Preslav Aleksandrov

pa511@cam.ac.uk +44-7401-005612 linkedin.com/in/preslav-aleksandrov
https://github.com/Helios113 preslavaleksandrov.com Google Scholar

Abstract

My research focuses on creating new foundational model architectures which are more parameter and data efficient, thus they require much less compute for training and inference The reduced parameter size leads to a higher throughput and lower latency and enables more capable edge model deployments. I specifically work on dynamic iterative transformer models which allow test time compute scaling and circuit composition, as I believe that future of foundational models, similar to biological brains, will have a dynamic compute path.

Keywords: Machine Learning; Deep Learning; Natural Language Processing; Computer Vision; Reinforcement Learning; AI Safety; Large Language Models; Generative AI; Multimodal Learning

Education

University of Cambridge 2024 – Current
PhD in Computer Science *Cambridge, UK*
Advisor: Prof. Nic Lane – Machine Learning Systems Group
Funder: ICASE (UKRI and IMEC)
Research Focus:

- **Novel Foundational Architectures:** Deep Learning, Machine Learning, Natural Language Processing, Computer Vision, Reinforcement Learning, Probabilistic Graphical Models, Advanced Algorithms, Statistical Learning Theory
- **Foundational Model Training:** Dean's List (all semesters), Research Excellence Award, AI Research Fellowship
- **Distributed Optimisation of Foundational models:** Dean's List (all semesters), Research Excellence Award, AI Research Fellowship
- **Systems modeling for ML workloads (bottleneck discovery):** Dean's List (all semesters), Research Excellence Award, AI Research Fellowship

University of Glasgow 2024 - Current
MSc in Computer Science, GPA: 3.9/4.0 *Cambridge, UK*

- **Relevant Coursework:** Deep Learning, Machine Learning, Natural Language Processing, Computer Vision, Reinforcement Learning, Probabilistic Graphical Models, Advanced Algorithms, Statistical Learning Theory
- **Honors:** Dean's List (all semesters), Research Excellence Award, AI Research Fellowship

Research Experience

Research Institution/Lab Name Month Year – Present
Research Intern/Assistant | Advisor: Prof. Name *City, State/Country*

- Developed novel transformer architecture achieving 15% improvement in BLEU score on WMT translation benchmark, reducing model parameters by 30% while maintaining performance
- Implemented and optimized distributed training pipeline using PyTorch and DeepSpeed, scaling to 128 GPUs with 85% efficiency, reducing training time from 2 weeks to 3 days
- Published first-author paper at NeurIPS/ICML/ICLR (specify conference) on efficient attention mechanisms for long-context language models with 100+ citations
- Collaborated with 5-person research team, contributing 10,000+ lines of production-quality research code to open-source repository (3,000+ GitHub stars)

Previous Research Lab/Company Month Year – Month Year
Machine Learning Research Intern | Mentor: Dr. Name *City, State/Country*

- Designed and trained multimodal vision-language model achieving state-of-the-art results on 3 benchmark datasets (VQA, COCO Captioning, Visual Reasoning)
- Conducted ablation studies analyzing attention patterns in vision transformers, identifying key architectural improvements adopted by 50+ follow-up papers
- Co-authored workshop paper at CVPR/ECCV/ICCV on zero-shot transfer learning for image classification

Publications

First Author Publications:

- **Your Name**, Co-author Names. "Descriptive Title About Your Research Contribution." *NeurIPS/ICML/ICLR* Year. [arXiv] [code]
- **Your Name**, Co-author Names. "Another Paper Title on Relevant AI Topic." *Workshop at Major Conference*, Year. [arXiv]

Contributing Author Publications:

- Co-author Names, **Your Name**, et al. "Paper Title Where You Made Significant Contributions." *Conference Name*, Year.

Technical Projects

Project Name: Descriptive AI/ML Focus Month Year – Month Year

- Built end-to-end deep learning system for [specific task] using PyTorch/TensorFlow, achieving [specific metric] performance
- Implemented novel [algorithm/architecture] based on recent research, demonstrating [percentage] improvement over baseline
- Open-sourced implementation with comprehensive documentation, gaining 500+ GitHub stars and used by 10+ research groups
- **Technologies:** PyTorch, Transformers, CUDA, Weights & Biases, Docker, AWS/GCP

Another Significant Project Name Month Year – Month Year

- Developed [specific AI system] processing [data scale] with [performance metric], deployed in production serving 10,000+ users
- Optimized inference latency from [X]ms to [Y]ms through model quantization, pruning, and efficient deployment strategies
- **Technologies:** Python, JAX/Flax, Ray, Kubernetes, MLflow, TensorBoard

Technical Skills

Programming Languages: Python (expert), C++ (proficient), Julia, R, CUDA, SQL
ML Frameworks: PyTorch, TensorFlow, JAX, Hugging Face Transformers, scikit-learn, XGBoost
AI/ML Specializations: Deep Learning, Natural Language Processing, Computer Vision, Reinforcement Learning, Generative Models, Large Language Models, Diffusion Models, Graph Neural Networks
Tools & Infrastructure: Git, Docker, Kubernetes, AWS/GCP/Azure, Weights & Biases, MLflow, Ray, DeepSpeed, SLURM
Libraries: NumPy, Pandas, SciPy, Matplotlib, OpenCV, NLTK, spaCy, LangChain
Research Skills: Experimental design, statistical analysis, technical writing, literature review, reproducible research

Awards & Honors

- Best Paper Award, [Conference Name], Year – Recognized for outstanding research contribution
- Research Fellowship, [Institution], Year – Competitive fellowship awarded to top 5% of applicants
- Kaggle Competition Winner/Grandmaster, [Competition Name], Year – Ranked 1st/Top 1% among

5,000+ teams
- Outstanding Student Award in AI/ML, [Institution], Year

Teaching & Mentorship

Teaching Assistant – Course Name (ML/DL/AI) Month Year – Month Year

- Led weekly discussion sections for 50+ students, achieving 4.8/5.0 teaching evaluation
- Developed course materials and programming assignments on neural networks, CNNs, RNNs, and transformers
- Mentored 3 undergraduate research students, resulting in 2 workshop papers

Service & Leadership

- Reviewer for [Conference Names: NeurIPS, ICML, ICLR, etc.], Year
- Organizer, AI Research Reading Group at University, Year – Present
- Volunteer, AI for Social Good Initiative, contributing ML expertise to non-profit projects