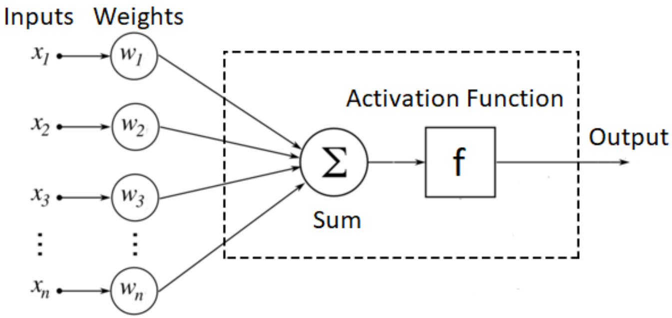
Lecture 2: Introduction to Neural Networks. Deep Cross-Entropy Method

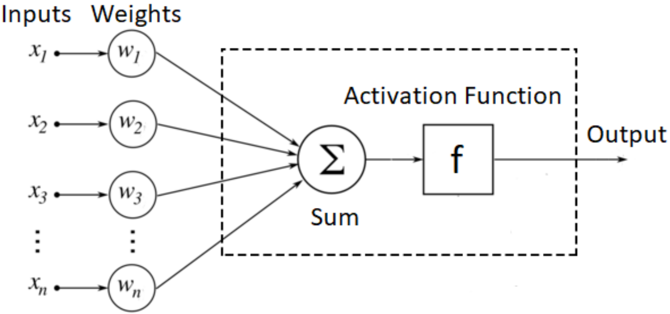
Anton Plaksin

Perceptron (Neuron)

Perceptron (Neuron)

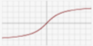


Perceptron (Neuron)

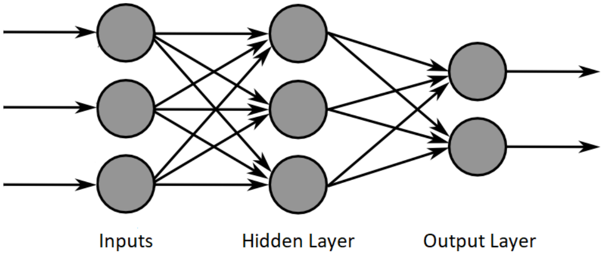


$$y = f\left(\sum_{i=1}^{n} w_i x_i\right) \quad \text{or} \quad y = f\left(b + \sum_{i=1}^{n} w_i x_i\right)$$

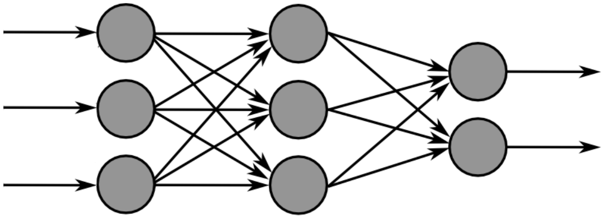
Activation Functions

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for} \quad x \neq 0 \\ ? & \text{for} \quad x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1-f(x))$
TanH		$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2+1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ x & \text{for} \quad x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for} \quad x < 0 \\ x & \text{for} \quad x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for} \quad x < 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for} \quad x < 0 \\ x & \text{for} \quad x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for} \quad x < 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1+e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$

