

構造化ドロップアウトによる追加学習

小澤 優太^{*1} 甲野 佑^{*1} 高橋 達二^{*2}

^{*1}東京電機大学 大学院 ^{*2}東京電機大学

1. はじめに

ニューラルネットワーク (NN) は画像認識, 音声認識や自然言語処理など幅広い分野で活用されている強力な計算モデルであり, 近年では, 深い階層構造で高い表現能力を実現する深層学習が注目されている. しかしながら, 学習中に新しいデータを追加すると既存の学習結果が失われてしまう, という追加学習の問題は解決していない. 既存の追加学習の実現方法は二つに大別される. 一つは, 一度学習した結合荷重とは異なる結合荷重を用いて学習を行うことで, 既存の学習結果を上書きすることなく, 新たな訓練データ集合を追加する方法である. もう一つは, 一度学習した訓練データとともに, 新たな訓練データを追加して再度学習を行うものである. どちらの方法もネットワークとは別の機構を必要としているため, ネットワークそのものの追加学習を行えているとは言えない. そこで本研究では, 深層学習においてより高い汎化能力を実現したドロップアウトと呼ばれる手法に着目し, 従来ランダムに行われてきたドロップアウトの構造化により結合荷重を一時的に保持する, 外部の機構を必要としない追加学習法を導入, 検証する.

2. ドロップアウト

ドロップアウトは NN のユニットを脱落させることで汎化性の改善を目的とした手法であり [Nitish 14], ネットワークの訓練の際に脱落ユニットを抜いたネットワークを構成する. ここで脱落ユニットはある確率 (多くの場合 $p = 0.5$) でランダムに選択されるが, 結合荷重の修正はユニットが削除されたネットワークに対して行われるため, 結局小規模化したネットワークを抽出し, 全ての重みパラメータの集合から複数のネットワーク部分集合に対する訓練を行っていることになる. ネットワーク出力は次のように計算される.

$$b_i = \text{Bernoulli}(p) \quad (1)$$

$$\tilde{h} = h * b \quad (2)$$

$$y = \tilde{h} * w^{(2)} \quad (3)$$

ここで p は各変数が 1 を取る確率, b は p に従ったベルヌーイ分布の変数から得られたベクトル, \tilde{h} はドロップアウトが適用された中間層の出力である. また, 中間層の 0 番目のユニットであるバイアス項は脱落させないものとする. 学習の際には, 学習の度にドロップアウトするユニットを判断するベクトル b を生成し, それを用いて小規模化されたネットワークに対して誤差逆伝播法を適用する. 学習したネットワークから予測する結果を得る際にはベクトル b は用いず, 全ての中間層を使用かつ重みを p 倍して出力を得る.

3. 提案手法

本研究の提案手法は, ある訓練データ集合による学習の後に別の訓練データ集合によって学習が行われる際, ネットワークの内部構造のみで追加学習を行う事を目的とする. 同様の試みとしては太田らが, 側抑制あるいは前シナプス抑制に基づくネットワークモデルを提案している [太田 06]. ここでは深層学習において過学習を回避し汎化を実現しているドロップアウトに着目した. 従来ランダムに行われるドロップアウトを構造化し, それによる抑制 (ユニットの脱落) の追加学習への影響を検証する. ここではネットワークを構成する各中間ユニットの学習度合いを意味する価値付けを行い, ユニットの学習状況に応じた抑制を行う. この学習度を参照してどのように抑制を行うかは様々な方法が考えられるが, 本稿では近年バンディット問題などに適用されている, ある内的な基準によって価値付けを変化させる満足化と呼ばれる概念による満足化価値関数 RS モデルをニューラルネットワーク上で用いて検証を行う.

ユニットの学習度合いの指標として, ネットワーク出力と教師信号のユークリッド距離を用いる. 中間層の j 番目のニューロンの価値を V_j とした時, 次式から V_j を計算する.

$$V_j(t+1) = V_j(t) + \alpha \left(\frac{1}{E} * s_j - V_j(t) \right) \quad (4)$$

$$s_j = \sum_k |\delta_k^{(2)} h_j| \quad (5)$$

ここで E はネットワークの出力と教師信号とのユークリッド距離である. 価値の初期値は $[0, 1]$ における一様な乱数とする. 学習中に, ユークリッド距離を用いたニューロンの価値から, ニューロンの価値 V_i が高いほど学習できているとして各状態でのニューロンの学習状況を推定する. ネットワークの出力と教師信号から得られるユークリッド距離はスカラー量であるため, そのままでは E 値が全ての h_j に対して同じになってしまう. そこで中間層の各ユニット h_j が E を出力した際の重みとして, 結合荷重の修正量の変化量の総和を取った値を乗算して h_j ごとの価値を決定する. RS モデルの参照点は満足化の振る舞いを左右する値である. もしニューロンの価値が低かった場合であっても, 参照点 R を下回る場合は試行回数 N_i が少ないニューロンほど RS 値が高くなる. 従って, 参照点 R を定めることで一定以上の学習ができていると判断できる上位のニューロンと, 一定以上の学習ができていないと判断することができる下位のニューロンが分割されることになる. よって, ネットワークの中間層におけるニューロンの価値を評価する際に, 指定した上位 $x\%$ を決定づけるような値を参照点 R とすることで, 学習できていると判断できるニューロンの $x\%$ を区分することが可能になる. RS の値は参照点 R とニューロンの価値 V_i を用いて, 次式で定義される.

$$RS(V_i) = N_i(V_i - R) \quad (6)$$

^{*}Incremental Learning by Structured Dropout”, Ozawa Yuta and Kohno Yu: Graduate school of Tokyo Denki University, Takahashi Tatsuji: Tokyo Denki University

このようなRS値の使用により、学習ができていると推測されるニューロンの中から使用回数が多いものを優先的に学習させることで分散的な学習を狭め、学習を偏らせることができると考えられる。これはRSの値を降順にした際、上位のユニットを優先的に抑制することで実現できる。これは中間ユニットにおけるRS値の下位50%のドロップアウトすることと同じである。RS-Dropoutモデルの適用にあたり、参照点Rの値を定める必要があるが、この値は中間層のユニットを脱落させる振る舞いを左右させる値であるため、上位x%のユニットを決定づけるように割合を変更させ、中間層のニューロンと学習能力の検証を行う。

4. シミュレーション

評価として、手書き文字認識の実験で広く用いられているMNIST Datasetを用いて手書き数字の追加学習を検証する。ここでは手書き数字の10種類のうち、4のみを除いた9種類について5000回の先行学習を行い、引き続いて4のみを5000回追加学習する。

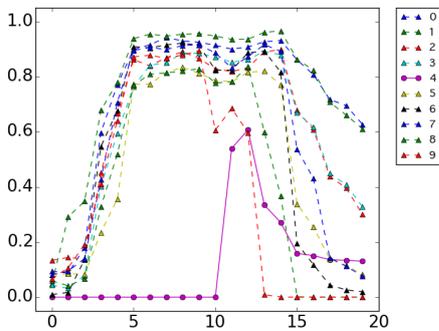


図1: MLPにおける10クラス追加学習

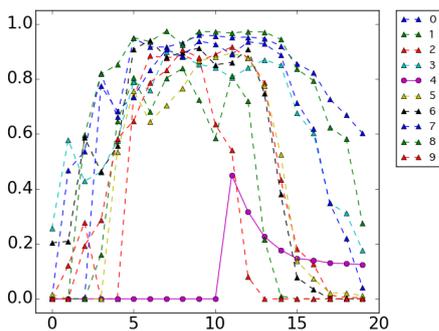


図2: MLP-Dropoutにおける10クラス追加学習

5. 考察

図1のRS-Dropoutによる追加学習後の推移に注目すると、追加学習による新たなクラスの学習と同時に学習構造の破壊が遅延していることがわかる。その後は通常が多層パーセプトロンと同様に構造の破壊が進んでしまうというを確認し

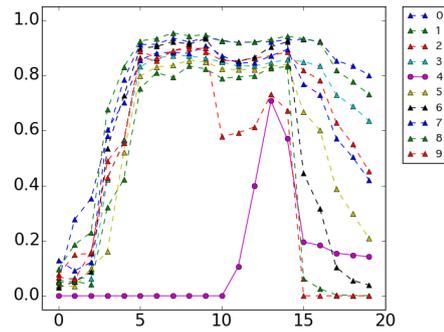


図3: MLP-RS Dropoutにおける10クラス追加学習

た。RS-Dropoutによって学習構造の破壊を遅延させながら一時的にクラス4の学習を追加的に学習することができた理由としては、学習度によるユニットの抑制が働いたことによると考えられる。未知であるクラス4をネットワークに入力した際に、先行学習によって獲得した構造の中における学習度の高いニューロンへの学習の上書きを抑制しながら、学習度の低いニューロンを優先して学習を行うことで分散された情報表現を弱めたことによると考えられる。

6. 結論

本研究では学習度を用いた構造化Dropoutによって、分散的な情報表現を弱め、未知データの入力に対して動的に学習の抑制を行いながら、影響の弱いニューロンによって学習を補うことで学習構造の破壊を遅延させることが可能であることが示唆された。この構造破壊の遅延を検出することで、未知データを未知と分類すること及び学習率の減少等による動的な対応が可能であると考えられる。今後の予定としては、学習度によってDropoutするユニットを選択する価値関数として、Softmax関数や他の関数による検証と、さらに大きな規模のネットワークでの検証を行っていく必要がある。

参考文献

- [太田 06] H. Ohta and Y.-P. Gunji, "Recurrent neural network architecture with pre-synaptic inhibition for incremental learning," *Neural Networks*, vol. 19, pp. 110-1119, 2006.
- [Nitish 14] Nitish, S., Geoffrey, H., Alex, K., Ilya, S., Ruslan, S., Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research* 15, 1929-1958 (2014).
- [篠原 07] 篠原修二, 田口亮, 桂田浩一, 新田恒雄. 因果性に基づく信念形成モデルとN本腕バンディット問題への適用, *人工知能学会論文誌*, Vol.22, No.1, pp.58-68, 2007.