

Validation of models with auditor: : CHEAT SHEET



Intro

Package **auditor** is a tool for model-agnostic validation. Implemented techniques facilitate assessing and comparing the goodness of fit and performance of models. In addition, they may be used for the analysis of the similarity of residuals and for the identification of outliers and influential observations.

The auditor is based on the DrWhy grammar, the idea that you can wrap up a model into explainer. Such explainer can be used for further analysis of model in a uniform way.

PREPARING EXPLAINER OF MODEL

The **DALEX::explain()** function creates model adapters: objects with standardised structure that are used by auditor functions for model validation and exploration.

```
DALEX::explain(model,  
data, y, label, predict_function, residual_function)
```

CALCULATION FUNCTIONS

Created explainer object can be passed to five calculation functions from the **auditor** package. Each function focuses on other validation aspect.

- Exploration of residuals (distribution, homoscedasticity, ...)

```
model_residual(explainer)
```

- Evaluation of classifiers (derivations from confusion matrix)

```
model_evaluation(explainer)
```

- Measuring a performance of model (MSE, MAE, AUC, ...)

```
model_performance(explainer, score, new_score)
```

- Influence of observations (Cook's distances)

```
model_cooksdistance(explainer)
```

- Verification of model fit (Half-Normal plots)

```
model_halfnormal(explainer, quant, ...)
```

GENERAL WORKFLOW

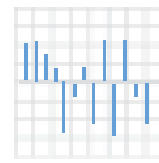
Explanations can be both plotted and printed.

```
explanation <- model %>%  
  explain %>%  
  <calculation function>
```

```
plot(explanation)  
plotD3(explanation)  
print(explanation)
```

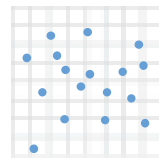
Exploration of residuals

model_residual(explainer)



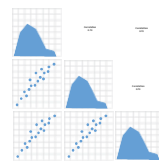
Autocorrelation Function of Residuals

```
plot(explanation, ..., type = "acf")  
plot_acf(explanation, ..., variable, alpha)  
plotD3_acf(explanation, ..., variable, alpha, scale_plot)
```



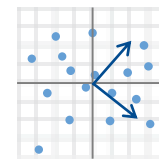
Autocorrelation of Residuals

```
plot(explanation, ..., type = "autocorrelation")  
plot_autocorrelation(explanation, ..., variable, smooth)  
plotD3_autocorrelation(explanation, ..., variable, points,  
smooth, point_count, single_plot, scale_plot, background)
```



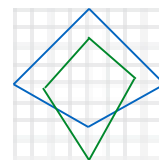
Correlation of residuals

```
plot(explanation, ..., type = "correlation")  
plot_correlation(explanation, ..., values)
```



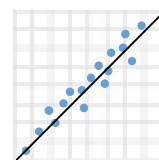
Principal Components of model's of residuals

```
plot(explanation, ..., type = "pca")  
plot_pca(explanation, ..., scale)
```



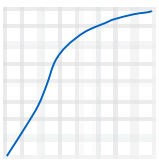
Model Ranking Radar Plot

```
plot(explanation, ..., type = "radar")  
plot_radar(explanation, ..., score, new_score)
```



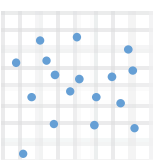
Predicted Response vs Actual or Variable Values

```
plot(explanation, ..., type = "prediction")  
plot_prediction(explanation, ..., variable, smooth, abline)  
plotD3_prediction(explanation, ..., variable, points, smooth,  
abline, point_count, single_plot, scale_plot, background)
```



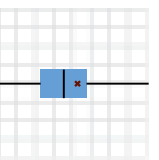
Regression Error Characteristic Curve (REC)

```
plot(explanation, ...)  
plot_rec(explanation, ...)  
plotD3_rec(explanation, ..., scale_plot)
```



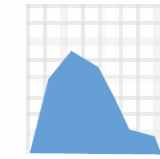
Plot Residuals vs Actual, Fitted or Variable Values

```
plot(explanation, ..., type = "prediction")  
plot_residual(explanation, ..., variable, smooth,  
std_residuals, nlabel)  
plotD3_residual(explanation, ..., points, std_residuals,  
nlabel, point_count, single_plot, scale_plot, background)
```



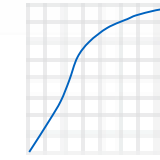
Residual Boxplot

```
plot(explanation, ..., type = "residual_boxplot")  
plot_residual_boxplot(explanation, ...)
```



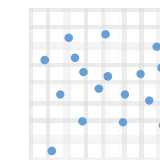
Residual Density

```
plot(explanation, ..., type = "residual_density")  
plot_residual_density(explanation, ..., split, variable)
```



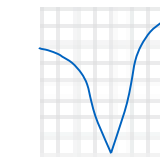
Regression Receiver Operating Characteristic (RROC)

```
plot(explanation, ..., type = "rroc")  
plot_rroc(explanation, ...)
```



Scale-Location Plot

```
plot(explanation, ..., type = "scalelocation")  
plot_scalelocation(explanation, ..., variable, smooth, peaks)  
plotD3_scalelocation(explanation, ..., variable, smooth,  
peaks, point_count, single_plot, scale_plot, background)
```

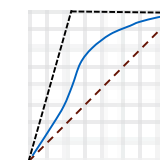


Two-sided Cumulative Distribution Function

```
plot(explanation, ..., type = "tsecdf")  
plot_tsecdf(explanation, ..., scale_error, outliers, residuals,  
reverse_y)
```

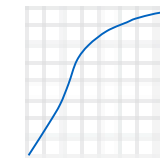
Evaluation of classifiers

model_evaluation(explainer)



LIFT Curve

```
plot(explanation, ..., type = "lift")  
plot_lift(explanation, ...)  
plotD3_lift(explanation, ..., scale_plot)
```

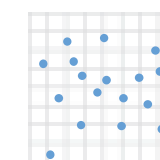


Receiver Operating Characteristic (ROC) Curve

```
plot(explanation, ..., type = "roc")  
plot_roc(explanation, ..., nlabel)
```

Influence of observations

model_cooksdistance(explainer)

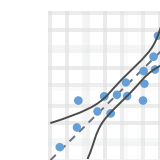


Cook's Distances of Observations

```
plot(explanation, ..., type = "cooksdistance")  
plot_cooksdistance(explanation, ..., nlabel)  
plotD3_cooksdistance(explanation, ..., nlabel, single_plot,  
scale_plot, background)
```

Verification of model fit

model_halfnormal(explainer, quant, ...)



Half-normal plots

```
plot(explanation, ..., type = "halfnormal")  
plot_halfnormal(explanation, ..., quantiles, sim)  
plotD3_halfnormal(explanation, ..., nlabel, single_plot,  
scale_plot, background)
```

