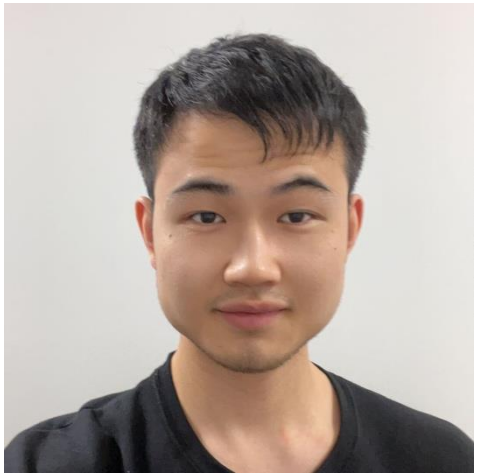
Advanced Machine Learning Overview

Yu Wang
Assistant Professor
Department of Computer Science
University of Oregon

Welcome

Welcome to the CS 610 – Advanced Machine Learning!



Yu (Jack) Wang
You

Knowledge Intelligence for Discovery and Decision-making (KIND) Lab

- Data Mining and Machine Learning
- Graph and Geometric Machine Learning
- Neural-Symbolic Learning
- Agent and LLM
- Application: Information Retrieval/Social Network Analysis

What to expect from this class?

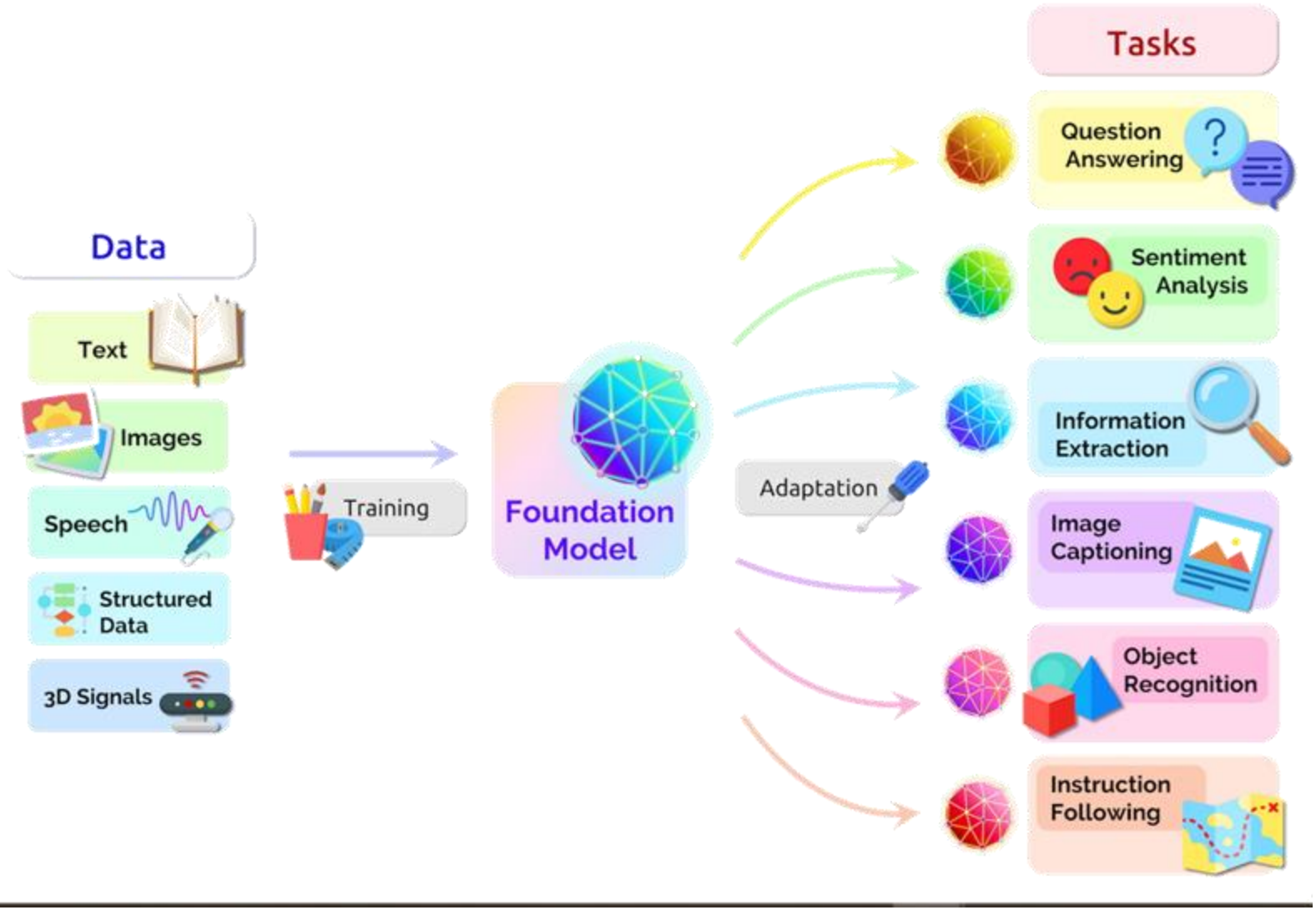
Neural Network User

