AMMI Deep Learning DIY:

Day 3

Timothée Lacroix, Pierre Stock, Alexandre Sablayrolles

Welcome to the high dimension

High dimension: what ? why ?

- Most "real world" data is in high dimensions

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?
- Multi-dimensional data: each axis represents a feature

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?
- Multi-dimensional data: each axis represents a feature
  - Food (sweetness, sourness, bitterness, saltiness, color, ...)

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?
- Multi-dimensional data: each axis represents a feature
  - Food (sweetness, sourness, bitterness, saltiness, color, ...)
- Some data has very very high dimensionality

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?
- Multi-dimensional data: each axis represents a feature
  - Food (sweetness, sourness, bitterness, saltiness, color, ...)
- Some data has very very high dimensionality
  - Images: $500 \times 300 \times 3 = 450,000$

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?
- Multi-dimensional data: each axis represents a feature
  - Food (sweetness, sourness, bitterness, saltiness, color, ...)
- Some data has very very high dimensionality
  - Images: $500 \times 300 \times 3 = 450,000$
  - Text: Vocabulary is $\approx 10,000$ words

High dimension: what ? why ?

- Most "real world" data is in high dimensions
- Why ?
- Multi-dimensional data: each axis represents a feature
  - Food (sweetness, sourness, bitterness, saltiness, color, ...)
- Some data has very very high dimensionality
  - Images: $500 \times 300 \times 3 = 450,000$
  - Text: Vocabulary is $\approx 10,000$ words
  - The list goes on!

The surprises of high dimension

- The high dimension is not what you think it is

The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in

The surprises of high dimension

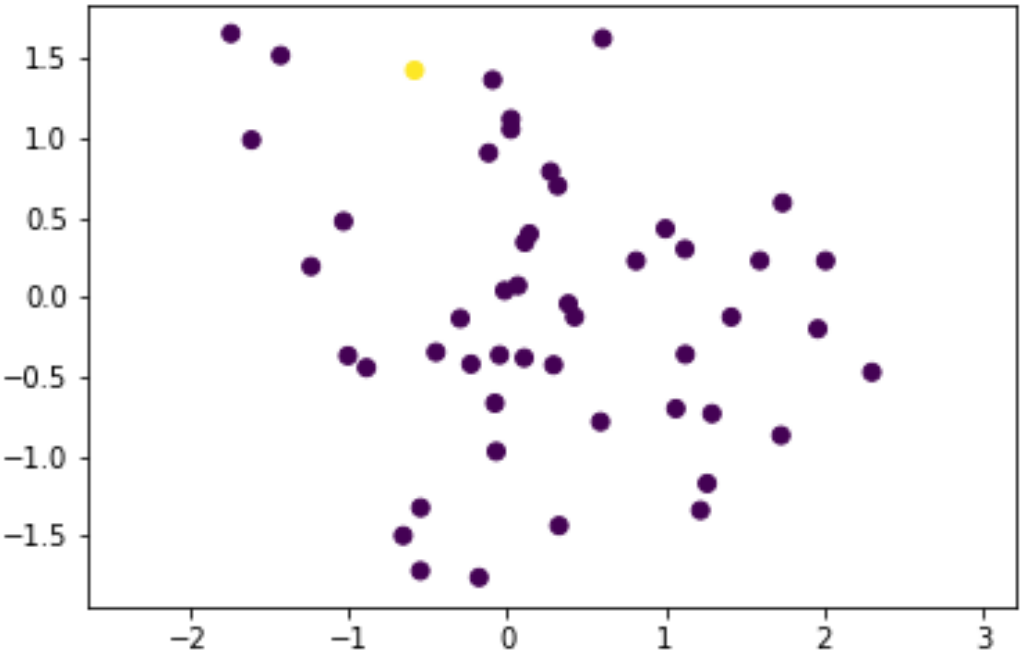
- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour

The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour

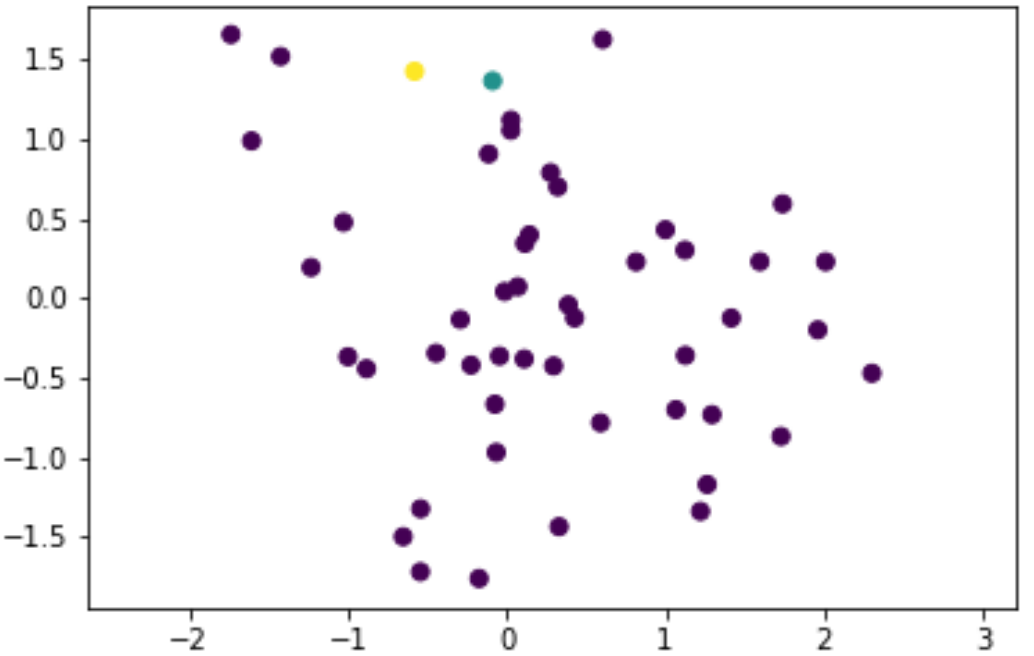
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



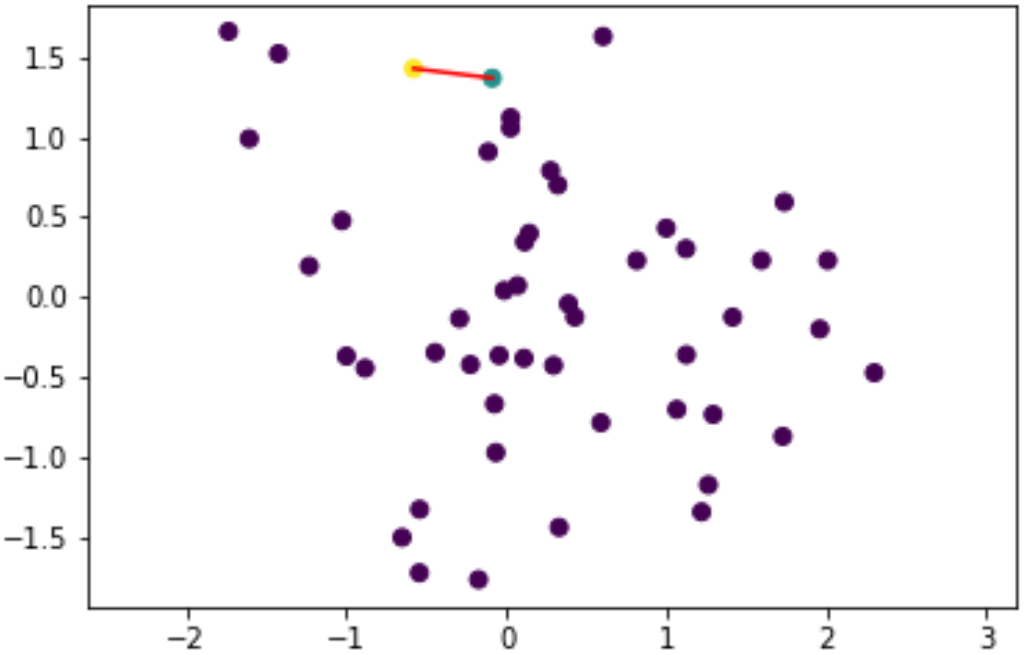
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



