09/28/2017

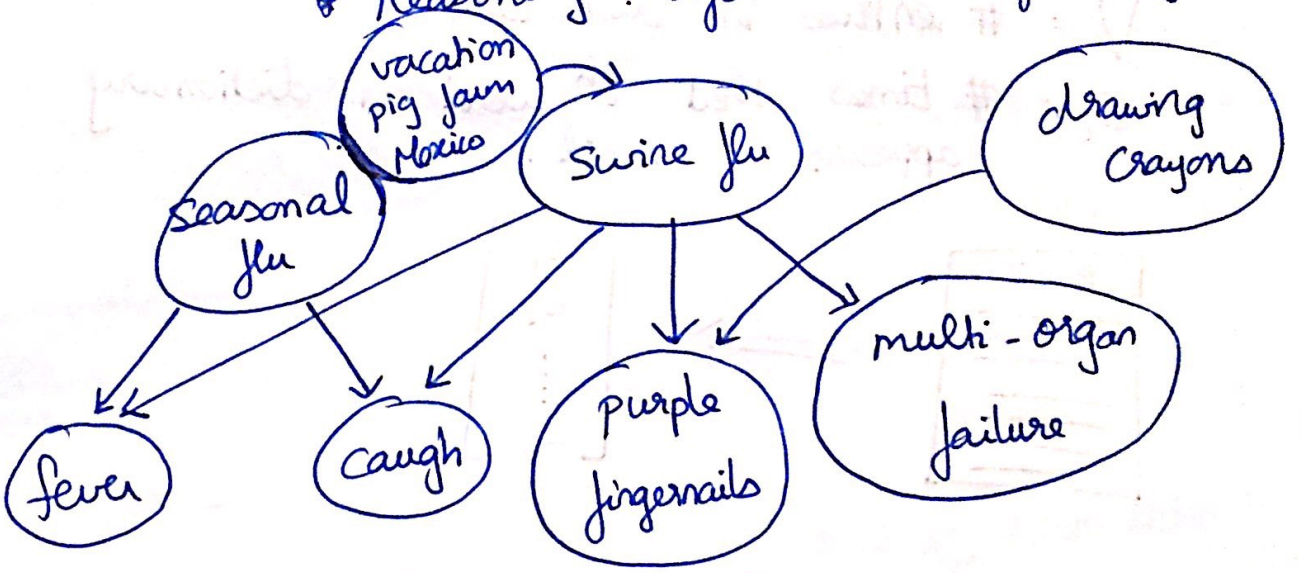
Scribe : Nivetha

Lecture - 1

Probabilistic reasoning

Ex : medical diagnosis

* knowledge representation : diseases cause symptoms
* Modeling uncertainity : some diseases, some symptoms more likely than others.
* Reasoning : infer diseases from symptoms



As more and more diagnosis comes in, inferences can change / flip.

How do graphs represent correlation, causation, independence?

Marriage of graph theory and probability

Prediction

Ex : spam filter

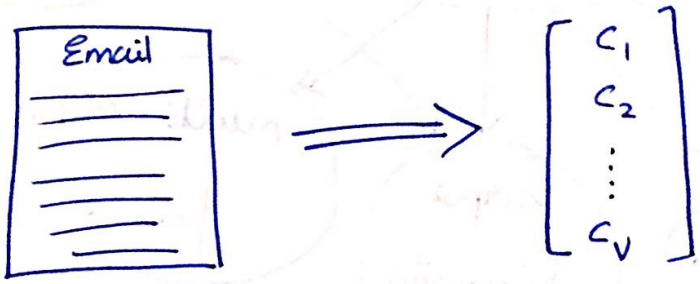
input : email message

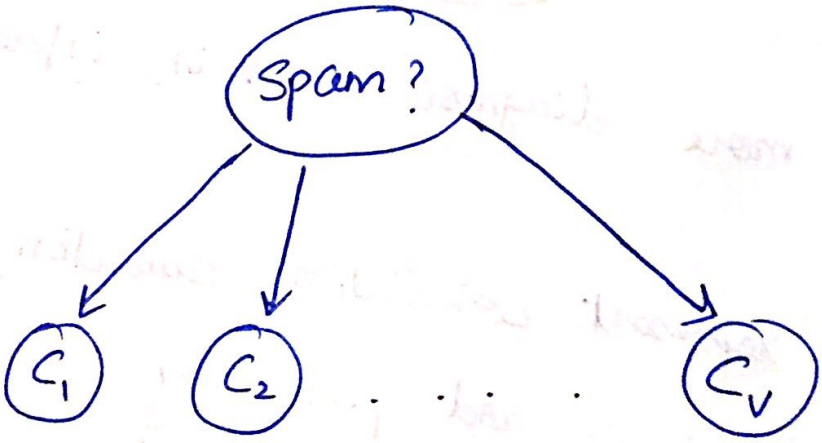
output : {spam, not spam}

How to represent input ? convert text to fixed-length vector of word counts

V = # entries in dictionary

c_i = # times that i^{th} word in dictionary appears in email.