- Off-the-shelf
- Fine-tuning
- Don't know much about the underlying principles
- …

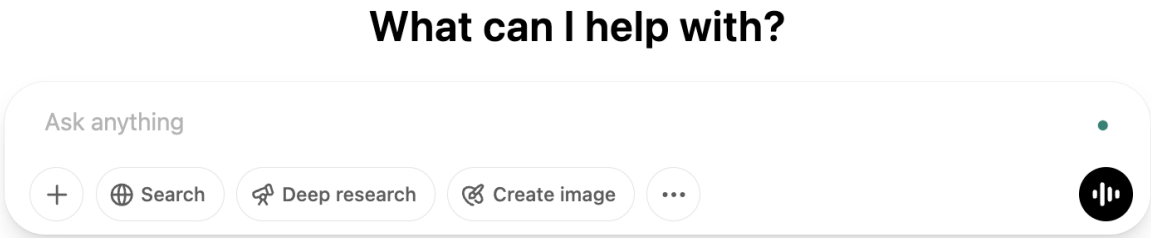
Neural Network Researcher

- **Understand mathematical underpinnings**
- Provide a holistic view that enables you to design and train custom architectures for a wide variety of problems
- …

Logistics

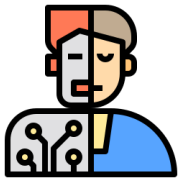

Data
Text
Images
Speech
Structured Data
3D Signals
Training
Foundation Model
Adaptation
Tasks
Question Answering
Sentiment Analysis
Information Extraction
Image Captioning
Object Recognition
Instruction Following

