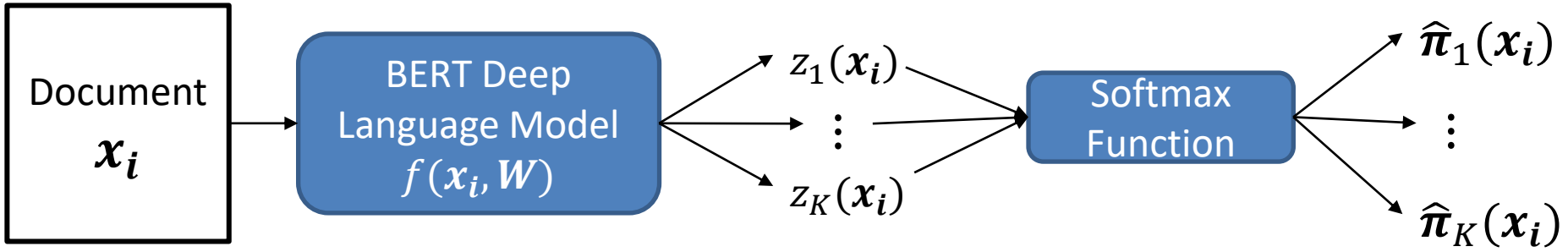
Correlated Belief Matching for Uncertainty Quantification in Text Classification

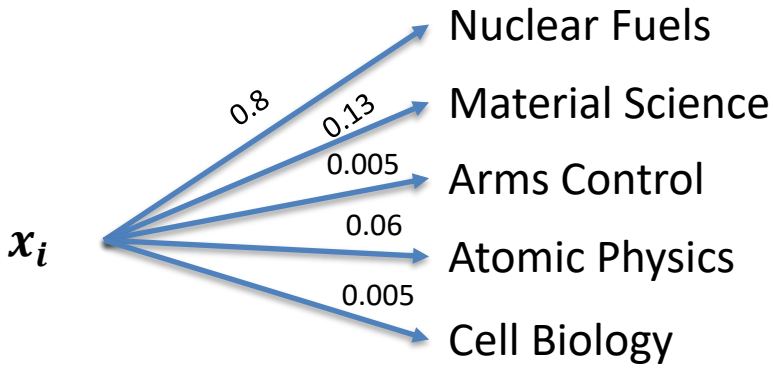
Andrew Hollis, North Carolina State University, Department of Statistics
Karl Pazdernik, Pacific Northwest National Laboratory
Alyson Wilson, North Carolina State University, Department of Statistics
Ralph Smith, North Carolina State University, Department of Mathematics

