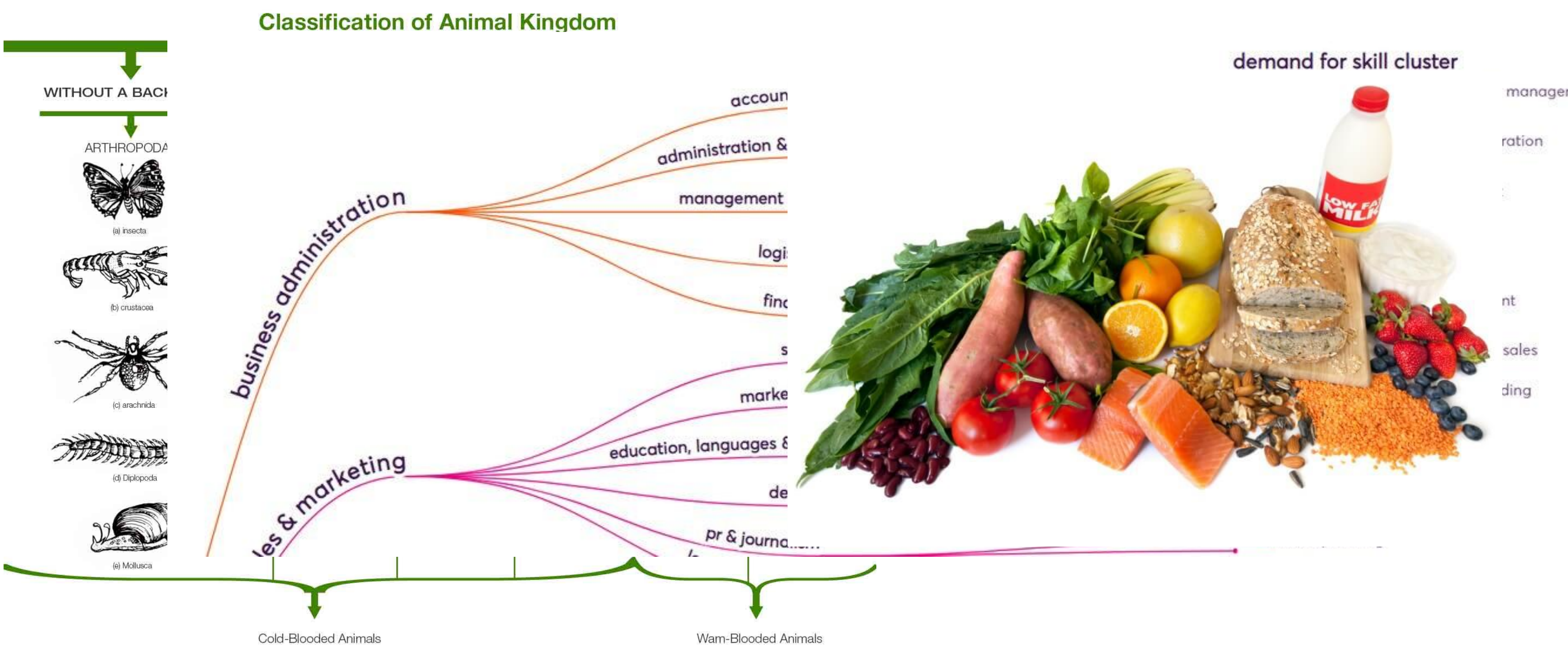
Learning from Regularities in the World

Aida Nematzadeh

Regularities in the World

People are great at learning from regularities.



Language Learning and Regularities

Learning word meanings:

A dax!

***dax* means dog**

Look at the dax!

Language Learning and Regularities

Anticipating the next words:

She is drinking coffee

lemonade

a chair

doogh

How Do We Learn These Regularities?

Probabilities: Modeling Regularities

How likely something is to happen?

½ or 50%

Possible outcomes: {heads, tails}
Current outcome: heads

Probability of an event:
# of ways it can happen / # of outcomes

Probabilities: Modeling Regularities

Flipped a coin 20 times:

What is the probability of observing heads?

