

Mixture of Expertsによる気温の予測

梶, 雄大 / KAJI, Yudai

(出版者 / Publisher)

法政大学大学院理工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

61

(開始ページ / Start Page)

1

(終了ページ / End Page)

4

(発行年 / Year)

2020-03-24

(URL)

<https://doi.org/10.15002/00022882>

Mixture of Experts による気温の予測

TEMPERATURE PREDICTING BY MIXTURE OF EXPERTS

梶雄大

Yudai KAJI

指導教員 平原誠

法政大学大学院理工学研究科応用情報工学専攻修士課程

Research has been conducted to predict temperature in Neural Network (NN). It has devised the use of vegetation index and distance from the coast for input. In this study, I focused on the division of learning data to consider the difference in the difference in temperature height and low depending on the season. If I split the learning data manually, the division will include the intention of the creator of the network. So I use Mixture of Experts (ME) to predict temperature and automatically divide the learning data.

Key Words : Mixture of Experts, Neural Network,

1. はじめに

ニューラルネットワーク(NN)による気温の予測において、入力に過去の気温、気圧、湿度などの気象データを用いることは一般的である。しかし吉川ら[1]や宮崎ら[2]は入力に植生指標や海岸距離を使用するなどの工夫をすることで気温を予測した。

本研究では季節による気温の高低差の違いなどを考慮するため学習データの分割に注目した。学習データを分割せずに学習した1個のNNで気温を予測するより、学習データを季節別、月別などに分割し、それぞれのデータごとに学習した複数のNNで予測することで精度が上がると考えた。

学習データの分割でNNの作成者が手動分割すると、分割に作成者の意図が入ってしまうことや、作成者が気づかなかった良い分割方法を見逃してしまう可能性がある。本稿ではMixture of Experts(ME)を用いて気温を予測し、学習データの自動分割の効果を検討する。

2. NNによる気温の予測と生じた問題

MEで気温を予測する前に通常の3層NNで気温を予測した。NNは1時間毎の気温、気圧、湿度、風向、風量の12時間分の入力から、1時間先の気温を出力する。風向は東西南北の4入力として扱った。入力層のユニット数は96個(8入力×12時刻分)、中間層のユニット数は15個、出力層のユニット数は1個である。学習には誤差逆伝播法を用いた。

学習データの分割の効果を確認するために、全季節が含まれるデータで学習したNN(通年NN)と、季節ごとに手動分割したデータで学習した4個のNN(季節NN)を用意した。

季節NNと通年NNのテストデータに対する季節ごとの絶対誤差の結果(図1)から、秋を除き季節NNの方が誤差は小さくなったことが分かる。

秋NNを確認してみると、秋の気温を予測する場合は誤差が小さかった。一方で、暦上は秋だが気温が冬並みという季節外れの気温が数日続く箇所(図2)では、秋NNの誤差は大きいことが分かった。そのため季節NNは季節外れの気温の予測は難しいと考えられる。

次に4個の季節NNの出力の中から最も誤差が小さくなる季節NNの出力を手動で選び、理想の誤差を計算した。その結果(図1)、通年NNや季節NNより精度が良くなることが分かった。

以上のことより、季節などに特化した複数個のネットワークの中から、最適な選択ができた場合、精度が高まると考えられる。

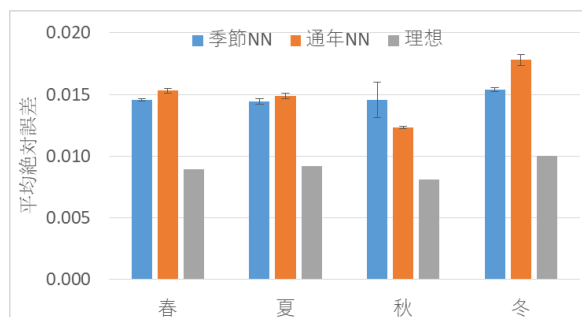


図1 季節NNと通年NNと理想の誤差

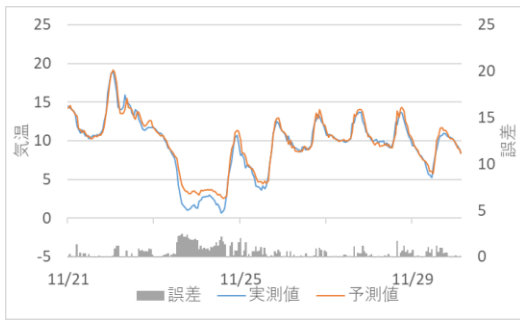


図2 秋 NN の拡大図

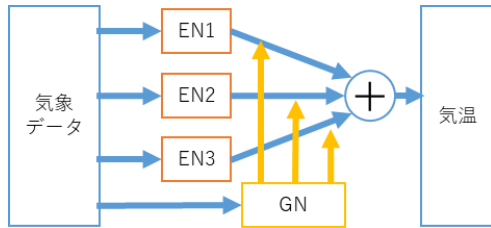


図3 気温予測の ME の構造

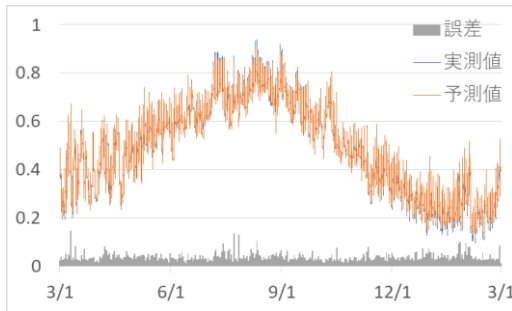


図4 季節 ME の予測値と実測値と誤差

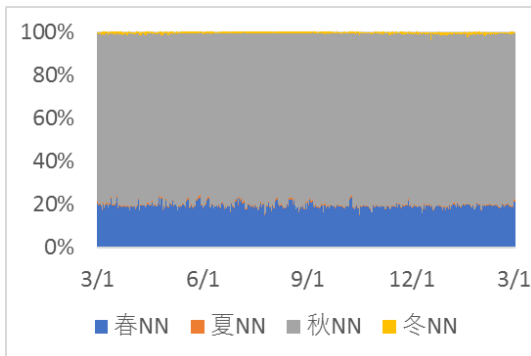


図5 季節 ME の GN の出力結果

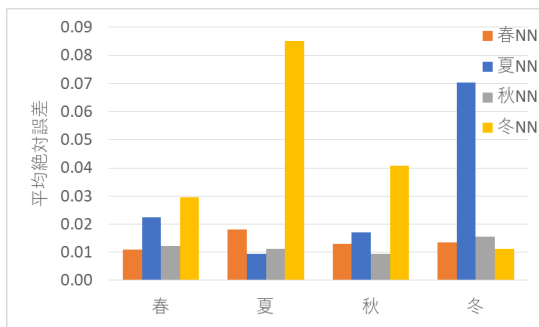


図6 季節 NN の季節ごとの誤差

3. Mixture of Experts(ME)の利用

3.1. ME の利用

ME は、図 3 に示す構造をしており、複数個のエキスパートネットワーク(EN)と 1 個のゲーティングネットワーク(GN)で構成されている。EN は予測を出力するネットワークで、GN は入力にふさわしい EN を選択するネットワークになっている[3][4]。

例えば EN が季節に特化しているとする。暦上では秋だが冬のような気温を予測する場合(図 2)において、GN が冬に特化した EN を選択するようになれば予測精度が高まると考えられる。

学習により何かに特化した EN の中から最適な EN を GN によって自動選択するため、本稿では ME を利用することにした。

3.2. GN の効果の確認

3.2.1. 季節 ME

ME を気温の予測に利用できるか確かめる。GN の効果を確認するため、EN に学習済みの季節 NN(春,夏,秋,冬)を設定し、GN の重みのみ更新するように ME を学習させた。この ME を季節 ME と呼ぶことにする。GN は 3 層構造で、入力層のユニット数は 96 個、中間層のユニット数は 20 個、出力層のユニット数は 4 個(EN の数)である。

EN を季節 NN に設定することで、GN が春なら春 NN を 10 割、夏なら夏 NN を 10 割で予測するようになることを期待した。

季節 ME の予測結果(図 4)より、気温の予測ができていることが分かる。学習後の GN の出力結果(図 5)より、通年で春 NN の出力を 2 割、秋 NN の出力を 8 割使うことで予測していることが分かる。夏 NN と冬 NN は全く使われていないことも分かる。

GN の結果(図 5)から、学習データを分割する効果は見られなかった。これは春 NN と秋 NN の季節ごとの誤差の変動(図 6)が小さいために通年で春 NN と秋 NN のみで予測することが良いと GN が学習したためと考えられる。また、春 NN と秋 NN は通年 NN と出力が似ている。春 NN と秋 NN は季節に特化していない NN であったため、学習データの分割が出来なかったのだろう。

3.2.2. 夏冬 ME

3.2.1.より、通年で誤差の変動が小さい春 NN と秋 NN は特化していない NN であるため、GN の学習データの分割が出来ないと考えられた。このことを確認するため、夏 NN と冬 NN を EN に設定した ME で学習させた。この ME を夏冬 ME と呼ぶことにする。

EN の重みを学習済みの夏 NN と冬 NN の重みに固定し、GN の重みのみ更新するように学習させた。GN の中間層のユニット数は 20 個である。

学習後の GN の出力結果(図 7)より、気温が高い 7 月から 9 月は夏 NN がほぼ 10 割で予測し、気温が低い 12 月から 2 月は冬 NN がほぼ 10 割で予測していることが分かる。その他の期間は夏 NN と冬 NN を混合して予測して

いることが分かる。

GN は、高気温に特化した夏 NN と低気温に特化した冬 NN を気温の高低によって選択していると考えられる。このため EN が何かに特化すれば、GN で最適な EN を選択することができることが分かる。

4. ME の学習

4.1. ME

3.2.2 より、EN を学習済み NN にすることで、GN による分割効果を確認できた。ここでは EN と GN の重みを乱数で設定し、EN と GN の重みの両方を更新するように ME を学習させた。EN 数は 12 個、EN の中間層のユニット数は 1 個、GN の中間層のユニット数は 20 個である。

学習後の GN の出力結果(図 8)より、学習データが分割されずに 1 個の EN のみで予測していることが分かる。これは初回の重み更新の際、誤差が小さい EN が選択されるように重みが更新されたことで、次の重み更新でもその EN が選択され続け、他の EN の重みが更新されないためであると考えられる。

4.2. 通年 ME

4.1 より EN と GN の重みを乱数で設定した場合、初回の重み更新から特定の EN が選択され続け、1 個の EN のみで予測してしまうことが分かった。1 個の EN で予測することは 1 個の NN で予測することと同じである。

このことを避けるために、EN の数だけ学習済みの通年 NN を用意して EN の重みに設定した。そのため学習前の EN の出力は気温に近い値になる。通年 NN は分割していないデータで学習しているため、季節 NN のように特化した NN ではない。この ME を通年 ME と呼ぶことにする。

各 EN の出力が気温に近いため、初回の重み更新で GN が誤差の小さい EN を選択するように重み更新しても、次の重み更新で他の EN が選択される可能性が重みを乱数に設定した場合より高くなる。そのため GN が複数の EN を選択するようになり、学習が進むにつれて GN が学習データを分割し、EN が特化していくことが期待できる。

EN は学習済みの通年 NN を設定し、GN は乱数で設定した。EN と GN の重みの両方を更新するように通年 ME を学習させた。EN 数は 4 個、EN の中間層のユニット数は 15 個、GN の中間層のユニット数は 20 個である。

学習後の GN の出力結果(図 9)より、EN1 は 6 月から 10 月に影響が大きくなっている。EN2 と EN4 は 12 月から 3 月に影響が大きくなっている。そのため EN1 は高気温に強い EN に、EN2 と EN4 は低気温に強い EN になったと考えられる。また、特化していない通年 NN から学習しても EN が特化していくことが分かった。

次に通年 ME の性能を確認する。通年 NN と通年 ME の誤差を比較した結果(図 10)より、通年 ME は通年 NN より精度が向上していることが分かる。

以上の結果より、EN の重みを通年 NN に変更することで、1 個の EN のみで予測することを避けることができ、通年 NN より精度が向上することを確認できた。

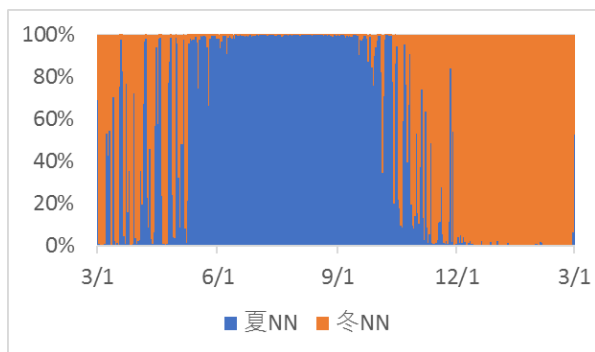


図 7 夏冬 ME の GN の出力結果

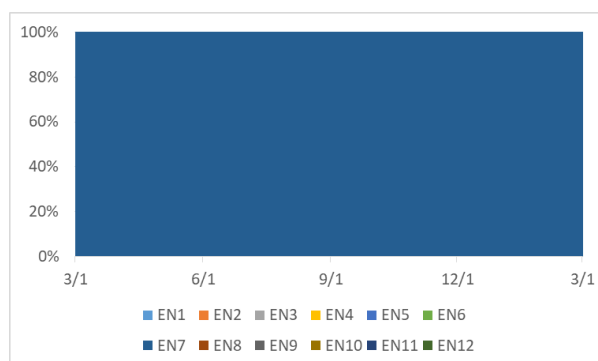


図 8 ME の GN の出力結果

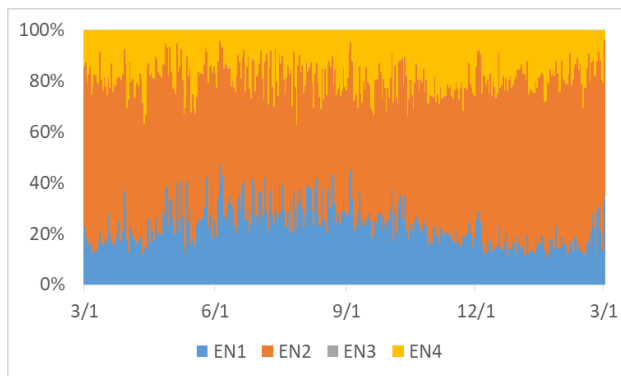


図 9 通年 ME の GN の出力結果

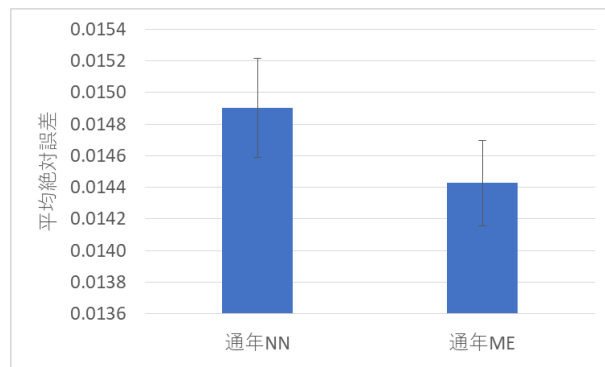


図 10 通年 NN と通年 ME の誤差の比較

4.3. EN の中間層のユニット数を減らす

4.2.で EN に設定した通年 NN はあらかじめ中間層のユニット数を調整していた。1 個の通年 NN で気温を予測できるため、GN によって明確な分割がされにくいと考えられる。そのため通年 NN の中間ユニットを 1 個に変更することにした。EN の中間層のユニット数を減らすことで EN の能力は低下することで、ME による学習データの分割が明確になると考えた。また、中間ユニット数が減少することで学習時間が短くなることも期待できる。

EN の重みは中間ユニット数を 1 個の学習済みの通年 NN の重みを設定し、GN の重みは乱数で設定した。EN と GN の重みの両方を更新するように ME を学習させた。EN 数は 4 個、GN の中間層のユニット数は 20 個である。

中間ユニット数が 1 個の通年 NN(1)と通年 ME(1)の誤差を比較した結果(図 11)より、通年 ME(1)は通年 NN(1)より精度が向上していることが分かる。

EN の中間層のユニット数が 1 個の通年 ME(1)と EN の中間層のユニット数が 15 個の通年 ME(15)の性能を比較した結果(図 12)より、中間層のユニット数を減らしても予測精度を保つことができたことが分かる。

通年 ME(1)の GN の出力結果(図 13)より、EN1 と EN3 は 6 月から 9 月に影響が高まり、EN2 は 12 月から 3 月に影響が高まること分かる。これは EN の中間ユニット数に関係なく分割されていることを示している。しかし学習データの分割は明確にならなかった。

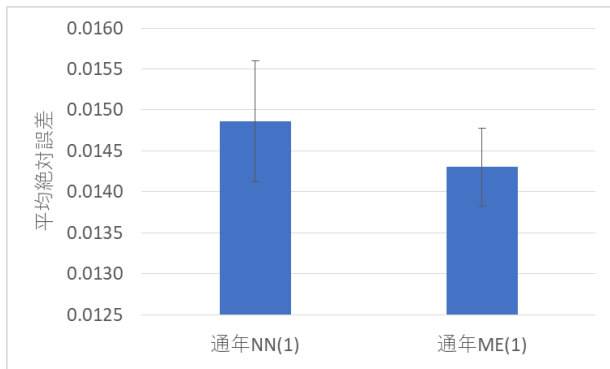


図 1 1 通年 NN(1)と通年 ME(1)の誤差

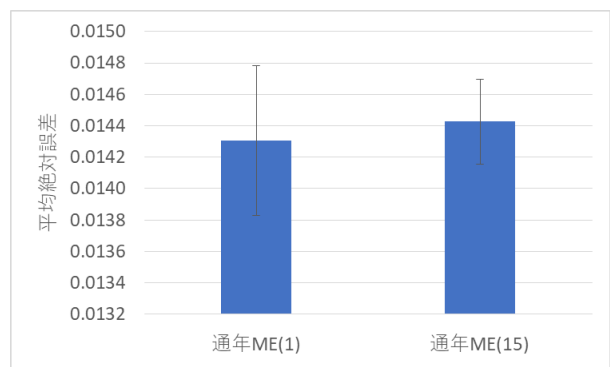


図 1 2 通年 ME(1)と通年 ME(15)の誤差

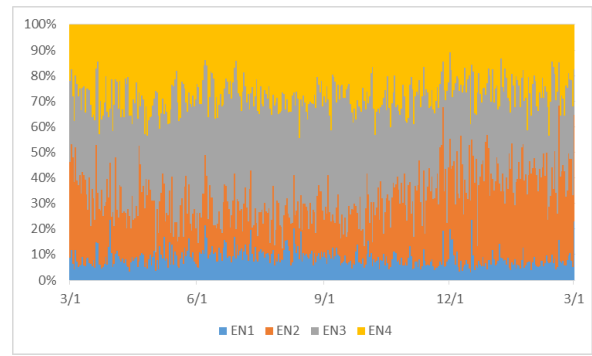


図 1 3 通年 ME(1)の GN の出力

5. 結論

4.1.と 4.2.より、ME で気温を予測する際は EN の重みは乱数で設定すると 1 個の EN で学習してしまうため、あらかじめ学習した NN を設定する方が良いことが分かった。また NN と ME の性能を比較すると、ME は NN より精度が高いことが分かった。

4.3.より EN の中間層のユニット数を減らしても精度を保つことができることが分かった。このことより EN の能力が低下しても ME による学習データの分割によって精度を保つことができる。

以上より気温を予測する際に ME は有効であると考えられる。

6. 今後の課題

EN に学習済みの NN を設定するとき、学習済みの NN は設定する EN の数だけ必要である。多くの学習済みの NN の用意は手間がかかる。そのため、1 個の NN の重みの一部を乱数に変えて複数の EN に設置するなどすることで、少数の学習済み NN から多くの EN を作り出す必要がある。

謝辞

本研究を進めるにあたり、ご指導いただいた修士論文指導教員の平原誠准教授に心より感謝いたします。

参考文献

- 1) 吉川敏則, 田中賢次, “階層型ニューラルネットを用いた局所的最低気温の予測,” 電子情報通信学会論文誌, Vol.J81-D2, pp.453-454, 1998.
- 2) 宮崎ひろ志, 森山正和, “ニューラルネットワークを用いた気温分布推定に関する研究,” 日本建築学会計画系論文集, Vol.543, pp.71-76, 2001.
- 3) Robert A. Jacobs, Michael I. Jordan, Steven J. Nowlan, Geoffrey E. Hinton, “Adaptive mixtures of local experts,” Neural Computation, Vol. 3, pp. 79-87, 1991.
- 4) Perry Moerland, “Some Methods for Training Mixtures of Experts,” IDIAP-RR 97-05, 1997.