

cases	doc_1		doc_2				decision	id		
			authors	<ul style="list-style-type: none">Marco PavoneNavid AzizanApoorva Sharma			DUPLICATES	111		
	title	Sketching Curvature for Efficient Out-of-Distribution Detection for Deep Neural Networks	publication_date	2021-02-24 00:00:00						
	source	SupportedSources.PAPERS_WITH_CODE	journal							
	volume		doi							
	urls	<ul style="list-style-type: none">https://arxiv.org/pdf/2102.12567v1.pdfhttps://github.com/StanfordASL/SCOD	id	id2100165434279377660						
	abstract	In order to safely deploy Deep Neural Networks (DNNs) within the perception pipelines of real-time decision making systems, there is a need for safeguards that can detect out-of-training-distribution (OoD) inputs both efficiently and accurately. Building on recent work leveraging the local curvature of DNNs to reason about epistemic uncertainty, we propose Sketching Curvature of OoD Detection (SCOD), an architecture-agnostic framework for equipping any trained DNN with a task-relevant epistemic uncertainty estimate. Offline, given a trained model and its training data, SCOD employs tools from matrix sketching to tractably compute a low-rank approximation of the Fisher information matrix, which characterizes which directions in the weight space are most influential on the predictions over the training data. Online, we estimate uncertainty by measuring how much perturbations orthogonal to these directions can alter predictions at a new test input. We apply SCOD to pre-trained networks of varying architectures on several tasks, ranging from regression to classification. We demonstrate that SCOD achieves comparable or better OoD detection performance with lower computational burden relative to existing baselines.			versions					
	versions									
	authors	<ul style="list-style-type: none">Apoorva SharmaNavid Azizan RuhiMarco Pavone								
	title	Sketching Curvature for Efficient Out-of-Distribution Detection for Deep Neural Networks								
	publication_date	2021-07-27 00:00:00								
	source	SupportedSources.OPENALEX								
	journal	Uncertainty in Artificial Intelligence								
	volume									
	doi	None								
	urls	<ul style="list-style-type: none">https://openalex.org/W3186504155								
id	id4320951299594195266									
abstract										
versions										