Data Mining II
Regression Analysis

Zijun Zhang

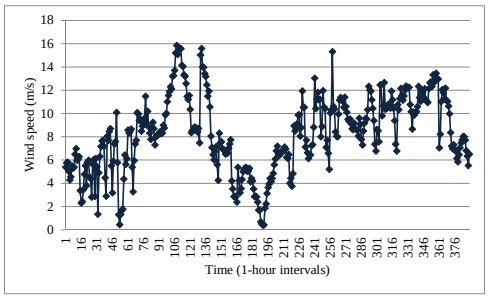
Content

- Prediction: Wind speed and wind power
- Training Power Curve Model
- Weka Software
- Extract Results

Prediction

- Time Series Data – a sequence of data points measured typically at successive time instants spaced at uniform time intervals.

Example



Prediction

- Basic Analysis Technique for Time Series Data – Exponential Smoothing
- Single Exponential and Double Exponential
- Successful application: Demand Forecast

Prediction

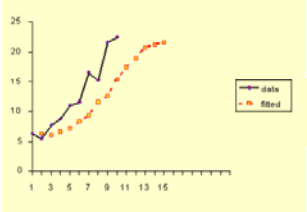
- Single Exponential Smoothing

$$s_0 = x_0$$

$$s_t = \alpha x_{t-1} + (1-\alpha)s_{t-1}$$

where α is the smoothing factor, $0 < \alpha < 1$

Example



Prediction

- Double Exponential Smoothing

$$\hat{\upsilon}_t = \alpha\upsilon_t + (1-\alpha)(\hat{\upsilon}_{t-T} + b_{t-T})$$

$$b_t = \gamma(\hat{\upsilon}_t - \hat{\upsilon}_{t-T}) + (1-\gamma)b_{t-T}$$

$$b_0 = (\sum_{t=-5}^{-2}\upsilon_t - \sum_{t=-4}^{-1}\upsilon_t)/4$$

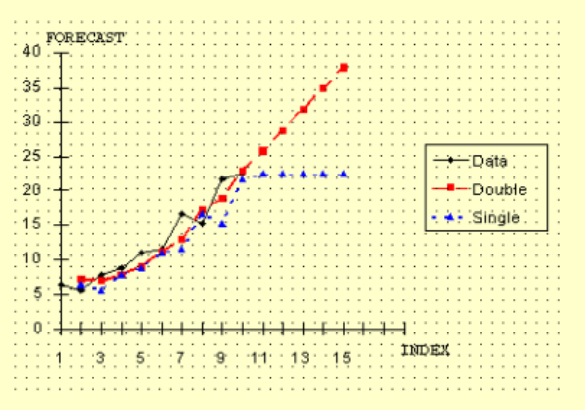
$$\hat{\upsilon}_0 = \upsilon_{-1}$$

$$\hat{\upsilon}_{t+T} = \hat{\upsilon}_t + b_t$$

$\hat{\upsilon}$	predicted wind speed
υ	observed wind speed
t	current time
T	sampling time
n	number of steps
α	smoothing constant
γ	smoothing constant
b	factor adjusting trend

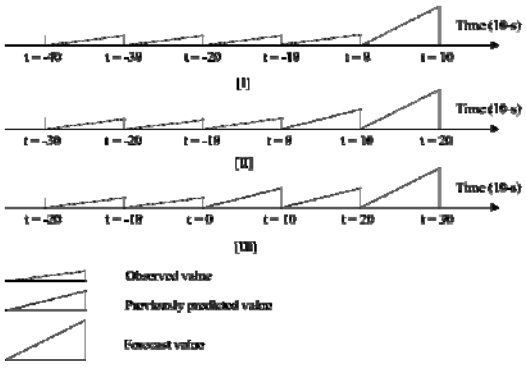
Prediction

Double Exponential Smoothing – Example



Prediction

- Data Driven Time Series Analysis



Prediction

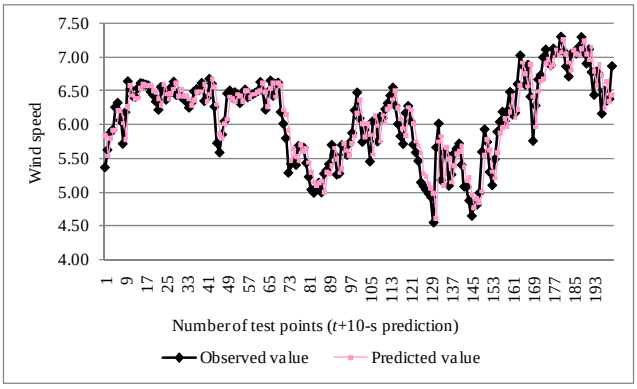
Equation:

$$\hat{\upsilon}_t = f_A(\upsilon_{t-T}, \upsilon_{t-2T}, \upsilon_{t-3T}, \ldots \upsilon_{t-nT})$$