Neuron Network



Neuron Network

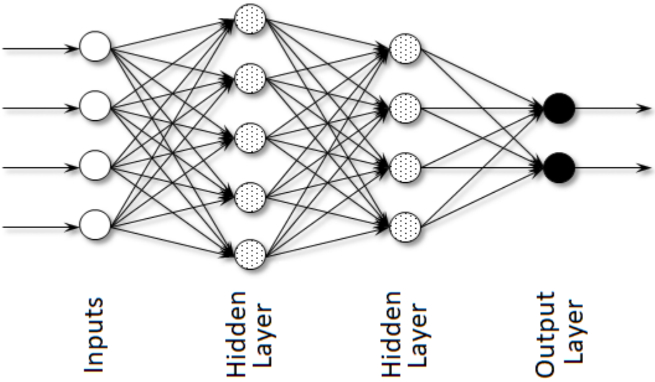


$$F_j^\theta(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

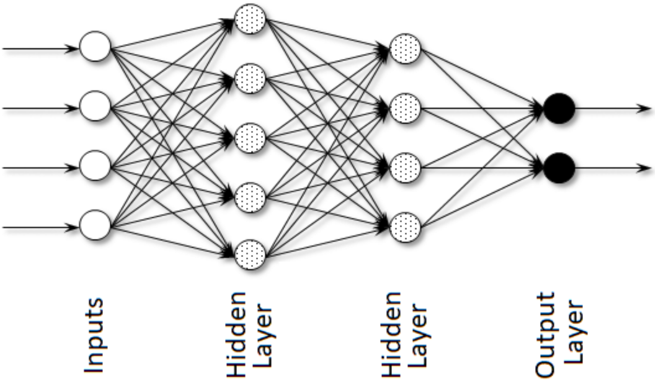
$$F^\theta(X) \in \mathbb{R}^2, \quad X = (x_1, x_2, x_3) \in \mathbb{R}^3,$$

$$\theta = (b_1, b_2, w_{1,1}, \dots, w_{2,3}, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{w}_{1,1}, \dots, \hat{w}_{3,3}) \in \mathbb{R}^{20}$$

Neuron Network



Neuron Network



$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{4} w_{j,k} f\left(\hat{b}_k + \sum_{l=1}^{5} \hat{w}_{k,l} f\left(\tilde{b}_l + \sum_{i=1}^{4} \tilde{w}_{l,i} x_i\right)\right)\right), \quad j \in \overline{1,2}.$$

$$F^{\theta}(X) \in \mathbb{R}^2, \quad X \in \mathbb{R}^4, \quad \theta \in \mathbb{R}^{59}$$

Regression

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\sum_{i=1}^{k} \|F(X_i) - Y_i\|^2 \to \min_{F \text{ is continuous}}$$

Regression

$$\{(X_1, Y_1), (X_2, Y_2), \dots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\sum_{i=1}^{k} \|F(X_i) - Y_i\|^2 \to \min_{F \text{ is continuous}}$$

Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0

Regression

$$\{(X_1,Y_1),(X_2,Y_2),\dots,(X_k,Y_k)\},\quad X_i\in\mathbb{R}^n,\quad Y_i\in\mathbb{R}^m$$

$$\Downarrow$$

$$\sum_{i=1}^{k}\|F(X_i)-Y_i\|^2\to\min_{F\text{ is continuous}}$$

Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0
- Define $Loss(\theta)=\frac{1}{k}\sum_{i=1}^{k}\|F^\theta(X_i)-Y_i\|^2$

Regression

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\sum_{i=1}^{k} \|F(X_i) - Y_i\|^2 \to \min_{F \text{ is continuous}}$$

Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0
- Define $Loss(\theta) = \frac{1}{k}\sum_{i=1}^{k} \|F^\theta(X_i) - Y_i\|^2$
- Solve the optimization problem $Loss(\theta) \to \min_{\theta}$

Regression

$$\{(X_1, Y_1), (X_2, Y_2), \dots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

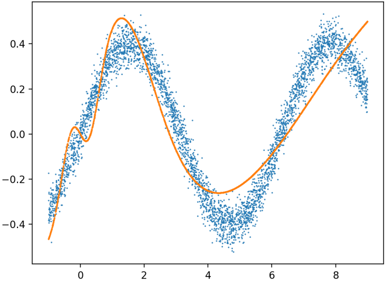
$$\Downarrow$$

$$\sum_{i=1}^{k} \|F(X_i) - Y_i\|^2 \to \min_{F \text{ is continuous}}$$

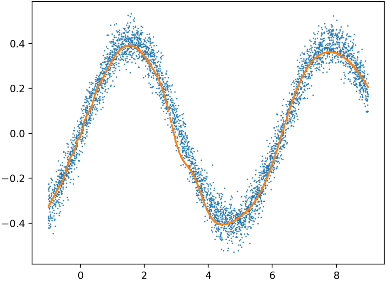
Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0
- Define $Loss(\theta) = \frac{1}{k} \sum_{i=1}^{k} \|F^\theta(X_i) - Y_i\|^2$
- Solve the optimization problem $Loss(\theta) \to \min_\theta$
  - Gradient Descent: $\theta_{j+1} = \theta_j - \eta \nabla_\theta Loss(\theta)$, $\eta > 0$

