FRAUDULENT
JOB
DESCRIPTIONS

by Imani Thompson

OUTLINE

1

BUSINESS
UNDERSTANDING

2

DATA
UNDERSTADNING

3

DATA
PREPARATION

4

RESULTS

5

RECOMMENDATIONS

6

FUTURE PLANS

BUSINESS UNDERSTANDING

- Working with Indeed
- Experiencing high fraudulent jobs rate

DATA UNDERSTANDING

- Obtained from Kaggle

- 18,000 rows

- Contains textual information
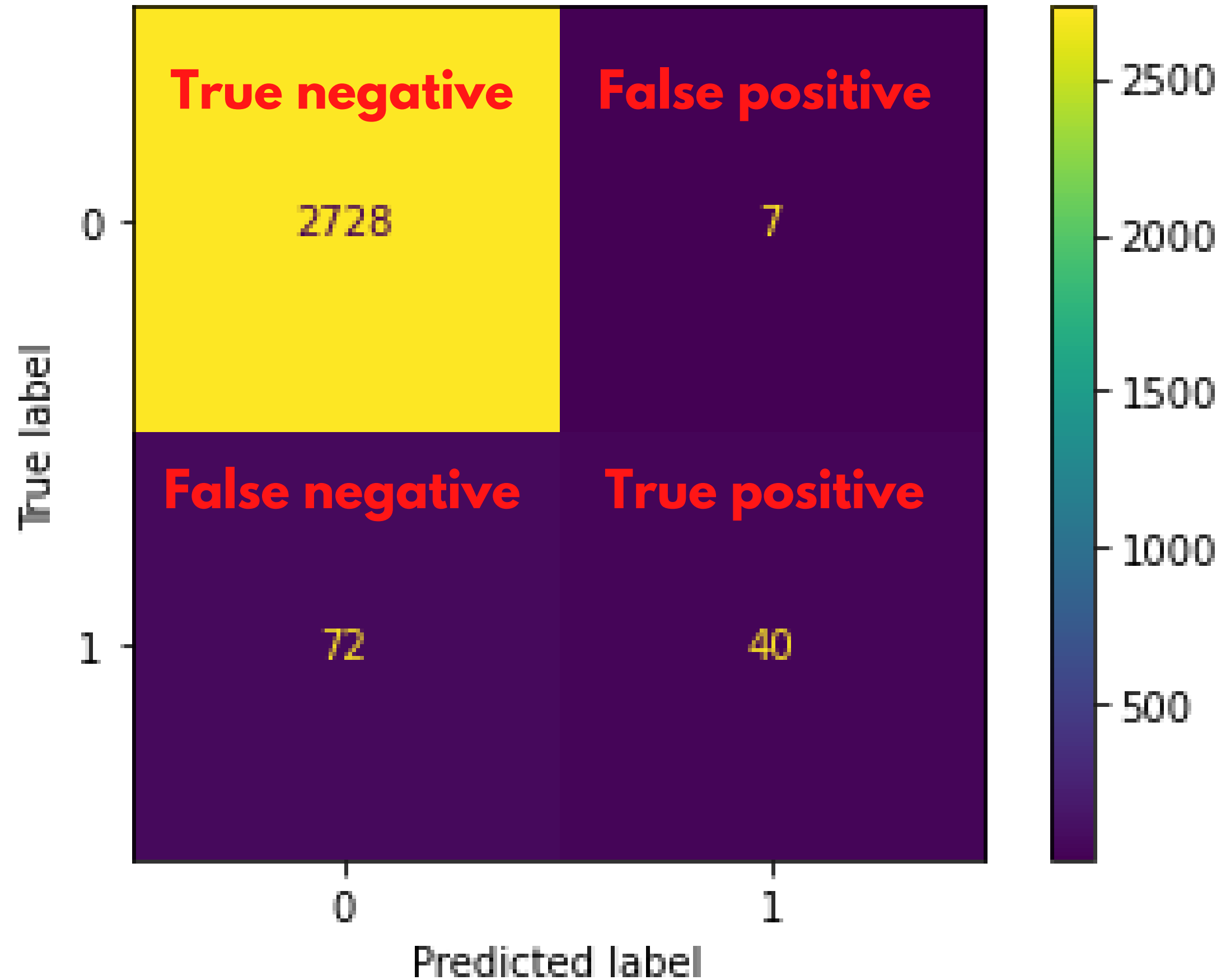
- Target variable fraudulent column

