第十七屆台灣電力電子研討會暨
第四十一屆中華民國電力工程研討會
台灣 台北市 2020 年 9 月 3-4 日

多項式模型校驗機器學習之負載預測分析結果

Analyze the machine learning forecast result using the polynomial equations

池欣慶 [1] 林韋辰 [1] 黃維澤 [1,*] 姚凱超 [1]
Hsin-Ching Chih Wei-Chen Lin Wei-Tzer Huang Kai-Chao Yao
李奕德 [2] 姜政綸 [2] 何元祥 [2] 蔡佳豪 [2]
Yih-Der Lee Jheng-Lun Jiang Yuan-Hsiang Ho Jia-Hao Cai

[1] 國立彰化師範大學工業教育與技術學系
Department of Industrial Education and Technology, National Changhua University of Education
vichuang@cc.ncue.edu.tw

[2] 核能研究所
The Institute of Nuclear Energy Research

摘要

本文以校園負載曲線為研究標的，採用長短期記憶法(LSTM)與極限梯度提升法(XGBoost)二種機械學習方法進行預測。並用賽局理論之沙普利值法(SHAP)檢驗預測模型，可得出預測週期內之各影響因子比例，供電力業者評估。然而，預測結果準確性常與模型訓練等參數有關，常困擾著相互間資料交換時的信心程度。本文提出以多項式模型，採大數理論為基礎所設計之資料格式，以相同資料大小驗證前述步驟，求得整體預測結果之各因子誤差水準與各因子間的差異值。

關鍵詞：負載預測、長短期記憶、極限梯度提升、沙普利值分析

Abstract

This study focuses on the campus load forecasting that utilizes the two machine learning methods, impact factors analysis, and polynomial model. The long short-term memory method (LSTM), eXtreme Gradient Boosting, XGBoost (XGBoost), SHapley Additive exPlanations (SHAP), and the polynomial model associates the work into reality. LSTM and XGBoost methods serve the load forecast work. SHAP method demystifies the forecast result in the explanation. Finally, the proposed polynomial model validates the specific confident level of the forecast result for data exchange.

Keywords: load forecasting, LSTM, XGBoost, SHAP.

I. 簡介

近年來電力負載預測之結果，應用在許多不同場合，如電力公司資訊網頁、電力潮流調度控制、配電系統調度維護最佳化[1]與綠色能源設備投資[2-4]等應用。台灣電力股份有限公司亦成立整合大數據系統，計算出當時負載調度曲線供電力調度處參考。除要求實際與預測出的負載曲線吻合，亦需要減低預測所需之計算資源並且分析其組成之因子。較低之預測時間可採用較低成本之計算硬體；瞭解組成因子能掌握負載變動的趨勢。

隨著資料儲存媒介與技術興起，電力設備與負載等資訊均能儲存供參考，傳統的統計法如迴歸分析如多元序列分析(Autoregressive Integrated Moving Average with Exogenous Inputs, ARIMAX)與狀態方程式法如 Kalman 等均有其實作上的挑戰，前者為無法得知其組成因子；後者為狀態方程式的推導等挑戰。然而，機器學習(Machine Learning, ML)與數據資料的結合能夠分析出之標的模型，如股價趨勢等應用。本文採用之機器學習法之長短期記憶(LSTM) [2-4]與極限梯度提升(XGBoost)[5]等方法設計負載預測。並且採用沙普利值法(SHAP)[6]分析預測模型之預測因子。本文提出以統一因子(Unified factor)將預測因子集合至單一標準，供使用者參考。

再者，不同的負載預測法則、預測資料庫多寡等設定均會影響統一因子結果。例如，較少的資料庫僅需較少的模型訓練時間，但降低結果之精確性；反之，較多的訓練資料，在計算硬體資源有其限制，無法均一適用。有其必要性以標準化模型驗證出關係因子之信心度，供最終資料交換時給使用者參考。本文提出以多項式模型取得統一因子結果，作為評鑑之標準。下圖一為預測資料、機器學習法則、分析處理與本文貢獻之關係圖。

