

超高密度気象観測データを用いたLSTMによる暑さ指数の予測

甲斐 健太[†] 橋田 修一^{††} 田村 慶一^{††}

[†] 広島市立大学情報科学部

^{††} 広島市立大学大学院情報科学研究科

1. はじめに

近年、熱中症対策として、熱中症になる危険度の目安となる暑さ指数を予測する研究が行われており、熱中症予防が重要となっている。環境省やテレビ報道において暑さ指数の予報を行っているが、予報は市町村などの行政区レベルであり、暑さ指数は環境が異なると極端に変化することが知られており、局所的な予測が必要であることが考えられる。そこで、本研究では、自立式の小型気象観測装置として注目されている POTEKA において観測された超高密度気象観測データを用いた LSTM モデルによる暑さ指数の予測手法を提案する。

2. 暑さ指数とその予測

暑さ指数(WBGT)とは、熱中症のなりやすさを示す温度の指標で単位は[°C]で表されるが、気温と同一のものではなく、下記の式(1)で計算される[1]。

$$\begin{aligned} \text{WBGT}(\text{°C}) = & 0.735 \times \text{Ta} + 0.0374 \times \text{RH} + 0.00292 \times \text{Ta} \\ & \times \text{RH} + 7.619 \times \text{SR} - 4.557 \times \text{SR}^2 \\ & - 0.0572 \times \text{WS} - 4.064 \quad \dots(1) \end{aligned}$$

Ta は気温(°C), RH は相対湿度(%), SR は全天日射量(km/m²), WS は平均風速(m/s)である。環境省では気象庁の数値予測に基づいて時間ごとの予測を行っている。

3. 提案モデル

POTEKA では、20 種類の気象データ観測を1分間隔で行い、携帯回線を使ってクラウド上にリアルタイムにデータを保存するシステムと連動してサービスを提供している。POTEKA を設置することで気象庁の観測地点に依存しないピンポイントでの気象観測が可能となる。1分間隔で各地の気象観測を行うため、データは超高密度に集めることができる。

本研究では、LSTM モデル(Long Short Term Memory)を用いて、暑さ指数の予測を行う。LSTM モデルとは、リカレントニューラルネットワーク(RNN)では不十分な長期的な依存関係を学習できるモデルである。我々は、過去の暑さ指数から暑さ指数を予測するだけでなく、超高密度気象観測データの気温、湿度、日射、風速、暑さ指数を多次元のデータとして扱い、多次元時系列データを用いてモデルの学習と予測を行う。

本研究では、 m 分間のデータを入力としてその後の n 分間の暑さ指数の推移を予測する。そこで、過去のデータをスライディング方式により抽出し、学習データを作成する。加えて、過去のデータが存在しない地点については近隣の観測地点のデータが利用可能かの検証を行う。

4. 評価実験

評価実験では、島根県の5地点の気象観測のデータのうち2年分(2016年11月~2018年9月)を用いて暑さ指数の予測を行った。学習データとして2016年11月~2017年10月のデータを用いる。検証データとして2018年5月~2018年9月のデータを用いる。入力データ項目として気温、湿度、日射、風速、暑さ指数を抽出し、欠損値処理、10分、20分間隔の平均化、標準化を行っている。

入力データ長を1日ごとに1日~7日まで、出力データ長を60分と、150分から330分まで60分ごとに変化させ、1層のLSTMで暑さ指数のみから暑さ指数を予測した。実

験結果から入力データ長は4日と5日が最適で、出力データ長は少ないほど暑さ指数の予測値が実データに一致しやすいことがわかった。図1と図2に暑さ指数のみで学習した際の暑さ指数の予測結果を示す。実験結果から観測データが通常とは違う動きをする日では予測ができていないことが分かった。

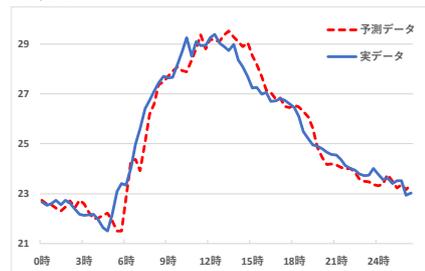


図1. 予測結果(誤差小)

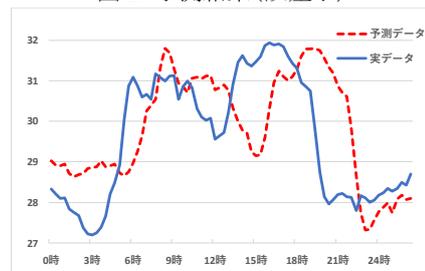


図2. 予測結果(誤差大)

多次元時系列データで予測する場合の暑さ指数の予測精度を検証した。多次元時系列データを使用する場合は{湿度、風速、暑さ指数}で予測する場合は他の要素を組み合わせた場合よりも実データに近い予測値を得られた。

また、島根県川本町役場の周辺4地点の過去のデータが存在しないと仮定し、川本町役場のデータで学習したモデルで4地点の暑さ指数を予測できるか検証を行った。その結果、川本町役場のデータで学習したモデルを使用したときの誤差の平均は1.0°C~2.0°Cで、4地点でそれぞれ学習したモデルの誤差の平均は0.5°C~1.2°Cだった。これにより、近隣の観測地点のデータを利用できることが分かった。

5. まとめ

本研究では超高密度気象観測データを用いたLSTMによる暑さ指数の予測手法を提案した。出力時間を抑えることで一定の精度で暑さ指数を予測することができた。しかし、深層学習だけでは安定した予測をすることができないため、今後の課題として、安定した予測を可能にする補正方法の検討が必要であると考えられる。

謝辞

本研究の一部は、JSPS 科研費 JP18K11320、広島市立大学・特定研究費とサタケ技術振興財団の助成により行われた。また、POTEKA データについては(株)テクノシステムから提供を受けた。

参考文献

- [1] 小野雅司, 登内道彦, 通常観測気象要素を用いた WBGT (湿球黒球温度)の推定, 日生気誌 50(4):147-157, 2014.