

# Machine Learning: CS 472/572

## **Introduction**

Instructor: Thien Huu Nguyen

Based on slides by Daniel Lowd, Vibhav Gogate, Pedro Domingos, and others.



# Logistics

- **Instructor:** Thien Huu Nguyen
  - Email: [thien@cs.uoregon.edu](mailto:thien@cs.uoregon.edu)
  - Office: Deschutes 330
  - Office hours: Tuesdays (15:00-17:00)
- **Teaching Assistant:** Steven Walton
  - Email: [swalton@uoregon.edu](mailto:swalton@uoregon.edu)
  - Office: Deschutes 245
  - Office hours: Thursday (13:30-14:30), Friday: (13:30-14:30)
- **Web:** <https://classes.cs.uoregon.edu/24S/cs472/>
- **Discussion Board:** Piazza (link on Canvas)  
Please use this for discussion!

# Evaluation

- 4 assignments (40%)
  - Some programming, some exercises
- One Midterm (30%)
  - 2/3rds of the way through
- One Project (30%)
  - Apply machine learning to a real problem of your choice  
(*Recommended: Participate in a contest on Kaggle.com.*)
  - Groups allowed (for undergrad students)
  - Code, written report

# Kaggle

**Active Competitions**








**All Competitions**

17 found, 17 active

All competitions  
 Enterable

**Status**  
 Active  
 Completed

**Sponsor**  
 InClass (student competition)

Competition Name	Reward	Teams	Deadline
 <b>National Data Science Bowl</b> Predict ocean health, one plankton at a time	\$175,000	312	2 months
 <b>Driver Telematics Analysis</b> Use telematic data to identify a driver signature	\$30,000	453	2 months
 <b>Helping Santa's Helpers</b> Jingle bells, Santa tells ...	\$20,000	439	2.1 days
 <b>Click-Through Rate Prediction</b> Predict whether a mobile ad will be clicked	\$15,000	1203	35 days
 <b>BCI Challenge @ NER 2015</b> A spell on you if you cannot detect errors!	\$1,000	161	50 days
 <b>Sentiment Analysis on Movie Reviews</b> Classify the sentiment of sentences from the Rotten Tomatoes dataset	Knowledge	675	54 days
 <b>Finding Elo</b> Predict a chess player's FIDE Elo rating from one game	Knowledge	100	2 months

# Source Materials

Primary source of readings:

- H. Daume III, ***A Course in Machine Learning v0.9***. <http://ciml.info>  
(Good intro, focuses on machine learning concepts before math.  
Free online. Not finished.)

Excellent supplements:

- K. Murphy, ***Machine Learning: A Probabilistic Perspective***, MIT Press, 2012. (Great reference and in-depth coverage.)
- T. Mitchell, ***Machine Learning***, McGraw-Hill, 1997. (Great intro, but old and expensive.)
- Ian Goodfellow and Yoshua Bengio and Aaron Courville, ***Deep Learning***, MIT Press, 2016. (Good source for deep learning, free online: <https://www.deeplearningbook.org/>)
- C. Bishop, ***Pattern Recognition and Machine Learning***, Springer, 2006
- R. Duda, P. Hart & D. Stork, ***Pattern Classification*** (2<sup>nd</sup> ed.), Wiley, 2000
- D. Barber, ***Bayesian Reasoning and Machine Learning***, Cambridge University Press, 2012. (Free online!)
- T. Hastie, R. Tibshirani, J. Friedman, ***The Elements of Statistical Learning***, Springer, 2009. (Free online!)

# Why Study Machine Learning: A Few Quotes

“A breakthrough in machine learning would be worth ten Microsofts.”

-Bill Gates, Microsoft

“Machine learning is the next Internet.”

-Tony Tether, Former Director, DARPA

“Machine learning is the hot new thing.”

-John Hennessy, President, Stanford

“Web rankings today are mostly a matter of machine learning.”

-Prabhakar Raghavan, Dir. Research, Yahoo

“Machine learning is going to result in a real revolution.”

-Greg Papadopoulos, CTO, Sun



# So What Is Machine Learning?

- Learning is acquiring and improving performance through experience.
- H. Simon: Any process by which a system improves its performance
- M. Minsky: Learning is making useful changes in our minds
- R. Michalsky: Learning is constructing or modifying representations of what is being experienced
- L. Viliant: Learning is the process of knowledge acquisition in the absence of explicit programming



# So What Is Machine Learning?

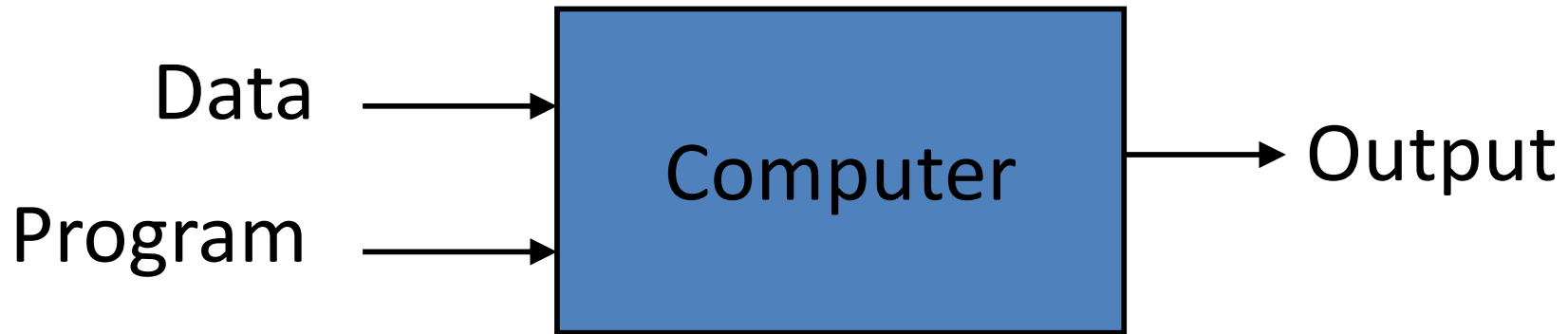
- What does it mean for a computer to learn? Why would we want them to learn? How do we get them to learn?
- We want computers to learn when it is too difficult or too expensive to program them directly to perform a task → Automating automation
- Get the computer to program itself by showing examples of inputs and outputs.
- In reality: we will write a “parameterized” program, and let the learning algorithm find the set of parameters that best approximates the desired function or behavior → Let the data do the work!



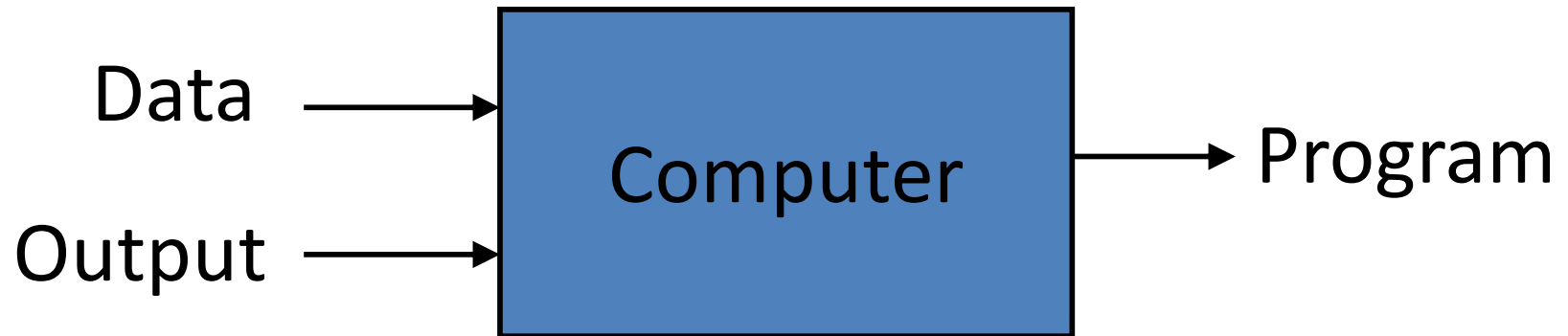


## Traditional Programming

- Automating automation
- Getting computers to program themselves



## Machine Learning



# Another view on machine learning

- Machine learning = automated science (sort of)
- Goal is to go from raw data to useful knowledge
- An ML algorithm finds a theory to fit the data and background knowledge as well as possible.
- A theory is good if it has good predictive accuracy.

# Related Fields

- Fields that use machine learning:
  - Artificial intelligence
  - Computer vision: object detection, tracking, segmentation, ...
  - Natural language processing: machine translation, question answering, text generation, ...
  - Computational biology: sequence alignment, information extraction, ...
  - Robotics: state estimation, self-driving, ...
  - Medicine: medical diagnosis, treatment prediction, drug design, ...
  - Financial Analysis: portfolio allocation, option pricing (many people on Wall Street use machine learning)



# Related Fields

- Fields with similar goals to machine learning:
  - Statistical Estimation
  - Data mining
  - Data science
  - Psychology (developmental, cognitive)
- Fields used by machine learning:
  - Information theory
  - Numerical optimization
  - Computational complexity

## Definition: Machine Learning!

- T. Mitchell: Improving performance via experience
  - A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at tasks in  $T$  as measured by  $P$ , improves with experience.

