**Explainable AI:**

The explainable AI concept was implemented in project to develop trust in AI model. The AI model before using XAI can be termed as black box.

There are three main parts of explainable ai:

1)Prediction Accuracy

(\*)Lime model is used for adding instance to prediction accuracy

2)Traceability

3)Decision Under

(\*)Applications on Dashboard design for business purpose

Model Investigation:

There are 4 parts of tracking:

(\*)Deployment status

(\*)Fairness

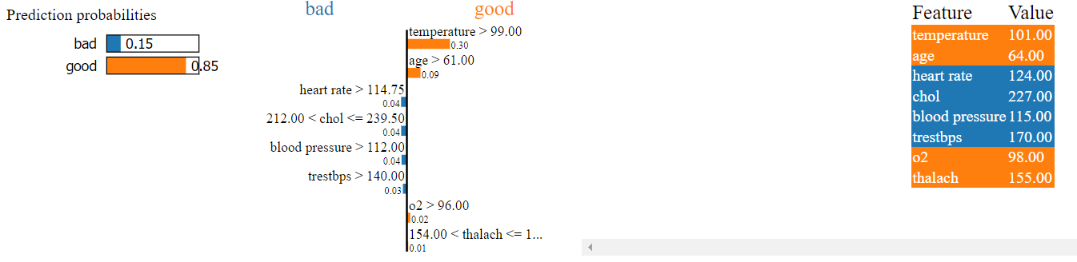
(\*)Quality

(\*)Drift

Implemented models in explainable ai are:

**Lime:**

* **Local Interpretable Model-agnostic Explanations (LIME)** is one of the most popular **Explainable AI (XAI)** methods used for explaining the working of machine learning and deep learning models.
* LIME can provide model-agnostic local explanations for solving both regression and classification problems and it can be applied with both structured datasets and even with unstructured datasets like text and images



Explaination:

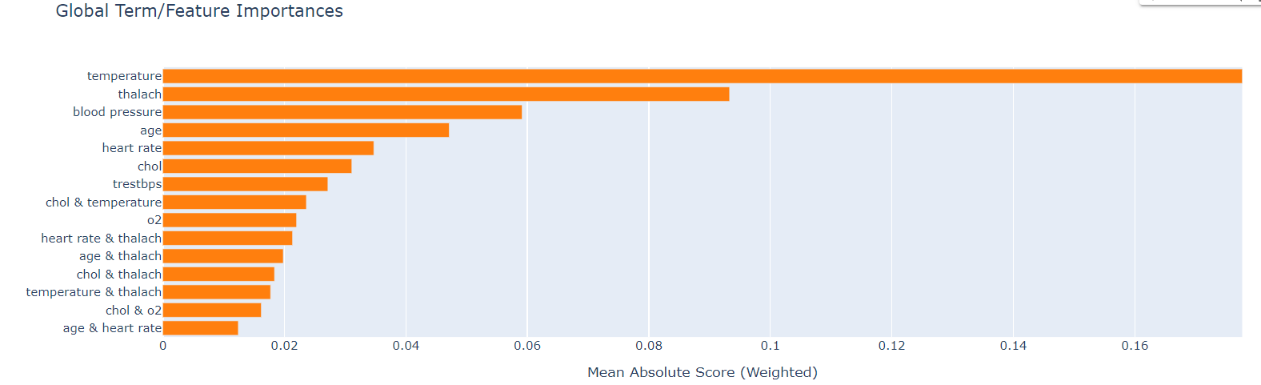
Initially 4 sample entries are imported with random classifier. The plot on the left hand side contains bad and good predictions of imported parameters/features like temperature,age,heart rate,choleastrol,trestbps,blood pressure,o2 and thalch.

The middle plot consist of good and bad interpret of input parameter. The blue color signifies how badly the certain input parameter impact on AI model. The orange color signifies how good or positive impact of input parameter on AI model.

Finally the last plot describes with score/accuracy that states about goodness in our model due to inputted model termed as feature.

Conclusion: Temperature,age,o2 and thalch are good result productive parameters in AI model.