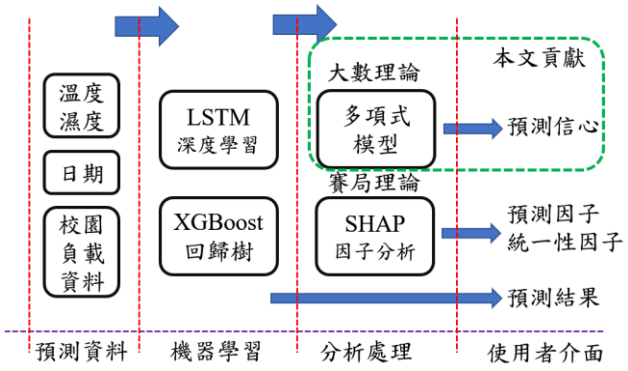

圖 1 預測資料、機器學習法則、分析處理與本文貢獻之關係圖

下章節將介紹本文採用的機器學習法則與因子分析原理。第三章則說明預測模型設計與相對性關係因子；第四章為校園負載預測結果比較與多項式檢驗模型統一因子分析。討論與結論則分別論述於第五章與第六章。

II. 機器學習預測

機器學習通常思考成：分析歷史資料，取得其規則，達成人工智慧的方法。本文中的歷史資料為電力負載資料，"取得其規則"之方式分別為 LSTM 法與 XGBoost，規則為 SHAP 法之組成因子，達成人工智慧的目的可當作後續的應用，如配電系統調度維護最佳化[1]等等。本章節將陸續介紹 LSTM 法、XGBoost 法與 SHAP 法之背景與基本理論。

2.1 長短期記憶(LSTM)

LSTM 法為 Hochreiter[7]於 1997 年所提出，屬於機械學習內的深度學習分枝，特點是在神經單元內加入遺忘門(forget gate)，免除梯度消失等因傳統遞歸神經網路(Recurrent Neural Network, RNN)問題，典型結構如下圖 2。Ospina 等人亦提出以此法搭配穩定小波轉換(Stationary Wavelet Transform, SWT)取得 PV 發電之預測結果，與傳統 LSTM 法相比可減少約 30%的 RMSE 誤差，但增加約一倍的運算時間[4]。本文則以校園負載為預測輸出結果；將日期時間參考等分離成獨立因子供 LSTM 法運算，取得較精細預測分析結果。

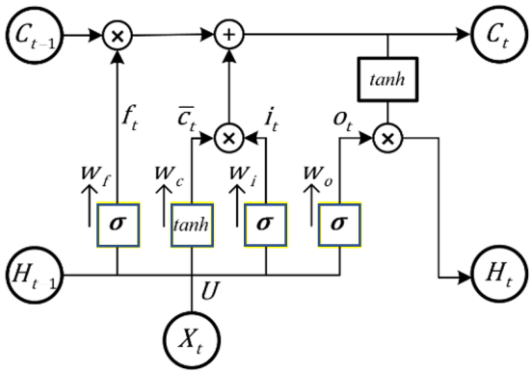

圖 2 典型 LSTM 神經網路

數學表達可由(1)至(6)式，其中b_i, b_f, b_o與b表示為偏壓向量，記憶體單元$\bar{c}_t$用來儲存經由激活函數所得出之長期記憶參數，狀態變數C_{t-1}與H_{t-1}表示為前次疊代之記憶體參數與輸出向量。

$$i_t = \sigma(w_i H_{t-1} + U_i X_t + b_i) \quad (1)$$
$$f_t = \sigma(w_f H_{t-1} + U_f X_t + b_f) \quad (2)$$
$$o_t = \sigma(w_o H_{t-1} + U_o X_t + b_o) \quad (3)$$
$$\bar{c}_t = tanh(w_c H_{t-1} + U X_t + b) \quad (4)$$
$$C_t = f_t C_{t-1} + i_t \bar{c}_t \quad (5)$$
$$H_t = o_t tanh(C_t) \quad (6)$$

