Data set name	Data set size	Number of attributes	Number of classes	Number of clusters	Testing method	Default success rate
Australian	690	15	2	17	10-fold CV	56.0
Breast cancer	286	10	2	21	11-fold CV	70.3
Diabetes	768	9	2	20	12-fold CV	65.0
DNA	3186	181	3	13	train&test	50.8
Glass	214	10	6	30	7-fold CV	40.7
Heart disease	270	14	2	8	9-fold CV	79.4
Hepatitis	150	20	2	9	5-fold CV	55.6
Iris	150	5	3	4	5-fold CV	33.3
Lymphography	148	19	4	19	5-fold CV	54.7
Primary tumor	339	18	21	21	10-fold CV	24.8

Table 1: Datasets and testing methods used in our classification experiments.

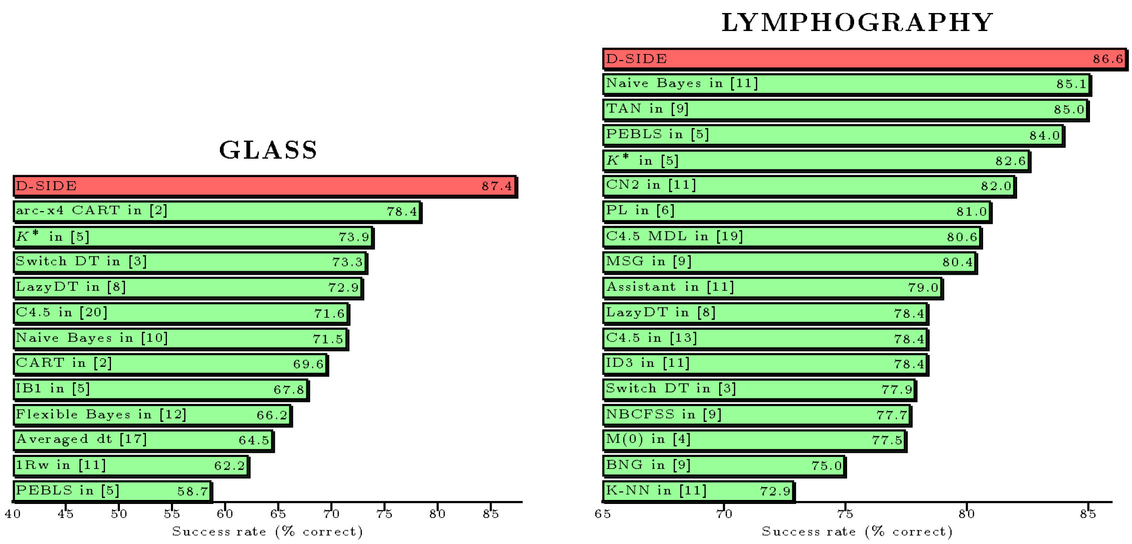

Figure 1: Experimental results on the Glass and Lymphography databases.