DATA PREPARATION

- Focus on description and requirement columns
- Data was very imbalanced
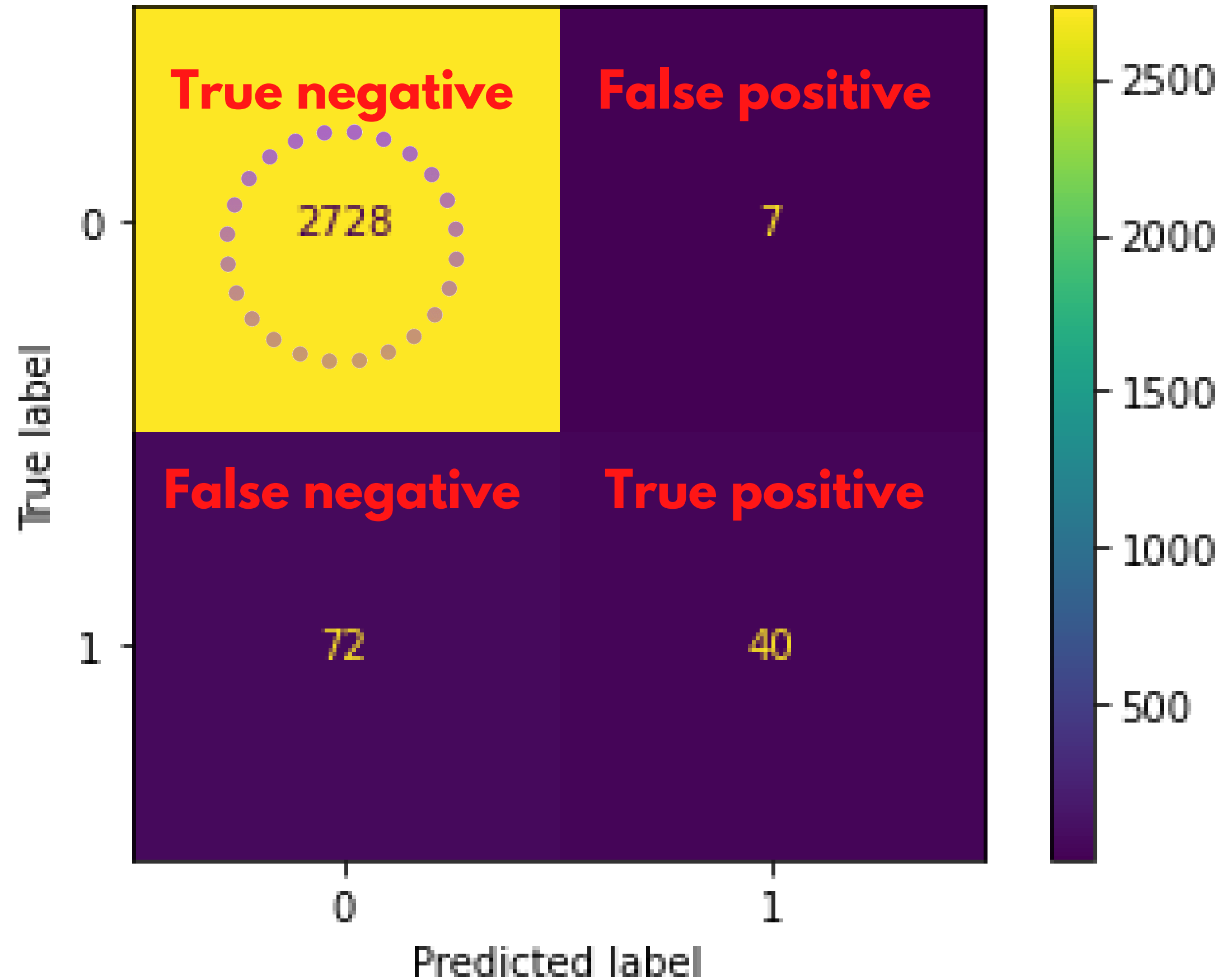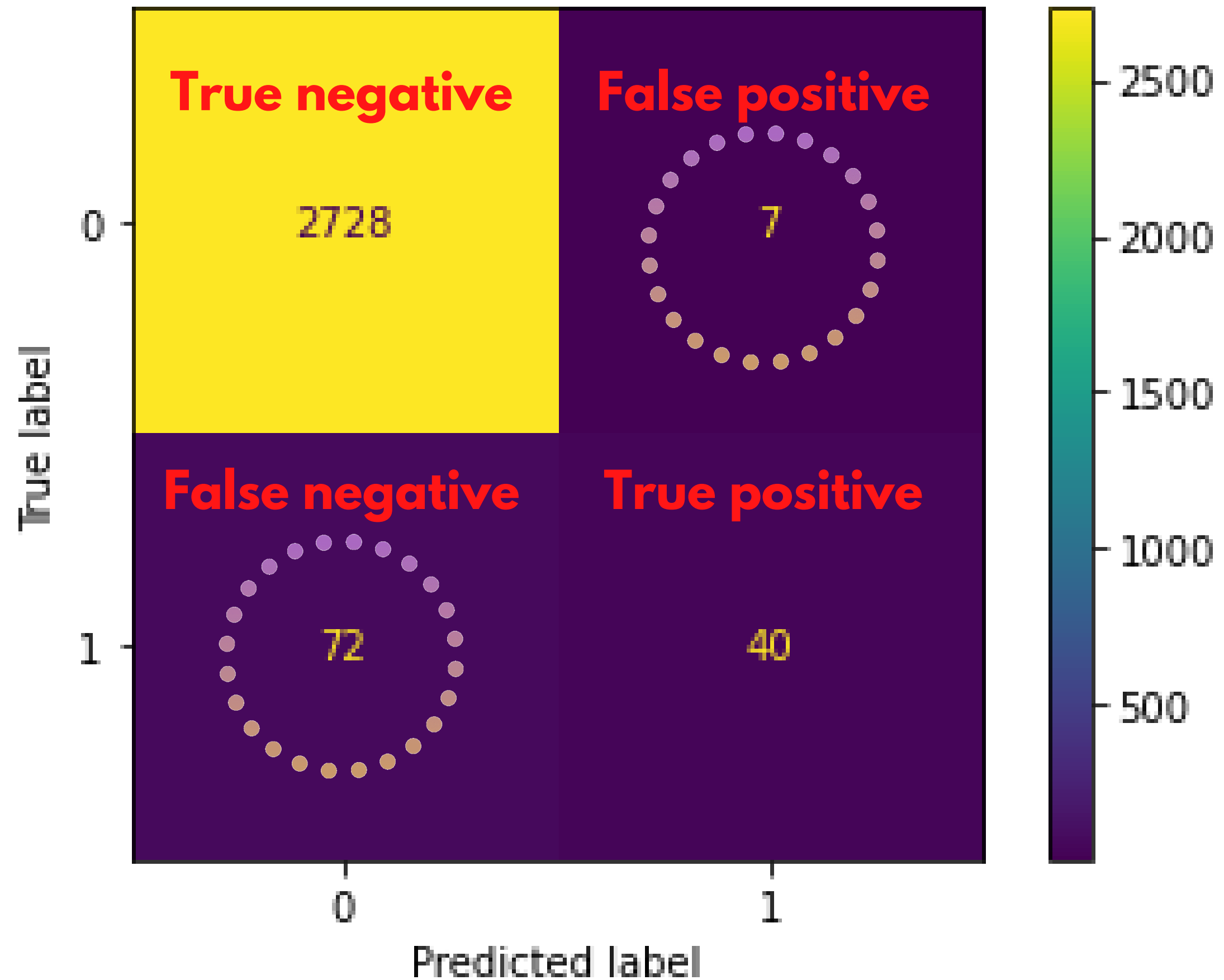- Predict not fraudulent 95% of time

