Novelty Assessment Report

Paper: Contextual Similarity Distillation: Ensemble Uncertainties with a Single Model
PDF URL: https://openreview.net/pdf?id=arms7s9dDK
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-01

Abstract

Uncertainty quantification is a critical aspect of reinforcement learning and deep learning, with numerous applications ranging from efficient exploration and stable offline reinforcement learning to outlier detection in medical diagnostics. The scale of modern neural networks, however, complicates the use of many theoretically well-motivated approaches such as full Bayesian inference. Approximate methods like deep ensembles can provide reliable uncertainty estimates but still remain computationally expensive. In this work, we propose contextual similarity distillation, a novel approach that explicitly estimates the variance of an ensemble of deep neural networks with a single model, without ever learning or evaluating such an ensemble in the first place. Our method builds on the predictable learning dynamics of wide neural networks, governed by the neural tangent kernel, to derive an efficient approximation of the predictive variance of an infinite ensemble. Specifically, we reinterpret the computation of ensemble variance as a supervised regression problem with kernel similarities as regression targets. The resulting model can estimate predictive variance at inference time with a single forward pass, and can make use of unlabeled target-domain data or data augmentations to refine its uncertainty estimates. We empirically validate our method across a variety of out-of-distribution detection benchmarks and sparse-reward reinforcement learning environments. We find that our single-model method performs competitively and sometimes superior to ensemble-based baselines and serves as a reliable signal for efficient exploration. These results, we believe, position contextual similarity distillation as a principled and scalable alternative for uncertainty quantification in reinforcement learning and general deep learning.

Core Task Landscape

This paper addresses: **Uncertainty Quantification in Deep Neural Networks**

A total of **50 papers** were analyzed and organized into a taxonomy with **24 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Foundational Frameworks and Comprehensive Reviews**
- **Bayesian and Probabilistic Approaches**
- **Ensemble-Based Uncertainty Quantification**
- **Aleatoric Uncertainty Modeling**
- **Joint Epistemic and Aleatoric Uncertainty Quantification**
- **Deterministic and Single-Pass Uncertainty Estimation**
- **Calibration and Confidence Refinement**
- **Distribution Shift and Out-of-Distribution Detection**
- **Multi-Fidelity and Surrogate Modeling**
- **Specialized Architectures and Model Classes**
- ... and 2 more categories

Complete Taxonomy Tree

- Uncertainty Quantification in Deep Neural Networks Survey Taxonomy
- Foundational Frameworks and Comprehensive Reviews
  - General Survey Literature (4 papers)
  - [1] A survey of uncertainty in deep neural networks (Gawlikowski, 2023) View paper
  - [2] Uncertainty quantification for deep learning (van Leeuwen, 2025) View paper
  - [3] Uncertainty quantification in deep learning (Stantchev, 2023) View paper
  - [5] A survey on uncertainty quantification methods for deep learning (He, 2025) View paper
  - Statistical and Theoretical Foundations (2 papers)
  - [12] Quantifying epistemic uncertainty in deep learning (Ziyi Huang, 2021) View paper
  - [25] Quantification of uncertainties in neural networks (Xinyang Wu, 2023) View paper
- Bayesian and Probabilistic Approaches
  - Bayesian Neural Networks and Variational Methods (4 papers)
  - [8] Uncertainty quantification using Bayesian neural networks in classification: Application to biomedical image segmentation (Yongchan Kwon, 2020) View paper
  - [11] Bayesian optimized deep ensemble for uncertainty quantification of deep neural networks: a system safety case study on sodium fast reactor thermal stratification â¦ (Z Abulawi, 2025) View paper
  - [23] Uncertainty quantification of deep neural network-based turbulence model for reactor transient analysis (Yang Liu, 2021) View paper
  - [43] Modeling epistemic and aleatoric uncertainty with bayesian neural networks and latent variables (Stefan Depeweg, 2019) View paper
  - Stochastic Regularization Techniques (2 papers)

