R2N

Oximetry reading bias solution

deepLearners

SaO2 vs SpO2

01

SaO2

is **measure of blood oxygen saturation**. Understanding it is crucial in identifying potential underlying respiratory illnesses or more fatally the likelihood of tissue hypoxia (insufficient oxygen to body tissues for it to function).

02

SpO2

is an **estimate** of SaO2. Taken by a pulse oximeter, which is less invasive of a testing method than conducting an ABG test required for SaO2. **However, taking SpO2 from pulse oximeters are less accurate than conventional ABG tests**.

THE PROBLEM

HOW OXIMETERS WORK

Oximeters work by transmitting light through your finger or earlobe and measuring how much light passes through. Having darker skin leads to more light being absorbed, indicating a higher oxygen saturation.

BIAS THAT OCCURS

Devices overestimate oxygen levels in patients with darker skin tones, causing ‘false negatives’; where patients who have low oxygen saturation are reported to have normal saturation levels.

Factors other than light such as excessive movement could also affect the readings. This means that patients who are older, or who suffer from brain disorders such as Parkinson’s disease, are more likely to receive unreliable readings.

OUR AIM

To improve the accuracy of pulse oximeter readings (SpO2) by compensating for the distortions due to skin tone.

BENEFITS

1. Better SpO2 reading accuracy = less reliance on invasive ABG tests for blood oxygen saturation.

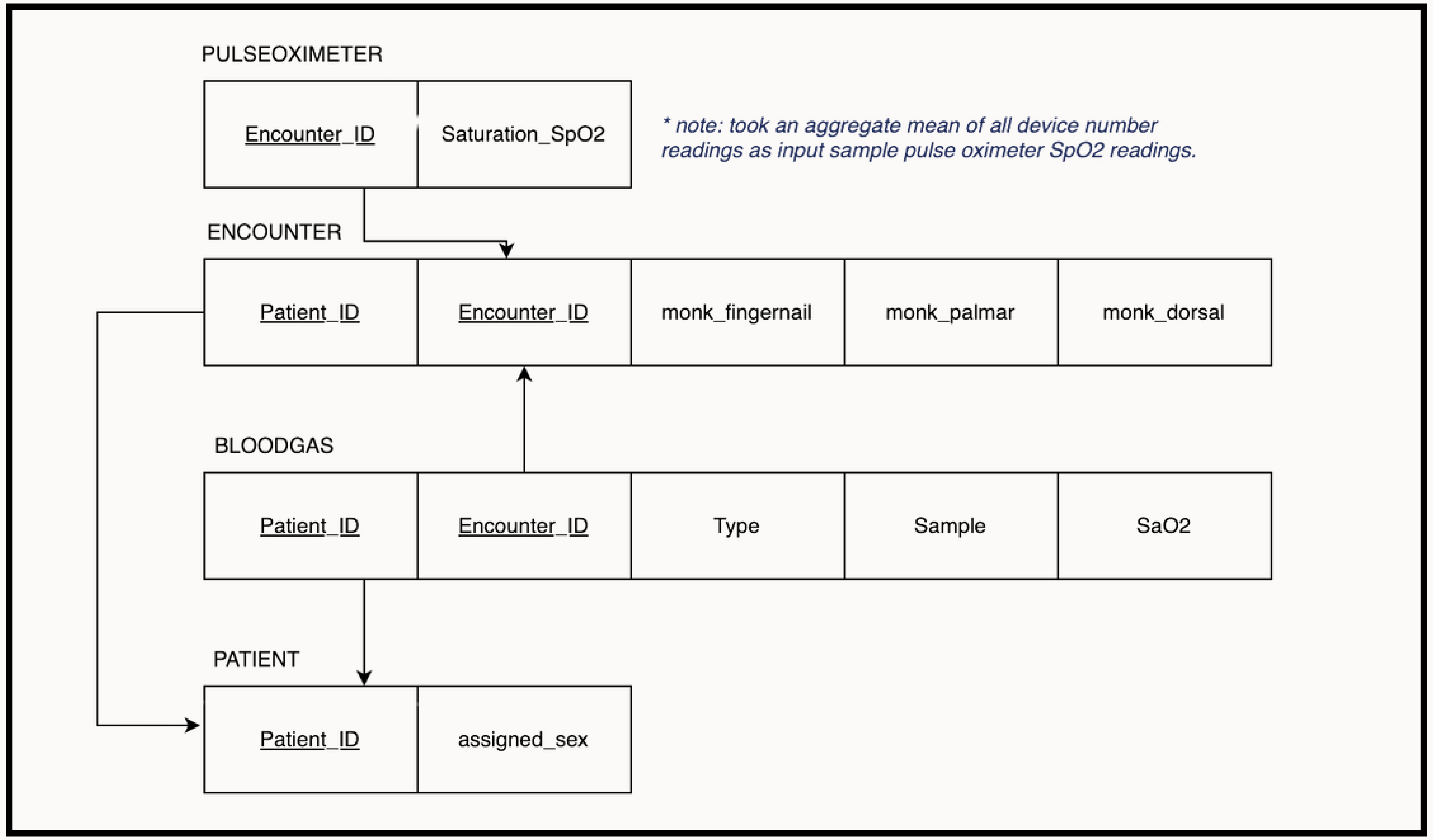
2. Currently, doctors perform clinical judgement in addition to the SpO2 data from pulse oximeters. Improving the accuracy of readings reduces need for human judgement which might not always be correct.

