RMIT UNIVERSITY

Multivariate Electricity Consumption Prediction with Extreme Learning Machine

Hui Song

Supervised by Dr. Flora Salim, Dr. Kai Qin

{hui.song,flora.salim,kai.qin}@rmit.edu.au

Computer Science and Information Technology
RMIT University

26th July 2016

Outline

Background

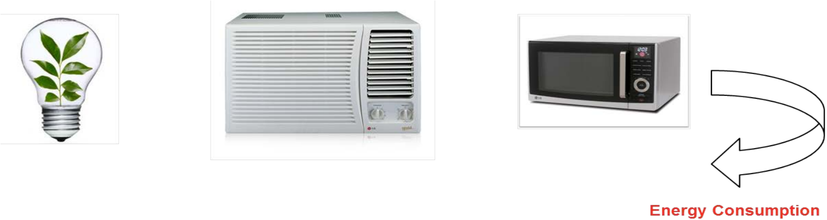

Figure 1: Appliances with Electricity consumption

Problems of electricity usage:

- Exhaustion of energy resources
- Heavy impact on the environment
- Threat for the sustainable development

RMIT UNIVERSITY

Background

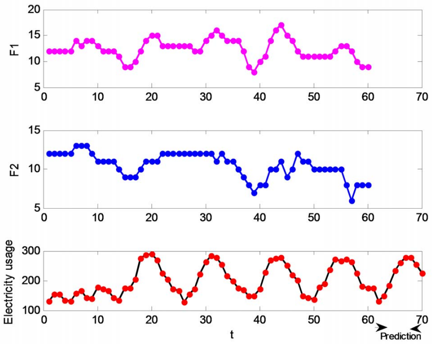

Figure 2: Multivariate time series for predicting electricity usage

Multivariate electricity consumption prediction:

- Thermal comfort
- Temperature, wind speed etc.
- Windows, floor etc.
- ...

Limitation and problems:

- Small buildings
- Feature overlapping
- Window size overlapping

Related works

- Methodologies for electricity consumption prediction
  - Statistical methods
  - Neural Network (NN)
  - Support Vector Machines (SVM)
  - ...
- Evolutionary algorithms for electricity consumption prediction
  - PSO used to optimize SVM
  - PSO applied for optimizing RBFN
  - GA utilized to tune all parameters for NN
  - ...

Problem Definition

Data collection

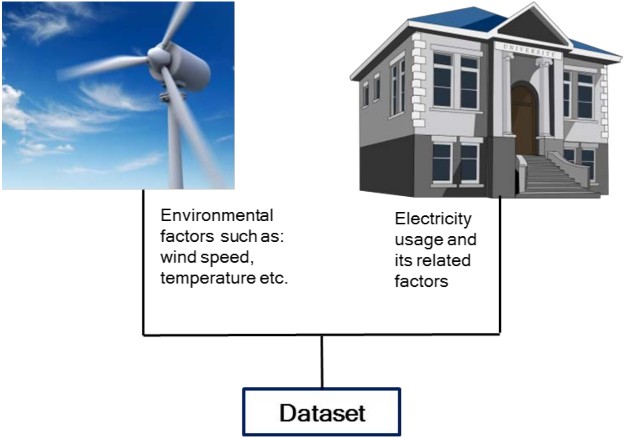

Figure 3: Multivariate time series for predicting electricity usage

Problem Definition

Data collection

- Electricity usage and its related factors
  - Sustainable Urban Precincts Program (SUPP)
  - RMIT University buildings with 15-minutes to 1-day
  - https://www.utiliview.com.au/login.aspx
- Environmental factors
  - Online weather station
  - Broadcast every 20-minutes to 1-hour
  - https://www.wunderground.com/history

Problem Definition

Proposed Problem

How to perform multiple time series prediction based on historical data and auxiliary factors?

1. Auxiliary factors enumeration
2. Optimal subset of auxiliary factors and respective parameters

Problem Definition

Challenge

1. Baseline setting
2. Prediction model selection
3. Time consuming with more parameter settings

Contribution

1. Explore the influence of auxiliary factors
2. Apply for other multiple time series prediction

Problem Definition

Extreme learning machine

- Theory

  $(\mathbf{x}_i, \mathbf{y}_i)$ with M distinct samples, satisfied $\mathbf{x}_i \in \mathcal{R}^{d1}$ and $\mathbf{y}_i \in \mathcal{R}^{d2}$, the structure of ELM with N hidden neurons perfectly approximates to the given output as:

$$\sum_{i=1}^{N} \beta_i \mathbf{f}(\mathbf{w}_i^T \mathbf{x}_j + b_i) = \mathbf{y}_j, 1 \leq j \leq M \tag{1}$$

  f: activation function; $\mathbf{w}_i$: the weight for connecting input layer and hidden layer; b_i: bias; β_i: output weight