GLOBAL INTERPRET:

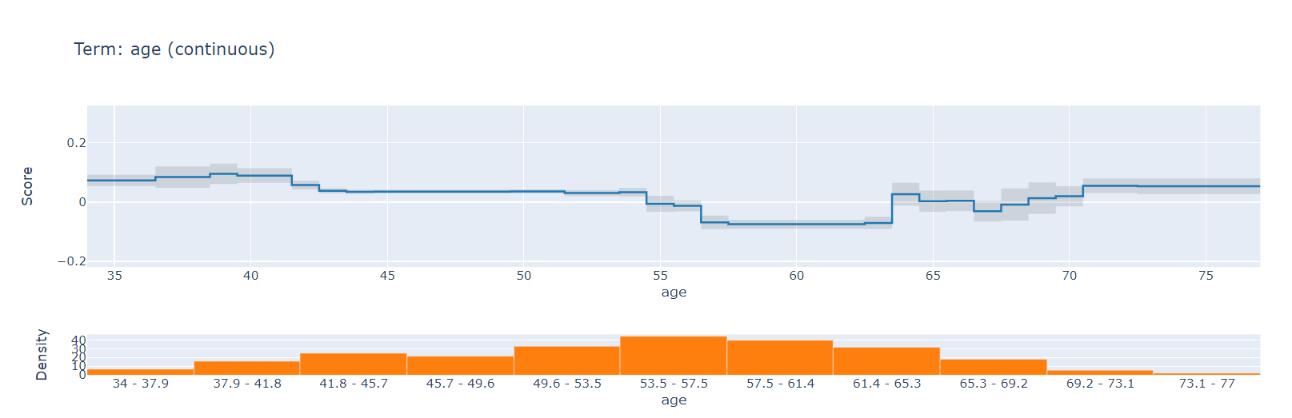


The most common ways of obtaining global interpretation is through:

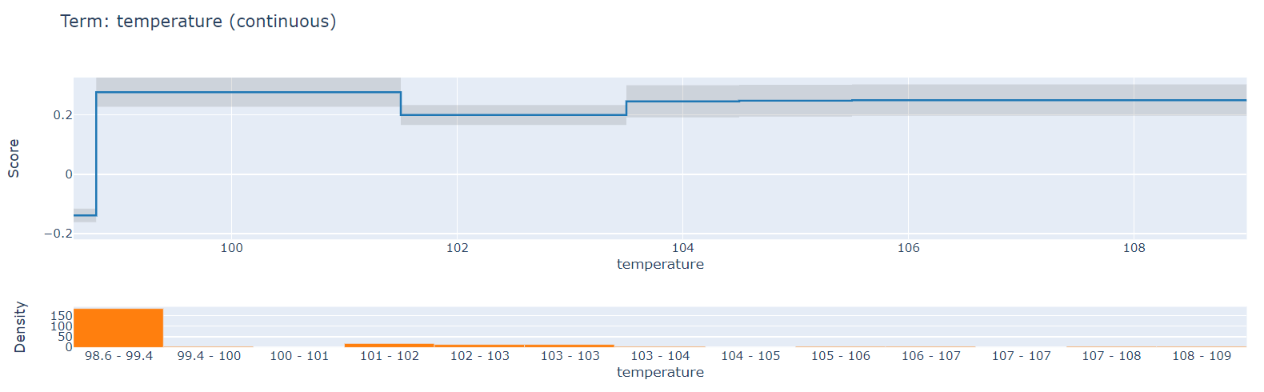
* variable importance measures
* partial dependence plots

Variable importance quantifies the global contribution of each input variable to the predictions of a machine learning model. Variable importance measures rarely give insight into the average direction that a variable affects a response function. They simply state the magnitude of a variable’s relationship with the response as compared to other variables used in the model.