若有$N(w)$個 LSTM 單元則整體所需記憶體，可以由(7)表示，用於計算機記憶體容量評估。其餘計算細節與程式編寫技巧，可參考 Brownlee 之說明[8]。

$$M = 4\big(N(w)(N(w)+1) + N(X_t)N(w)\big) + N(H_t)(N(w)+1) \quad (7)$$

2.2 極限梯度提升(XGBoost)

如前述之 LSTM 法功能，XGBoost 法亦擔任負載預測任務，但為監督式(Supervised)學習問題法則。利用多個特徵之輸入x_i來預測目標$\hat{y}_i$並以(8-9)式表示。為一組分類與回歸樹(classification and regression trees, CART)[9]

$$\hat{y}_i = \phi(x_i) = \sum_{j=1}^{K} f_j(x_i),\ f_j \in \mathcal{F}, \quad (8)$$

$\mathcal{F}$屬於回歸樹樹枝q與之權重$w_q(x)$。

$$\mathcal{F} = \{f(x) = w_q(x)\} \quad (9)$$

目標函數包含訓練誤差與其正則化，

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (10)$$

$l(.)$ 為誤差損失函數與$\Omega(.)$回歸樹之複雜度處罰函數[10]。利用預測資料與實際負載訓練出回歸樹，立即將訓練資料進行預測並與實際負載取得差異值，第二顆樹則以差異值訓練並得出更新差異值，進行第三顆樹等建立。優點為先以簡單模型近似出結果，並漸漸地增加模型的複雜度，以學習率參數調整個回歸樹之重要性，期望每顆樹平均地提供相對的影響。

2.3 沙普利值分析(SHapley Additive exPlanations, SHAP)

沙普利值分析[11]被用於分解並分析預測模型中每個因子的影響，SHAP 是經濟學門之賽局理論(Game theory)中採用的一種技術，用於解析合作賽局(Cooperative game theory)中每個玩家對其結果的貢獻程度[6, 12]。該理論由 Lloyd Shapley 於 1953 年提出，其以公平、合理的方式給予賽局理論嚴格的公式描述，使得賽局中擁有唯一公平的利益分配。而後沙普利值被用於解釋機器學習模型中每個局部輸入因子。

使用簡單的可解釋模型$g(z')$，局部的解釋任意的複雜模型$f(z)$，如(11)所示

$$f(z) = g(z') = \phi_0 + \sum_{j=1}^{M} \phi_j z_j' \quad (11)$$

本文有 7 種輸入因子，即M值為 7。ϕ_j值為欲求出之相對於第j項之分析因子，ϕ_0為常數項。若存在一線性模型$\hat{f}(z)$

$$\hat{f}(z) = \beta_0 + \beta_1 z_1 + \cdots + \beta_j z_j \quad (12)$$

，β為特徵權重。定義在第j項之特徵貢獻為

$$\phi_j = \beta_j z_j - E(\beta_j z_j) = \beta_j z_j - \beta_j E(z_j) \quad (13)$$

$E(\beta_j z_j)$為第j項之平均影響期望值。若預測樣本數為p個，可改寫(12)(13)式為(14)，意義為所有樣本z之因子貢獻之和等於預測值減去平均預測值。彩券博弈為例，若有j項獎項(頭獎、一獎、與普獎等等)，則該j項獎項可以賺進的利潤，可為預測購買銷售數字減去該獎項發出獎金期望值。

$$\sum_{j=1}^{p} \phi_j \hat{f}(z) = \sum_{j=1}^{p} \left(\beta_j z_j - E(\beta_j z_j)\right) = \left(\beta_0 + \sum_{j=1}^{p} \beta_j z_j\right) - \left(\beta_0 + \sum_{j=1}^{p} E(\beta_j z_j)\right) = \hat{f}(z) - E\left(\hat{f}(z)\right) \quad (14)$$

