**Summary**

Machine learning is the science of getting computers to learn, without being explicitly programmed.

## Supervised

### Linear regression

Cost

Gradients

### Logistic regression

Cost

Gradients

### SVM

Cost

SVM with kernels, also called *gaussian kernel*

## Unsupervised

### K-means

For i=1 to m

denotes the index of cluster centroids closet to x(i)

denotes the average(mean) of points assigned to cluster j

A close up of a map

Description automatically generated

**Random initialization**

**For i = 1 to 100{**

**Initial different centroids**

initial\_centroids = random.sample(X.tolist(), K)

**cost function**

*Return centroids of the smallest J.*

**}**

### PCA

**Dimensionality Reduction**

1. **Data processing** 
   1. **Feature scaling (mean normalization)**

U, S, V = numpy.linalg.svd(sigma)

Ureduce = U[:, 0:K].T

Z = Ureduce\*X = X\_norm \* U[:, 0:K]

X\_approximate = X\_recovered = Z \* U[:, 0:K].T

**Choosing K:**

**Pick the smallest value of k,**

**99% of variance retained.**

### Anomaly detection

**Gaussian (Normal) distribution**

#### Original model

Density estimation:

Training set:

Each example is:

#### Multivariate Gaussian

**Comparison between these two model**

**F1-score**

****

**Precision**

**Recall**

**F1**

### Collaborative filtering

**Recommender system**

**Minimize**  and

**cost**

Gradients

For user j, on movie i predict:

A screenshot of a cell phone

Description automatically generatedA close up of a map

Description automatically generated

Plot matrix

plt.imshow(Y, aspect='equal', origin='upper', extent=(0, Y.shape[1], 0, Y.shape[0]/2.0))