P(heads) = 1 / 20 = 5 %

Language Learning and Probabilities

Anticipating the next words:

She is drinking hot cocoa lemonade

In a cold winter day, she is drinking hot cocoa

Using knowledge given the situation: hot drink is better on a cold day.

lemonade

Combining Evidence & Knowledge

In a cold winter day, she is drinking

evidence

hypothesis 1

hot cocoa

hypothesis 2

lemonade

The Bayes rule:

likelihood prior

posterior

$$P(h|e) = \frac{P(e|h)P(h)}{\sum_{h'}(e|h')P(h')}$$

hypothesis evidence

P(*e*)

Evidence & Knowledge in Action

In a cold winter day, she is drinking hot cocoa h_1

evidence

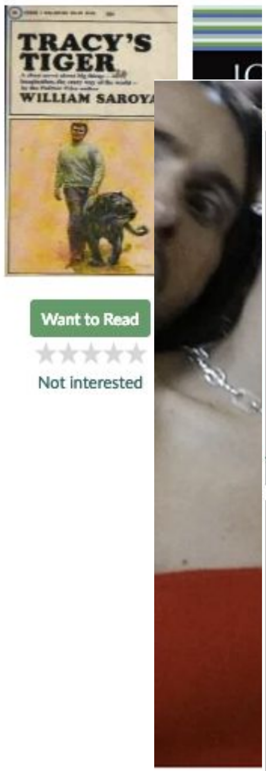
lemonade h_2

Let's calculate the probability of each hypothesis!

Can Computers Learn These Regularities?

Machine Learning

Machine Learning of Regularities



When Hypotheses Are Classes

H_1: the cat class

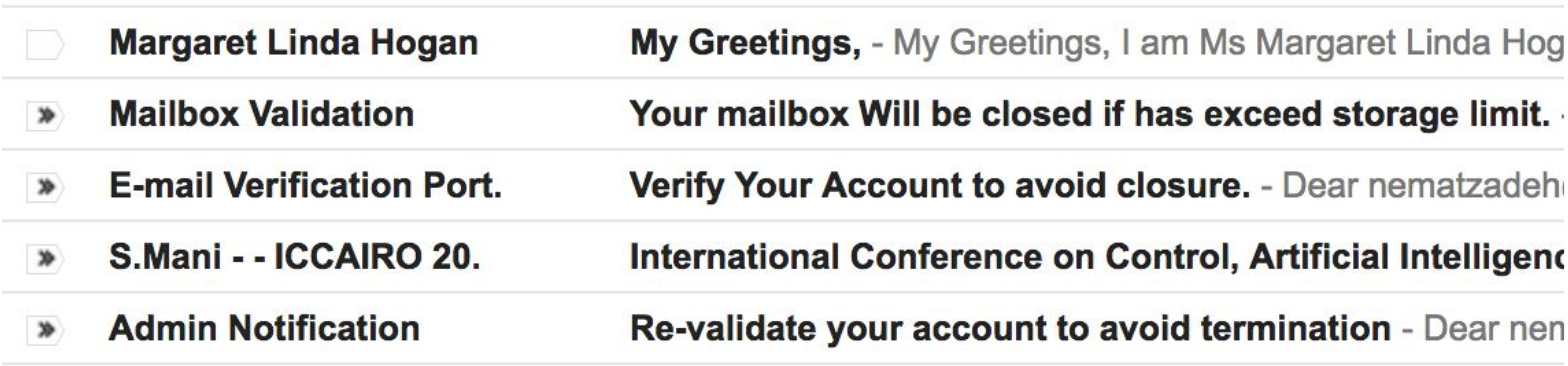
H_2: the dog class

classification

Many language processing tasks are classification!