本文實驗值之p值為 168，且以上推導定義於線性系統$\hat{f}(z)$，故 XGBoost 法並不適用[13]。在下列限定的 3 個性質下，SHAP 值會有唯一解。分別為局部準確性(Local accuracy)，表示每個因子的重要度之和等於$f(z)$

的重要程度，如(11)式表示；其次為缺失性(Missingness)，表示缺失值對於因子的重要度沒有貢獻。最後為一致性(Consistency)，表示改變模型不會對特徵的重要度造成改變。滿足上述三種性質的可解釋模型$g(z')$會有唯一解，其解的形式即為 Shapley 值。

III. 預測模型設計與相對性關係因子

基於 SHAP 理論之一致性原則，本文提出以多項式模型進行預測與分析，可得出與電力負載分析一致之結果。目的為設定各預測法之張量(tensor)、疊代次數、學習率與學習樹深度等參數。再者，比較各多項式權位(因子)值，與標準值相比較，即可得公平地比較出各方法之效率與信心度。

3.1 校園用電模型設定

以校園某教學大樓之用電為模型，輸入的參數有年(2018-2019)、月(1-12)、日(1-31)、小時(0-23)、星期(1-7)、溫度(5.4-36.5˚C)與濕度(0.21-0.99%)等共計有 7 種，輸出的為負載實功率；資料時間約為 18 個月份之每小時記錄(計 13,104 筆)。並分成預測資料與測試資料，前者作為模型訓練；後者驗證預測功率。在真實的運轉預測狀況，測試資料可由學校行事曆與中央氣象局預報資料取得。均採 LSTM 與 XGBoost 法為預測法則；LSTM 法之因子分析則以 SHAP 法；XGBoost 法即可提供因子分析結果。下圖 2 為整體關係示意。

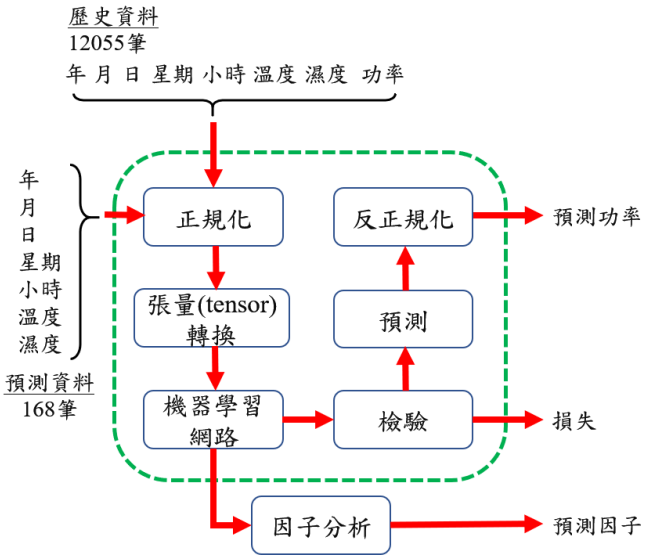

圖 3 預測程序、輸入與輸出關係

3.2 多項式模型設定

基於前述之局部準確性與缺失性，設計此模型用於驗證預測正確性與因子關係，若存在於如(15)之向量$\boldsymbol{y_j}$

$$\boldsymbol{y_j} = \begin{bmatrix} y_1 \\ \vdots \\ y_j \end{bmatrix} = \begin{bmatrix} w_{11} & \cdots & w_{1j} \\ \vdots & \ddots & \vdots \\ w_{i1} & \cdots & w_{ij} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_j \end{bmatrix} \tag{15}$$

y_j為第j個狀態時的總和，由相對的i個取樣次數與x_j個的亂數值所組成，其中w_{ij}為權位值。基於大數原則，當取樣數i越值多時，則期望值將接近於平均值，故第j個狀態變數的期望值$E[x_j]$，在有i個樣數樣本時，滿足於下式。

