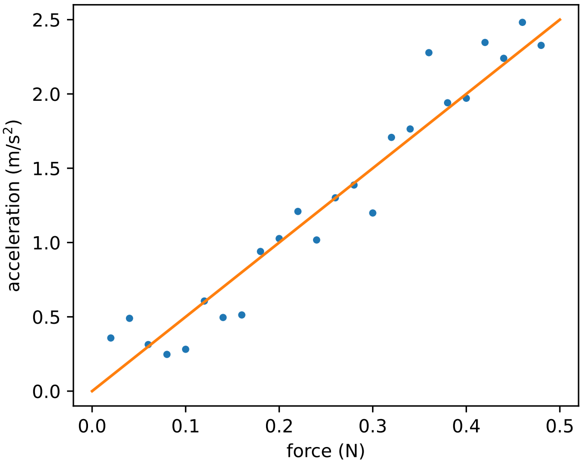
Machine Learning

Lecture 6: Tasks

Hao Tang

October 3, 2022

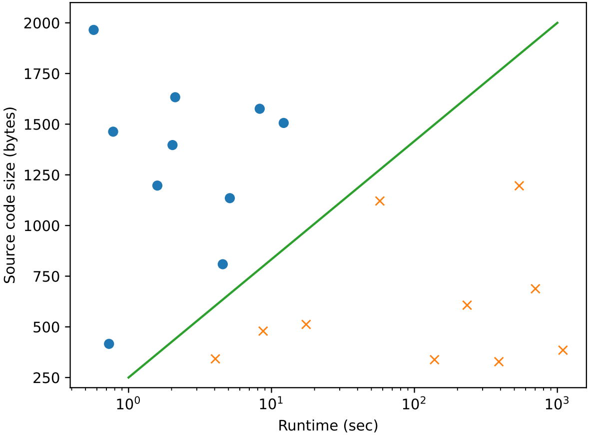
Linear regression



Linear regression

- We have learned how to write a program $f(x) = w^\top \phi(x)$ with data.
- More abstractly, we have learned how to write a program $f : \mathbb{R}^d \to \mathbb{R}$ with data.
- If we have k regression models, we can learn a program $f : \mathbb{R}^d \to \mathbb{R}^k$ with data.
- In fact, we can write $f(x) = W^\top \phi(x)$ where W is now a matrix instead of a vector to implement $\mathbb{R}^d \to \mathbb{R}^k$. (What's the optimal solution in this case?)

Linear classification

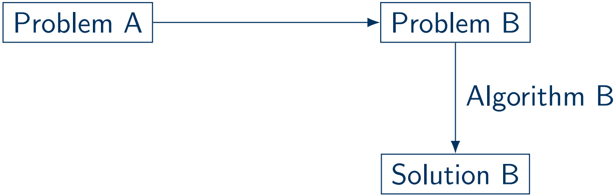


Linear classification

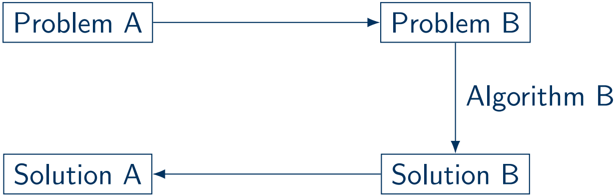
- We have learned how to write a program $f(x) = w^\top \phi(x)$ with data.
- The input is of type $\mathbb{R}^d$.
- The output is of type $\mathcal{Y}$, where $\mathcal{Y}$ is a finite, discrete set.
- When $\mathcal{Y}$ is $\{+1, -1\}$, the task is called binary classification; otherwise, it is multiclass classification.