# Example 1: A Chess learning problem

- Task T: playing chess
- Performance measure P: percent of games won against opponents
- Training Experience E: playing practice games against itself

# Example 2: Autonomous Vehicle Problem

- Task T: driving on a public highway/roads using vision sensors
- Performance Measure P: percentage of time the vehicle is involved in an accident
- Training Experience E: a sequence of images and steering commands recorded while observing a human driver

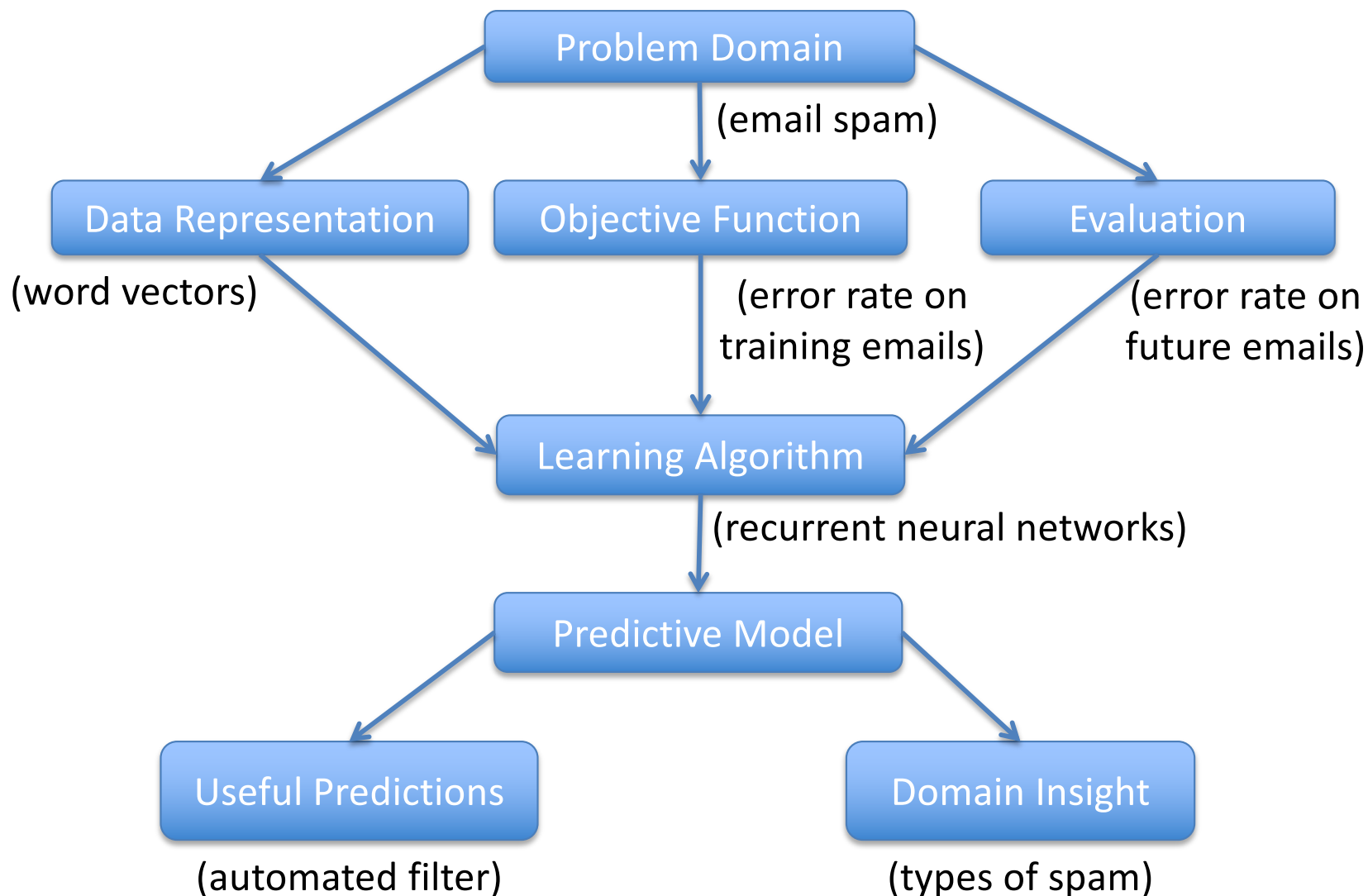


# Example 3: Machine Translation

- Task T: translating sentences from one language to another language
- Performance Measure P: the similarities between the machine-translated sentences and the human-translated sentences (BLUE scores)
- Training Experience E: pairs of translated sentences between languages provided by human

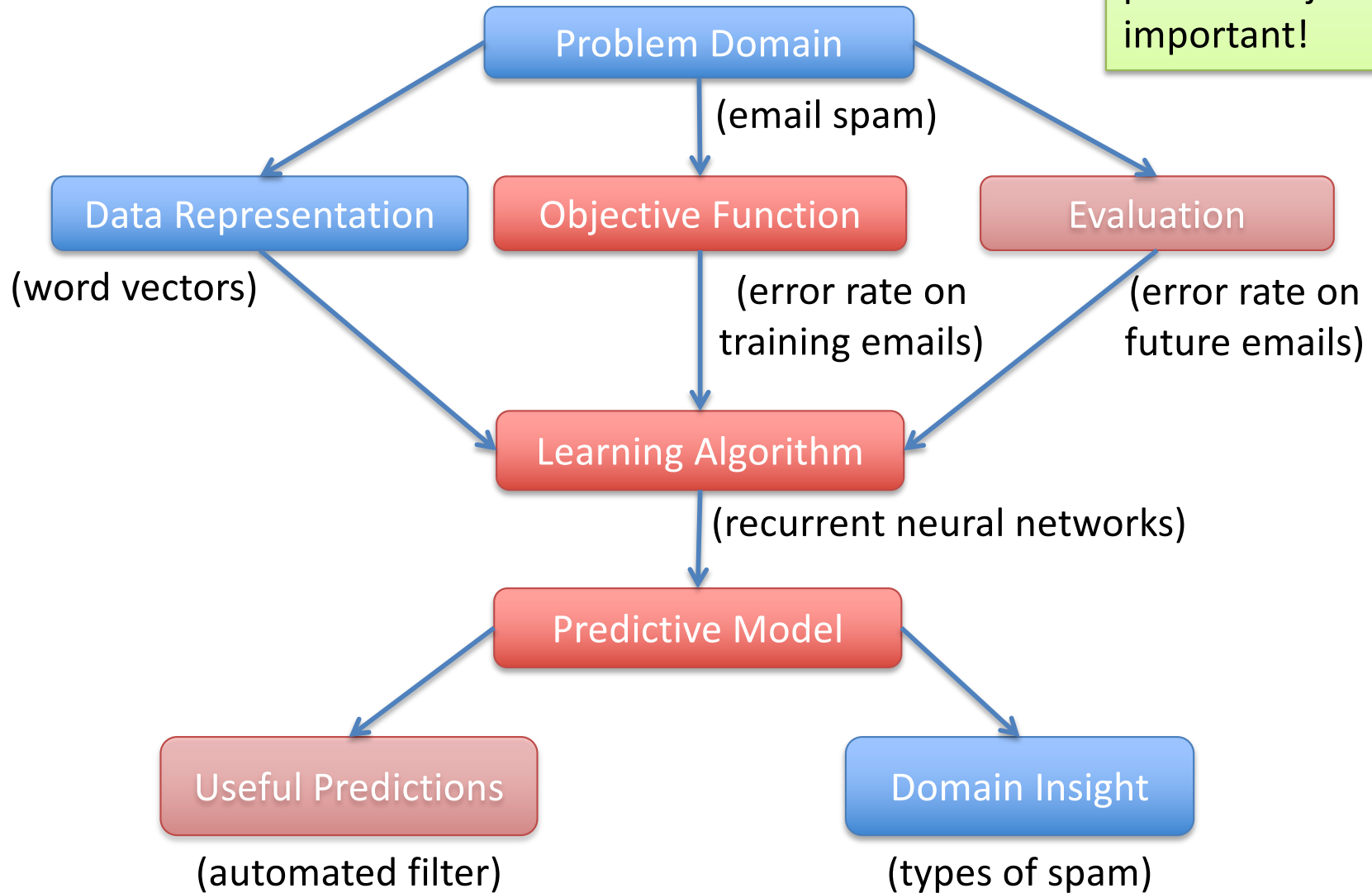


# The Big Picture



This class mostly focuses on the red pieces, but the blue pieces are just as important!

# The Big Picture



# ML in a Nutshell

- Tens of thousands of machine learning algorithms
- Hundreds new every year
- Some algorithms are specific to application domains/data
- Every machine learning algorithm has three components:
  - **Model Representation**
  - **Evaluation**
  - **Optimization**

# Model Representation

- Decision trees
- Instance-based models
- Linear function (hyperplane)
- Support vector machines
- Kernel methods
- Neural networks/Deep learning (Transformers)
- (Model ensembles)
- (Sets of rules / Logic programs)
- (Graphical models (Bayes/Markov nets))
- Etc.