April 26, 2023

PNNL-SA-184001

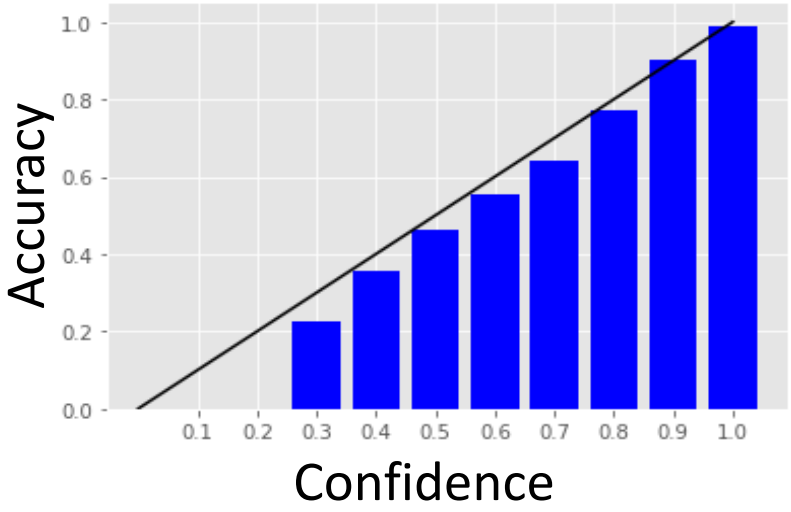
Text Classification



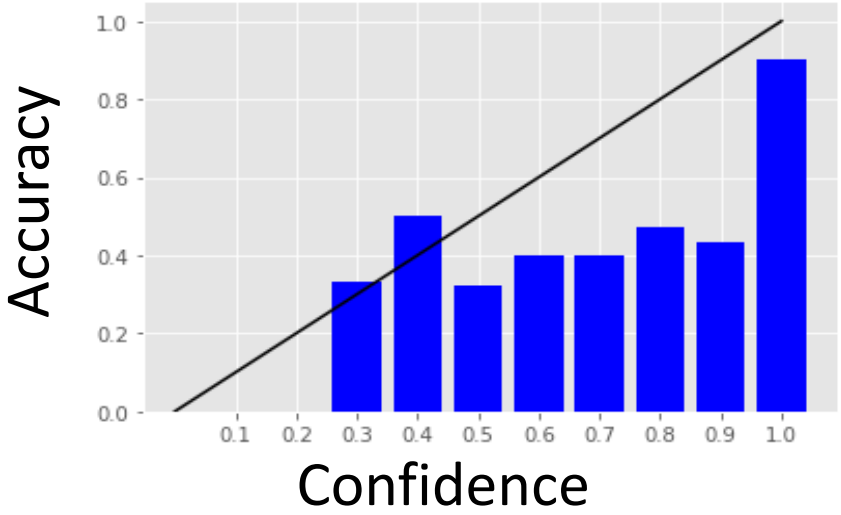


Calibrated: Expected Calibration Error (ECE) [1]

Well-Calibrated

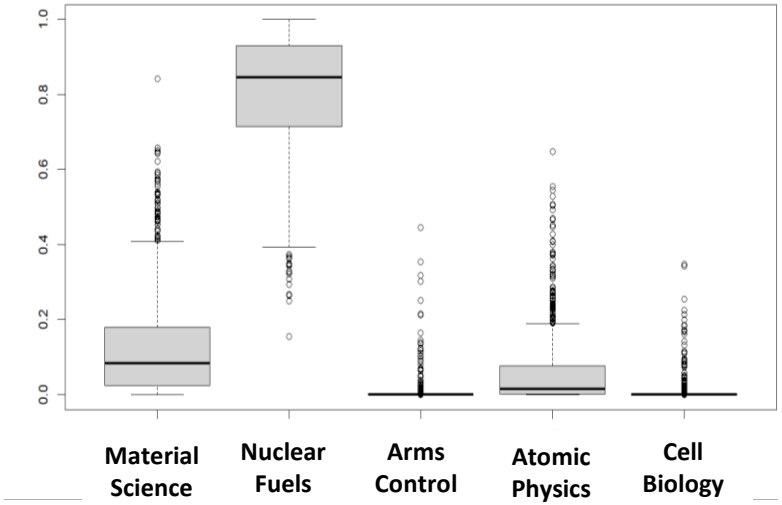


Poorly-Calibrated

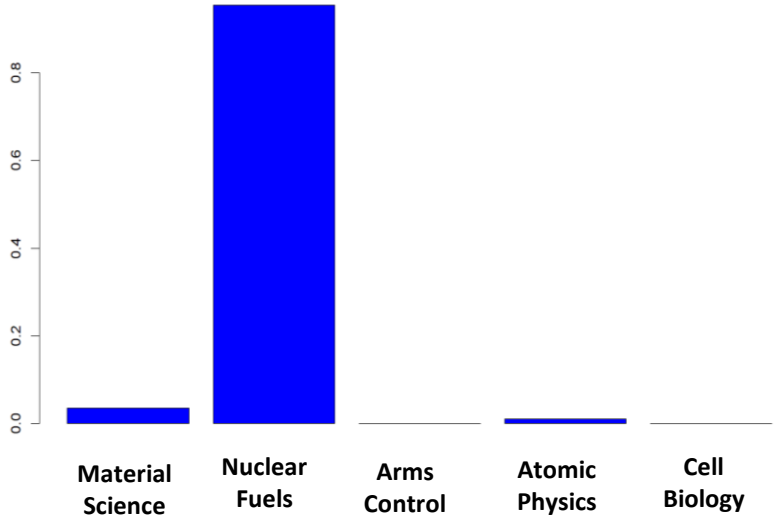


Coverage: Posterior Predictive Coverage (CVG) [2]

Posterior Distribution for $\pi(x)$



Posterior Predictive for $\hat{y}$



$\text{cred}_{0.95}(\boldsymbol{x_i})$

Interpretability: Ranking Accuracy (RA) [3]

Ranking by Similarity to Nuclear Fuels

(Nuclear Fuels, Atomic Physics, Material Science, Arms Control, Cell Biology)

VS.

