Lecture 23: AI, ML, NLP, ...

Artificial intelligence, machine learning, natural language processing, ...

- buzzwords, hype, real accomplishments, wishful thinking
  - big data, deep learning, neural networks, ...
- brief history
- examples
  - classification (spam detection)
  - prediction (future prices)
  - recommendation systems (Netflix, Amazon, Goodreads, ...)
  - natural language processing (sentiment analysis, translation, generation)
  - games (chess, Go)
- issues and concerns

- Beware: on this topic,
  I am even less of an expert than normal.

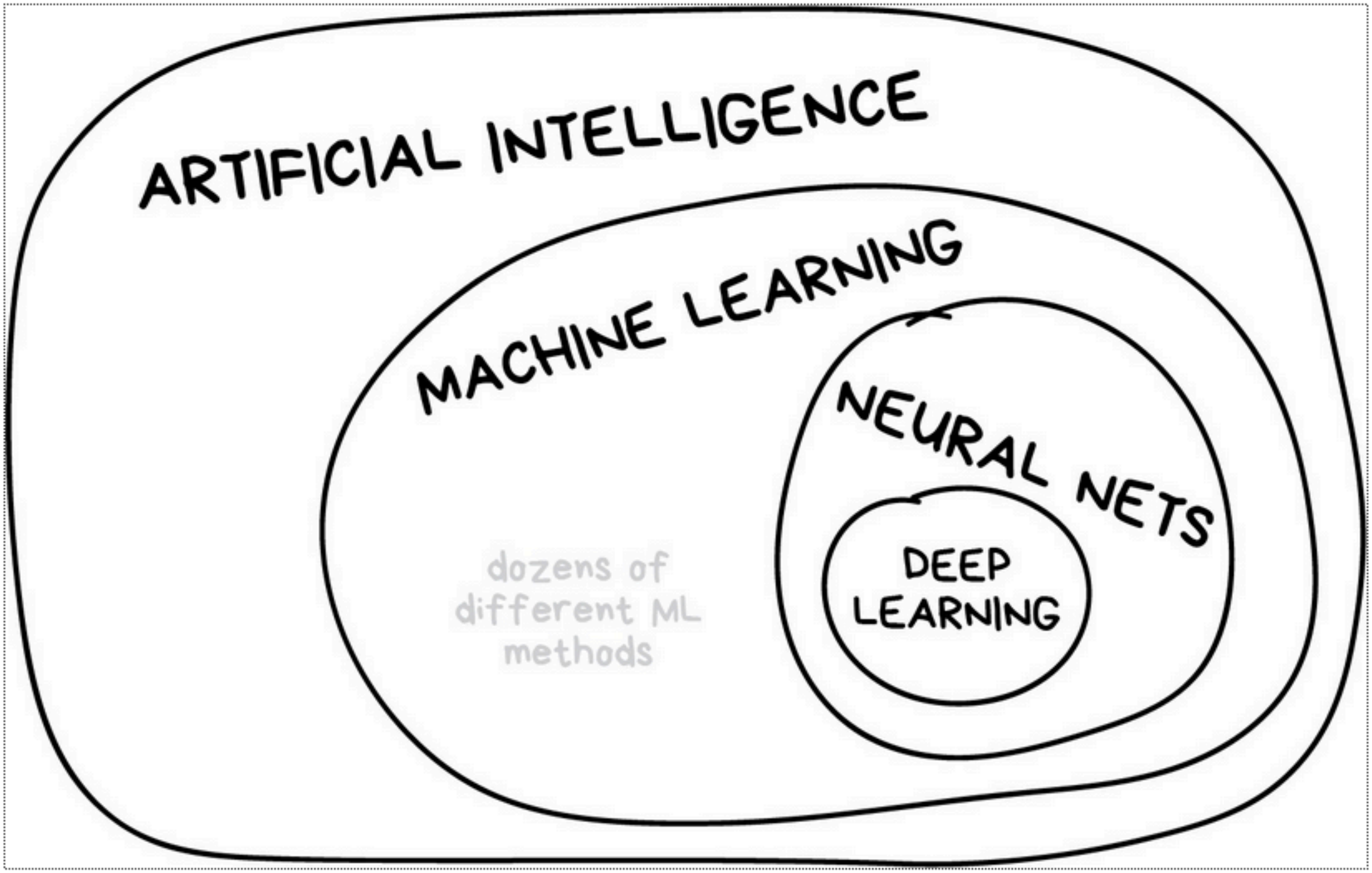
Revisionist history (non-expert perspective)