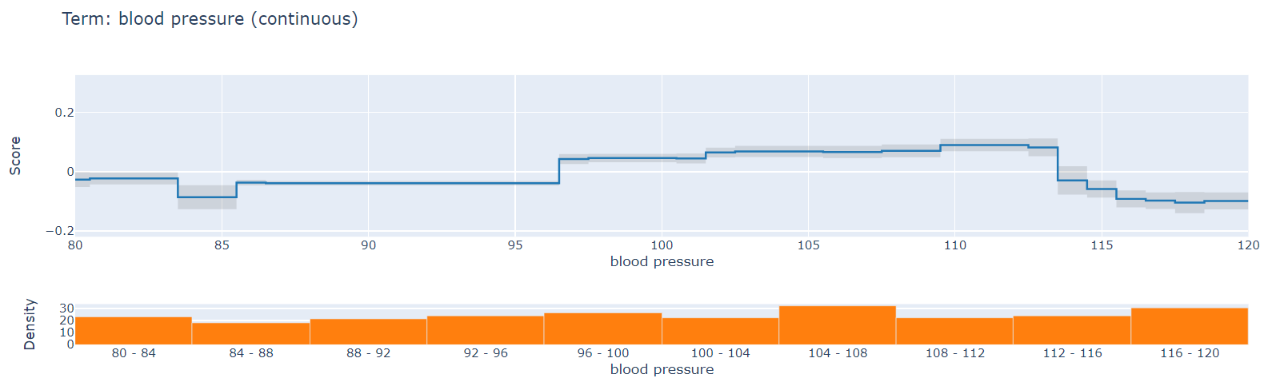
The interpret was made by making input parameter as global features. In this plot,temperature parameter contributing good probability for our model so no doubt in accuracy of temperature parameter and similarly each parameter is compared individually the plotted accordingly.



Explain: The component graph is implemented with all input parameter. The age vs score continuous plot states that range of patient who are normal or abnormal and accuracy /score is computed according to range of ages.

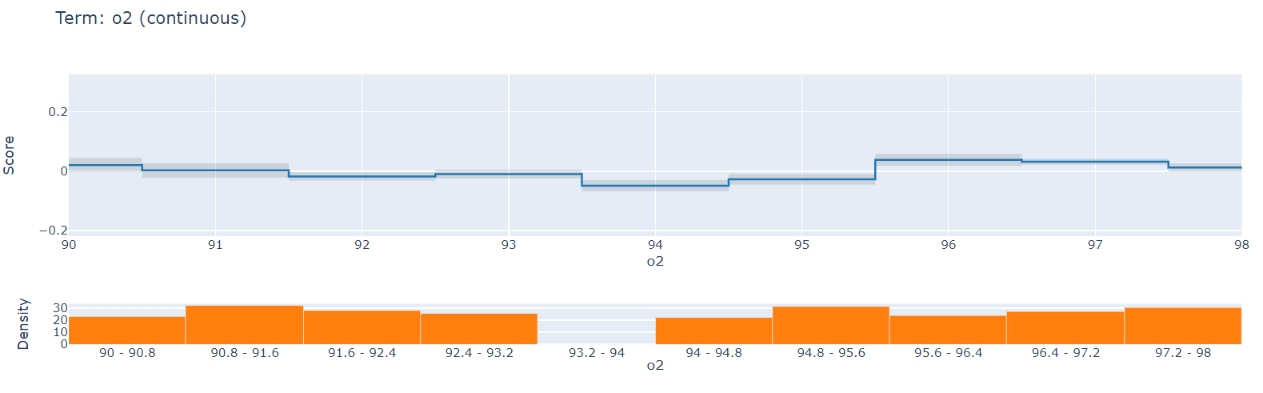


Explain: There are certain range of temperature that states at most temperature suffered by patient and by using this graph doctor can diagnose certain person disease with indication of temperature. by this we can interpret that at what maximum temperature did the patient suffered.



Explain:

The bp vs score graph states the peak of blood pressure measure of a patient and how much accurately it was predicted. The patient can be easily diagnosed by looking to this plot of average raise in range of bp measure.

