



Conclusion

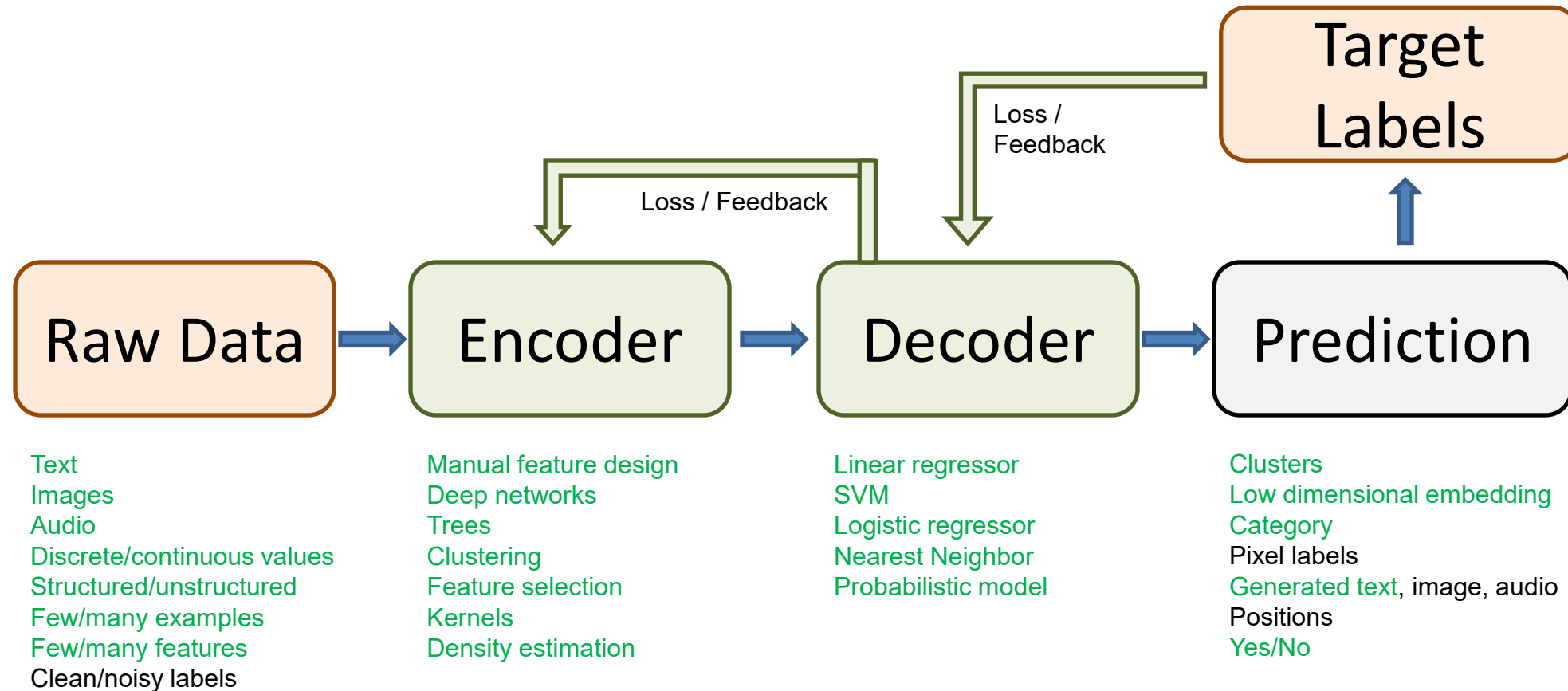
Applied Machine Learning
CS 441

Derek Hoiem

This class: Conclusion

- Birds-eye view of machine learning, recap
- Where to learn more
- Trends and future of machine learning
- Feedback and closing remarks

