cases	doc_1		doc_2		decision	id
			authors	• Shreyas Padhy • Zachary Nado • Jie Ren • Jeremiah Liu • Jasper Snoek • Balaji Lakshminarayanan	DUPLICATES	139
	authors	• Shreyas Padhy • Zachary Nado • Jie Ren • Jeremiah Zhe Liu • Jasper Snoek • Balaji Lakshminarayanan	**title**	Revisiting One-vs-All Classifiers for Predictive Uncertainty and Out-of-Distribution Detection in Neural Networks		
	title	Revisiting One-vs-All Classifiers for Predictive Uncertainty and Out-of-Distribution Detection in Neural Networks	**publication_date**	2020-07-10 00:00:00		
	publication_date	2020-07-10 00:00:00	**source**	SupportedSources.INTERNET_ARCHIVE		
	source	SupportedSources.OPENALEX	**journal**			
	journal	arXiv (Cornell University)	**volume**			
	volume		**doi**			
	doi	None	**urls**	• https://web.archive.org/web/20200721104909/https://arxiv.org/pdf/2007.05134v1.pdf		
	urls	• https://openalex.org/W3041775211	**id**	id-5464527606838449503		
	id	id5324055575912532880	**abstract**	Accurate estimation of predictive uncertainty in modern neural networks is critical to achieve well calibrated predictions and detect out-of-distribution (OOD) inputs. The most promising approaches have been predominantly focused on improving model uncertainty (e.g. deep ensembles and Bayesian neural networks) and post-processing techniques for OOD detection (e.g. ODIN and Mahalanobis distance). However, there has been relatively little investigation into how the parametrization of the probabilities in discriminative classifiers affects the uncertainty estimates, and the dominant method, softmax cross-entropy, results in misleadingly high confidences on OOD data and under covariate shift. We investigate alternative ways of formulating probabilities using (1) a one-vs-all formulation to capture the notion of "none of the above", and (2) a distance-based logit representation to encode uncertainty as a function of distance to the training manifold. We show that one-vs-all formulations can improve calibration on image classification tasks, while matching the predictive performance of softmax without incurring any additional training or test-time complexity.		
	abstract		**versions**			
	versions					