

令和3年度 卒業論文概要					
所属	物理情報工学科・電子物理工学コース				
学生番号	20224004	学生氏名	橋口駿亮	指導教員	小田部荘司
論文題目	ビジネス・社会の問題解決に向けた AI によるシステム開発				

1. 緒言

近年 AI 技術の発展により、AI を使ったシステム開発が広まっており、その1つとして電力需要予測がある。電力需要は IT 技術の進展やエアコンの普及などによって、全世界において顕著である。しかしそれに伴い、地球温暖化や CO2 排出などの社会・環境問題に対して、世界各国で持続可能な環境づくりを目指す動きがあり、経済成長と関係する部分が増えてきている。この社会・環境問題の解決に、発電量を調整することを目的として、電力需要データや天候などのデータを使って電力需要モデルを構築し、電力データを予測する研究が進められている。しかし、電力需要モデルをビジネスや社会に適用するためのシステム作りをするためにも、電力需要モデルの複雑化や入力データの次元数増加、学習に対する時間・メモリなどのコスト増大は回避しなければならない。本研究では、電力需要モデルを機械学習システムとして導入することを踏まえて、学習に使用する入力データの次元を小さくし、単純なモデルで翌日の電力需要予測を行った。

2. 理論

2.1. LSTM (Long Short-Time Memory)

LSTM は RNN (Recurrent Neural Network) の長期記憶を保持するのが難しいという問題点を改善した時系列モデルである。LSTM はゲートと呼ばれる仕組みを導入することで、過去の情報を忘れるか忘れないかを判断しながら、必要な情報だけを次の時刻に引き継ぐことができる。

2.2. 主成分分析 (PCA : Principal Component Analysis)

主成分分析とは、代表的な次元削減手法の1つである。次元削減は、データの次元を、人間が理解しやすい2次元、3次元に削減することを目的としている。主成分分析はデータから分散が大きいベクトル(固有ベクトル)を抽出し、ベクトルを並べた行列と元データとの内積を取ることで、データの次元を圧縮できる。

2.2. 選択的アンサンブル平均法

選択的アンサンブル平均法は Zhou et.al. が提唱したアンサンブル学習で、複数のモデルを学習させ、そのモデルの中から評価誤差が小さいモデルをいくつか選択し、アンサンブル平均をとる手法である。これにより予測性能・分類性能の向上の可能性を示した。

4. 前処理

前処理として、電力需要データは規格化と次元圧縮、最高気温・最低気温のデータは規格化、月 (month) は自然数値、祝日・休暇はラベル「1」(残りはラベル「0」)、天候は one-hot ベクトル、曜日は0~6の番号を割り当てる。

3. 従来手法

従来手法 [1] のモデル図を Fig. 1 に示す。従来手法では、3つの全結合層 NN (Neural Network) を使っている。Fig. 1 のモデル図より、左のモデルから順に、翌日 0 時~7 時までの電力データの平均を予測するモデル、翌日の電力データの主成分得点を予測するモデル、翌日 24 時間の電力データの予測をするモデルを表している。また、翌日の電力データの主成分得点を予測するモデルで、主成分得点を予測して、電力需要の近似曲線を得ているが、この近似曲線は、主成分分析で用いた $d-1$ 日までの電力データの時間ごとの平均値と、電力データの第 4 主成分と翌日の主成分得点との線形結合の和で求めている。

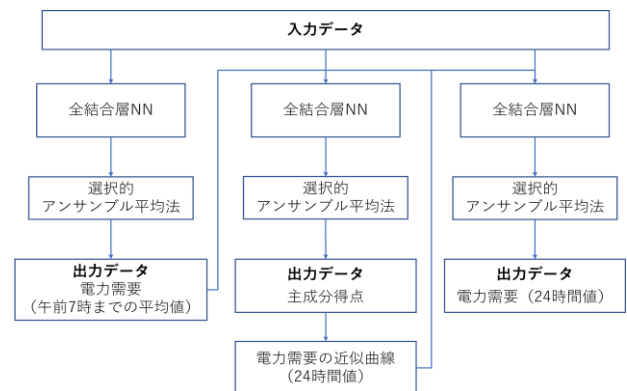


Fig. 1 従来手法のモデル

5. 提案手法

提案手法のモデル図を Fig. 2 に示す。モデル図は、入力データ（次元圧縮された電力データと暦データ、気象データ）を時系列モデルの LSTM に入力し、選択的アンサンブル平均法を行い、翌日の電力需要を予測している。



Fig. 2 提案手法のモデル

5. 学習条件

今回のモデルを学習する際のパラメータや関数を Table 1 に示す。

Table 1 学習条件

条件	従来手法	提案手法
エポック数	100	
バッチサイズ	16	
アンサンブル	10	
最適化関数	Adam	
誤差関数	RMSE	
活性化関数	softplus	tanh

また学習の際、EarlyStopping と呼ばれる Callback 関数を用意し、評価誤差が連続で大きくなったら学習を止め、評価誤差が最も小さいときのモデルの重みを保持する。これにより、過学習が発生することなく、最適なモデルの重みで予測することができる。よって、エポック数は

100 epochs に達することなく、学習が終了する場合がある。

データに関しては、電力データは東京電力の電力データを使用し、気象データは気象庁が公開している気象データを使用する。学習データは 2016 年 4 月 1 日～2017 年 6 月 30 日、テストデータは 2017 年 7 月 1 日～2018 年 6 月 30 日までの 1 年間としてモデルの学習・予測を行った。入力データの総次元数は 13 次元である。また、学習データから 2 割、評価データとして学習を行った。

5. 実験結果

提案手法によるテストデータの 2018 年 6 月 10 日～2018 年 6 月 16 日の 1 週間の間の翌日の電力需要予測を Fig. 3 に、従来手法と提案手法のテストデータの RMSE (Root Mean Square Error) を Table 2 に示す。Fig. 3 の origin は実績値、pred. は予測値を表しており、グラフから正確な予測ができていることが確認できる。また、Table 2 から従来手法に比べ良い性能となった。

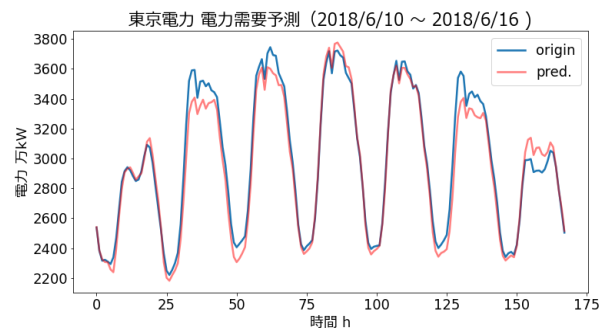


Fig. 3 提案手法による電力需要予測

Table 2 東京電力の電力需要予測における RMSE の結果

	従来手法	提案手法
RMSE 万 kW	196.8	154.0

6. 結論

本研究では、入力の次元数を小さくし、単純なモデルで電力需要予測を行った。結果として、正確な電力需要予測を行うことができ、従来に比べ良い性能となった。今後の展望としては、他の電力データで評価を行い、今回のモデルが実用的か調査する必要があると考えられる。また、東芝グループが研究として発表したスパースモデリングによる気象地点選択を考慮した電力需要予測[2]と組み合わせることで、より強固な電力需要予測を構築できると考えられる。

参考文献

- [1] 松尾雄司ら. 人工ニューラルネットワークを用いた短期電気需要予測 -国内 10 地域を対象とした検討-. (2018). <http://eneken.ieej.or.jp/data/8106.pdf>
- [2] 気象予測データと機械学習を用いた高精度な電力需要予測手法, September 2019, 東芝レビュー, 74, 5