

Machine Learning

Lecturer: Doctor Bui Thanh Hung

Data Science Laboratory

Faculty of Information Technology

Industrial University of Ho Chi Minh city

Email: hung.buithanhcs@gmail.com (buithanhhung@iuh.edu.vn)

Website: <https://sites.google.com/site/hungthanhbui1980/>

Tự cài đặt giải thuật Gaussian Naïve Bayes biết rằng xác suất phân lớp theo giải thuật này như sau:

$$P(x | C_1) = \frac{1}{\sqrt{2\pi}S_1} e^{-\frac{(x - m_1)^2}{2S_1^2}}$$

Trong đó:

- μ (mean) và σ (standard deviation)

Viết các hàm (không sử dụng thư viện):

- 1- Đọc dữ liệu
- 2- Tính mean
- 3- Tính Standard deviation
- 4- Tính xác suất
- 5- Phân lớp
- 6- Tính độ chính xác
- 7- Áp dụng giải thuật này cho bộ dữ liệu sau:

Day	Outlook	Humidity	Wind	PlayTennis
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No