Problem reduction

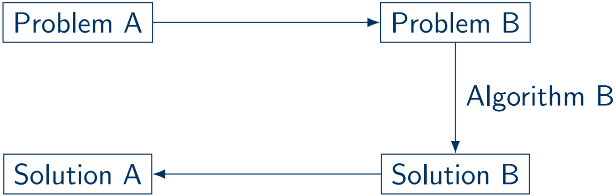
Problem reduction



Problem reduction

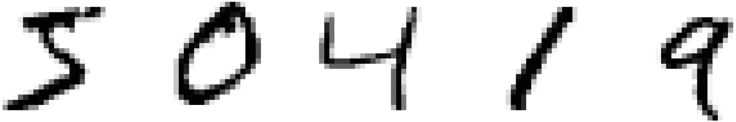


Problem reduction



Problem B is harder than Problem A

Digit recognition



Digit recognition

- The input is of type $\mathbb{R}^{28\times 28}$.
- The output is of type $\{0, 1, \ldots, 9\}$.
- The goal is to learn a function $\mathbb{R}^{28\times 28} \to \{0, 1, \ldots, 9\}$.
- The task is multiclass classification.

Multiclass classification

- Recall that we use

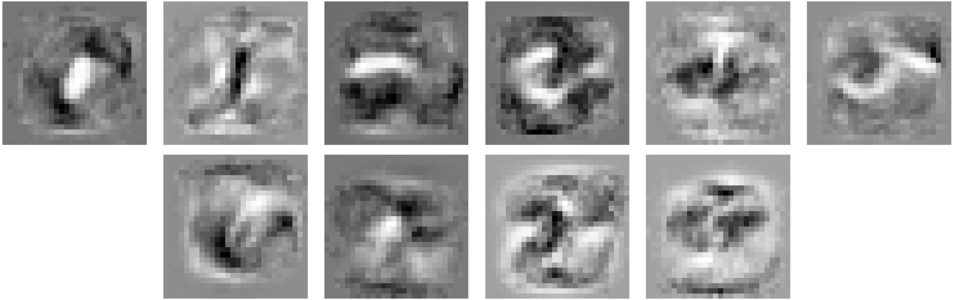
$$f(x) = \operatorname*{argmax}_{y \in \mathcal{Y}} w_y^\top \phi(x) \tag{1}$$

  to make prediction.

- Think of w_y as a "template" of the class y, and think of dot product as a measure of similarity.

- The function f can be interpreted as finding the class that is the most similar to the input.

Learned w



Face recognition

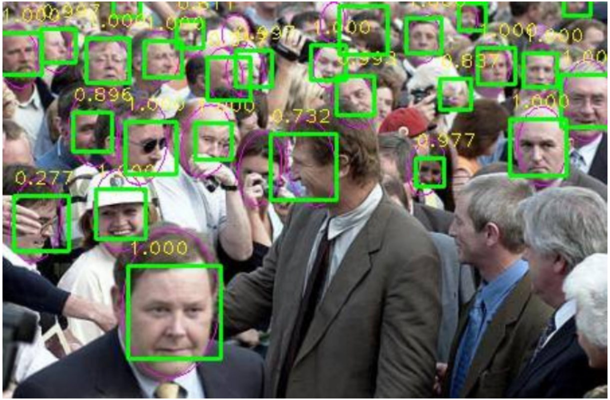
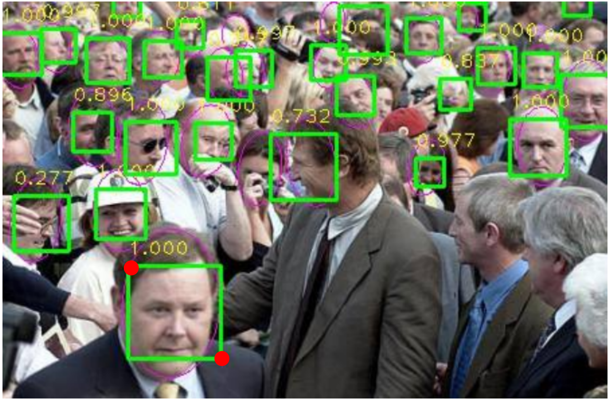


Image credit: (Triantafyllidou and Tefas, 2016)

Face recognition

- Is there a face in the image?
- The input is of type $\mathbb{R}^{m \times n}$, the set of images of $m \times n$ pixels.
- The output is of type $\{+1, -1\}$.
- This is binary classification.