$$E[x_j] = \frac{1}{i}\sum_{j=1}^{100} w_{ij}x_j \tag{16}$$

配合本文所提之負載預測實驗，狀態j為 7，取樣個數i設定為 100 次，各權位值w_{ij}均設定為 1。顯見期望值$E[x_j]$應接近於 1/7，表示各狀態變數x_j均勻地組成向量$\boldsymbol{y_j}$。相對性統一因子(Unified factor)U_j則可表示為，即可表示出各狀態變數之整體貢獻度。

$$U_j = \frac{E[x_j]}{\sum_{n=1}^{j} E[x_n]} \tag{17}$$

再者，設計權位值差異測試項目，當狀態j為奇數時，權位值w_{ij}設計為$\frac{2}{7}$；偶數則為$\frac{6}{7}$。依據局部準確性原則與(17)式定義，所得出之相對性統一因子U_j應分別為 0.08 與 0.23，實驗結果詳述於後續章節。

此多項式模型亦可提供分析因子的正確性參考值。再者，經由取樣數n與期望值結果，可調整如輸入資料批次(Batch)、疊代次數、學習率與學習樹深度等參數的設定。

IV. 實驗結果

4.1 校園負載預測

任選某二段之校園負載測試資料，進行負載預測與相對性重要關係U_j。預測法則之細節設定則整理入下表 1。

表 1 預測法則之細節設定比較

LSTM	訓練輸入 張量維度	訓練輸出 矩陣維度	疊代次數
	12051x4x7	12051x1	50
XGBoost	訓練輸入 矩陣維度	訓練輸出 矩陣維度	樹個數/ 學習率
	12055x7	12055x1	32000/ 0.0080

下圖 4 與圖 5 分別為 168 小時之校園負載實際值、LSTM 法與 XGBoost 法之預測結果。與實際值相比，此二種方法都能達成預測但無法完全相符。時間區域 1，各方法之方均根偏移(Root mean square deviation)分別為 56.43 與 63.56；時間區域 2，則分別為 94.41 與 52.35。

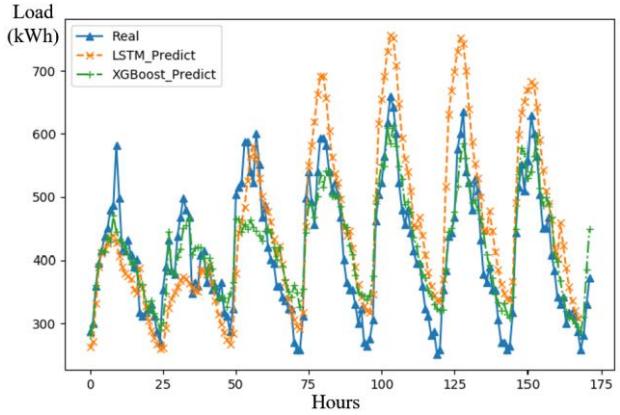

圖 4 實際負載(藍色)、LSTM 法預測(橘色)與 XGBoost 法預測之比較

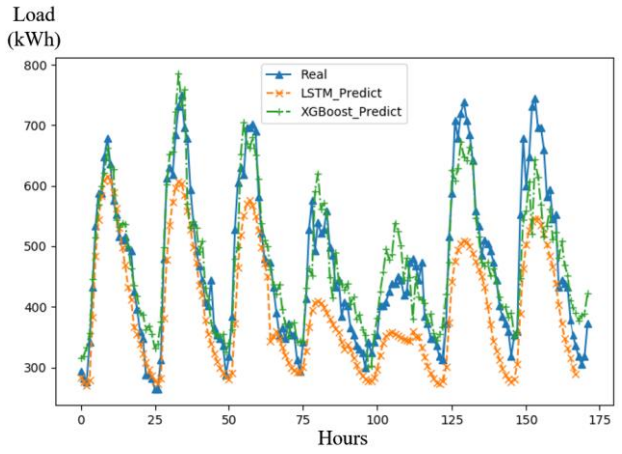

圖 5 實際負載(藍色)、LSTM 法預測(橘色)與 XGBoost 法預測之比較

相對性關係因子U_j分析結果於圖 6 與圖 7 表示。於時間區域 1 與 2 內，LSTM 法之最大統一因子為小時(Hour) 且分別為 0.45 與 0.42；XGBoost 法則為月份(Month)且分別均為 0.23。較不相同的是 LSTM 法完全無任何年(Year)與月(Month)的關聯性；XGBoost 法則其統一因子則有較多的月份因子考慮。

