

Rejection Classification Index (RC-Index) can evaluate how well uncertainty functions as a predictor of incorrectness.

 Katherine E. Brown, PhD, Steve Talbert, PhD, Douglas A. Talbert, PhD

Assessing the Quality of Uncertainty Calibration

INTRO

- Standard empirical evaluation of uncertainty involves generating and visually inspecting rejection-classification plots.

Rejection Classification Index

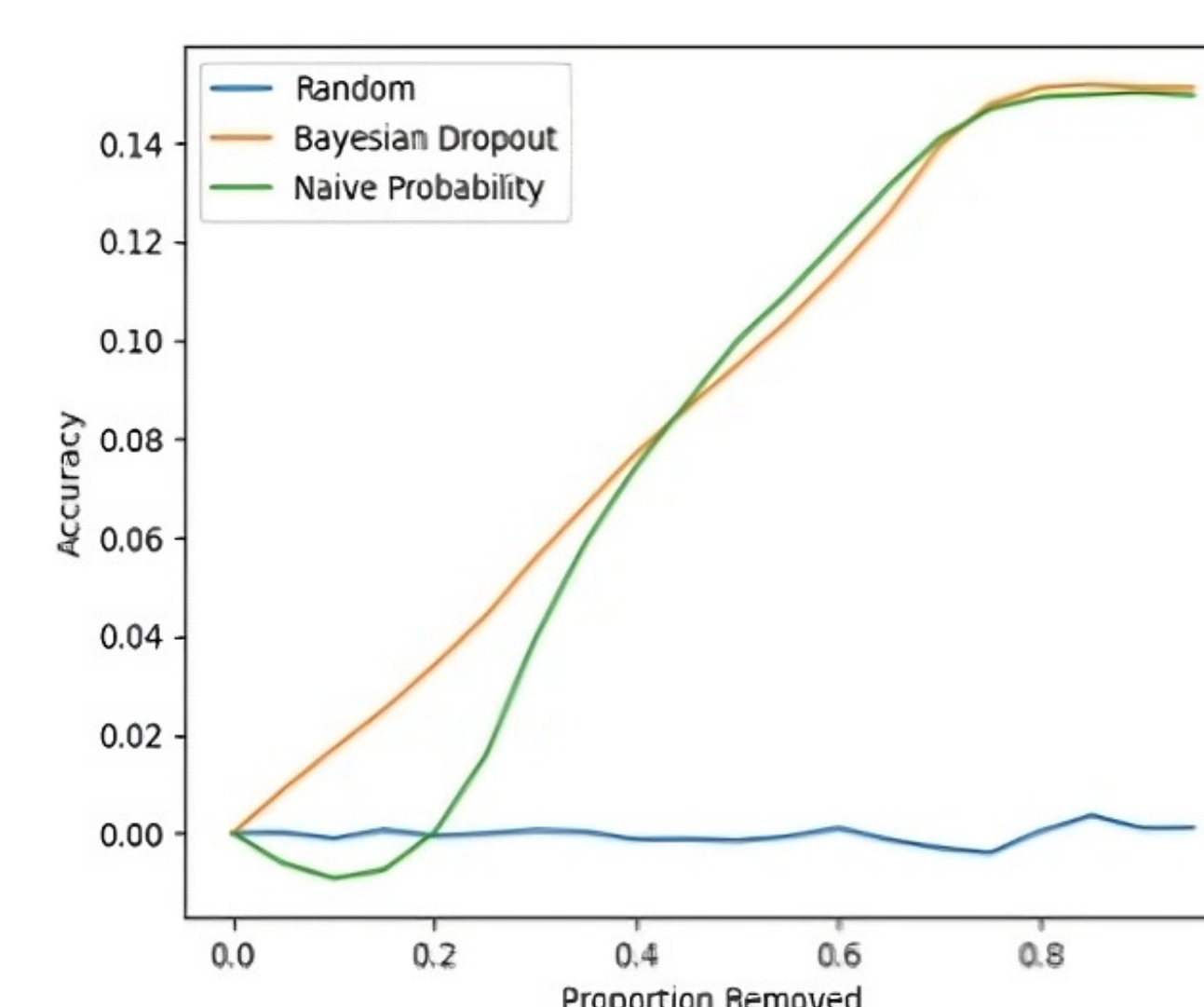
- Let x_i indicate the proportion of most uncertain data and x_i indicate model's accuracy with this data removed
- RC-Index is calculated as the area under the curve given by $\{(x_0, 0), (x_1, y_1 - y_0), \dots, (x_m, y_m - y_0)\}$
- Enforces initial accuracy starting at 0 to emphasize magnitude and direction of accuracy change
- Possible range of values is -1 to 1.

Methods

- All values and figures are averaged across 10-fold CV.
- Visualization produced by vertically averaging accuracy across 10-fold CV.
- Dataset: Level 1 Trauma Center In-Field Trauma Triage Registry ($n = 52,888$)
- Classification Model: Neural Network
- Uncertainty Measures:
 - Epistemic Uncertainty: Bayesian Dropout
 - Softmax Uncertainty: Derived from output probabilities
- Proposed UQ evaluation techniques:
 - Rejection Classification Curves
 - RC-Index

RESULTS

Uncertainty Measure	RC-Index
Bayesian Uncertainty	0.084
Softmax Uncertainty	0.076
Random Control	-0.0007



DISCUSSION

- Epistemic and Softmax uncertainty produce usable rejection-classification compared to random control.
- Epistemic uncertainty produces a more consistent, monotonic increase than Softmax uncertainty
- Behavior of rejection classification curves reflected in RC-Index values

VANDERBILT  UNIVERSITY
MEDICAL CENTER

Department of
Biomedical Informatics



Take a picture to
download the full text