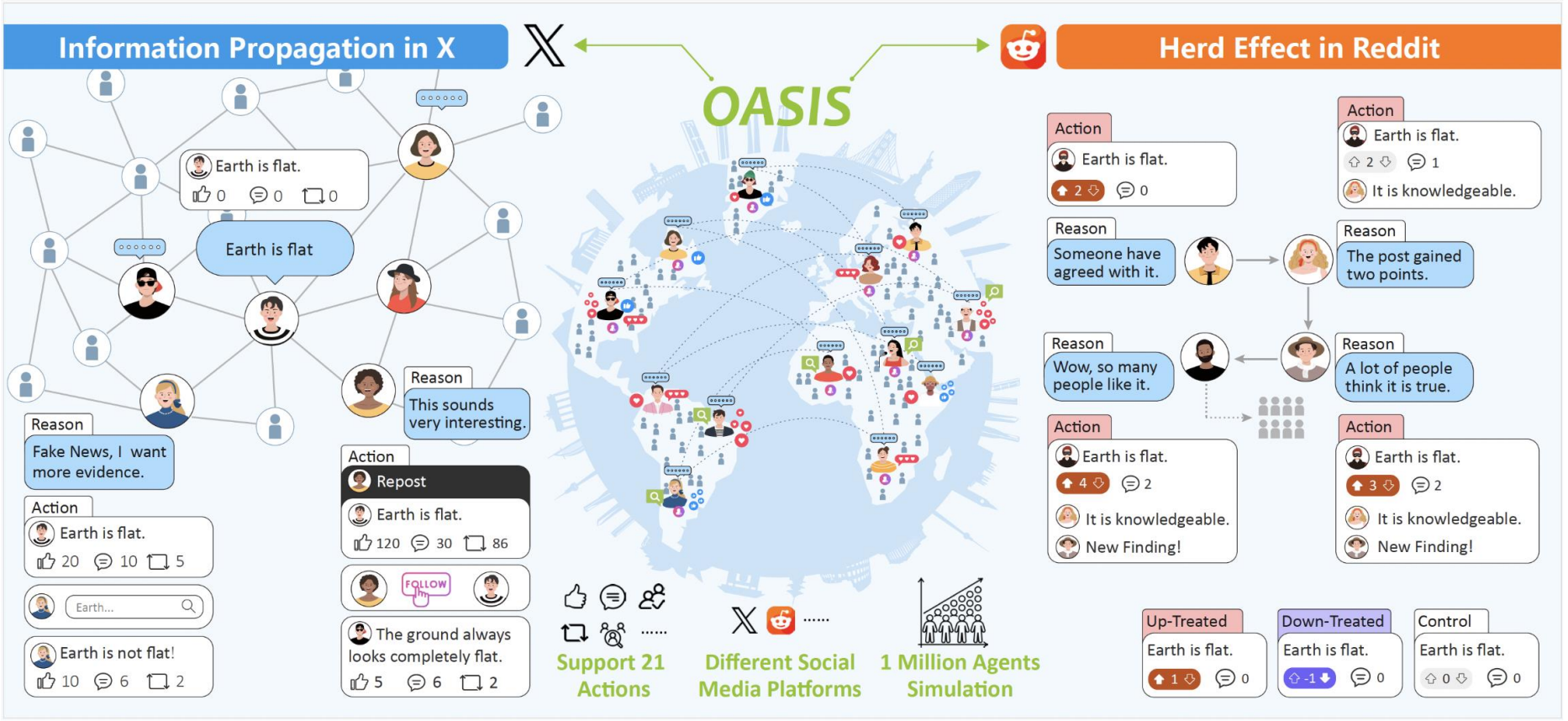
Logistics


What can I help with?
Ask anything
Search
Deep research
Create image


Adobe
Adobe

Logistics



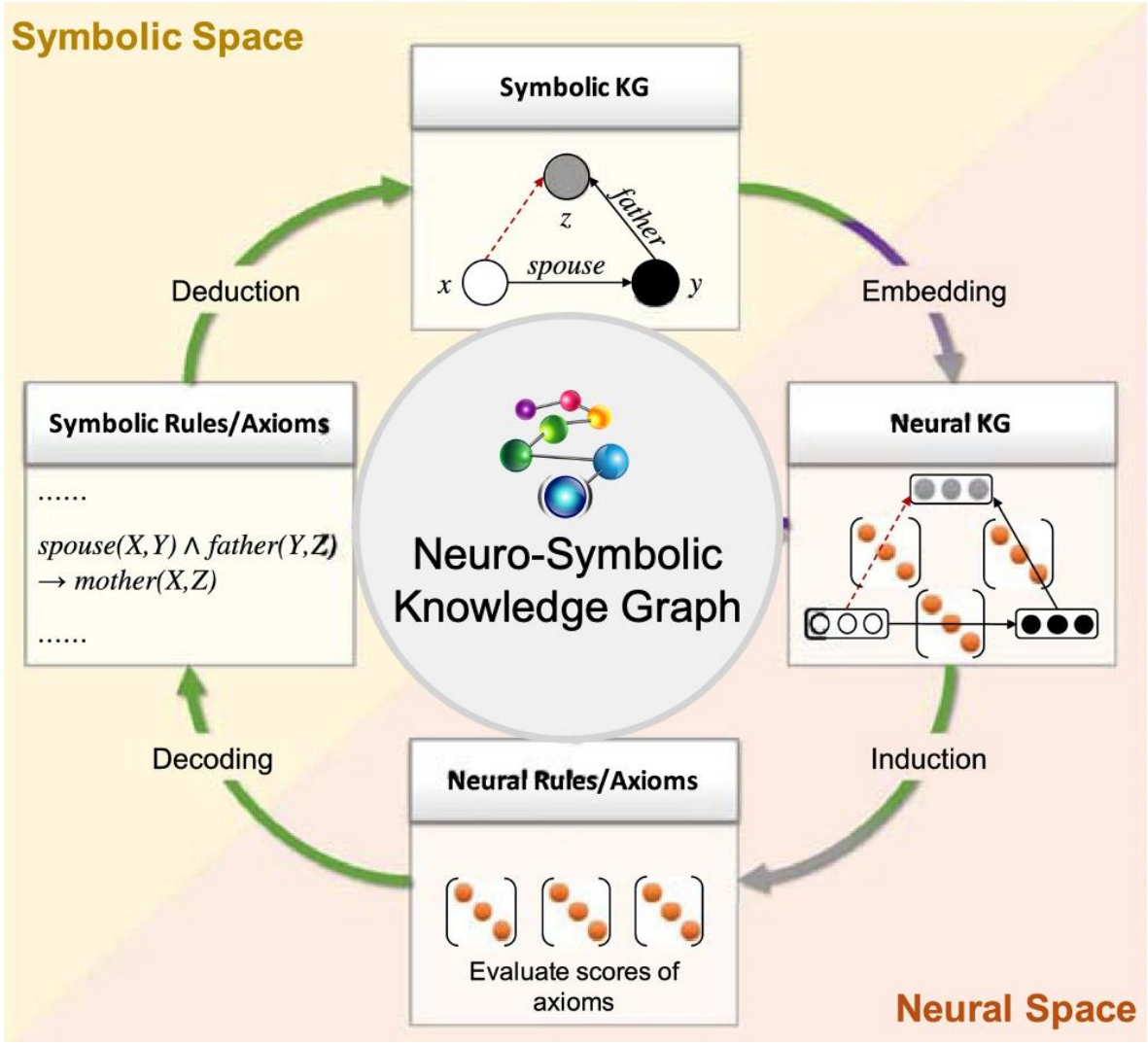

Information Propagation in X
OASIS
Herd Effect in Reddit
Earth is flat.
0 0 0
Earth is flat
Reason
Fake News, I want more evidence.
Reason
This sounds very interesting.
Action
Earth is flat.
20 10 5
Earth...
Earth is not flat!
10 6 2
Action
Repost
Earth is flat.
120 30 86
FOLLOW
The ground always looks completely flat.
5 6 2
Support 21 Actions
Different Social Media Platforms
1 Million Agents Simulation
Action
Earth is flat.
2 0
Reason
Someone have agreed with it.
Action
Earth is flat.
2 1
It is knowledgeable.
Reason
The post gained two points.
Reason
Wow, so many people like it.
Reason
A lot of people think it is true.
Action
Earth is flat.
4 2
It is knowledgeable.
New Finding!
Action
Earth is flat.
3 2
It is knowledgeable.
New Finding!
Up-Treated
Earth is flat.
1 0
Down-Treated
Earth is flat.
-1 0
Control
Earth is flat.
0 0

Logistics

Large Language Models Won't Achieve AGI

Sam Altman says “we are now confident we know how to build AGI”

Logistics


Symbolic Space
Symbolic KG
z
father
spouse
x
y
Deduction
Embedding
Symbolic Rules/Axioms