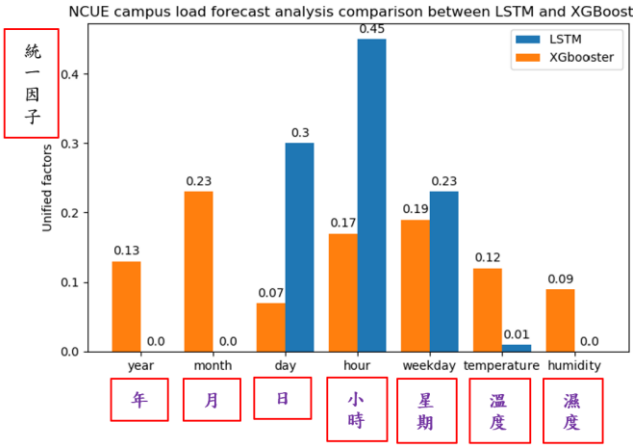

圖 6 LSTM 法與 XGBoost 法之分析因子結果於時間區域 2

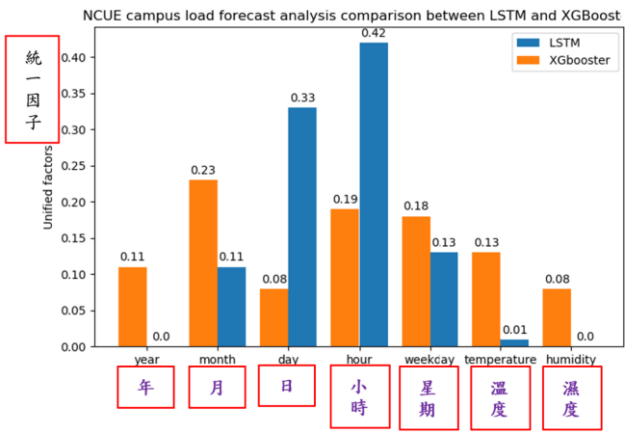

圖 7 實際負載(藍色)、LSTM 法預測(橘色)與 XGBoost 法預測之比較於時間區域 2

4.2 多項式函數測試

以亂數產生出與上表 1 相同之輸入資料維度與設定，並且進行 100 次的估測實驗。每次實驗需重新產生資料以確保其獨立性。擷取某段之真實(Real)且由亂數產生之多項式與預測法結果於圖 8 所示，二種估測法與真實多項式結果非常相似，曲線幾乎重疊且無明顯差異。