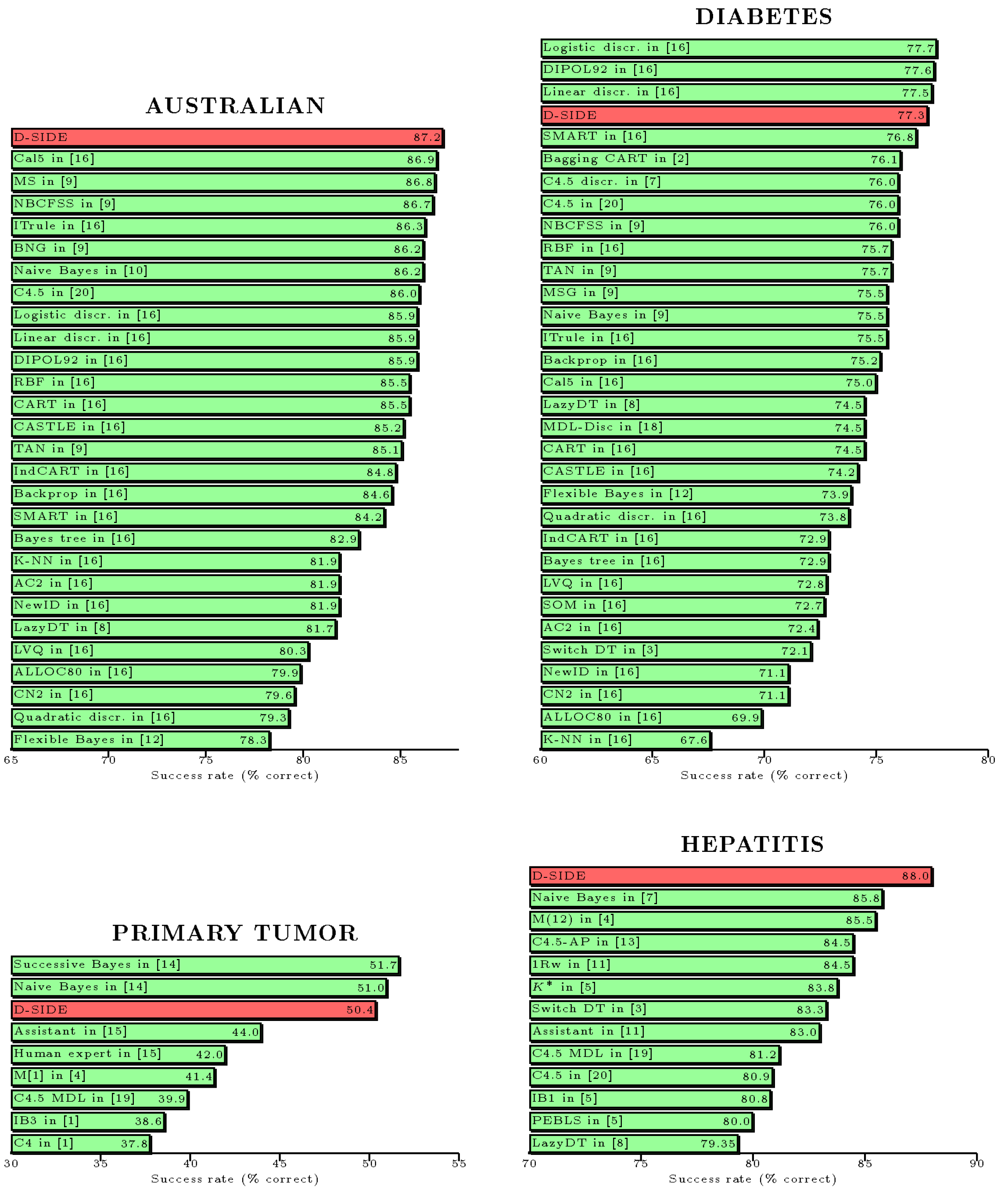

Figure 2: Experimental results on the Australian, Diabetes, Primary tumor and Hepatitis datasets.