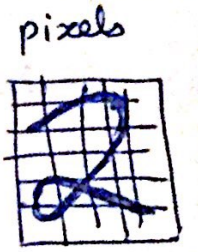




Certain words more / less likely to appear in spam.

How to quantify / estimate ?

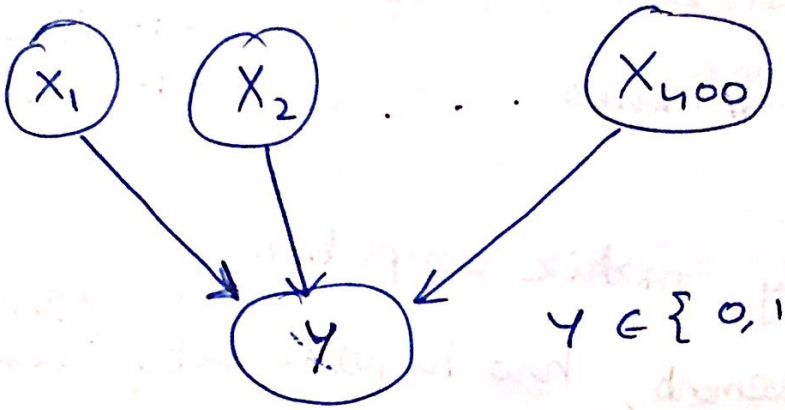
Ex: handwriting recognition



inputs : grayscale images (20×20)

outputs : labels $\{0, 1, 2, \ldots 9\}$

Represent image by vector $\vec{x} \in \mathbb{R}^{400}$ with one element per pixel.



Pattern analysis and discovery

Ex : topic modeling

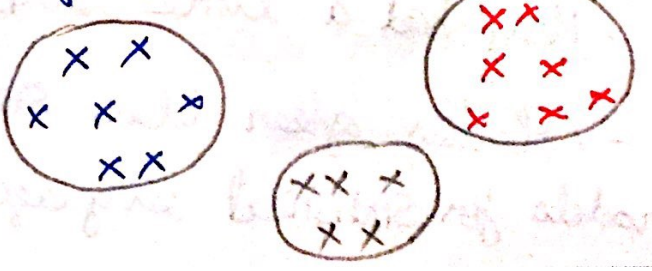
- how to organize large collection of (unlabeled) documents ?

* more generally, clustering

inputs $\{\vec{x}_1, \vec{x}_2, \ldots \vec{x}_N\}$ $\vec{x}_i \in \mathbb{R}^D$

How to group inputs when no labels are provided ?

Map inputs to discrete label $y \in \{1, 2, \ldots C\}$

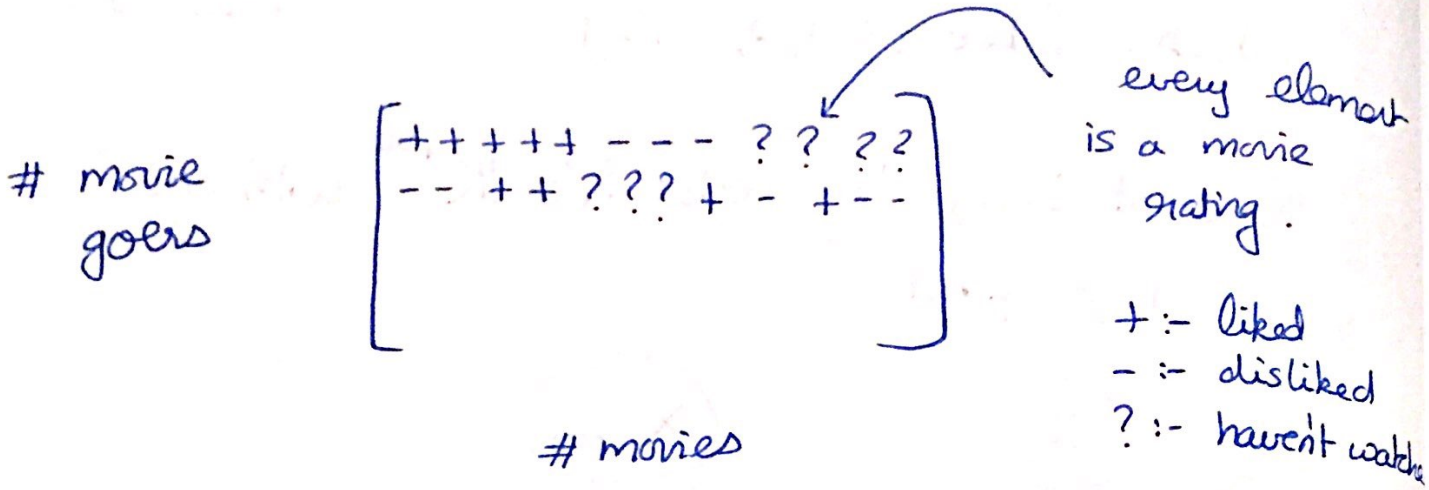


$D = 2$

$C = 3$

Ex : collaborative filtering

How to build a movie recommendation system?



