UNIVERSITY OF GEORGIA

CSCI 8945 | Fall 2024

Advanced Representation Learning

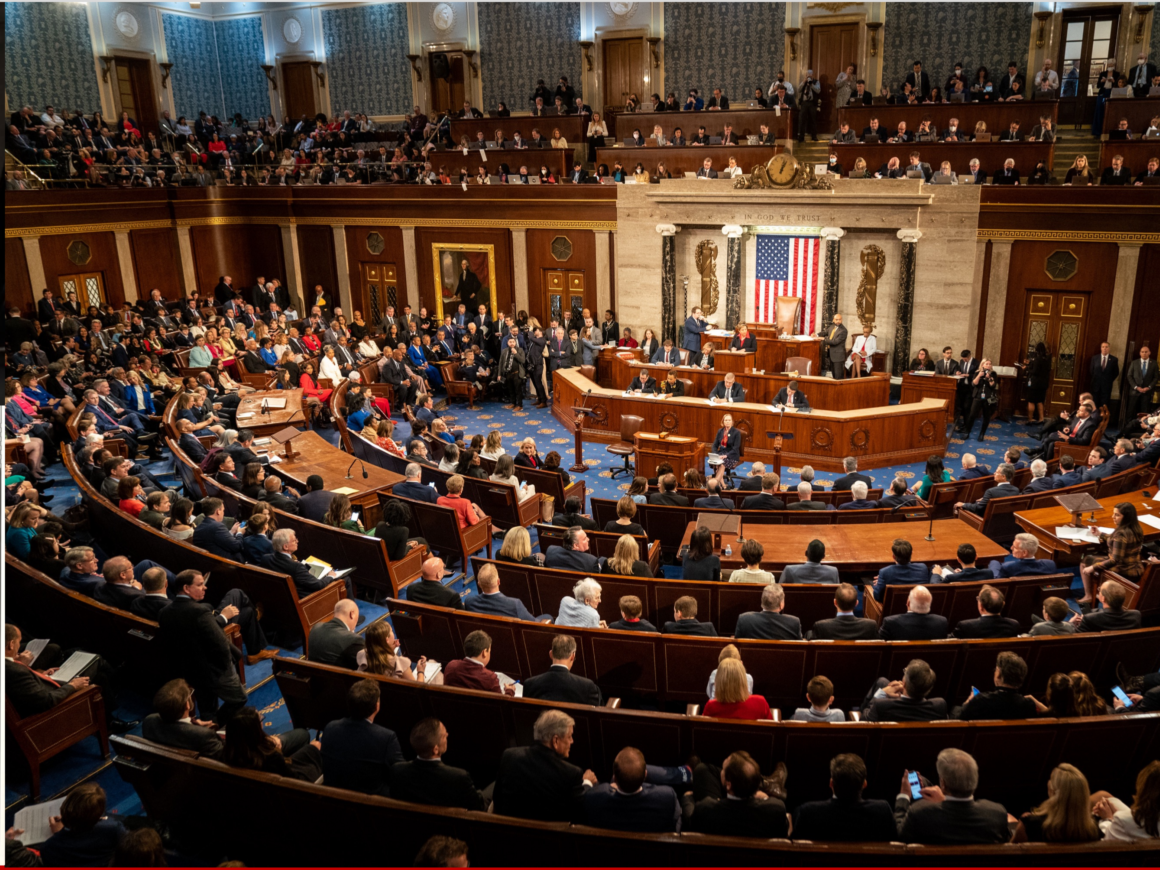
Jin Sun, PhD
School of Computing

Lec 1: Introduction

Outline

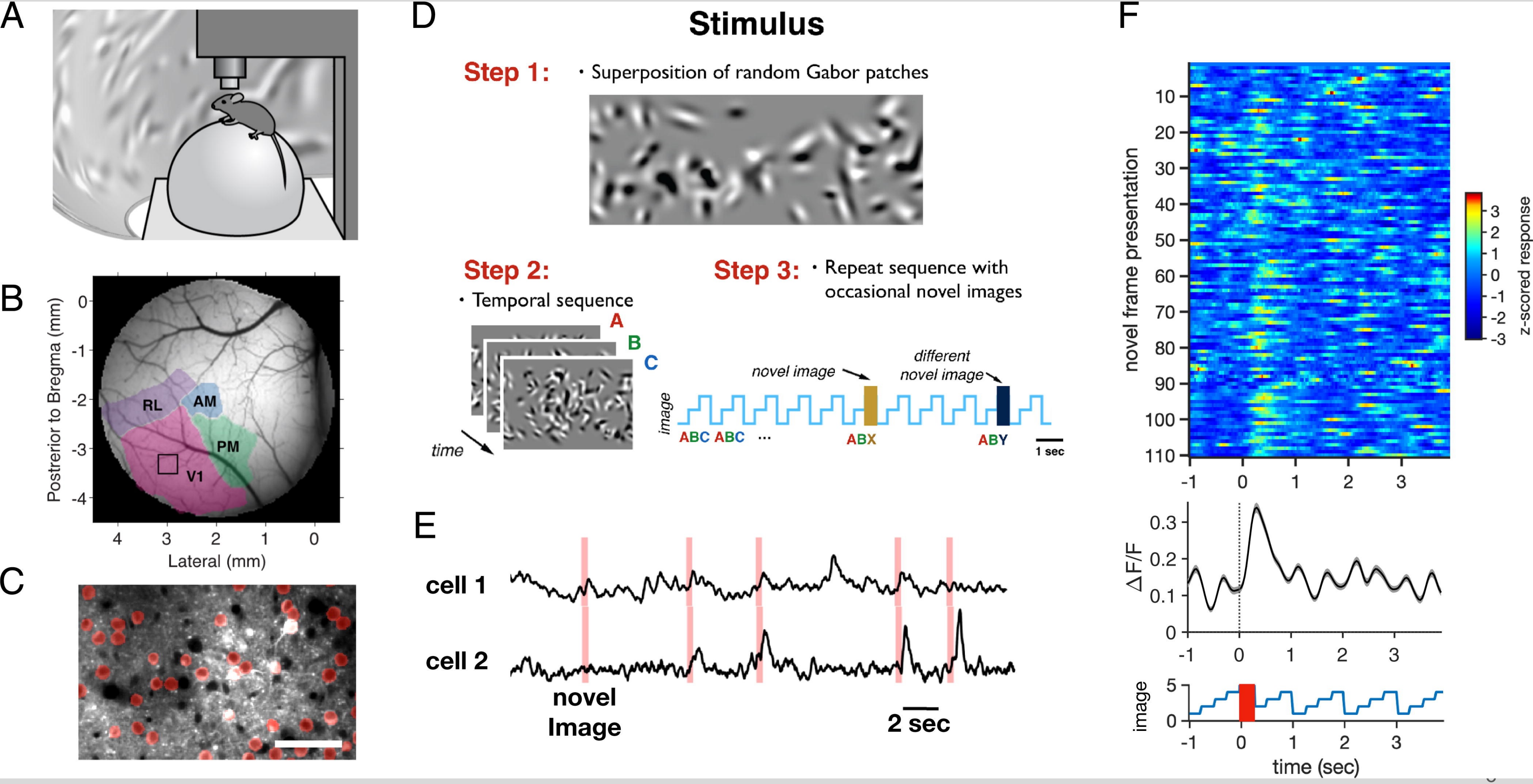
- What are representations
- Why should we care about learning representations
- What makes a representation good
- How to build (deep) representations
- Structure of the class
    - Homeworks and exam
    - Project
    - Format
    - Topics

Representations

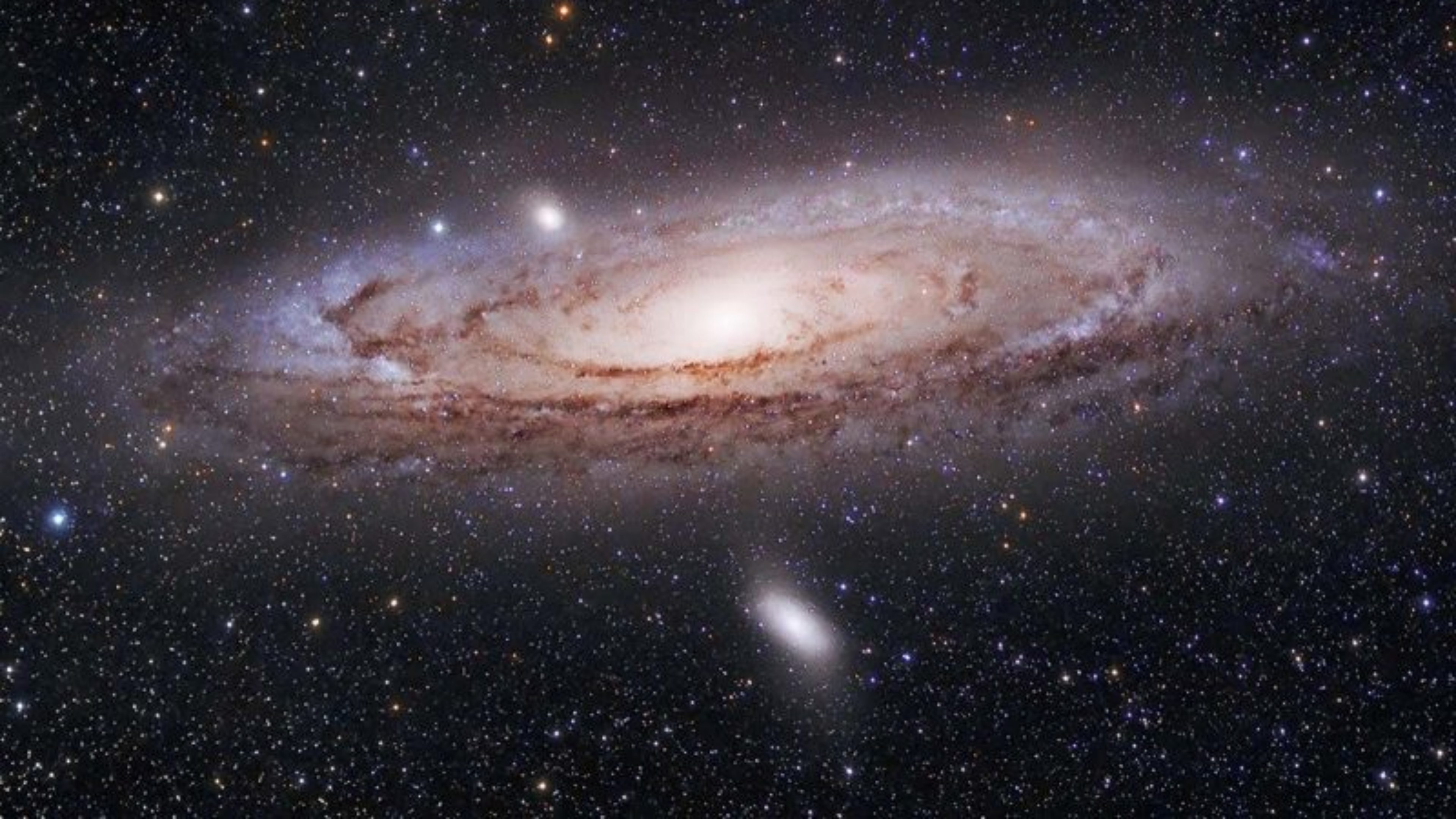


Representations

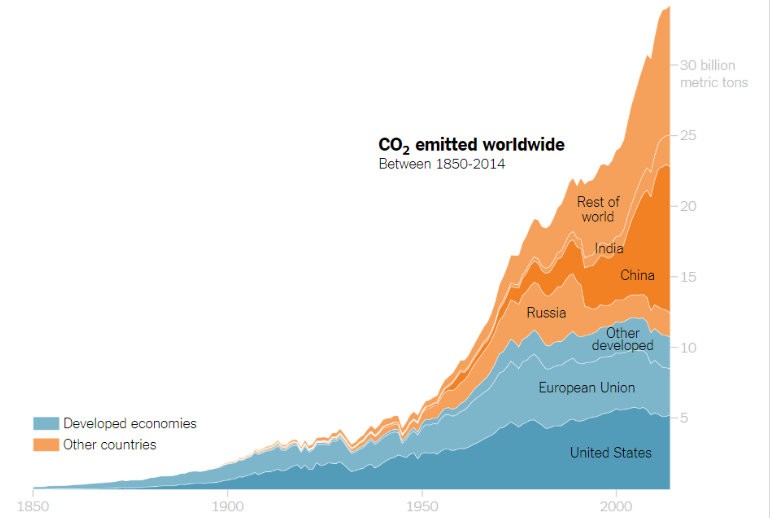
Representations

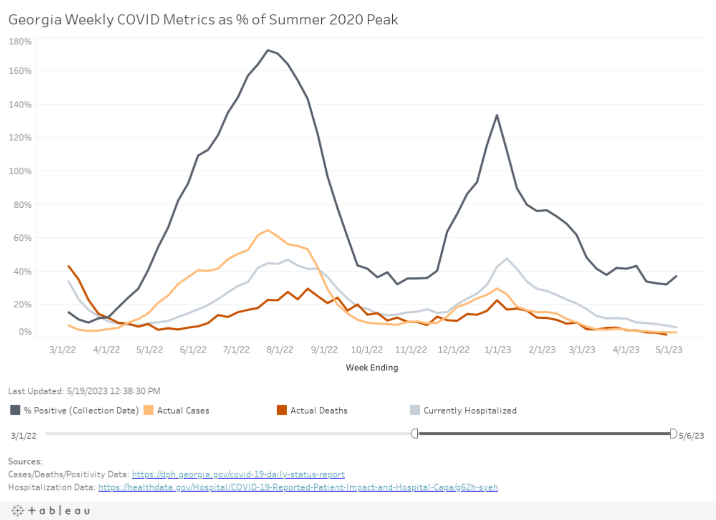


A
B
Posterior to Bregma (mm)
0
-1
-2
-3
-4
RL
AM
PM
V1
4
3
2
1
0
Lateral (mm)
C
D
Stimulus
Step 1:
• Superposition of random Gabor patches
Step 2:
• Temporal sequence
A
B
C
time
Step 3:
• Repeat sequence with occasional novel images
novel image
different novel image
image
ABC ABC ...
ABX
ABY
1 sec
E
cell 1
cell 2
novel Image
2 sec
F
novel frame presentation
10
20
30
40
50
60
70
80
90
100
110
-1
0
1
2
3
z-scored response
3
2
1
0
-1
-2
-3
ΔF/F
0.3
0.2
0.1
0
image
5
0
time (sec)