$\hat{\upsilon}$	predicted wind speed
υ	observed wind speed
t	current time
T	sampling time
n	number of steps
$f_A(\bullet)$	data-driven model
A	data-mining algorithms

Prediction

The first 300 points of t + 10-s predictions by a neural network model



Power Curve Model

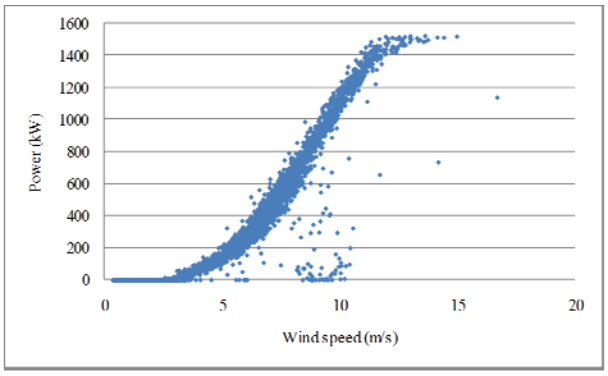
Parametric model of power output:

$$P = \frac{1}{2}\rho\pi R^2 C_p(\lambda, \beta)v^3$$

Data-derived model of power output:

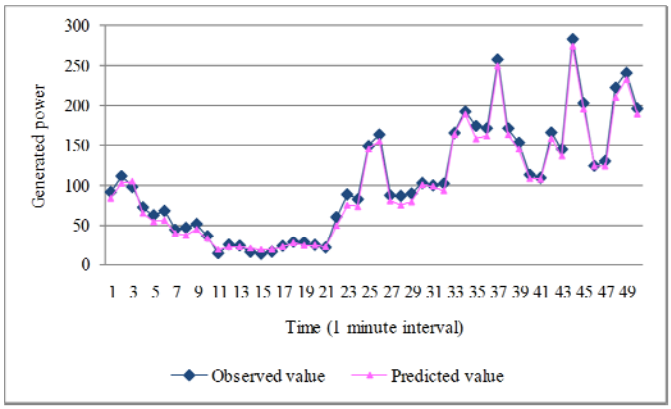
$$y_3(t) = f_3(v_1(t), v_1(t-1), x_1(t), x_1(t-1), x_2(t), x_2(t-1))$$

