DBMS Praktikum
AI Tools Lab
SS 2020

Lime Team

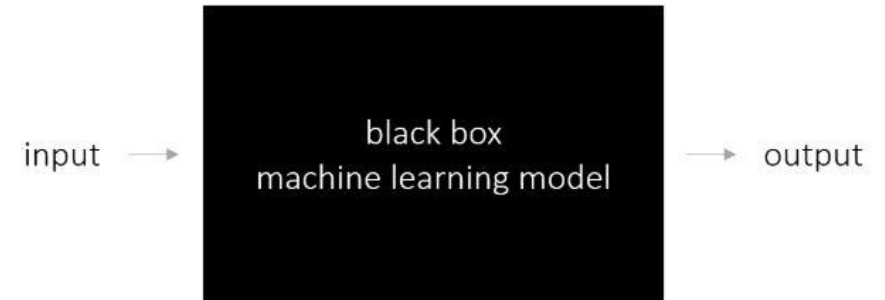Takak & parnet

Version : 16.06.2020

Contents

- LIME
- Datasets
- MS IML Implementation
- IBM AIX360 Implementation
- Evaluate and Compare
- Source

LIME (short version)

Some Important Properties

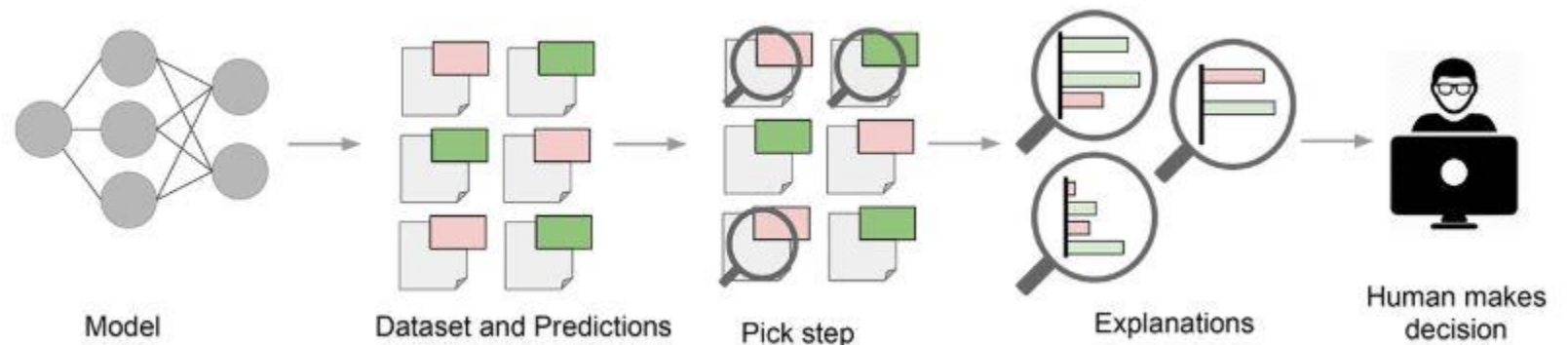*LIME* can be applied to any machine learning model

LIME try to understand the model by perturbing the input of data samples and understanding how the predictions change.

LIME provides local model interpretability.

Local

Interpretable

Model-Agnostic

Explanation

Model   Dataset and Predictions   Pick step   Explanations   Human makes decision

Dataset (1)