- [18] Uncertainty Quantification in Deep Residual Neural Networks (Wandzik Lukasz, 2022) View paper
    - [30] Adaptive sparse dropout: Learning the certainty and uncertainty in deep neural networks (Yuanyuan Chen, 2021) View paper
    - Neural Stochastic Differential Equations (2 papers)
    - [27] Neural SDE-Based Epistemic Uncertainty Quantification in Deep Neural Networks (Aabila Tharzeen, 2024) View paper
    - [39] Sde-net: Equipping deep neural networks with uncertainty estimates (Lingkai Kong, 2020) View paper
- Ensemble-Based Uncertainty Quantification
    - Deep Ensemble Methods (2 papers)
    - [15] Uncertainty Quantification for Traffic Forecasting Using Deep-Ensemble-Based Spatiotemporal Graph Neural Networks (Tanwi Mallick, 2024) View paper
    - [26] Uncertainty quantification of spectral predictions using deep neural networks. (Sneha Verma, 2023) View paper
    - Ensemble Distillation and Approximation ★ (2 papers)
    - [0] Contextual Similarity Distillation: Ensemble Uncertainties with a Single Model (Anon et al., 2026) View paper
    - [33] Estimating Epistemic and Aleatoric Uncertainty with a Single Model (Matthew Chan, 2024) View paper
- Aleatoric Uncertainty Modeling (2 papers)
    - [14] One step closer to unbiased aleatoric uncertainty estimation (Zhang Wang, 2023) View paper
    - [32] Test-time data augmentation for estimation of heteroscedastic aleatoric uncertainty in deep neural networks (Murat SeÃ§kin Ayhan, 2018) View paper
- Joint Epistemic and Aleatoric Uncertainty Quantification (4 papers)
    - [13] A Bayesian deep learning RUL framework integrating epistemic and aleatoric uncertainties (Gaoyang Li, 2020) View paper
    - [16] A general framework for quantifying aleatoric and epistemic uncertainty in graph neural networks (Sai Munikoti, 2023) View paper
    - [22] Bayesian Neural Networks for Satellite Fog Detection: Quantifying Epistemic and Aleatoric Uncertainties (Prasad Deshpande, 2024) View paper
    - [24] Remaining Useful Life Prediction Accounting for Epistemic and Aleatoric Uncertainties (Wei-Jun Xu, 2024) View paper
- Deterministic and Single-Pass Uncertainty Estimation
    - Depth and Architectural Uncertainty (2 papers)
    - [6] Depth uncertainty in neural networks (Javier AntorÃ¡n, 2020) View paper
    - [45] Probabilistic Skip Connections for Deterministic Uncertainty Quantification in Deep Neural Networks (Jimenez Felix, 2025) View paper
    - Feature-Based and Latent Space Methods (3 papers)
    - [41] Quantifying aleatoric and epistemic uncertainty using density estimation in latent space (Postels, 2020) View paper
    - [44] Deep k-nearest neighbors: Towards confident, interpretable and robust deep learning (Papernot, 2018) View paper
    - [46] Predicting neural network confidence using high-level feature distance (Jie Wang, 2023) View paper
    - Learned Confidence and Failure Prediction (1 papers)
    - [38] Addressing failure prediction by learning model confidence (CorbiÃ¨re, 2019) View paper
- Calibration and Confidence Refinement
    - Training-Time Calibration Enhancement (1 papers)
    - [28] On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks (Thulasidasan, 2019) View paper
    - Calibration Evaluation and Benchmarking (3 papers)
    - [21] Being confident in confidence scores: calibration in deep learning models for camera trap image sequences (Gaspard Dussert, 2024) View paper
    - [37] Quantitative evaluation of confidence measures in a machine learning world (Matteo Poggi, 2017) View paper
    - [50] A Survey on Confidence Calibration of Deep Learning-Based Classification Models Under Class Imbalance Data (Jinzong Dong, 2025) View paper
- Distribution Shift and Out-of-Distribution Detection (1 papers)
    - [19] Predicting with confidence on unseen distributions (Devin Guillory, 2021) View paper
- Multi-Fidelity and Surrogate Modeling (2 papers)
    - [4] Aleatory uncertainty quantification based on multi-fidelity deep neural networks (Zhihui Li, 2024) View paper
    - [17] Uncertainty quantification analysis using multi-fidelity deep neural network (Luogeng Lv, 2024) View paper
- Specialized Architectures and Model Classes (2 papers)
    - [35] Multidimensional Uncertainty Quantification for Deep Neural Networks (Zhao, 2023) View paper
    - [47] Density regression and uncertainty quantification with Bayesian deep noise neural networks (Daiwei Zhang, 2023) View paper
- Uncertainty Propagation and Forward Analysis (1 papers)
    - [7] Uncertainty propagation through deep neural networks (Abdelaziz, 2015) View paper
- Domain-Specific Applications
    - Medical and Healthcare Applications (6 papers)
    - [29] Advancing EEG prediction with deep learning and uncertainty estimation (Mats Tveter, 2024) View paper
    - [31] Alzheimer's Disease Classification Confidence of Individuals using Deep Learning on Heterogeneous Data (Afolabi Salami Alausa, 2024) View paper
    - [34] Uncertainty Estimation in Deep Neural Networks for Dermoscopic Image Classification (M. Combalia, 2020) View paper
    - [42] Explainable AI for raising confidence in deep learning-based tumor tracking models. (D. Grama, 2025) View paper
    - [48] A Reliable Deep Learning Model for ECG Interpretation: Mitigating Overconfidence and Direct Uncertainty Quantification (Xuedong Li, 2025) View paper
    - [49] Uncertainty quantification in neural-network based pain intensity estimation. (Burcu Ozek, 2024) View paper
    - Earth Observation and Environmental Monitoring (1 papers)
    - [10] Uncertainty quantification in deep neural networks for multi-sensor Earth observation (Shubhankar Roy, 2025) View paper
    - Physical Sciences and Engineering Simulation (1 papers)
    - [40] Joint learning equation of state surfaces with uncertainty-aware physically regularized neural networks (Dongyang Kuang, 2025) View paper
    - Spatiotemporal and Sequential Prediction (1 papers)
    - [20] Enhancing Maritime Safety: Estimating Collision Probabilities with Trajectory Prediction Boundaries Using Deep Learning Models (Robertas Jurkus, 2025) View paper