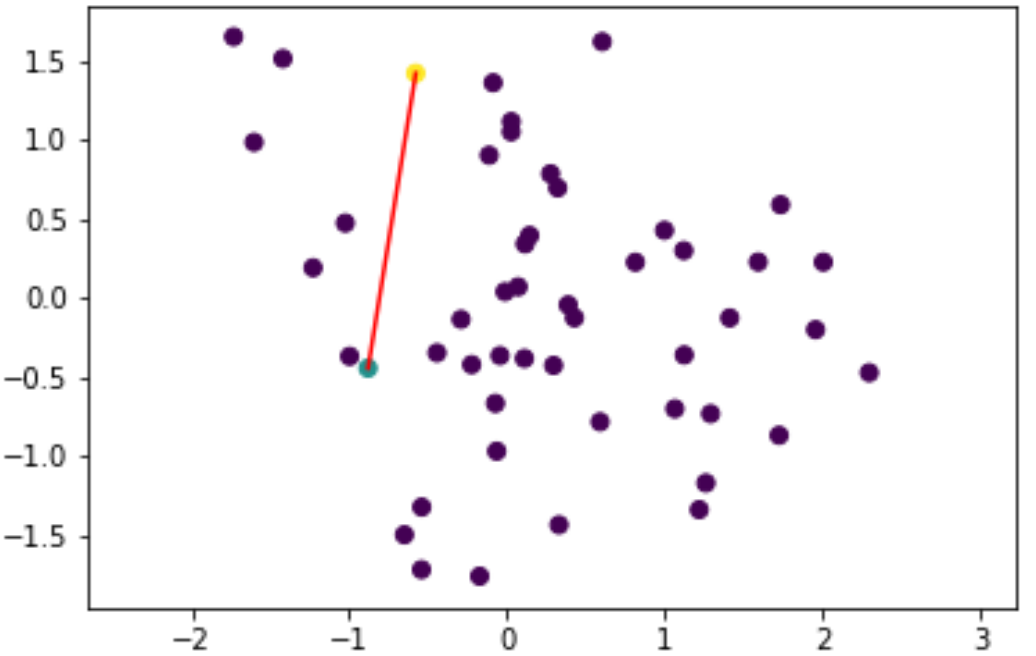
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



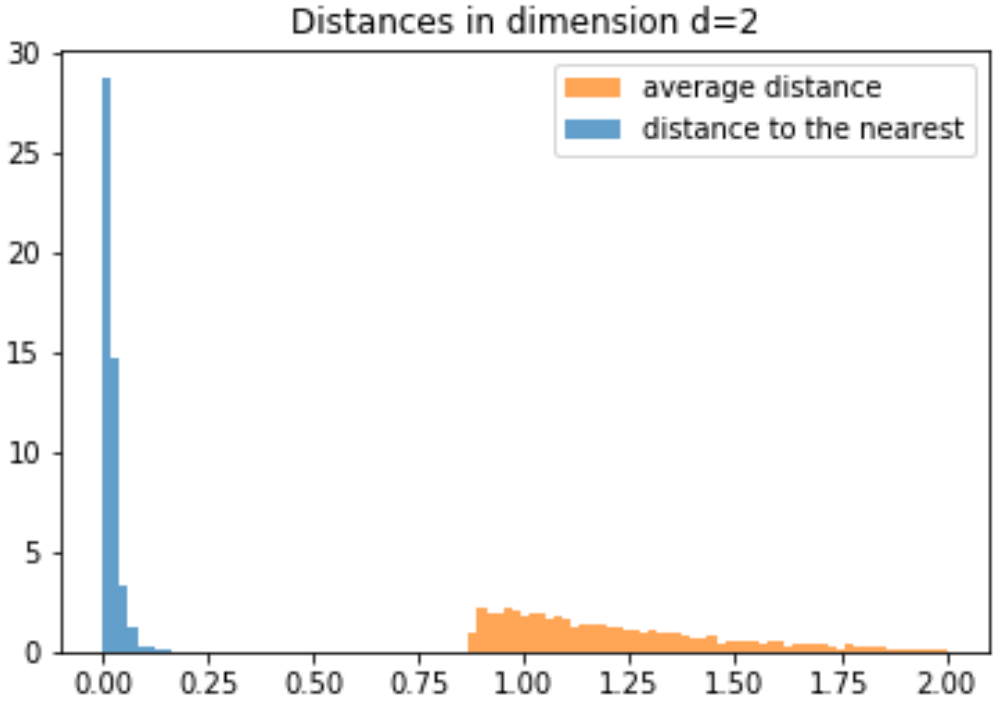
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