Mobile Price Classification

 classify mobile price range

Goal   **predict price range of an mobile phone**

            price range indicating how high the price is

            depending on individual parts such as battery power and memory

Price Range	
0	very cheap
1	cheap
2	expensive
3	very expensive

Source : https://www.kaggle.com/iabhishekofficial/mobile-price-classification

MS IML Explanation Mobile Price

- Trainings data 67%

- Test data 33%
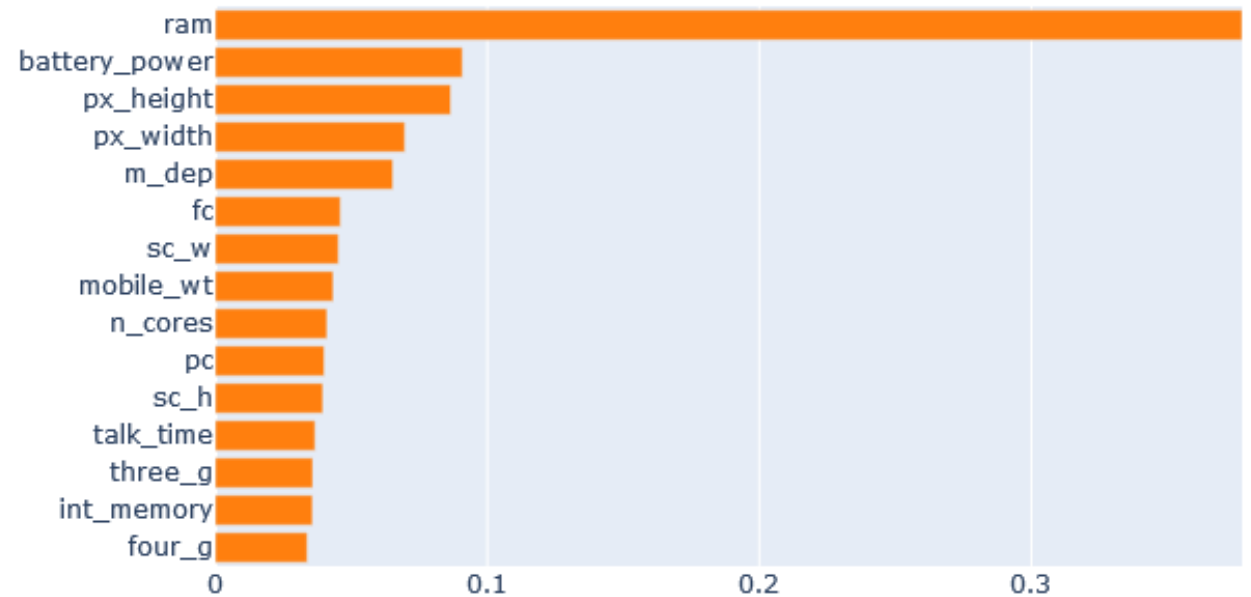
- Linear Regression

$$R^2 \approx 0.91$$

- Random Forest Classifier

$$R^2 \approx 0.95$$

$(R^2$ coefficient of determination)
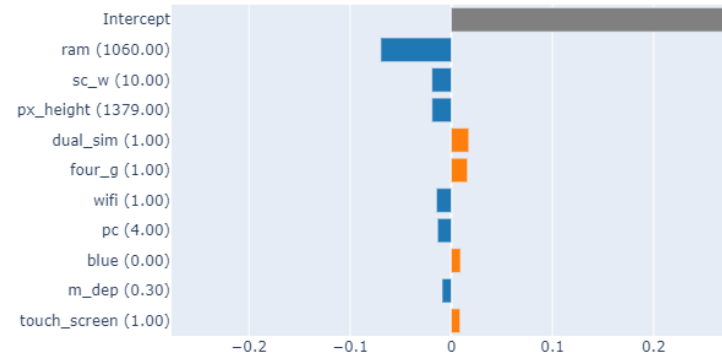

Morris Sensitivity
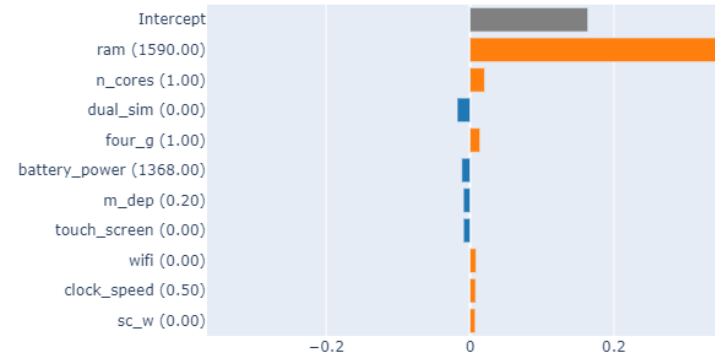Convergence Index: 0.060