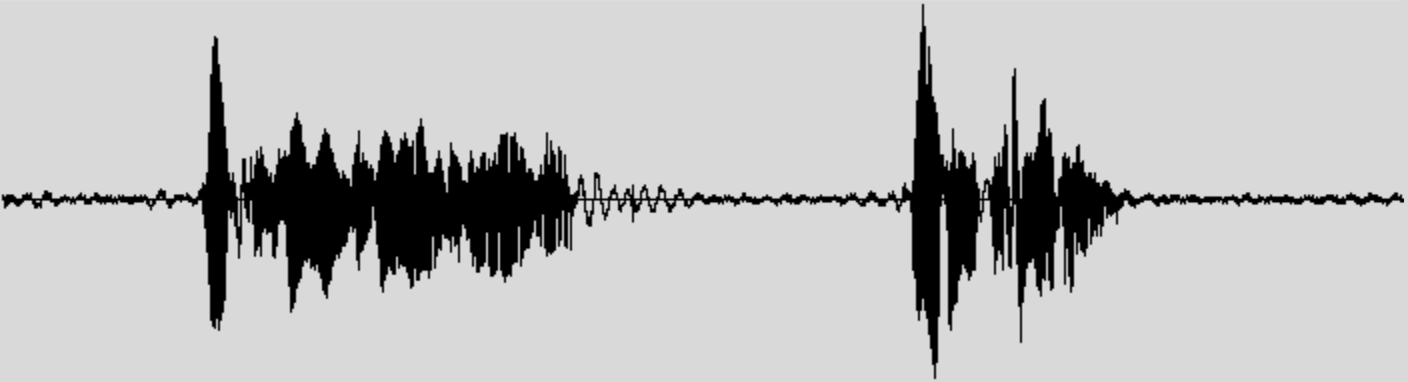


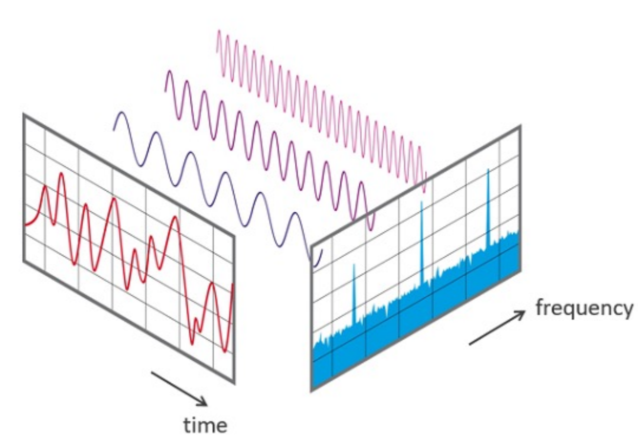
Temporal Data



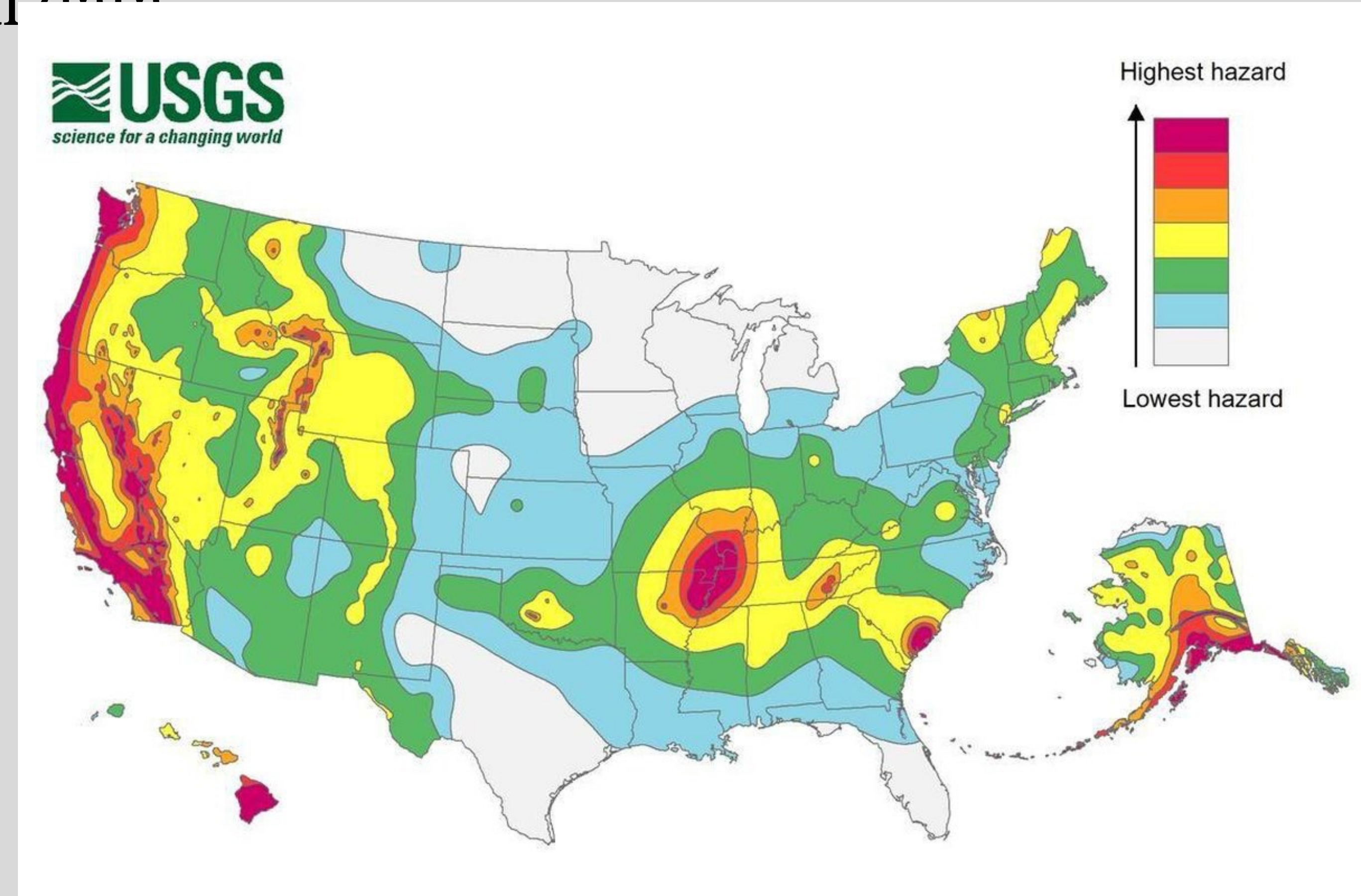


Speech and Music



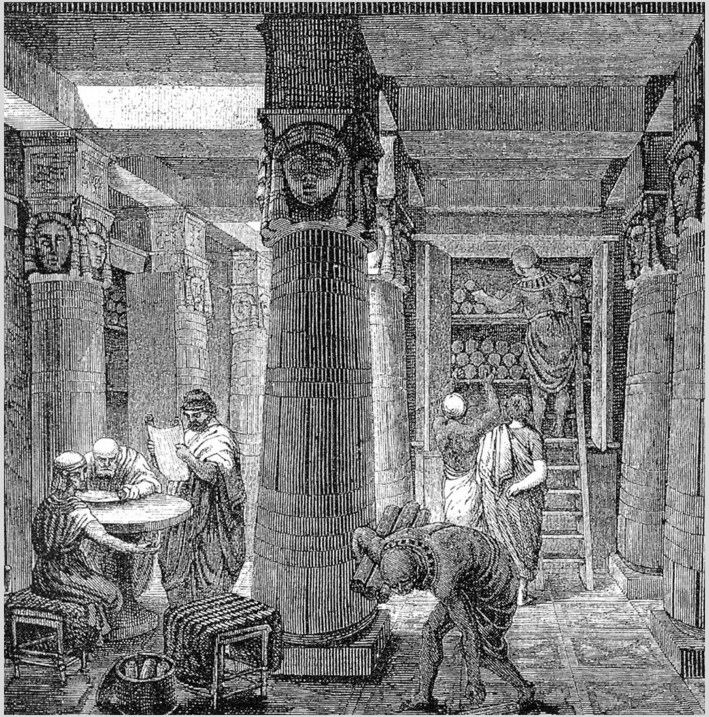


Spatial data

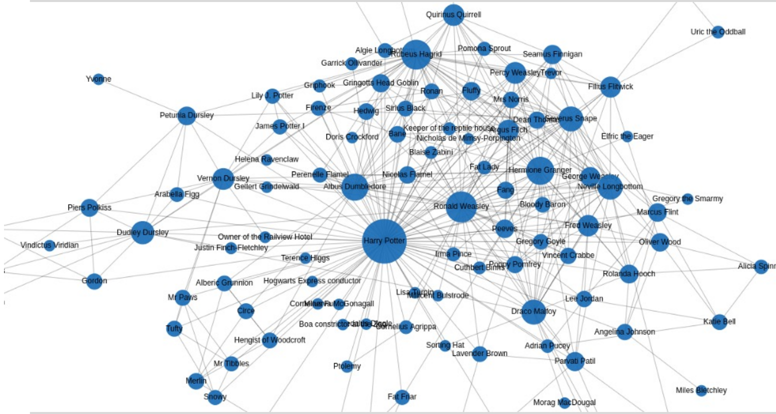


Language


Hello!
Salut!
Hallo!
Bonjour!
ПРИВЕТ!
Ciao!
مرحبا
नमस्ते
您好
Hi!
Salut!
Hola!
Olá!
Hi!



Graph data



Quirinus Quirrell
Uric the Oddball
Rubeus Hagrid
Pomona Sprout
Seamus Finnigan
Garrick Ollivander
Percy Weasley
Trevor
Yvonne
Griphook
Gringotts Head Goblin
Ronan
Fluffy
Filius Flitwick
Lily J. Potter
Mrs Norris
Petunia Dursley
Firenze
Hedwig
Sirius Black
Severus Snape
James Potter I
Argus Filch
Doris Crockford
Elfric the Eager
Nicholas de Mimsy-Porpington
Blaise Zabini
Helena Ravenclaw
Fat Lady
Hermione Granger
Vernon Dursley
Perenelle Flamel
Nicolas Flamel
Gellert Grindelwald
Albus Dumbledore
Neville Longbottom
Arabella Figg
Fang
Gregory the Smarmy
Piers Polkiss
Ronald Weasley
Bloody Baron
Marcus Flint
Fred Weasley
Dudley Dursley
Peeves
Owner of the Railview Hotel
Harry Potter
Gregory Goyle
Oliver Wood
Vindictus Viridian
Justin Finch-Fletchley
Irma Pince
Vincent Crabbe
Terence Higgs
Poppy Pomfrey
Rolanda Hooch
Gordon
Hogwarts Express conductor
Lee Jordan
Mr Paws
Draco Malfoy
Circe
Angelina Johnson
Katie Bell
Tufty
Hengist of Woodcroft
Sorting Hat
Adrian Pucey
Lavender Brown
Parvati Patil
Mr Tibbles
Ptolemy
Merlin
Miles Bletchley
Snowy
Fat Friar
Morag MacDougal

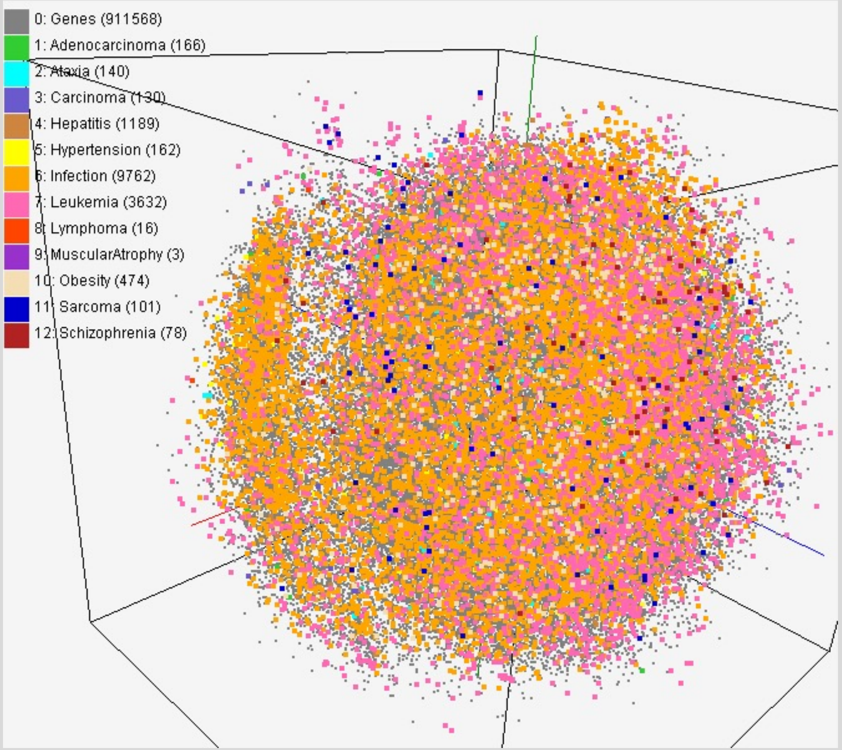
Tabular Data

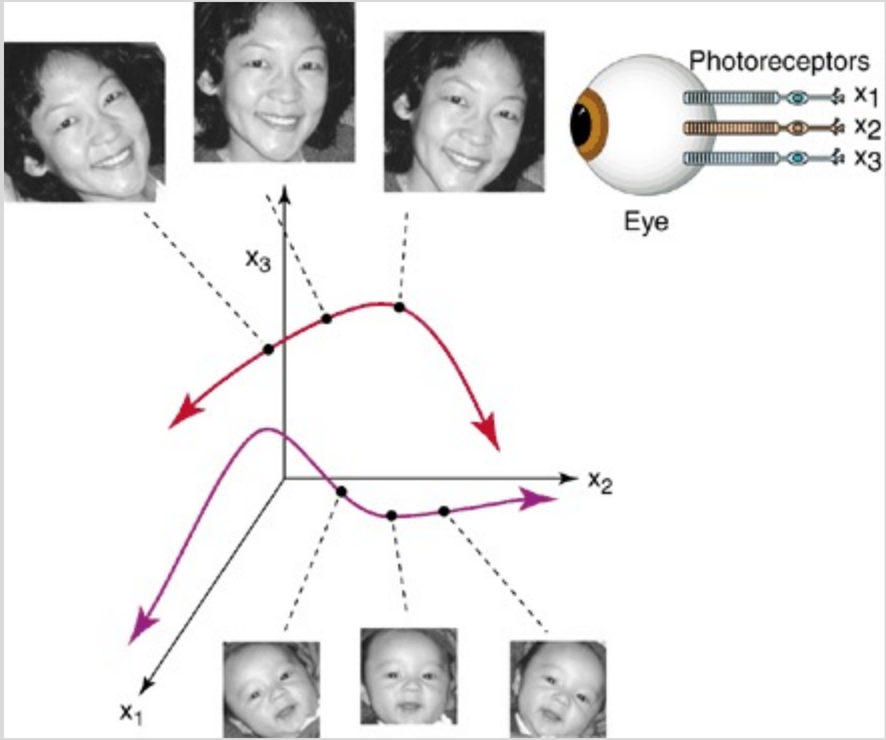
Bl.xlsx - Excel

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW VIEW DEVELOPER Tableau LOAD TEST POWER QUERY POWERPIVOT

