

1 Section1_入力層~中間層

(1) 問題 1

以下の文章を読み、次の問題に答えよ。

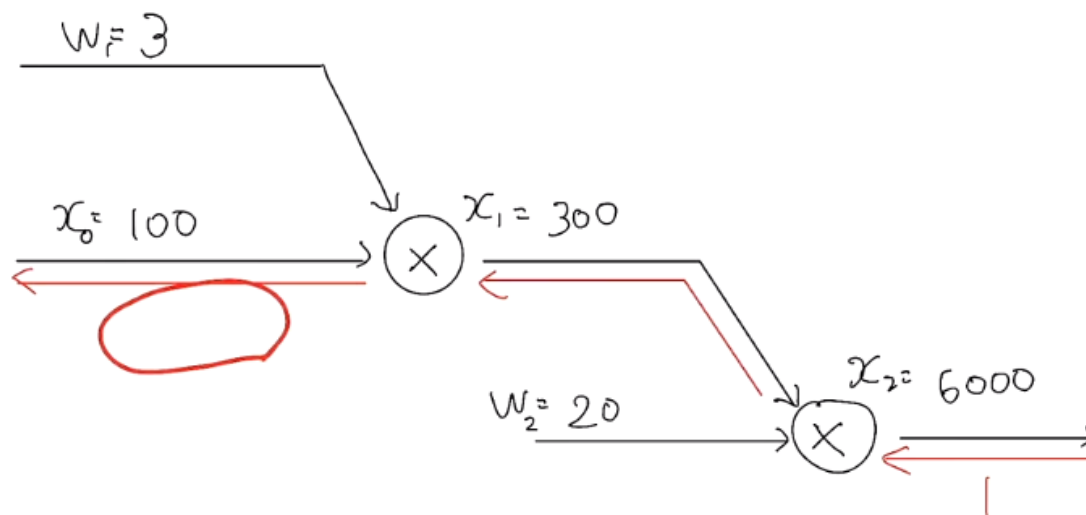
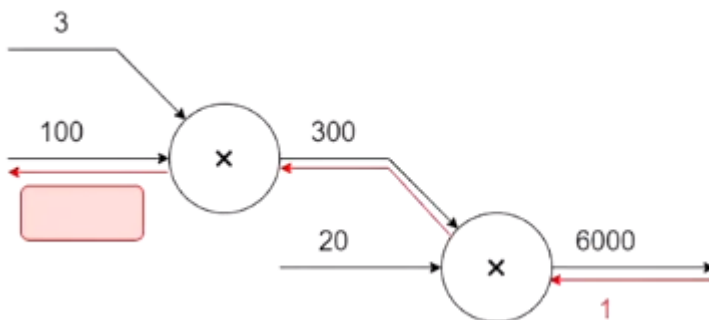
機械学習では主にニューラルネットワークにおいて、出力層から入力層に向かって順番に計算していく過程を「逆伝播」といい、パラメータの更新に必要な重要な過程である。

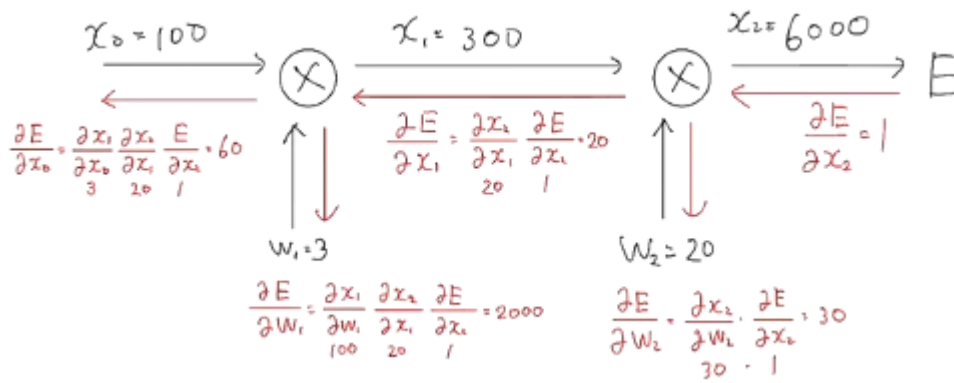
問題

以下の【図1】のような掛け算ユニットがある計算グラフで逆伝播を行ったとき、赤枠に当てはまる数字はどれか。次の選択肢の中から選びなさい。

なお、計算グラフは、計算の過程をグラフによって表したものであり、 \times は乗算を、 $+$ は加算を表している。また、赤い矢印は逆伝播による微分値の伝達を表している。

1. 60
2. 600
3. 100
4. 30000





$$\delta x_2 / \delta x_1 = 6000 / 300 = 20$$

$$\delta x_1 / \delta x_0 = 300 / 100 = 3$$

$$3 \cdot 20 \cdot 1 = 60$$

答) 60

(2) 問題2 ベイズの定理

以下の文章を読み、次の問題に答えよ。

クラス分類とは、与えられた入力空間を相異なる K 個の空間に分割し、それぞれの空間をクラス C_k とラベルをつけることである。訓練データ x が与えられたときに、推論段階と決定段階を経て各クラスに割り当てる。訓練データからどのようにして分類するかによって、生成モデル・識別モデル・識別関数に分けられる。