Ram is dominant
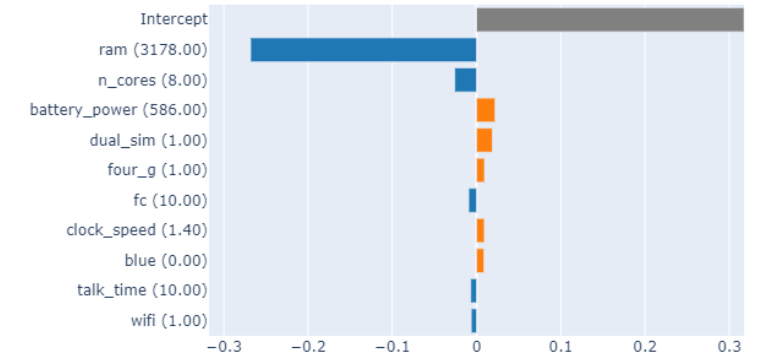
MS IML Explanation Mobile Price
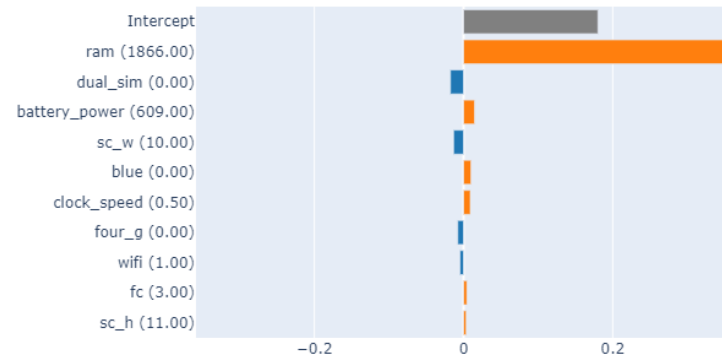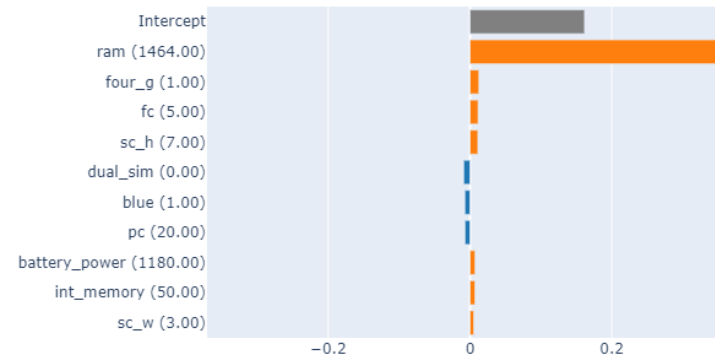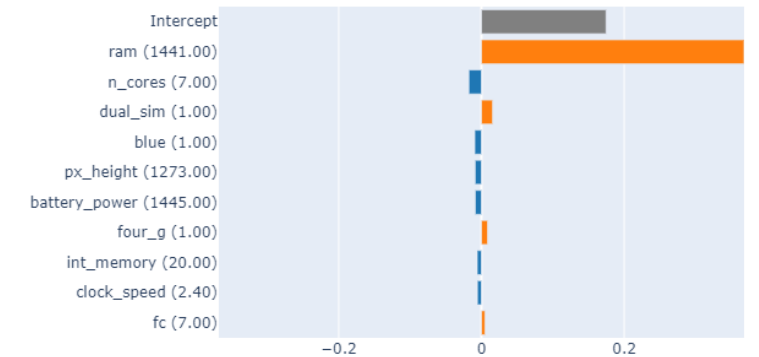
Data relate to class 1 (cheap)

Ram is dominant

MS IML Explanation Mobile Price

Explainable Boosting Machine
How probably is the assignment to a class depending on RAM

Partial Dependence
Average response of each attribute depending on RAM

Data quality

data is not realistic

IBM AIX360 Explanation Mobile Price

Direct comparison between multiple classes at once

Prediction probabilities
- very cheap: 0.00
- cheap: 1.00
- expensive: 0.00
- very expensive: 0.00

