

Các nghiên cứu ứng dụng AI vào bài toán dự báo KTTV

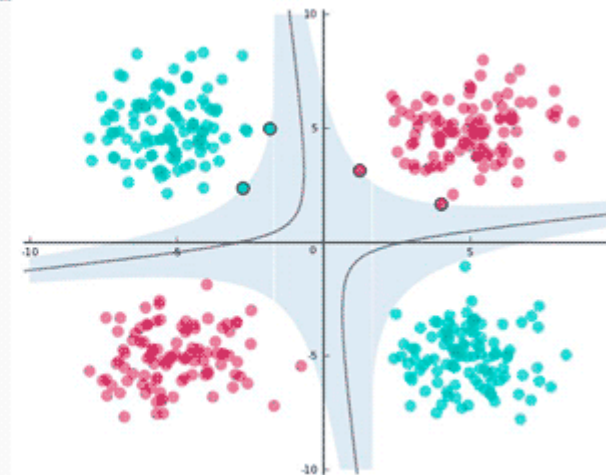
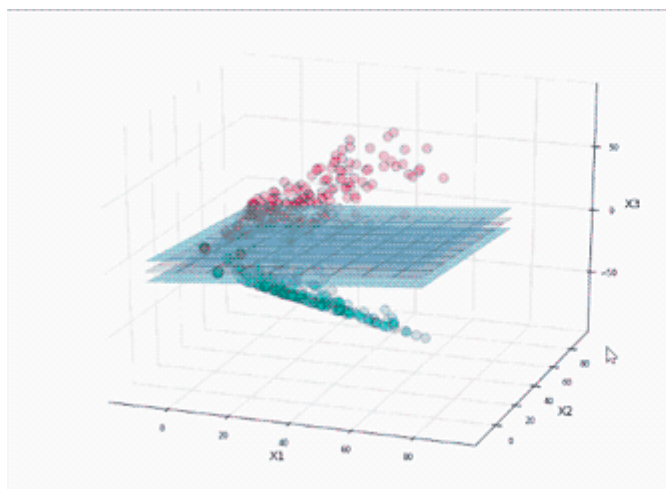
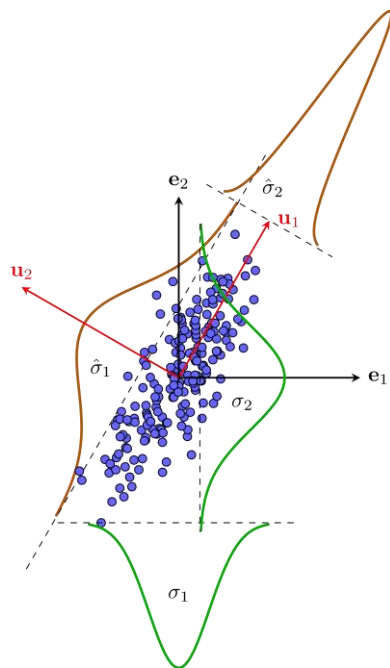


ARS Viet Nam – Tháng 8, 2019

- **Optimal parameters of the SVM for temperature prediction**
 - XIAOYU SHI, QIANG HUANG, JIANXIA CHANG, YIMIN WANG, JIANGQUN LEI & JING ZHAO State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an, 710048, P.R. China
- **Bài toán:** Dự đoán nhiệt độ trung bình tại trạm Foping station (China)
- **Data:** Dữ liệu hàng tháng từ 1960 đến 2010
 - Dữ liệu khí hậu quy mô toán cầu: United States Prediction Center (NCEP) (độ phân giải 2.5x2.5)
 - Dữ liệu khí hậu khu vực: Dữ liệu đo lưu vực sông Vị (Wei River)
 - AIR850, AIR500, HGT850, **HGT500**, Omega850, Omega500, PRES, **SLP**, VWND
 - Nhiệt độ trung bình hàng tháng tại trạm Foping station (China)

- Kỹ thuật

1. Ứng dụng PCA để giảm chiều dữ liệu, tìm ra những yếu tố thời tiết chính.
2. Ứng dụng SVM để mô hình hóa dữ liệu nhiệt độ hàng tháng và hàng năm



$$\begin{aligned} X_1 &= x_1^2 \\ X_2 &= x_2^2 \\ X_3 &= \sqrt{2}x_1x_2 \end{aligned}$$