We've learned a lot



We've covered what you need to get started for most ML applications

- Solid foundation in core ML concepts
 - Classification, clustering, dimensionality reduction, objectives, losses, regularization, generalization, experimental setups
- Intuition, some math, and application of classic methods
 - Logistic regression, Linear Regression, SVM, KNN, Neural Nets, Naïve Bayes, Boosting, Trees, Ensembles, PCA, K-means, Kernel Density Estimation, EM algorithm, PCA
- Exposure to state of the art recent methods
 - Deep convolutional networks, transformers, CLIP, GPT, UMAP
- Data representation and application domains
 - Images, text, audio, general data
- Practical considerations for deployment of ML applications
 - Societal impact, bias and mitigation, process for creating ML applications

Predictors: Three basic kinds

1. Linear (Regression, Logistic Regression, SVM)
 - General purpose for classification/regression
 - Benefits from feature learning
2. Instance-based (KNN)
 - Zero training time, flexible
 - Benefits from feature learning
3. Probabilistic:
 - Works well if a parametric form is known, or with tree ensembles

Feature learning: Two basic kinds

- Tree
 - Learn how to partition input space to group examples together in a discriminative way
 - Followed by instance/linear/probabilistic prediction
 - Easily combine unnormalized continuous and discrete features
 - Especially powerful with ensembles, as in Random Forests or Boosted Trees
- Neural network
 - Jointly train feature representation with linear predictor
 - Major advantages for structured data with convolutional or transformer architectures, or use for retrieval/similarity
 - Representations can be learned on one problem and tuned or applied for another problem in the same domain

Go-to approaches for **classification**

- K-NN
 - Super easy and sometimes surprisingly effective, good sanity check
- Linear SVM or linear logistic regression
 - Easy to optimize, often works well
- Random forest or boosted trees
 - Highly effective with minimal fuss, very flexible
- Linear probe or fine-tune deep network
 - Best for common structured data like images, text, audio

Go-to approaches for **regression**

- Linear regression
 - Interpretable and often works well
 - Sometimes important to transform features or targets so they are better captured by a linear model
- Random forest
 - Highly effective with minimal fuss
- K-NN
 - Especially helpful when predicting multiple correlated values, e.g. which patch can be used to fill a hole in the image

In supervised learning, always remember

- **Clean experiments**
 - training set to learn model parameters
 - validation set to select method and hyperparameters
 - test set for final performance evaluation
- **Bias/variance trade-off**
 - Avoidable error is due to challenge in fitting parameters (variance) and inability to perfectly fit the data (bias)
 - Model designs and hyperparameters often trade off between these, e.g. increasing model complexity can increase variance but reduce bias
 - Ensembles work around this trade-off, and modern deep networks often act like ensembles
- People and algorithms use mental shortcuts that lead to a kind of **bias that can be harmful to society** – transparency and fairness require conscientious sourcing, development, and evaluation