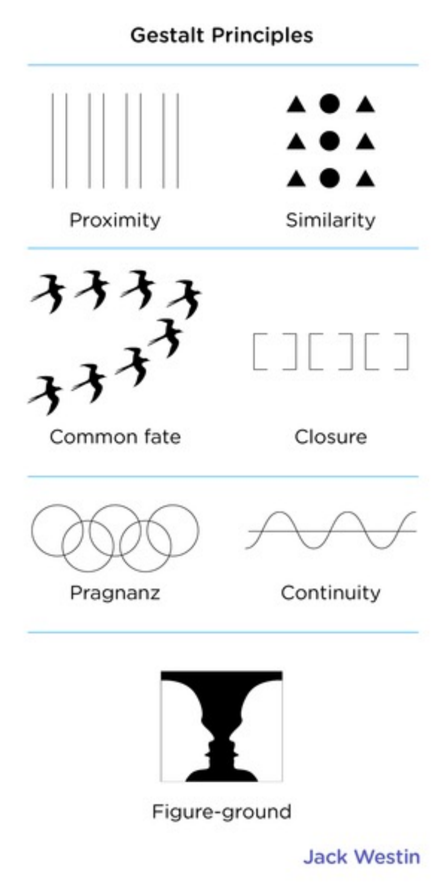
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	C1	C2	C3	C4	C5	C6	C7	C8	S1	S2	S3	S4	S5	S6	S7	S8
2	21st centu	Reputatio	Sales	Human be	Economic	Object	Reputatio	Psycholog	0.944344	0.679675	0.65679	0.650183	0.633821	0.632812	0.63147	0.611786
3	Abbreviat	Acronym a	Problem s	Informati	Research	Abbreviat	Intelligen	Word	0.980795	0.757834	0.642039	0.601544	0.597469	0.567183	0.520521	0.51937
4	Academia	Online pa	Systems th	Social me	Research	Knowledg	Scientific	Interdisci	0.950885	0.81972	0.78258	0.772868	0.772475	0.74008	0.695616	0.663366
5	Academic	Managem	Higher ed	Competiti	Strategic r	Intelligen	Psycholog	Education	0.969353	0.780938	0.686018	0.578373	0.494522	0.461092	0.455948	0.452912
6	Access con	Database	Business i	Authoriza	Authentic	Database	Computer	Corporati	0.97301	0.942225	0.838212	0.674035	0.644832	0.62018	0.604592	0.592976
7	Actuarial s	Statistics	Insurance	Business i	Data	Business	Analytics	Economic	0.975186	0.944509	0.734045	0.561082	0.51668	0.500373	0.476394	0.453732
8	Agile soft	Business i	Extreme F	Project ma	Flexible p	Managem	Project	Grounded	0.94732	0.686218	0.667628	0.652256	0.605085	0.595524	0.568934	0.564779
9	Agile soft	Business i	Data analy	Systems D	Scientific	Extreme F	Waterfall	Methodol	0.954089	0.644952	0.610109	0.547888	0.545839	0.527249	0.496478	0.495703
10	Algebra	Variables	Variable	Polynomi	Euclidean	Variable	Dimensio	Scatter pl	0.989955	0.866679	0.793393	0.786708	0.718503	0.702535	0.644944	0.601248
11	Algorithm	Tabu sear	Methodol	Heuristic	Text mini	Improve	Business i	Statistical	0.9786	0.688966	0.640741	0.483909	0.478563	0.428156	0.425666	0.417563
12	Algorithm	Approxim	Optimizat	Optimizat	Heuristic	Operation	Consisten	Optimizat	0.967184	0.838888	0.751672	0.711391	0.595888	0.541111	0.532653	0.532646
13	Algorithm	Computer	Computer	Data ware	Data mini	Profiling	Fuzzy logi	Software	0.967553	0.711902	0.606711	0.534577	0.492062	0.46977	0.450693	0.443457
14	Algorithm	Regressio	Data	Data mini	Data analy	Java	Statistics	Play	0.97989	0.85421	0.825393	0.803136	0.76979	0.715599	0.697516	0.682136
15	Algorithm	Algorithm	Relational	Operation	Relational	Optimizat			0.965917	0.929522	0.927793	0.910567	0.850657	0.798374		
16	Alternativ	Pharmaco	Relational	Database	Botany	Traditiona	SQL	Traditiona	0.967788	0.890986	0.846344	0.819472	0.800314	0.789922	0.779625	0.741703

