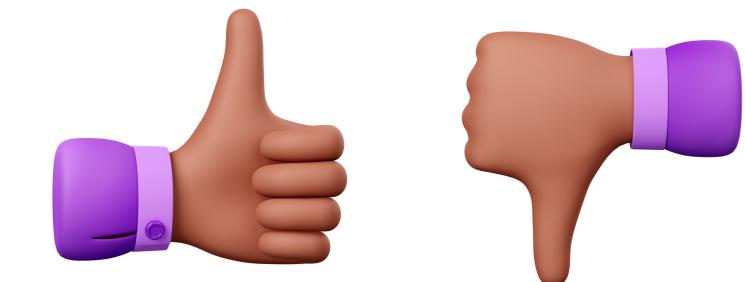
STRATEGY N°3

Evaluate Feature Importance and Interpretability



STRATEGY N°4

Validate Findings with External Data



STRATEGY N°5

Promote Transparency and Documentation