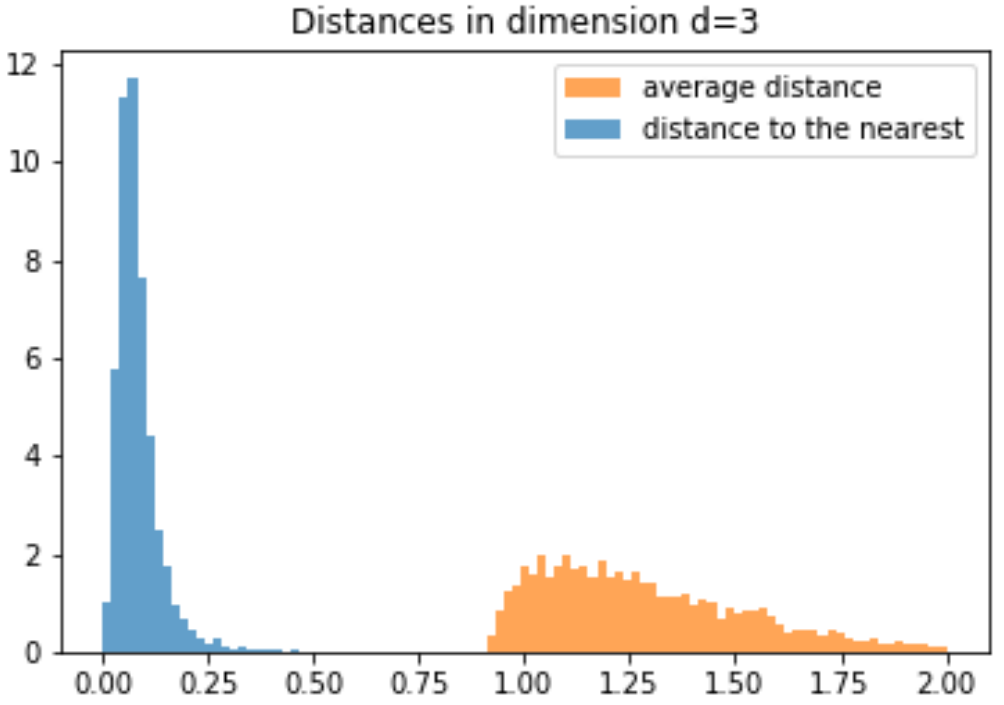
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



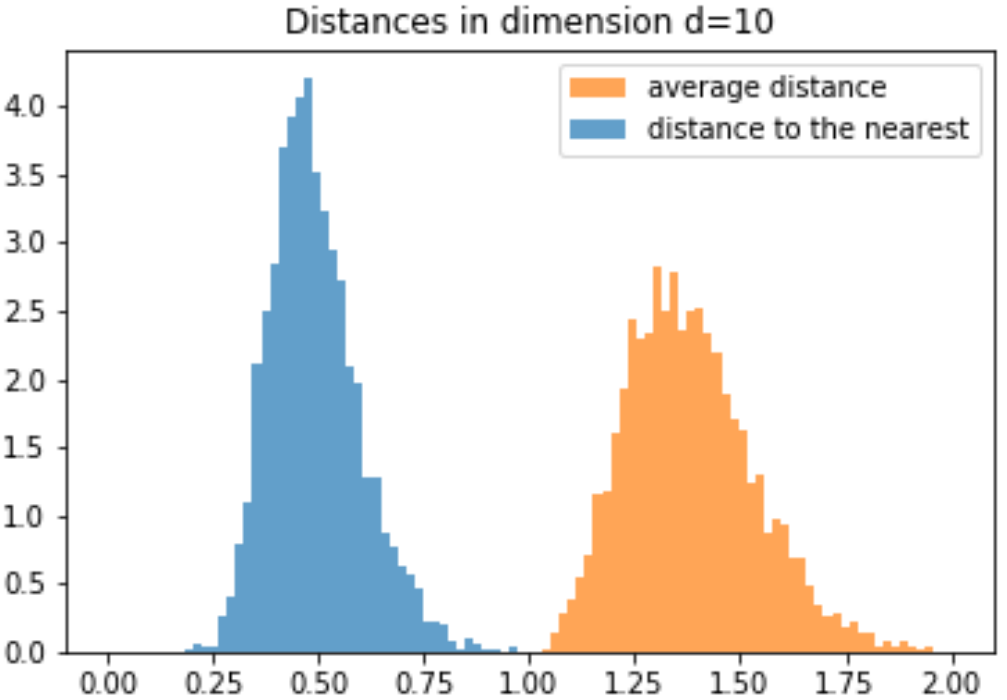
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