• Kết quả

1. Độ đo đánh giá kết quả mô hình dự báo

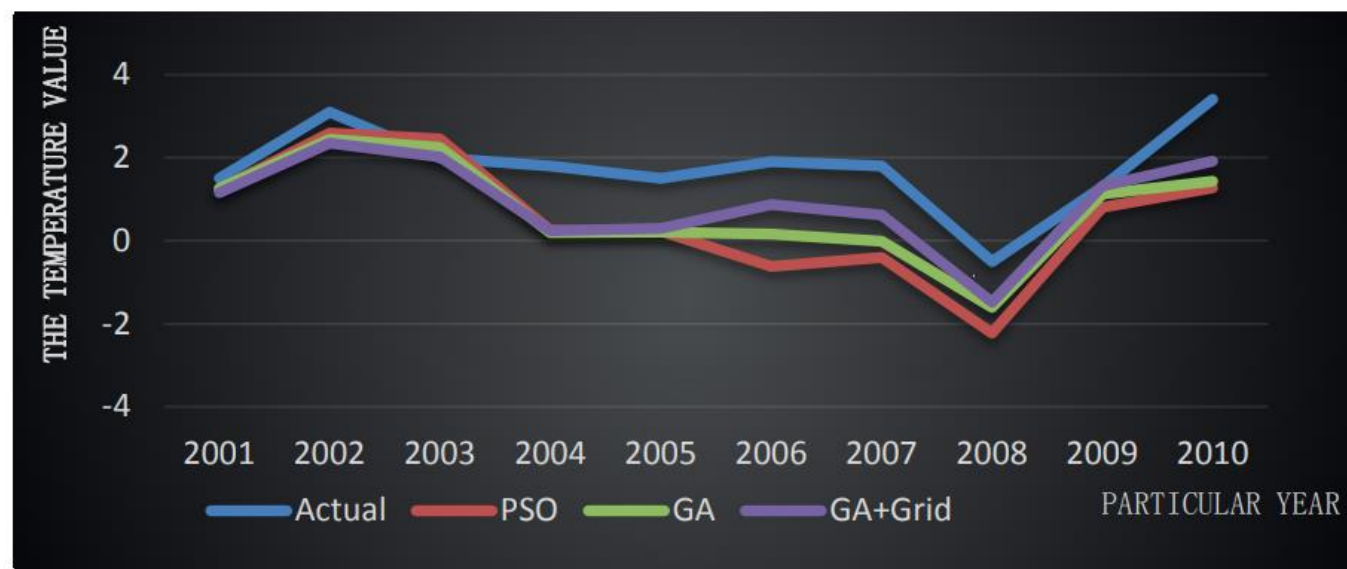
- x_i là kết quả dự đoán của mô hình
- y_i là giá trị nhiệt độ trung bình đo được trong thực tế tại trạm Foping

2. Kết quả

- Đánh giá 3 thuật toán tối ưu tham số của SVM:
 - Particle Swarm Optimization
 - Genetic Algorithm
 - Genetic Algorithm + Grid Search

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_{10} - y_{10})^2}$$

$$d_{PSO} = 4.02933, d_{GA} = 3.383041, d_{GA+grid} = 2.67989$$



- **Rainfall Prediction: A Deep Learning Approach**

- Emilcy Hernandez

- **Bài toán:** Dự báo lượng mưa trong ngày tiếp theo tại Manizales (Colombia) sử dụng dữ liệu thời tiết các ngày trước đó.