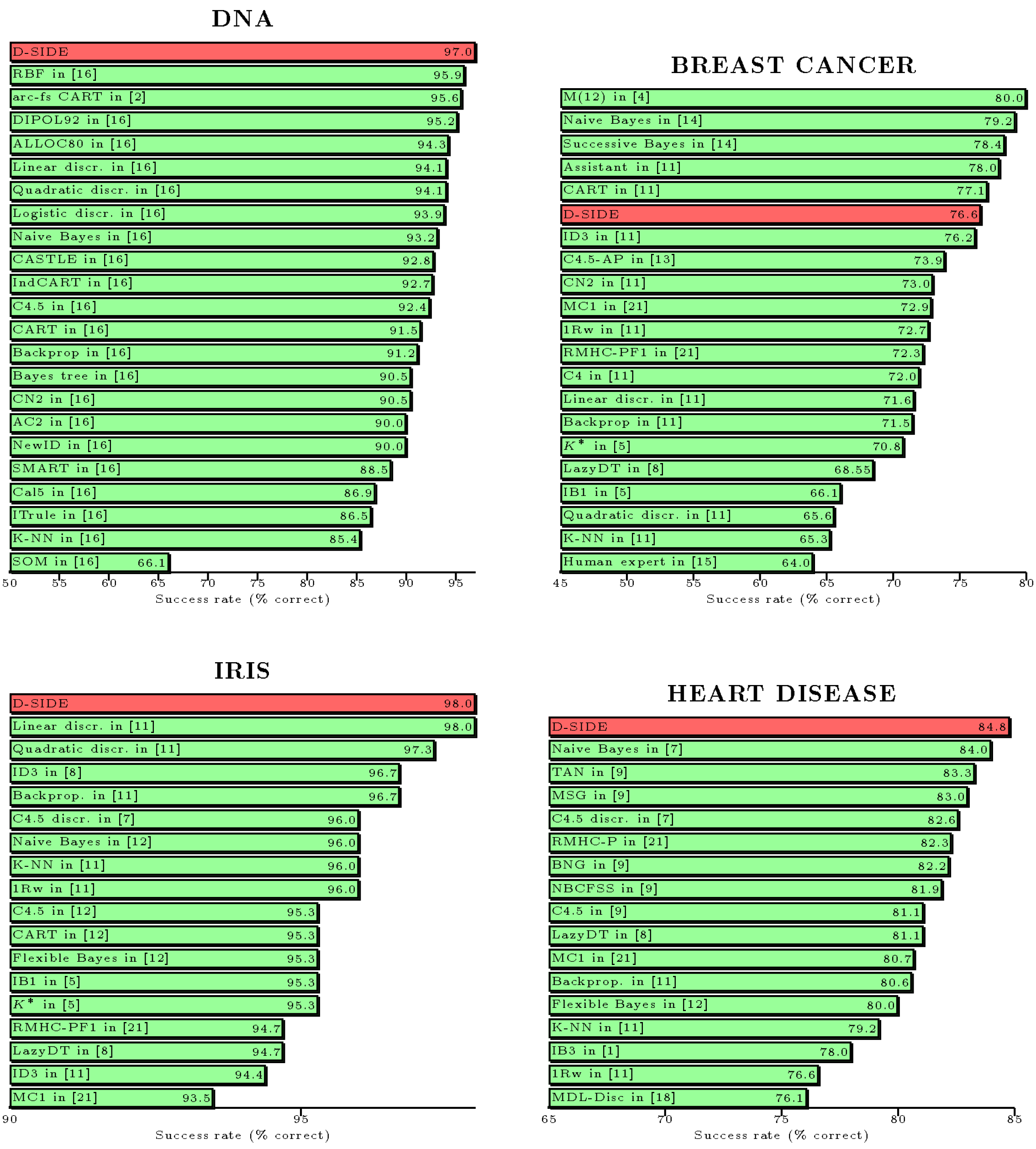

Figure 3: Experimental results on the DNA, Breast cancer, Iris, and Heart disease databases.

References

[1] D. Aha, D. Kibler, and M. Albert. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.

[2] L. Breiman. Bias, variance, and arcing classifiers. Technical Report TR 460, University of California, Statistics Department, April 1996.

[3] C.E. Brodley. Automatic selection of split criterion during tree growing based on node location. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 73–80. Morgan Kaufmann Publishers, 1995.

[4] B. Cestnik and I. Bratko. On estimating probabilities in tree pruning. In Y. Kodratoff, editor, *Machine Learning EWSL-91*, pages 138–150. Springer-Verlag, 1991.

[5] J.G. Cleary and L.E. Trigg. k^*: An instance-based learner using and entropic distance measure. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 108–114. Morgan Kaufmann Publishers, 1995.

[6] P. Datta and D. Kibler. Learning prototypical concept descriptions. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 158–166. Morgan Kaufmann Publishers, 1995.

[7] J. Dougherty, R. Kohavi, and M. Sahami. Supervised and unsupervised discretization of continuous features. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 194–202. Morgan Kaufmann Publishers, 1995.

[8] J.H. Friedman, R. Kohavi, and Y. Yun. Lazy decision trees. In *Proceedings of AAAI-96 (to appear)*, 1996.

[9] N. Friedman and M. Goldszmidt. Building classifiers using Bayesian networks. In *Proceedings of AAAI-96 (to appear)*, 1996.

[10] N. Friedman and M. Goldszmidt. Discretizing continuous attributes while learning Bayesian networks. In L. Saitta, editor, *Machine Learning: Proceedings of the Thirteenth International Conference (to appear)*. Morgan Kaufmann Publishers, 1996.

[11] R.C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11:63–91, 1993.

[12] G.H. John and P. Langley. Estimating continuous distributions in Bayesian classifiers. In P. Besnard and S. Hanks, editors, *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, pages 338–345. Morgan Kaufmann Publishers, 1995.

[13] R. Kohavi and G.H. John. Automatic parameter selection by minimizing estimated error. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 304–312. Morgan Kaufmann Publishers, 1995.

[14] I. Kononenko. Successive naive Bayesian classifier. *Informatica*, 17:167–174, 1993.

[15] I. Kononenko and I. Bratko. Information-based evaluation criterion for classifier's performance. *Machine Learning*, 6:67–80, 1991.

[16] D. Michie, D.J. Spiegelhalter, and C.C. Taylor, editors. *Machine Learning, Neural and Statistical Classification.* Ellis Horwood, London, 1994.

[17] J.O. Oliver and D.J. Hand. On pruning and averaging decision trees. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 430–437. Morgan Kaufmann Publishers, 1995.

[18] B. Pfahringer. Compression-based discretization of continuous attributes. In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 456–463. Morgan Kaufmann Publishers, 1995.

[19] J.R. Quinlan. Mdl and categorical theories (continued). In A. Prieditis, editor, *Machine Learning: Proceedings of the Twelfth International Conference*, pages 464–470. Morgan Kaufmann Publishers, 1995.

[20] J.R. Quinlan. Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Research*, 4:77–90, 1996.

[21] D. Skalak. Prototype and feature selection by sampling and random mutation hill climbing algorithms. In *Machine Learning: Proceedings of the Eleventh International Conference*, pages 293–301, 1994.