# Batch:

Small size in total batch

SGD: one value at a time for one epoch

Mini Bach GD: small batch for one epoch

Batch GD: Complete batch pre epoch

# Gradient Decent:

Used to find the weights of modal parameters

Optimizer:

-It is impossible to calc weights for 1000 of neurons across 100 of layers. So, we use the optimizers

-Optimizer will calc gradients and update model perameters

ADAM – Adaptive Moment Estimation

# Preventing Overfitting:

## Regularization:

Penalize complex models

## Cross Validation:

Distinct training and validation phases

## Dropouts (NN only):

Intermediately turning off some neurons during training