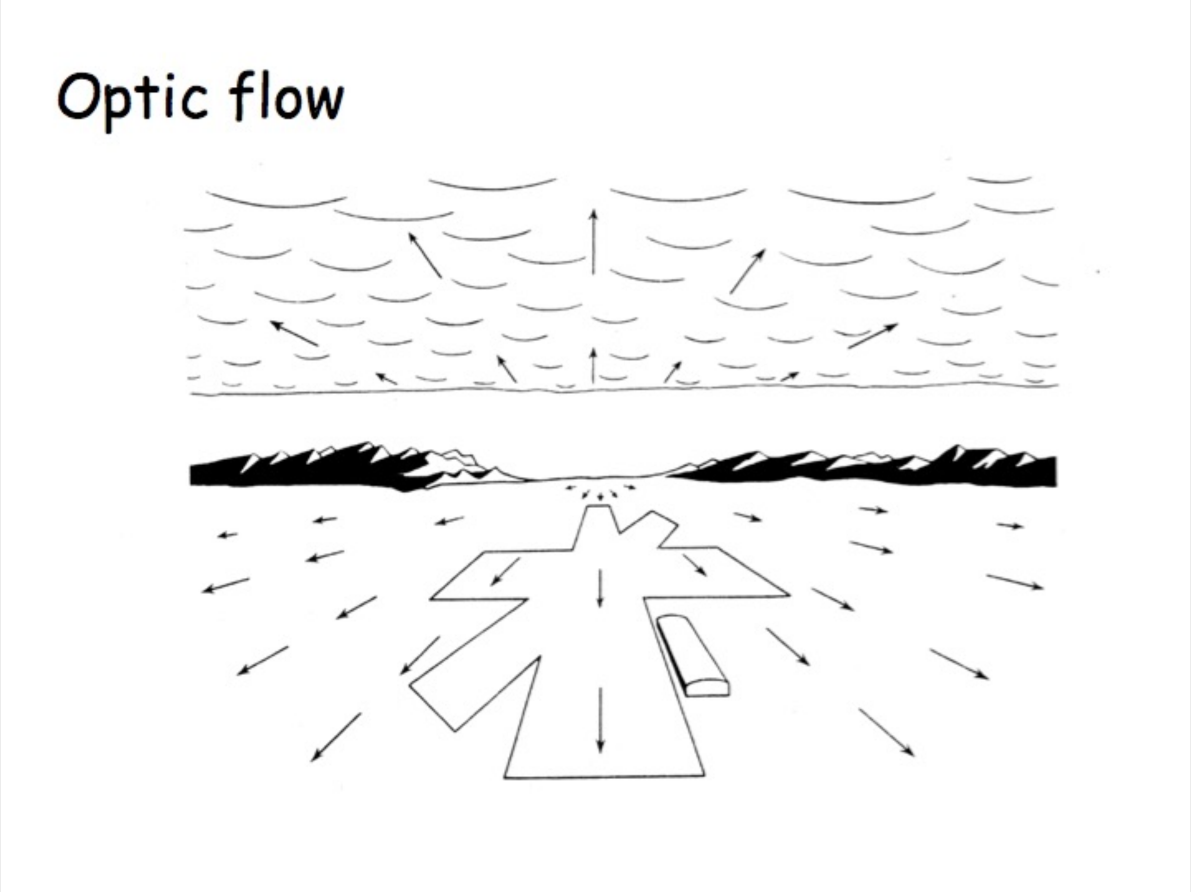
Data, dimensions, and the space



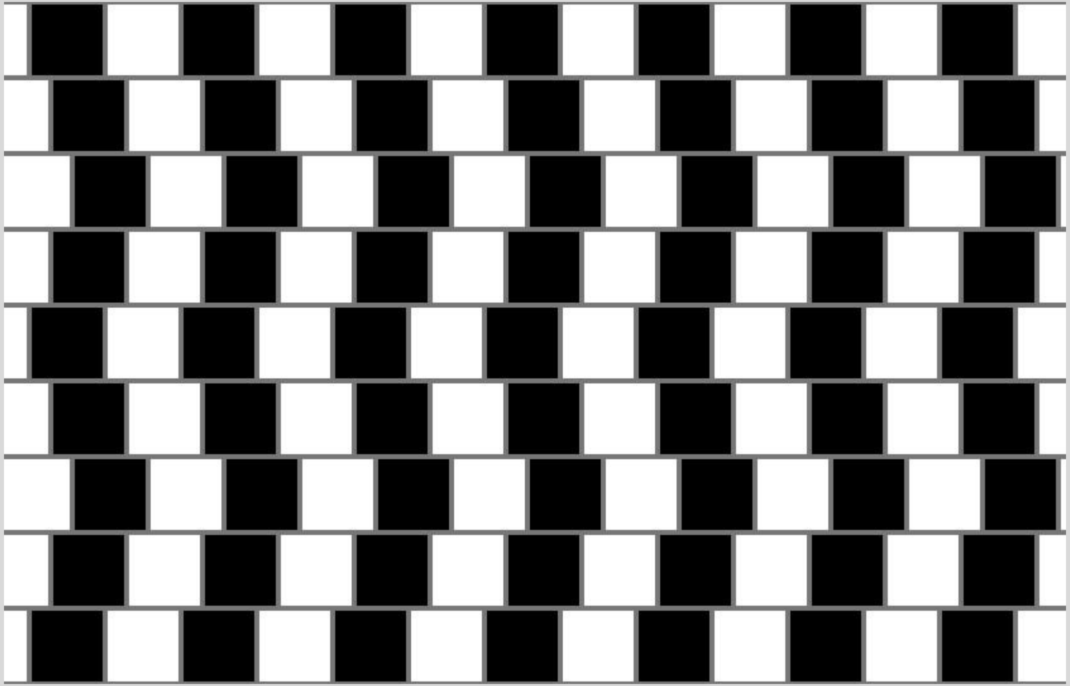


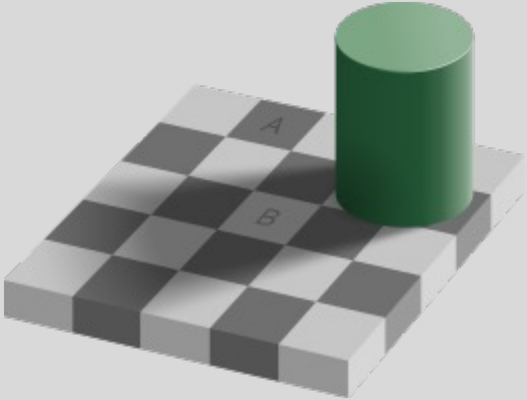
Human perception





Human perception

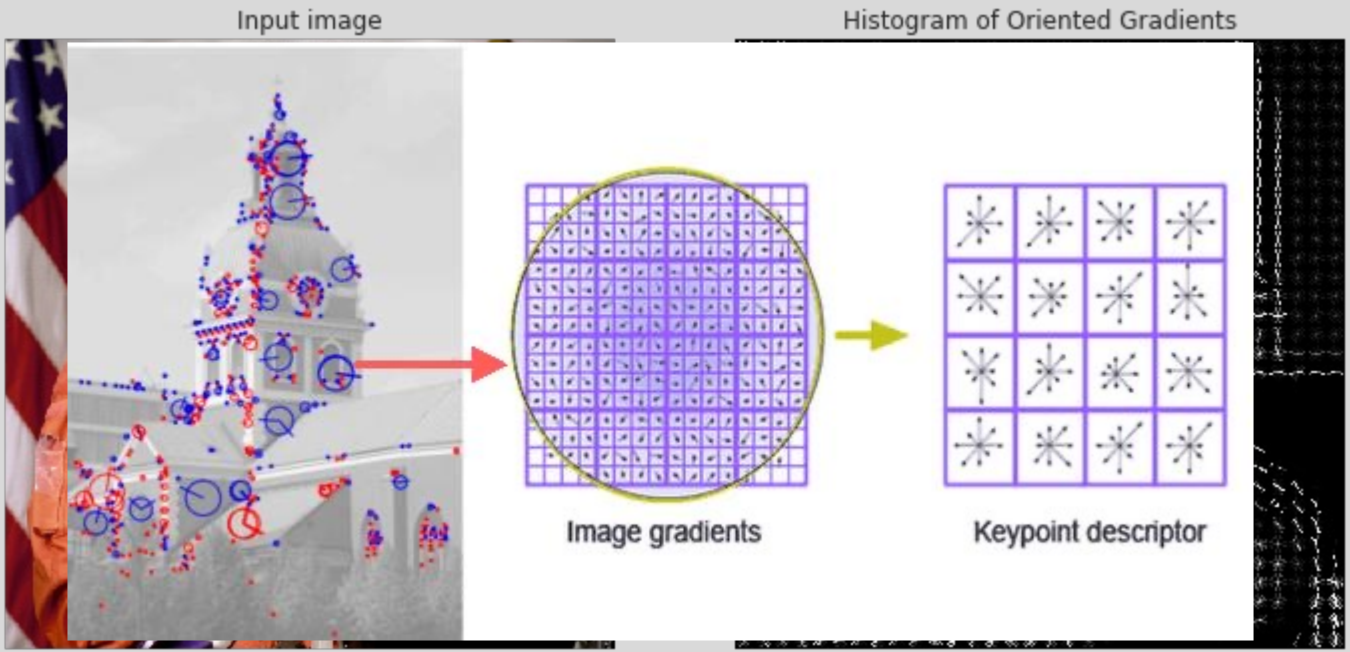




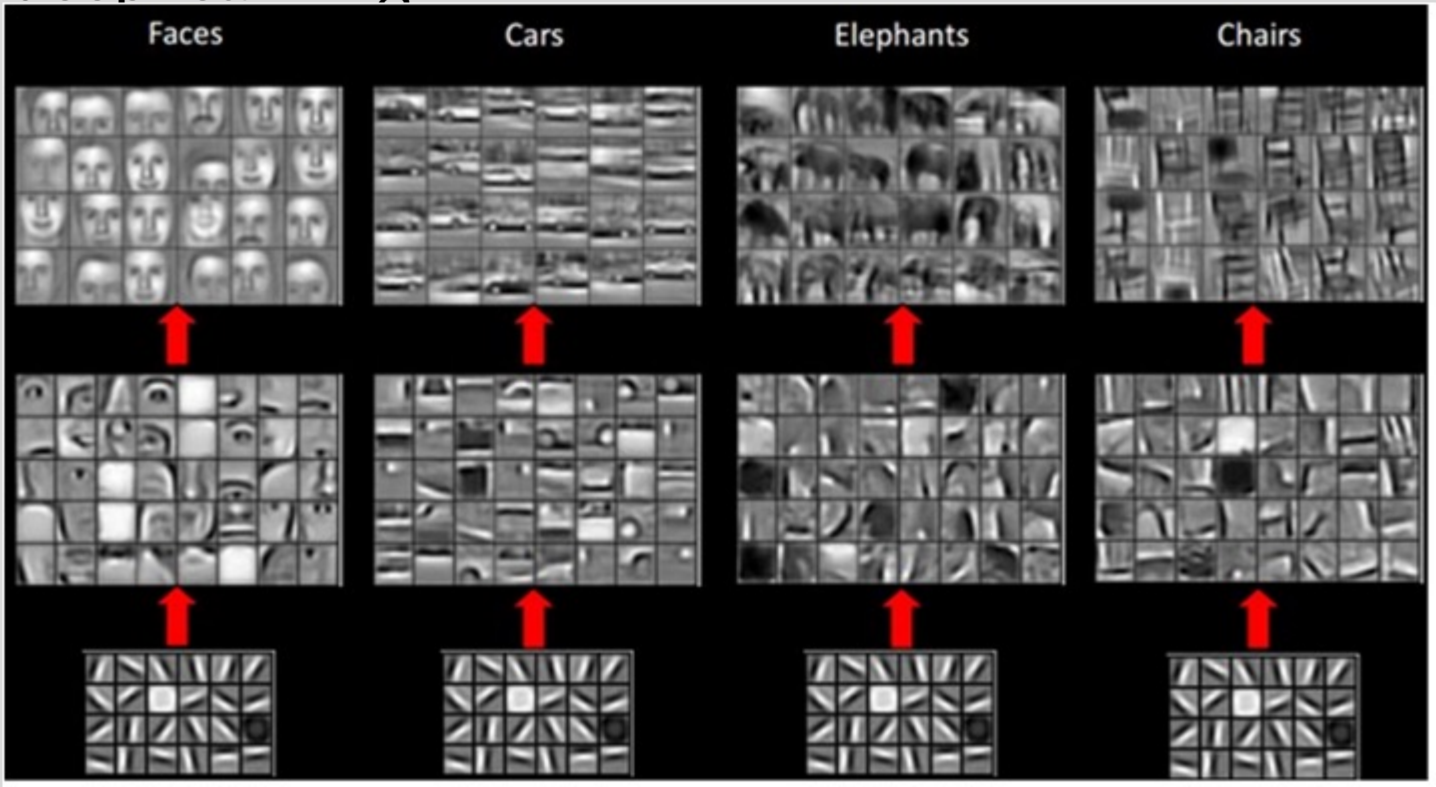
Outline

- What are representations
- Why should we care about learning representations
- What makes a representation good
- How to build (deep) representations
- Structure of the class
  - Homeworks and exam
  - Project
  - Format
  - Topics

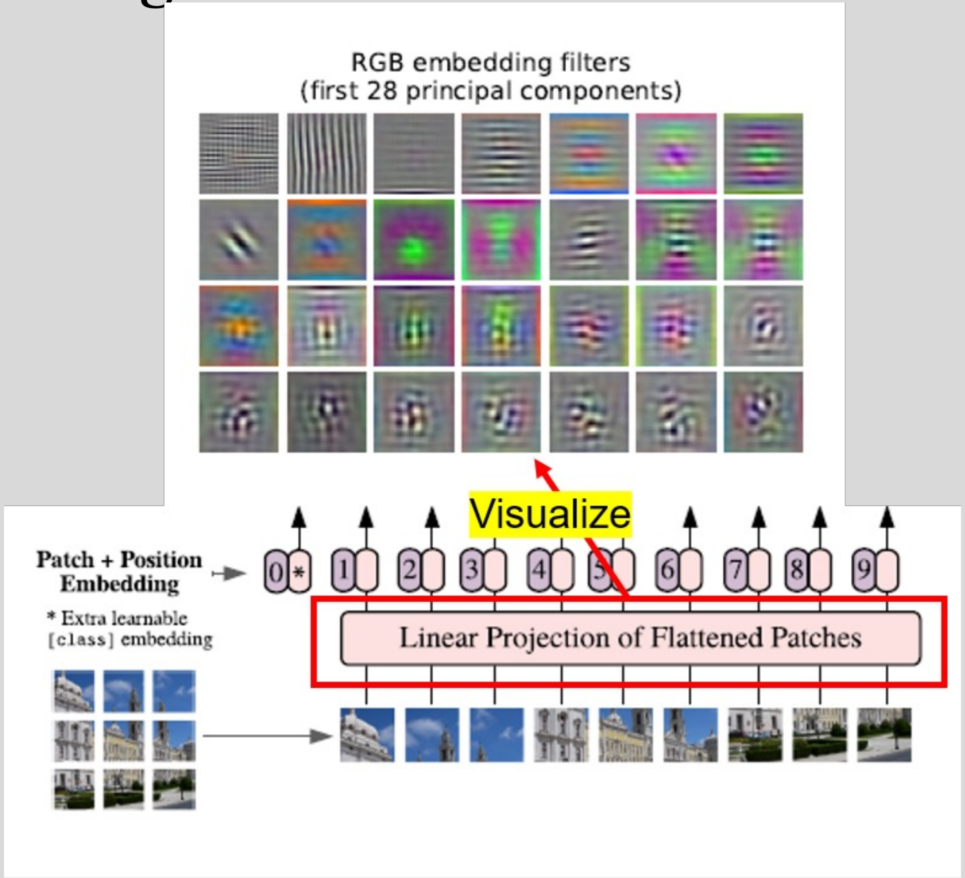
Before deep learning



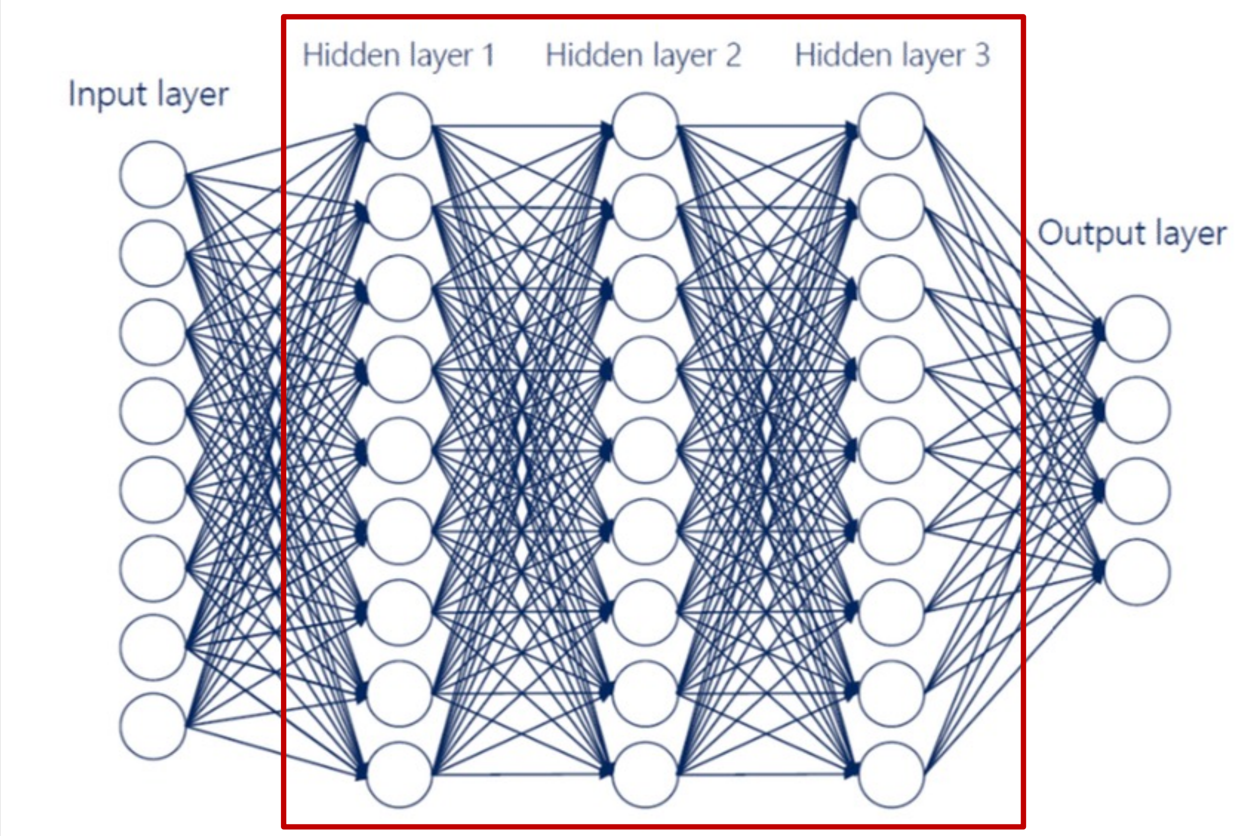
After deep learning



After deep learning



What is deep learning about?



Representation and predictive model

- Do we need "deep" learning models if we are given very good representations to start with?

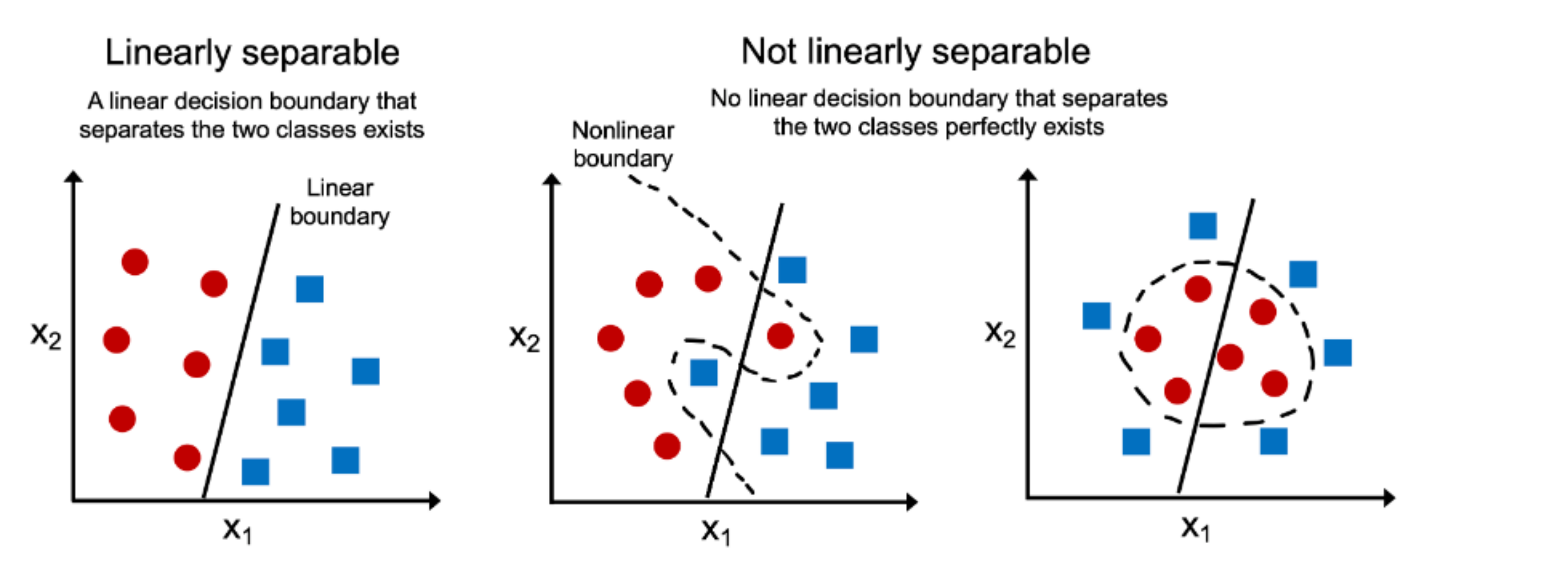
[1,0,0,0,0]

[1,0,0,0,0]

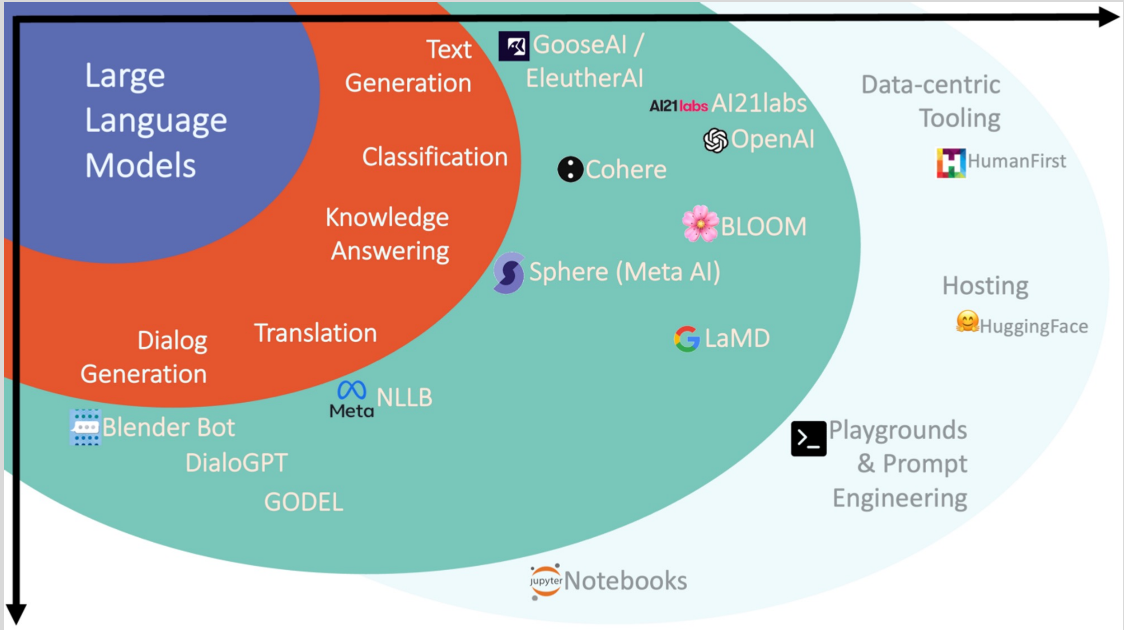
[0,0,0,0,1]

Representation and predictive model

- Do we need “deep” learning models if we are given very good representations to start with?



