Neural Network (Cont.)

Petchara Pattarakijwanich

Introduction to Data Science, 2 December 2022

x_1	x_2	x_3	y

แล้วดูโมเดล

- MSE ของ training
- MSE ของ test data

- AIC } ทำแทน
- BIC } MSE ของ test ตอนหมดมุ้

$$AIC = \frac{1}{\sigma^2}\left(MSE + \frac{2d\sigma^2}{n}\right)$$

$$BIC = \frac{1}{\sigma^2}\left(MSE + \frac{\sigma^2 d\ln n}{n}\right)$$

Best subset selection

8 model

- $y = \beta_0$

- $y = \beta_0 + \beta_1 x_1$

- $\cdot\cdot$

- $\cdot\cdot$

  } 1 ตัวแปร

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

- $\cdot\cdot$

- $\cdot\cdot$

  } 2 ตัวแปร

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ 3 ตัวแปร

Ridge & Lasso

ต้อง normalize

$$y_{model} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Linear Regression : Minimize $(y - y_{model})^2$

Ridge : Minimize $(y - y_{model})^2 + \lambda(\beta_1^2 + \beta_2^2 + \beta_3^2)$

ระวังค่า β

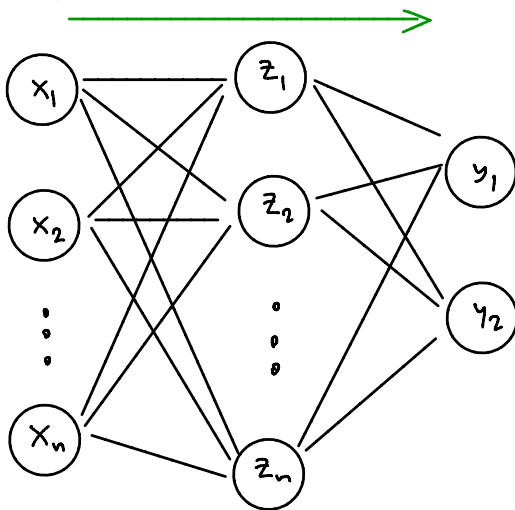Lasso : Minimize $(y - y_{model})^2 + \lambda(|\beta_1| + |\beta_2| + |\beta_3|)$

Parameter selection
β บางตัว = 0

Goal of this week

- Recurrent Neural Network

- Back Propagation and Fitting

- How to Prevent Overfitting

- Summary of the Course

Recurrent Neural Network

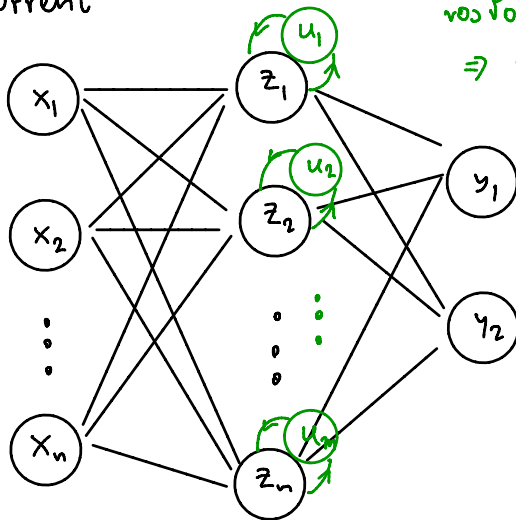# Recurrent Neural Network

Recurrent

Input   Hidden   Output
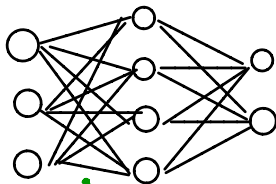
ตัว u เป็น "ความจำ" ของข้อมูลก่อนหน้า
⇒ เหมาะกับ time-series

Back Propagation and Fitting

Prediction ข้างขวา

Fitting ข้างซ้าย

weight α_i

หา weight ทั้งหมด ที่ทำให้ฟิตดีที่สุด ?

ฟิตดี $\Longleftrightarrow$ MSE น้อยสุด

หา $\dfrac{\partial(MSE)}{\partial\alpha_i}$ ได้ $\Longrightarrow$ หาได้ว่าต้องปรับ α_i เท่าไหร่ MSE ถึงน้อยลง
(Back-Propagation)

Back Propagation and Fitting
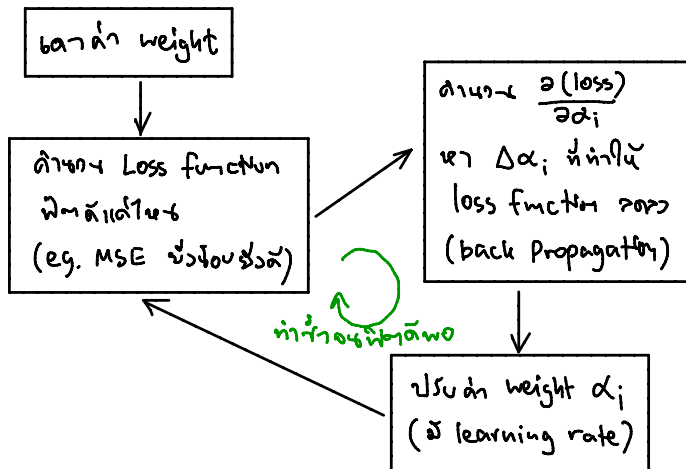
How to Prevent Overfitting

