COVID-19 SYMPTOM PREDICTOR MODEL

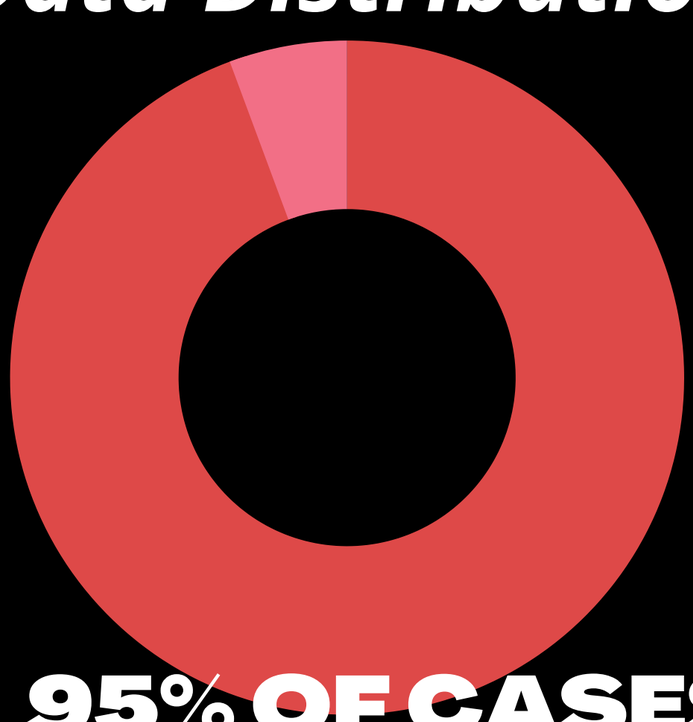
HOW RELIABLE IS YOUR MODEL?

Many of the machine learning models are advertised as having high accuracies. But, does it mean the model's prediction is always reliable? No. Such model reliability can be assessed thoroughly via Precision, Recall, F-1 score measurements along with the model accuracies.

ID3: Iterative Dichotomiser Algorithm

Israel's 2020 Corona Test Dataset over 200000 individuals

Data Distribution



95% OF CASES

are **NEGATIVE**

Accuracies

original dataset:	***96 %***
1:3 (positive:negative):	***89 %***
1:1 (positive:negative):	***50 %***

- *Modified dataset's proportions based on positive test cases.*

Model Features : Common COVID-19 Symptoms

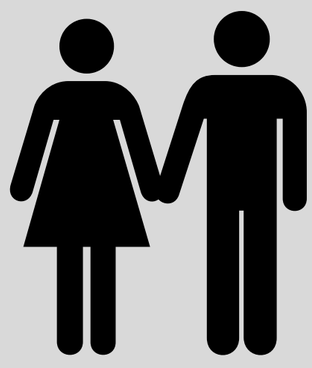
-Cough-

-Fever-

-Sore Throat-

-Headache-

-Breadth Shortness-

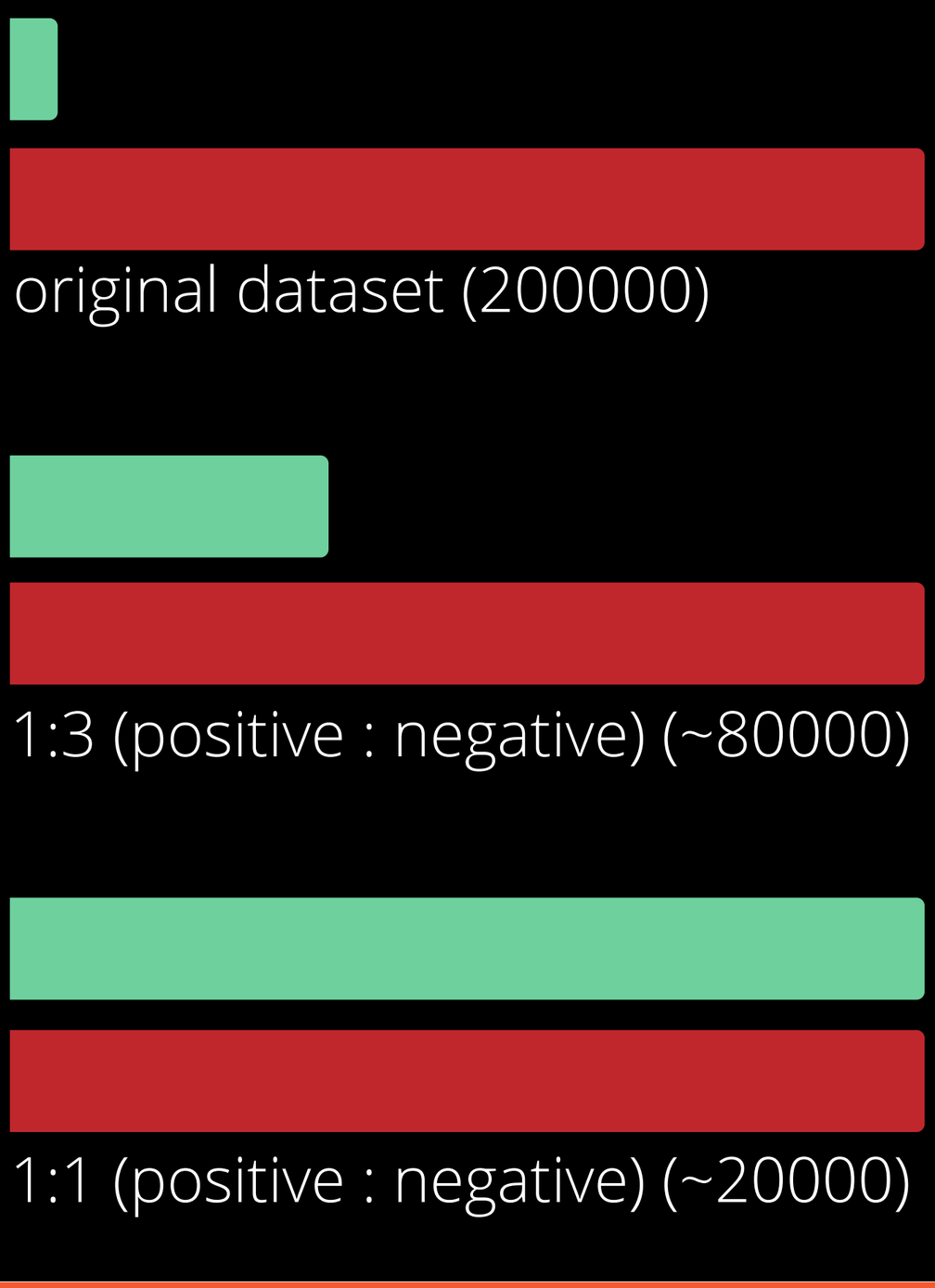


-Gender-

-Age 60 Over-

-Indication-

Data Distribution



Evaluation Metrics

F-1 Score:	***67 %***
Precision:	***82 %***
Recall:	***56 %***

F-1 Score:	***77 %***
Precision:	***88 %***
Recall:	***68 %***

F-1 Score:	***0 %***
Precision:	***0 %***
Recall:	***0 %***

CONCLUSION

Opposed from the accuracies, F-1 score measurement offers more dynamic analysis towards assessing the model performance.

Good models hit above F-1 scores 80 percent.

Reliablility cannot be soley determined by the model's accuracy.

SOURCES

https://www.researchgate.net/publication/350640627
https://medium.com/analytics-vidhya/accuracy-vs-f1-score-6258237beca2
https://github.com/yeonhak-kim/machine-learning/blob/main/covid19_classification_final.ipynb