Example



128 neurons



$32 \times 128 \times 32$ neurons

Cybenko Theorem (1989)

Theorem

For every continuous function $G\colon [0,1]^n \mapsto \mathbb{R}^m$ and every $\varepsilon > 0$, there exist N, w_i, $\hat{b}_i$, $\hat{w}_{i,j}$, $i \in \overline{1,N}$, $j \in \overline{1,n}$ such that

$$\|F^\theta(X) - G(X)\| \le \varepsilon, \quad X \in [0,1]^n,$$

where

$$F^\theta(X) = \sum_{i=1}^{N} w_i \sigma\left(\hat{b}_i + \sum_{j=1}^{n} \hat{w}_{i,j} x_j\right),$$

$$X = (x_1, \ldots, x_n), \quad \theta = (w_1, \ldots w_N, \hat{b}_1, \ldots, \hat{b}_N, \hat{w}_{1,1}, \ldots, \hat{w}_{N,n})$$

Gradient Calculation

$$Loss(\theta) = \sum_{i=1}^{k} \|F^{\theta}(X_i) - Y_i\|^2,$$

$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

Gradient Calculation

$$Loss(\theta) = \sum_{i=1}^{k} \|F^{\theta}(X_i) - Y_i\|^2,$$

$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

$$\frac{\partial Loss(\theta)}{\partial \hat{w}_{2,3}} = ???$$

Gradient Calculation

$$Loss(\theta) = \sum_{i=1}^{k} \|F^{\theta}(X_i) - Y_i\|^2,$$

$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

$$\frac{\partial Loss(\theta)}{\partial \hat{w}_{2,3}} = 2\sum_{i=1}^{k} \langle F^{\theta}(X_i) - Y_i, \frac{\partial F^{\theta}(X_i)}{\partial \hat{w}_{2,3}} \rangle$$

Gradient Calculation

$$\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} \approx \frac{F_j^{(\ldots,\hat{w}_{2,3}+\Delta w,\ldots)}(X) - F_j^{\theta}(X)}{\Delta w}$$

Gradient Calculation

$$\cancel{\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} \approx \frac{F_j^{(\ldots,\hat{w}_{2,3}+\Delta w,\ldots)}(X) - F_j^{\theta}(X)}{\Delta w}}$$

Gradient Calculation

$$\cancel{\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} \approx \frac{F_j^{(\dots,\hat{w}_{2,3}+\Delta w,\dots)}(X) - F_j^{\theta}(X)}{\Delta w}}$$

$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

Gradient Calculation

$$\cancel{\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} \approx \frac{F_j^{(\dots,\hat{w}_{2,3}+\Delta w,\dots)}(X) - F_j^{\theta}(X)}{\Delta w}}$$

$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

$$\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} = ???$$

Gradient Calculation

$$\cancel{\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} \approx \frac{F_j^{(\dots,\hat{w}_{2,3}+\Delta w,\dots)}(X) - F_j^{\theta}(X)}{\Delta w}}$$

$$F_j^{\theta}(X) = f_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right), \quad j \in \overline{1,2}.$$

$$\begin{aligned}\frac{\partial F_j^{\theta}(X)}{\partial \hat{w}_{2,3}} \quad &= \quad f'_{out}\left(b_j + \sum_{k=1}^{3} w_{j,k} f\left(\hat{b}_k + \sum_{i=1}^{3} \hat{w}_{k,i} x_i\right)\right) \\ &\times \quad w_{j,2} f'\left(\hat{b}_2 + \sum_{i=1}^{3} \hat{w}_{2,i} x_i\right) x_3\end{aligned}$$

