Interpretable Machine Learning for COVID-19:

An Empirical Study on Severity Prediction Task
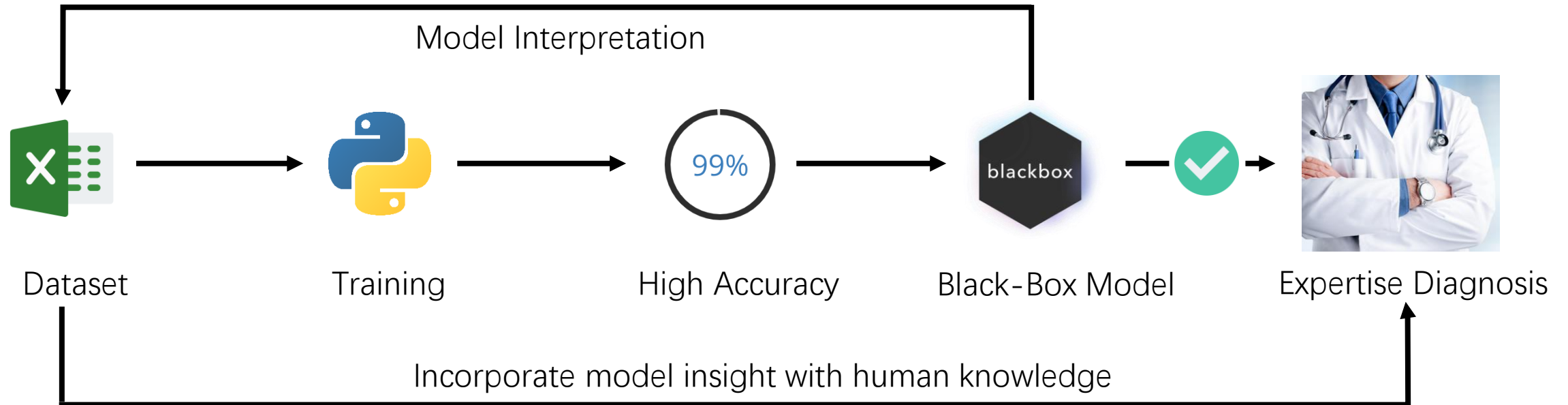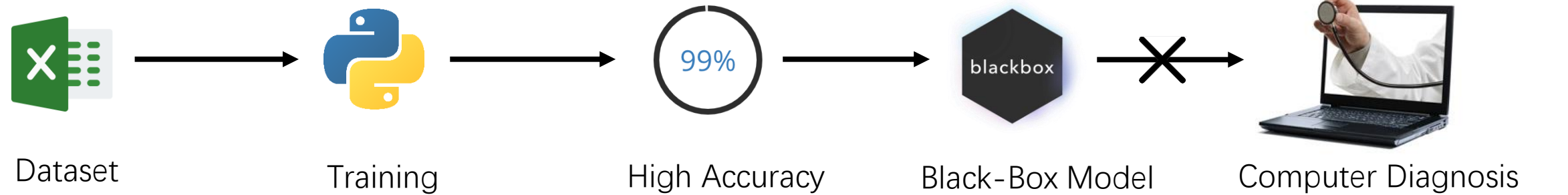
Confirmed Cases :   83,107,700
Global Population: 7800,000,000
                                1 / 100

Background

The usual workflow of machine learning.

Who is responsible for the 1% error, the laptop?

Dataset → Training → High Accuracy (99%) → Black-Box Model ✗→ Computer Diagnosis

Model Interpretation

Dataset → Training → High Accuracy (99%) → Black-Box Model → ✓ → Expertise Diagnosis

Incorporate model insight with human knowledge

Dataset & Models

Dataset Description

Source Code

Preprocessing

1. Remove clinical irrelevant columns ['MedNum', 'No']
2. Remove columns that have no data ['LVEF', 'SO2', 'PO2', 'YHZS', 'RML', 'RUL', 'RLL', 'LUL', 'LLL']
3. Remove columns with fewer records ['Onset2Admi', 'Onset2CT1', 'Onset2CTPositive1', 'Onset2CTPeak']
4. Remove patients that have incomplete records. [Height == "", cTnI == ""]

Train Test Split (86 patients, 21 severe, 65 normal, with 55 features)

1. 90% as Training Set, 10% as Testing Set
2. Five folds Cross Validation

In this research, we focus on model **interpretation**. **Interpretation** gives us valuable insights even if these are not 99% perfect models.

Classifier	Cross Validation	Test Set			95% confidence-interval
-------------------------------------	F1	Precision	Recall	F1	---------------------------------
Decision Tree	0.56	0.67	0.50	0.57	0.307
Random Forest	0.64	0.56	0.25	0.33	0.324
XGBoost	0.62	0.78	1.00	0.80	0.271
Neural Network	0.53	0.78	1.00	0.80	0.271

Patient No.7 - Severe

Feature	Value
LYM2	0.46
CRP2	22.20
Phlegm=False	True
ALB2	35.30
NauseaNVomit=False	True
Fatigue=False	True
CKMB	8.30
cTnI	0.01
AST	19.70
Arrythmia=False	True

Feature	Patients Data	Normal Range	Conclusion
LYM <= 0.82	0.46	(1.1, 3.2)	↓
CRP (7.00, 33.25)	22.20	(0, 5)	↑
No Phlegm	True	-	No Symptom
ALB <= 35.50	35.30	(35, 55)	Normal
No Nausea and Vomit	True	-	No Symptom
CKMB <= 10.30	8.30	(0, 18)	Normal
AST (19.30, 25.70)	19.70	(29, 35)	↓

What is the role of T cells in COVID-19 infection? Oxford COVID-19 Evidence Service Team

Prediction probabilities

normal 0.02
severe 0.98
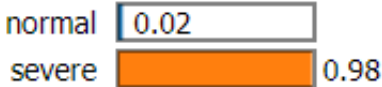
No Phlegm, decrease of the number of lymphocytes in the blood is more likely to turn severe?

This may seem counterintuitive, but corroborated by medical research months later after the outbreak of the pandemic.

Patient No.3 - Normal

Feature	Value
NTproBNP	475.00
LYM2	0.81
CRP2	78.76
Phlegm=False	True
ALB2	37.60
SoreThroat=False	True
CKMB	17.90
PCT2	0.00
LDH	263.00
Diarrhea=False	True

Feature	Patients Data	Normal Range	Conclusion
NTproBNP > 338.00	475.00	(0, 300)	↑
LYM <= 0.82	0.81	(1.1, 3.2)	↓
CRP > 33.25	78.76	(0, 5)	↑
No Phlegm	True	-	
PCT	0.00	(0, 0.5)	Normal
LDH	263.00	(109, 245)	↑

Prediction probabilities

normal 0.42
severe 0.58

A normal patient is diagnosed as sever by our models.

But actually he is indeed severe even diagnosed by doctors.

Z. Xu, L. Shi, Y. Wang, J. Zhang, L. and etc. "Pathological findings of COVID-19 associated with acute respiratory distress syndrome," Lancet Respir Med, vol. 8, no. 4, pp. 420–422, 04 2020

Conclusion

> Elderly people are more likely to turn severe.

The average age of severity level [0 – 3]: 36.8 --> 47.5 --> 54.3 --> 69.4

Interpretable Machine Learning

- Inflammation Indicators： High CRP (Virus Infection)
- Blood Routing Examination： Low LYM
- Biomarker： High LDH, Low ALB

- Symptoms： No Phlegm， No diarrhea