- ◦ Molecular and Chemical Property Prediction (1 papers)
- ◦ [36] Evaluating scalable uncertainty estimation methods for deep learning-based molecular property prediction (Gabriele Scalia, 2020) View paper
- ◦ Natural Language Processing (1 papers)
- ◦ [9] A survey of confidence estimation and calibration in large language models (Jiahui Geng, 2024) View paper

Narrative

Core task: uncertainty quantification in deep neural networks. The field has matured into a rich landscape organized around several complementary strategies. Foundational frameworks and comprehensive reviews (e.g., Survey Uncertainty Deep Networks[1], Uncertainty Quantification Deep Learning[2][3], Survey Uncertainty Quantification Methods[5]) establish the conceptual underpinnings, distinguishing epistemic uncertainty (model ignorance) from aleatoric uncertainty (inherent data noise). Bayesian and probabilistic approaches leverage posterior inference to capture model uncertainty, while ensemble-based methods aggregate predictions from multiple models or training runs to estimate epistemic confidence. Deterministic and single-pass techniques offer computational efficiency by extracting uncertainty from a single forward pass, and calibration methods refine raw confidence scores to align with empirical accuracy. Specialized branches address distribution shift and out-of-distribution detection, multi-fidelity modeling that fuses information across data sources (e.g., Aleatory Multi-Fidelity[4], Multi-Fidelity Analysis[17]), and domain-specific applications spanning medical imaging, autonomous systems, and scientific computing.

Within the ensemble-based branch, a central tension exists between the high fidelity of full ensembles and the computational cost they impose at inference time. Ensemble distillation and approximation methods seek to compress ensemble knowledge into a single, efficient model while preserving uncertainty estimates. Contextual Similarity Distillation[0] exemplifies this direction by distilling ensemble predictions through contextual similarity mechanisms, aiming to retain predictive diversity without maintaining multiple networks. This contrasts with Single Model Estimation[33], which pursues uncertainty from a lone model via architectural or training innovations, and with test-time augmentation strategies (Test-Time Augmentation[32]) that generate pseudo-ensembles on the fly. The original work thus occupies a pragmatic middle ground: it inherits the representational richness of ensembles yet targets deployment scenarios where resource constraints favor a streamlined architecture.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Estimating Epistemic and Aleatoric Uncertainty with a Single Model

Authors: Matthew Chan, Christopher Metzler, M. A. Chan, MarÃa Molina, Maria J. Molina, et al. (6 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

Estimating and disentangling epistemic uncertainty, uncertainty that is reducible with more training data, and aleatoric uncertainty, uncertainty that is inherent to the task at hand, is critically important when applying machine learning to high-stakes applications such as medical imaging and weather forecasting. Conditional diffusion models' breakthrough ability to accurately and efficiently sample from the posterior distribution of a dataset now makes uncertainty estimation conceptually strai...

Relationship Analysis

Both papers belong to the Ensemble Distillation and Approximation category, focusing on efficiently approximating ensemble behavior with reduced computational cost. They overlap in their goal of estimating ensemble uncertainty without training full ensembles, both leveraging theoretical frameworks (NTK for the original paper, diffusion models for the candidate). The key difference is that the original paper uses contextual similarity distillation based on NTK Gaussian Processes to directly estimate ensemble variance through a supervised regression approach, while the candidate paper employs Bayesian hyper-networks to generate diffusion model weights, enabling joint estimation of both epistemic and aleatoric uncertainty through a generative modeling framework.

