

**Wickramarachchi, D. C.; Robertson, B. L.; Reale, M.; Price, C. J.; Brown, J.**  
**HHCART: an oblique decision tree.** (English) Zbl 06918560  
*Comput. Stat. Data Anal.* 96, 12-23 (2016).

**Summary:** Decision trees are a popular technique in statistical data classification. They recursively partition the feature space into disjoint sub-regions until each sub-region becomes homogeneous with respect to a particular class. The basic Classification and Regression Tree (CART) algorithm partitions the feature space using axis parallel splits. When the true decision boundaries are not aligned with the feature axes, this approach can produce a complicated boundary structure. Oblique decision trees use oblique decision boundaries to potentially simplify the boundary structure. The major limitation of this approach is that the tree induction algorithm is computationally expensive. Hence, as an alternative, a new decision tree algorithm called HHCART is presented. The method uses a series of Householder matrices to reflect the training data at each non-terminal node during tree construction. Each reflection is based on the directions of the eigenvectors from each class' covariance matrix. Considering of axis parallel splits in the reflected training data provides an efficient way of finding oblique splits in the unreflected training data. Experimental results show that the accuracy and size of HHCART trees are comparable with some benchmark methods. The appealing feature of HHCART is that it can handle both qualitative and quantitative features in the same oblique split.

**MSC:**

62 Statistics

**Keywords:**

oblique decision tree; data classification; statistical learning; Householder reflection; machine learning

**Software:**

CARTopt; UCI-ml

**Full Text:** [DOI](#)

**References:**

- [1] Amasyah, M.; Ersoy, O., Cline: A new decision-tree family, *IEEE Trans. Neural Netw.*, 19, 2, 356-363, (2008)
- [2] Bache, K., Lichman, M., 2013. UCI machine learning repository. URL: <http://archive.ics.uci.edu/ml>.
- [3] Breiman, L.; Friedman, J.; Stone, C. J.; Olshen, R. A., *Classification and regression trees*, (1984), CRC Press New York · [Zbl 0541.62042](#)
- [4] Gama, J.; Brazdil, P., Linear tree, *Intell. Data Anal.*, 3, 1, 1-22, (1999) · [Zbl 1059.68606](#)
- [5] Heath, D., Kasif, S., Salzberg, S., 1993. Induction of oblique decision trees. In: *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, pp. 1002-1007.
- [6] Hyafil, L.; Rivest, R. L., Constructing optimal binary decision trees is NP-complete, *Inform. Process. Lett.*, 5, 1, 15-17, (1976) · [Zbl 0333.68029](#)
- [7] Ittner, A., Schlosser, M., 1996. Non-linear decision trees-NDT. In: *13th International Conference on Machine Learning*. Citeseer, Italy, pp. 252-257.
- [8] Iyengar, V. S., HOT: heuristics for oblique trees, (11th International conference on Tools with Artificial Intelligence, (1999), IEEE New York), 91-98
- [9] Kolakowska, A.; Malina, W., Fisher sequential classifiers, *IEEE Trans. Syst. Man Cybern. B*, 35, 5, 988-998, (2005)
- [10] Li, Y.; Dong, M.; Kothari, R., Classifiability-based omnivariate decision trees, *IEEE Trans. Neural Netw.*, 16, 6, 1547-1560, (2005)
- [11] Li, X.-B.; Sweigart, J. R.; Teng, J. T.; Donohue, J. M.; Thombs, L. A.; Wang, S. M., Multivariate decision trees using linear discriminants and tabu search, *IEEE Trans. Syst. Man Cybern. A*, 33, 2, 194-205, (2003)
- [12] Loh, W.-Y.; Shih, Y.-S., Split selection methods for classification trees, *Statist. Sinica*, 7, 4, 815-840, (1997) · [Zbl 1067.62545](#)
- [13] López-Chau, A.; Cervantes, J.; López-García, L.; Lamont, F. G., Fisher's decision tree, *Expert Syst. Appl.*, 40, 16, 6283-6291,

(2013)

- [14] Manwani, N.; Sastry, P., Geometric decision tree, *IEEE Trans. Syst. Man Cybern. B*, 42, 1, 181-192, (2012)
- [15] Murthy, S.K., Kasif, S., Salzberg, S., 1993. The oc1 decision tree software system. URL: <http://www.cbc.umd.edu/salzberg/announce-oc1.html>.
- [16] Murthy, S. K.; Kasif, S.; Salzberg, S., A system for induction of oblique decision trees, *J. Artificial Intelligence Res.*, 2, 1, 1-32, (1994) · [Zbl 0900.68335](#)
- [17] Quinlan, J. R., Induction of decision trees, *Mach. Learn.*, 1, 1, 81-106, (1986)
- [18] Robertson, B.; Price, C.; Reale, M., Cartopt: a random search method for nonsmooth unconstrained optimization, *Comput. Optim. Appl.*, 56, 2, 291-315, (2013)
- [19] Rokach, L.; Maimon, O., Top-down induction of decision trees classifiers—a survey, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, 35, 4, 476-487, (2005)
- [20] Sokolova, M.; Lapalme, G., A systematic analysis of performance measures for classification tasks, *Inf. Process. Manage.*, 45, 4, 427-437, (2009)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.