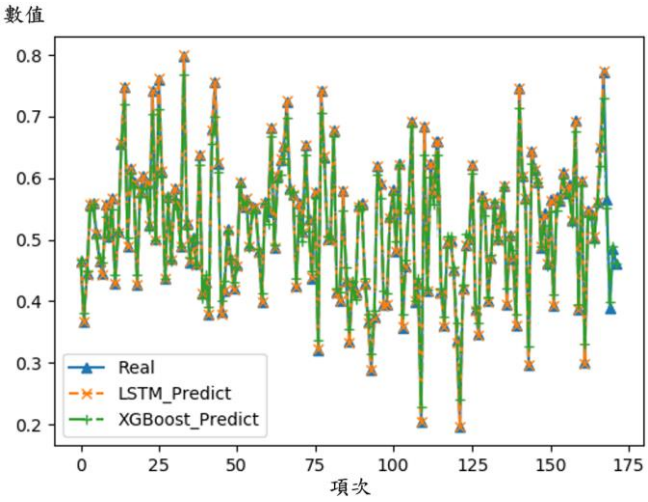

圖 8 多項式函數(藍色)、LSTM 法預測(橘色)與 XGBoost 法預測之比較

此多項式模型不僅可驗證由 LSTM、XGBoost 與 SHAP 計算出的預測模型，亦可提供分析因子的正確性參考值。再者，經由取樣數n與期望值結果，可調整如輸入資料批次(Batch)、疊代次數、學習率與學習樹等參數的設定。表 2 整理出模型訓練與因子分析時間，LSTM 法需要平均 1015.009 秒完成模型訓練；較 XGBoost 法多出 654.279 秒。因子分析亦多需要 5.526 秒時間。測試之各狀態變數的平均貢獻度等分析數值整理於表 3。可知 LSTM 法與 XGBoost 法均能提供約 14.1%誤差之分析，但 LSTM 法提供較一致性，即低標準差之分析結果。

表 2 100 次多項式計算平均時間

	模型訓練與估測時間(秒)	因子分析時間(秒)	合計(秒)
LSTM	1015.009	5.856	1020.865
XGBoost	360.730	0.330	361.060

圖 9 為以不同權位函數測試多項式模型法之結果，奇數權位設定為 0.08；偶數權位設定為 0.23。以同樣之資料格式與訓練模型參數，可得 LSTM 法分別得出奇數與偶數之平均統一性因子為 0.08 與 0.23；同樣地，XGBoost 法法分別得出奇數與偶數之平均統一性因子為 0.045 與 0.27。相較於設定值，LSTM 法較 XGBoost 法能求得較精確之統一性因子數值。

表 3　100 次多項式計算測試之各因子之貢獻度、平均值、誤差比與標準差

方法	x_1	x_2	x_3	x_4	x_5	x_6	x_7	平均值	誤差比(%)	標準差
LSTM	1.137	1.145	1.137	1.139	1.149	1.149	1.141	1.141	14.1	0.00516
XGBoost	1.112	1.177	1.113	1.148	1.037	1.243	1.154	1.143	14.3	0.06376

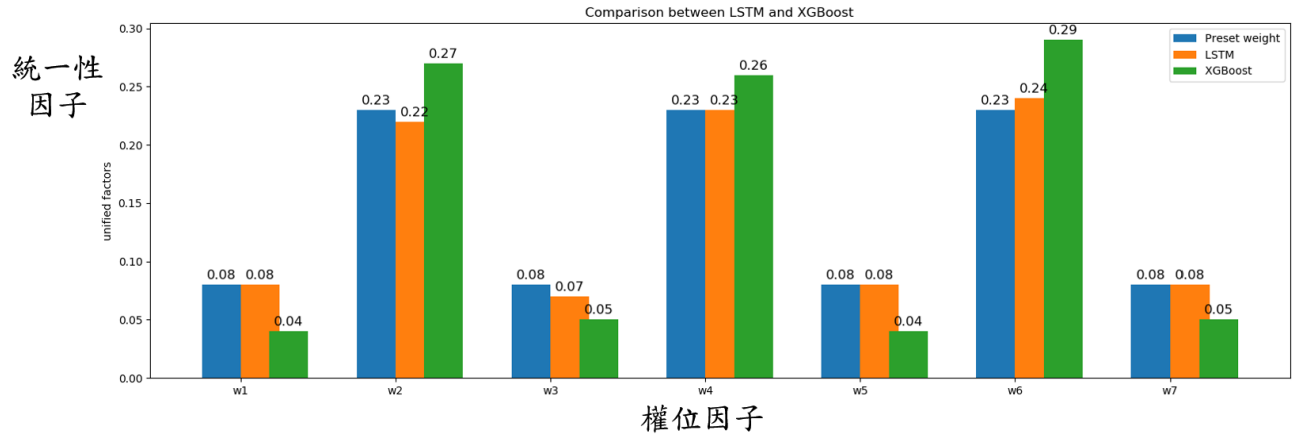

圖 9 多項式函數(藍色)、LSTM 法預測(橘色)與 XGBoost 法預測之不同權位值測試

V. 結論

本文首先說明 LSTM 法與 XGBoost 法之基本原理；隨之，完成校園電力負載預測的目的，並且以 SHAP 法分析所組成因子條件。本文提出以多項式模型驗證因子的誤差值，以提供資料參考之信度。