Data organization

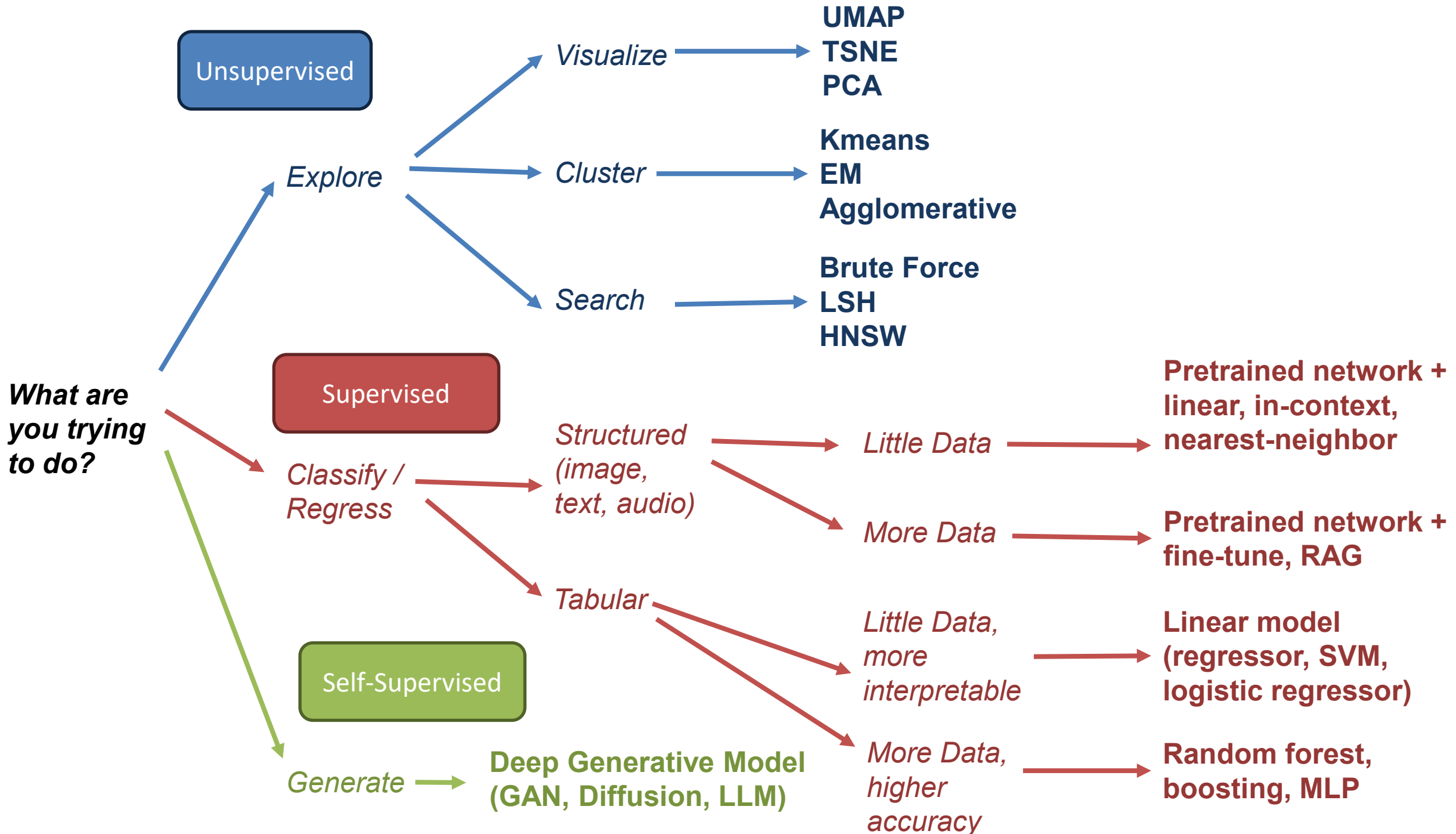
- PCA is used to compress a vector into fewer values in a way that can be decompressed with minimal mean squared difference
- Clustering reveals common modes of data
 - K-means is an essential algorithm
- Search for similar data items is an important application and is the computational foundation for clustering
 - FAISS is a very useful library for efficient search
 - Approximate search, e.g. trees and LSH, are needed when speed is a priority
 - Representation learning (e.g. with deep networks) is needed to make similarity meaningful

It's all about the data

- Model architectures and computational techniques get all the attention, but there are often many reasonable choices that perform similarly
- To get best performance, data requires more thought and effort than algorithms
- Need creative ways to get freely supervised data, as well as careful curation of evaluation sets

Building machine learning applications is easier and more effective than ever

- Great libraries like sklearn and pytorch
- Great models like CLIP, GPT, BERT
- Leverage existing models whenever possible to minimize development cost



Choosing the right tool for the job

<https://tinyurl.com/441-L26-fa24>



Choosing the right tool for the job: answers

Predict the sale price (\$) of a house given location, square footage, year built, and other information typically included in real estate listings. Training data has ~50K samples.

- Nearest neighbor classifier
- Naive Bayes Classifier
- Logistic regression
- Linear regression
- Random forest regressor
- Random forest classifier
- Convolutional network

Identify the plant type from a photo of a seed on a white background. You have 100 examples per plant type in training data.