Power Curve Model



Power Curve Model

- Example – Prediction Results



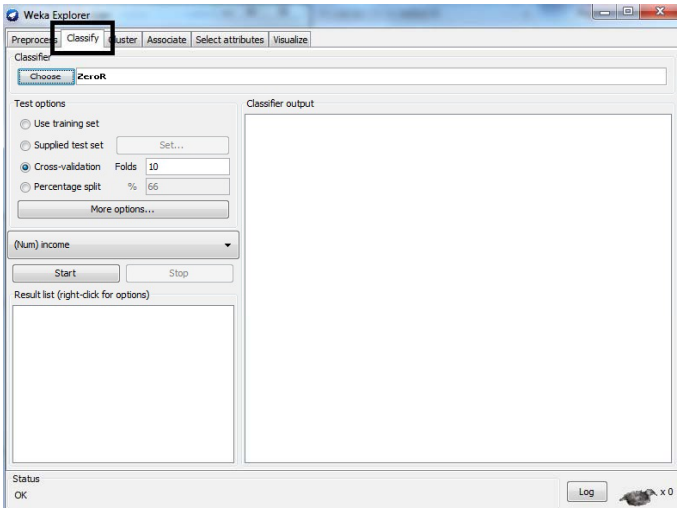
Weka Software

Three considered data-mining algorithms

- Neural Network
- Support Vector Machine
- K-Nearest Neighbor

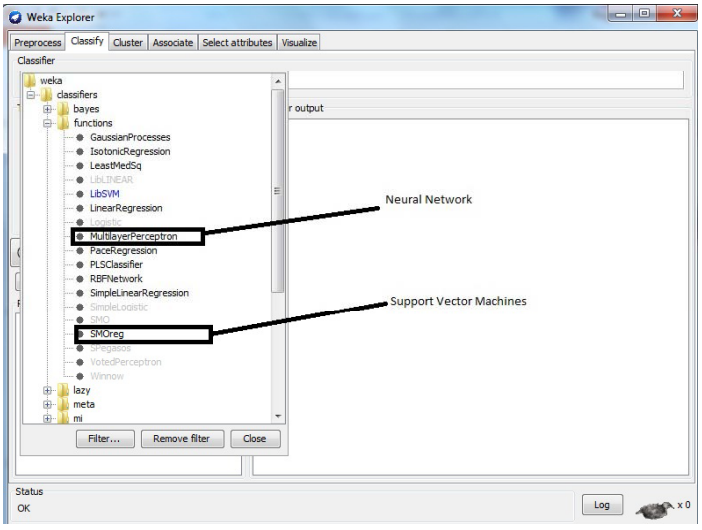
Weka Software

Step 1



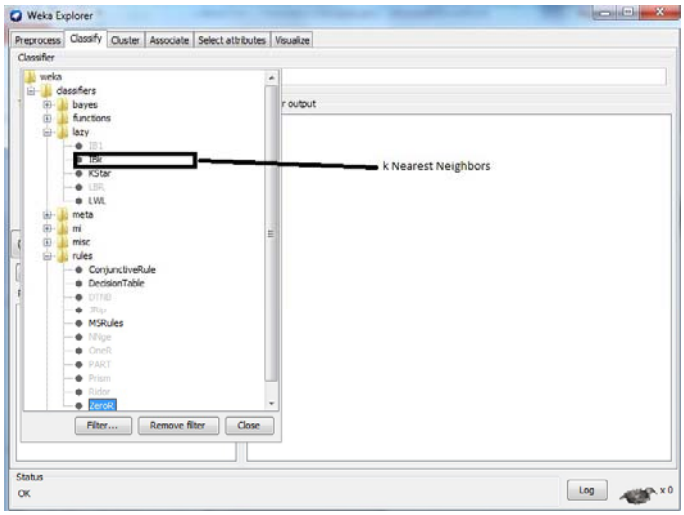
Weka Software

Step 2.1



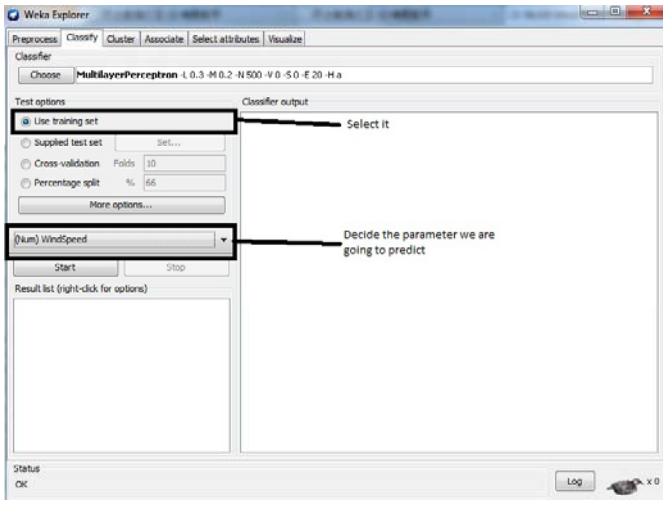
