Transformers: Open Knowledge Models for Robotics

Nikolaus Correll

About me

- Studied Electrical Engineering in Munich and Zurich
- Phd in Computer Science in Lausanne
- Faculty since 2009
- Startup from 2016-2022
- Research on robotics manipulation and robotic materials

Class overview

- https://canvas.colorado.edu/courses/122969

Goals of this class

1. Fundamentally understand transformer-based models with applications to robotics and their limitations
2. Learn about current research in the field, summarizing it and presenting to others
3. Advance the current state of the art by performing an independent research project
4. *Improve your writing skills and learn how to organize your research*

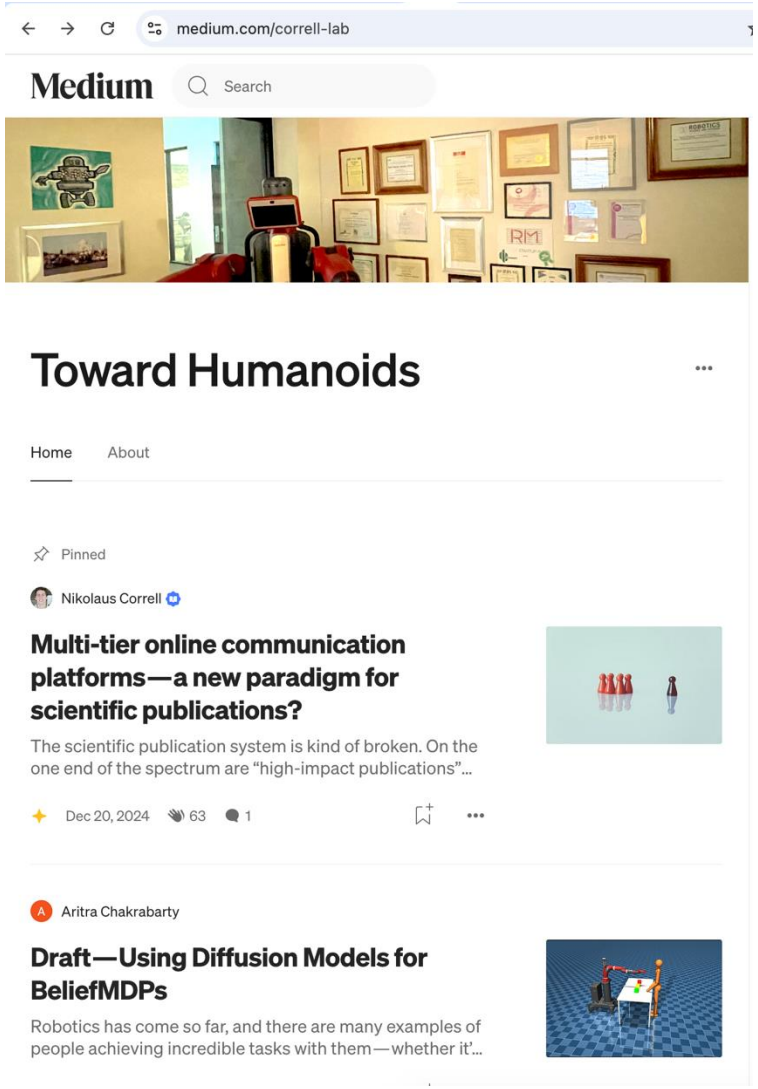
Outline (tentative)

- Week 1: Introduction / “Makemore”
- Week 2: Transformer encoder-decoder: Chat GPT
- Week 3: Building a vision transformer from scratch
- Week 4: Building CLIP from scratch
- Week 5: Behavior Cloning
- Week 6: Reinforcement learning
- Week 7: -
- Week 8-10: Paper reading
- Week 11-15: Final project / student-driven lectures

Assessments / Grading

- Final project -> 25%
- Implement transformer decoder-encoder on own data and blog article -> 20%
- Paper review and blog article -> 15%
- Attendance: 16*2.5% -> 40%
- Participate in peer reviews
- Bonus: quizzes