- Nearest neighbor classifier
- Naive Bayes Classifier
- Logistic regression
- Random forest classifier
- Convolutional network, trained from scratch
- Convolutional network, fine-tuned based on an ImageNet pretrained model

Random forest regressor

- Regression problem
- Mixed continuous and categorical features
- Lots of training data available

Fine-tuned ConvNet

- Image classification (use deep network to take advantage of structured image data)
- Limited data, so fine-tuning from pre-trained is much better than training from scratch

Choosing the right tool for the job: answers

Given the text of a review, assign a score from 1 to 5 indicating how positive the review is. You have 1M training samples.

- Use Naive Bayes Classifier on word count
- Use GPT-4
- Fine-tune a BERT model
- Boosted decision tree on word count

Automatically identify which students are present from a photo of the classroom. Training data is one headshot per student. Face detector is available to localize faces.

- Use nearest neighbor on an ImageNet pretrained model
- Train a face classifier on Labeled Faces in the Wild (LFW), and then use nearest neighbor on the trained encoder features
- Train a face classifier on LFW and then fine-tune on the training set
- Train a MLP from scratch
- Use PCA on the patch of face pixels and then nearest neighbor

Fine-tune BERT

- Using a deep network language model gives more effective features than word count
- Plenty of data for fine-tuning
- GPT-4 is slow/expensive and not easy to tune

NN w/ LWF-pretrained model

- Training a deep network face classifier on a large dataset can provide good features
- Only one example per face available, so use nearest neighbor (fine-tuning will overfit)

Choosing the right tool for the job: answers

Problem that involves a combination of image and text analysis, with lots of training data available.

- Nearest neighbor
- Logistic regression
- Convolutional network model
- Transformer model

Image classification problem with 10 classes and ~10 examples per class available for training.

- Deep network, trained from scratch
- Deep network with pre-trained model and linear classifier ("linear probe") on features
- Fine-tuned deep network

Transformer

- Both images and text can be represented as tokens that can be processed together with the transformer
- Plenty of data for training

Pre-trained model w/ linear classifier

- Use good features from pre-trained model
- Probably not enough examples to effectively fine-tune and validate

Where to learn more: courses

Many related 400-level courses, e.g.

- CS 412: Data Mining
- CS 442: Trustworthy ML
- CS 443: Reinforcement Learning (SP)
- CS 444: Deep Learning for Vision (SP)
- CS 445: Computational Photography (less focus on learning)
- CS 446: Machine Learning
- CS 447: NLP (SP)
- CS 448: Audio Computing Lab (SP)
- CS 449: Robot Perception

Also many 500-level courses, often requiring background in ML and some domain, e.g.

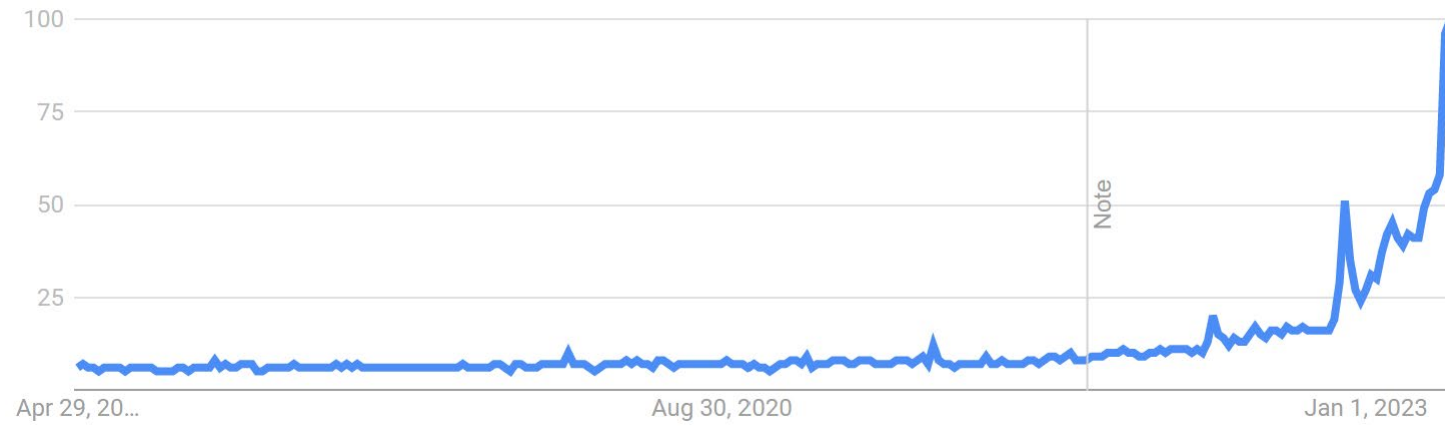
- CS 543 (Computer Vision)
- CS 545 (ML for Signals)
- CS 598 (Adv Robotics)
- CS 562 (Topics in Security and Privacy)
- Many relevant 598 options