Weka Software

Step 2.2



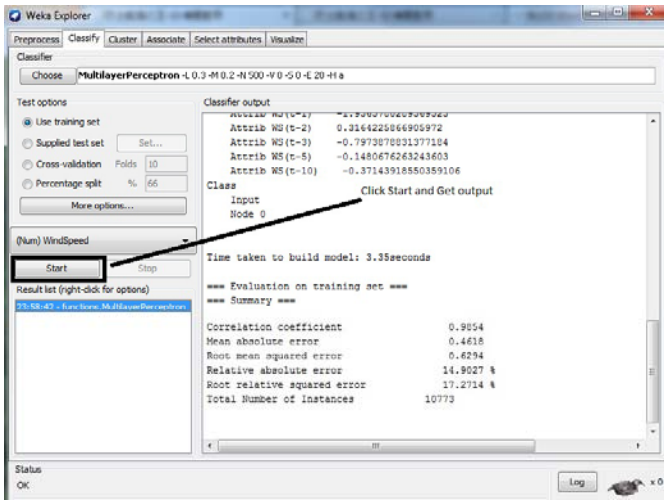
Weka Software

Step 3



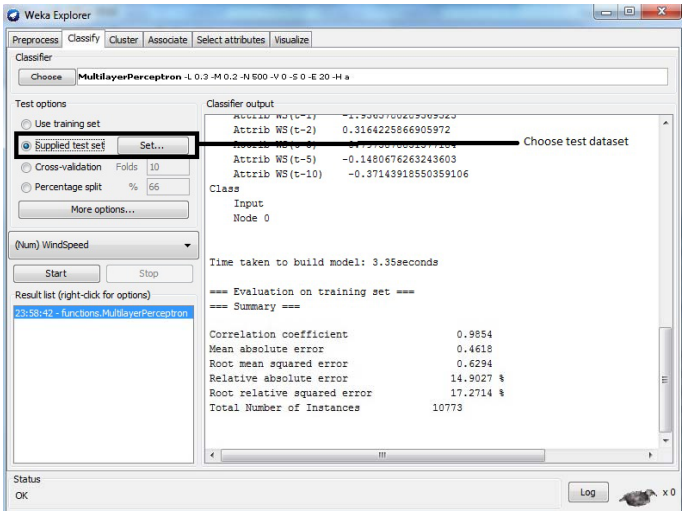
Weka Software

Step 4



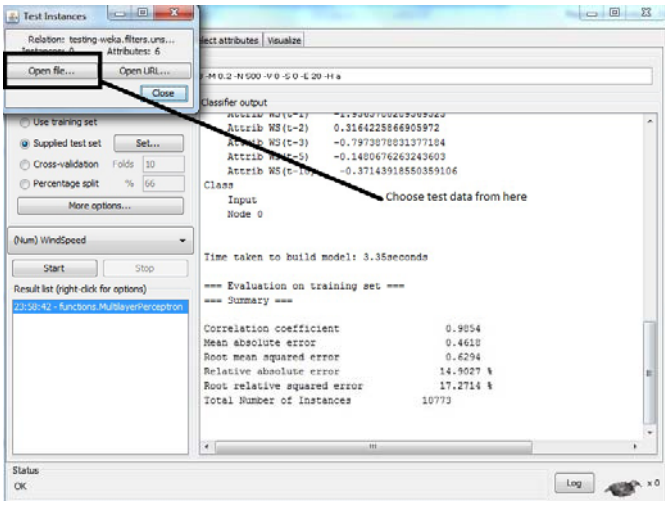
Weka Software

Step 5



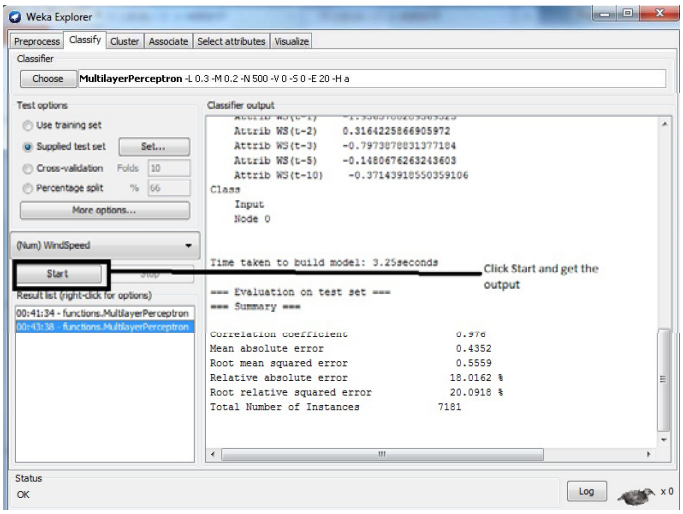
Weka Software

Step 6



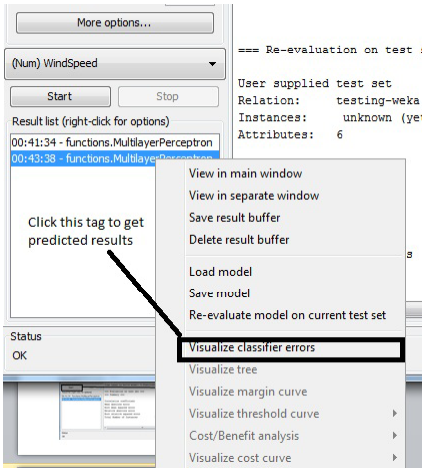
Weka Software

Step 7



Extract Results

Step 1



Extract Results

Step 2