Trích chọn đặc trưng là gì ?

Là phương pháp giảm biến đầu vào cho mô hình của bạn bằng cách chỉ sử dụng dữ liệu có liên quan và loại bỏ nhiễu trong dữ liệu.

Trích chọn đặc trưng( feature selection ) :

* Có giám sát ( supervise )
* Không có giám sát ( unsupervise )

\* Có giám sát ( supervise ) : được sử dụng với các bộ dữ liệu có nhãn .

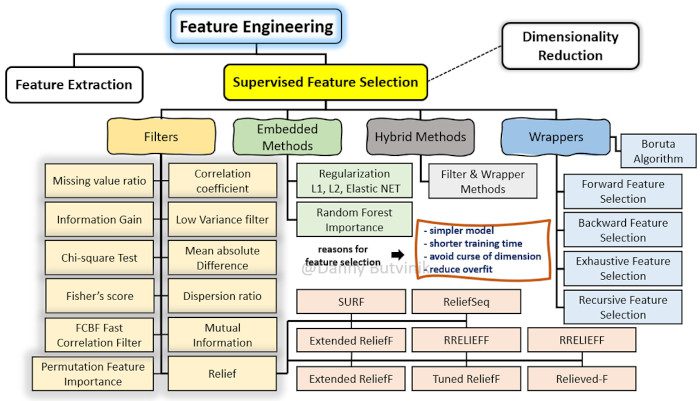
* Phương pháp có giám sát được chia thành 4 loại dựa trên sự tương tác đối với mô hình :

- Phương pháp lọc ( filter method )

- Phương pháp gói ( wrapper method )

- Phương pháp kết hợp ( hybrid method )

- Phương pháp nhúng ( embedded method )



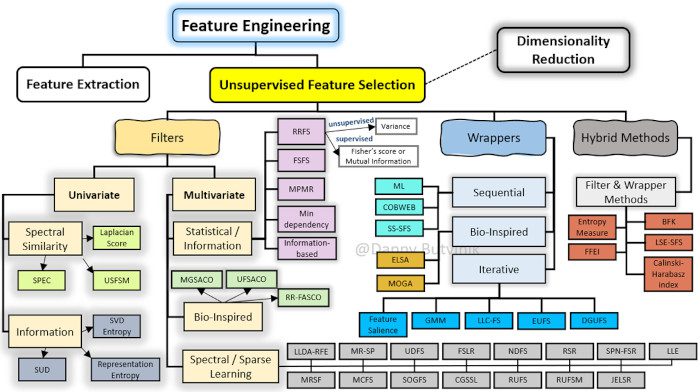
\* Không có giám sát ( unsupervise ) : sử dụng cho các bộ dữ liệu không có nhãn.

Phương pháp có giám sát được chia thành 3 loại dựa trên sự tương tác đối với mô hình :

- Phương pháp lọc ( filter method )

- Phương pháp gói ( wrapper method )

- Phương pháp kết hợp ( hybrid method )



tài liệu tham khảo:

<https://www.kdnuggets.com/2021/06/feature-selection-overview.html>