Where to learn more: try things

- Many datasets, github repos, and libraries available
- Easy to check out state-of-art in many problems or to try out different related algorithms

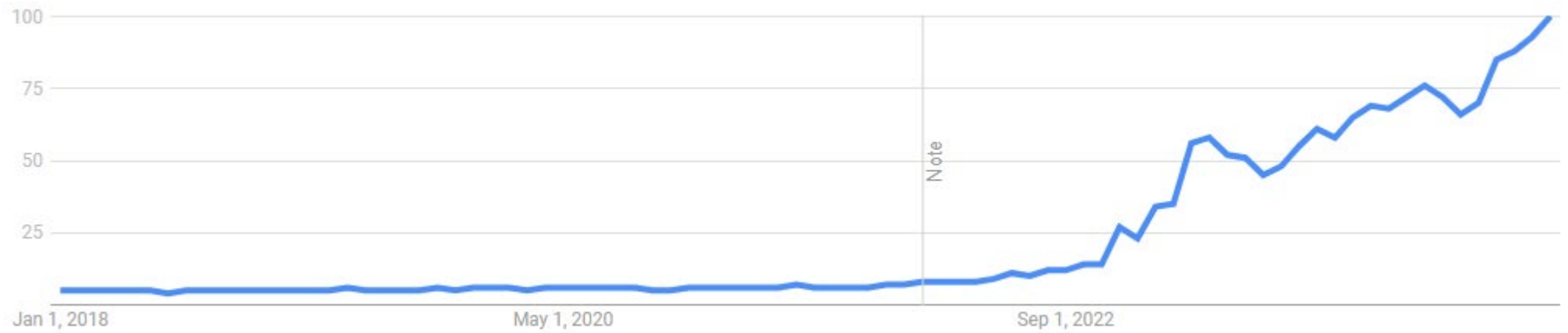
ML is changing

Google Trends: AI (2018-2023)