Stochastic Gradient Descent

Stochastic Gradient Descent

- Randomly choose $\{(X, Y)\}$

Stochastic Gradient Descent

- Randomly choose $\{(X, Y)\}$
- Define $Loss(\theta) = \|F^{\theta}(X) - Y\|^2$

Stochastic Gradient Descent

- Randomly choose $\{(X, Y)\}$
- Define $Loss(\theta) = \|F^{\theta}(X) - Y\|^2$
- Update parameters $\theta_{j+1} = \theta_j - \eta \nabla_{\theta} Loss(\theta)$

Classification

$$\big\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\big\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\big|\{i \in \overline{1,k} \mid F(X_i) \neq Y_i\}\big| \to \min_F$$

Classification

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\left|\{i \in \overline{1,k} \mid F(X_i) \neq Y_i\}\right| \to \min_F$$

$$S_i = \mathrm{Softmax}(Z)_i = \frac{e^{z_i}}{\sum\limits_{j=1}^{m} e^{z_j}}, \quad i \in \overline{1,m}, \quad Z = (z_1, \ldots, z_m)$$

Classification

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\left|\{i \in \overline{1,k} \mid F(X_i) \neq Y_i\}\right| \to \min_F$$

$$S_i = \mathrm{Softmax}(Z)_i = \frac{e^{z_i}}{\sum\limits_{j=1}^{m} e^{z_j}}, \quad i \in \overline{1,m}, \quad Z = (z_1, \ldots, z_m)$$

$$S_i \in (0,1),\ i \in \overline{1,m}, \quad \sum_{i=1}^{m} S_i = 1.$$

Classification

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\left|\{i \in \overline{1,k} \mid F(X_i) \neq Y_i\}\right| \to \min_F$$

$$S_i = \mathrm{Softmax}(Z)_i = \frac{e^{z_i}}{\sum\limits_{j=1}^{m} e^{z_j}}, \quad i \in \overline{1,m}, \quad Z = (z_1, \ldots, z_m)$$

Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0

Classification

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$
$$\Downarrow$$
$$\left|\{i \in \overline{1,k} \mid F(X_i) \neq Y_i\}\right| \to \min_F$$

$$S_i = \mathrm{Softmax}(Z)_i = \frac{e^{z_i}}{\sum\limits_{j=1}^{m} e^{z_j}}, \quad i \in \overline{1,m}, \quad Z = (z_1, \ldots, z_m)$$

Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0
- Define $Loss(\theta) = -\frac{1}{k}\sum\limits_{i=1}^{k} \ln \mathrm{Softmax}(F^\theta(X_i))_{Y_i}$

Classification

$$\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_k, Y_k)\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

$$\left|\{i \in \overline{1,k} \mid F(X_i) \neq Y_i\}\right| \to \min_F$$

$$S_i = \text{Softmax}(Z)_i = \frac{e^{z_i}}{\sum\limits_{j=1}^{m} e^{z_j}}, \quad i \in \overline{1,m}, \quad Z = (z_1, \ldots, z_m)$$

Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0
- Define $Loss(\theta) = -\frac{1}{k} \sum\limits_{i=1}^{k} \ln \text{Softmax}(F^\theta(X_i))_{Y_i}$
- Solve the optimization problem $Loss(\theta) \mapsto \min\limits_\theta$

Classification

$$\big\{(X_1, Y_1), (X_2, Y_2), \dots, (X_k, Y_k)\big\}, \quad X_i \in \mathbb{R}^n, \quad Y_i \in \mathbb{R}^m$$

$$\Downarrow$$

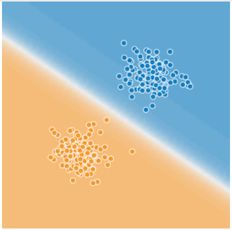
$$\big|\{i \in \overline{1,k} \,|\, F(X_i) \neq Y_i\}\big| \to \min_F$$

$$S_i = \mathrm{Softmax}(Z)_i = \frac{e^{z_i}}{\sum\limits_{j=1}^{m} e^{z_j}}, \quad i \in \overline{1,m}, \quad Z = (z_1, \dots, z_m)$$