Contributions Analysis

Overall novelty summary. The paper proposes contextual similarity distillation (CSD), a method to estimate ensemble variance using a single model without training or evaluating the ensemble. It resides in the 'Ensemble Distillation and Approximation' leaf, which contains only two papers including this one. This sparse population suggests the specific approach of distilling ensemble variance via kernel-based regression targets is relatively underexplored. The taxonomy shows ensemble-based uncertainty quantification is a well-established branch, but the distillation subfield remains narrow compared to broader Bayesian or deterministic categories.

The taxonomy reveals neighboring leaves include 'Deep Ensemble Methods' (full ensemble training) and 'Bayesian Neural Networks' (probabilistic weight inference). CSD bridges these directions by leveraging neural tangent kernel theory—typically associated with theoretical foundations—to approximate ensemble behavior without Bayesian sampling or multiple training runs. The 'Deterministic and Single-Pass Uncertainty Estimation' branch offers alternative efficiency strategies (feature-based metrics, learned confidence), but CSD's kernel-similarity regression formulation diverges by explicitly targeting ensemble variance rather than implicit confidence proxies. This positioning suggests the work synthesizes theoretical insights with practical ensemble approximation goals.

Among 25 candidates examined, the theoretical framework based on neural tangent kernel shows overlap with prior work (3 refutable candidates out of 10 examined for this contribution). The CSD method itself and the contextualized regression formulation appear more novel within this limited search scope (0 refutable candidates across 15 examined). The statistics indicate that while the kernel-theoretic foundation connects to existing literature, the specific distillation mechanism and unlabeled-data regression strategy have less direct precedent among the top-25 semantically similar papers. This pattern suggests incremental theoretical grounding combined with a more distinctive methodological contribution.

Based on the limited search scope (25 candidates from semantic retrieval), the work appears to occupy a sparsely populated niche within ensemble approximation. The taxonomy context confirms that distillation-based ensemble compression is less crowded than full ensemble or Bayesian methods. However, the analysis does not cover exhaustive citation networks or domain-specific ensemble literature, so the novelty assessment remains provisional. The kernel-theoretic overlap suggests the work builds on established theory while introducing a new application pathway for variance estimation.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Contextual Similarity Distillation (CSD) method

Description: The authors introduce a method that approximates the predictive variance of an infinite ensemble of neural networks using only a single model. CSD reframes ensemble variance computation as a supervised regression problem where labels correspond to kernel similarities, enabling efficient uncertainty quantification without training multiple models.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Deep confidence: a computationally efficient framework for calculating reliable prediction errors for deep neural networks

URL: View paper

Brief Assessment

Deep Confidence[59] focuses on computing confidence intervals for predictions using snapshot ensembles and conformal prediction in drug discovery applications, not on approximating ensemble variance through kernel similarity regression as in the original paper's CSD method.

---

2. The diversified ensemble neural network

URL: View paper

Brief Assessment

Diversified Ensemble[51] focuses on training multiple neural network modules jointly with diversity constraints for ensemble learning, not on approximating ensemble variance with a single model through kernel similarity regression as in the original paper's CSD method.

---

3. ST-TransNet: A Spatiotemporal Transformer Network for Uncertainty Estimation from a Single Deterministic Precipitation Forecast

URL: View paper

Brief Assessment

ST-TransNet[55] addresses precipitation forecast uncertainty estimation from ensemble forecasts using spatiotemporal transformers, not ensemble variance approximation with a single neural network model through kernel similarity regression as in the original paper's CSD method.

---

4. Ensemble solar forecasting and post-processing using dropout neural network and information from neighboring satellite pixels

URL: View paper

Brief Assessment

Ensemble Solar Forecasting[56] uses dropout neural networks with Monte Carlo sampling for ensemble forecasting, which is fundamentally different from CSD's approach of approximating ensemble variance through kernel similarity regression without requiring multiple models or stochastic sampling.

---

5. Probabilistic binary neural networks

URL: View paper

Brief Assessment

Probabilistic Binary Networks[57] focuses on training binary neural networks with stochastic weights and activations for resource-constrained environments, not on estimating ensemble variance with a single model for uncertainty quantification.

---

6. Estimating Epistemic and Aleatoric Uncertainty with a Single Model

URL: View paper

Brief Assessment

Single Model Estimation[33] focuses on uncertainty quantification in diffusion models using Bayesian hyper-networks, not on approximating ensemble variance through kernel similarity regression as in CSD. The technical approaches and problem formulations are fundamentally different.

---

7. Deep ensembles work, but are they necessary?

URL: View paper

Brief Assessment

Deep Ensembles Necessary[54] focuses on comparing deep ensembles with single larger models for uncertainty quantification, not on approximating ensemble variance through kernel similarity regression as CSD does.