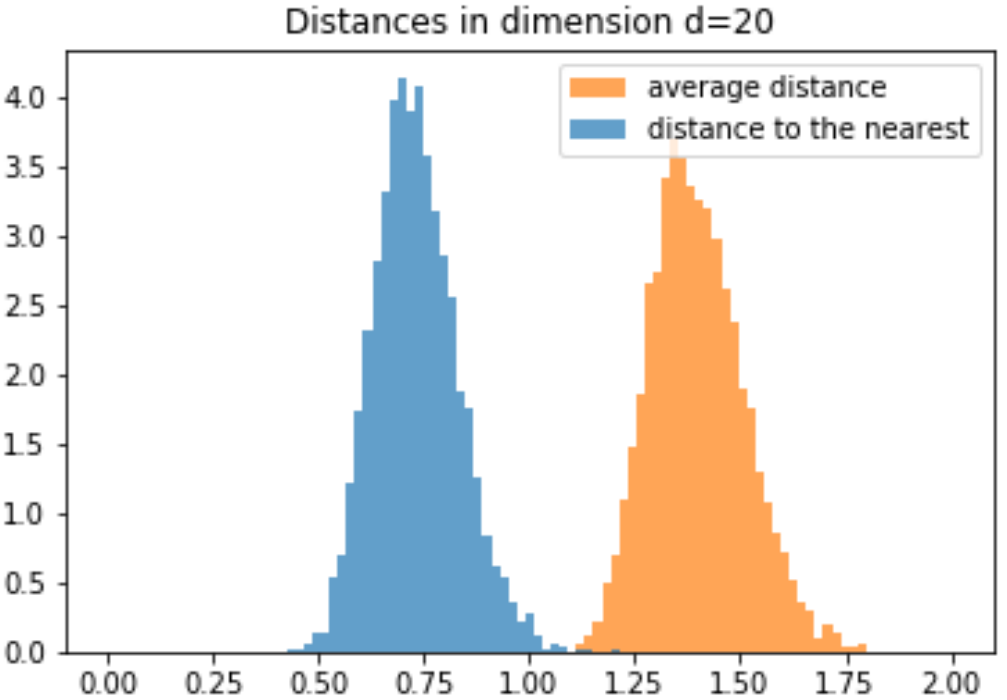
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



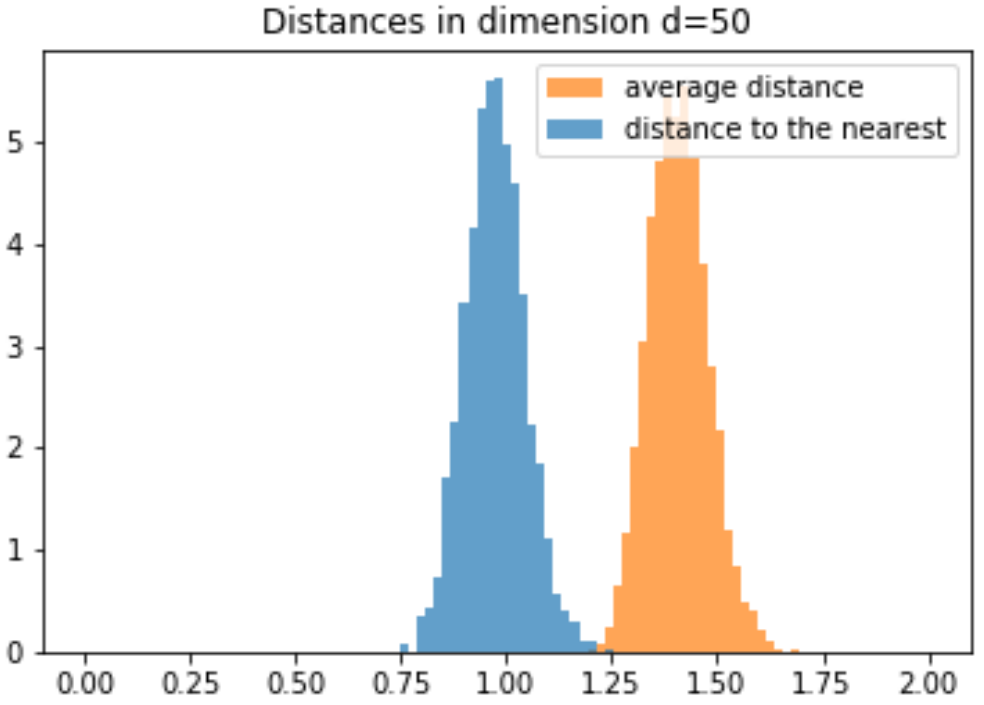
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



