Name	Instances	Min	Max	Features	Bags	+ bags	- bags
Musk 1 [1]	476	2	40	166	92	47	45
Musk 2 [1]	6598	1	1044	166	102	39	63
Mutagenesis 1 (easy) [2]	10486	28	88	7	188	125	63
Mutagenesis 2 (hard) [2]	2132	26	86	7	42	13	29
Protein [3]	26611	35	189	8	193	25	168
Elephant [4]	1391	2	13	230	200	100	100
Fox [4]	1302	1	13	230	200	100	100
Tiger [4]	1220	2	13	230	200	100	100
Corel, African [5]	7947	2	13	9	2000	100	1900
Corel, Antique [5]	7947	2	13	9	2000	100	1900
Corel, Battleships [5]	7947	2	13	9	2000	100	1900
Corel, Beach [5]	7947	2	13	9	2000	100	1900
Corel, Buses [5]	7947	2	13	9	2000	100	1900
Corel, Cars [5]	7947	2	13	9	2000	100	1900
Corel, Desserts [5]	7947	2	13	9	2000	100	1900
Corel, Dinosaurs [5]	7947	2	13	9	2000	100	1900
Corel, Dogs [5]	7947	2	13	9	2000	100	1900
Corel, Elephants [5]	7947	2	13	9	2000	100	1900
Corel, Fashion [5]	7947	2	13	9	2000	100	1900
Corel, Flowers [5]	7947	2	13	9	2000	100	1900
Corel, Food [5]	7947	2	13	9	2000	100	1900
Corel, Historical [5]	7947	2	13	9	2000	100	1900
Corel, Horses [5]	7947	2	13	9	2000	100	1900
Corel, Lizards [5]	7947	2	13	9	2000	100	1900
Corel, Mountains [5]	7947	2	13	9	2000	100	1900
Corel, Skiing [5]	7947	2	13	9	2000	100	1900
Corel, Sunset [5]	7947	2	13	9	2000	100	1900
Corel, Waterfalls [5]	7947	2	13	9	2000	100	1900
UCSB Breast Cancer [6]	200	21	40	708	58	26	32
Newsgroups 1, alt.atheism [7]	5443	22	76	200	100	50	50
N.g. 2, comp.graphics [7]	3094	12	58	200	100	49	51
N.g. 3, comp.os.ms-windows.misc [7]	5175	25	82	200	100	50	50
N.g. 4, comp.sys.ibm.pc.hardware [7]	4827	19	74	200	100	49	51
N.g. 5, comp.sys.mac.hardware [7]	4473	17	71	200	100	50	50
N.g. 6, comp.windows.x [7]	3110	12	54	200	100	49	51
N.g. 7, misc.forsale [7]	5306	29	84	200	100	50	50
N.g. 8, rec.autos [7]	3458	15	39	200	100	50	50
N.g. 9, rec.motorcycles [7]	4730	22	73	200	100	50	50
N.g. 10, rec.sport.baseball [7]	3358	15	58	200	100	50	50
N.g. 11, rec.sport.hockey [7]	1982	8	38	200	100	50	50
N.g. 12, sci.crypt [7]	4284	20	71	200	100	50	50
N.g. 13, sci.electronics [7]	3192	12	58	200	100	47	53
N.g. 14, sci.med [7]	3045	11	54	200	100	50	50
N.g. 15, sci.space [7]	3655	16	59	200	100	50	50
N.g. 16, soc.religion.christian [7]	4677	21	71	200	100	50	50
N.g. 17, talk.politics.guns [7]	3558	13	59	200	100	50	50
N.g. 18, talk.politics.mideast [7]	3376	15	55	200	100	50	50
N.g. 19, talk.politics.misc [7]	4788	21	75	200	100	50	50
N.g. 20, talk.religion.misc [7]	4606	25	79	200	100	49	51

Table 1. MIL benchmark data sets.

Name	Instances	Min	Max	Features	Bags	+ bags	- bags
Web recommendation 1 [8]	2212	4	131	5863	75	17	58
Web recommendation 2 [8]	2212	5	200	6519	75	18	57
Web recommendation 3 [8]	2212	5	200	6306	75	14	61
Web recommendation 4 [8]	2291	4	200	6059	75	55	20
Web recommendation 5 [8]	2546	5	200	6407	75	61	14
Web recommendation 6 [8]	2462	4	200	6417	75	59	16
Web recommendation 7 [8]	2400	4	200	6450	75	39	36
Web recommendation 8 [8]	2183	4	200	5999	75	35	40
Web recommendation 9 [8]	2321	5	200	6279	75	37	38
Birds, Brown creeper [9]	10232	2	43	38	548	197	351
Birds, Chestnut-backed chickadee [9]	10232	2	43	38	548	117	431
Birds, Dark-eyed junco [9]	10232	2	43	38	548	20	528
Birds, Hammonds flycatcher [9]	10232	2	43	38	548	103	445
Birds, Hermit thrush [9]	10232	2	43	38	548	15	533
Birds, Hermit warbler [9]	10232	2	43	38	548	63	485
Birds, Olive-sided flycatcher [9]	10232	2	43	38	548	90	458
Birds, Pacific slope flycatcher [9]	10232	2	43	38	548	165	383
Birds, Red-breasted nuthatch [9]	10232	2	43	38	548	82	466
Birds, Swainsons thrush [9]	10232	2	43	38	548	79	469
Birds, Varied thrush [9]	10232	2	43	38	548	89	459
Birds, Western tanager [9]	10232	2	43	38	548	46	502
Birds, Winter Wren [9]	10232	2	43	38	548	109	439

Table 1 continued.

References

[1] T. G. Dietterich, R. H. Lathrop, T. Lozano-Pérez, Solving the multiple instance problem with axis-parallel rectangles, Artificial intelligence 89 (1) (1997) 31–71.

[2] A. Srinivasan, S. Muggleton, R. King, Comparing the use of background knowledge by inductive logic programming systems, in: Proceedings of the 5th International Workshop on Inductive Logic Programming, Department of Computer Science, Katholieke Universiteit Leuven, 1995, pp. 199–230.

[3] Q. Tao, S. Scott, N. Vinodchandran, T. T. Osugi, Svm-based generalized multiple-instance learning via approximate box counting, in: Proceedings of the twenty-first international conference on Machine learning, ACM, 2004, p. 101.

[4] S. Andrews, I. Tsochantaridis, T. Hofmann, Support vector machines for multiple-instance learning, in: Advances in Neural Information Processing Systems 15, MIT Press, 2003, pp. 561–568.

[5] Y. Chen, J. Bi, J. Z. Wang, Miles: Multiple-instance learning via embedded instance selection, Pattern Analysis and Machine Intelligence, IEEE Transactions on 28 (12) (2006) 1931–1947.

[6] M. Kandemir, C. Zhang, F. A. Hamprecht, Empowering multiple instance histopathology cancer diagnosis by cell graphs, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014, Springer, 2014, pp. 228–235.

[7] Z.-H. Zhou, Y.-Y. Sun, Y.-F. Li, Multi-instance learning by treating instances as non-iid samples, in: Proceedings of the 26th annual international conference on machine learning, ACM, 2009, pp. 1249–1256.

[8] Z.-H. Zhou, K. Jiang, M. Li, Multi-instance learning based web mining, Applied Intelligence 22 (2) (2005) 135–147.

[9] F. Briggs, B. Lakshminarayanan, L. Neal, X. Z. Fern, R. Raich, S. J. Hadley, A. S. Hadley, M. G. Betts, Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach, The Journal of the Acoustical Society of America 131 (6) (2012) 4640–4650.