KLASIFIKASI VARIETAS BIJI MENGGUNAKAN DECISION TREE & ENSEMBLE METHODS

Fahmi Noor Figri

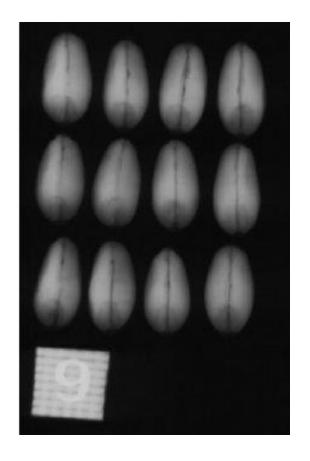
By Kodesiana.com



Dataset Seeds

Dataset seeds merupakan dataset yang berisi informasi ukuran biji padi dari 3 varietas biji.

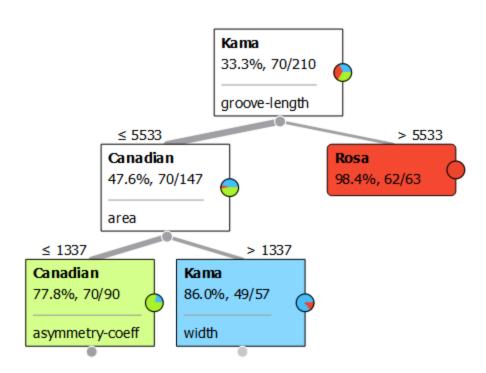
Tujuan dari dataset ini adalah membuat model yang dapat melakukan klasifikasi varietas biji berdasarkan ukuran bijinya.



Sumber: https://archive.ics.uci.edu/ml/datasets/seeds
https://www.researchgate.net/publication/234073
581 Seeds dataset/figures

Decision Tree

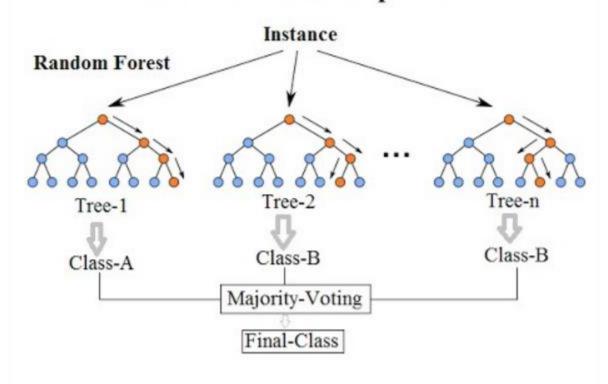
Decision tree merupakan metode klasifikasi prediktif yang menggunakan observasi pada *node* untuk menemukan target pada *leaves*.



Random Forest

Random forest merupakan metode klasifikasi yang menggabungkan beberapa classification tree untuk meningkatkan akurasi dan mengurangi overfiitting.

Random Forest Simplified

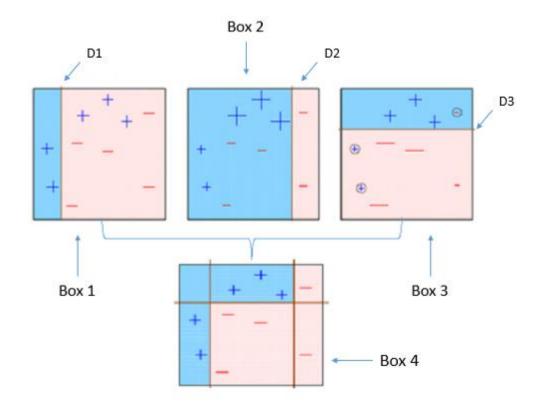


Sumber:

https://commons.wikimedia.org/wiki/File:Ra ndom_forest_diagram_complete.png

AdaBoost

AdaBoost merupakan metode klasifikasi yang menggabungkan beberapa *classification tree* untuk menemukan pola yang rumit dari data.



Sumber: https://towardsdatascience.com/understan ding-adaboost-2f94f22d5bfe

AdaBoost v.s. Random Forest

Random Forest

- Merupakan metode bagging.
- 2. Melatih *learner* secara non-sekuensial.
- 3. Pada setiap iterasi, *learner* akan semakin berkembang.

AdaBoost

- 1. Merupakan metode **boosting**.
- 2. Melatih *learner* secara sekuensial.
- Target kelas diambil dari kelas yang paling populer dari seluruh tree.

Bagging: Parallel tree growing with subsamples

X1, y1

X3, y3

X2, y2

X3, y3

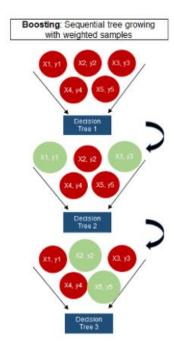
X1, y1

X4, y4

Decision
Tree 1

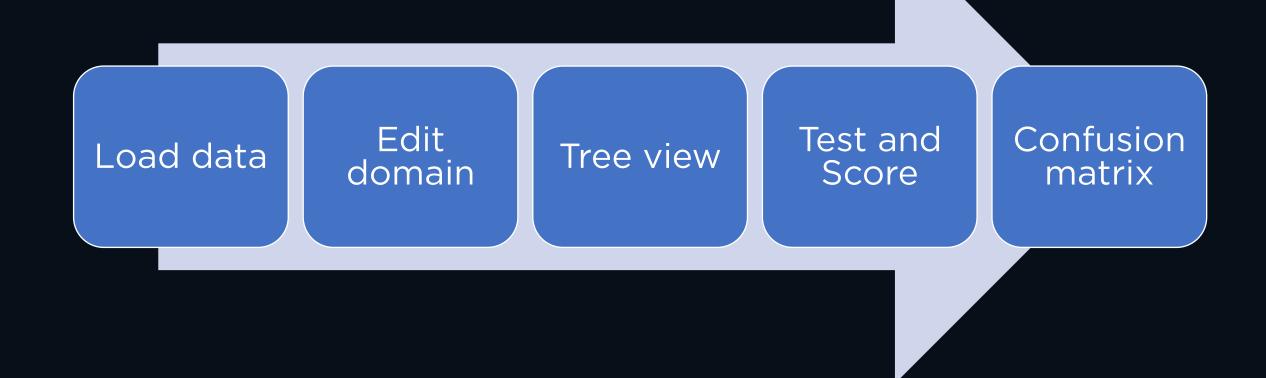
Decision
Tree 2

Decision
Tree 3



Sumber: https://towardsdatascience.com/theultimate-guide-to-adaboost-random-forests-andxgboost-7f9327061c4f

Proses Data Mining



KLASIFIKASI VARIETAS BIJI MENGGUNAKAN DECISION TREE & ENSEMBLE METHODS

Fahmi Noor Fiqri

By Kodesiana.com