NOT cheap | cheap
- 1191.00 < ram <= 20... 0.65
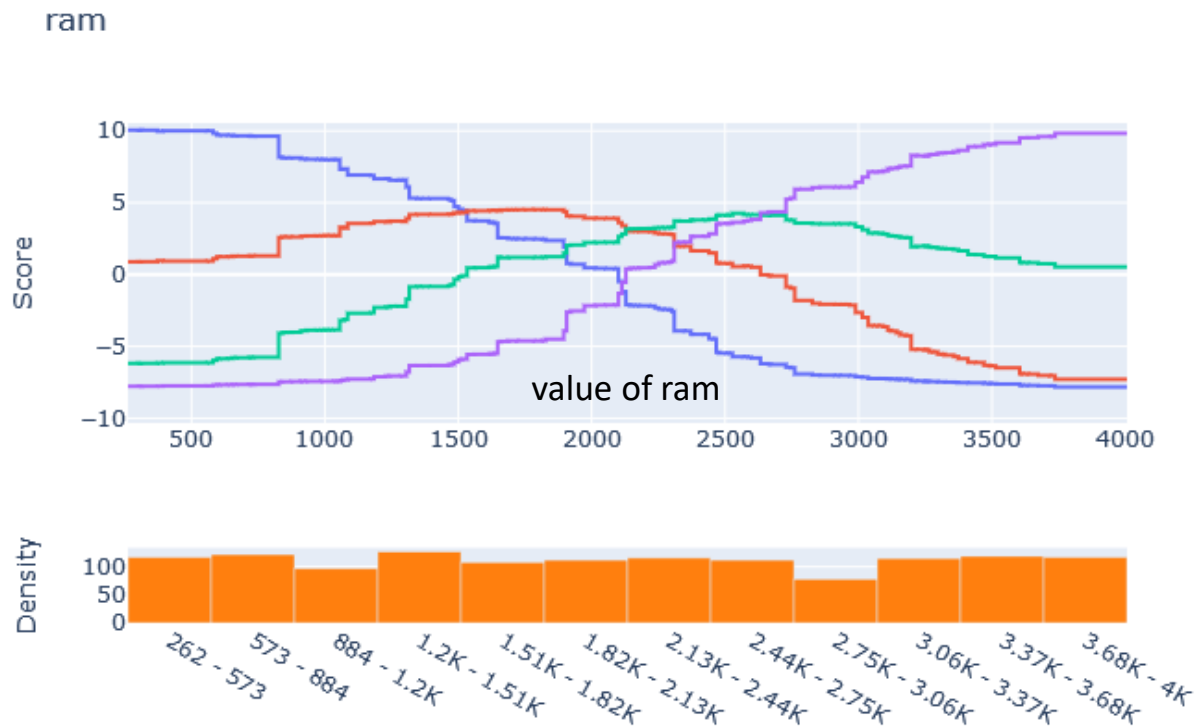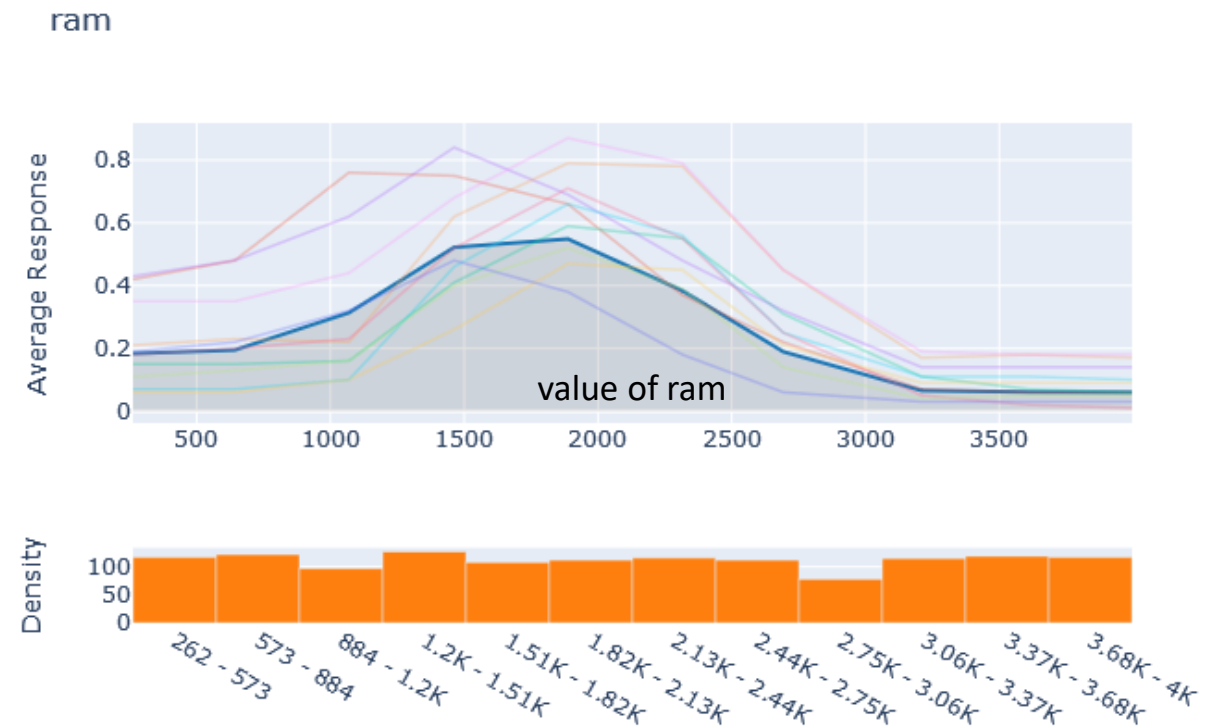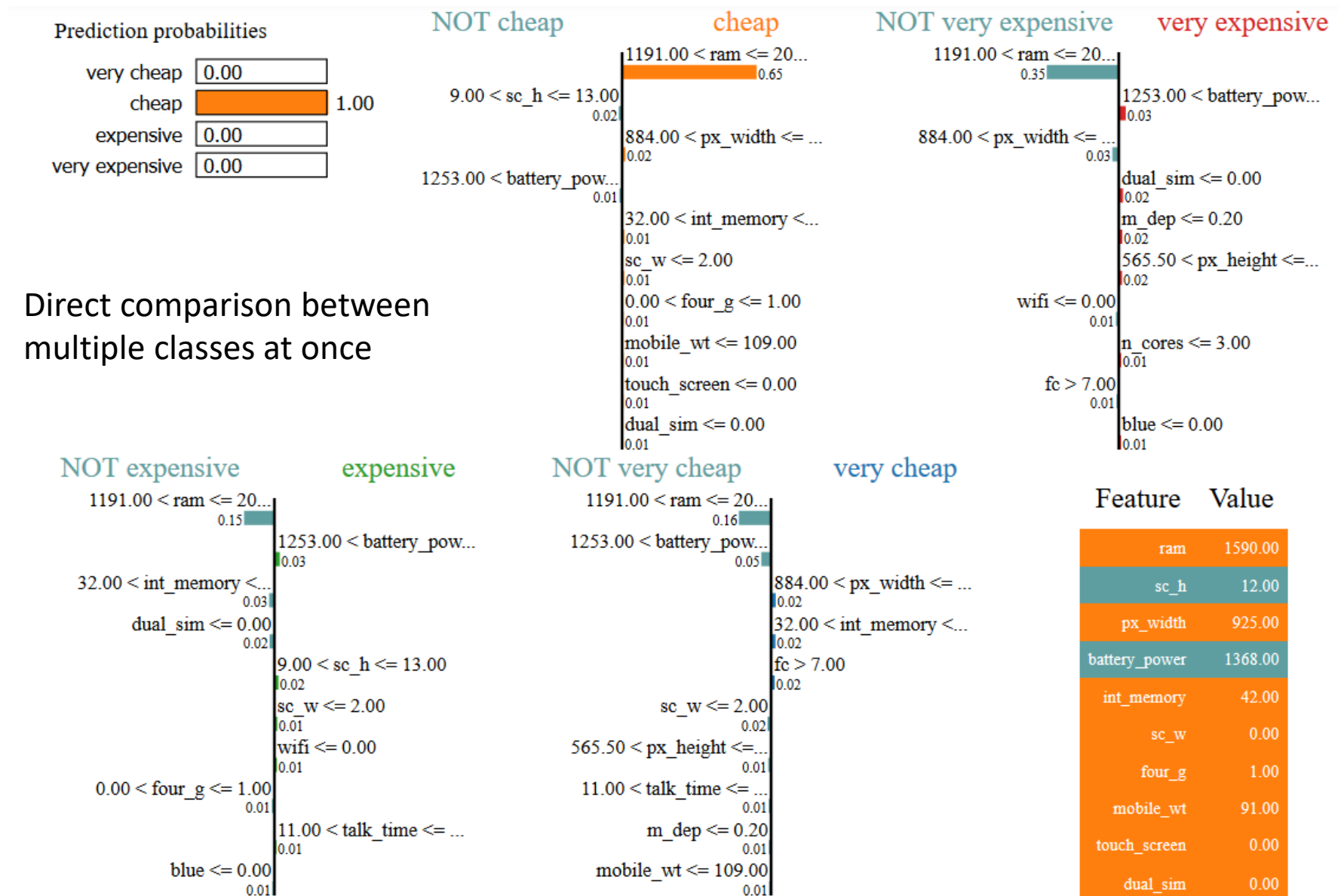- 9.00 < sc_h <= 13.00 0.02
- 884.00 < px_width <= ... 0.02
- 1253.00 < battery_pow... 0.01
- 32.00 < int_memory <... 0.01
- sc_w <= 2.00 0.01
- 0.00 < four_g <= 1.00 0.01
- mobile_wt <= 109.00 0.01
- touch_screen <= 0.00 0.01
- dual_sim <= 0.00 0.01