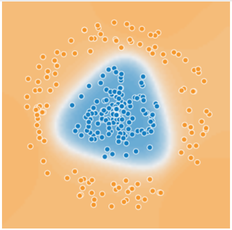
Neural Network Approach

- Set a NN structure F^θ and initial parameters θ_0
- Define $Loss(\theta) = -\frac{1}{k}\sum\limits_{i=1}^{k} \ln \mathrm{Softmax}(F^\theta(X_i))_{Y_i}$
- Solve the optimization problem $Loss(\theta) \mapsto \min\limits_\theta$
  - Gradient Descent: $\theta_{j+1} = \theta_j - \eta \nabla_\theta Loss(\theta)$

Examples

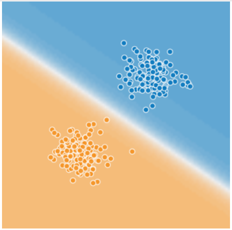


1 layer $\times$ 1 neuron

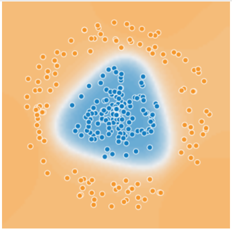


1 layer $\times$ 3 neuron

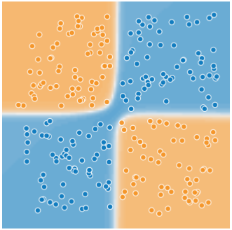
Examples



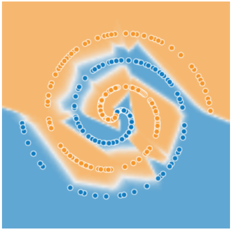
1 layer × 1 neuron



1 layer × 3 neuron

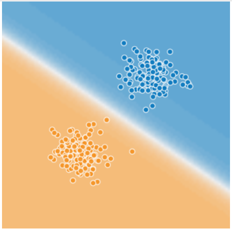


2 layer × 3 neuron

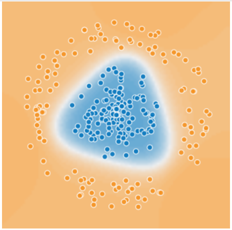


4 layer × 8 neuron

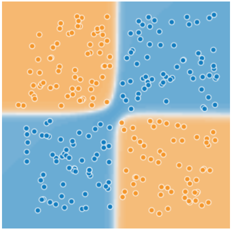
Examples



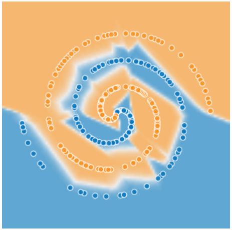
1 layer × 1 neuron



1 layer × 3 neuron



2 layer × 3 neuron



4 layer × 8 neuron

https://playground.tensorflow.org

Neural Network Elements

Neural Network Elements

- Structure (number of layers and neurons, activation functions)

Neural Network Elements

- Structure (number of layers and neurons, activation functions)
- Loss Functions (MSE, L1, Cross Entropy, ...)

Neural Network Elements

- Structure (number of layers and neurons, activation functions)
- Loss Functions (MSE, L1, Cross Entropy, ...)
- Gradient Descent Methods (SGD, Adam, Adadelta, ...)

Neural Network Elements

- Structure (number of layers and neurons, activation functions)
- Loss Functions (MSE, L1, Cross Entropy, ...)
- Gradient Descent Methods (SGD, Adam, Adadelta, ...)
- Weight initialization

Neural Network Elements

- Structure (number of layers and neurons, activation functions)
- Loss Functions (MSE, L1, Cross Entropy, ...)
- Gradient Descent Methods (SGD, Adam, Adadelta, ...)
- Weight initialization
- Regularization (L2, L1, Batch normalization, Dropout, ...)