- **1950s, 1960s: naive optimism about artificial intelligence**
  - checkers, chess, machine translation, theorem proving, speech recognition, image recognition, vision, ...
  - almost everything proved to be much harder than was thought

- **1980s, 1990s: expert or rule-based systems**
  - domain experts write down lots of rules, computers apply them to make decisions
  - it's too hard to collect the rules, and there are too many exceptions
  - doesn't scale to large datasets or new problem domains

- **2010s: machine learning, big data, ...**
  - provide a "training set" with lots of examples correctly characterized
  - define "features" that might be relevant
  - write a program that "learns" from its successes and failures on the training data (basically by figuring out how to combine feature values)
  - turn it loose on new data

The big picture (vas3k.com/blog/machine_learning)



Examples of ML applications (a tiny subset)

- **classification**
  - spam detection, digit recognition, optical character recognition, authorship, ...
  - image recognition, face recognition, ...
- **prediction**
  - house prices, stock prices, credit scoring, resume screening, ...
  - tumor probabilities, intensive care outcomes, ...
- **recommendation systems**
  - e.g., Netflix, Amazon, Goodreads, ...
- **natural language processing (NLP)**
  - language translation
  - text to speech; speech to text
  - sentiment analysis
  - text generation
- **games**
  - checkers, chess, Go

Types of learning algorithms

- **supervised learning (labeled data)**
  - teach the computer how to do something with training examples
  - then let it use its new-found knowledge to do it on new examples
- **unsupervised learning (unlabeled data)**
  - let the computer learn how to do something without training data
  - use this to find structure and patterns in data
- **reinforcement learning**
  - some kind of "real world" system to interact with
  - feedback on success or failure guides/teaches future behavior
- **recommender systems**
  - look for similarities in likes and dislikes / behaviors / ...
  - use that to predict future likes / behaviors

Classification example: spam detection

- rule-based: look for odd words & spellings, known bad sources, etc.
  - V1/\6R/\, M0l\lE`/, spamRus.com, ...

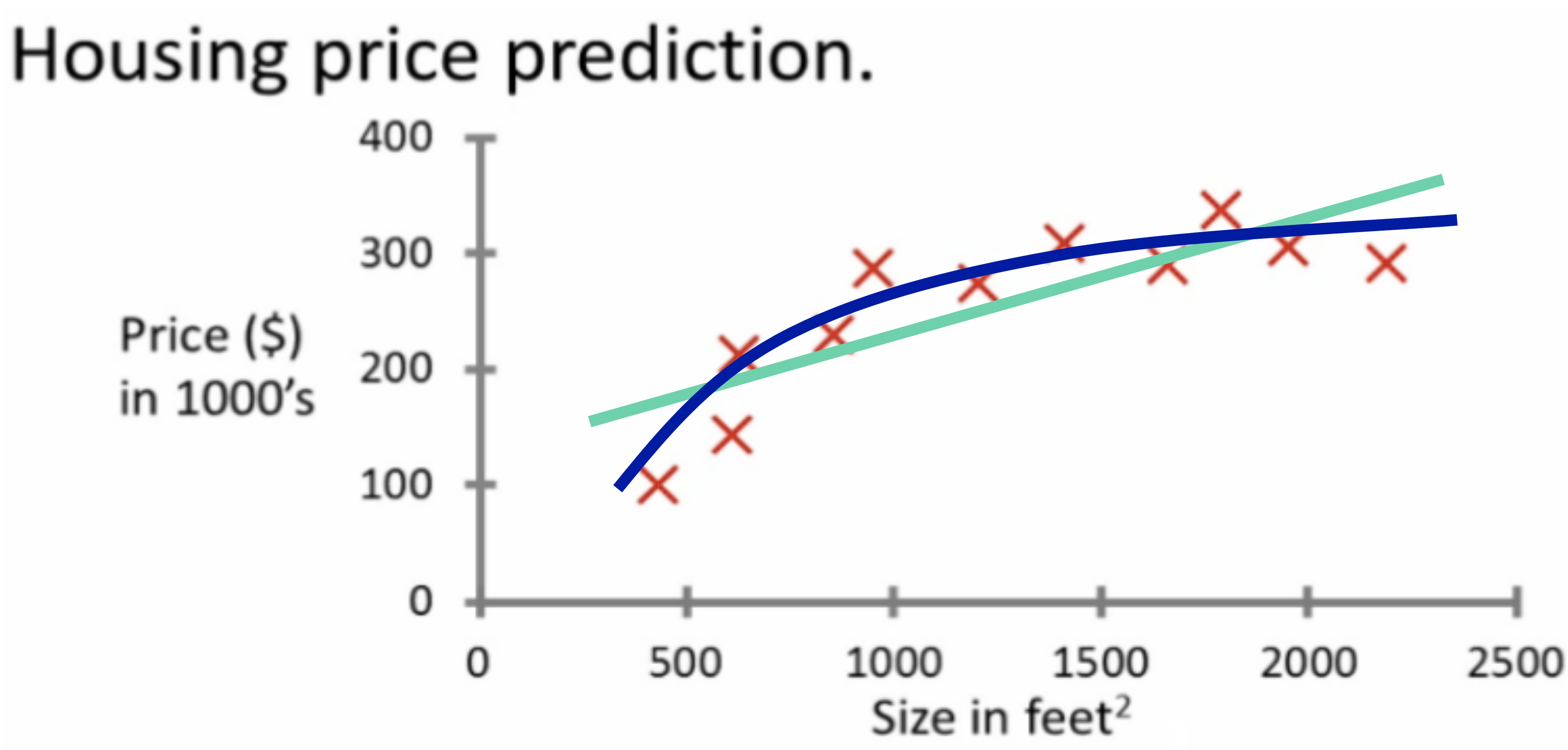
- machine learning: choose a set of features like
  - odd spelling, weird characters, language and grammar, origin, length, ...
- provide a training set of messages that are marked "spam" or "not spam"

