

深層学習 day3 確認テスト

確認テスト 1-1

「サイズ 5×5 の入力画像を、サイズ 3×3 のフィルタで畳み込んだ時の出力画像のサイズを答えよ。なおストライドは 2, パディングは 1 とする。」
3×3

確認テスト 1-2

「RNN のネットワークには大きく分けて 3 つの重みがある。...」
前回の中間層の出力を現在の中間層に加える際にかけられる重み

確認テスト 1-3(演習チャレンジ)

「以下は再帰型ニューラルネットワークにおいて...」

(1)

↑

(正解) (2)

前後の単語の特徴を保持したまま結合したい。
concatenate() 関数で、特徴量をそのまま結合する。
重みを掛け合わせることで、結合した特徴量の数を調整する。

確認テスト 1-4

「連鎖律の原理を使い、dz/dx を求めよ。」

$$z = t^2$$

$$t = x + y$$

$$dz/dx = dz/dt * dt/dx$$

$$= 2t * 1$$

$$= 2(x + y)$$

確認テスト 1-5

「下図の y1 を x, z0, z1, W(in), W(out) を用いて数式で表せ。」

$$z1 = g(xW(in) + z0W + b)$$

$$y1 = g(z1W(out) + c)$$

※b, c はバイアス

確認テスト 1-6(コード演習問題)

「左の図は BPTT を行うプログラムである。...」

(3)

↑

(正解) (2)

これまでの説明の z の部分に当たる。

確認テスト 2-1

「シグモイド関数を微分したとき、入力値が0の時に最大値をとる。
その値として正しいものを選択肢から選べ。」

(2)

確認テスト 2-2(コード演習問題)

「RNNや深いモデルでは勾配の消失または爆発が起こる傾向がある。勾配爆発を防ぐために

勾配のクリッピングを行うという手法がある。具体的には勾配のノルムが閾値を超えたら、

勾配のノルムを閾値に正規化するというものである。以下は勾配のクリッピングを行う関数である。

(さ) に当てはまるのはどれか。」

(3)

↑

(正解) (1)

勾配のノルムを閾値に成果する際は、
閾値/勾配のノルム の値で正規化する。

確認テスト 2-3

「以下の文章を...どのゲートが作用すると考えられるか。」

「映画おもしろかったね。ところで、とてもお腹が空いたから何か____。」

忘却ゲート

確認テスト 2-4

「以下のプログラムはLSTMの順伝播を行うプログラムである。...」

(3)

確認テスト 3-1

「LSTMとCECが抱える課題について、それぞれ簡潔に述べよ。」

LSTM : パラメータ数が多く、計算負荷が高くなる。

CEC :

↑

(正解)

CEC : 勾配が1となり、学習能力がない

確認テスト 3-2

「GRU(Gated Recurrent Unit) もLSTMと同様にRNNの一種であり、単純なRNNにおいて問題となる

勾配消失問題を解決し、長期的な依存関係を学習することができる。LSTMに比べ変数の数やゲートの数が

少なく、より単純なモデルであるが、タスクによってはLSTMより良い性能を発揮する。

以下のプログラムはGRUの順伝播を行うプログラムである。…」

(4)

確認テスト 3-3

「LSTMとGRUの違いを簡潔に述べよ。」

- ・ LSTMでは記憶を保持するCECと入力ゲート、忘却ゲート、出力ゲートを持つ。
GRUではリセットゲート、更新ゲートを持つ。
GRUの方がパラメータ数が少なく、計算効率が高い。

確認テスト 3-4(演習チャレンジ)

「以下は双方向RNNの順伝播を行うプログラムである。…」

(3)

↑

(正解) (4)

時間的に同じ場所のものは同じ配列の場所に配置する。

確認テスト 5-1

「下記の選択肢から、seq2seqについて説明しているものを選び。」

(2)

確認テスト 5-2(演習チャレンジ)

「機械翻訳タスクにおいて、入力は複数の単語からなる文（文章）であり、…」

(1)

確認テスト 5-3

「VAEに関する下記の説明文中の空欄に当てはまる言葉を答えよ。」

確率分布

確認テスト 5-4

「seq2seqとHRED, HREDとVHREDの違いを簡潔に述べよ。」

seq2seqでは一問一答形式でしか入出力を行えないが、
HREDでは一文一文に対してcontext RNNを適用することで文脈を反映できる。

HREDでは確率的な多様性に乏しいことから会話の流れが乏しくなることがあるが、
VHREDはHREDにVAEを適用することで確率的な多様性を付与できる。

確認テスト 6-1

「RNNとword2vec、seq2seqとAttentionの違いを簡潔に述べよ」

RNNでは入力データとして単語の特徴ベクトル表現を使用する。
単語を特徴ベクトル表現に変換する方法の一つに word2vec がある。

seq2seq は固定長のベクトルに変換するため長い文章への対応が難しいが、
Attention は単語間の関連度を学習させる形式をとり長い文章への対応を可能にしている。