RMIT UNIVERSITY

Problem Definition

Eq. 1 can be written as $\mathbf{HB} = \mathbf{Y}$, $\mathbf{H}$ can be represented as:

$$\mathbf{H} = \begin{pmatrix} f(\mathbf{w}_1^T \mathbf{x}_1 + b_1) & \cdots & f(\mathbf{w}_N^T \mathbf{x}_1 + b_N) \\ \cdots & \cdots & \cdots \\ f(\mathbf{w}_1^T \mathbf{x}_M + b_1) & \cdots & f(\mathbf{w}_N^T \mathbf{x}_M + b_N) \end{pmatrix} \tag{2}$$

$\mathbf{B} = (\beta_1^T, \beta_2^T, ..., \beta_N^T)^T$ and $\mathbf{Y} = (y_1^T, y_2^T, ..., y_M^T)^T$.

Problem Definition

- Characteristics
  - One parameter to be adjusted
  - Extremely fast speed
- Applications
  - Electricity price forecasting
  - Wind power density prediction
  - ...

Problem Definition

Approach for solving the first subproblem

- Baseline setting with purely historical electricity consumption

$$x_{t+1} = f(x_t, x_{t-1}, ..., x_{t-q-1}) \tag{3}$$

- Find which factors influence the prediction accuracy the most among all electricity-related factors

$$x_{t+1} = f(x_t, x_{t-1}, ..., x_{t-q-1}, r_t^i, r_{t-1}^i, ..., r_{t-q-1}^i), \quad i = 1, 2, ..., m \tag{4}$$

RMIT UNIVERSITY

Problem Definition

Approach for solving the first subproblem

- Find which of them help to improve the prediction accuracy for this problem with the influence of electricity-related factors

$$x_{t+1} = f(x_t, x_{t-1}, ..., x_{t-q-1}, r^c_t, r^c_{t-1}, ..., r^c_{t-q-1}; \\ z^j_t, z^j_{t-1}, ..., z^j_{t-q-1}), j = 1, 2, ..., n \qquad (5)$$

q: window size; $f()$: predictor
m and n:number of electricity-related factors and environmental factors respectively
c: the most important electricity-related factors

Problem Definition

Approach for solving the second subproblem

- Discrete Dynamic Multi-Swarm Particle Swarm Optimization (DDMS-PSO)

  Velocity and position are updated as:

$$v_{id}^{k+1} = \omega {v_{id}}^{k} + c_1 r_1 (pbest_{id}^{k} - X_{id}^{k}) + c_2 r_2 (lbest_{pd}^{k} - X_{id}^{k}),$$
$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \quad (6)$$

  Then the position will be mapped to the real values:

$$x_{id} = round(x_{id}), 0 < d =< d_1$$
$$x_{id} = (X_{Trmax} - X_{Trmin})x_{id} + X_{Trmin}, d_1 < d < D$$
$$x_{id} = 50fix(((X_{rmax} - X_{rmin})x_{id} + X_{rmin})/50), d = D \quad (7)$$

Problem Definition

Approach for solving the second subproblem

- Dynamically find optimal subset of auxiliary factors and its parameters

$$x_{t+1} = f(x_t, x_{t-1}, ..., x_{t-q-1}, r_t^i, r_{t-1}^i, ..., r_{t-q_i-1}^i; \\ z_t^j, z_{t-1}^j, ..., z_{t-q_j-1}^j) \qquad (8)$$

q, q_i and q_j are different window sizes and $i \in [1, m]$, $j \in [1, n]$.

Experiment

Data description

- All datasets from 01.06.2014 to 31.05.2015
- 12 points in one day
- Normalized to $[0, 1]$
- Sampled according to the number of points in one day with different window sizes

Evaluation method

- Cross validation
- 10 10-fold

Experiment

Factors expression

- Electricity usage: E
- Electricity-related factors: A_1 (apparent power), A_2 (power factor)
- Environmental factors: F_1 (temperature), F_2 (dew point), F_3 (humidity), F_4 (wind speed), F_5 (sea level)
- 1 and 0: selected or not

Experiment

Experiment 1 (for solving the first subproblem)

- Parameter setting
  1. ELM parameter settings
     The number of hidden neurons: 50, 100, 200, 300, 400, 500,800,1000
     Activation function: Sigmoid function
  2. SVR parameter settings
     Kernel type: Radial Basis Function
     C (Cost): 2^{-5} 1 2^{5} 2^{10} 2^{15}
     ϵ: 0.01 0.03 0.05 0.08 0.1 0.15 0.18 0.2

Experiment

- Experimental result

Table 1: The comparison of elm and svr with only historical electricity consumption