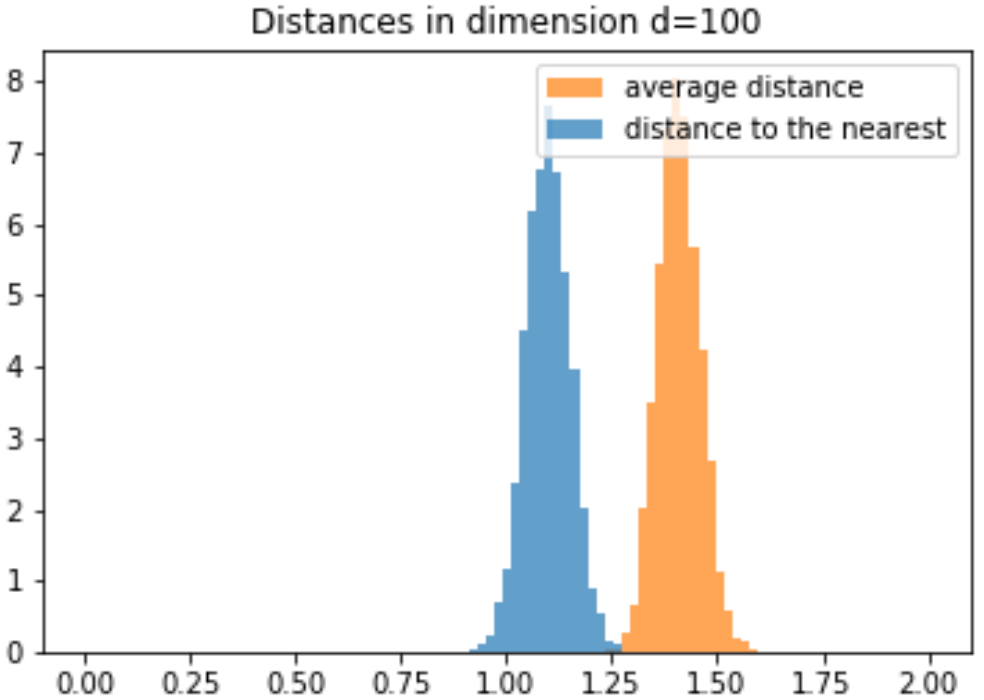
The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour



The surprises of high dimension

- The high dimension is not what you think it is
  - It is not like the 2D or 3D world we live in
- Distribution of distances to the nearest neighbour

- In high dimensions, the "nearest" neighbor is actually very far
  - It almost is as far as the all the other points

Embeddings

- We talked about high-dimensional data that is **interpretable**

Embeddings

- We talked about high-dimensional data that is **interpretable**
  - Food: different features

Embeddings

- We talked about high-dimensional data that is **interpretable**
  - Food: different features
  - Images: colors values of the pixels

Embeddings

- We talked about high-dimensional data that is **interpretable**
  - Food: different features
  - Images: colors values of the pixels
  - Text: words of the vocabulary

Embeddings

- We talked about high-dimensional data that is **interpretable**
  - Food: different features
  - Images: colors values of the pixels
  - Text: words of the vocabulary
- Sometimes we have vectors that are not interpretable

Embeddings

- We talked about high-dimensional data that is **interpretable**
  - Food: different features
  - Images: colors values of the pixels
  - Text: words of the vocabulary
- Sometimes we have vectors that are not interpretable
  - activations of a neural network