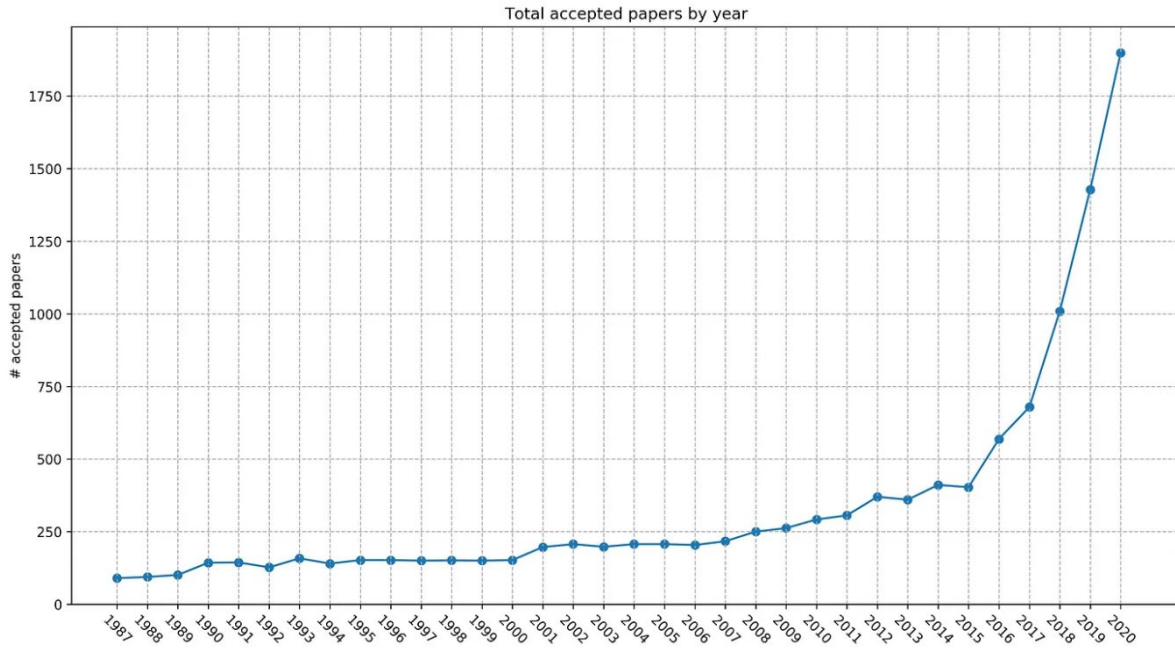


ML is changing

Google Trends: 2018-2024



ML is changing



Papers published in NeurIPS

- 2,672 in 2022
- 3,584 in 2023
- 4,033 in 2024

10x growth in 10 years



Unique Authors in NeurIPS

Increased from 1,064 in 2014 to 13,012 in 2023

13x growth in 9 years

ML is changing

- Machine learning has been used for decades, e.g. reading addresses for the post office, but it's now everywhere
- Major increases in conference attendance, press coverage, investment by companies, number of authors per paper
- Much better infrastructure for ML in terms of libraries and models, as well as tutorials and other online resources
- Short-term value in AI is more clear, causing many companies to focus their investments

AI/ML will change the world in unpredictable ways

- We have seen transformations of agriculture and manufacturing – AI will transform services
 - In a three sector breakdown, the first two sectors of raw materials (including agriculture) and production have already been transformed by technology
 - In consequence, the majority of the workforce has move into services in advanced countries, as well as people working less on average
 - AI may dramatically transform services, but it's not clear where the workforce can move
- Previous products were predictable; AI, not so
 - Objects and computer software are under direct control of their users and operate in predictable ways
 - ML-based products depend on data in non-transparent ways

How to prepare yourself for a fast-changing tech world?

Master the fundamentals

- Tools and algorithms change, but the ideas behind them are relatively constant
- Fundamentals are math, ability to frame problems computationally, ability to draw insights and predict from data, communication, and empathy
- These abilities enable you to learn the latest, design solutions, and apply effectively

Current state of AI/ML

- Fundamental ideas of classification and regression are well established and haven't changed in many years
- Recent advances enable integrating lots of data to create powerful predictors/generators

Generative AI: Text

Write a one paragraph fairytale synopsis about a girl that finds a magical flower that saves her ailing father.

Once upon a time, in a quaint village nestled at the edge of an enchanted forest, lived a young girl named Elara. Her father, the village blacksmith, fell gravely ill, and no remedy seemed to help. Desperate, Elara ventured into the mystical woods, guided by tales of a legendary flower with unparalleled healing powers. After braving countless trials and befriending magical creatures, she discovered the radiant Moonblossom, blooming under the light of the full moon. With unwavering hope, she brought the flower back and brewed a potion from its petals, curing her father and restoring joy to their home. The village celebrated Elara's bravery and the magical bond between a daughter's love and nature's wonders.

Generative AI: Images

Generate a series of three images that illustrates this story.



Elara standing at the edge of an enchanted forest, determined to find the magical flower.



Elara in the middle of the enchanted forest, reaching out to the glowing Moonblossom surrounded by magical creatures.



Elara back in her village, standing beside her now healthy father, with villagers celebrating her success and bravery.

Generative AI: Audio

Write a short catchy song about this story with three verses and no chorus.

Verse 1:

In a village by the forest deep,
Lived a girl named Elara, dreams she'd keep,
Her father sick, hope fading fast,
She ventured forth, her heart steadfast.

Verse 2:

Through enchanted woods, with creatures fair,
Moonlight guiding, magic in the air,
Found a flower, glowing bright,
Moonblossom's petals, cure in sight.

