

HEADED FOR A BREAKUP

Predicting the winning day of the Nenana Ice Classic

Presented By David Walkup

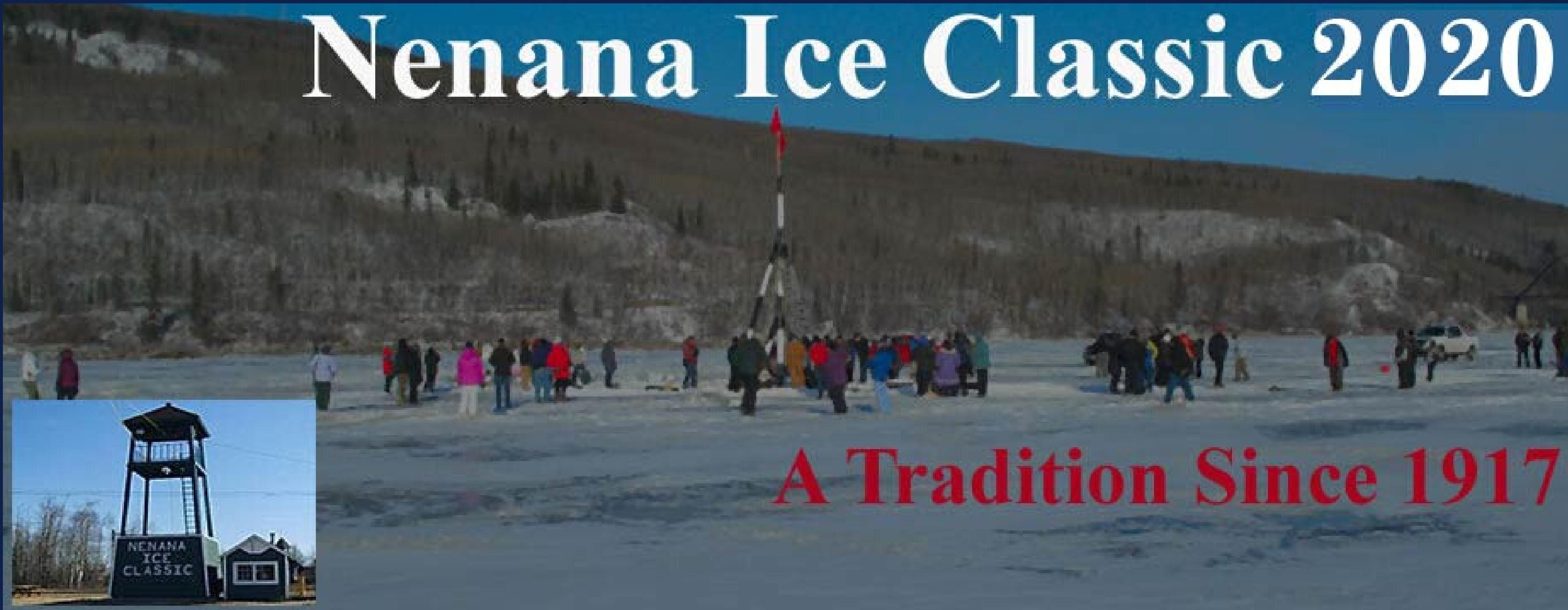
PRESENTATION AGENDA

- What Is The Nenana Ice Classic?
- Problem Definition
- Modeling Approach
- Modeling Results
- Going Forward



WHAT IS THE NENANA ICE CLASSIC?

Nenana Ice Classic 2020



A Tradition Since 1917

"The Nenana Ice classic is a non-profit corporation formed by the residents of Nenana, Alaska, for the purpose of conducting an annual ice pool contest under special statutory authority enacted the first state legislature in 1960."

-- from the Nenana Ice Classic website

<https://www.nenanaakiceclassic.com/organization.htm>



SOME BACKGROUND

- Started in 1917 with AK Railroad surveyors
 - First jackpot was \$800.00
 - Since 2009, jackpot average > \$300,000
 - Average winning payout ~\$116,000
 - Minimum payout ~\$6,400
 - Maximum payout = \$350,000
 - Ice generally breaks between April 20 and May 20

How It WORKS

- 26' "tripod" erected first weekend of March.
- Tripod connected to clock on shore by tripwire.
- When tripod moves 100' down river, tripwire trips, setting winning time.
- Ticket holders closest to month, day, hour, and minute of winning time split jackpot.
- Tickets must be purchased in Alaska.



