Applied Machine Learning

Practical Performance Evaluation
Lift Charts & Calibration Curves

Learning goals

- How to visualize model performance beyond standard metrics?
- What are lift charts and how do we interpret them?
- How do we assess probability calibration?

LIFT CHARTS

- Lift charts visualize a model's ability to detect events.
- Samples are ranked according to their scores (probability for an event to take place). We then evaluate the cumulative event rate.
- In the optimal case, we would suspect the m highest ranked samples to contain all m events.
- For a non-informative model, the fraction $\frac{k}{n}$ of the highest ranked samples contains k events on average.
- For a more informative model, the fraction $\frac{k}{n}$ of the highest ranked samples contains more than k events.
- The **lift** corresponds to the amount of events detected by a model above a completely random selection of samples.

LIFT CHART EXAMPLE

- We plot the cumulative event rate against the percentage of screened observations.
- Here, we face a 50% event rate. For a perfect model, the 50% highest ranked observations contain 50% of all events.
- We create a lift curve for a non-informative model which runs along the diagonal line.
- Another model is able to perfectly detect events. After evaluating 50% of the observations associated with the highest event probabilities, we have captured all events.

Figure: Lift chart from Applied Predictive Modeling (Max Kuhn).

CALIBRATION CURVES

- Calibration curves plot the true frequency of the positive label against its predicted probability for binned predictions.
- The horizontal axis represents the average predicted probability in each bin. The vertical axis indicates the the proportion of predicted positive labels.
- A well-calibrated classifier is associated with a diagonal clibration curve.

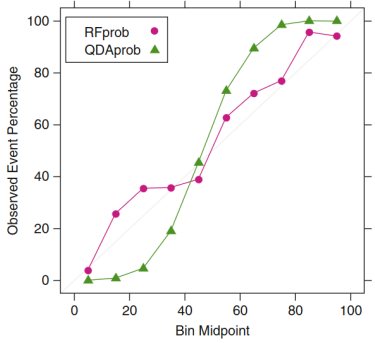

Figure: Calibration plot from Applied Predictive Modeling (Max Kuhn).