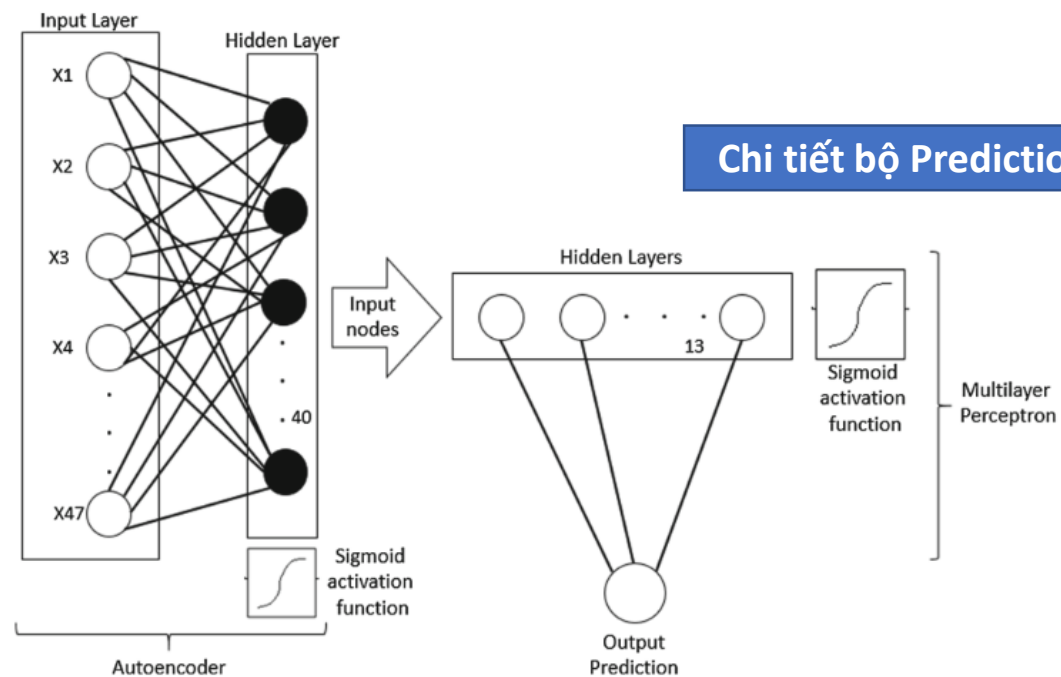
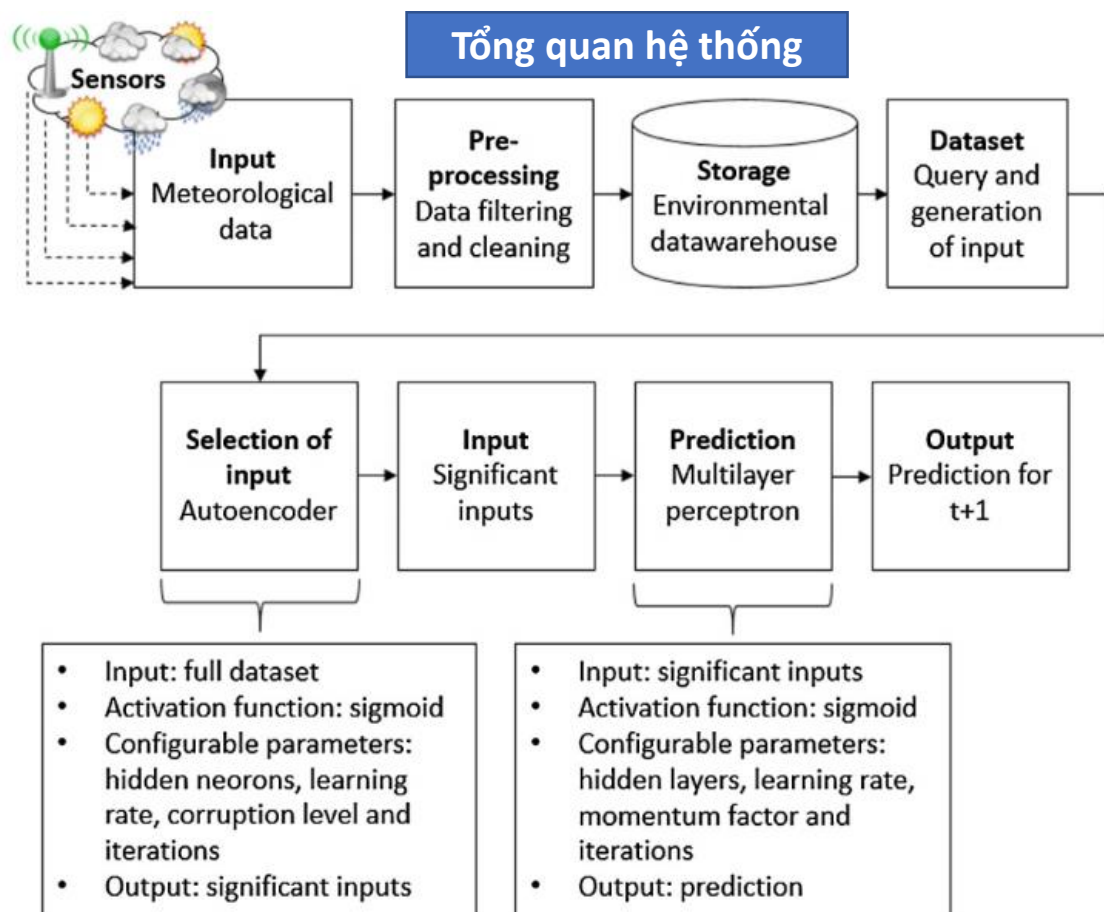
- **Data:** Dữ liệu thời tiết hàng ngày từ 2002 đến 2013 tại trạm đo ở Manizales (Colombia).

