An Improved Algorithm based on KNN and Random Forest

Authors: Jun Liang, Qin Liu, Nuihua Nie, Biqing Zeng, Zanbo Zhang | Authors Info & Claims

CSAE '19: Proceedings of the 3rd International Conference on Computer Science and Application Engineering
Article No.: 74, Pages 1 - 6 https://doi.org/10.1145/3331453.3360963

Published: 22 October 2019 Publication History

Abstract

This paper gives an improved algorithm called RFDKNN based on an enhanced KNN (K-Nearest Neighbor) and random forest. First, RFDKNN sorts features based on importance through Gini index and a random forest algorithm. Then it deletes some unimportant features based on this sort in a certain proportion r. Finally, it uses an enhanced KNN algorithm dynamically selecting the optimal nearest neighbor number and distance function to make the distance between two samples closer to true value. Experiments are carried out on the 20 data sets from UCI Machine Learning repository. The results show that compared with other r values, RFD of r=0.7 can obtain a relatively satisfactory classification accuracy. Compared Naive Bayes, Adaboost, Random Forest, RRSB, W-KNN, dwh-FNN and LI-KNN,

References

[1] Z Li, K Zhang, B Chen, Y Dong and L Zhang (2018). Driver identification in intelligent vehicle systems using machine learning algorithms. IET Intelligent Transport Systems, 13(1), 40--47.

Crossref | Google Scholar

[2] J B Kisiel, B A Dukek, R VSA Kanipakam, et al. (2019). Hepatocellular Carcinoma Detection by Plasma Methylated DNA: Discovery, Phase I Pilot, and Phase II Clinical Validation. Hepatology, 69(3), 1180--1192.

Google Scholar

Show all references

Cited By

View all

Tri Wibowo A, Utami Putri A, Reza Tribosnia M, Putawara R and Mujiya Ulkhaq M. (2023). A Machine Learning Application to Predict Customer Churn: A Case in Indonesian Telecommunication Company. Advanced Mathematical Applications in Data Science. 10.2174/978981512484212301013. (144-161). Online publication date: 22-Aug-2023.

https://doi.org/10.2174/978981512484212301013

Josephine H, Rajan D and Joy J. (2023). Forecasting a FastMoving Consumer Goods (FMCG) Cor Customer Repurchase Behavior via Classification Machine Learning Models. Proceedings of the 5th International Conference on Information Management & Machine Intelligence. 10.1145/3647444.3647840. (1-5). Online publication date: 23-Nov-2023.

https://dl.acm.org/doi/10.1145/3647444.3647840

CSAE

Index Terms

An Improved Algorithm based on KNN and Random Forest

Computing methodologies

Machine learning

Learning paradigms

Supervised learning

Supervised learning by classification

Machine learning algorithms

Ensemble methods

Boosting

Feature selection

Recommendations

Reinforced random forest

ICVGIP '16: Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image P

Reinforcement learning improves classification accuracy. But use of reinforcement learning is relatively une case of random forest classifier. We propose a reinforced random forest (RRF) classifier that exploits reinforcement…

Read More

Rotation Forest: A New Classifier Ensemble Method

We propose a method for generating classifier ensembles based on feature extraction. To create the training data for a base classifier, the feature set is randomly split into K subsets (K is a parameter of the algorithm) and Principal…

- Propose an improved random forest based on the improvement of decision trees....

Abstract

CSAE

Comments

DL Comment Policy

Comments should be relevant to the contents of this article, (sign in required).

Got it

0 Comments

Share **Best** **Newest** **Oldest**

Nothing in this discussion yet.

Download PDF

View Table Of Contents

PDF

Help