NOT very expensive | very expensive
- 1191.00 < ram <= 20... 0.35
- 1253.00 < battery_pow... 0.03
- 884.00 < px_width <= ... 0.03
- dual_sim <= 0.00 0.02
- m_dep <= 0.20 0.02
- 565.50 < px_height <=... 0.02
- wifi <= 0.00 0.01
- n_cores <= 3.00 0.01
- fc > 7.00 0.01
- blue <= 0.00 0.01

NOT expensive | expensive
- 1191.00 < ram <= 20... 0.15
- 1253.00 < battery_pow... 0.03
- 32.00 < int_memory <... 0.03
- dual_sim <= 0.00 0.02
- 9.00 < sc_h <= 13.00 0.02
- sc_w <= 2.00 0.01
- wifi <= 0.00 0.01
- 0.00 < four_g <= 1.00 0.01
- 11.00 < talk_time <= ... 0.01
- blue <= 0.00 0.01

NOT very cheap | very cheap
- 1191.00 < ram <= 20... 0.16
- 1253.00 < battery_pow... 0.05
- 884.00 < px_width <= ... 0.02
- 32.00 < int_memory <... 0.02
- fc > 7.00 0.02
- sc_w <= 2.00 0.01
- 565.50 < px_height <=... 0.01
- 11.00 < talk_time <= ... 0.01
- m_dep <= 0.20 0.01
- mobile_wt <= 109.00 0.01

Feature	Value
ram	1590.00
sc_h	12.00
px_width	925.00
battery_power	1368.00
int_memory	42.00
sc_w	0.00
four_g	1.00
mobile_wt	91.00
touch_screen	0.00
dual_sim	0.00

Dataset (2)

NASA: Asteroids Classification

classify dangerousness of an Asteroid

Goal   **predict whether an Asteroid is dangerous or not**

finding potential hazardous and non-hazardous asteroids

depending on individual parts such as asteroid size, speed and orbit

Hazardous	
True	**dangerous**
False	**safe**

Source : https://www.kaggle.com/shrutimehta/nasa-asteroids-classification

MS IML Explanation Asteroids

- Trainings data 67%
- Test data 33%
- Standardized

- Linear Regression
  $$R^2 \approx 0.40$$

- Random Forest Classifier
  $$R^2 \approx 0.89$$

(R^2 coefficient of determination)
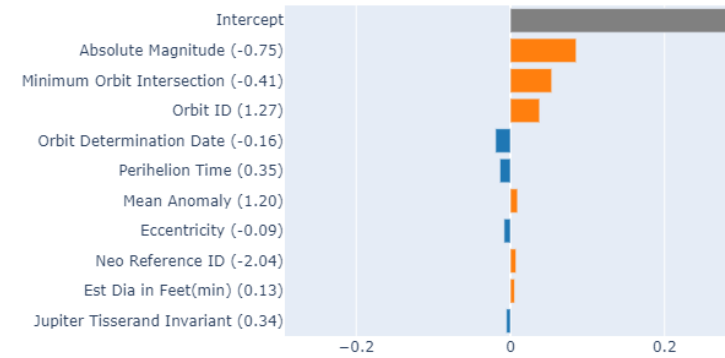
Morris Sensitivity
Convergence Index: 0.080

Minimum Orbit Intersection	
Orbit ID	
Inclination	
Absolute Magnitude	
Eccentricity	
Perihelion Distance	
Perihelion Time	
Orbital Period	
Orbit Determination Date	
Mean Motion	
Aphelion Dist	
Epoch Osculation	
Jupiter Tisserand Invariant	
Orbit Uncertainty	
Relative Velocity km per hr	
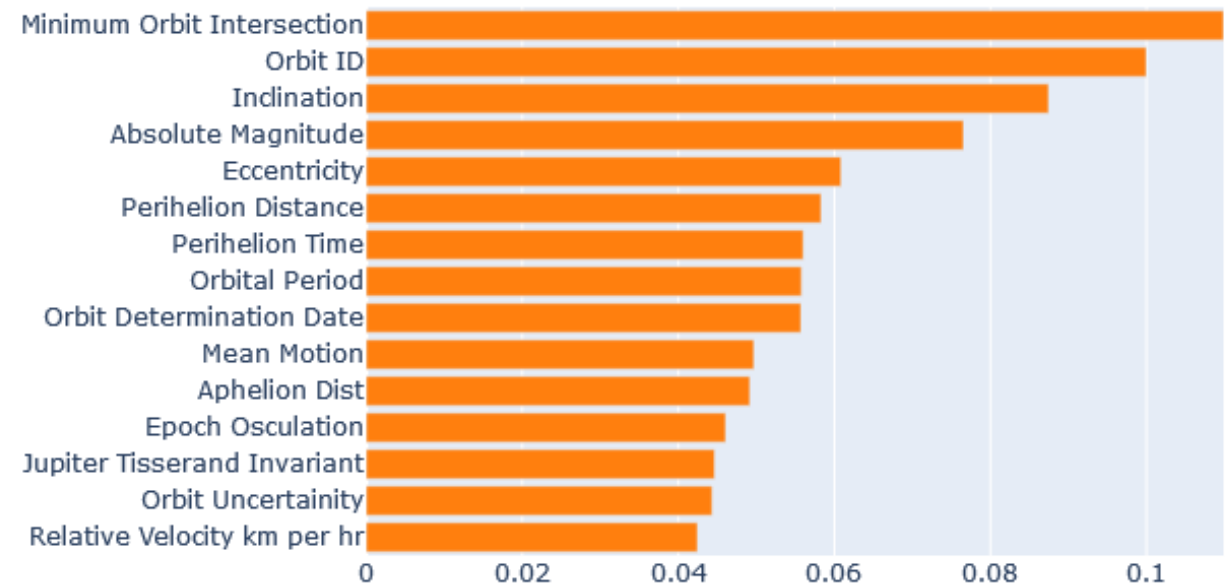
No linear relationships basing on single attribute