① Learning Rate (ปรับอัตราการเปลี่ยนค่าต่าง ๆ แบบ α_i)

       Slow learning ผักว่าดี   Fast learning มักจะ overfit

② Regularization (เพิ่มราคาของตัวแปรที่ใช้)

    minimize $\nearrow$ loss function $+ \lambda \sum (weight)^2$

             $\searrow$ loss function $+ \lambda \sum |weight|$

③ Drop out (ถ้า weight บางตัว ⟹ ตั้งให้เป็น 0)

     − ประหยัดเวลาขึ้น?

     − Decorrelate ตัวแปรต่าง ๆ ($\approx$ Random Forest)

Course Summary

Supervised Machine Learning

- Input $\Rightarrow$ Model $\Rightarrow$ ทำนาย output

- Model รับจาก training data ↰ รู้คำตอบ

- Model ทดสอบ กับ test data (+ cross-validation)

    (ทดสอบ กับ training data ไม่ได้ $\Rightarrow$ overfit)

- เอา Model ไปใช้

Course Summary

Data Science

- Empirical (เชิงประจักษ์)
  ไม่มีวิธีที่ "ถูก" ตายตัว

  ทดสอบ ความดีของโมเดล
  ด้วย test data

Physics

- มีวิธีที่ "ถูก" ตายตัว

Course Summary

Bias Variance Trade-off
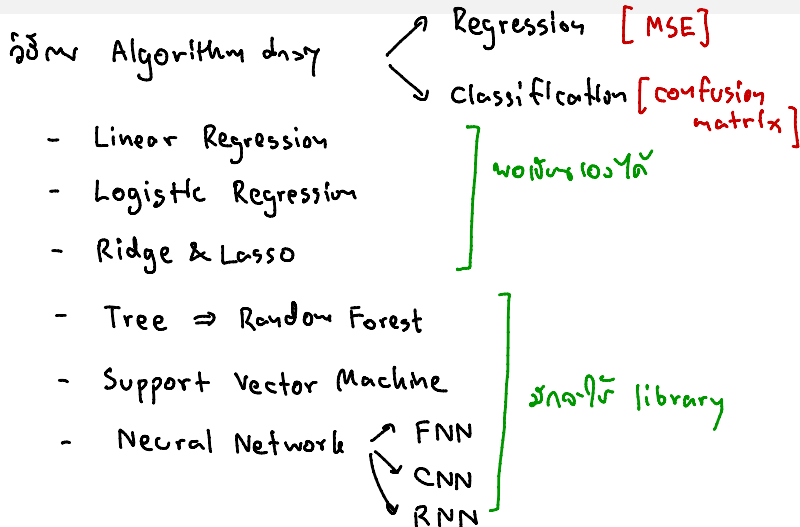
$[Bias]^2 + variance + \sigma^2$

variance

$[Bias]^2$

ความซับซ้อนของโมเดล (MSE)

Test Data

จับข้อมูลน้อยเกินไป
Bias สูง variance ต่ำ

จับข้อมูลพอดี
หาด้วย test data

จับข้อมูลมากเกินไป
Bias ต่ำ variance สูง
(overfit กับ training data)

Course Summary

วิธีการ Algorithm ต่างๆ

Regression [MSE]

Classification [confusion matrix]

- Linear Regression
- Logistic Regression
- Ridge & Lasso

พอเขียนเองได้

- Tree ⇒ Random Forest
- Support Vector Machine
- Neural Network → FNN
  → CNN
  → RNN

ต้องใช้ library

Course Summary

ความเชี่ยวชาญ →

ผู้ใช้องประยุกต์

- ใช้ library เป็น Blackbox
- ข้อมูล → library → ผล
- พารามิเตอร์ ดีไหม? ไม่ดี? ปรับได้ยังไง? ไม่รู้

ข้อดี ของ [ML] = ใช้ค่าตอบ เร็ว
ข้อเสีย ของ [ML] = ใช้ค่าตอบ เร็ว

เป้าหมายของวิชา
ผู้ใช้ที่ดี

- ใช้โค้ด library อย่างเข้าใจ ง่องขึ้น
- Diagnose ปัญหาได้ (บอกได้ว่าผลลัพธ์ดีมั้ย? ปรับยังไง ลงดีมั้ย?)
- เขียน algorithm ง่ายๆได้
- แกะ library ได้ระดับนึง
- ล้วงจำเป็น ⇒ ปรับไปเป็น กลุ่มบนได้

ผู้สร้าง

- สร้าง algorithm ใหม่
- เขียน library
- ผู้ใช้ขั้นสูงมาก