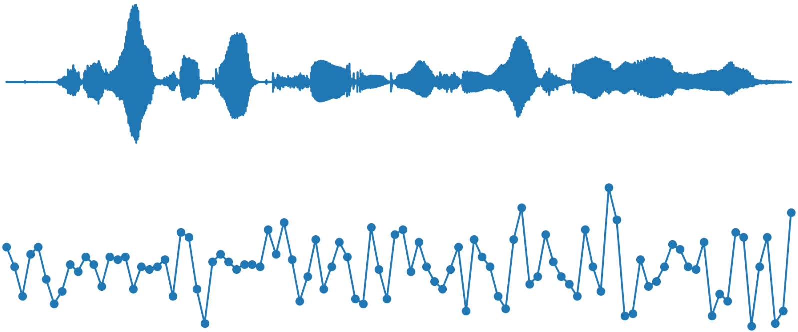
Face recognition



Face recognition

- Where is the face?
- Instead of asking where, we can ask if a bounding box is good.
- The input is of type $\mathbb{R}^{m\times n} \times \mathbb{N}^2 \times \mathbb{N}^2$.
- The output is of type $\{+1, -1\}$.
- This is binary classification.

Speaker identification



Speaker identification

- Who is this?
- The input is of type $\mathbb{R}^T$, the set of T-sample waveforms.
- The output is of type $\{1, \ldots, K\}$, say, if we have K speakers.
- This is multiclass classification.

Speaker recognition

Speaker recognition

- Is this the same person?
- The input is of type $\mathbb{R}^T \times \mathbb{R}^T$, pairs of T-sample waveforms.
- The output is of type $\{+1, -1\}$.
- This is binary classification.

Speech recognition

She had your dark suit in greasy wash water all year.

Speech recognition

- The input is of type $\mathbb{R}^T$, a T-sample waveform.
- The output is of type V^*, where V is a set of possible words.
- The type V^* is called the Kleene closure of V, meaning zero or more items concatenated.
- One common approach is to cast this problem as a sequence of multiclass classification.

Speech synthesis

She had your dark suit in greasy wash water all year.

Speech synthesis

- The input is of type V^*, a set of sentences.
- The output is of type $\mathbb{R}^T$, a set of T-sample waveforms.
- One common approach is to cast this problem as a sequence of regression.

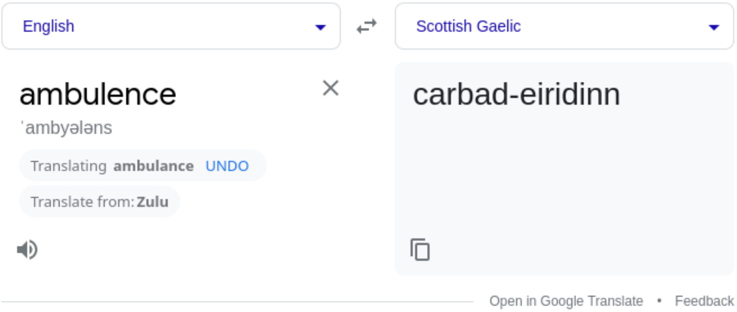
Sentiment analysis

Meets all expectations. Solid build. Quick charge. Great multi coloured lights. Even the touch pad ain't too bad. Not as good as a £500 laptop obviously but does the job.

Sentiment analysis

- The input is of type V^*, the set of sentences.
- A word can be represented as a one-hot $\mathbb{R}^{|V|}$ vector.
- The output is of type $\{1, 2, \ldots, 5\}$.
- This is a multiclass classification.

Machine translation



Machine translation

Machine translation

- The input is of type U^*, the set of sentences in one language.
- The output is of type V^*, the set of sentences in another language.
- One common approach is to cast this problem as a sequence of multiclass classification.

Example representations

- SIFT and HOG features for images
- Co-occurrences and n-grams for texts
- Spectrograms for speech
- Representations learned by neural networks