By Darren Giles - Own work, CC BY-SA 3.0,
<https://commons.wikimedia.org/w/index.php?curid=27753734>

DEFINING THE PROBLEM

"What day and time will the ice break in the spring on the Tanana River by Nenana, AK?"

Answering this question is complex for a couple of reasons:

Most of the data used is weather information.

Very small number of winning days relative to the total number of days.

SO, I chose to simplify the question to:

"What day will the ice break in the spring on the Tanana River by Nenana, AK?"



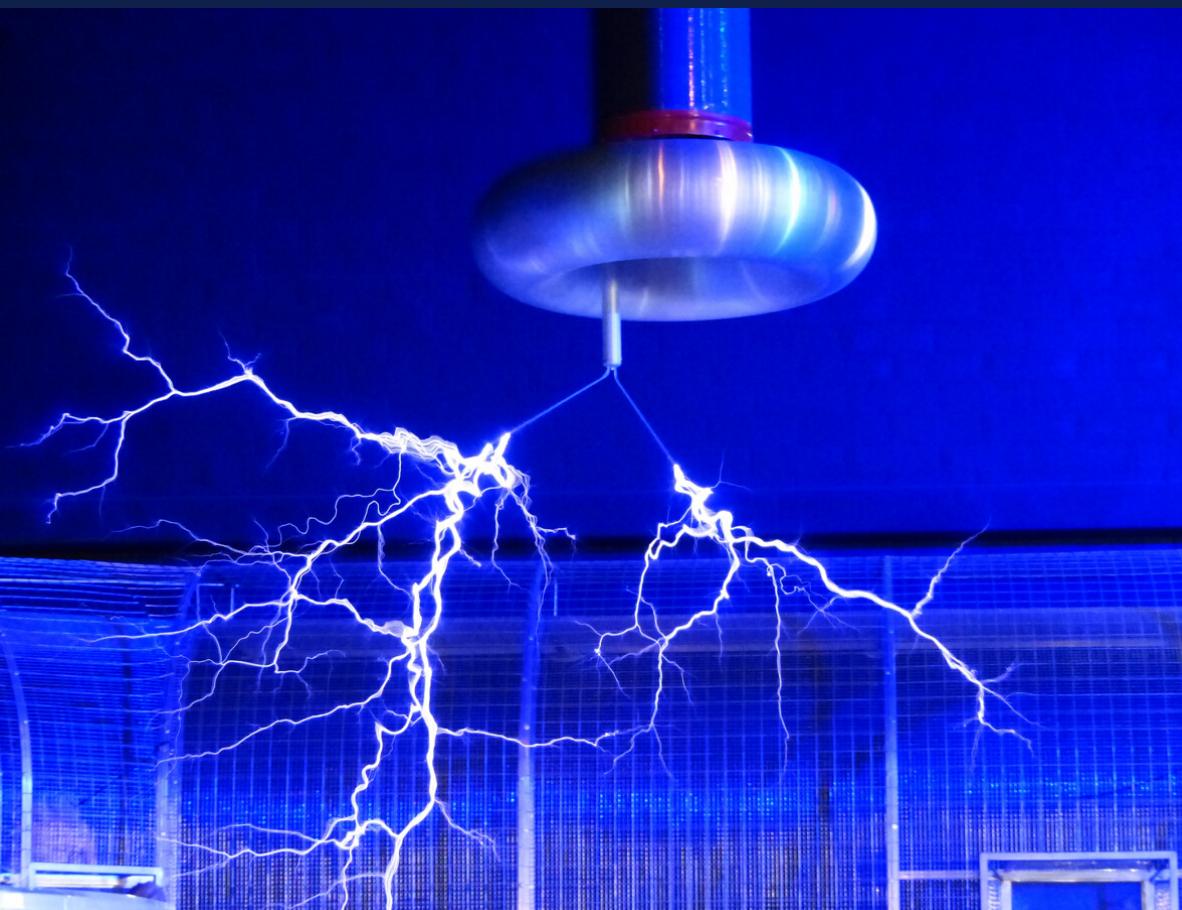
GATHERING DATA

What Information Did I Think Would Be Useful?

- What was the weather like?
- How thick was the ice?
- When did the ice break in the past?
- How does the ice break?

My MODELING APPROACH

Decide on a metric



Perform baseline experiment

Data selection and feature engineering

Perform additional experiments with different algorithms and groups of data

Refine the best model



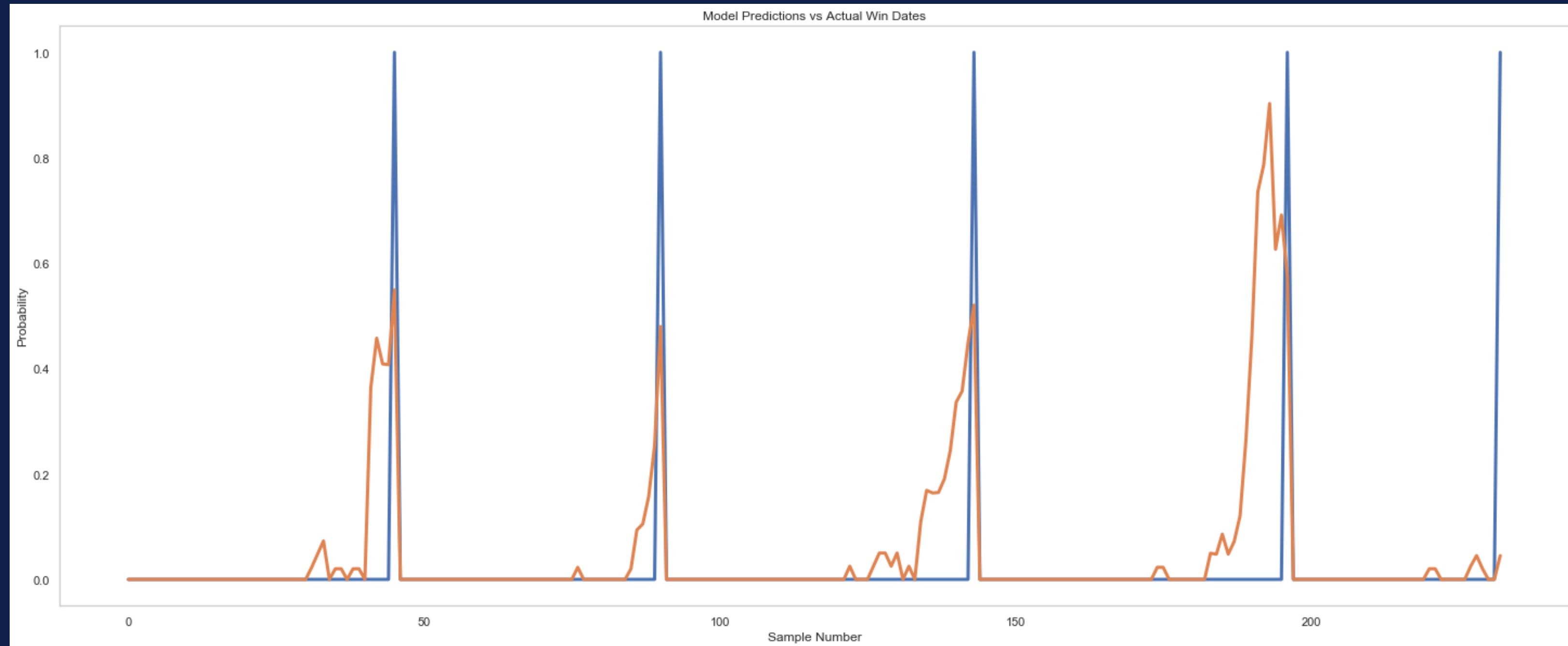
RESULTS