Representation learning and LLM



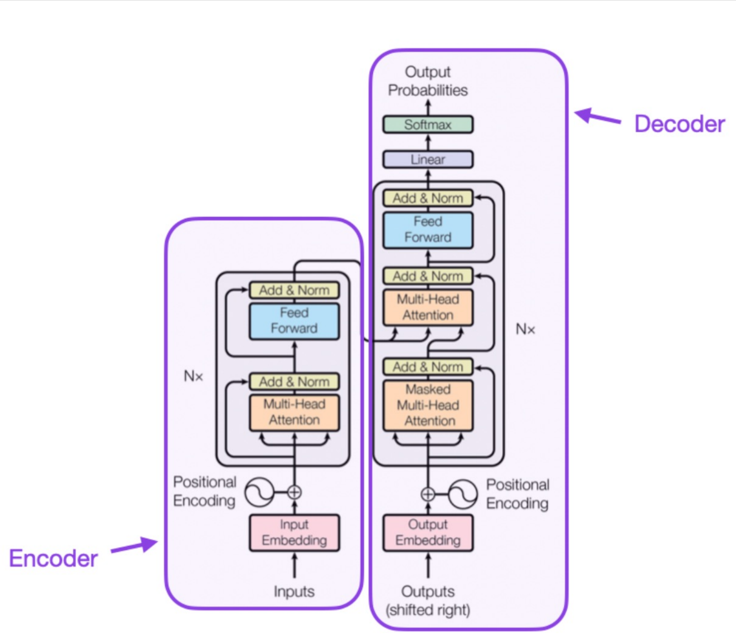


Figure 1: The Transformer - model architecture.

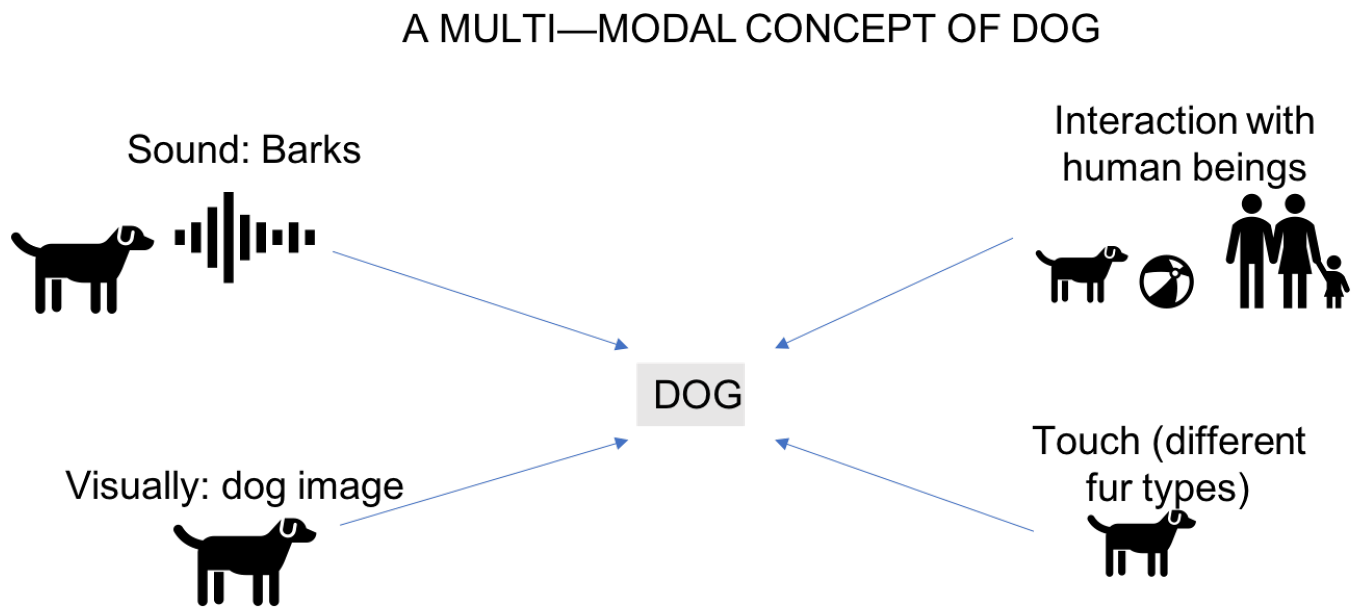
Transfer Learning

Representation learning in different domains.

Transfer Learning

Representation learning in different domains.

Multi-modality Representations

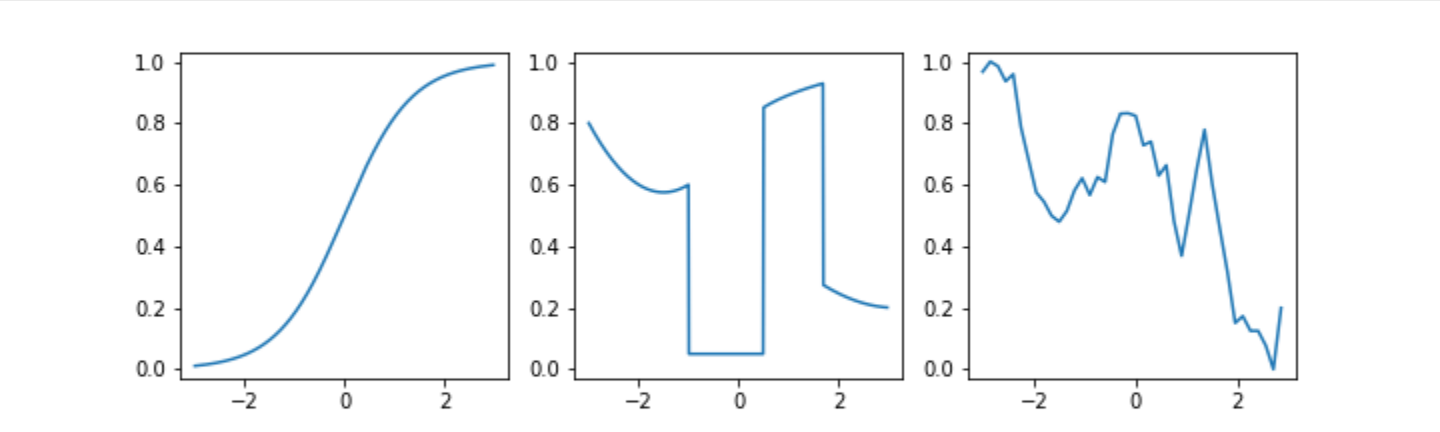


Multi-modality Representations

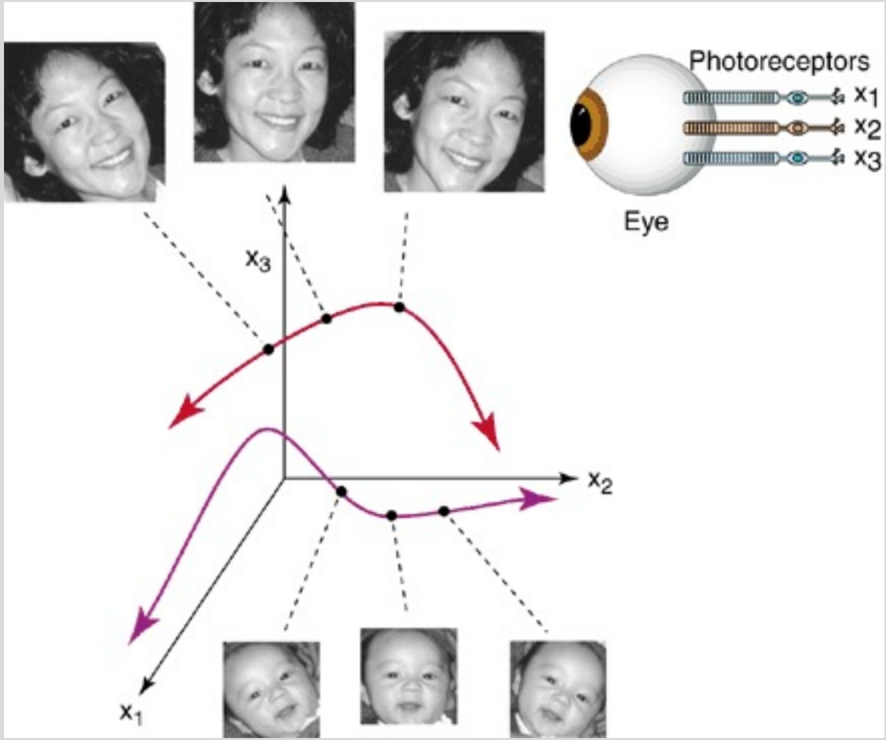
Outline

- What are representations
- Why should we care about learning representations
- What makes a representation good
- How to build (deep) representations
- Structure of the class
    - Homeworks and exam
    - Project
    - Format
    - Topics

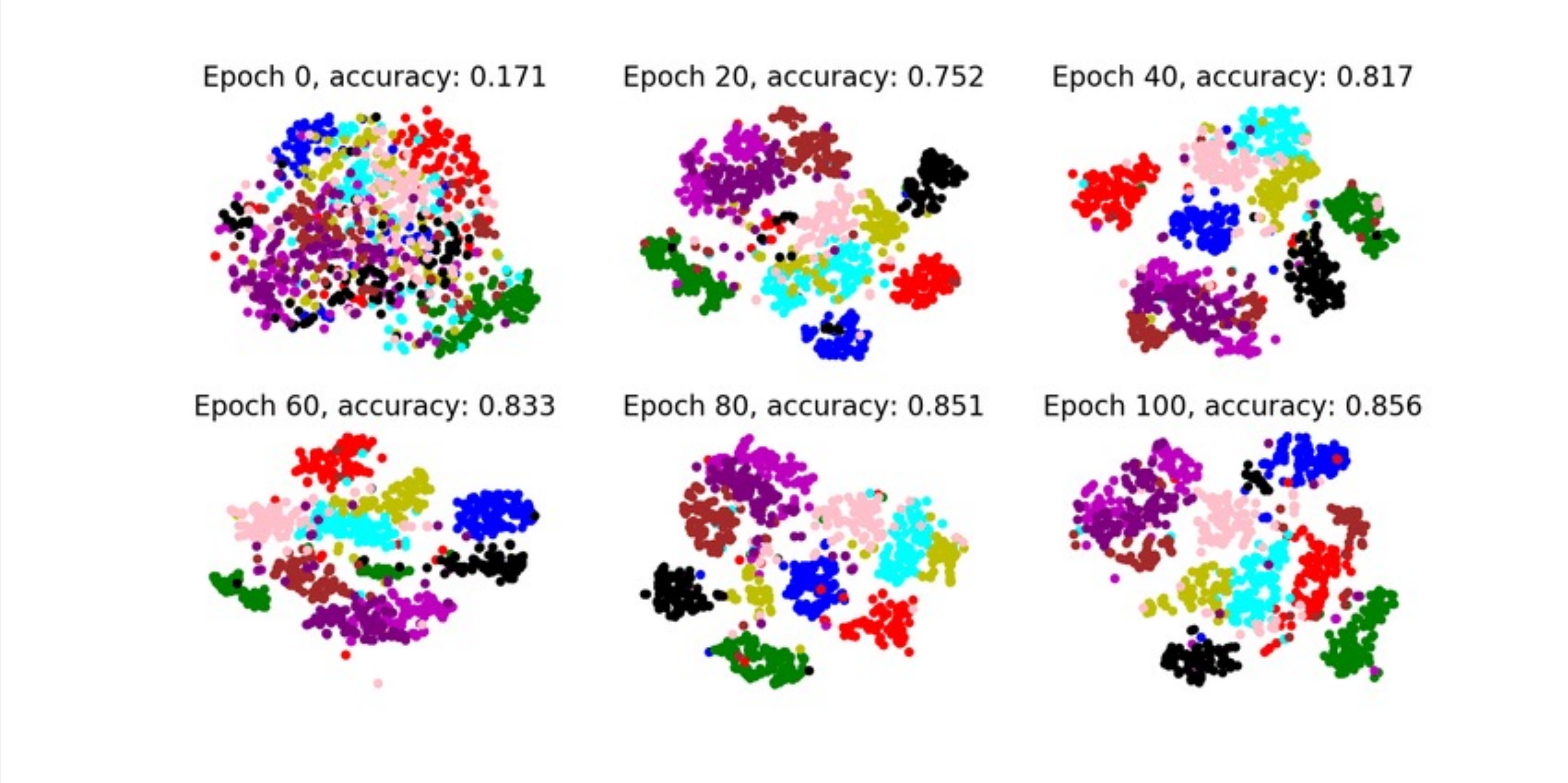
Smoothness



Smoothness

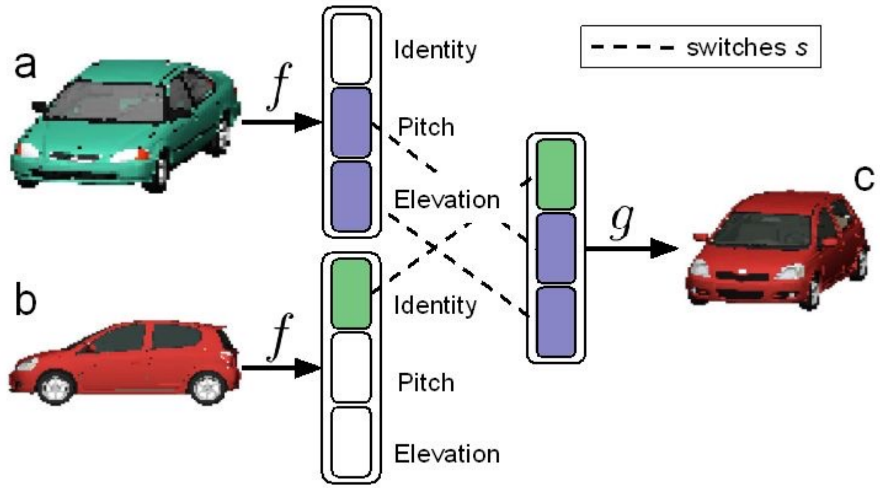


Disentanglement



Disentanglement

Learning a disentangled representation



$$\mathcal{L}_{dis} = \sum_{a,b,c,s \in \mathcal{D}} ||c - g(s \cdot f(a) + (1 - s) \cdot f(b))||_2^2$$

Outline

- What are representations
- Why should we care about learning representations
- What makes a representation good
- How to build (deep) representations
- Structure of the class
    - Homeworks and exam
    - Project
    - Format
    - Topics

Deep neural networks

This class

Outline

- What are representations
- Why should we care about learning representations
- What makes a representation good
- How to build (deep) representations
- Structure of the class
    - Homeworks and exam
    - Project
    - Format
    - Topics

Class webpage

https://jinsungit.github.io/teaching/Fall2024/

Jin Sun

about publications projects teaching

CSCI 8945 Advanced Representation Learning

Advanced Representation Learning is a course designed to delve deeper into the fundamental concepts of representation learning and its applications. In this class, students will explore various representation learning techniques, including both classical and deep learning methods, and learn how to apply these techniques to solve complex problems in computer vision, natural language processing, audio, and other areas. By working on the research project component of the course, the students will develop novel methods and theories about representation learning and prepare manuscripts describing their findings. By the end of this course, the students will have a solid understanding of the state-of-the-art in representation learning and be able to apply these techniques to solve real-world problems.