Neural Network Elements

- Structure (number of layers and neurons, activation functions)
- Loss Functions (MSE, L1, Cross Entropy, ...)
- Gradient Descent Methods (SGD, Adam, Adadelta, ...)
- Weight initialization
- Regularization (L2, L1, Batch normalization, Dropout, ...)

С. Николенко, Е. Архангельская «Глубокое обучение. Погружение в мир нейронных сетей»

Markov Decision Process

Markov Property

$$\mathbb{P}[S_{t+1}|S_t, A_t] = \mathbb{P}[S_{t+1}|S_1, A_1, S_2, A_2 \ldots, S_t, A_t]$$

$$\mathbb{P}[R_t|S_t, A_t] = \mathbb{P}[R_t|S_1, A_1, S_2, A_2 \ldots, S_t, A_t] = 1$$

Markov Decision Process

Markov Property

$$\mathbb{P}[S_{t+1}|S_t, A_t] = \mathbb{P}[S_{t+1}|S_1, A_1, S_2, A_2 \dots, S_t, A_t]$$

$$\mathbb{P}[R_t|S_t, A_t] = \mathbb{P}[R_t|S_1, A_1, S_2, A_2 \dots, S_t, A_t] = 1$$

Markov Decision Process $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$

- $\mathcal{S}$ — a ~~finite ($|\mathcal{S}| = n$)~~ state space
- $\mathcal{A}$ — a ~~finite ($|\mathcal{A}| = m$)~~ action space
- $\mathcal{P}$ — a deterministic transition probability function

$$\mathcal{P}(s'|s, a) = \mathbb{P}[S_{t+1} = s'|S_t = s, A_t = a]$$

- $\mathcal{P}_0$ — a deterministic initial state function
- $\mathcal{R}$ — a reward function

$$\mathcal{R}(s, a) = R_t \quad \Leftrightarrow \quad \mathbb{P}[R_t|S_t = s, A_t = a] = 1$$

- $\gamma \in [0, 1]$ — discount coefficient

Cross-Entropy Method. Case of $\mathcal{S} \subset \mathbb{R}^n$ и $\mathcal{A} \subset \mathbb{R}^m$

Let $\pi^\theta \colon \mathbb{R}^n \mapsto \mathbb{R}^m$ be a neural network, θ_0 be initial parameters,
N be a number of iterations, K be a number of trajectories,
$q \in (0,1)$ be a parameter for defning «elite» trajectories,
$\eta > 0$ be a learning rate, $\varepsilon = 1$ be an exploration parameters.
For each $n \in \overline{1,N}$, do

Let $\pi^\theta \colon \mathbb{R}^n \mapsto \mathbb{R}^m$ be a neural network, θ_0 be initial parameters, N be a number of iterations, K be a number of trajectories, $q \in (0,1)$ be a parameter for defning «elite» trajectories, $\eta > 0$ be a learning rate, $\varepsilon = 1$ be an exploration parameters.
For each $n \in \overline{1,N}$, do

- (Policy evaluation) Acting in accordance with the policy

$$\pi_n(s) = \left[\pi^{\theta_n}(s) + Noise(\varepsilon)\right]_{\mathcal{A}},$$

get K trajectories τ_k and total rewards $G(\tau_k)$. Evaluate π_n:

$$\mathbb{E}_{\pi_n}[G] \approx V_n := \frac{1}{K}\sum_{k=1}^{K} G(\tau_k)$$

Cross-Entropy Method. Case of $\mathcal{S} \subset \mathbb{R}^n$ и $\mathcal{A} \subset \mathbb{R}^m$

Let $\pi^\theta \colon \mathbb{R}^n \mapsto \mathbb{R}^m$ be a neural network, θ_0 be initial parameters, N be a number of iterations, K be a number of trajectories, $q \in (0,1)$ be a parameter for defning «elite» trajectories, $\eta > 0$ be a learning rate, $\varepsilon = 1$ be an exploration parameters.
For each $n \in \overline{1, N}$, do

- (Policy improvement) Select «elite» trajectories $\mathcal{T}_n = \{\tau_k, k \in \overline{1,K} \colon G(\tau_k) > \gamma_q\}$, where γ_q is a q-quantile of the numbers $G(\tau_k)$, $k \in \overline{1,K}$. Define

$$Loss(\theta) = \frac{1}{|\mathcal{T}_n|} \sum_{(a|s) \text{«} \in \text{»} \mathcal{T}_n} \|\pi^{\theta_n}(s) - a\|^2$$

  and update parameters by the (stochastic) gradient descent:

$$\theta_{n+1} = \theta_n - \eta \nabla_\theta Loss(\theta_n).$$

  Reduce ε ($\varepsilon = 1/N$).