Verse 3:

Back to the village, Elara ran,
Her father's smile, health began,
Villagers cheered, tales to tell,
Bravery and love cast a healing spell.

Folk music: gentle acoustic guitar, fiddle, and tambourine, female vocal, irish lilt, fast pace



Current state of AI/ML

- AI capability is still far inferior to humans
 - Usually requires clearly defined tasks
 - Cannot perform complex tasks
 - Requires many examples
 - Difficult to adapt to new tasks

Research directions in machine learning

- Using ML to perform complex tasks
 - Perform software tasks, such as image editing, drafting slides based on an outline, or summarizing and reporting information from diverse sources
 - Learn from watching how people perform tasks and through natural training interfaces
- Integrating ML into physical devices
 - IoT, sensors
 - Robots
- Multimodal and broadly knowledgeable AI
 - Models increasingly defined in terms of input/output modalities rather than narrow tasks
 - Use vision, language, situation-specific, and broader knowledge to summarize and recommend
- Complementary learning systems
 - Integrate predictive/generative models and retrieval of individual memories and associations
 - Integrate with planning, task decomposition, and prioritization
- World models
 - Learn to predict how observations will change given action

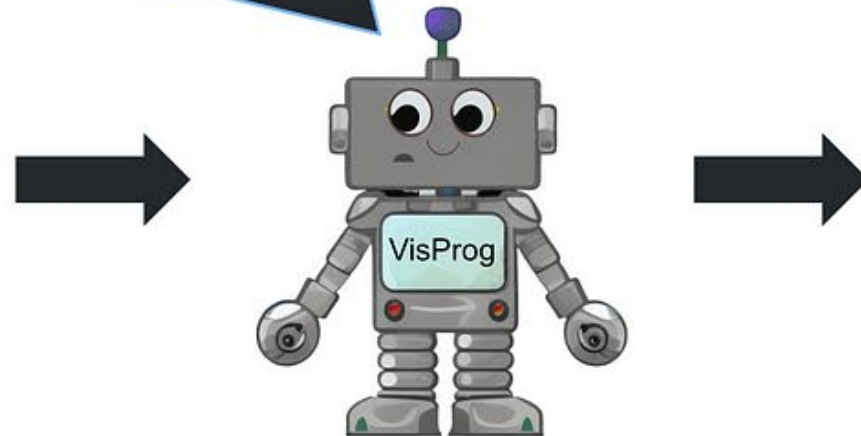
One example: Visual Programming

- Use LLM to generate code that calls pretrained vision models and functions

Instruction: Replace the ground with white snow and the bear with a white polar bear



```
OBJ0=Seg(image=IMAGE)
OBJ1=Select(image=IMAGE,object=OBJ0,query='ground')
IMAGE0=Replace(image=IMAGE,object=OBJ1,prompt='white snow')
OBJ2=Seg(image=IMAGE0)
OBJ3=Select(image=IMAGE0,object=OBJ2,query='bear')
IMAGE1=Replace(image=IMAGE0,object=OBJ3,prompt='white polar bear')
RESULT=IMAGE1
```

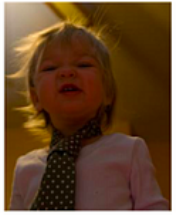


Prediction:



Compositional Visual Question Answering

IMAGE:



Question: Are there both ties and glasses in the picture?

Program:

```

BOX0=Loc(image=IMAGE, object='ties')
ANSWER0=Count(box=BOX0)
BOX1=Loc(image=IMAGE, object='glasses')
ANSWER1=Count(box=BOX1)
ANSWER2=Eval("{'yes' if {ANSWER0} > 0 and {ANSWER1} > 0 else 'no'")
RESULT=ANSWER2
Prediction: no

```

Natural Language Visual Reasoning

LEFT:



RIGHT:



Statement: The left and right image contains a total of six people and two boats.

