Date:

Conceptual part

Covariance: used to identify the relationship between two features i.e positive trend, negative trend, no trend.

PCA : Principal component Analysis

① find origin of data
② draw a random line crossing through that origin
③ project all data points onto that line
④ find the line that maximizes the distances from the projected points to the origin OR find a line that minimizes the distances of each data point to the origin.
⑤ sum of squared distances, minimize/maximize that value accordingly.
⑥ After we have our principal component no. 1 or first eigen vector we find our second eigen vector. which is perpendicular to the pc1.
⑦ Total number of PCs depend on number of variables/features
⑧ After we have all our PCs we check which PC/s can accumulate most of variables, we choose them and plot the data on a decreased dimensioned plane.

Tips: ① Scale the data ② Centre the data ③

Date:

correlation and convolution on image:

* convolution combines two functions and shows how they convolve together. Let the first function be an image and the second be the guassian blur function if you combine both, convolutional filter would apply the second function i.e the guassian blur function on each pixel of the image passed as the first function. As an end result, we will receive a smoother picture.

* cross-correlation, is wally used for object detection or similarity detection. we give two functions, an image and another image which may be smaller. The cross-corr slides the smaller image accross the larger one and calculates on each slide how similar that part is to the small image i.e correlation.

Date:

Bag of Words:

① we make vectors of each sentence in a corpus using the 0s and 1s as unpresent and present word of a corpus in that specific sentence.

TF-IDF:

① we make a table of number of each term in each document

② use formula:

$$tfidf = tf \times \log\left(\frac{N}{df}\right)$$

- tf: term frequency
- N: number of documents
- df: number of documents containing that term.

N-grams:

→ set of words containing n elements.

→ calculate probability of a word with the history of the word (predict the next word)

→ calculate probability of a sentence

$$P(w|h) = \frac{\text{count}(w) + \text{count}(h)}{\text{count}(w)}$$

Date:

Data Refinement techniques:

we can clean/refine our data in following ways:

1. Remove duplicate tuples / Nulls / irrelevant
2. Fix structural errors (spellings, typos etc)
3. Filter unwanted outliers.
4. Handle missing data

Graph & when to use them:

1. Bar Chart: use when * ~~tabular data.~~ * more than 10 features in data
2. Line Chart: * show trends * relationships of features
3. Scatter plot: * show relationship between variables
4. Pie Chart: * show part-to-whole comparison * identify smallest + largest parts
5. Bubble Chart: * show relation between more than 2 features

Date:

* Convolutional Neural Networks: CNNs are often used for image classification, they can identify objects in images by learning and recognizing important features.
→ CNN can be used for image data, classification prediction problems and regression prediction problem.

* Multi-layer perceptrons: can be used for tabular data ~~text, image, time series, and other types of~~ data. It is the classical neural network with hidden layers.

* Recurrent Neural network: used for text, speech, classification, regression, generative models. They don't work on image or tabular data.

Types of Gradient descent:

* Stochastic: one example at a time
* mini-batch: apply GD with many examples at a time but still less than the total dataset of examples.
* batch: → if we input entire dataset at a time and update weight, that batch GD.

epoch → go through all examples once.

Date:

* Types of activation functions in neural networks:

→ An activation function decides whether a single neuron in a network should be activated / should participate or not.

→ The ultimate outcome of using activation functions is non-linearity in our decision boundary.

① Sigmoid : $f(x) = \frac{1}{1+e^{-x}}$ range $(0, 1)$

② Hyperbolic tangent / tanh function :

slowest

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$ output range $(-1, 1)$

③ Softmax : a combo of multiple sigmoids.

$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

$z_1, z_2, \cdots, z_n$.

$$f(z_1) = \frac{e^{z_1}}{e^{z_1} + e^{z_2} + \cdots + e^{z_n}}$$

$$f(z_i) = \frac{e^{z_i}}{e^{z_1} + e^{z_2} + \cdots + e^{z_n}}$$

Date:

Image Filtering:

① Lowpass filters (Smoothing)
→ removes high spatial frequency noise

② High pass filters (edge detection, sharpening)
→ make image sharp
→ emphasize details

* Mean filter / Average
→ used to smooth images, reduce noise

$$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

* ~~Median~~ Guassian filter
→ used to remove noise (especially ~~salt & pepper~~ guassian noise)

$$\begin{bmatrix} 2 & 1 & 2 \\ 1 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix} \quad \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

linear.

* Median filter
→ removes noise, salt & pepper.
→ arrange elements into order & choose middle element & replace.. non-linear