---

8. Single-model uncertainties for deep learning

URL: View paper

Brief Assessment

Single-Model Uncertainties[53] focuses on estimating aleatoric and epistemic uncertainty through simultaneous quantile regression and orthonormal certificates, not on approximating ensemble variance through kernel similarity regression as in CSD.

---

9. Prune and tune ensembles: low-cost ensemble learning with sparse independent subnetworks

URL: View paper

Brief Assessment

Prune and Tune[58] focuses on creating diverse ensembles through pruning and fine-tuning subnetworks, not on approximating ensemble variance with a single model using kernel similarities as the original paper does.

---

10. Uncertainty estimation using a single deep deterministic neural network

URL: View paper

Brief Assessment

Single Deep Deterministic[52] focuses on RBF networks with gradient penalties for OOD detection in classification tasks, not on approximating ensemble variance through kernel similarity regression as in CSD.

Contribution 2: Theoretical framework based on Neural Tangent Kernel

Description: The authors develop a theoretical foundation grounded in the Neural Tangent Kernel (NTK) theory to derive an analytical expression for ensemble uncertainties. This framework characterizes deep ensembles through the NTK Gaussian Process and enables the derivation of their single-model approximation method.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Epistemic uncertainty and observation noise with the neural tangent kernel

URL: View paper

Brief Assessment

Epistemic Observation Noise[68] focuses on computing posterior mean and covariance for NTK-GPs with non-zero aleatoric noise, while the original paper uses NTK theory to derive ensemble variance approximations for uncertainty quantification in RL. The candidate addresses a different problem (GP inference with observation noise) rather than ensemble uncertainty distillation.

---

2. Deep Learning for High-Dimensional Decision Making and Uncertainty Quantification

URL: View paper

Brief Assessment

High-Dimensional Decision Making[70] only briefly mentions NTK in passing as a regression method, without developing any theoretical framework for ensemble uncertainty quantification or deriving analytical expressions for ensemble variance.

---

3. Fed-ensemble: Ensemble Models in Federated Learning for Improved Generalization and Uncertainty Quantification

URL: View paper

Brief Assessment

[Final Audit Failure] The model insisted on a refutation claim but failed to provide verifiable evidence after multiple retries. Marked as cannot_refute for safety. Please manually verify the candidate text.

---

4. Disentangling the Predictive Variance of Deep Ensembles through the Neural Tangent Kernel

URL: View paper

Prior Art Analysis

Disentangling Predictive Variance[74] demonstrates that prior work exists on using Neural Tangent Kernel theory to derive analytical expressions for ensemble uncertainties. The candidate paper develops a theoretical framework grounded in NTK theory to characterize deep ensembles through the NTK Gaussian Process and derives analytical expressions for predictive variance. Both papers leverage the same theoretical foundation (NTK theory by Jacot et al. 2018 and Lee et al. 2020) to analyze ensemble behavior, and both derive analytical expressions for ensemble variance using the NTK framework. The candidate paper explicitly provides first-order approximations and decompositions of predictive variance terms, demonstrating that this theoretical approach was established before the original paper's submission.

Evidence

Evidence 1 - **Rationale**: Both papers explicitly state they build on NTK theory to understand ensemble predictive distributions, establishing that this theoretical approach was already developed in the candidate paper. - **Original**: our method builds on the predictable learning dynamics of wide neural networks, governed by the neural tangent kernel, to derive an efficient approximation of the predictive variance of an infinite ensemble - **Candidate**: we leverage this tractable description of trained neural networks and take a first step towards understanding the predictive distribution of neural networks ensembles with large but finite width. building on top of the various studies mentioned, we do so by studying the case where these networks can b...

---

5. Bayesian deep ensembles via the neural tangent kernel

URL: View paper

Prior Art Analysis

Bayesian Deep Ensembles[65] demonstrates that prior work exists establishing theoretical frameworks for ensemble uncertainties using Neural Tangent Kernel theory. The candidate paper explicitly derives analytical expressions for ensemble variance through the NTK Gaussian Process framework, characterizing deep ensembles through NTK theory. Both papers leverage the same theoretical foundation from Jacot et al. (2018) and Lee et al. (2020) to derive analytical expressions for ensemble uncertainties, with the candidate paper providing the mathematical framework (equations 6-7) that the original paper builds upon. The candidate's work on characterizing ensemble variance through NTK GP (equation 8) and deriving posterior predictive distributions predates and establishes the theoretical groundwork that the original paper claims as novel.

Evidence

Evidence 1 - **Rationale**: Both papers explicitly use NTK theory to derive analytical expressions for ensemble uncertainties. The candidate establishes the theoretical framework for characterizing deep ensembles through NTK in the infinite width limit. - **Original**: our method builds on the predictable learning dynamics of wide neural networks, governed by the neural tangent kernel, to derive an efficient approximation of the predictive variance of an infinite ensemble - **Candidate**: previous work has shown that even in the infinite width limit, when nns become gps, there is no gp posterior interpretation to a deep ensemble trained with squared error loss. we introduce a simple modification to standard deep ensembles training, through addition of a computationally-tractable, rando...

Evidence 2 - **Rationale**: The candidate paper derives the analytical expression for ensemble variance through NTK GP framework, establishing the mathematical foundation that the original paper uses. The candidate's equation (8) provides the posterior predictive distribution that characterizes ensemble uncertainties. - **Original**: the variance of an ensemble over infinite random initializations is given by v[f(x, θ∞)] = θ(x, x)-θ(x,x)θ(x,x) -1θ(x, x) - **Candidate**: thus, ˜f1(x0) d ⇠n (⇥x0x ⇥ 1 xx y, ⇥x0x0 ⇥x0x ⇥ 1 xx ⇥xx 0) on a test setx0, in the infinite width limit. note that eq. (8) is the gp posterior using prior kernel⇥ and noiseless observations˜f1(x)=y, which we will refer to as thentkgp posterior predictive

---

6. Universal Value-Function Uncertainties

URL: View paper

Brief Assessment

Universal Value-Function[73] applies NTK theory to temporal difference learning in reinforcement learning, deriving closed-form solutions for value function ensembles trained with TD losses. The original paper applies NTK theory to supervised learning with squared losses to characterize deep ensemble uncertainties. These are distinct problem domains with different theoretical challenges.

---

7. Single Model Uncertainty Estimation via Stochastic Data Centering

URL: View paper

Brief Assessment

Stochastic Data Centering[72] uses NTK theory to analyze how domain shifts affect the kernel, focusing on shift-invariance properties. The original paper uses NTK to characterize ensemble uncertainties through Gaussian Process interpretation and derives single-model approximations via similarity regression, which is a fundamentally different approach.

---

8. Uncertainty quantification from ensemble variance scaling laws in deep neural networks

URL: View paper

Brief Assessment

Ensemble Variance Scaling[67] focuses on computing ensemble variance statistics (mean and variance of test loss) in the infinite-width NTK limit for uncertainty quantification, rather than developing methods to estimate ensemble variance with a single model as in the original paper.

9. No-regret bandit exploration based on soft tree ensemble model

URL: View paper

Brief Assessment

Soft Tree Ensemble[69] focuses on tree-based bandit algorithms using Tree Neural Tangent Kernel (TNTK) for soft tree ensembles, not on ensemble uncertainty quantification or deriving single-model approximations from NTK theory as in the original paper.

10. Uncertainty quantification with the empirical neural tangent kernel

URL: View paper

Prior Art Analysis

Empirical Neural Tangent[66] demonstrates that prior work exists establishing theoretical frameworks connecting Neural Tangent Kernel theory to uncertainty quantification in neural networks. Both papers develop theoretical foundations grounded in NTK theory to characterize deep ensembles and derive analytical expressions for ensemble uncertainties. The candidate paper explicitly shows that under certain conditions, their method generates samples from a Gaussian Process with an empirical NTK kernel, providing a theoretical connection between neural networks, Gaussian processes, and the NTK for uncertainty quantification - a framework that predates the original paper's contribution.

Evidence

Evidence 1 - **Rationale**: Both papers develop methods to estimate predictive variance/uncertainty using neural network training approaches, with the candidate demonstrating a prior sampling-based method for uncertainty quantification. - **Original**: the predictive variance of a deep ensemble for a known query point xt by training a single nn on a regression task using gradient descent and a carefully designed label function dependent on xt - **Candidate**: we propose a monte-carlo sampling based uq method to approximate the predictive distribution of nns, called neural uncertainty quantification by linearized sampling (nuqls). our method is lightweight, post-hoc, numerically stable, and embarrassingly parallel.

Evidence 2 - **Rationale**: Both papers establish that neural networks can be characterized as Gaussian processes with NTK kernels, demonstrating the same theoretical foundation for uncertainty quantification through NTK-GP frameworks. - **Original**: further extending this framework, he et al. (2020) demonstrate that by introducing suitable function priors on f(x, θ_0), akin to the well-known randomized prior functions by osband et al. (2019), the post-training function is described by a gaussian process (gp, rasmussen and williams, 2006): f(xt, ... - **Candidate**: therefore, f($\theta_\star$ z,x) follows a gp with an ntk kernel. conditioning on f($\theta_\star$ z,x) =y is reasonable, since by construction f($\theta_\star$ z,x)≈ ef($\theta_\star$ z,x) , and under the full-rank assumption of jx , we have ef($\theta_\star$ z,x) =y.

Contribution 3: Contextualized regression formulation with unlabeled data

Description: The authors formulate a contextualized regression model that extends their approach to work efficiently for arbitrary query points. This formulation enables the method to leverage unlabeled data from target domains or data augmentations to improve uncertainty estimates, a capability not easily incorporated in standard ensemble methods.

This contribution was assessed against **5 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Semiâ□□Supervised Hybrid Local Kernel Regression for Soft Sensor Modelling of Rubberâ□□Mixing Process

URL: View paper

Brief Assessment

Rubber-Mixing Process[64] focuses on semi-supervised soft sensor modeling for chemical processes using kernel regression with time-windows, not on uncertainty quantification or ensemble variance estimation in deep learning contexts.

2. Semi-supervised Deep Kernel Learning: Regression with Unlabeled Data by Minimizing Predictive Variance

URL: View paper

Brief Assessment

Semi-Supervised Deep Kernel[62] focuses on semi-supervised regression using Gaussian processes with unlabeled data to minimize predictive variance, while the original paper addresses ensemble variance estimation through contextualized regression with kernel similarities. The technical approaches and problem formulations differ fundamentally.

3. 3d semi-supervised learning with uncertainty-aware multi-view co-training

URL: View paper

Brief Assessment

Semi-Supervised Multi-View[60] focuses on 3D medical image segmentation using multi-view co-training with uncertainty estimation, not on contextualized regression for ensemble variance estimation or kernel similarity methods.

4. TSceneJAL: Joint Active Learning of Traffic Scenes for 3D Object Detection

URL: View paper

Brief Assessment

TSceneJAL[61] focuses on active learning for 3D object detection in autonomous driving, using mixture density networks for uncertainty estimation. It does not address contextualized regression with kernel similarity or ensemble variance estimation through unlabeled data refinement as described in the original paper.

5. A Physics-Informed Multiview Collaborative Semisupervised Framework for Battery Lifespan Early Prediction

URL: View paper

Brief Assessment

Battery Lifespan Prediction[63] focuses on battery degradation prediction using physics-informed Gaussian processes with domain-specific features (aging modes, phase transitions), not general uncertainty quantification or kernel similarity regression for neural networks.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Contextual Similarity Distillation: Ensemble Uncertainties with a Single Model View paper
- [1] A survey of uncertainty in deep neural networks View paper
- [2] Uncertainty quantification for deep learning View paper
- [3] Uncertainty quantification in deep learning View paper
- [4] Aleatory uncertainty quantification based on multi-fidelity deep neural networks View paper
- [5] A survey on uncertainty quantification methods for deep learning View paper
- [6] Depth uncertainty in neural networks View paper
- [7] Uncertainty propagation through deep neural networks View paper
- [8] Uncertainty quantification using Bayesian neural networks in classification: Application to biomedical image segmentation View paper
- [9] A survey of confidence estimation and calibration in large language models View paper
- [10] Uncertainty quantification in deep neural networks for multi-sensor Earth observation View paper
- [11] Bayesian optimized deep ensemble for uncertainty quantification of deep neural networks: a system safety case study on sodium fast reactor thermal stratification â¦ View paper
- [12] Quantifying epistemic uncertainty in deep learning View paper
- [13] A Bayesian deep learning RUL framework integrating epistemic and aleatoric uncertainties View paper
- [14] One step closer to unbiased aleatoric uncertainty estimation View paper
- [15] Uncertainty Quantification for Traffic Forecasting Using Deep-Ensemble-Based Spatiotemporal Graph Neural Networks View paper
- [16] A general framework for quantifying aleatoric and epistemic uncertainty in graph neural networks View paper
- [17] Uncertainty quantification analysis using multi-fidelity deep neural network View paper
- [18] Uncertainty Quantification in Deep Residual Neural Networks View paper
- [19] Predicting with confidence on unseen distributions View paper
- [20] Enhancing Maritime Safety: Estimating Collision Probabilities with Trajectory Prediction Boundaries Using Deep Learning Models View paper
- [21] Being confident in confidence scores: calibration in deep learning models for camera trap image sequences View paper
- [22] Bayesian Neural Networks for Satellite Fog Detection: Quantifying Epistemic and Aleatoric Uncertainties View paper
- [23] Uncertainty quantification of deep neural network-based turbulence model for reactor transient analysis View paper
- [24] Remaining Useful Life Prediction Accounting for Epistemic and Aleatoric Uncertainties View paper
- [25] Quantification of uncertainties in neural networks View paper
- [26] Uncertainty quantification of spectral predictions using deep neural networks. View paper
- [27] Neural SDE-Based Epistemic Uncertainty Quantification in Deep Neural Networks View paper
- [28] On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks View paper
- [29] Advancing EEG prediction with deep learning and uncertainty estimation View paper
- [30] Adaptive sparse dropout: Learning the certainty and uncertainty in deep neural networks View paper
- [31] Alzheimer's Disease Classification Confidence of Individuals using Deep Learning on Heterogeneous Data View paper
- [32] Test-time data augmentation for estimation of heteroscedastic aleatoric uncertainty in deep neural networks View paper
- [33] Estimating Epistemic and Aleatoric Uncertainty with a Single Model View paper
- [34] Uncertainty Estimation in Deep Neural Networks for Dermoscopic Image Classification View paper
- [35] Multidimensional Uncertainty Quantification for Deep Neural Networks View paper
- [36] Evaluating scalable uncertainty estimation methods for deep learning-based molecular property prediction View paper
- [37] Quantitative evaluation of confidence measures in a machine learning world View paper
- [38] Addressing failure prediction by learning model confidence View paper
- [39] Sde-net: Equipping deep neural networks with uncertainty estimates View paper
- [40] Joint learning equation of state surfaces with uncertainty-aware physically regularized neural networks View paper
- [41] Quantifying aleatoric and epistemic uncertainty using density estimation in latent space View paper
- [42] Explainable AI for raising confidence in deep learning-based tumor tracking models. View paper
- [43] Modeling epistemic and aleatoric uncertainty with bayesian neural networks and latent variables View paper
- [44] Deep k-nearest neighbors: Towards confident, interpretable and robust deep learning View paper
- [45] Probabilistic Skip Connections for Deterministic Uncertainty Quantification in Deep Neural Networks View paper
- [46] Predicting neural network confidence using high-level feature distance View paper
- [47] Density regression and uncertainty quantification with Bayesian deep noise neural networks View paper
- [48] A Reliable Deep Learning Model for ECG Interpretation: Mitigating Overconfidence and Direct Uncertainty Quantification View paper
- [49] Uncertainty quantification in neural-network based pain intensity estimation. View paper
- [50] A Survey on Confidence Calibration of Deep Learning-Based Classification Models Under Class Imbalance Data View paper
- [51] The diversified ensemble neural network View paper
- [52] Uncertainty estimation using a single deep deterministic neural network View paper
- [53] Single-model uncertainties for deep learning View paper
- [54] Deep ensembles work, but are they necessary? View paper
- [55] ST-TransNet: A Spatiotemporal Transformer Network for Uncertainty Estimation from a Single Deterministic Precipitation Forecast View paper
- [56] Ensemble solar forecasting and post-processing using dropout neural network and information from neighboring satellite pixels View paper
- [57] Probabilistic binary neural networks View paper
- [58] Prune and tune ensembles: low-cost ensemble learning with sparse independent subnetworks View paper
- [59] Deep confidence: a computationally efficient framework for calculating reliable prediction errors for deep neural networks View paper
- [60] 3d semi-supervised learning with uncertainty-aware multi-view co-training View paper
- [61] TSceneJAL: Joint Active Learning of Traffic Scenes for 3D Object Detection View paper
- [62] Semi-supervised Deep Kernel Learning: Regression with Unlabeled Data by Minimizing Predictive Variance View paper

- [63] A Physics-Informed Multiview Collaborative Semisupervised Framework for Battery Lifespan Early Prediction View paper
- [64] Semiâ□□Supervised Hybrid Local Kernel Regression for Soft Sensor Modelling of Rubberâ□□Mixing Process View paper
- [65] Bayesian deep ensembles via the neural tangent kernel View paper
- [66] Uncertainty quantification with the empirical neural tangent kernel View paper
- [67] Uncertainty quantification from ensemble variance scaling laws in deep neural networks View paper
- [68] Epistemic uncertainty and observation noise with the neural tangent kernel View paper
- [69] No-regret bandit exploration based on soft tree ensemble model View paper
- [70] Deep Learning for High-Dimensional Decision Making and Uncertainty Quantification View paper
- [71] Fed-ensemble: Ensemble Models in Federated Learning for Improved Generalization and Uncertainty Quantification View paper
- [72] Single Model Uncertainty Estimation via Stochastic Data Centering View paper
- [73] Universal Value-Function Uncertainties View paper
- [74] Disentangling the Predictive Variance of Deep Ensembles through the Neural Tangent Kernel View paper