ID JOURNAL INPUT

1 Title:

Supervised Machine Learning Algorithms: Classification and Comparison

Author:

Osisanwo F.Y., Akinsola J.E.T., Awodele O., JHinmikaiye J. O., Olakanmi O., Akinjobi J.

Year: 2017

Journal:

International Journal of Computer Trends and Technology (IJCTT)

+ Seven different machine learning algorithms:

Decision Table, Random Forest (RF), Naïve Bayes (NB), Support Vector Machine (SVM), Neural Networks (Perceptron), JRip and Decision Tree (J48)

- + use diabetes dataset
- + SVM>Naïve Bayes>Random Forest
- + Claim that:
 - a) Common use supervised algorithm for classification is Linear Classifiers, Logistic Regression, Naïve Bayes Classifier, Perceptron, Support Vector Machine; Quadratic Classifiers, K-Means Clustering, Boosting, Decision Tree, Random Forest (RF); Neural networks, Bayesian Networks and so on.
- + Found that:
 - a) SVM & NB require larger dataset for higher accuracy
 - b) SVM & RF high accuracy for lower dataset
 - c) for smaller dataset, NB is the fastest but SVM highest accuracy
 - d) DT does not perform well with higher dataset

2 Title:

A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease

Author:

Morshedul Bari Antor, 1 A. H. M. Shafayet Jamil, 1 Maliha Mamtaz, 1 Mohammad Monirujjaman Khan, 1 Sultan Aljahdali, 2 Manjit Kaur, 3 Parminder Singh, 4 and Mehedi Masud 2

Year: 2021

DOI: doi.org/10.1155/2021/9917919

+ Use dataset from Magnetic Resonance Imaging (MRI) from OASIS to predict dementia from different attributes of patient.

- + Claim that:
 - Usual use ML model in medical diagnosis include SVM, RF,DT,LR
- + Found that:
 - a) SVM operate outlier well
 - b) SVM perform better than RF
 - c) DT better at collinearity than LR
 - d) RF better than DT
 - e) SVM > RF>DT>LR

3 Title:

Comparing different supervised machine learning algorithms for disease prediction

Author:

Shahadat Uddin1*, Arif Khan1,2, Md Ekramul Hossain1 and Mohammad Ali Moni3

Year: 2019

DOI:

doi.org/10.1186/s12911-019-1004-8

- + Data gain from Scopus and PublicMed databases.
- + Data use include: cancer incidence & survival, medical image and gene data
- + apply cross validation at 2/5/10 k fold for each algortihm
- + Found that:
 - a) SVM found to be applied most, 2nd is NB
 - b) RF show superior accuracy > SVM
 - c) SVM perform best at 10 fold& 5fold > RF