LSTM 法屬於深度學習，加入記憶與遺忘單元之神經網路實現；XGBoost 法為決定樹方式實現。二種分析法均能達到 56.43、63.56、94.41 與 52.35 之方均根偏移值。XGBoost 法平均僅需 360.730 秒即可完成模型訓練；反之，LSTM 法需要 1015.009 秒的時間。各因子貢獻度分析，二種方法均有約 1.14 之多項式因子平均值，但以 LSTM 法有較小之標準值表現，且權位值差異測試，亦得到 LSTM 法與設定值相差甚小。

本文所得結果有助於未來進一步研究影響負載變化主要因素，進而提供更準確的預測技術。

誌謝

本研究承蒙核能研究所計畫編號：109A010 之經費補助研究計畫之經費補助，特此感謝。

參考文獻

[1] Y.-D. Lee, J.-L. Jiang, Y.-H. Ho, W.-C. Lin, H.-C. Chih, and W.-T. Huang, "Neutral Current Reduction in Three-Phase Four-Wire Distribution Feeders by Optimal Phase Arrangement Based on a Full-Scale Net Load Model Derived from the FTU Data," *Energies 2020,* vol. 13, no. 1844. doi: https://doi.org/10.3390/en13071844. p. 20, 2020.

[2] M. Dong and L. Grumbach, "A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction," *IEEE Transactions on Smart Grid,* vol. 11, no. 1. doi: 10.1109/TSG.2019.2924183. pp. 470-482, 2020.

[3] Y. Hong, J. J. F. Martinez, and A. C. Fajardo, "Day-Ahead Solar radiation Forecasting Utilizing Gramian Angular Field and Convolutional Long Short-Term Memory," *IEEE Access,* vol. 8, doi: 10.1109/ACCESS.2020.2967900. pp. 18741-18753, 2020.

[4] J. Ospina, A. Newaz, and M. O. Faruque, "Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model," *IET Renewable Power Generation,* vol. 13, no. 7. doi: 10.1049/iet-rpg.2018.5779. pp. 1087-1095, 2019.

[5] Y. Liu, H. Luo, B. Zhao, X. Zhao, and Z. Han, "Short-Term Power Load Forecasting Based on Clustering and XGBoost Method," in *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, 2018. doi: 10.1109/ICSESS.2018.8663907. pp. 536-539.

[6] M. Wang, K. Zheng, Y. Yang, and X. Wang, "An Explainable Machine Learning Framework for Intrusion Detection Systems," *IEEE Access,* vol. 8, doi: 10.1109/ACCESS.2020.2988359. pp. 73127-73141, 2020.

[7] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation,* vol. 9, no. 8. doi: 10.1162/neco.1997.9.8.1735. pp. 1735-1780, 1997.

[8] J. Brownlee, *Long Short-term Memory Networks with Python: Develop Sequence Prediction Models with Deep Learning*. 2017.

[9] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*. 2016, pp. 785-794.

[10] S. Jansen, "Hands-On Machine Learning for Algorithmic Trading," (in Undetermined), doi: 2018.

[11] S. Lundberg, G. Erion, H. Chen, A. DeGrave, J. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, *Explainable AI for Trees: From Local Explanations to Global Understanding*. 2019.

[12] M. Sundararajan and A. Najmi, *The many Shapley values for model explanation*. 2019.

[13] M. Ancona, E. Ceolini, C. Oztireli, and M. Gross, "Towards better understanding of gradient-based attribution methods for Deep Neural Networks," presented at the 6th International Conference on Learning Representations (ICLR 2018), Vancouver, BC, Canada, April 30 - May 03, 2018, 2018. doi: 10.3929/ethz-b-000249929. Available: https://openreview.net/forum?id=Sy21R9JAW