Cross-Entropy Method. Case of $\mathcal{S} \subset \mathbb{R}^n$ и $|\mathcal{A}| = m$

Let $F^\theta \colon \mathbb{R}^n \mapsto \mathbb{R}^m$ be a neural network. Define

$$\pi^\theta(i|s) = \mathrm{Softmax}(F^\theta(s))_i, \quad \pi_{\text{uniform}}(i|s) = 1/m, \quad i \in \overline{1,m}.$$

Let θ_0 be initial parameters, N be a number of iterations,
K be a number of trajectories, $q \in (0,1)$ be a parameter for defning «elite» trajectories, $\eta > 0$ be a learning rate,
$\varepsilon = 1$ be an exploration parameters.
For each $n \in \overline{1,N}$, do

Cross-Entropy Method. Case of $\mathcal{S} \subset \mathbb{R}^n$ и $|\mathcal{A}| = m$

Let $F^\theta \colon \mathbb{R}^n \mapsto \mathbb{R}^m$ be a neural network. Define

$$\pi^\theta(i|s) = \mathrm{Softmax}(F^\theta(s))_i, \quad \pi_{\text{uniform}}(i|s) = 1/m, \quad i \in \overline{1,m}.$$

Let θ_0 be initial parameters, N be a number of iterations,
K be a number of trajectories, $q \in (0,1)$ be a parameter for defning «elite» trajectories, $\eta > 0$ be a learning rate,
$\varepsilon = 1$ be an exploration parameters.
For each $n \in \overline{1,N}$, do

- (Policy evaluation) Acting in accordance with the policy

$$\pi_n(\cdot|s) = (1-\varepsilon)\pi^{\theta_n}(\cdot|s) + \varepsilon\pi_{\text{uniform}}(\cdot|s),$$

  get K trajectories τ_k and total rewards $G(\tau_k)$. Evaluate π_n:

$$\mathbb{E}_{\pi_n}[G] \approx V_n := \frac{1}{K}\sum_{k=1}^{K} G(\tau_k)$$

Cross-Entropy Method. Case of $\mathcal{S} \subset \mathbb{R}^n$ и $|\mathcal{A}| = m$

Let $F^\theta : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a neural network. Define

$$\pi^\theta(i|s) = \text{Softmax}(F^\theta(s))_i, \quad \pi_{\text{uniform}}(i|s) = 1/m, \quad i \in \overline{1,m}.$$

Let θ_0 be initial parameters, N be a number of iterations,
K be a number of trajectories, $q \in (0,1)$ be a parameter for defning «elite» trajectories, $\eta > 0$ be a learning rate,
$\varepsilon = 1$ be an exploration parameters.
For each $n \in \overline{1,N}$, do

- (Policy improvement) Select «elite» trajectories $\mathcal{T}_n = \{\tau_k, k \in \overline{1,K} : G(\tau_k) > \gamma_q\}$, where γ_q is a q-quantile of the numbers $G(\tau_k)$, $k \in \overline{1,K}$. Define

$$Loss(\theta) = -\frac{1}{|\mathcal{T}_n|} \sum_{(a|s) \text{«}\in\text{»} \mathcal{T}_n} \ln \pi^\theta(a|s)$$

  and update parameters by the (stochastic) gradient descent:

$$\theta_{n+1} = \theta_n - \eta \nabla_\theta Loss(\theta_n)$$

  Reduce ε ($\varepsilon = 1/N$).

QUESTIONS?