問題

生成モデルではあるデータ x について、各 C_i における x となる確率、すなわち $p(x|C_i)$ と C_i が出現する確率 $p(C_i)$ を求めることで、 x がある C_i に分類される確率を求める。ある x_i について3つのクラスに分けたい、 $p(x_i|C_0) = 0.1, p(x_i|C_1) = 0.3, p(x_i|C_2) = 0.2, p(C_0) = 0.7, p(C_1) = 0.2, p(C_2) = 0.1$ であるとき x_i が C_1 に属する確率を求めよ。

1. 0.2
2. 0.3
3. 0.4
4. 0.5

事後確率はベイズの定理より $p(C_1|x) = \frac{p(x|C_1)p(C_1)}{p(x)}$ で求めることができる。また $p(x)$ は $p(x) = \sum p(x|C_k)p(C_k)$ である。問題文の条件から求める値は $\frac{0.3 \times 0.2}{0.1 \times 0.7 + 0.3 \times 0.2 + 0.2 \times 0.1} = 0.4$ となる。

(3) 問題3

問題

識別モデルは直接、事後確率 $p(C_k|x)$ を決める推論問題を解く。球の入った箱があり、球の色は白か赤、大きさは大きいものと小さいものがある。球の数の分布が次のような表のようになっていることが分かっている。

このとき $p(C_k = \text{白} | x = \text{大})$ を求めよ。ただし、 C_k は白か赤の二値、 x は”大”と”小”の二値しかとらないものとする。

	大	小
白	12個	28個
赤	4個	20個

1. 0.25
2. 0.3
3. 0.7
4. 0.75

$$12/16=0.75$$

(4) 問題4

以下の文章を読み、次の問題に答えよ。

未知のデータに対する性能（汎化性能）を評価することで、モデルの性能を正しく検証することができる。

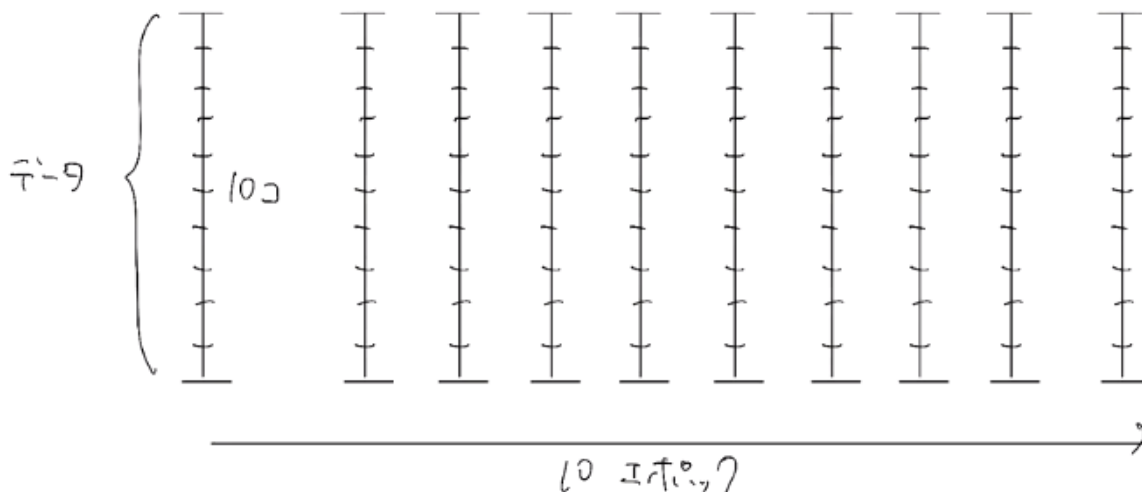
問題

k-分割検証法でモデルの性能を検証したい。「データをk個に分割し、分割したk個の各データセットを学習に回す（k回学習を繰り返す）」という手順を1回行うことを1epochとする。k=10, epoch=10のとき、評価の計算を行う回数は何回か。

次の選択肢の中から1つ選びなさい。

ただし、1回の学習で1回の評価を行うものとする。

1. 10
2. 100
3. 9
4. 90



$$10 \times 10 = 100$$

(5) 問題5

以下の文章を読み、次の問題に答えよ。

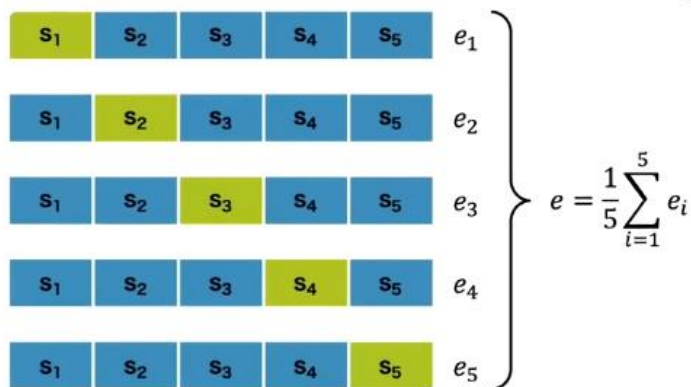
未知のデータに対する性能（汎化性能）を評価することで、モデルの性能を正しく検証することができる。