Spam Detection

Classify email to spam and non-spam



“Dear winner”; “Click this link”; “Urgent: send me your credit card information”; …

Authorship Attribution

Find the text author and author's characteristics:

- Gender, age, etc

Study claims Agatha Christie had Alzheimer's

Textual analysis detects signs of sharply declining faculties towards the end of beloved mystery writer's life

An in-depth analysis of Agatha Christie's novels has suggested that the much-loved author of more than 80 mysteries was suffering from Alzheimer's disease.

Academics at the University of Toronto studied a selection of Christie's novels written between the ages of 28 and 82, counting the numbers of different words, indefinite nouns and phrases used in each.

Sentiment Analysis

Positive/negative orientation (sentiment) of text:

- Book, restaurant, movie and product reviews
- Political text

It's just gorgeous, like a flipbook made of dreamy vintage postcards that are somehow about contemporary life in Los Angeles.

The cinematography and special effects are fantastic, but don't actually compensate for a weak storyline, and forgettable musical numbers.

Classification: Formalism

Given an input & fixed classes $C=c_1, c_2, \ldots, c_M$, find:

- the probability of each class
- the predicted class c_i

Supervised training: uses data points & their gold-standard labels, $(x_1, c_1), (x_2, c_2), \ldots, (x_N, c_N)$

Goal: Find the correct class for the new data point

Classification: Example

data x

class c

training

Featuring a case that jumps around like a jack-rabbit, look like a return to form.

The result is a season of television that seems at once overstuffed and thinned out - long on character and events, but short on any sense of mounting tension, urgency or consequences.

It's never as cutting edge or plainly cool as it so desperately wants to be.

Exhausted, and exhausting.

It is joyful. It leaps off the screen.

---

test

prediction

It's not a bad series, in fact it is entertaining, but I fear that it may move on to being forgotten. [Full Review in Spanish]

?