			RMSE result under different window sizes							
			5	10	15	20	25	30	35	40
ELM	Number of hidden neurons	50	0.0698	0.0645	0.0632	0.0633	0.0639	0.0645	0.0651	0.0658
		100	0.0702	0.0644	0.0624	**0.0621**	0.0627	0.0630	0.0631	0.0633
		200	0.0835	0.0663	0.0629	**0.0622**	0.0630	0.0630	0.0631	0.0632
		300	0.1211	0.0697	0.0648	0.0637	0.0643	0.0638	0.0641	0.0640
		400	0.2316	0.0757	0.0680	0.0657	0.0663	0.0654	0.0657	0.0652
		500	0.4544	0.0860	0.0725	0.0684	0.0690	0.0674	0.0675	0.0670
SVR	Different parameters settings	H_1	0.0735	0.0671	0.0643	0.0639	0.0640	0.0642	0.0640	0.0638
		H_2	0.0726	0.0662	0.0638	0.0633	0.0636	0.0636	0.0634	0.0633
		H_3	0.0721	0.0656	0.0635	0.0632	0.0632	0.0632	0.0631	0.0629
		H_4	0.0721	0.0652	0.0635	0.0631	0.0632	0.0633	0.0630	0.0630
		H_5	0.0774	0.0686	0.0669	0.0662	0.0662	0.0662	0.0657	0.0657
		H_6	0.0828	0.0728	0.0708	0.0702	0.0701	0.0702	0.0692	0.0699

Compared with SVR, ELM shows a better prediction ability on this prediction problem.

RMIT UNIVERSITY

Experiment

Experiment 2 (for solving the first subproblem)

- Experimental result

Table 2: Subset of electricity-related factors with the best prediction performance

		Electricity-related factor		Number of hidden neurons	RMSE
		A_1	A_2		
Window sizes	15	1	1	200	**0.0536**
	20	1	1	200	**0.0536**

Each of the electricity-related factors can help improve the prediction accuracy but when they are combined, it will lead to better performance.

RMIT UNIVERSITY

Experiment

Experiment 3 (for solving the first subproblem)

- Experimental result

Table 3: Subsets of environmental factors for improving prediction accuracy mostly

		Environmental factors					Number of hidden neurons	RMSE
		F_1	F_2	F_3	F_4	F_5		
Window sizes	15	1	1	0	1	0	200	**0.0491**
	15	1	1	0	1	0	300	**0.0492**

Under the influence of electricity-related factors, a subset of environmental factors can further and mostly improve the prediction performance.

Experiment

Experiment 4 (for solving the second subproblem)

- parameter settings
  DDMS-PSO parameter settings
  $\omega = 0.729$, $c_1 = c_2 = 1.49445$, $R = 10$, $n = 30$, $ns = 10$,
  $p = 3$, $Max\_FEs = 1000$, $D = 16$
  X_{Trmax} and X_{Trmin}: 5 and 35 respectively
  X_{rmax} and X_{rmin}: 50 and 600 with an interval of 50

RMIT UNIVERSITY

Experiment

Experiment 4

- Experimental result

Table 4: optimal subset of all factors, respective window sizes and number of hidden neurons based on DDMS-PSO

	Auxiliary factors							Window sizes								Number of hidden neurons	RMSE
	F_1	F_2	F_3	F_4	F_5	A_1	A_2	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8		
1	1	1	0	1	0	1	0	21	9	0	15	0	19	0	12	250	**0.0483**
2	1	1	0	1	0	1	0	22	8	0	13	0	19	0	12	250	**0.0483**
3	1	1	0	1	0	1	0	21	6	0	16	0	19	0	12	250	**0.0483**
4	1	1	0	1	0	1	0	28	6	0	15	0	19	0	12	250	**0.0483**
5	1	1	0	1	0	1	0	21	11	0	15	0	19	0	12	250	**0.0483**
6	1	1	0	1	0	1	1	20	8	0	10	0	18	23	12	250	**0.0483**
7	1	1	0	1	0	1	0	20	7	0	14	0	19	0	12	250	**0.0483**
8	1	1	0	1	0	1	0	23	5	0	14	0	19	0	12	250	**0.0483**
9	1	1	0	1	0	1	0	21	6	0	21	0	19	0	13	250	**0.0483**
10	1	1	0	1	0	1	0	26	19	0	14	0	19	0	12	250	**0.0484**

DDMP-PSO can find an optimal subset of auxiliary factors, their respective window sizes and the number of hidden neurons in ELM which can lead to the best prediction performance.

RMIT UNIVERSITY

Conclusion

- ELM shows superiority over SVR on electricity consumption prediction.
- Compared with purely electricity usage, electricity-related and environmental factors were demonstrated to improve prediction.
- The proposed DDMS-PSO can find the optimal subset of electricity-related factors and environmental factors and their parameters respectively for obtaining the best prediction accuracy.
- We will use feature extraction for further improving the prediction performance in the future.