RESULTS

- Multinomial Naive Bayes
- 97% accurate
- True negative 2728
- False positive 7
- False negative 72

RESULTS

- Multinomial Naive Bayes
- 97% accurate
- True negative 2728
- False positive 7
- False negative 72

RESULTS

- Multinomial Naive Bayes
- 97% accurate
- True negative 2728
- False positive 7
- False negative 72

RECOMMENDATIONS

- Indeed focus job description and requirements
- Using my website

NEXT STEPS

More text cleaning

Modeling with Indeed data

Increase true positives

Fraudulent Job Detection!

Is It Fake?

Enter Job Description

Our full-time, project-based data science program covers the full pipeline (planning, acquisition,
preparation, exploration, modeling, delivery), with a focus on Python, SQL, ML, NLP, Distributed
Data, and Data Storytelling. We teach students how to think like a seasoned data scientist, what
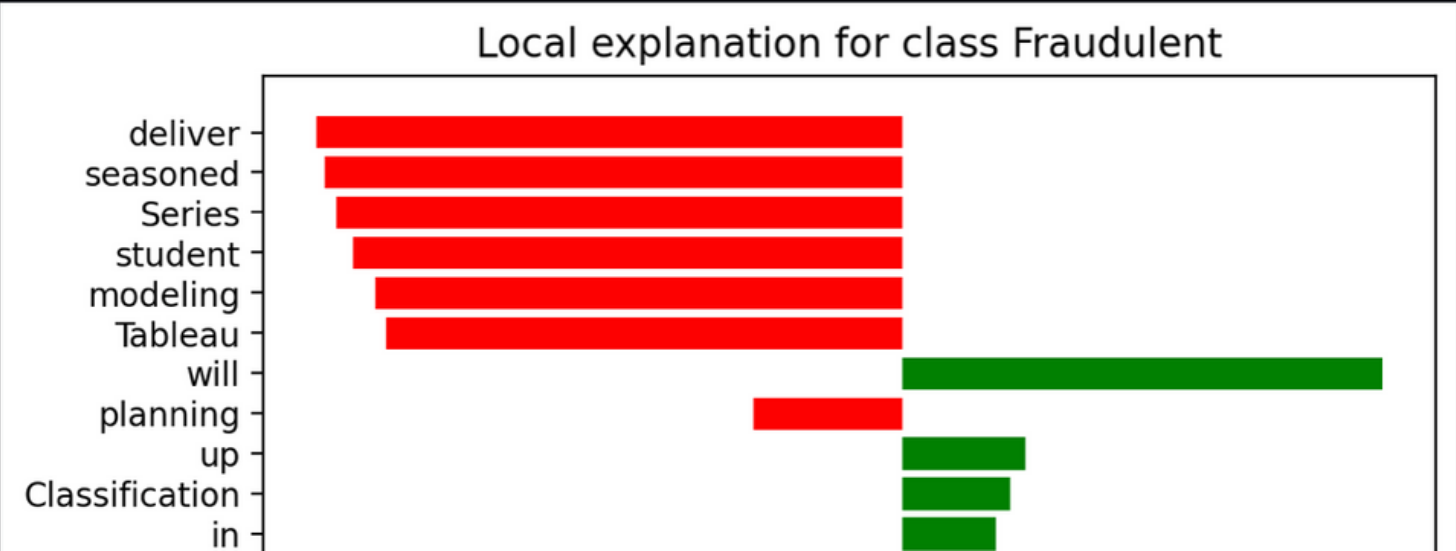
Enter Job Requirements

2+ years industry experience in data science
Fluency in Python, SQL, Data Visualization (matplotlib, seaborn, Tableau), and Machine Learning
(Regression, Classification, Clustering, Time Series Analysis, Anomaly Detection)

	words
0	Our full-time, project-based data science program cov…

Generate

The job is not predicted to be fraudulent

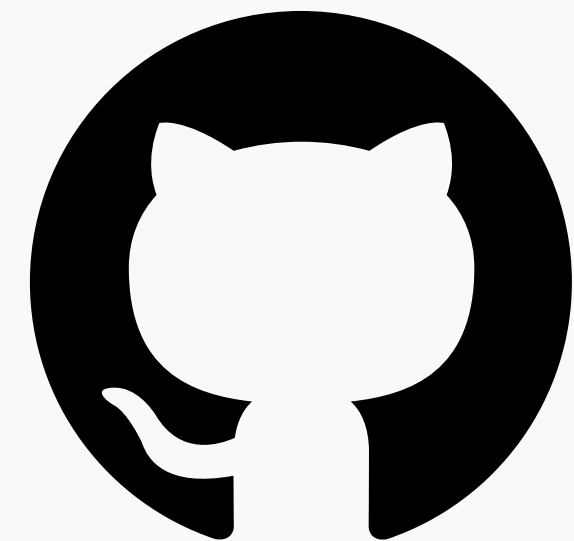Predicted probability to not be fraudulent: 4.73 %

Local explanation for class Fraudulent

THANK YOU!

ijthomps@bsc.edu

https://github.com/ImaniT12