Naive Approach
98.27% Accuracy



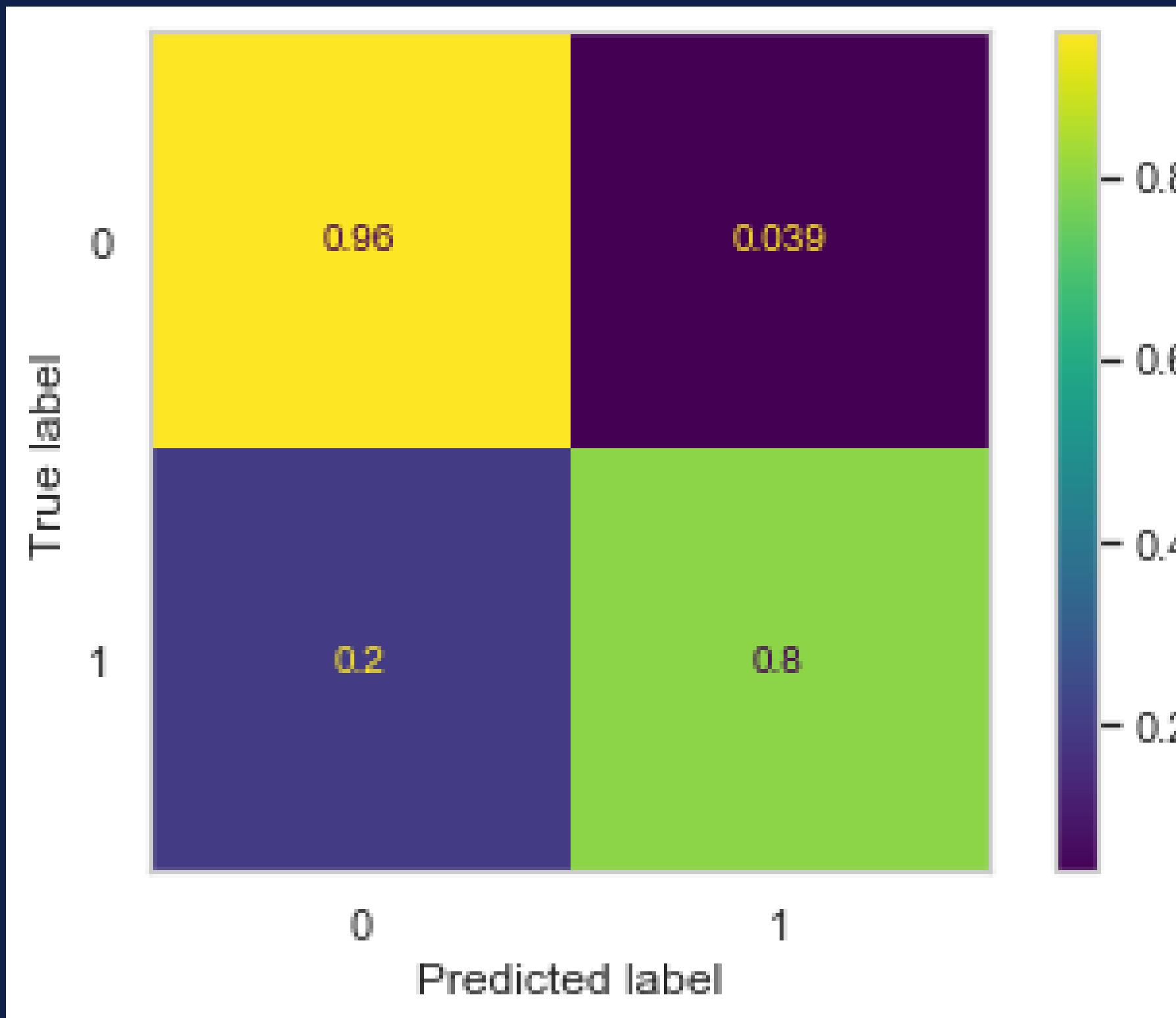
Model Performance
95.71% Accuracy

... BUT WAIT! THERE'S MORE!



The model is picking up the signal!

... BUT WAIT! THERE'S MORE!



My model successfully predicted 80%
(4 out of 5) of the winning days in the
testing set of data.

GOING FORWARD



Gather more data
Additional tuning of current model
Other models