問題

検証法は様々な種類があるが、k-分割検証法の利点は何か。
適当なものを選択肢の中から1つ選びなさい。

1. 計算時間が比較的短い。
2. 他の手法と比べ、とても優れた性能を発揮するモデルの学習を行うのに有効な手法である。
3. サイズの小さいデータセットを活用して、繰り返し学習と評価を行うことで、過学習を防ぐ。
4. データセットのサイズが大きいとき、効率よく学習を行う。

問題10 k分割検証法 下記では1つのモデルを5分割して
5回学習と検証を繰り返す



解答 3

(6) 問題6

以下の文章を読み、次の問題に答えよ。

確率は身の回りの様々な場所で用いられている。例えば、自動運転の分野では自分の位置を推定するために用いられている。確率の計算を行うにあたって、基になっているものにベイズの定理がある。

問題

ベイズの定理を構成する要素として誤っているものはどれか。選択肢の中から1つ選びなさい。なお、 A は事象 a の確率変数を、 B は事象 b の確率変数を表す。

1. $p(A)$ は事象 a が起こる確率を表す。
2. $p(B)$ は事象 b が起こる確率を表す。
3. $p(A | B)$ は事象 a が起こるという条件のもとで事象 b が起こる確率を表す。
4. $p(B | A)$ は事象 a が起こるという条件のもとで事象 b が起こる確率を表す。

解答 3

2 Section2_活性化関数

ニューラルネットワークでは、ネットワークを構成する各ニューロンの出力を決定するために、ニューロンの入力を非線形な関数に通す。この関数のことを活性化関数と呼ぶ。

(1) 非線形な関数を通さないと、ニューラルネットワークは線形回帰と同様のモデルとなる。

(2) シグモイド関数

①公式

$$F(x)=1/(1+\exp(-x))$$

②値の範囲

$$0 \sim 1$$

(3) 双曲線正接関数

①公式

$$F(x)=\tanh(x)$$

②値の範囲

$$-1 \sim 1$$

3 Section3_出力層

学習時におけるニューラルネットワークの出力の良し悪しを評価する指標に、損失関数がある。
回帰問題では二乗和誤差、分類問題では交差エントロピー誤差がある。
出力層のノードの数を K とし、 k 番目のノードの出力及び正解の値をそれぞれ y_k 、 t_k とする。

(1) 二乗和誤差

$$\frac{1}{2} \sum (y_k - t_k)^2$$

予測値と正解値の差を二乗することにより、差の値を正に変換し、誤差として機能させている。

$$\delta L / \delta y_k = y_k - t_k$$

(2) 交差エントロピー誤差

$$-\sum t_k \log(y_k)$$

$$\delta L / \delta y_k = t_k / y_k$$

4 Section4_勾配降下法

ニューラルネットワークにおける代表的なパラメータ探索方法が、勾配降下法である。

目的関数を各パラメータで偏微分した値が必要となるが、この偏微分値を効率よく求める方法がバックプロパゲーション（誤差逆伝播法）である。

勾配降下法とは、ある関数を最小化することを目的とするアルゴリズムの1つである。

深層ニューラルネットワークの学習においては、誤差を最小化する重みを求めたいが、一般に重みは次元が高く、且つ複雑であるため、局所最適解や停留点にトラップされてしまう問題がある。

重みの更新タイミングによる分類

（1）逐次学習

訓練データ 1 つに対して 1 回更新する。

- ・ 確率的勾配降下法

（2）ミニバッチ学習

いくつかの訓練データからランダムにサンプリングミニバッチデータすべてに対して更新する。

- ・ ミニバッチ勾配降下法

（3）バッチ学習

訓練データ全ての誤差を計算し、重みを 1 回更新する。

- ・ 勾配降下法

5 Section5_誤差逆伝播法

深層ニューラルネットワークの学習は、教師データを用いて以下の流れで行う。

- (1) 教師データを用いて予測を行い、その予測値と正解ラベルを比較して誤差を計算する。
まず教師データを用いて予測計算を行い、その際の予測値と正解ラベルと比較して誤差を計算する。
この際は入力から出力へ伝わり、順伝播となる。
- (2) 勾配降下法に基づいて各枝の重みを更新する。
これにより誤差を最小化することを目指す。
累計された誤差の値が小さくなるように、勾配降下法を用いて各枝の重みを更新する。
枝の重みは出力から入力の方へ順番に更新される。これを誤差逆伝播法という。

※予測値は順伝播

※重みの更新は逆伝播