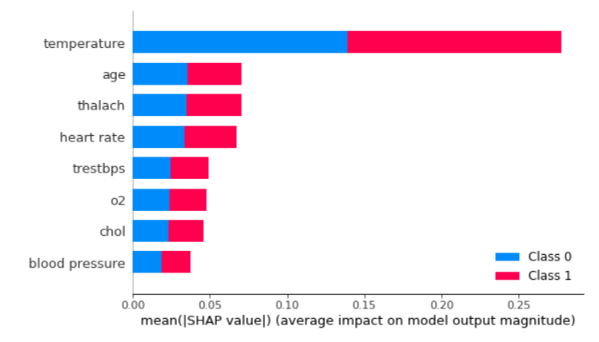
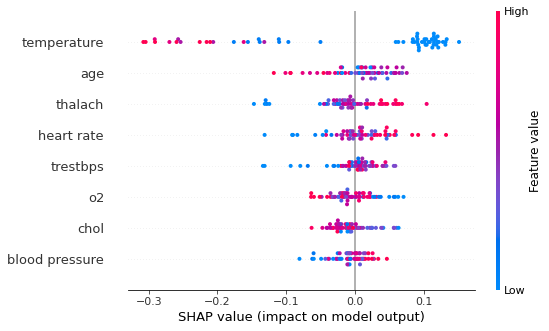


Explain:

The oxygen level vs score graph states the peak of oxygen level of a patient and how much accurately it was predicted. The patient can be easily diagnosed by looking to this plot of average raise in range of oxygen level.

SHAP:

**SHAP (SHapley Additive exPlanations)** is a game theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the classic Shapley values from game theory and their related extensions.

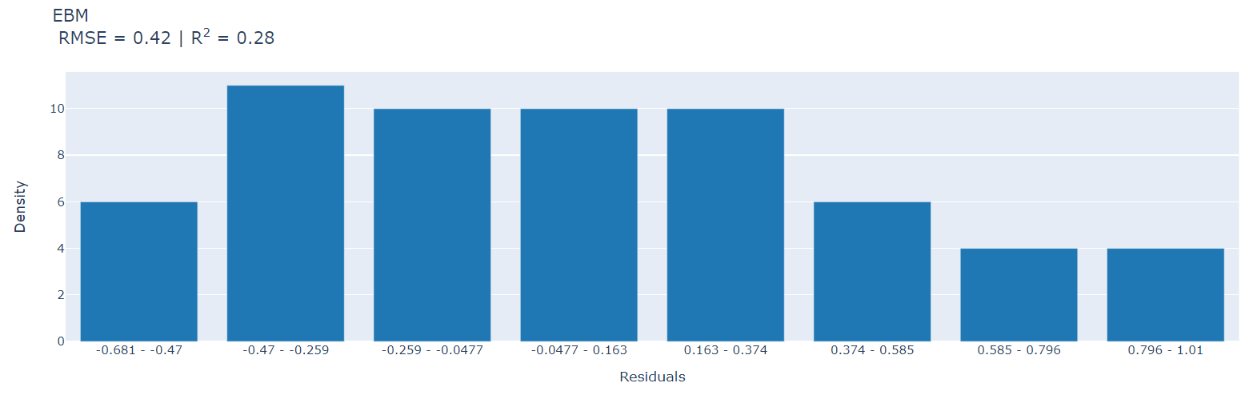
 

Explain:

The input parameter like temperature,age,thalch,heart rate,trestbps,o2,chol and blood pressure are interpreted with the model. It states that temperature parameter yields an better accuracy in AI model comparable to all exsisting parameter. The class 0 and class 1 used to indicate good and bad probability of prediction. Overall all input parameter are producing equal number of good and bad result and model stability can be defined.

EBM:

EBM is a glassbox model, designed to have accuracy comparable to state-of-the-art machine learning methods like Random Forest and BoostedTrees, while being highly intelligible and explainable. EBMs are a form of Generalized Additive Models (GAM), formalized by Trevor Hastie and Robert Tibshirani.



Explain:

The input parameters like temperature,age,heart rate,blood pressure,o2 and thalch are concatenated and mean was calculated . Use of the mean the root was taken on value and that was plotted with certain ranges.

Conclusion:

By using XAI, we can analyze our model and generate alerts when models deviate from intended outcomes and perform inadequately.

With health care ,XAI can accelerate diagnostics and image processing and streamline the pharmaceutical approval process.

By implementing XAI,we can ensure that our model/system is working as expected.