- **Time and location:**
  - Tue & Thu, 12:45pm-2pm, Boyd 222
  - Wed, 12:45pm-1:30pm, Boyd 222
- **References:**
  - Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville. **Free**
  - Dive into Deep Learning by Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola. **Free**
  - Computer Vision: Algorithms and Applications by Richard Szeliski. **Free**
  - "Machine Learning: a Probabilistic Perspective" by Kevin Murphy.
  - "Foundations of Data Science" by Avrim Blum, John Hopcroft, and Ravindran Kannan.
- **Syllabus:**
- **Learning outcomes**
  1. Demonstrate understanding of machine learning and deep neural network fundamentals.
  2. Gain experience deploying deep learning models in computer vision, natural language processing, and audio domains.
- **Contact:** Annoucements will be made on eLC. You can also send an email to me at *jinsun@uga.edu*.

Class Schedule

Homeworks

- A mix of pen and paper and coding questions.
- 2 weeks to finish.
- HW1: machine learning basics, linear algebra, probability, PCA and MDS, sparse coding
- HW2: image operations, camera model, image editing, word embeddings, NLP tasks
- HW3: Audio generation, graph neural nets, contrastive learning, multi-modal learning

Mid-term Exam

- Pen and paper
- No coding
- Closed-book
- In-class

- More details later in the semester

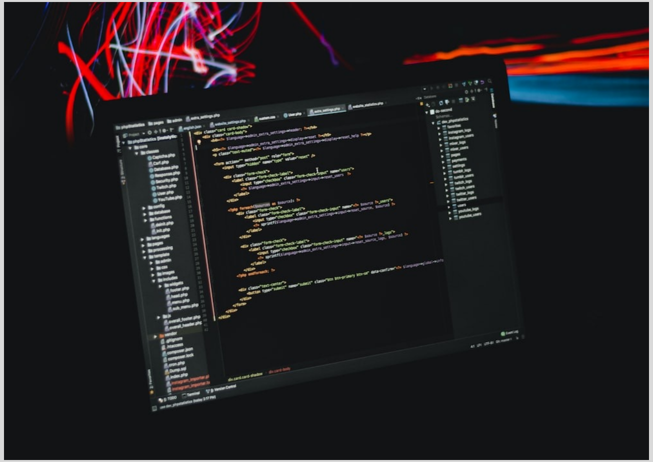
Project

- ~4 people
- Work on the research frontier in representation learning
- Will discuss in groups about the project during the Wed sessions
- Deliverables:
  - Proposal
  - Milestone
  - Final report

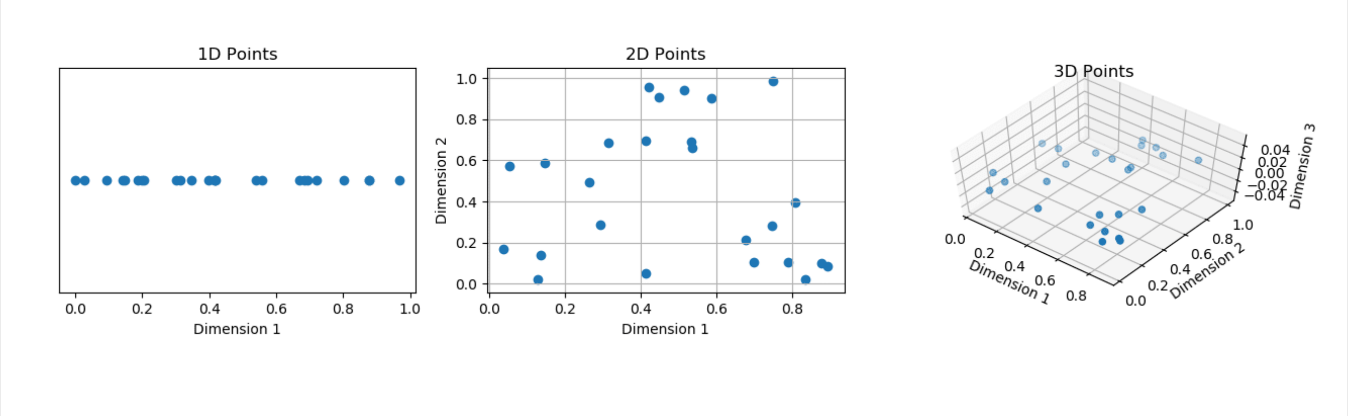
Format

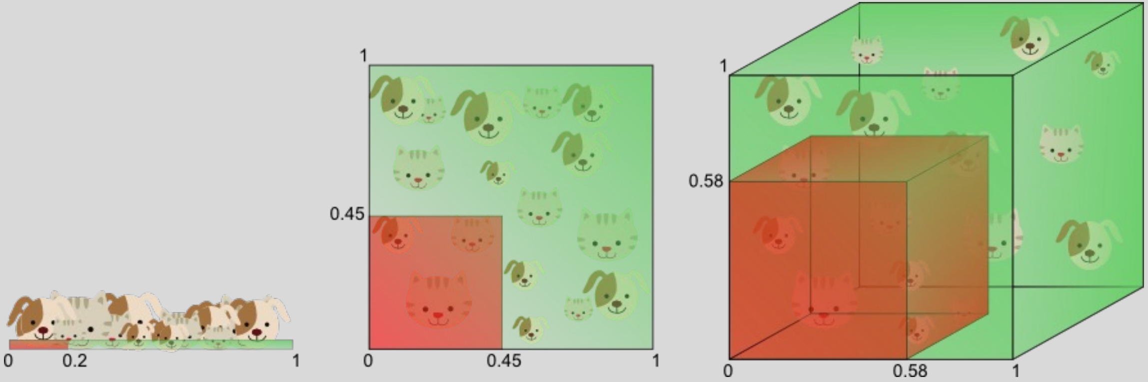
- Tue and Thu sessions:
  - Lecture
  - Paper discussion

- Wed sessions:
  - Project status update and discussion
  - Useful tools and workflows
  - Tips for research: how to read a paper, how to write a paper, how to do research?



1 Introduction and background

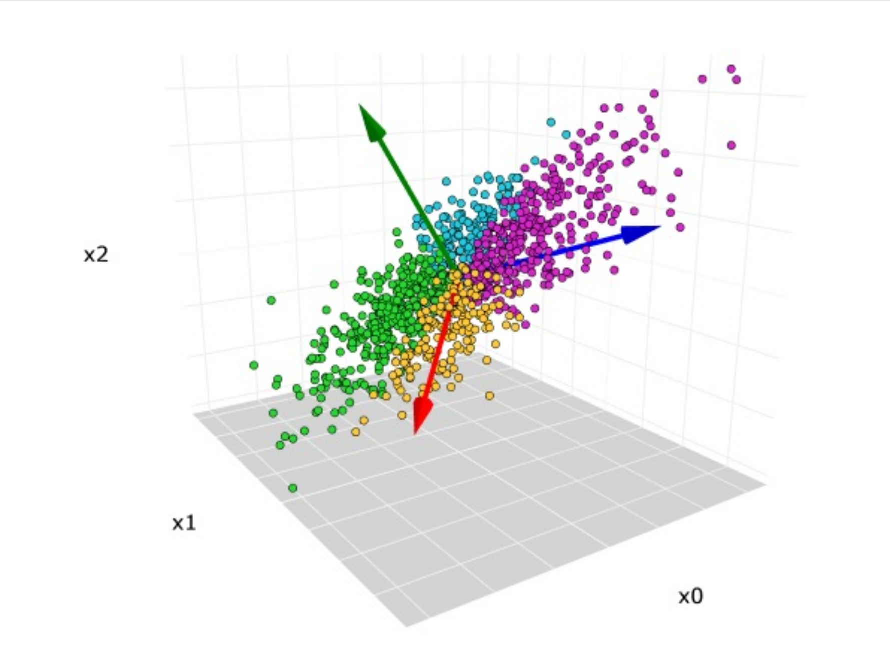




Also essential math concepts

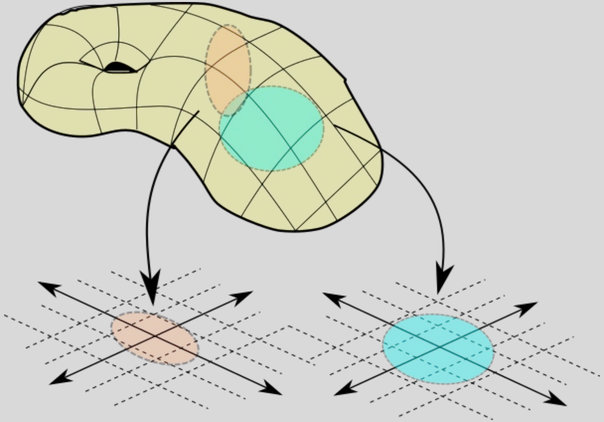
2 Data representation space and structures

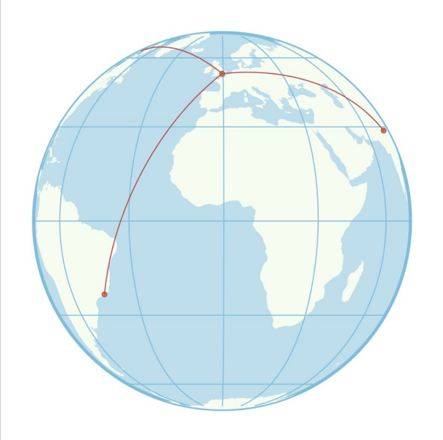
- Dimension reduction
- PCA
- MDS
- Metric learning
- Distances in high dimensional space
- Manifolds
- Subspaces
- Sparse coding



2 Data representation space and structures

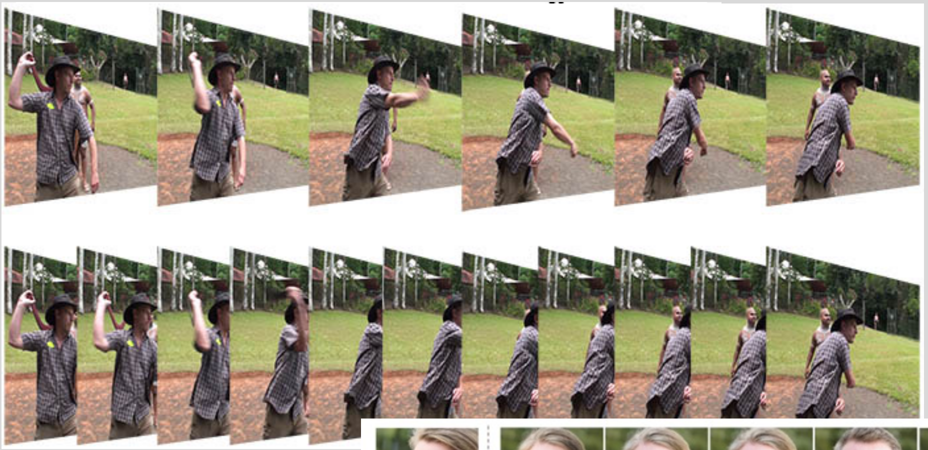
- Dimension reduction
- PCA
- MDS
- Metric learning
- Distances in high dimensional space
- Manifolds
- Subspaces
- Sparse coding

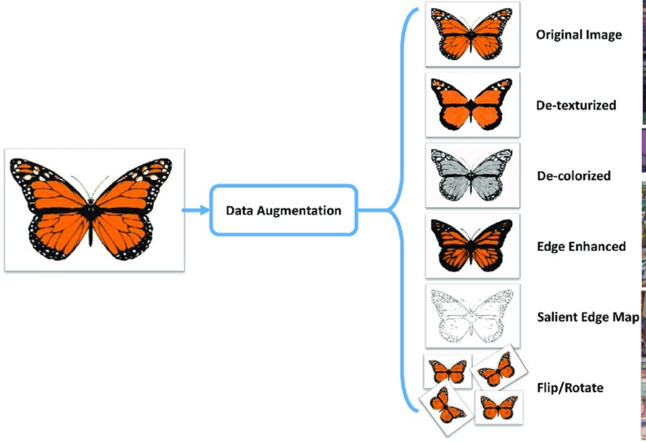


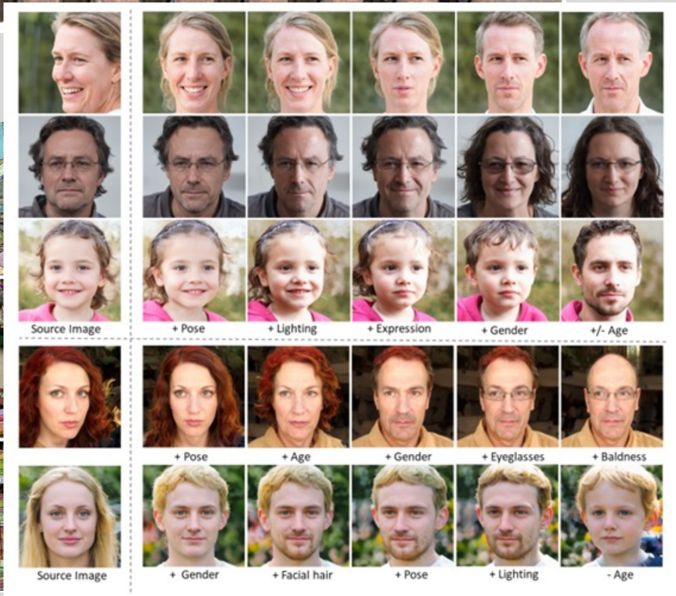


3 Visual representations

- Pixels, 3D points, and cameras
- Image operations
- Visual semantics
- Videos
- Image subspaces