.....
spouse(X,Y) ∧ father(Y,Z) → mother(X,Z)
.....
Neuro-Symbolic Knowledge Graph
Neural KG
Decoding
Induction
Neural Rules/Axioms
Evaluate scores of axioms
Neural Space

Logistics

EVENT	DATE	DESCRIPTION	COURSE MATERIAL
Paper Presentation	03/31/2025 04:30 Monday	Topic of Paper Release.	
Assignment	03/31/2025 Monday	**Project released!**	[Project]
Lecture	04/04/2025 Friday	Overview Syllabus	**Course Materials:** ◦ Slides
Lecture	04/11/2025 Friday	Kernel Density Estimation	**Course Materials:** ◦ Slides ◦ Codebook
Lecture	04/18/2025 Friday	Autoencoder	**Course Materials:** ◦ Slides ◦ Codebook
Lecture	04/25/2025 Friday	Variational Autoencoder	**Course Materials:** ◦ Slides
Lecture	05/02/2025 Friday	Generative Adversarial Network - Video	**Course Materials:** ◦ Slides
Lecture	05/09/2025 Friday	Diffusion - Video	**Course Materials:** ◦ Slides
Lecture	05/16/2025 Friday	Reinforcement Learning1	**Course Materials:** ◦ Slides
Lecture	05/23/2025 Friday	Reinforcement Learning 2	**Course Materials:** ◦ Slides
Lecture	05/30/2025 Friday	Neural Symbolic Learning 1	**Course Materials:** ◦ Slides
Lecture	06/06/2025 Friday	Neural Symbolic Learning 2	**Course Materials:** ◦ Slides
Due	**06/06/2025 23:59 Friday**	**Project Report Due**	