Class deliverables: Blog articles



Lifetime

Most read

April 4, 2024 – Today · Updated daily

Apr 2024	Matt Nguyen **Building a Vision Transformer Model From Scratch** 12 min read · Apr 4, 2024 · View story	3.6K Views	1.93K Reads
May 2024	Matt Nguyen **Building CLIP From Scratch** 19 min read · May 16, 2024 · View story	1.7K Views	956 Reads
Dec 2024	Maxlconway **Embodied AI For Block Stacking Tasks** 5 min read · Dec 17, 2024 · View story	1.5K Views	746 Reads
Oct 2024	Aritra Chakrabarty **Deep Transformer Q Networks—A paper analysis** 10 min read · Oct 31, 2024 · View story	1.1K Views	513 Reads
Dec 2024	Naren Dasan **Using Waypoints to Stitch Diffusion Generated Sub-Tr...** 10 min read · Dec 20, 2024 · View story	682 Views	319 Reads

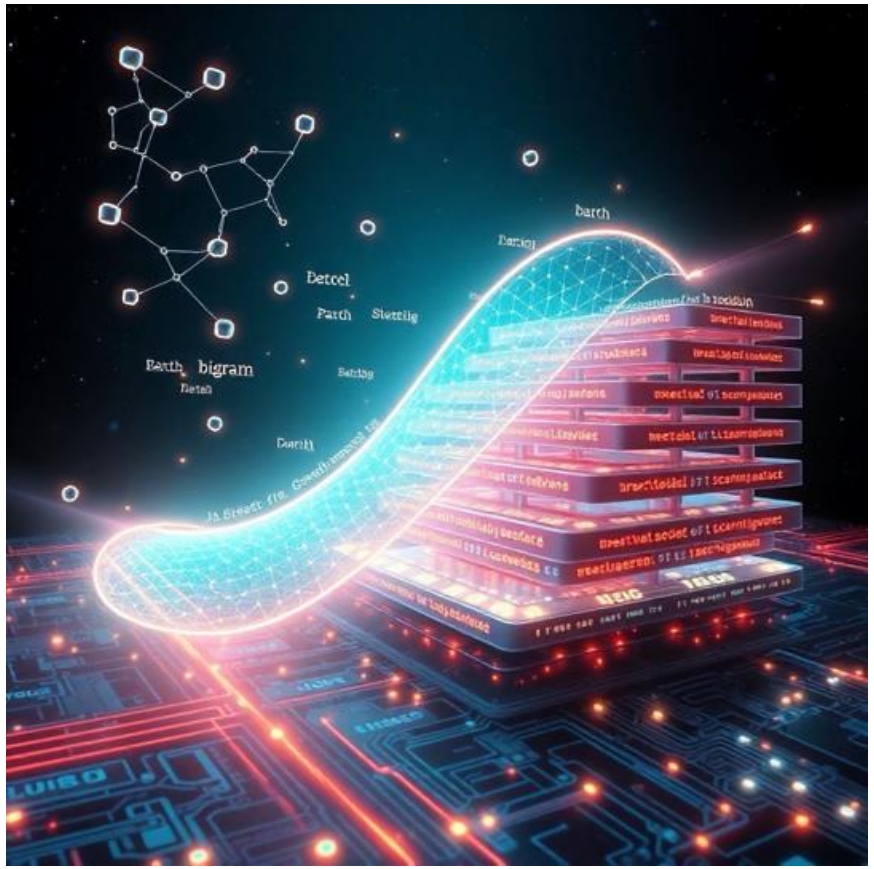
You!

Week 1: Introduction

- “Manipulate Anything” : https://arxiv.org/pdf/2406.18915
- “OK-Robot: What really matters in integrating open-knowledge models in robotics” : https://arxiv.org/abs/2401.12202
- Generative AI for trajectory generation: https://robotics-transformer-x.github.io/
- Brush up on language models

What you need to know

- Basic bigram models
- Tokenization
- Autograd
- One-Hot Encoding
- SoftMax
- Sampling from a multinomial distribution
- Cross-entropy Loss
- Gradient Descent
- Batch-processing
- Padding



https://medium.com/towards-data-science/gradient-descent-and-batch-processing-for-generative-models-in-pytorch

Colab

“Makemore”

- Brush-up on Torch / ML: “Building Makemore” The spelled-out intro to language modeling: building makemore
- Quiz on Canvas