MS IML Explanation Asteroids
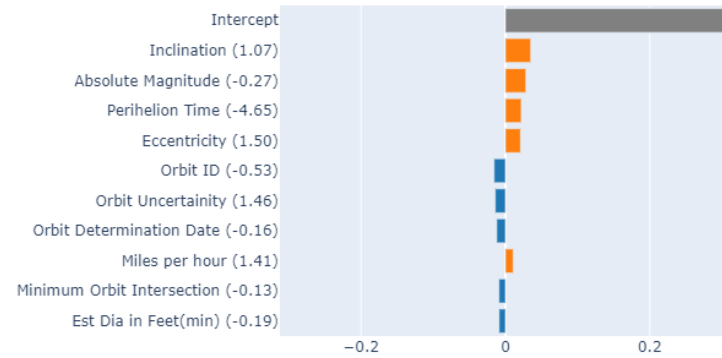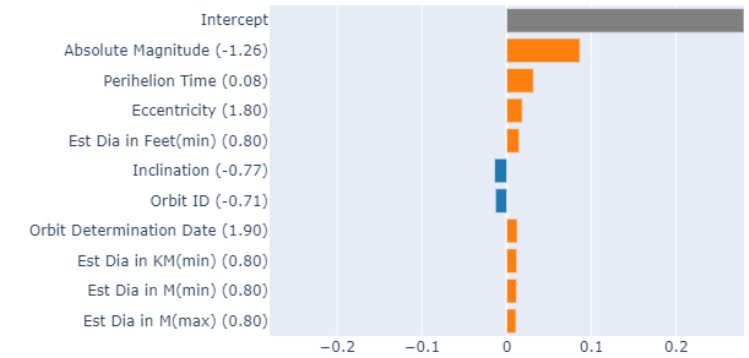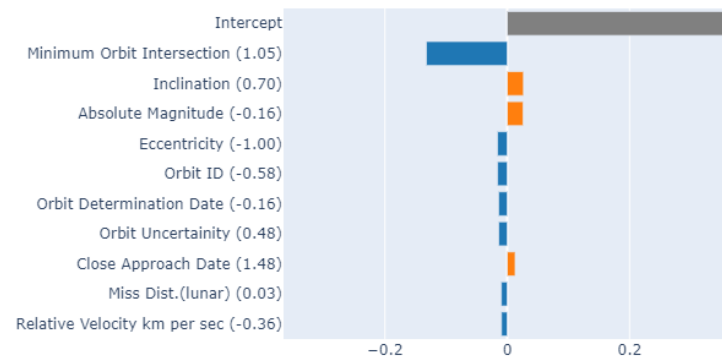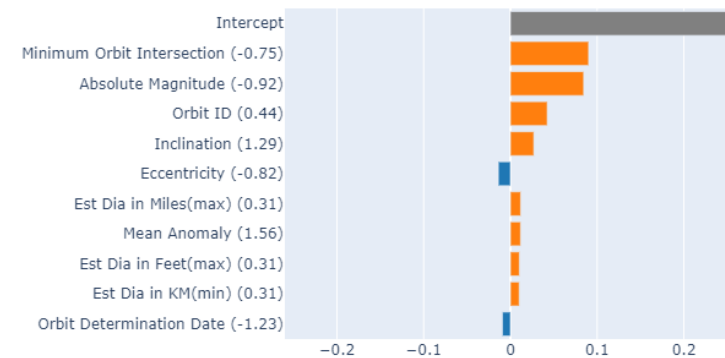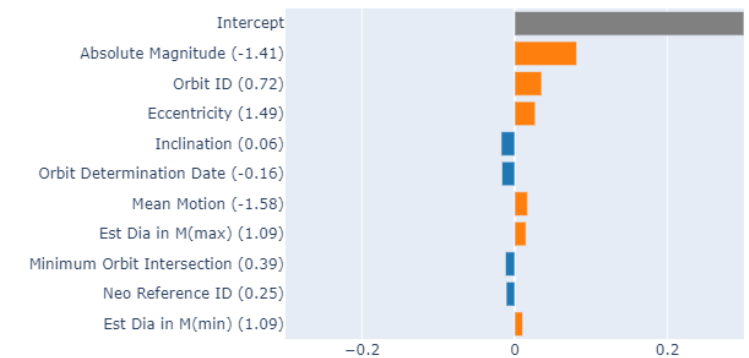
Data relate to class 1 (safe)

