

**38th Annual Conference of Neural Information Processing Systems  
(NeurIPS)  
10-15 December 2024**

**Fact Sheet**

**2024 Location:** British Columbia, Canada, The Vancouver Convention Centre

**Registration numbers:**

- Total: 19,756
- In-person registrations: 16,777
- Virtual registrations: 2978
- In-person registration was up 27% (the lottery prevented this from being even higher) overall registration was up 21% while virtual registrations were down 4%

**Past attendance and location:**

- 16,382 Hybrid - 13,307 in-person, 3075 virtual - New Orleans, Louisiana, United States 2023
- 15,390 Hybrid - 9,835 in-person and 5,555 virtual - New Orleans, Louisiana, United States 2022
- 17,091 Virtual Conference 2021
- 22,823 Virtual Conference 2020
- 13,000 Vancouver, British Columbia, Canada 2019
- 8,648 Montreal, Quebec, Canada 2018
- 8,008 Long Beach, California, United States 2017
- 5,231 Barcelona, Spain 2016
- 3,852 Montreal, Quebec, Canada 2015
- 2,581 Montreal, Quebec, Canada 2014
- 1,994 Lake Tahoe, California, United States 2013
- 1,676 Lake Tahoe, California, United States 2012
- 1,452 Granada, Spain 2011
- 1,354 Vancouver, British Columbia, Canada 2010

2025 Location: San Diego California, United States

**2024 Program:**

- 11 conference tracks
  - 6 [Creative AI](#)
  - 7 [Invited Talks](#)
  - 8 [Socials](#)
  - 11 [Affinity Events](#)
  - 14 [Tutorials](#)
  - 16 [Competitions](#)
  - 56 [Workshops](#)
  - 69 [Oral Sessions](#)
  
- Paper reviews
  - 4497 total accepted combined papers
    - 4,037 main conference track
    - 460 datasets and benchmarks track
  - 17,491 total submissions
    - 15,671 main conference track
    - 1,820 datasets and benchmarks (about double the previous year's 987 submissions)
  - Paper acceptance rate:
    - 25.8% main conference track
    - 25.3% datasets and benchmark
  - Reviewers
    - 1,393 main conference Area Chairs
    - 195 main conference Senior Area Chairs
    - 13,640 main conference reviewers
  
    - 203 datasets and benchmark Area Chairs
    - 36 datasets and benchmark Senior Area Chairs
    - 1,844 datasets and benchmark reviewers
  
    - 377 Ethics reviewers - 958 papers ( 5.48% of all submissions) were flagged for ethics review, increased from 502 papers ( 3.77% of all submissions) in 2023
  - Papers are available in the NeurIPS Proceeding - <https://proceedings.neurips.cc/>
  
- Invited Keynote Speakers 2024
  - Alison Gopnik, [The Golem vs. Stone Soup: Understanding How Children Learn Can Help Us Understand And Improve AI](#)
  - Sepp Hochreiter, [Toward Industrial Artificial Intelligence](#)
  - Fei-Fei Li, [From Seeing to Doing: Ascending the Ladder of Visual Intelligence](#)
  - Lidong Zhou, [A Match Made in Silicon: The Co-Evolution of Systems and AI](#)
  - Arnaud Doucet, [From Diffusion Models to Schrödinger Bridges](#)
  - Danica Kragic, [Learning for Interaction and Interaction for Learning](#)
  - Rosalind Picard, [How to optimize what matters most?](#)
  
- New - NeurIPS High School Projects
  - Launched to get the next generation excited and thinking about how Machine Learning can benefit society
    - Theme: Machine learning for social impact

- 330 total project submissions
  - 21 Spotlited
  - 4 Awards Given
- Expo
  - 58 Expo Talks / Panels / Workshops / Demonstrations for Exhibitors
- Nine [Affinity Groups](#) Represented
  - Indigenous in AI/ML - [Indigenous in AI website](#)
  - Muslims in ML - [Muslims in ML website](#)
  - New in ML workshop - [New in ML website](#)
  - Neurodivergent Community - [Workshop's website](#)
  - Global South in AI - [website](#)
  - Women in ML - [Women in ML website](#)
  - Queer in AI - [Queer in AI website](#)
  - LatinX in AI - [Latinx in AI website](#)
  - Black in AI - [Black in AI website](#)

**Invited Keynote Speaker Image:**



## Award Recipients:

### Test of Time Awards

Abstract:

[Generative Adversarial Nets](#) *Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio*

Presented by Ian Goodfellow and David Warde-Farley

Paper Abstract: We propose a new framework for estimating generative models via adversarial nets, in which we simultaneously train two models: a generative model  $G$  that captures the data distribution, and a discriminative model  $D$  that estimates the probability that a sample came from the training data rather than  $G$ . The training procedure for  $G$  is to maximize the probability of  $D$  making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions  $G$  and  $D$ , a unique solution exists, with  $G$  recovering the training data distribution and  $D$  equal to  $1/2$  everywhere. In the case where  $G$  and  $D$  are defined by multilayer perceptrons, the entire system can be trained with backpropagation. There is no need for any Markov chains or unrolled approximate inference networks during either training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.