Ranking by Predicted Class Probability

(Nuclear Fuels, Material Science, Atomic Physics, Arms Control, Cell Biology)

Correlated Belief-Matching

- Assumed Prior:

$$\boldsymbol{z}(\boldsymbol{x}) \sim N(\boldsymbol{0}, \sigma_0^2 \boldsymbol{P}_0)$$

- Assumed Posterior:

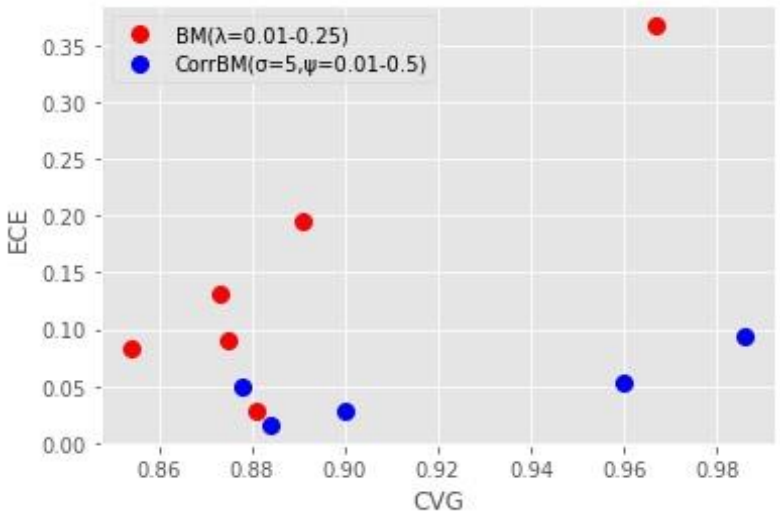
$$\boldsymbol{z}(\boldsymbol{x}) \sim \boldsymbol{N}\left(\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{W}), \begin{pmatrix} h_1(\boldsymbol{x}, \boldsymbol{W}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_K(\boldsymbol{x}, \boldsymbol{W}) \end{pmatrix} \boldsymbol{P} \begin{pmatrix} h_1(\boldsymbol{x}, \boldsymbol{W}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_K(\boldsymbol{x}, \boldsymbol{W}) \end{pmatrix}\right)$$

- Variational Inference Evidence Lower Bound:

$$\mathcal{L}_{CorrBM}(\boldsymbol{\theta}; \boldsymbol{x}, y) = -E_{q_{\boldsymbol{z}(\boldsymbol{x})}(\boldsymbol{\theta})} \log \pi_y(\boldsymbol{x}) - E_{q_{\boldsymbol{z}(\boldsymbol{x})}(\boldsymbol{\theta})} \log p\big(\boldsymbol{z}(\boldsymbol{x})\big) - \psi H(q_{\boldsymbol{z}(\boldsymbol{x})}(\boldsymbol{\theta}))$$

Small Data Experiment, Comparison with Original BM [4] and Dropout [5]

	BM	CorrBM	Dropout
Accuracy	**86.7%**	85.1%	82.4%
ECE	0.029	0.029	0.030
CVG	88.1%	90.0%	**93.7%**
RA	0.645	**0.921**	0.844
Sampling Time	0.66 minutes	**0.38 minutes**	51.9 minutes



Funding

- OUSD/R&E National Defense Education Program, BA-1, Basic Research.

- NNSA Office of Defense Nuclear Nonproliferation Research and Development, U.S. Department of Energy

References

[1] Gawlikowski, Jakob, et al. "A Survey of Uncertainty in Deep Neural Networks." arXiv preprint arXiv:2107.03342 (2021).
[2] Kompa, Benjamin, Jasper Snoek, and Andrew L. Beam. "Empirical Frequentist Coverage of Deep Learning Uncertainty Quantification Procedures." Entropy 23.12 (2021): 1608.
[3] Tong, Xuezhi, et al. "Semantic Correlations Loss: Improving Model Interpretability for Multi-Class Classification." 2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019.
[4]] Joo, Taejong, Uijung Chung, and Min-Gwan Seo. "Being Bayesian About Categorical Probability." International Conference on Machine Learning. PMLR, 2020.
[5] Gal, Yarin, and Zoubin Ghahramani. "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning." International Conference on Machine Learning. PMLR, 2016.