3. Allows for continuous assessment of a patient’s blood oxygen saturation as an advantage over the ABG test.

DATA PREPROCESSING

Data obtained from [OpenOximetry Repository](#).
Files involved in analysis: *bloodgas.csv, encounter.csv, patient.csv, pulseoximeter.csv*

Performed data cleaning and pre-processing to obtain files.
Extraction of relevant fields, with relationships visualized:



Merged Datasets and performed categorical variable encoding.
Used sklearn LabelEncoder to encode categorical variables in alphabetical order.
Columns processed: *monk_fingernail, monk_dorsal, monk_palmar*

DATASET

	sample	SaO2	monk_fingernail	monk_dorsal	monk_palmar	SpO2
count	25234.000000	25234.000000	25234.000000	25234.000000	25234.000000	25234.000000
mean	13.191210	85.482107	3.380518	5.413569	3.856464	86.634411
std	7.484846	9.111963	1.976522	2.149801	1.732978	9.174980
min	1.000000	58.100000	1.000000	1.000000	1.000000	50.666667
25%	7.000000	77.300000	2.000000	4.000000	2.000000	79.000000
50%	13.000000	85.900000	3.000000	5.000000	4.000000	87.500000
75%	19.000000	94.200000	5.000000	7.000000	5.000000	95.220000
max	32.000000	100.700000	8.000000	10.000000	8.000000	144.000000

LIMITATIONS

1. **Lack of robustness in dataset.**
   Other factors (eg. PI readings, tremors, ambient light, skin temperature etc.) can affect readings, and should also be taken into consideration.

2. **Smaller sample size due to inability to access higher-restricted databases (time-sensitivity).**
   Working size, after filtering for complete records: n = 193 patients, with multiple readings at different points in time.

XGB REGRESSOR MODEL

01

CROSS-VALIDATION WITH HYPERPARAMATER TUNING

We search for best hyperparameters using GridSearchCV and evaluate each combination using 5-fold cross-validation. Negative MAE is employed as the scoring metric.

02

REGRESSION

Our output data—predicted SpO_2 (SAO_2) readings—are continuous in nature, meaning they take on a range of values rather than discrete categories. Hence, we chose to adopt a linear regression model.