# Evaluation

- Accuracy
- Precision and recall
- Squared error
- Likelihood
- Posterior probability
- Cost / Utility
- Margin
- Entropy
- KL divergence
- Etc.

# Optimization

- Combinatorial optimization
  - E.g.: Greedy search
- Convex optimization
  - E.g.: Gradient descent
- Constrained optimization
  - E.g.: Linear programming

# Types of Learning

- **Based on the information available**
  - **Supervised (inductive) learning**
    - Training data includes desired outputs
  - **Unsupervised learning**
    - Training data does not include desired outputs
    - Find hidden structure in data
  - **Semi-supervised learning**
    - Training data includes a few desired outputs
  - **Reinforcement learning**
    - The learner interacts with the world via “actions” and tries to find an optimal policy of behavior with respect to “rewards” it receives from the environment
- **Based on the role of the learner**
  - **Passive Learning**
    - Training solely relies on data
  - **Active Learning**
    - Training additionally involves interactions with human

# Types of Supervised Learning Problems

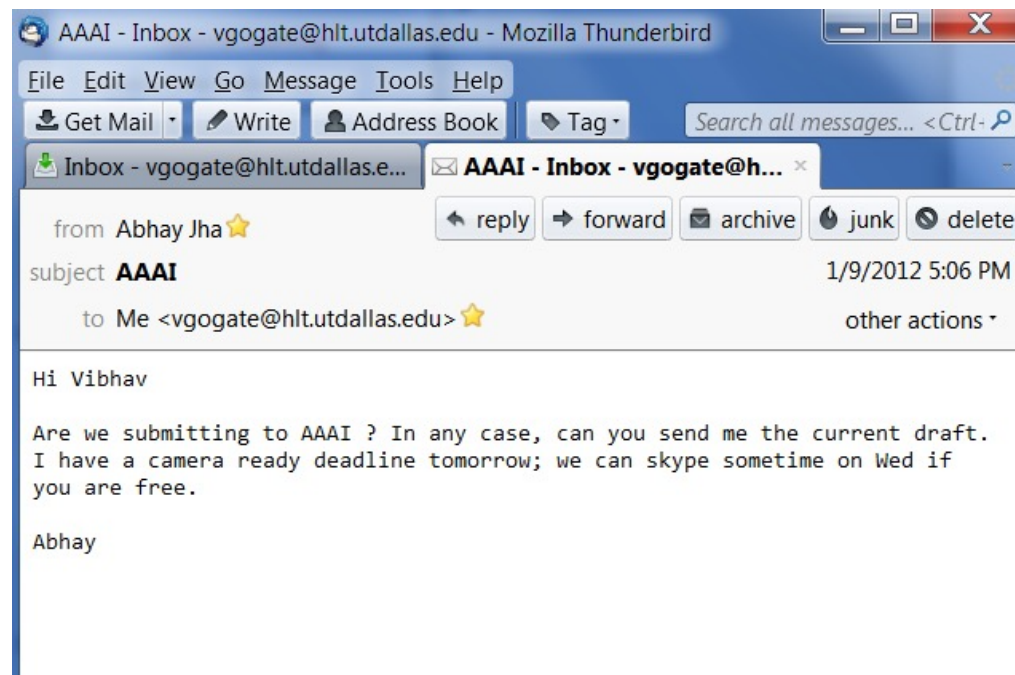
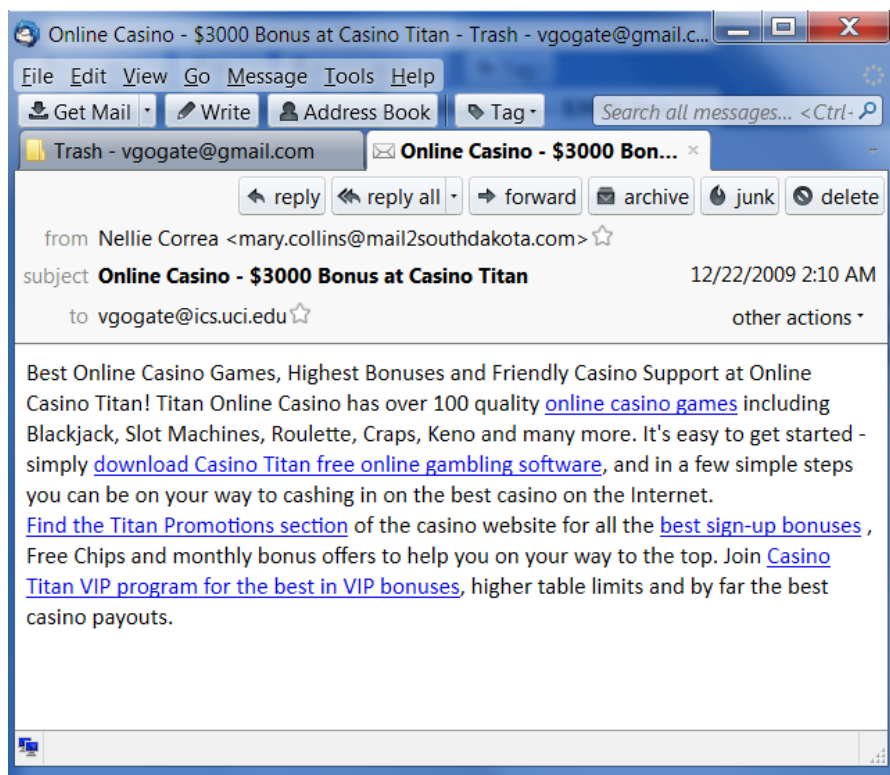
- **Classification:** predict a discrete/categorical value from a predefined set of values
- **Regression:** predict a continuous/real value
- **Structured prediction:** predict a complex output, such as a sequence or tree



# Machine Learning: Applications

Examples of what you will study in  
class in action!

# Classification Example: Spam Filtering



Classify as “Spam” or “Not Spam”

# Classification Example: Weather Prediction



# Regression example: Predicting Gold/Stock prices



**Good ML can make you rich (but there is still some risk involved).**

**Given historical data on gold prices, predict tomorrow's price!**

# Similarity Determination

The image displays two sets of search results for image similarity determination. Each result card includes a thumbnail image and two search options: '1. Find similar by Theme' and '2. Find similar by Color / Texture'. The search mode is indicated at the top of each card.

**Left Set (Theme Search):**

- Card 1: Search mode: Theme. Image: City skyline at sunset. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 2: Search mode: Theme. Image: City skyline at sunset. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 3: Search mode: Theme. Image: City skyline at sunset. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 4: Search mode: Theme. Image: City skyline at sunset. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 5: Search mode: Theme. Image: Sunset over a field. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 6: Search mode: Theme. Image: Sunset over a field. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 7: Search mode: Theme. Image: Sunset over a field. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 8: Search mode: Theme. Image: Sunset over a field. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 9: Search mode: Theme. Image: City skyline at sunset. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.

**Right Set (Color/Texture Search):**

- Card 1: Search mode: Color / Texture. Image: City skyline at sunset. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 2: Search mode: Color / Texture. Image: Sunflower. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 3: Search mode: Color / Texture. Image: Placeholder (THIS PHOTO IS CURRENTLY UNAVAILABLE). Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 4: Search mode: Color / Texture. Image: Red car with 'Vriendschap' text. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 5: Search mode: Color / Texture. Image: Boy lighting a match. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 6: Search mode: Color / Texture. Image: Two lit candles. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 7: Search mode: Color / Texture. Image: Sunset over a field. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 8: Search mode: Color / Texture. Image: Lava lamp. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.
- Card 9: Search mode: Color / Texture. Image: Crystal. Options: 1. Find similar by Theme, 2. Find similar by Color / Texture.

# Collaborative Filtering

The screenshot displays the Netflix user interface for a user named Vibhav Gogate. The top navigation bar includes 'Watch Instantly', 'Just for Kids', 'Browse DVDs', 'Your Queue', and 'Suggestions For You'. A search bar is located on the right. Below the navigation, there are tabs for 'Suggestions (4,663)', 'Rate Shows & Movies', 'Taste Preferences', and 'What You've Rated (316)'. A 'Suggestions In:' dropdown menu is set to 'All Genres', and a 'RATINGS 316' badge is visible.

The main content area is divided into three sections:

- Suggestions to Watch Instantly:** This section features three movie cards:
  - Bob the Builder: Three Musketrucks:** Recommended because the user enjoyed 'Caillou: Caillou's World of Wonder', 'Super Why!: Jack and the Beanstalk', and 'Care Bears: To the Rescue: The Movie'. It has a 'Play' button and a 'Not Interested' button.
  - Thomas & Friends: Carnival Capers:** Recommended because the user enjoyed 'Caillou: Caillou's World of Wonder', 'Clifford's Really Big Movie', and 'Dragon Tales: Easy as 1-2-3'. It has a 'Play' button and a 'Not Interested' button.
  - Angelina Ballerina: The Silver Locket:** Recommended because the user enjoyed 'Clifford's Really Big Movie', 'Super Why!: Jack and the Beanstalk', and 'Care Bears: To the Rescue: The Movie'. It has a 'Play' button and a 'Not Interested' button.
- New Suggestions:** This section features three movie cards:
  - Ben 10: Alien Force:** Recommended because the user enjoyed 'Astro Boy'. It has a 'Choose Discs' button and a 'Not Interested' button.
  - Handy Manny: Manny's Motorcycle Adventure:** Recommended because the user enjoyed 'Astro Boy', 'Caillou: Caillou's World of Wonder', and 'Toy Story 2'. It has an 'Add' button and a 'Not Interested' button.
  - Peep's New Friends:** Recommended because the user enjoyed 'Caillou: Caillou's World of Wonder', 'Clifford's Really Big Movie', and 'Dragon Tales: Easy as 1-2-3'. It has an 'Add' button and a 'Not Interested' button.
- Action & Adventure:** This section features one movie card:
  - Gladiator: Extended Edition:** Recommended because the user enjoyed 'The Patriot', 'Braveheart', and 'A Beautiful Mind'. It has a 'Rate more Action & Adventure' button with the text 'So we can give you more'.