Conference Travel Video/Virtual Onlne

Grade Scale

Course Assessment and Grading Scale

Category	Percentage
Project	60%
Paper Presentation	30%
Attendance	10%

Grade	Range
A	A+: 98-100, A: 93-97, A-: 90-92
B	B+: 87-89, B: 83-86, B-: 80-82
C	C+: 77-79, C: 73-76, C-: 60-72
F	F: <60

Project

Project

https://ml-graph.github.io/spring-2025/project/

1. Introduction and Background - 10%

- **General Background**:
  - What is the general background of the problem you are working on?
- **Specific Problem**:
  - Under the general topic, what specific problem is your project addressing?

2. Motivation and Objective - 10%

- **Problem Statement**:
  - What are the limitations of existing methods in addressing this problem?
- **Contribution and Novelty**:
  - Given the previous limitations, what is your unique contribution, and how does it provide novelty in solving the problem?

3. Data Collection and Analysis - 30%

- **Dataset**:
  - What dataset are you working on to solve the problem?
- **Dataset Introduction and Analysis**:
  - Provide a basic introduction to and analysis of the dataset.

4. Method - 30%

- **Algorithm Design/Implementation**:
  - What data mining/machine learning algorithms are you designing or applying to tackle the problem?

5. Experiment and Discussion - 20%

- **Experimentation**:
  - Conduct experiments to verify that the proposed method works.
- **Discussion**:
  - Analyze and discuss the results of the experiments.

Here is an example of the one-course project.

- Project Report Template

Paper

Paper

Paper Presentation Details

You can either collaborate with a team or present individually. The choice of topic is entirely up to you.

- Introduction and Background – What is the general impact and background of the topic?
- Motivation and Problem – What is the core research problem, and why do we study it?
- Related Work and Challenges – How did previous works address this problem, and what are some of the challenges?
- Proposed Solutions/Methods and Rationale – What are the proposed methods/techniques, and why are they proposed? What specific reasons that solving this problem would require these proposed(1) methods/techniques?
- Experimental Setting, Results, and Analysis – What experiments are designed to verify the proposed method? How are results being discussed and analyzed? Are there any interesting findings?
- Conclusion and Future Work

Generative Models

RAG Systems

From Isolated Conversations to Hierarchical Schemas: Dynamic Tree Memory Representation for LLMs [Paper]
LLM Alignment as Retriever Optimization: An Information Retrieval Perspective [Paper]
RAS: Retrieval-And-Structuring for Knowledge-Intensive LLM Generation [Paper]
Wikipedia Contributions in the Wake of ChatGPT [Paper]
Towards Knowledge Checking in Retrieval-augmented Generation: A Representation Perspective [Paper]

Agentic AI

HETEROGENEOUS SWARMS: Jointly Optimizing Model Roles and Weights for Multi-LLM System [Paper]
OASIS: Open Agents Social Interaction Simulations on One Million Agents [Paper]
Why Do Multi-Agent LLM Systems Fail? [Paper]
A Multi-LLM Debiasing Framework [Paper]
TrendSim: Simulating Trending Topics in Social Media Under Poisoning Attacks with LLM-based Multi-agent System [Paper]
Preference Leakage: A Contamination Problem in LLM-as-a-judge [Paper]
G-Designer: Architecting Multi-agent Communication Topologies via Graph Neural Networks [Paper]
GPTSwarm: Language Agents as Optimizable Graphs [Paper]

Neural-Symbolic Learning

Self-Discover: Large Language Models Self-Compose Reasoning Structures [Paper]
ThinkPatterns-21k: A Systematic Study on the Impact of Thinking Patterns in LLMs [Paper]
Stop Overthinking: A Survey on Efficient Reasoning for Large Language Models [Paper]
Do LLMs Think Fast and Slow? A Causal Study on Sentiment Analysis [Paper]

Complex Topology and Graph Structure

PolyhedronNet: Representation Learning for Polyhedra with Surface-attributed Graph [Paper]
Slow Perception: Let's Perceive Geometric Figures Step-by-step [Paper]
Representation Learning of Geometric Trees [Paper]
Deep Identification of Propagation Trees [Paper]
Network Tomography with Path-Centric Graph Neural Network [Paper]

https://ml-graph.github.io/spring-2025/paper/