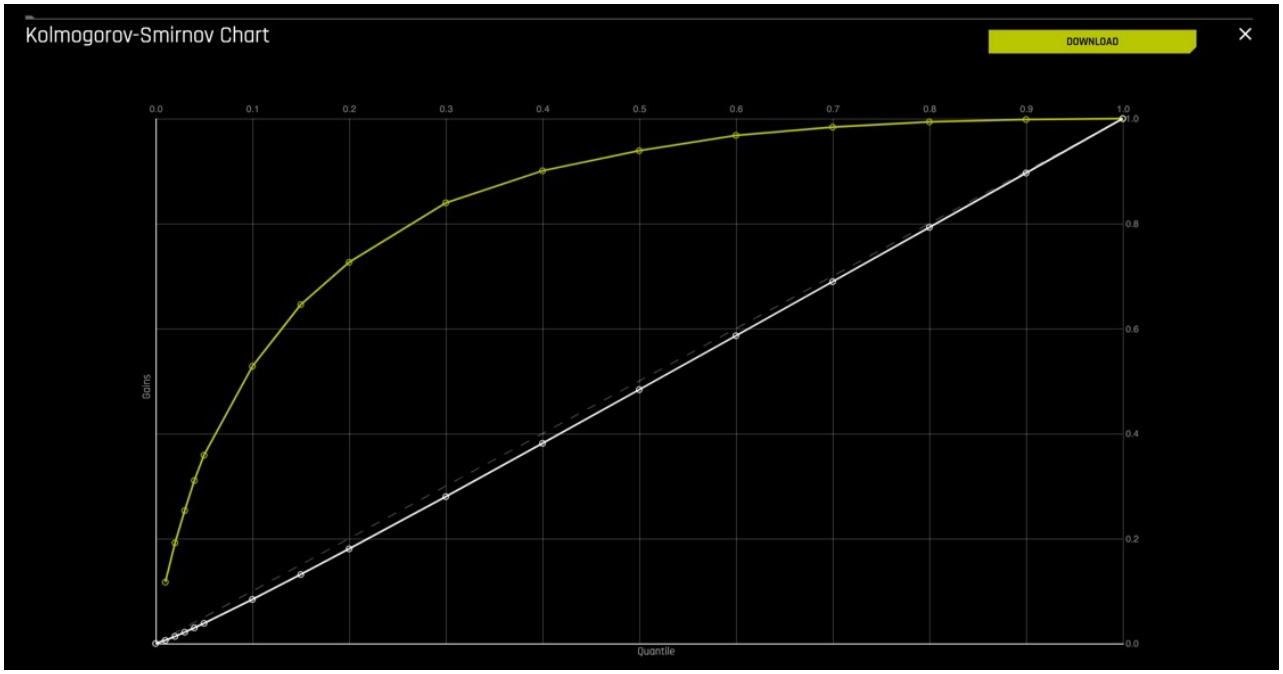
Machine Learning Experiment Scoring and Analysis Tutorial - Financial Focus

12. Task 10: Kolmogorov-Smirnov Chart

Continuing on the diagnostics page, select the **KS** chart. The K-S chart should look similar to the one below:



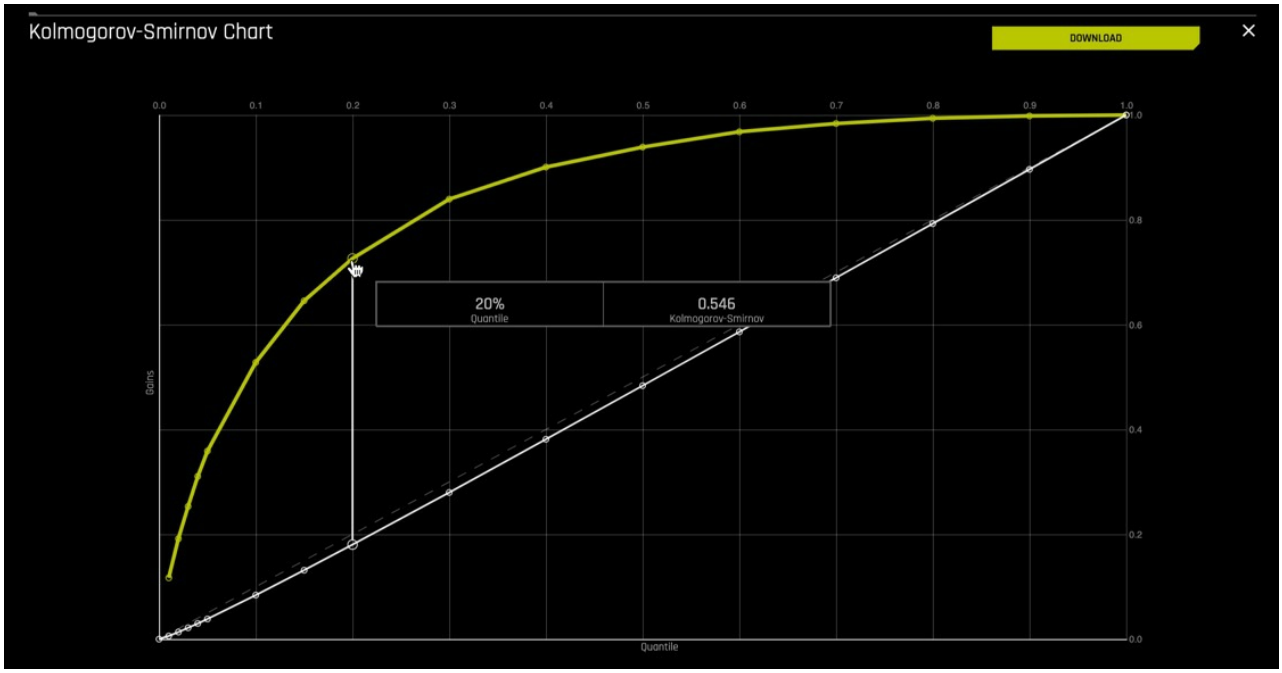
Remember that for the K-S chart:

- K-S measures the performance of classification models by measuring the degree of separation between positives and negatives for validation or test data.
- The K-S is 100 if the scores partition the population into two separate groups in which one group contains all the positives and the other all the negatives
- If the model cannot differentiate between positives and negatives, then it is as if the model selects cases randomly from the population and the K-S would be 0
- The K-S range is between 0 and 1
- The higher the K-S value, the better the model is at separating the positive from negative cases

Note: The y-axis of the plot has been adjusted to represent quantiles, this allows for focus on the quantiles that have the most data and therefore the most impact.

1. Hover over the various quantile points on the Lift chart to view the quantile percentage and cumulative lift values

2. What is the cumulative lift at 1%, 2%, 10% quantiles?



For this K-S chart, if we look at the top 20% of the data, the at-chance model (the dotted diagonal line) tells us that only 20% of the data was successfully separate between positives and negatives (defaulted and not defaulted). However, with the model it was able to do .546 or about 55% of the cases were successfully separated between positives and negatives.

3. Based on the K-S curve(yellow) and the baseline (white diagonal dashed line) is this a good model?

4. Exit out of the K-S chart by clicking on the **x** located at the top-right corner of the plot, next to the **Download** option

Deeper Dive and Resources

BackNext