Program:

```

ANSWER0=Vqa(image=LEFT, question='How many people are in the image?')
ANSWER1=Vqa(image=RIGHT, question='How many people are in the image?')
ANSWER2=Vqa(image=LEFT, question='How many boats are in the image?')
ANSWER3=Vqa(image=RIGHT, question='How many boats are in the image?')
ANSWER4=Eval('{ANSWER0} + {ANSWER1} == 6 and {ANSWER2} + {ANSWER3} == 2')
RESULT=ANSWER4
Prediction: False

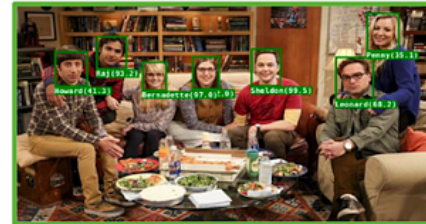
```

Factual Knowledge Object Tagging

IMAGE:



Prediction: IMAGE0



Instruction: Tag the 7 main characters on the TV show Big Bang Theory

Program:

```

OBJ0=FaceDet(image=IMAGE)
LIST0=List(query='main characters on the TV show Big Bang Theory', max=7)
OBJ1=Classify(image=IMAGE, object=OBJ0, categories=LIST0)
IMAGE0=Tag(image=IMAGE, object=OBJ1)
RESULT=IMAGE0

```

Natural Language Image Editing

IMAGE:



Prediction: IMAGE1



Instruction: Hide Daniel Craig with 8) and Sean Connery with ;)

Program:

```

OBJ0=FaceDet(image=IMAGE)
OBJ1=Select(image=IMAGE, object=OBJ0, query='Daniel Craig', category=None)
IMAGE0=Emoji(image=IMAGE, object=OBJ1, emoji='smiling_face_with_sunglasses')
OBJ2=Select(image=IMAGE, object=OBJ0, query='Sean Connery', category=None)
IMAGE1=Emoji(image=IMAGE0, object=OBJ2, emoji='winking_face')
RESULT=IMAGE1

```

IMAGE:



Prediction: IMAGE0



Instruction: Replace desert with lush green grass

Program:

```

OBJ0=Seg(image=IMAGE)
OBJ1=Select(image=IMAGE, object=OBJ0, query='desert', category=None)
IMAGE0=Replace(image=IMAGE, object=OBJ1, prompt='lush green grass')
RESULT=IMAGE0

```

IMAGE:



Prediction: IMAGE0



Instruction: Create a color pop of Barack Obama (person)

Program:

```

OBJ0=Seg(image=IMAGE)
OBJ1=Select(image=IMAGE, object=OBJ0, query='Barack Obama', category='person')
IMAGE0=ColorPop(image=IMAGE, object=OBJ1)
RESULT=IMAGE0

```

High-impact applications and what is needed to make them really work

- Personalized tutors, e.g. [Khanmigo](#)
 - Episodic memory, to have a history with the student
 - Personalization, to learn preferences by cues and previous responses
 - Improved ability to illustrate concepts
- Tool usage through language --- anything that is currently done through keyboard/mouse (shopping, finding information, creating illustrations, using spreadsheets) can be done with LLMs
- Robots (embodied agents)
 - Continual learning: gaining and using information about the environment, people, and concepts in real time
 - General purpose skill learning, maybe with world models as a foundation
 - Lighter, cheaper, easily repaired hardware

Finishing up the semester

- Dec 15: Final project due at 11:59pm
- Dec 5-10: Final exam
 - Covers all material for semester
 - See CampusWire pinned post for details
- Last office hour is Dec 13
 - We will still monitor and respond to CampusWire until Dec 15

ICES Feedback

- We put much time and energy into this course
- Please take some time to provide ratings and feedback
 - go.illinois.edu/ices-online
- Good things to know about ICES in general
 - Considered mainly as a reflection of the instructor and considered for promotions and course improvements
 - Reflect aspects of the course that are controllable by instructor, e.g. content, structure, and behavior
 - Unconscious bias that negatively affects women, people of color, and other nationalities is well documented; try to consciously counter-act it

Thank you to the course staff for all of their hard work to create materials, grade, and answer questions.

Thank you for your hard work and engagement!