Approximating Full Conformal Prediction for Neural Network Regression with Gauss-Newton Influence

Dharmesh Tailor*[1], Alvaro H.C. Correia[2], Eric Nalisnick[3], Christos Louizos[2]

*PhD Student

[1]University of Amsterdam, [2]Qualcomm AI Research, [3]Johns Hopkins University

Qualcomm AI Research is an initiative of Qualcomm Technologies, Inc.

Problem formulation

Setting

- Training data $D_N = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)\}$
- Unseen test point $(\mathbf{x}_{N+1}, y_{N+1})$
- Point prediction $\hat{y}_{N+1} = f(\mathbf{x}_{N+1}; \boldsymbol{\theta}_*)$

Goal

- construct prediction interval C_α that contains y_{N+1} with high probability

  <u>marginal coverage</u> $\mathcal{P}\left(y_{N+1} \in C_\alpha(\mathbf{x}_{N+1})\right) \geq 1 - \alpha$ given miscoverage rate $\alpha \in (0, 1)$

Desiderata

- Distribution-free (no assumptions on parametric form of data distribution)
- Efficiency (C_α should be as tight as possible to be informative as a measure of uncertainty)
- No data-splitting (use all available data for training)

(Full) conformal prediction

Repeat for new test point $\mathbf{x}_{N+1}$

Repeat for every $y \in \mathbb{R}$

- Postulate target for test input y

$$\mathcal{D}_{N+1}(y) = \mathcal{D}_N \cup \{(\mathbf{x}_{N+1}, y)\}$$

Exchangeability assumption

- Fit model on $\mathcal{D}_{N+1}(y)$ leading to $\boldsymbol{\theta}_*^+(y)$

Symmetrical algorithm
(test point treated in same way as training data)

& compute residuals

$$R_i(y) = |y_i - f(\mathbf{x}_i; \boldsymbol{\theta}_*^+(y))| \quad \forall i = 1, \ldots, N \qquad R_{N+1}(y) = |y - f(\mathbf{x}_{N+1}; \boldsymbol{\theta}_*^+(y))|$$

- Compute rank: $\pi(y) = \sum_{i=1}^{N+1} \mathbb{1}\{R_i(y) \leq R_{N+1}(y)\}$

Check if $\pi(y) \leq \lceil (1-\alpha)(N+1) \rceil$

if Yes: include y in $C_\alpha(\mathbf{x}_{N+1})$

if No: discard y

High computational cost!

Accelerating conformal prediction

- Split-CP as a special case of Full-CP
- Certain model classes (e.g., ridge[1], Lasso[2], k-NN[3]) lead to computational shortcuts
- Approaches based on homotopy continuation techniques[4] and algorithmic stability[5]
- Approaches that trade-off validity for efficiency by approximating retraining step from a single trained model on $\mathcal{D}_N$ [6]

Influence function [Jaeckel, 1972; Koh & Liang, 2017]

$$\boldsymbol{\theta}_*^+(y) \approx \boldsymbol{\theta}_* - \mathbf{H}_*^{-1} \nabla \ell_{N+1}^{y}(\boldsymbol{\theta}_*)$$

Limitation: for regression, need to use a finite grid imposing computation-precision trade-off

This work: adapts [Martinez et al., 2023] for regression by extending conformal ridge regression[1]

[1]Nouretdinov et al., 2001 [2]Lei, 2019 [3]Papadopoulos et al., 2011 [4]Ndiaye & Takeuchi, 2019 [5]Ndiaye, 2022
[6]Martinez et al., 2023

Approximating FCP via Gauss-Newton influence (ACP-GN)

- Recovers conformal ridge regression as special case unlike [Martinez et al., 2023]

Newton-step influence [Pregibon, 1981; Beirami et al., 2017]
with Gauss-Newton approximation ("GN-influence")

$$\boldsymbol{\theta}_*^+(y) \approx \boldsymbol{\theta}_* + \frac{e_{N+1}(y)}{1 + h_{N+1}} \mathbf{H}_{\mathrm{GN}}^{-1} \nabla_{\boldsymbol{\theta}} f_{N+1}(\boldsymbol{\theta}_*)$$

- Approximate scores by piecewise linear function of postulated label

$$R_i(y) = |y_i - f(\mathbf{x}_i; \boldsymbol{\theta}_*^+(y))| \approx |a_i + b_i y|$$

$$R_{N+1}(y) = |y - f(\mathbf{x}_{N+1}; \boldsymbol{\theta}_*^+(y))| \approx |a_{N+1} + b_{N+1} y|$$