- Lượng mưa, nhiệt độ, áp suất, độ ẩm, tốc độ gió, lượng ánh sáng mặt trời

- Giá trị đo đạc trong **3 ngày trước** và **trung bình của 5 ngày trước**

- Giá trị **chênh lệch trong ngày** (4:00 và 24:00)

• Kỹ thuật



Chi tiết bộ Auto Encoder

• Kết quả

- Độ đo đánh giá kết quả mô hình dự báo Mean Square Error (MSE) và Root Mean Square Error (RMSE)

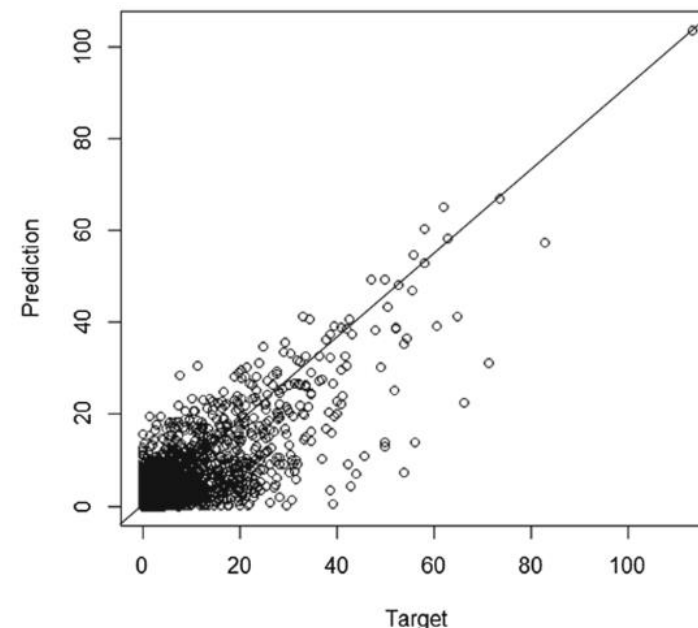
- \hat{Y}_i là kết quả dự đoán của mô hình
- Y_i là giá trị thực tế

2. Kết quả

Method	MSE	RMSE
Autoencoder and MLP	40.11	6.33
MLP	42.34	6.51
Naive 1	132.82	11.52
Naive 2	88.43	9.40
Abhishek et al. (BP) [11]	93.51	9.67
Abhishek et al. (LR) [11]	81.72	9.04
Abhishek et al. (CBP) [11]	81.36	9.02
Beltran et al. [10]	98.73	9.94

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

$$RMSE = \sqrt{MSE}$$



[10] Beltrán-Castro, J., Valencia-Aguirre, J., Orozco-Alzate, M., Castellanos-Domínguez, G., Travieso-González, C.M.: Rainfall forecasting based on ensemble empirical mode decomposition and neural networks.

[11] Abhishek, K., Kumar, A., Ranjan, R., Kumar, S.: A rainfall prediction model using artificial neural network.

- ❑ Tiềm năng ứng dụng AI vào các bài toán dự báo KTTV là rất lớn, như dự báo lượng mưa, dự báo nhiệt độ, dự báo đường đi, cường độ của bão.
- ❑ Tuy nhiên số lượng các nghiên cứu về việc ứng dụng các mô hình Machine Learning/Deep Learning vào bài toán dự báo KTTV còn chưa nhiều do:
 - Giới hạn về kiến thức chuyên gia: Do đây là bài toán khó, tác động bởi nhiều yếu tố môi trường và thời tiết nên cần nhiều kiến thức chuyên gia trong việc lựa chọn đặc trưng dữ liệu và xây dựng mô hình.
 - Giới hạn về dữ liệu: Khối lượng dữ liệu là rất lớn nhưng việc tiếp cận và sử dụng dữ liệu còn gặp nhiều khó khăn, cũng như đòi hỏi cơ chế tổ chức và tổng hợp dữ liệu một cách tối ưu.



Thank you!