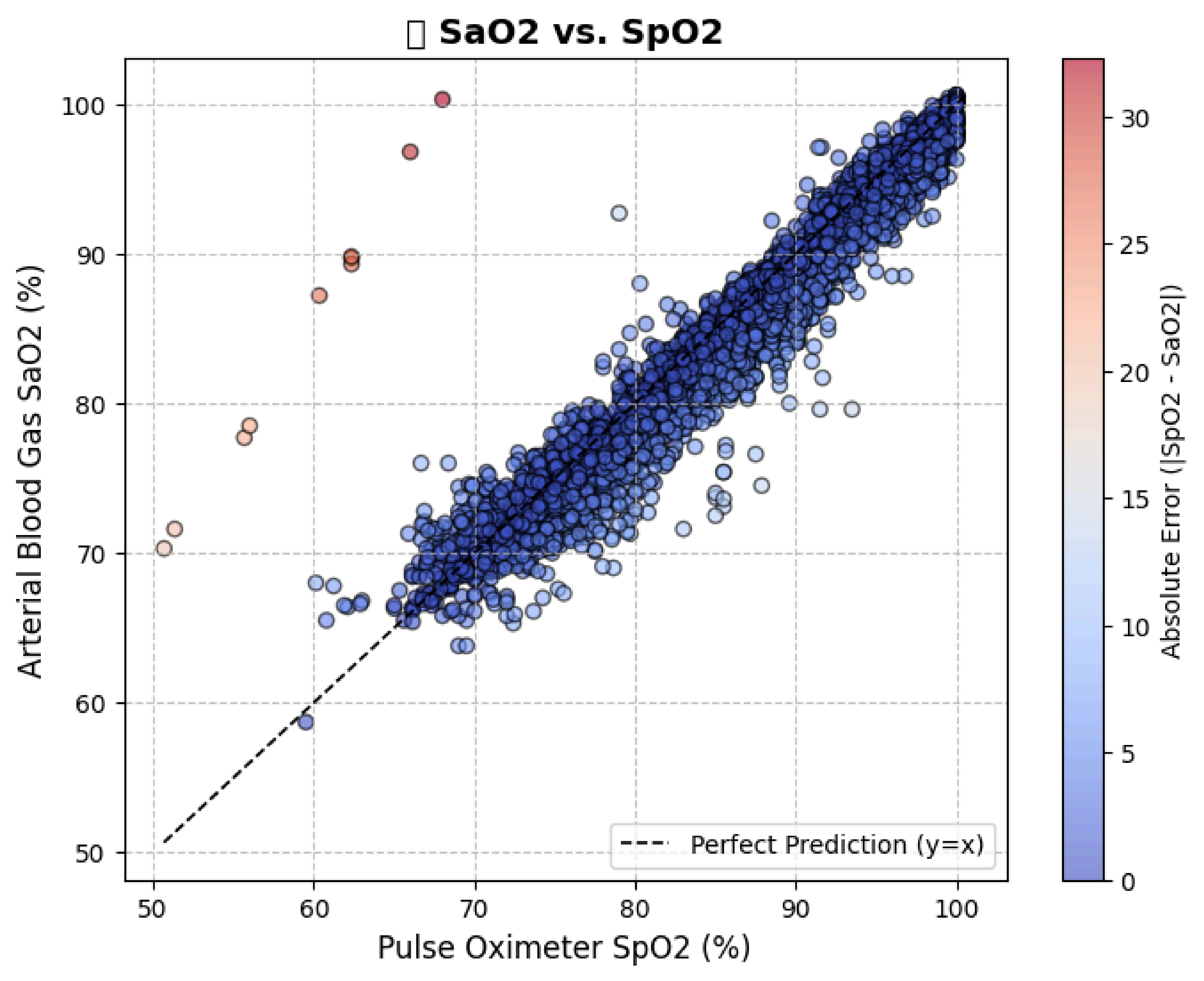
03

EVALUATION

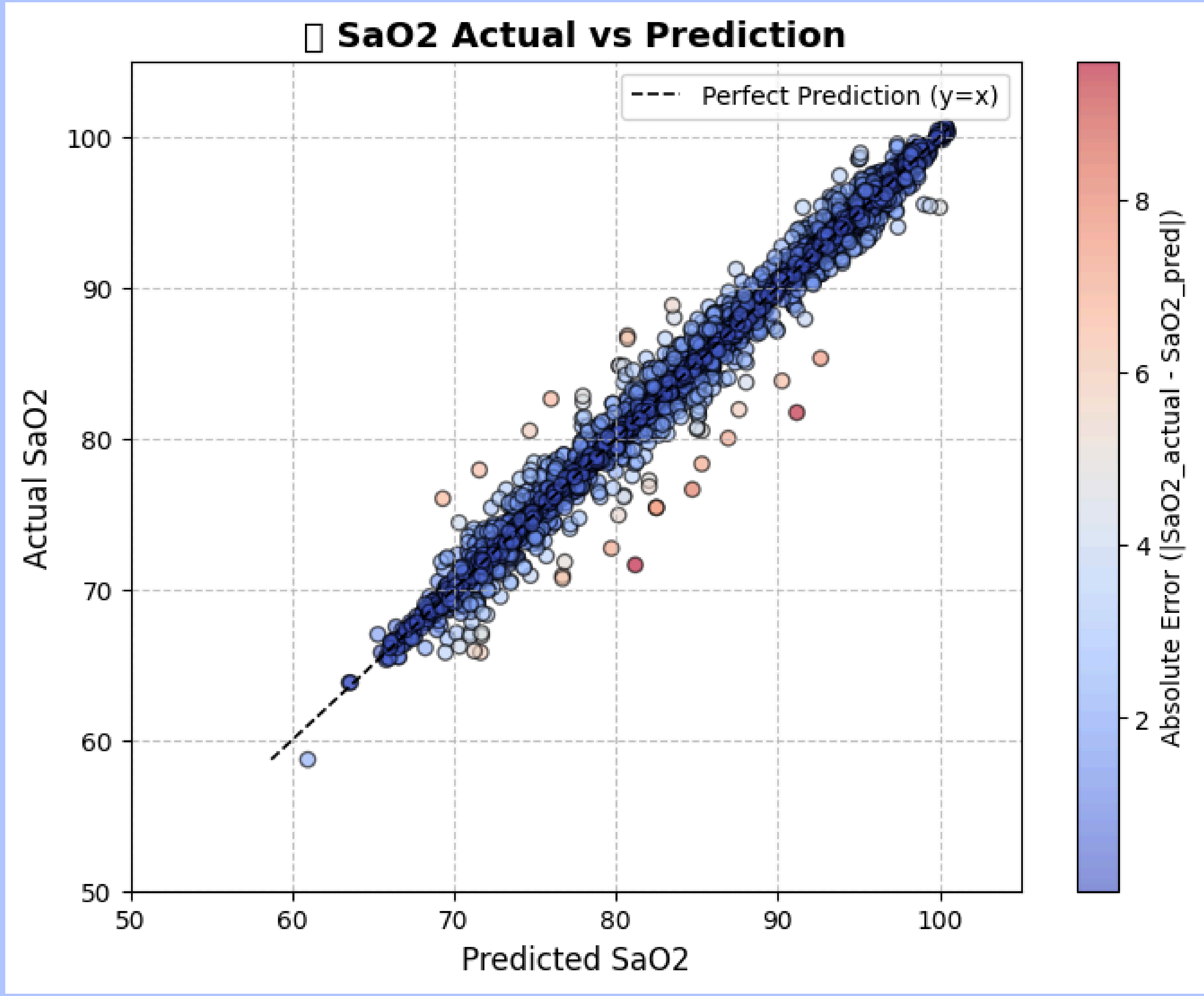
Because our predictions are continuous, appropriate evaluation metrics must be chosen. Instead of classification-based metrics, we focus on:

- Mean Absolute Error (MAE) – measures the average absolute difference between predicted and actual SAO_2 values.
- Mean Squared Error (MSE) – gives more weight to larger errors, useful when minimizing extreme deviations is a priority.
- Root Mean Squared Error (RMSE) – similar to MSE but in the same unit as the original data, making interpretation easier.
- **R^2** (Coefficient of Determination) – evaluates how well the predicted values explain variability in actual SAO_2 readings.

RESULTS



□ SaO2 vs. SpO2
Arterial Blood Gas SaO2 (%)
Pulse Oximeter SpO2 (%)
Absolute Error (|SpO2 - SaO2|)
Perfect Prediction (y=x)
50
60
70
80
90
100
0
5
10
15
20
25
30



□ SaO2 Actual vs Prediction
Actual SaO2
Predicted SaO2
Absolute Error (|SaO2_actual - SaO2_pred|)
Perfect Prediction (y=x)
50
60
70
80
90
100
2
4
6
8

OUR HOPE

By proving that it is indeed possible to compensate for the distortions in SpO2 readings of pulse oximeters, we can further improve its reliability by creating a more robust model that can take into consideration other significant factors, such as ambient light interference and vibrations.

REFERENCES

Dataset from: Fong, N., Lipnick, M., Bickler, P., Feiner, J., & Law, T. (2025). OpenOximetry Repository (version 1.1.1). PhysioNet. https://doi.org/10.13026/be2e-cn29.

Matos, J., Struja, T., Gallifant, J. et al. BOLD: Blood-gas and Oximetry Linked Dataset. Sci Data 11, 535 (2024). https://doi.org/10.1038/s41597-024-03225-z.

Rauniyar, Nabin & Pujari, Shyam & Shrestha, Pradeep. (2020). Study of Oxygen Saturation by Pulse Oximetry and Arterial Blood Gas in ICU Patients: A Descriptive Cross-sectional Study. Journal of Nepal Medical Association. 58. DOI:10.31729/jnma.5536.