Neural networks

- A neural network is a stack of linear layers and activation functions

Neural networks

- A neural network is a stack of linear layers and activation functions
  - the linear layer is a weight matrix W and a bias b:

$$h = Wx + b$$

Neural networks

- A neural network is a stack of linear layers and activation functions
  - the linear layer is a weight matrix W and a bias b:

$$h = Wx + b$$

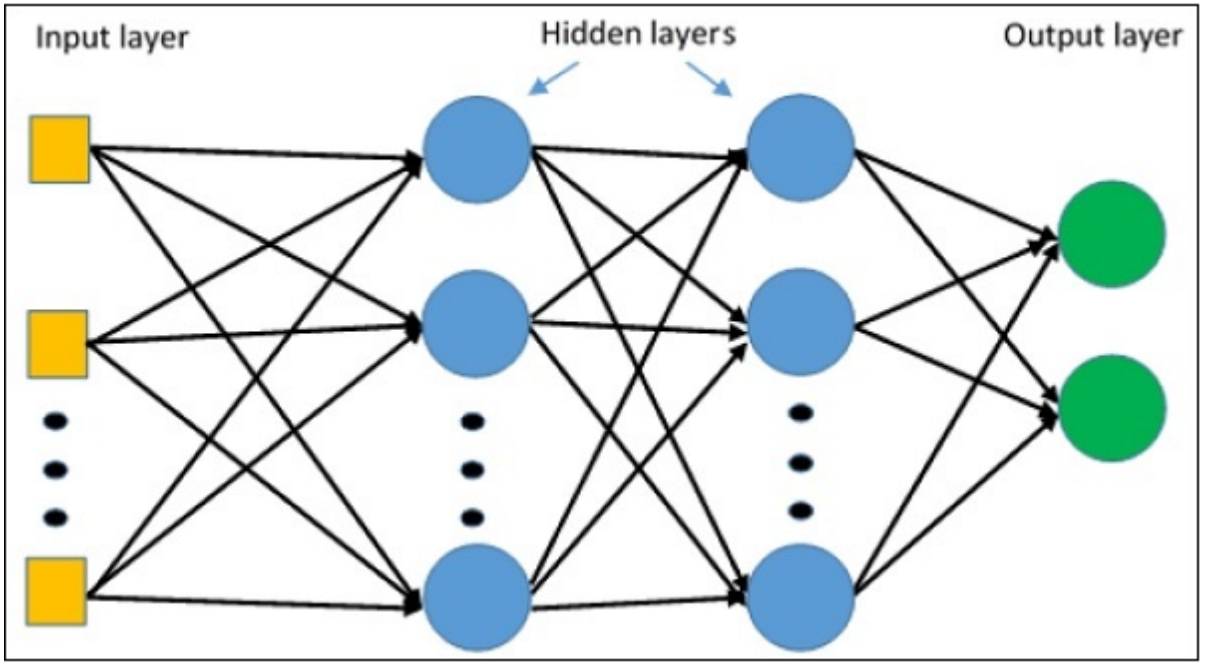
  - the activation function can be σ (sigmoid), ReLU

$$h = \sigma(x) = 1/(1 + \exp(-x))$$

$$h = ReLU(x) = \max(x, 0)$$

Neural networks

- A neural network is a stack of linear layers and activation function
  - Input layer: typically image, text, etc.
  - Output layer: our target (typically a class)
  - Hidden layers: intermediate representations learned by the neural net



Neural networks: analysis

- How can we understand what happens in the neural network ?

Neural networks: analysis

- How can we understand what happens in the neural network ?
  - We can compute the hidden layer activations for a bunch of inputs

Neural networks: analysis

- How can we understand what happens in the neural network ?
  - We can compute the hidden layer activations for a bunch of inputs
  - This gives us **vectors** in some dimension (e.g. 256, 512)

Neural networks: analysis

- How can we understand what happens in the neural network ?
  - We can compute the hidden layer activations for a bunch of inputs
  - This gives us **vectors** in some dimension (e.g. 256, 512)
  - This vectors are not easily interpretable: their dimension do not correspond to intuitive quantities like pixels, etc.

Neural networks: analysis

- How can we understand what happens in the neural network ?
  - We can compute the hidden layer activations for a bunch of inputs
  - This gives us **vectors** in some dimension (e.g. 256, 512)
  - This vectors are not easily interpretable: their dimension do not correspond to intuitive quantities like pixels, etc.
  - We cannot understand what each dimension does but we can understand what the vector represents **globally**

Understanding high dimensional data ?

1. Visualization: PCA (linear)
2. Visualization: t-SNE
3. Metrics (black-box approach)