- Obtain exact form of prediction set by applying ridge regression confidence machine[1] procedure on $\{(a_i, b_i)\}_{i=1}^{N+1}$

$$\pi(y) = \sum_{i=1}^{N+1} \mathbb{1}\{y \in S_i\} \qquad \text{with} \quad S_i = \{y : |a_i + b_i y| \leq |a_{N+1} + b_{N+1} y|\}$$

[1] Nouretdinov et al., 2001

ACP-GN gains in limited-data regimes

		Avg. Width 90%	Avg. Width 95%	Avg. Width 99%	Avg. Coverage 90%	Avg. Coverage 95%	Avg. Coverage 99%
yacht N=308 I=6	**LA**	1.690±0.017	2.014±0.020	2.647±0.027	88.73±0.61 (✓)	90.78±0.59 (✗)	93.89±0.60 (✗)
	SCP	2.553±0.093	4.001±0.115	10.018±0.361	89.56±0.66 (✓)	94.07±0.39 (✓)	99.32±0.08 (✓)
	CRF	2.526±0.092	3.947±0.115	9.674±0.294	89.53±0.64 (✓)	94.10±0.38 (✓)	99.29±0.10 (✓)
	CQR	4.090±0.105	5.845±0.187	18.650±0.484	89.94±0.42 (✓)	94.42±0.32 (✓)	99.02±0.17 (✓)
	ACP-GN	**1.594**±0.016	**2.385**±0.029	**6.915**±0.067	87.36±0.58 (✓)	92.56±0.68 (✓)	99.03±0.11 (✓)
boston N=506 I=13	**LA**	9.398±0.046	**11.199**±0.055	14.718±0.072	91.24±0.31 (✓)	94.34±0.22 (✓)	97.53±0.11 (✗)
	SCP	10.635±0.123	14.509±0.171	36.272±1.847	89.56±0.42 (✓)	94.64±0.32 (✓)	99.11±0.13 (✓)
	CRF	11.932±0.605	16.073±0.862	40.690±3.333	90.01±0.33 (✓)	94.77±0.22 (✓)	99.30±0.08 (✓)
	CQR	11.692±0.129	15.115±0.213	31.628±1.822	90.10±0.33 (✓)	95.12±0.24 (✓)	99.07±0.14 (✓)
	ACP-GN	**9.182**±0.046	12.111±0.038	**20.512**±0.057	90.64±0.26 (✓)	95.49±0.16 (✓)	99.11±0.08 (✓)
energy N=768 I=8	**LA**	1.502±0.006	1.790±0.007	2.353±0.009	88.96±0.35 (✓)	92.92±0.33 (✗)	96.95±0.23 (✗)
	SCP	1.942±0.032	2.486±0.046	3.772±0.093	89.44±0.28 (✓)	94.80±0.20 (✓)	99.18±0.08 (✓)
	CRF	1.923±0.031	2.454±0.046	3.728±0.092	89.39±0.28 (✓)	94.78±0.22 (✓)	99.14±0.08 (✓)
	CQR	4.670±0.030	5.139±0.029	6.438±0.120	90.08±0.26 (✓)	95.24±0.21 (✓)	98.96±0.09 (✓)
	ACP-GN	**1.462**±0.006	**1.884**±0.008	**3.076**±0.015	88.28±0.33 (✓)	93.69±0.33 (✓)	98.88±0.11 (✓)

Thank you

Nothing in these materials is an offer to sell any of the components or devices referenced herein.

© Qualcomm Technologies, Inc. and/or its affiliated companies. All Rights Reserved.

Qualcomm and Snapdragon are trademarks or registered trademarks of Qualcomm Incorporated.
Other products and brand names may be trademarks or registered trademarks of their respective owners.

References in this presentation to "Qualcomm" may mean Qualcomm Incorporated, Qualcomm Technologies, Inc., and/or other subsidiaries or business units within the Qualcomm corporate structure, as applicable. Qualcomm Incorporated includes our licensing business, QTL, and the vast majority of our patent portfolio. Qualcomm Technologies, Inc., a subsidiary of Qualcomm Incorporated, operates, along with its subsidiaries, substantially all of our engineering, research and development functions, and substantially all of our products and services businesses, including our QCT semiconductor business.

Snapdragon and Qualcomm branded products are products of Qualcomm Technologies, Inc. and/or its subsidiaries. Qualcomm patented technologies are licensed by Qualcomm Incorporated.

Follow us on:

For more information, visit us at qualcomm.com & qualcomm.com/blog