# Collaborative Filtering

Amazon.com: Recommended for You - Microsoft Internet Explorer

Address: <http://www.amazon.com/gp/yourstore/002-8908355-5636015?rGroup=all&>

amazon.com Sue's Store See All 32 Product Categories Your Account | Cart | Wish List | Help | \$

Search Amazon.com GO! Web Search GO!


**Recommended for Sue Yeon Syn** (If you're not Sue Yeon Syn, [click here.](#))

Narrow by Event: [Your Watch List](#) (Beta)


Narrow by Category: [Apparel & Accessories](#), [Baby](#), [Beauty](#), [Books](#), [Camera & Photo](#), [Computer & Video](#), [Games](#), [Computers](#), [DVD](#), [Electronics](#), [Health & Personal Care](#), [Jewelry & Watches](#), [Kitchen & Housewares](#), [Magazine Subscriptions](#), [Music](#), [Outdoor Living](#), [Software](#), [Sports & Outdoors](#), [Tools & Hardware](#), [Toys & Games](#), [Video](#), [Select Favorites](#)

Recommendations for you are based on [items you own](#) and more. [More results](#)


view: [All](#) | [New Releases](#) | [Coming Soon](#)

- 


**When Things Start to Think**  
by Gershenfeld Neil  
Average Customer Review: ★★★★★  
Publication Date: February 15, 2000  
Our Price: **\$11.20** [Used & new](#) from \$2.00 [Add to cart](#) [Add to Wish List](#)

Rate this item  ★★★★★  I own it  Not interested  
Recommended because you added [The Unfinished Revolution](#) to your Shopping Cart ([edit](#))
- 

**Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web**  
by Tim Berners-Lee  
Average Customer Review: ★★★★★  
Publication Date: November 1, 2000  
Our Price: **\$10.20** [Used & new](#) from \$2.71 [Add to cart](#) [Add to Wish List](#)

Rate this item  ★★★★★  I own it  Not interested  
Recommended because you added [The Unfinished Revolution](#) to your Shopping Cart ([edit](#))
- 

**Perl Cookbook, Second Edition**  
by Tom Christiansen, Nathan Torkington  
Average Customer Review: ★★★★★  
Publication Date: August 21, 2003  
Our Price: **\$32.97** [Used & new](#) from \$15.64 [Add to cart](#) [Add to Wish List](#)

Rate this item  ★★★★★  I own it  Not interested  
Recommended because you added [Programming Perl \(2nd Edition\)](#) to your Wish List and more ([edit](#))
- 

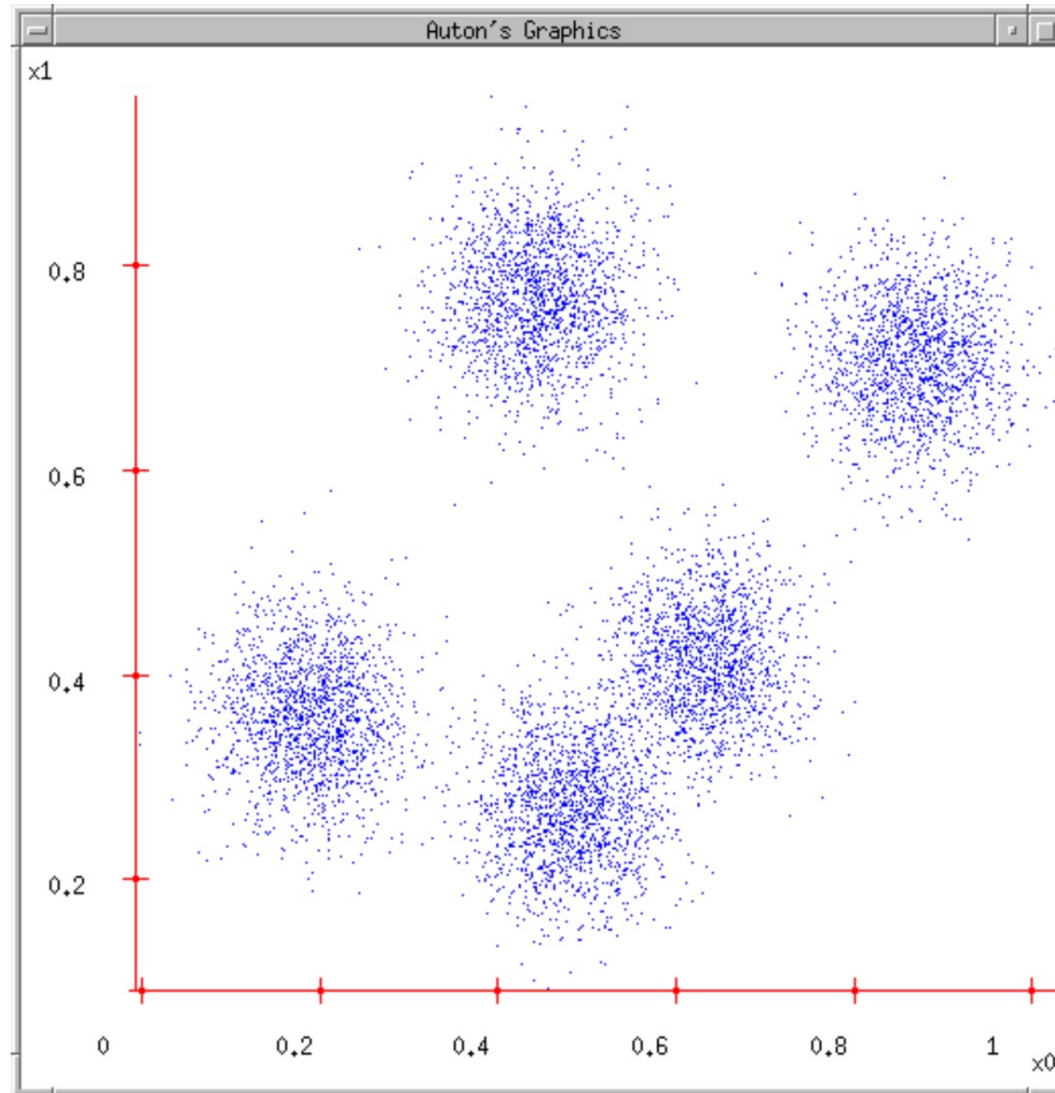
**Network Analysis, Architecture and Design, Second Edition (The Morgan Kaufmann Series in Networking)**  
by James D. McCabe  
Average Customer Review: ★★★★★  
Publication Date: April 1, 2003  
Our Price: **\$58.46** [Used & new](#) from \$46.77 [Add to cart](#) [Add to Wish List](#)

Rate this item  ★★★★★  I own it  Not interested

[Improve Your Recommendations](#)  
Update your Amazon history to improve your recommendations  
[Items you own](#)  
[Rated items](#)  
[Not Interested](#)