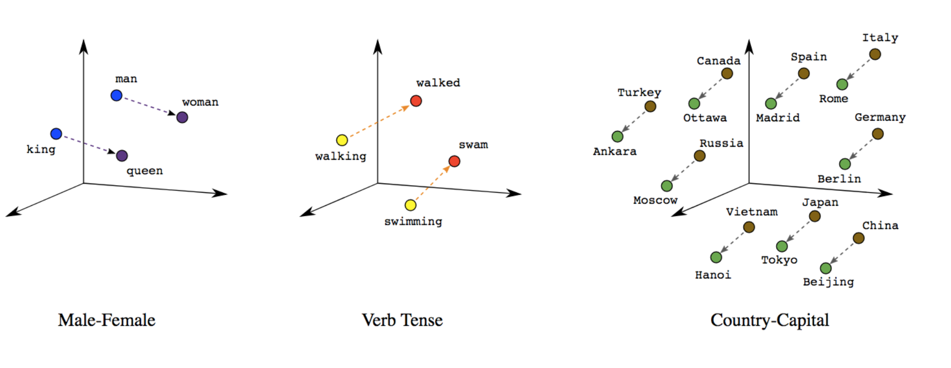


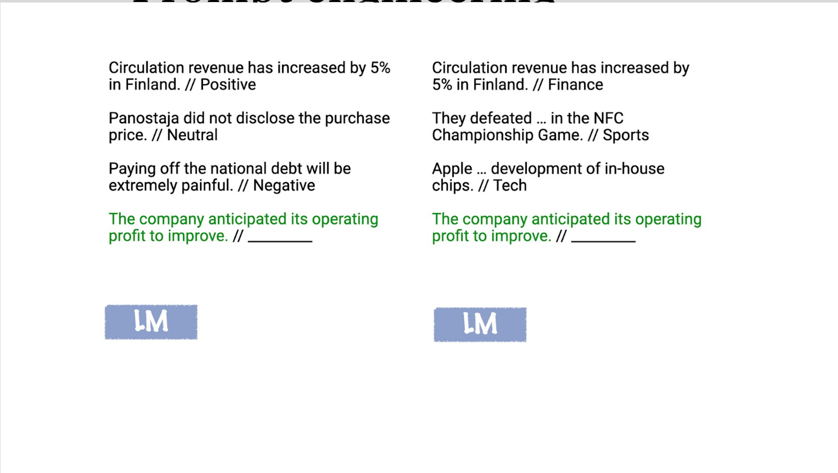


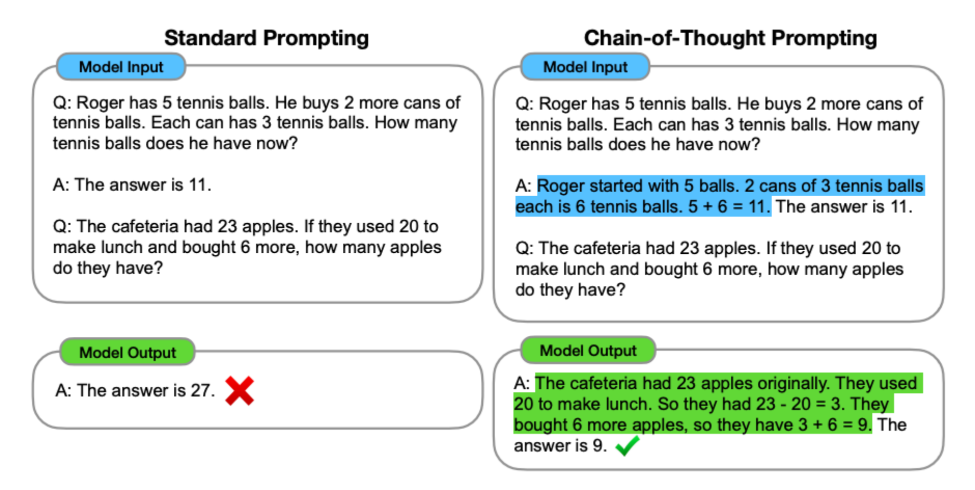


4 Language representations

- Representing words and sentences
- Language model pretraining
- NLP tasks
- Zero-shot and in-context learning
- Prompt engineering

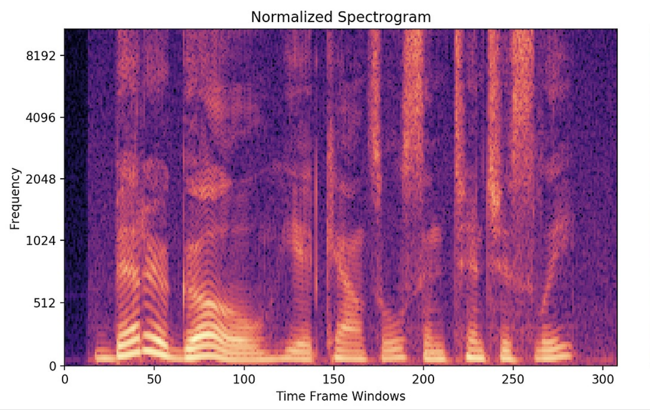






5 Audio representations

- Representing sound
- Audio generation
- Audio editing



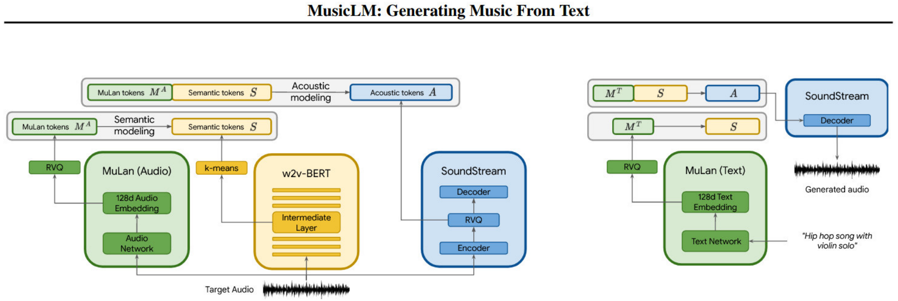
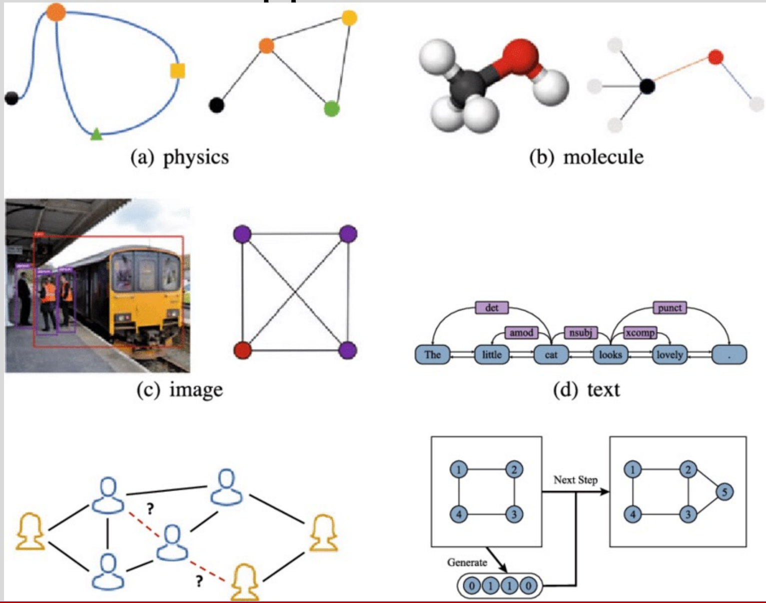
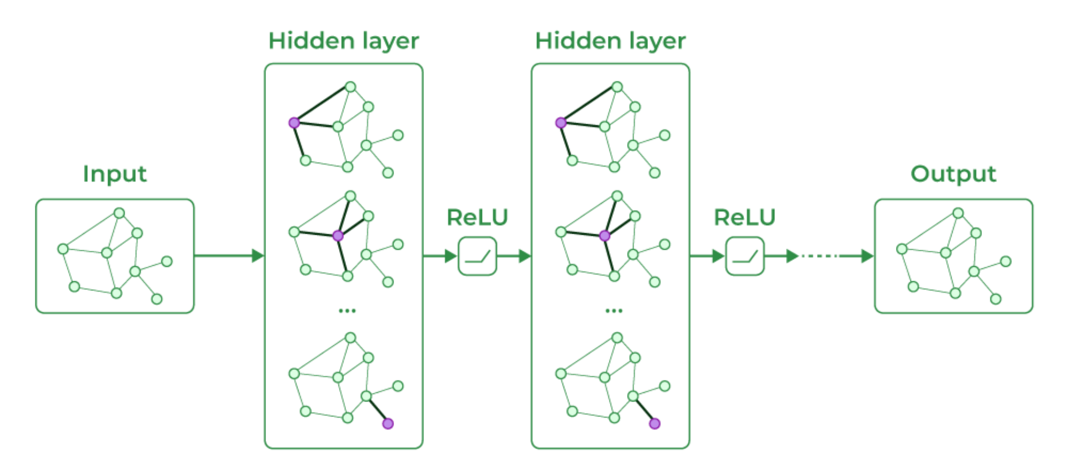


Figure 2. Left: During training we extract the MuLan audio tokens, semantic tokens, and acoustic tokens from the *audio-only* training set. In the semantic modeling stage, we predict semantic tokens using MuLan audio tokens as conditioning. In the subsequent acoustic modeling stage, we predict acoustic tokens, given both MuLan audio tokens and semantic tokens. Each stage is modeled as a sequence-to-sequence task using decoder-only Transformers. Right: During inference, we use MuLan text tokens computed from the text prompt as conditioning signal and convert the generated audio tokens to waveforms using the SoundStream decoder.

6 Graphs

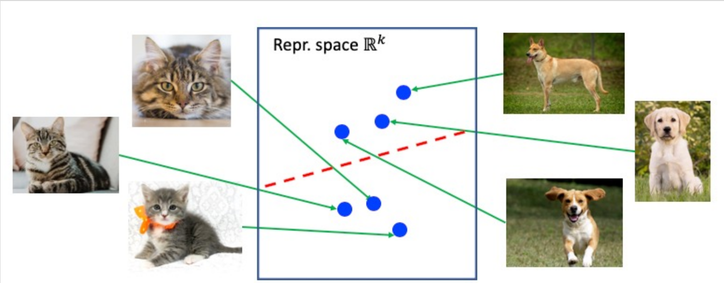
- Graphs and neural networks
- Graph operations and process
- GNN applications

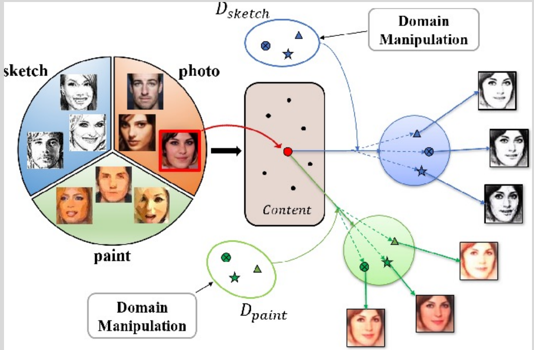


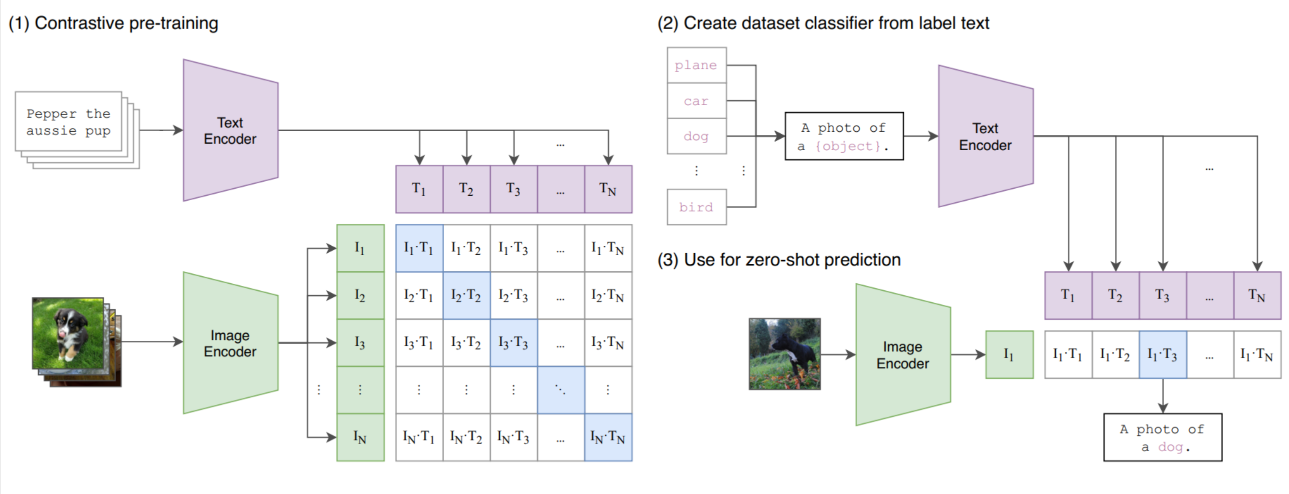


7 Multi-modal representations

- Contrastive learning scheme
- Vision+language learning
- Multiview learning
- Multitask learning
- Cross-modality



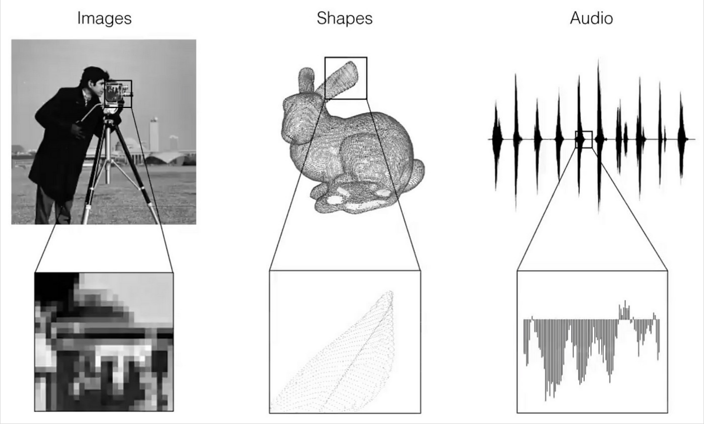




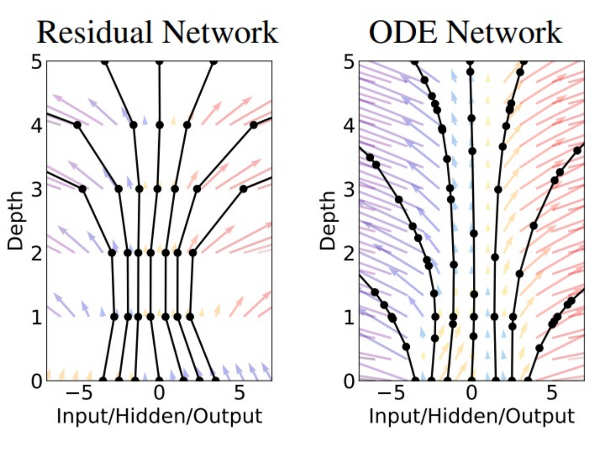
8 Advanced Topics – Implicit representations