- ML algorithm figures out parameter settings that let it do the best job of separating spam from not spam in the training set
- then apply that to real data

- potential problems:
  - training set isn't good enough or big enough
  - creating it is probably done manually
  - "over-fitting": does a great job on training set but little else
  - spammers keep adapting so we always need new training material

Prediction example: house prices

- only one feature here: square footage
- straight line? ("linear regression")
- some kind of curve?



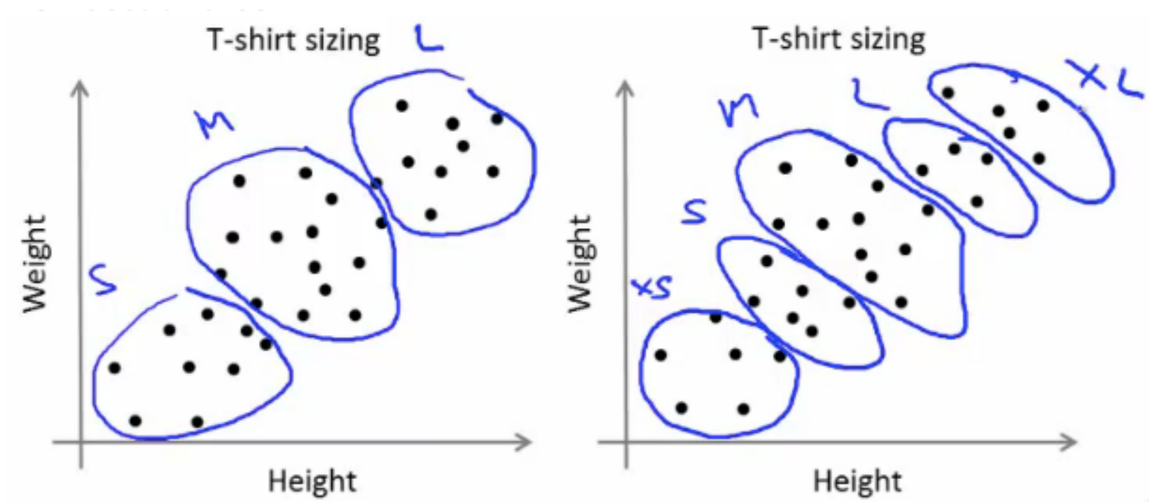
Clustering (learning from unlabeled data)