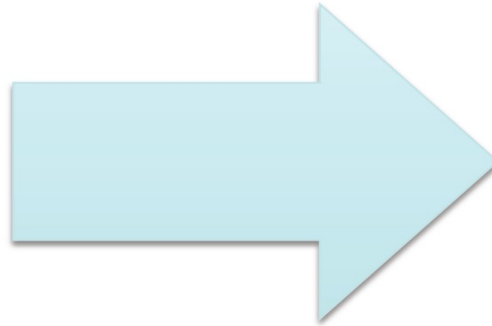
Internet

# Clustering: Discover Structure in data



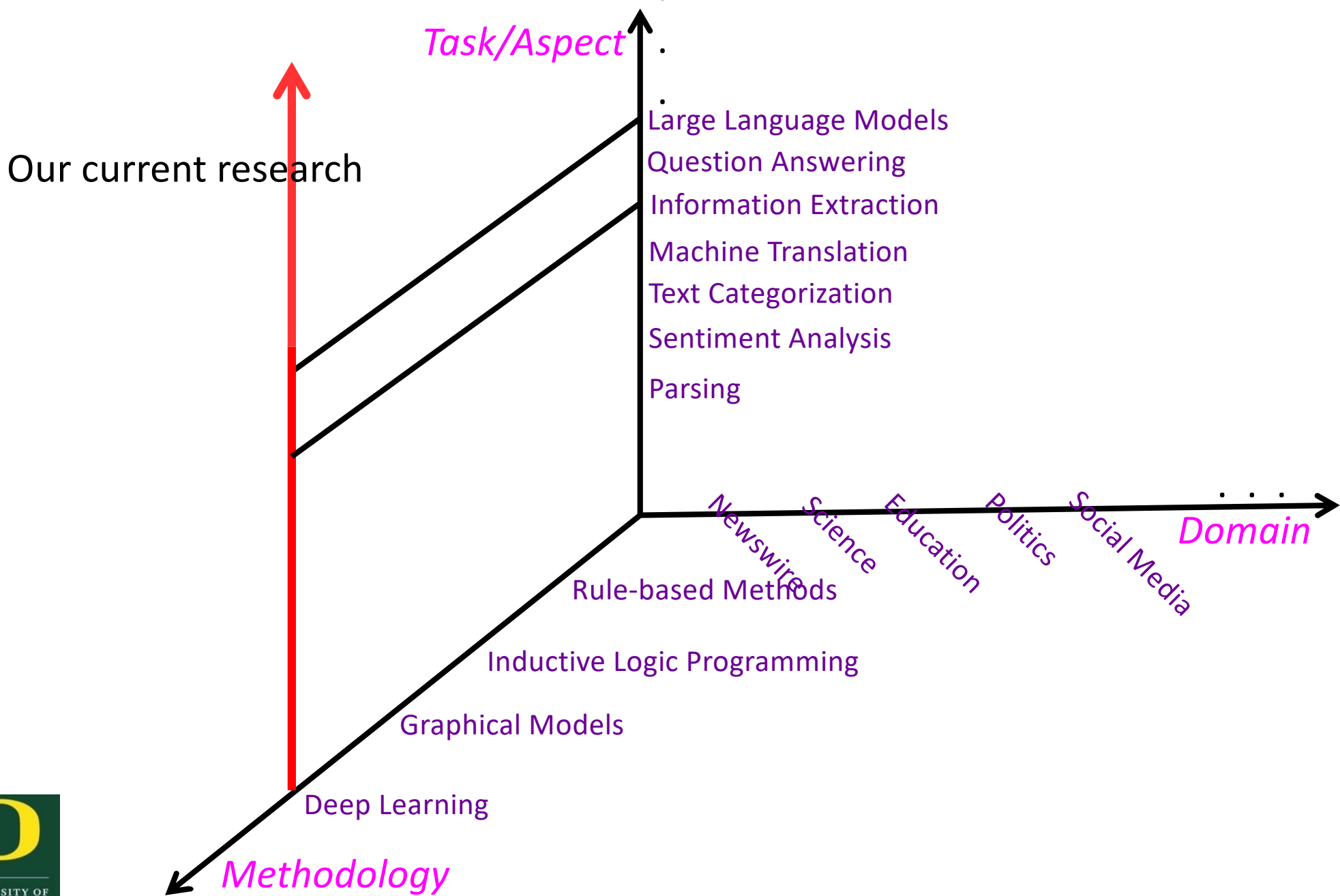


# Clustering images



[Goldberger et al.]

# Research in the NLP group at UO



# Information extraction



Giuliani, 58, proposed to Nathan, a former nurse, during a business trip to Paris \_ five months after he finalized his divorce from Donna Hanover in July after 20 years of marriage.



In interviews last year, Giuliani said Nathan gave him ``tremendous emotional support'' through his treatment for prostate cancer and as he led New York City during the Sept. 11, 2001, terror attacks.

Relation Knowledge Base

Name	leaderOf	...
Giuliani	New York City	
.....		

Corpora      Entity Recognition      Relation Extraction      Coreference Resolution      Entity Linking

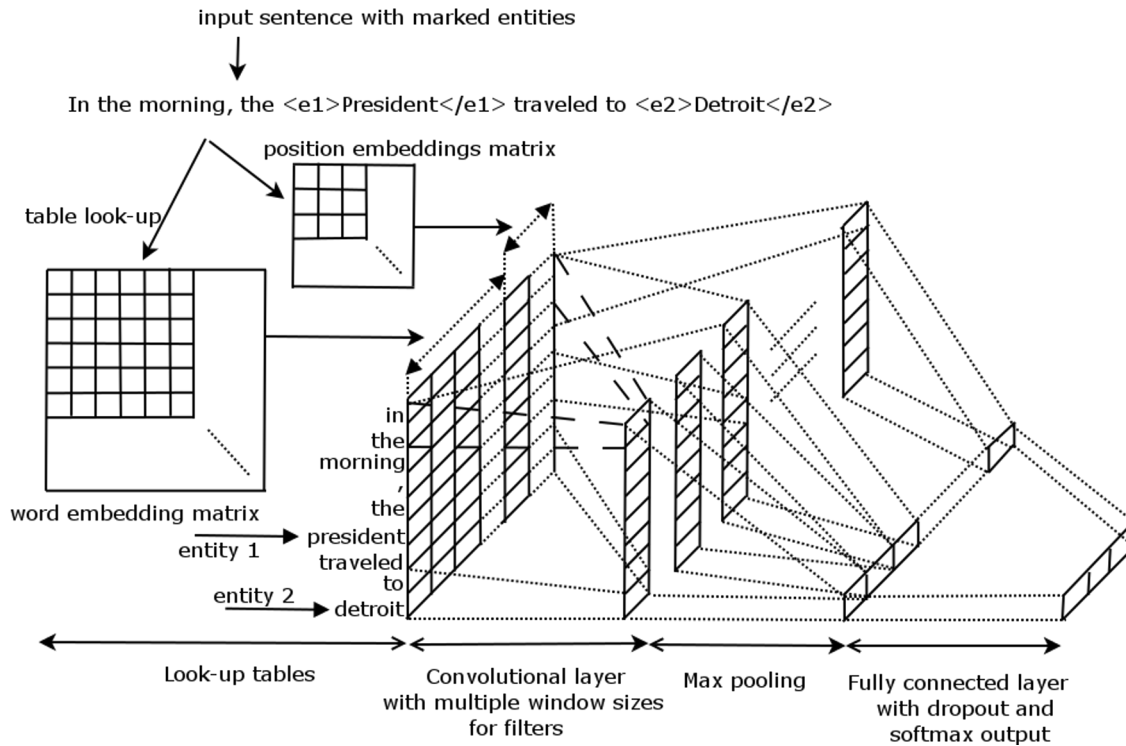
Event Knowledge Base

Trigger Prediction      Argument Prediction

Trigger	Type	Person1	Person 2	Time
divorce	Divorce	Giuliani	Donna Hanover	July
.....				

# Neural Networks for Relation Extraction

SemEval 2010 Dataset



A Convolutional Neural Network (CNN) for Relation Extraction

Nguyen and Grishman, 2015a

Classifier	Features	F
MaxEnt	POS, WordNet, morphological features, noun compound system, thesauri, Google n -grams	77.6
SVM	POS, WordNet, prefixes and other morphological features, dependency parse, Levin classes, PropBank, FrameNet, NomLex-Plus, Google n -grams, paraphrases, TextRunner	82.2
CNN (Zeng et al., 2014)	WordNet	82.7
CNN (Nguyen and Grishman, 2015a)	-	82.8

# Event Extraction

In **Baghdad**, a **soldier** *died* when his tank *fired* on the street.

Type: Died

Role

Place: **Baghdad**

Victim: **soldier**

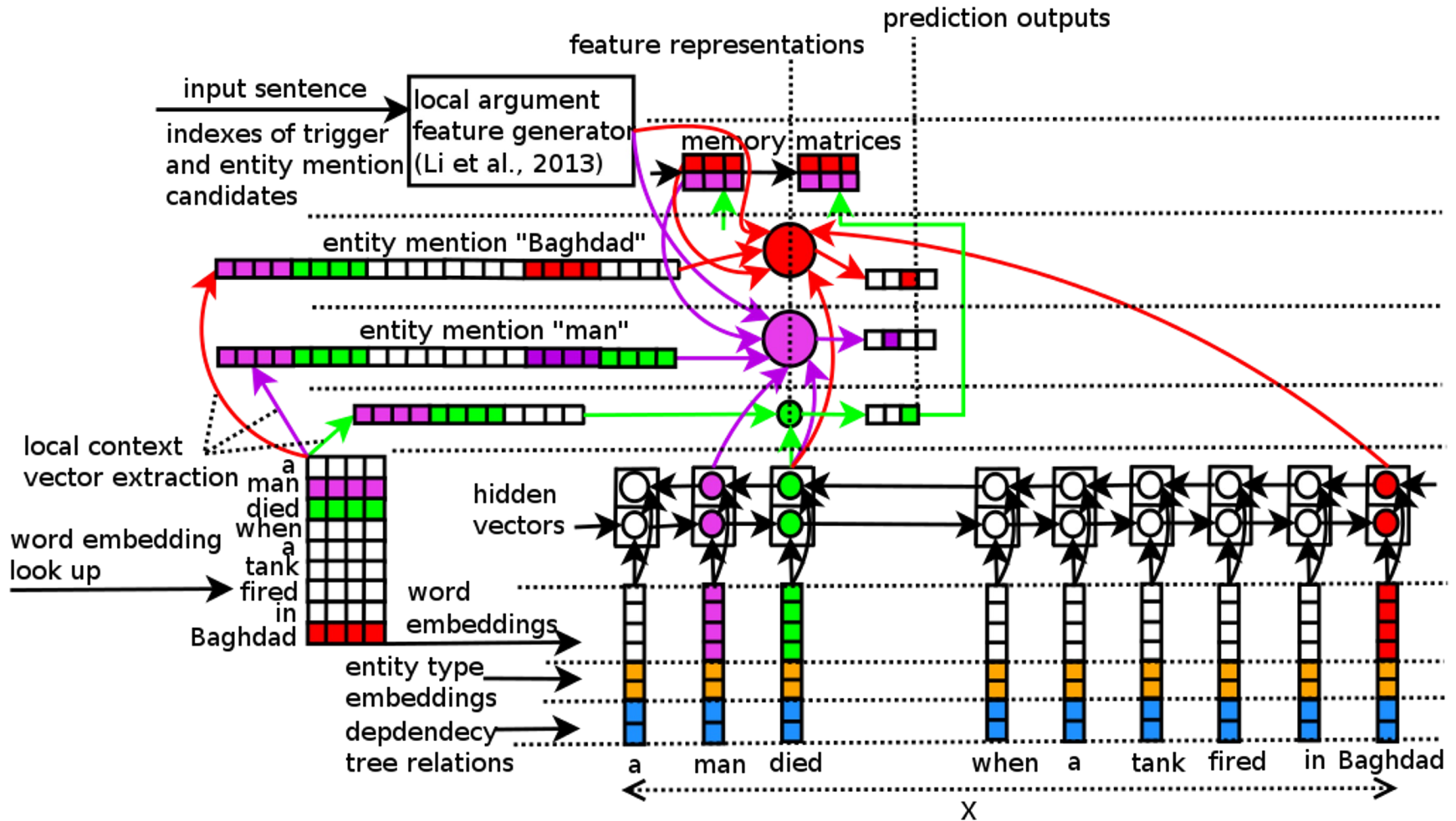
Type: Attack

Role

Place: **Baghdad**

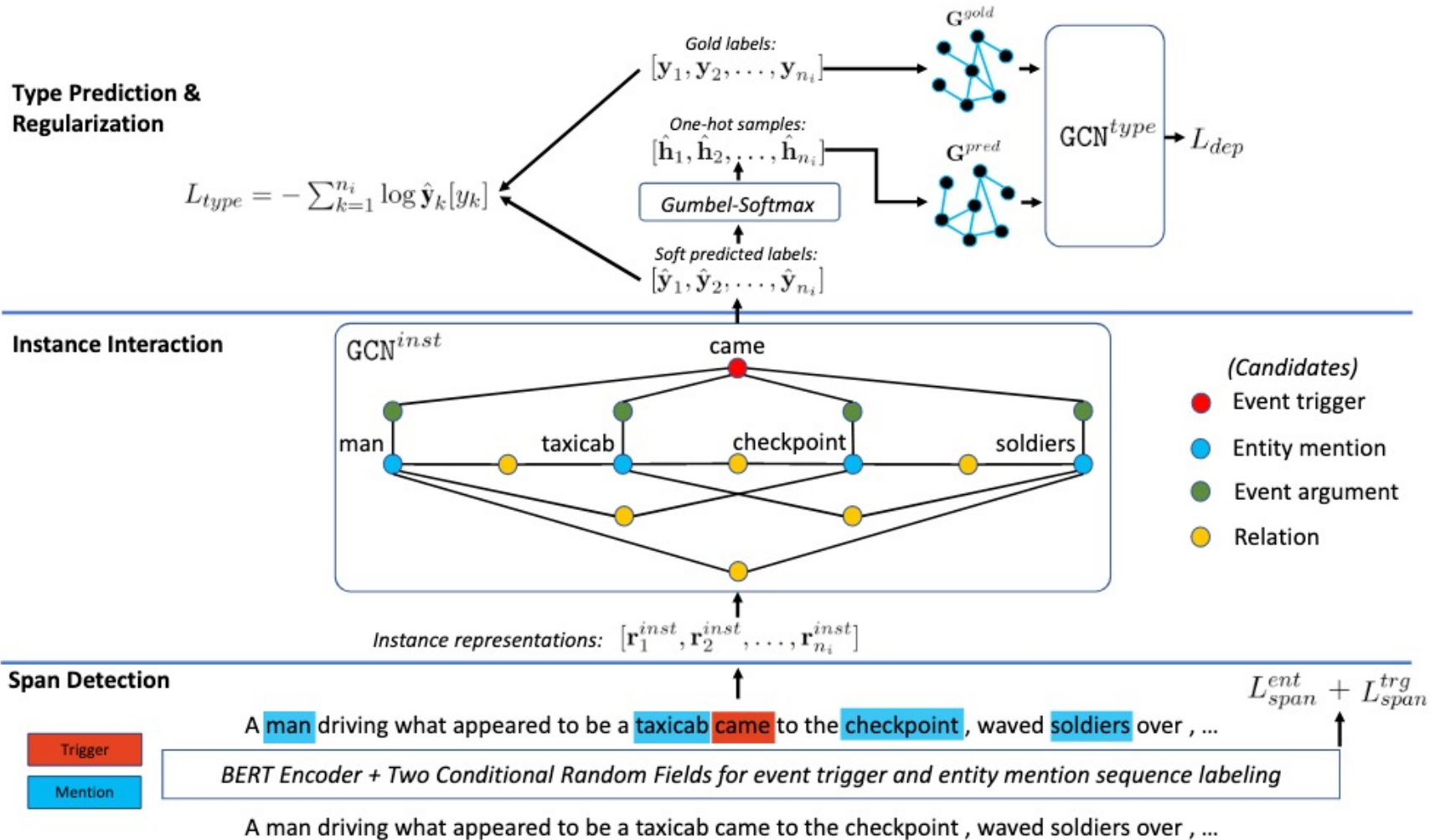
Target: **soldier**

# Joint Inference for Event Extraction



Nguyen et al., 2016 (NAACL 2016)

# Joint Inference for Event Extraction



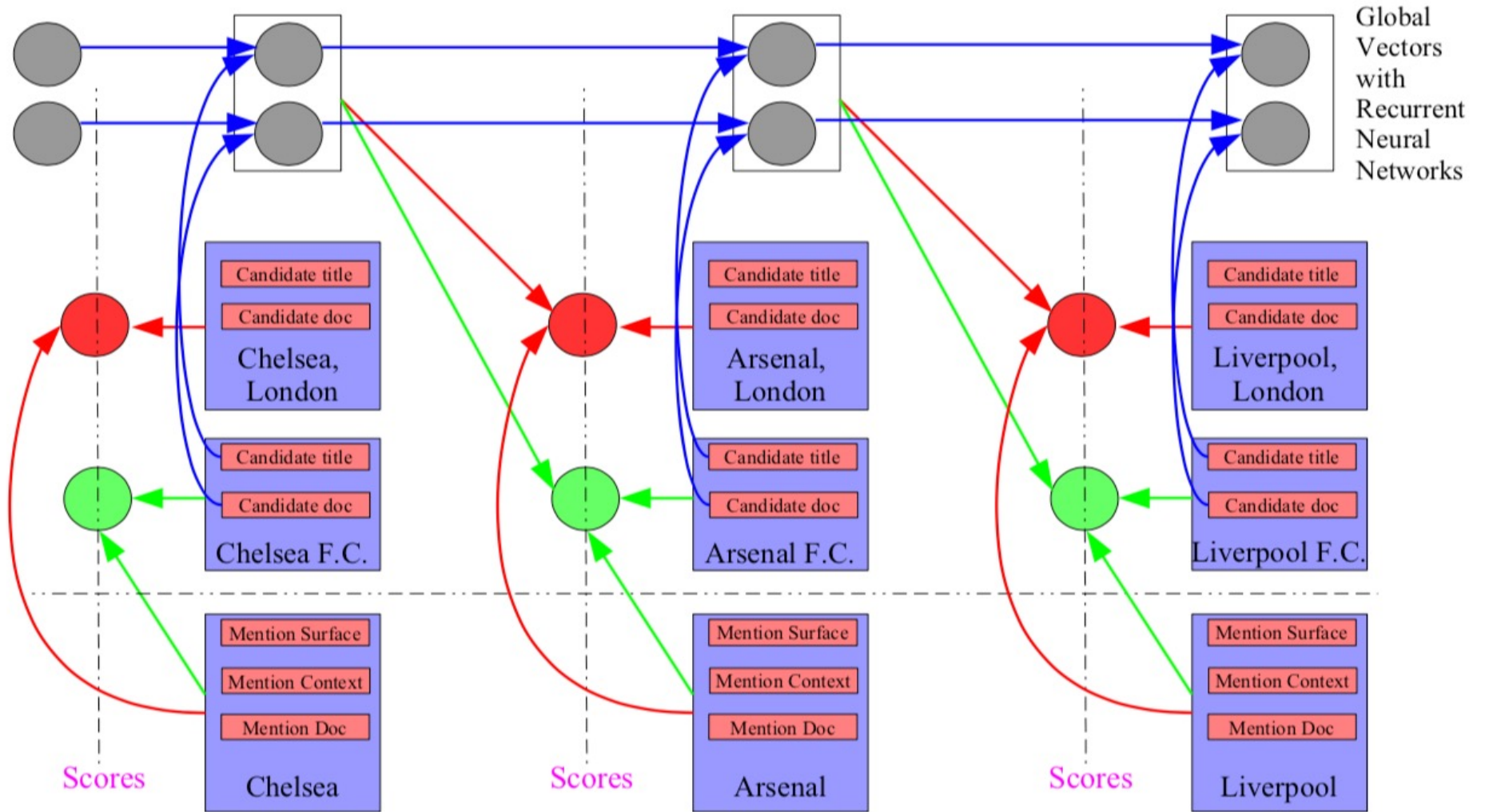
# Entity Linking



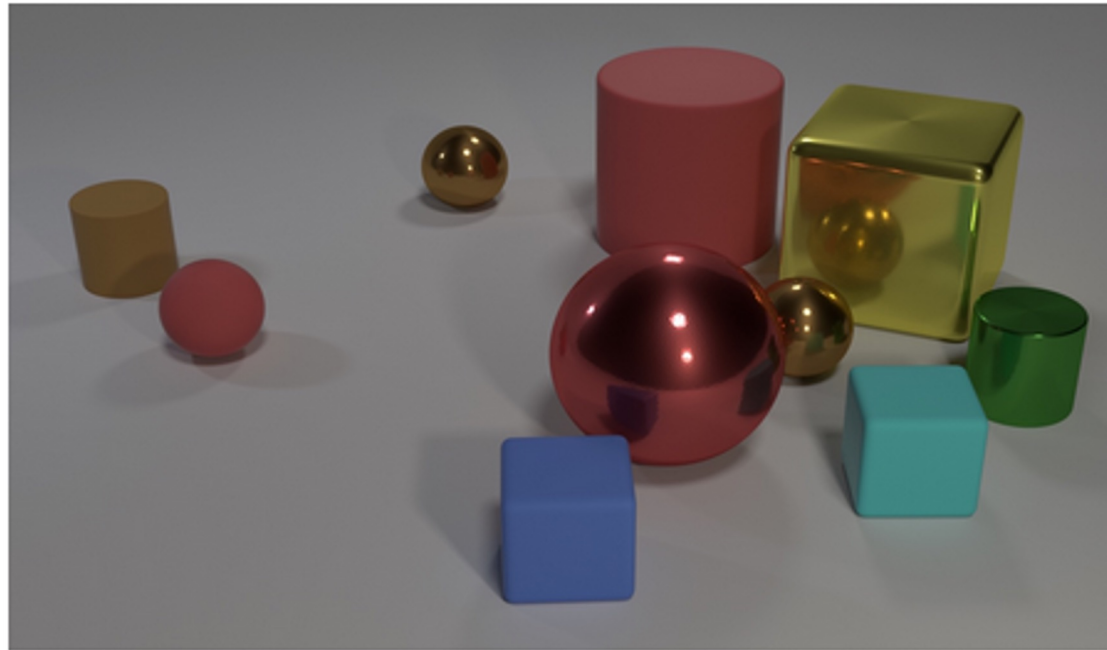
Chelsea have long-standing rivalries with North London clubs Arsenal and Tottenham Hotspur. A strong rivalry with Leeds United dates back to several heated and controversial matches in the 1960s and 1970s.



# Joint Inference for Entity Linking



# Visual Question Answering (VQA)



Q: Are there an **equal number** of **large things** and **metal spheres**?

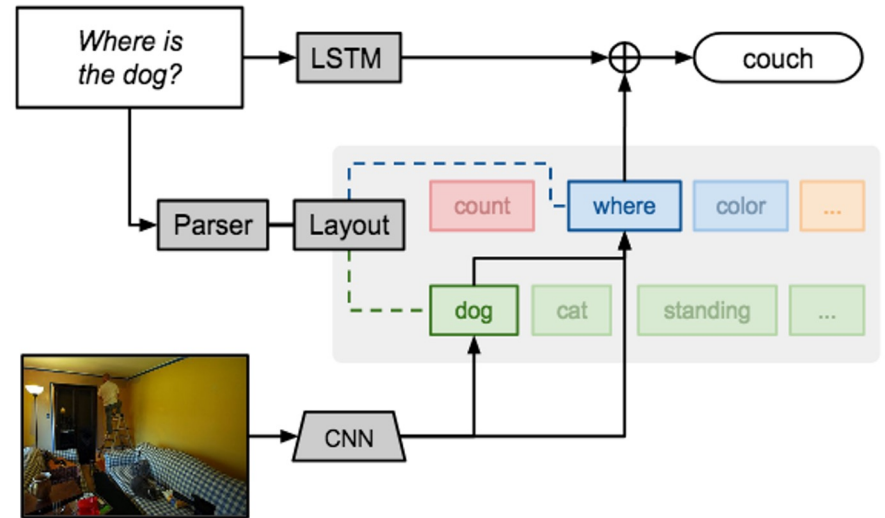
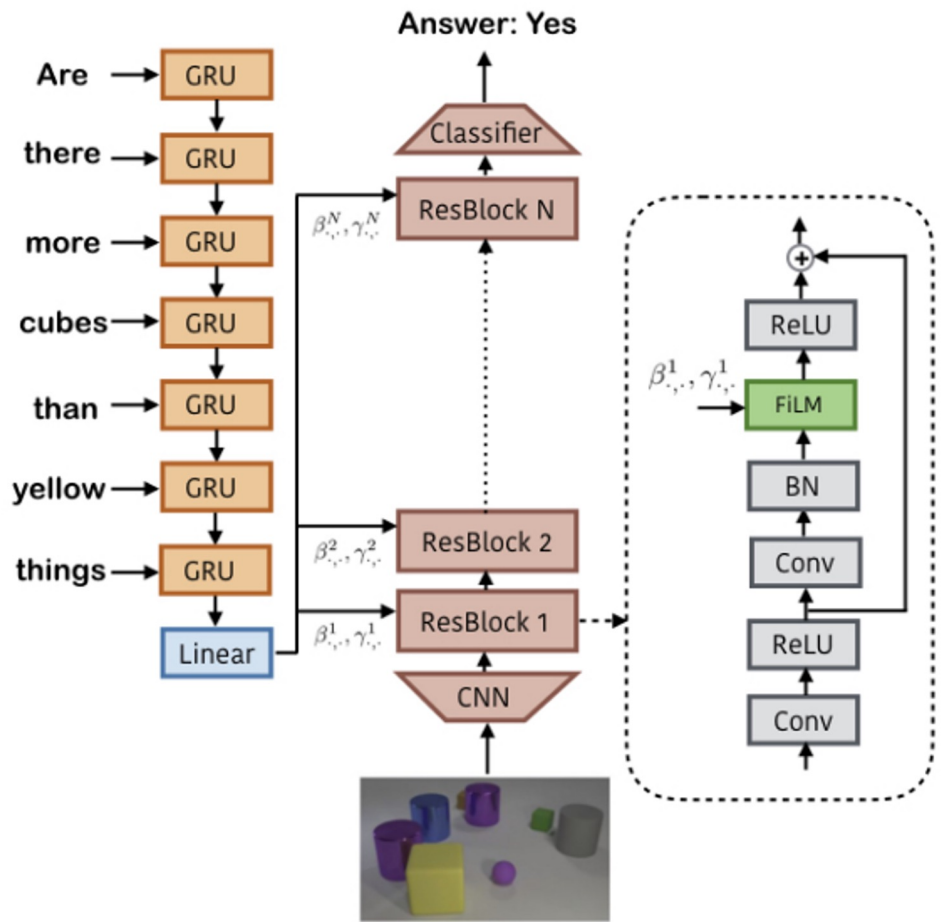
Q: **What size** is the **cylinder that is left of** the **brown metal** thing **that is left of** the **big sphere**?

Q: There is a **sphere** with the **same size as** the **metal cube**; is it **made of the same material as** the **small red sphere**?

Q: **How many** objects are **either small cylinders** or **red** things?

Johnson et al., 2017

# VQA Models



Neural Module Networks  
Jacob et al., 2017

FiLM: Ethan et al., 2018

# ChatGPT

TH

if  $A \sim \text{Normal}(1, 2)$  and  $B \sim \text{Normal}(2, 3)$ , and  $A$  and  $B$  are independent, what is the expected value of  $A + B$



The expected value of the sum of two random variables is equal to the sum of their individual expected values. Therefore, we can find the expected value of  $A + B$  by adding the expected values of  $A$  and  $B$ :

$$E(A + B) = E(A) + E(B)$$

Since  $A \sim \text{Normal}(1, 2)$ , we know that the expected value of  $A$  is 1, and since  $B \sim \text{Normal}(2, 3)$ , we know that the expected value of  $B$  is 2. Therefore:


$$E(A + B) = E(A) + E(B) = 1 + 2 = 3$$

So the expected value of  $A + B$  is 3.



# ChatGPT

**TH** if  $A \sim \text{Normal}(1, 2)$  and  $B \sim \text{Normal}(2, 3)$ , and  $A$  and  $B$  are independent, what is the expected value of  $A + B$

 The expected value of the sum of two random variables is equal to the sum of their individual expected values. Therefore, we can find the expected value of  $A + B$  by adding the expected values of  $A$  and  $B$ :


$$E(A + B) = E(A) + E(B)$$

Since  $A \sim \text{Normal}(1, 2)$ , we know that the expected value of  $A$  is 1, and since  $B \sim \text{Normal}(2, 3)$ , we know that the expected value of  $B$  is 2. Therefore:

$$E(A + B) = E(A) + E(B) = 1 + 2 = 3$$

So the expected value of  $A + B$  is 3.

**TH** What if  $A$  and  $B$  are not independent?

 If  $A$  and  $B$  are not independent, then we cannot simply add their expected values to get the expected value of their sum.

To find the expected value of  $A + B$  when  $A$  and  $B$  are not independent, we need to know the joint distribution of  $A$  and  $B$  or the covariance between  $A$  and  $B$ . If we have the joint distribution, we can use the following formula:

$$E(A + B) = E(A) + E(B) + \text{Cov}(A, B)$$

where  $\text{Cov}(A, B)$  is the covariance between  $A$  and  $B$ .

If we do not have the joint distribution of  $A$  and  $B$ , but we know their covariance, then we can use the following formula:


$$E(A + B) = E(A) + E(B) + \text{Cov}(A, B)$$

$$= E(A) + E(B) + \rho(A, B)\sigma(A)\sigma(B)$$

where  $\rho(A, B)$  is the correlation coefficient between  $A$  and  $B$ , and  $\sigma(A)$  and  $\sigma(B)$  are the standard deviations of  $A$  and  $B$ , respectively.

In summary, if  $A$  and  $B$  are not independent, we need to consider their joint distribution or their covariance to calculate the expected value of  $A + B$ .

**TH** **You** If  $A \sim \text{Normal}(1, 2)$  and  $B \sim \text{Normal}(2, 3)$ , and  $A$  and  $B$  are independent. What is the expected value of  $A + B$ ?

 **ChatGPT** If  $A$  and  $B$  are independent random variables, then the expected value of their sum is simply the sum of their individual expected values.


Given that  $A \sim \text{Normal}(1, 2)$  and  $B \sim \text{Normal}(2, 3)$ , we have:

- $E(A) = 1$  (the mean of  $A$ )
- $E(B) = 2$  (the mean of  $B$ )

So, the expected value of  $A + B$  is:

$$E(A + B) = E(A) + E(B) = 1 + 2 = 3.$$

**TH** **You** What if  $A$  and  $B$  are not independent?

 **ChatGPT** If  $A$  and  $B$  are not independent, we need more information about their joint distribution or the covariance between them to determine the expected value of their sum.

Without knowledge of their joint distribution or covariance, we cannot simply add their individual expected values to find the expected value of their sum. In cases where  $A$  and  $B$  are not independent, the expected value of their sum depends on the specific relationship between them, which requires further information beyond their marginal distributions.

# Text-to-Video

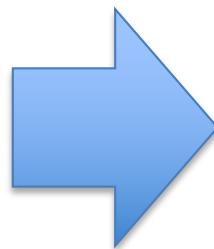
- Sora (OpenAI)

Images taken from generated videos

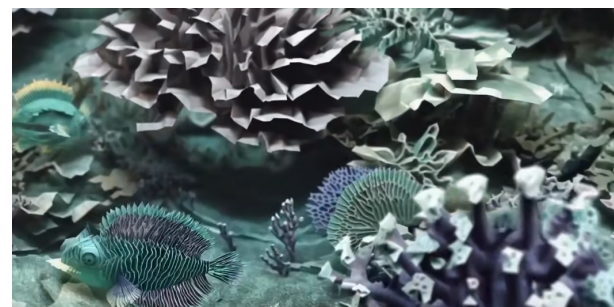
A beautiful homemade video showing the people of Lagos, Nigeria in the year 2056. Shot with a mobile phone camera.



A drone camera circles around a beautiful historic church built on a rocky outcropping along the Amalfi Coast, the view showcases historic and magnificent architectural details and tiered pathways and patios, waves are seen crashing against the rocks below as the view overlooks the horizon of the coastal waters and hilly landscapes of the Amalfi Coast Italy, several distant people are seen walking and enjoying vistas on patios of the dramatic ocean views, the warm glow of the afternoon sun creates a magical and romantic feeling to the scene, the view is stunning captured with beautiful photography.



A gorgeously rendered papercraft world of a coral reef, rife with colorful fish and sea creatures



# Machine learning has grown in leaps and bounds

- The main approach for
  - Speech Recognition
  - Robotics
  - Natural Language Processing
  - Computational Biology
  - Sensor networks
  - Computer Vision
  - Web
  - ...and many more each year...

# What We'll Cover

- **Supervised learning:** Decision tree induction, Instance-based learning, Neural networks/Deep Learning, Support vector machines, Linear regression, Learning theory, Transformers, etc.
- **General machine learning concepts and techniques:** Feature selection, cross-validation, algorithm evaluation ...
- **Unsupervised learning:** Clustering, Dimensionality reduction, Auto-Encoder, VAE, GAN, Diffusion (if time permitted)

## Not covering:

- Reinforcement learning
- Probabilistic graphical models (410/510)
- Structured prediction (e.g., machine translation, image segmentation, multi-label classification)



# What We'll Cover: Comparison

## Computer Science

- **Core concepts:**  
Variables, conditionals, loops, functions, etc.
- **Key algorithms:**  
Mergesort, linked lists, binary search trees, breadth-first search, etc.
- **Process:**  
Debugging, software engineering, etc.

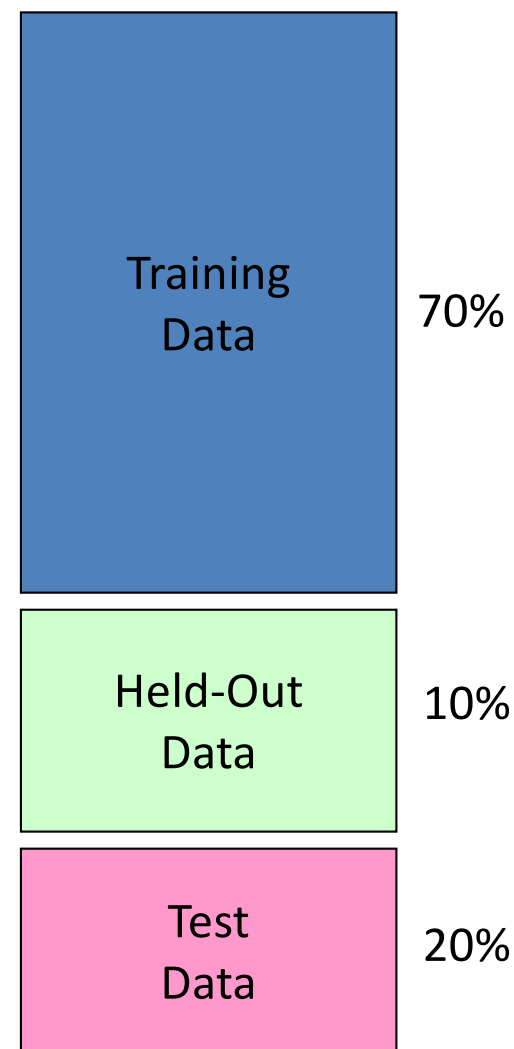
## Machine Learning

- **Core concepts:**  
Classification, overfitting, underfitting, training set, etc.
- **Key algorithms:**  
Decision trees, nearest neighbor, linear models, etc.
- **Process:**  
Designing and debugging ML systems



# Important Concepts

- Data: labeled instances, e.g. emails marked spam/ham
  - Training set
  - Held out set
  - Test set
- Features: attribute-value pairs which characterize each  $x$
- Experimentation cycle
  - Learn parameters (e.g. model probabilities) on training set
  - (Tune hyperparameters on held-out set)
  - Compute accuracy of test set
  - Very important: never “peek” at the test set!
- Evaluation
  - Accuracy: fraction of instances predicted correctly
- Overfitting and generalization
  - Want a classifier which does well on *test* data
  - Overfitting: fitting the training data very closely, but not generalizing well
  - We’ll investigate overfitting and generalization more formally in a few lectures



**Goal:** Use course attributes to predict the student's course rating (+2 = loved it, -2 = hated it).

How should we model this?

Rating	Easy?	AI?	Sys?	Thy?	Morning?
+2	y	y	n	y	n
+2	y	y	n	y	n
+2	n	y	n	n	n
+2	n	n	n	y	n
+2	n	y	y	n	y
+1	y	y	n	n	n
+1	y	y	n	y	n
+1	n	y	n	y	n
0	n	n	n	n	y
0	y	n	n	y	y
0	n	y	n	y	n
0	y	y	y	y	y
-1	y	y	y	n	y
-1	n	n	y	y	n
-1	n	n	y	n	y
-1	y	n	y	n	y
-2	n	n	y	y	n
-2	n	y	y	n	y
-2	y	n	y	n	n
-2	y	n	y	n	y

(Data is from CIML appendix.)



# The Big Picture