* more generally, matrix completion

Given some elements, how to fill in rest? (In the above example, the ?)

Sequential modeling

* How to model systems whose "state" changes over time or it has other extended representation?

Ex: text (written language)

"States" = words

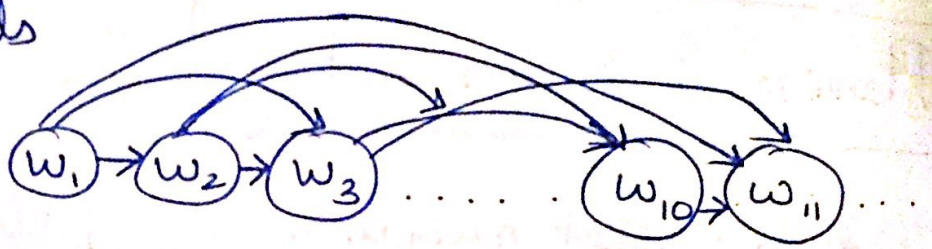
Which sentence is more likely?

- Mary had a little lamb
- Colorless green ideas sleep furiously

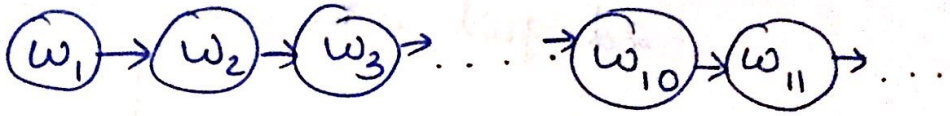
⇒ Markov models for statistical language processing

Graphical models

model A



model B



model A is richer but hard to estimate

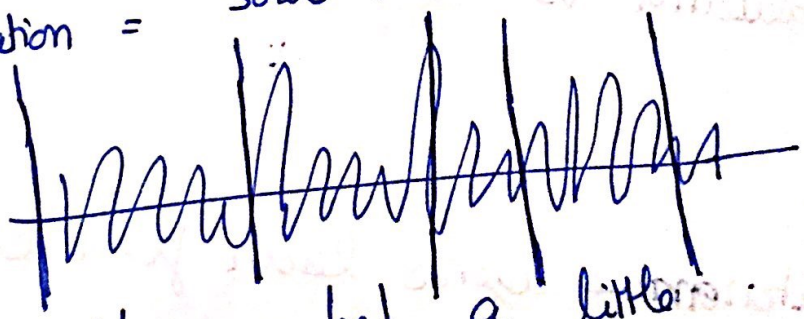
model B is wrong but very easy to estimate
(too simple)

Look for the simplest model that gets the job done!

Ex: speech (spoken language)

States = words (or syllables, or smaller units of speech)

Observation = sounds, waveforms



Mary had a little ...

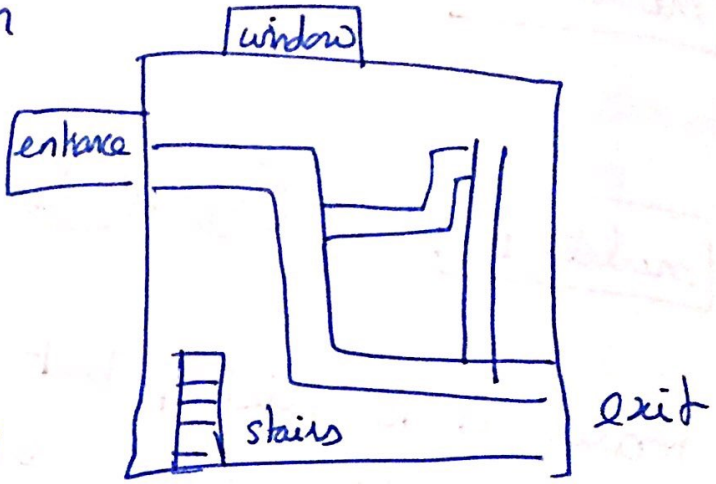
How do we infer words from waveforms?

⇒ hidden Markov models
for speech recognition

Planning & decision-making

* Ex: Robot navigation

2d grid world



"states" = cells on 2d grid

"actions" = attempts to move N, S, E, W

* noisy dynamics (moves are not deterministic, there is uncertainty)
* rewards = feedback from environment
  - delayed vs immediate rewards
  - evaluative vs instructive

More generally,

how can autonomous agents learn from experience?

⇒ Markov decision processes, reinforcement learning

Other "embodied" agents :

self-driving cars, self-flying drones

Other "embedded" agents :

game-playing agents (AlphaGo)

Themes of class

1) Probabilistic models of uncertainty

2) Principles vs heuristics

Inference as calculations
Learning as optimizations
} vs rules-of-thumb

3) Power vs tractability : how to develop compact representations of complex worlds ?