[Sequence to Sequence Learning with Neural Networks](#) *Ilya Sutskever, Oriol Vinyals, Quoc V. Le*

Paper Abstract: Deep Neural Networks (DNNs) are powerful models that have achieved excellent performance on difficult learning tasks. Although DNNs work well whenever large labeled training sets are available, they cannot be used to map sequences to sequences. In this paper, we present a general end-to-end approach to sequence learning that makes minimal assumptions on the sequence structure. Our method uses a multilayered Long Short-Term Memory (LSTM) to map the input sequence to a vector of a fixed dimensionality, and then another deep LSTM to decode the target sequence from the vector. Our main result is that on an English to French translation task from the WMT-14 dataset, the translations produced by the LSTM achieve a BLEU score of 34.8 on the entire test set, where the LSTM's BLEU score was penalized on out-of-vocabulary words. Additionally, the LSTM did not have difficulty on long sentences. For comparison, a phrase-based SMT system achieves a BLEU score of 33.3 on the same dataset. When we used the LSTM to rerank the 1000 hypotheses produced by the aforementioned SMT system, its BLEU score increases to 36.5, which is close to the previous state of the art. The LSTM also learned sensible phrase and sentence representations that are sensitive to word order and are relatively invariant to the active and the passive voice. Finally, we found that reversing the order of the words in all source sentences (but not target sentences) improved the LSTM's performance markedly, because doing so introduced many short term dependencies between the source and the target sentence which made the optimization problem easier.

### Best Paper Awards for the main track:

[Visual Autoregressive Modeling: Scalable Image Generation via Next-Scale Prediction](#)

Authors: Keyu Tian, Yi Jiang, Zehuan Yuan, BINGYUE PENG, Liwei Wang

Abstract: This paper introduces a novel visual autoregressive (VAR) model that iteratively predicts the image at a next higher resolution, rather than a different patch in the image following an arbitrary ordering. The VAR model shows strong results in image generation, outperforming existing autoregressive models in efficiency and achieving competitive results with diffusion-based methods. At the core of this contribution lies an innovative multiscale VQ-VAE implementation. The overall quality of the paper presentation, experimental validation and insights (scaling laws) give compelling reasons to

experiment with this model.

[Stochastic Taylor Derivative Estimator: Efficient amortization for arbitrary differential operators](#)

Authors: Zekun Shi, Zheyuan Hu, Min Lin, Kenji Kawaguchi

Abstract: This paper proposes a tractable approach to train neural networks (NN) using supervision that incorporates higher-order derivatives. Such problems arise when training physics-informed NN to fit certain PDEs. Naive application of automatic differentiation rules are both inefficient and intractable in practice for higher orders  $k$  and high dimensions  $d$ . While these costs can be mitigated independently (e.g. for large  $k$  but small  $d$ , or large  $d$  but small  $k$  using subsampling) this paper proposes a method, stochastic taylor derivative estimator (STDE) that can address both. This work opens up possibilities in scientific applications of NN and more generally in supervised training of NN using higher-order derivatives.

**Runners-up for the main track:**

[Not All Tokens Are What You Need for Pretraining](#)

Authors: Zhenghao Lin, Zhibin Gou, Yeyun Gong, Xiao Liu, yelong shen, Ruochen Xu, Chen Lin, Yujiu Yang, Jian Jiao, Nan Duan, Weizhu Chen

Abstract: This paper presents a simple method to filter pre-training data when training large language models (LLM). The method builds on the availability of a high-quality reference dataset on which a reference language model is trained. That model is then used to assign a quality score for tokens that come from a larger pre-training corpus. Tokens whose scores have the highest rank are then used to guide the final LLM training, while the others are discarded. This ensures that the final LLM is trained on a higher quality dataset that is well aligned with the reference dataset.

[Guiding a Diffusion Model with a Bad Version of Itself](#)

Authors: Tero Karras, Miika Aittala, Tuomas Kynkäänniemi, Jaakko Lehtinen, Timo Aila, Samuli Laine

Abstract: This paper proposes an alternative to classifier free guidance (CFG) in the context of text-2-image (T2I) models. CFG is a guidance technique (a correction in diffusion trajectories) that is extensively used by practitioners to obtain better prompt alignment and higher-quality images. However, because CFG uses an unconditional term that is independent from the text prompt, CFG has been empirically observed to reduce diversity of image generation. The paper proposes to replace CFG by Autoguidance, which uses a noisier, less well-trained T2I diffusion model. This change leads to notable improvements in diversity and image quality.

**Best Paper Award for Datasets & Benchmarks track:**

[The PRISM Alignment Dataset: What Participatory, Representative and Individualised Human Feedback Reveals About the Subjective and Multicultural Alignment of Large Language Models](#)

Authors: Hannah Rose Kirk, Alexander Whitefield, Paul Röttger, Andrew Michael Bean, Katerina Margatina, Rafael Mosquera, Juan Manuel Ciro, Max Bartolo, Adina Williams, He He, Bertie Vidgen, Scott A. Hale

Abstract: Alignment of LLMs with human feedback is one of the most impactful research areas of today, with key challenges such as confounding by different preferences, values, or beliefs. This paper

introduces the PRISM dataset providing a unique perspective on human interactions with LLMs. The authors collected data from 75 countries with diverse demographics and sourced both subjective and multicultural perspectives benchmarking over 20 current state of the art models. The paper has high societal value and enables research on pluralism and disagreements in RLHF.

Best Paper Award committee for main track: Marco Cuturi (Committee Lead), Zeynep Akata, Kim Branson, Shakir Mohamed, Remi Munos, Jie Tang, Richard Zemel, Luke Zettlemoyer

Best Paper Award committee for dataset and benchmark track: Yulia Gel, Ludwig Schmidt, Elena Simperl, Joaquin Vanschoren, Xing Xie.

###