Classification Algorithms

generative models

prior

$$P(c|d) = \frac{P(d|c)P(c)}{P(d)}$$

posterior

can use prior info

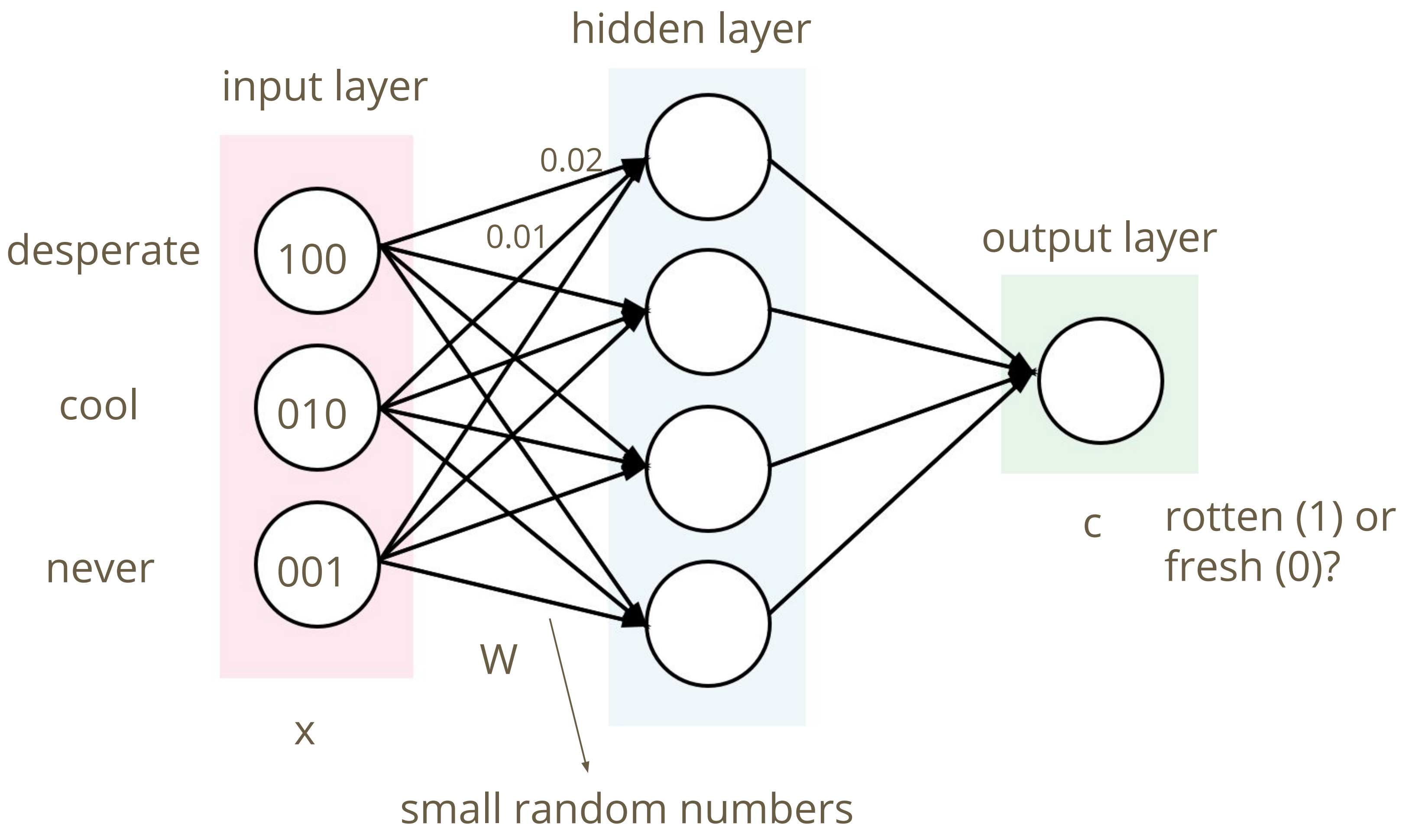
discriminative models

$$P(d|c).$$

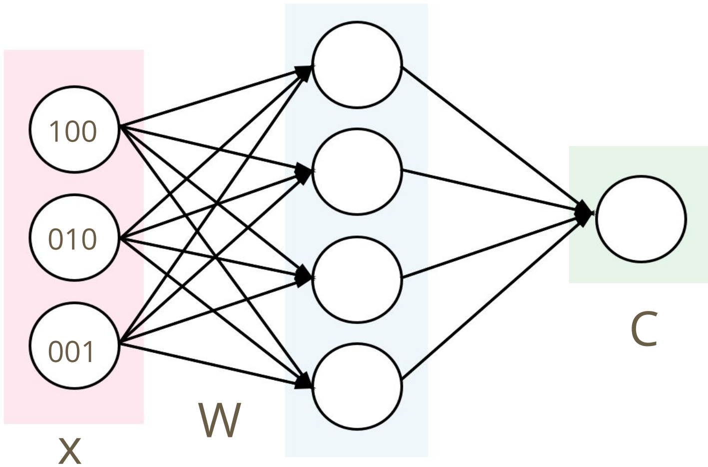
likelihood

easier to train

Neural Networks as Classifiers



Neural Networks as Classifiers



Y = Wx

C = f(Y)

Compare Y with the “real” output

Calculate an error: how similar the calculated and real output are.

Update W based on the error.

Neural Networks as Classifiers

$Y = Wx$

$C = f(Y)$

Compare Y with the “real” output

Calculate an error: how similar the calculated and real output are. **(forward pass)**

Update W based on the error. **(backpropagation)**

Neural Network Playground!

Let's watch a neural network being trained!

Evaluating Classifier Performance

Consider a binary detection task.

- Label text *positive* or *negative*.

Gold labels: human labels used as ground truth.

- Gold label is either positive or negative.

Need *metrics* to quantify classifier's performance.

Evaluation Metrics: Precision/Recall

Precision: correctly labeled + / all labeled +.

Recall: correctly labeled + / all truly +.

F-measure combines both: harmonic mean.

- Weighs min of two more heavily.

Questions?

Improving Sentiment Analysis

Dealing with negation:

I didn't love the food **vs** *I loved the food.*

Add a prefix after negation (n't, not, no, never)

I didn't NOT-love NOT-the NOT-food: