

電気使用量予測のための深層学習手法

1 月 18 日 (火)

小松 大起

1 はじめに

1.1 研究背景

2011 年 3 月 11 日に発生した東日本大震災以降、低コストかつ効率的な発電方法であり全体の 2 割から 4 割を占めていた原子力発電は安全性に問題があるとされ 2021 年 5 月現在では停止中の 2 基も合わせて合計 9 基の原子力発電所が稼働するのみとなっている [?]. それにより、原子力発電よりコストの高い火力発電での発電量が増え発電コストの増加による電気料金の値上げが生じ、我々消費者の負担増加に繋がっている [?].

1.2 研究目的

電気使用量は人間の認知は時間経過による視覚世界の変化の予測が可能である。近年では実際に予測動画を作る研究も行われてきている.[sankou2] 電力使用量を予測することによる、電気料金の予測が可能になると考えられる。本研究では、電力使用量を主データとし、天気や気温が与える電力使用量の変化を考慮した電気使用量の予測を行うことを目的とする。

加重移動平均 (Weighted Moving Average) は WMA と呼ばれ、一定の量ずつ重みを線形に減らしていき平滑化を行う手法のことである。n 個のデータからの WMA を求めたい場合には、直近のデータの重みを n とし、その前のデータの重みを n-1 のように重みを減らしていき、最終的な重みを 0 とすることにより古いデータに対して重みを減らしていくことで、古いデータの影響を弱くする手法である。

指数移動平均 (Exponential Moving Average) は EMA と呼ばれ、指数加重移動平均 (Exponentially Weighted Moving Average) や指数平滑移動平均 (Exponentially Smoothed Moving Average) と呼ばれることもある。EMA は WMA では線形に減らしていた重みを指数関数的に減らしていく手法である。本研究では、指数移動平均を用いた時系列データの平滑化を行なった。期間 10 日ごとで移動標準偏差 (Moving Standard Deviation) の計算を行い、10 日ごとでの期間で EMA を求め、その後求めた EMA の値が移動標準偏差から n 倍以上大きい値や小さい値をそれぞれ外れ値として検出を行なった。それぞれ 1.3 倍, 1.4 倍, 1.5 倍, 2.0 倍の時の検出結果を Fig. から Fig. に示す。

2 外れ値検出手法

時系列データを平滑化する手法の一つに移動平均が存在しており、移動平均は大きく分けて単純移動平均、加重移動平均、指数移動平均の 3 種類の手法が存在している。移動平均は主に金融分野でのテクニカル分析や、気象の分析などで用いられている手法である。

単純移動平均 (Simple Moving Average) は SMA と呼ばれ、直近 n 個のデータの平均を単純に重み付けをすることなくとり、平滑化を行う手法である。単純移動平均のデメリットとして、重み付けを行わない手法であるため古いデータの影響を受けすぎてしまうという点が挙げられる。

3 修論

修論 90 % 39 枚

1 はじめに 100 % 2p

1.1 研究背景 100 %

1.2 研究目的 100 %

2 機械学習手法について 80 % 17p

2.1 教師あり学習手法 100 %

2.1.1 線形回帰 100 %

2.1.2 ロジスティック回帰 100 %

2.1.3 K 近傍法 100 %

2.1.1 SVM (Support Vector Machine) 100 %

2.1.2 決定木 100 %

2.1.3 アンサンブル学習 100 %

2.2 教師なし学習手法 90 %

2.2.1 K 平均法 100 %

2.2.1 主成分分析 100 %

2.2.2 アソシエーション分析 100 %

2.2.3 自己組織化マップ 50 %

2.3 半教師あり学習 100 %

2.4 強化学習 10 %

2.4.1 動的計画法 100 %

2.4.1 モンテカルロ法 100 %

2.4.2 TD 学習 100 %

2.5 形式ニューロン 100 %

2.7 ニューラルネットワーク %

2.7.7 損失関数 100 %

2.7.8 活性化関数 100 %

2.7.9 学習アルゴリズム 100 %

2.7.1 パーセプトロン 100 %

2.7.1 オートエンコーダ 100 %

2.7.1 GAN (Generative Adversarial Networks) 100 %

2.7.2 CGAN (Conditional GAN) 100 %

2.7.3 DCGAN (Deep Convolutional GAN) 100 %

2.7.4 CNN 100 %

2.7.5 RNN 80 %

2.7.6 LSTM 80 %

2.7.10 深層学習モデルを用いたマルチモーダル予測

50 %

3 対象データとその特徴分析 80 % 14p

3.1 電力使用量 100 %

3.2 気温 100 %

3.3 降水量 100 %

3.4 天気 100 %

3.5 各データ間の関係性 100 %

3.5.1 相関係数 100 %

3.5.2 相互情報量 100 %

3.6 平均気温の平均からの絶対値 50 %

3.6 外れ値検出手法 70 %

4 深層学習手法による電力使用量予測 100 % 2p

4.1 RNN による単入力予測 100 %

4.2 LSTM による単入力予測 100 %

5 マルチモーダル入力による電力使用量予測 75 %

4p

5.1 LSTM による多入力予測 100 %

5.2 処理済み気温データを用いた LSTM による多入力予測 100 %

5.3 外れ値検出結果と電力使用量を用いた LSTM による多入力予測 100 %

5.4 構造化 LSTM による多入力予測 0 %

6 考察 70 % 2p

7 まとめ 0 %

7.1 まとめ 0 %

7.2 今後の課題 0 %

4 先週の作業

- 終わり次第構造化 LSTM をやる.

- 気温の実データと平均の差の絶対値（処理済みの気温データ）を使った電力使用量予測をした. グラフ作りと考察を行う.
- 外れ値のほうも進める.
- 修論の加筆と修正

5 今週の作業

- 外れ値を用いた予測精度が高かった. 他の倍率での結果でも出してみる.
- 修論の加筆と修正

Table 7: 時間ごとの各実験での予測精度上位 2 項目

時間ごと	入力項目	層数・素子数	結果	入力項目	層数・素子数	結果
実験 1	電力使用量のみ	2 層・100 個	0.0136	電力使用量のみ	2 層・200 個	0.0146
実験 2	電力使用量のみ	2 層・100 個	0.0136	電力使用量のみ	3 層・100 個	0.0137
実験 3	降水量	2 層・100 個	0.0130	降水量	2 層・100 個	0.0137
実験 4	3 項目	3 層・100 個	0.0115	気温 & 降水量	3 層・300 個	0.0119

Table 8: 日にちごとの各実験での予測精度上位 2 項目

日にちごと	入力項目	層数・素子数	結果	入力項目	層数・素子数	結果
実験 1	電力使用量のみ	3 層・100 個	0.0912	電力使用量のみ	2 層・200 個	0.0914
実験 2	電力使用量のみ	2 層・300 個	0.0919	電力使用量のみ	2 層・200 個	0.0924
実験 3	気温 & 降水量	2 層・300 個	0.0939	天気	2 層・300 個	0.0944
実験 4	気温	3 層・300 個	0.0857	気温 & 降水量	1 層・100 個	0.0858
実験 5	外れ値	1 層・100 個	0.0793	外れ値	1 層・300 個	0.0805

Table 9: 週ごとの各実験での予測精度上位 2 項目

週ごと	入力項目	層数・素子数	結果	入力項目	層数・素子数	結果
実験 1	電力使用量のみ	2 層・100 個	0.142	電力使用量のみ	2 層・200 個	0.143
実験 2	電力使用量のみ	1 層・300 個	0.167	電力使用量のみ	2 層・100 個	0.169
実験 3	3 項目	2 層・200 個	0.0972	3 項目	2 層・300 個	0.0989
実験 4	気温 & 天気	3 層・300 個	0.150	気温 & 天気	1 層・100 個	0.157

Fig. 1: 上位 2 つづつの結果