MS IML Explanation Asteroids

Receiver Operating Characteristics

ROC Curve: Blackbox
AUC = 0.9710

Area Under The Curve

Higher the AUC, better the model is at predicting

True Positive Rate
Sensitivity

TPR

1

0.5

0

0   0.2   0.4   0.6   0.8   1

FPR
False Positive Rate
1-Specificity
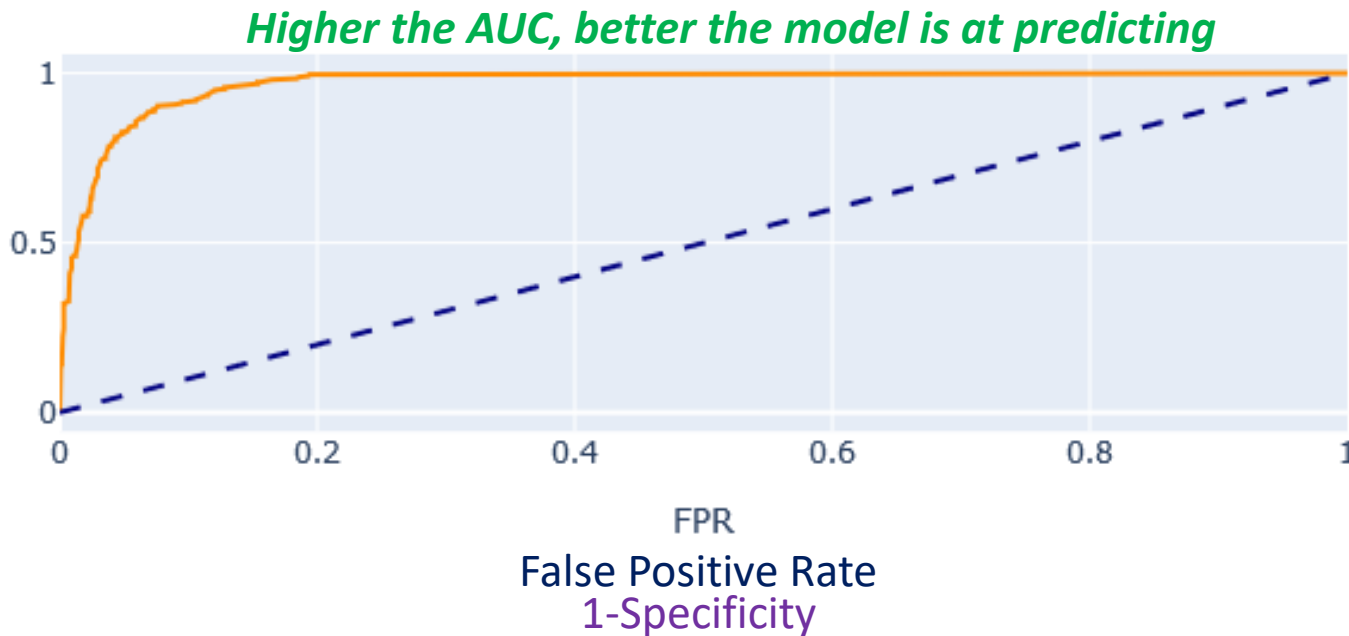
Review the performance of a classification model

IBM AIX360 Explanation Asteroids

Evaluate and Compare

Interpret ML

- Better visualization and user interaction

- more easy global explanation features

- restrictet to tabular data

- Summarization of different explainability methods

AIX360

- Many example notebooks

- comparison between multiple classes

- processing with images, texts, tabular, voices etc. is available

Source

- https://www.oreilly.com/content/introduction-to-local-interpretable-model-agnostic-explanations-lime/

- https://towardsdatascience.com/understanding-model-predictions-with-lime-a582fdff3a3b

- https://www.youtube.com/watch?v=hUnRCxnydCc&feature=emb_logo

- https://www.statistik-nachhilfe.de/ratgeber/statistik/deskriptive-statistik/visualisierung-von-daten/roc-kurve

- https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5

- Datasets
    - https://www.kaggle.com/iabhishekofficial/mobile-price-classification
    - https://www.kaggle.com/shrutimehta/nasa-asteroids-classification

- Project Files
    - https://github.com/parnet/Data-Science-Praktikum/