- Implicit neural representations

Input	Output supervised by	Implicit Formulation Find Φ that minimizes $\mathcal{L}$
$\mathbf{x} \in \mathbb{R}^2$ spatial coords.	$f(\mathbf{x}) \in \mathbb{R}^3$ RGB values	$\mathcal{L}_{\text{img}} = \int_{\Omega} \lVert \Phi(\mathbf{x}) - f(\mathbf{x}) \rVert \, d\mathbf{x}$



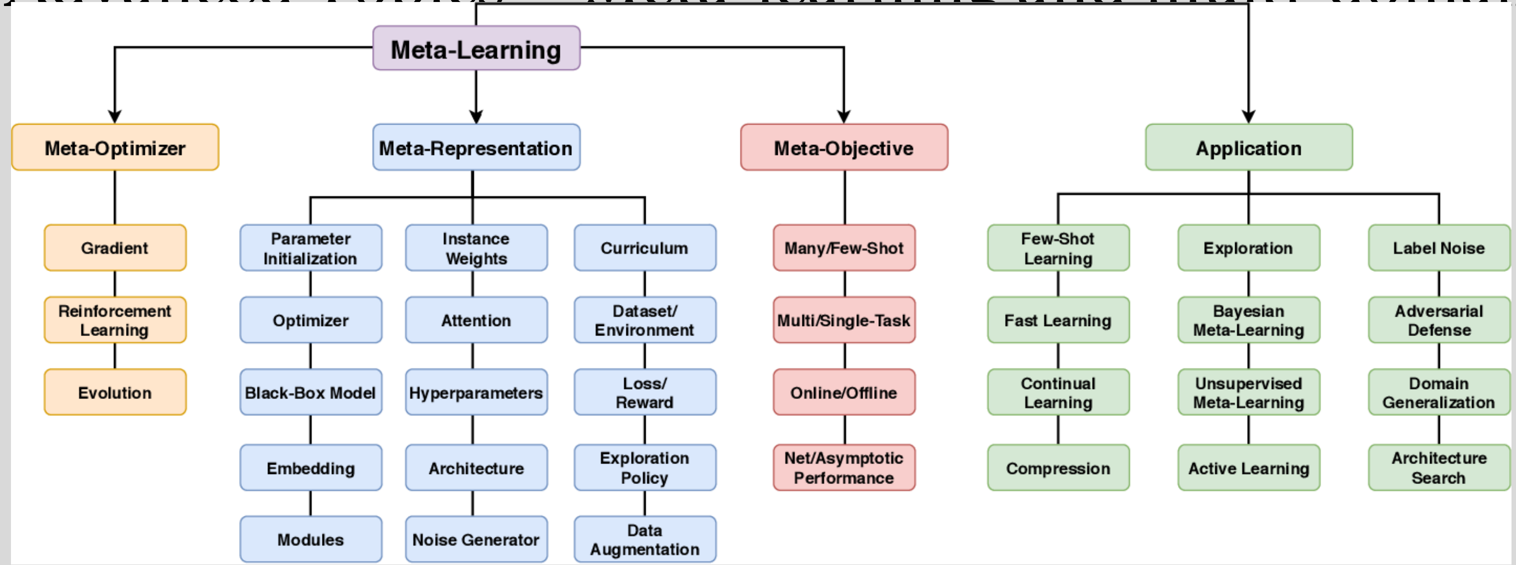
- Neural ODE

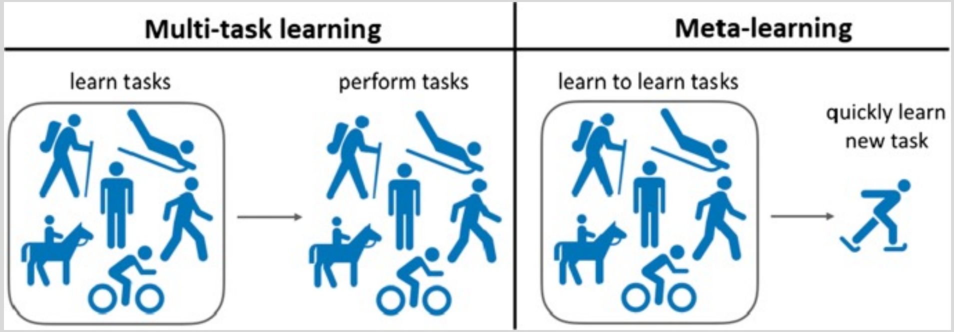


$$\mathbf{h}_{t+1} = \mathbf{h}_t + f(\mathbf{h}_t, \theta_t)$$

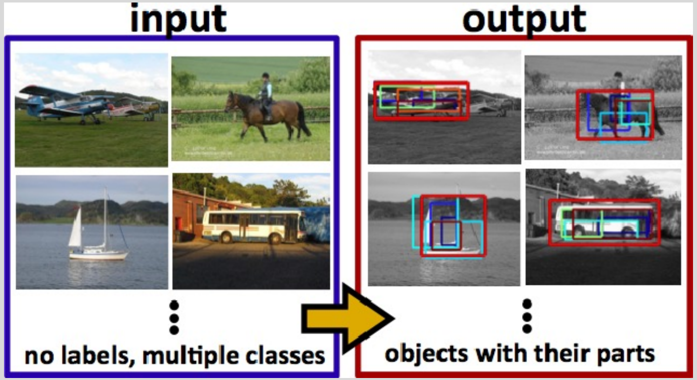
$$\frac{d\mathbf{h}(t)}{dt} = f(\mathbf{h}(t), t, \theta)$$

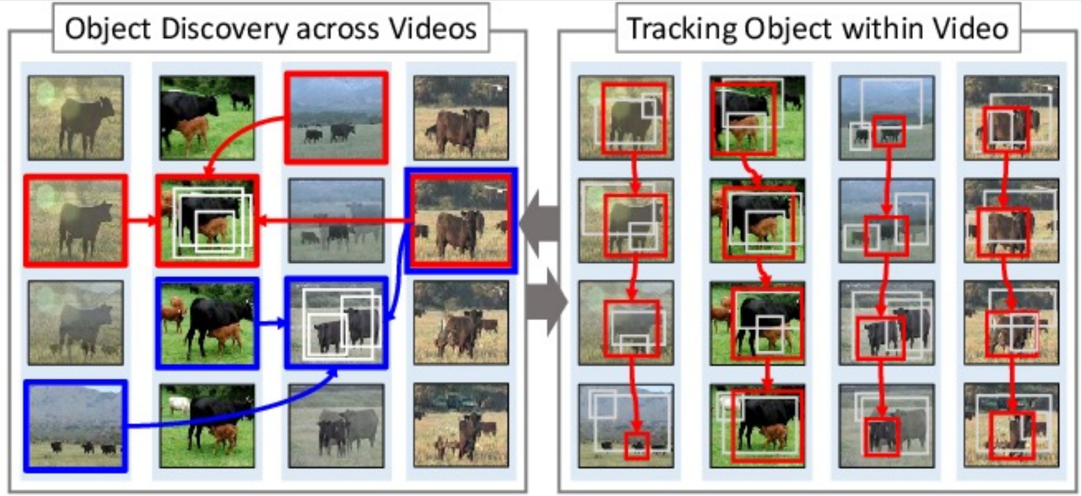
9 Advanced Topics – Meta-learning and multi-domains



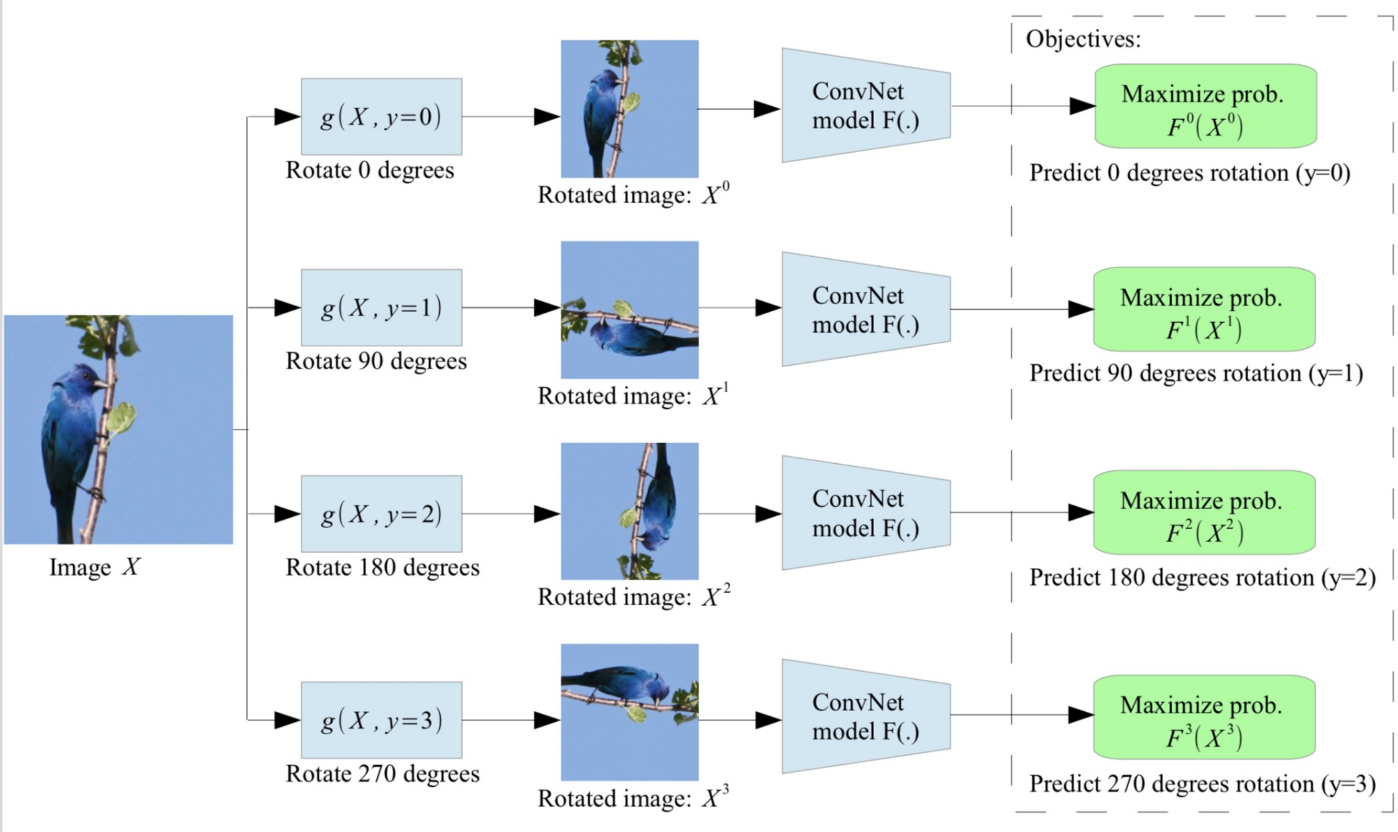


10 Advanced Topics – Unsupervised and self-supervised

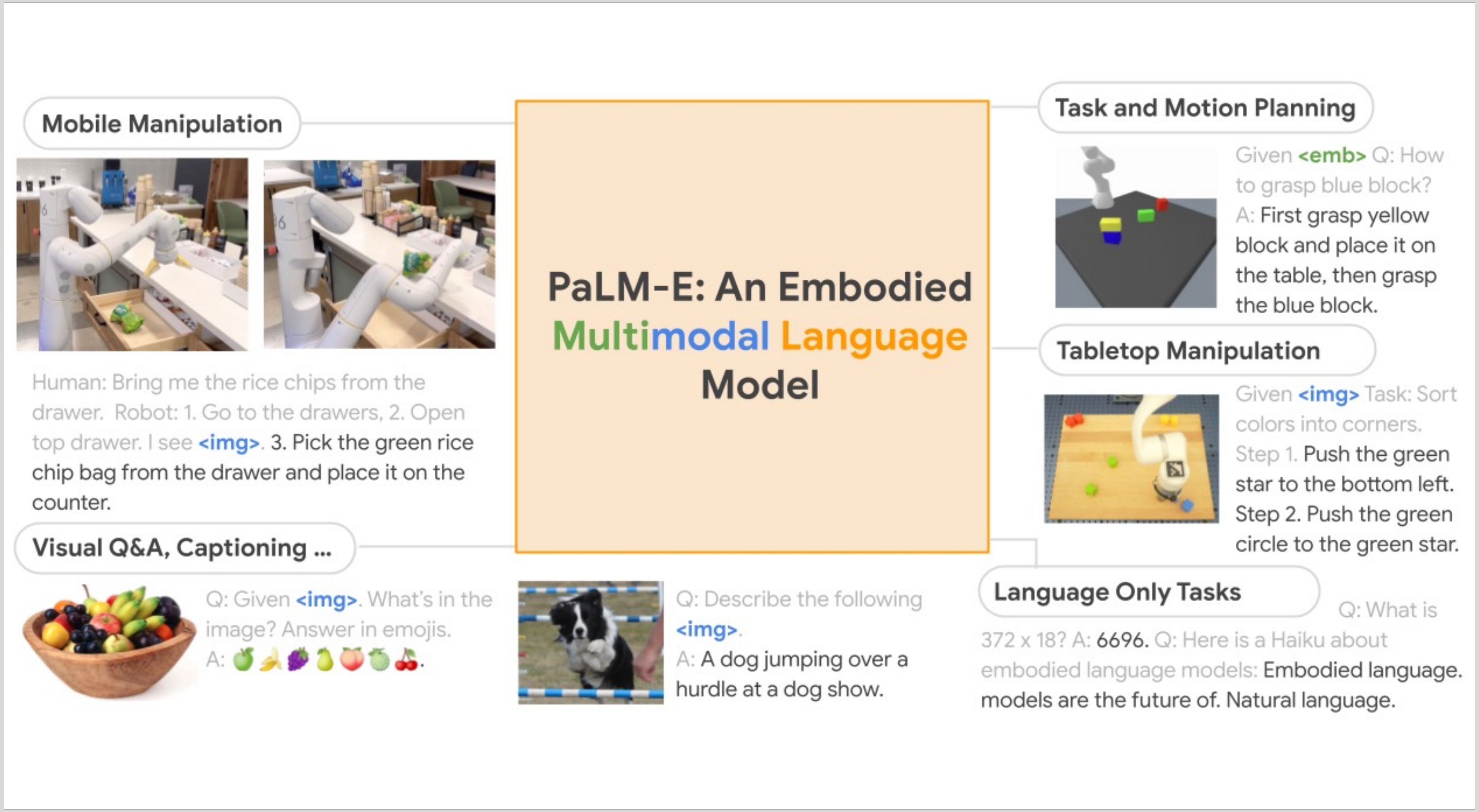




10 Advanced Topics – Unsupervised and self-supervised



11 Advanced Topics – Beyond perception



12 Project Presentation

- Expectation:
  - High quality
  - High novelty
  - Publication-oriented