- contrast with supervised learning
  - supervised learning

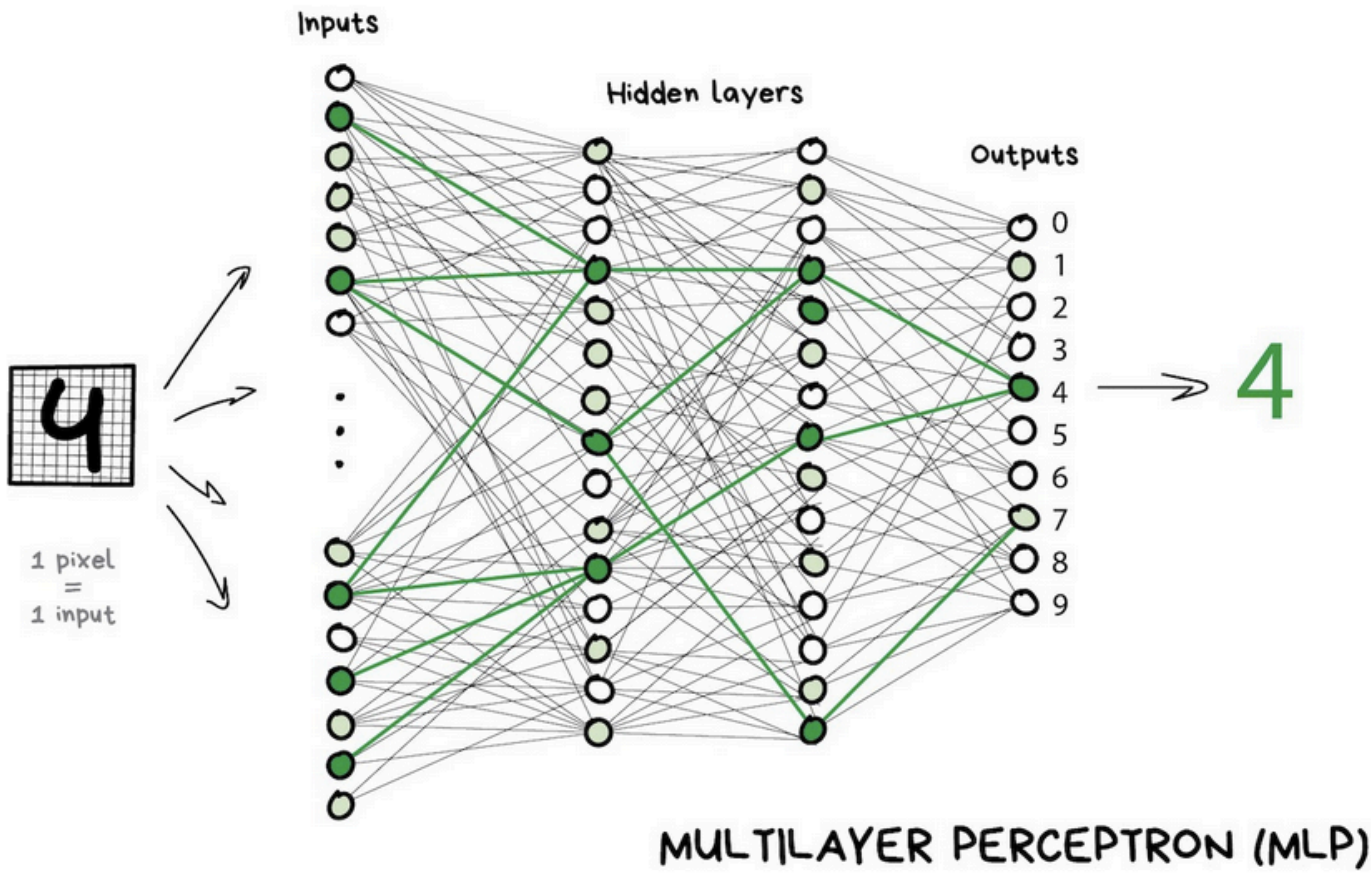
    given a set of labels, fit a hypothesis to it
  - unsupervised learning

    try and determine structure in the data

    clustering algorithm groups data together based on data features
- good for
  - market segmentation - group customers into different market segments
  - social network analysis - Facebook "smartlists"
  - topic analysis
  - authorship



Neural networks (from vas3k.com/blog/machine_learning)



Natural language processing (NLP)

- understanding text
  - parsing, syntactic structure
  - topic modeling
  - sentiment analysis
- text generation
- text to speech
- speech to text
- translation

ML / AI issues

- algorithmic fairness
  - results can't be better than training data
  - if that has implicit or explicit biases, results are biased
  - can we detect and eliminate bias?
- accountability and explainability
  - what is the algorithm really doing?
  - can its results be explained
- appropriate uses?
  - prison sentencing
  - drone strikes
  - weapon systems
  - resume evaluation
  - medical decisions
  - ...
- to learn more:

  https://fairmlbook.org