

Contents

Preface	i
Acknowledgment	ii
Executive Summary	iii
1 Introduction	3
2 What is Dedicated Collaborative Interactive Learning?	5
2.1 Characterization of Dedicated Collaborative Interactive Learning	6
2.2 Challenges for D-CIL Research	8
3 Contributions	10
4 Outline of the Thesis	11
5 Literature Review	12
5.1 Active Learning with Omniscient Oracles	12
5.2 Active Learning in Practice	16
6 AL with Uncertain Oracles	19
6.1 Challenges in AL with Uncertain Human Oracles	19
6.2 Active Learning with Uncertain Human Oracles	20
6.3 AL Strategies with Uncertain Human Oracles	23
6.4 Conclusion	41
7 Simulation of Error-prone Oracles	43
7.1 Brief Analysis of Related Work	44
7.2 Motivation – The Problem	44
7.3 Human Expertise	45
7.4 Simulate Error-prone Annotators	46
8 Practical Application of Simulation with Examples	51
8.1 Pool-based Active Learning Cycle with Multiple Annotators	51
8.2 Base and Standard Annotators	55
8.3 Clusters as Knowledge Areas of Annotators	57
8.4 Classifiers as Annotators	61
8.5 Estimate Sample Difficulty to Simulate Annotators	65
8.6 Annotators with Dynamic Labeling Performance	69
8.7 Active Learning with Multiple Types of Annotators	73
8.8 Conclusion	75

9	Active Learning with One Uncertain Annotator	76
9.1	Related Work	78
9.2	Labeling Strategy – Own Approach	79
9.3	Experimental Results	80
9.4	Lessons Learned	86
10	Active Learning with Multiple Uncertain Oracles	89
10.1	Foundations	90
10.2	Meritocratic Learning	96
10.3	Evaluation	109
10.4	Conclusion	120
11	Discussion and Future Work	122
11.1	Challenge 3: Alternative Query Types	125
11.2	Challenge 4: Complex Cost Schemes for Queries	127
11.3	Challenge 5: True Collaboration of Human Experts and Interaction with an Active Learner	128
11.4	Challenge 6: Online Learning for Time-Variant Learning Tasks	129
11.5	Further Challenges	129
12	Final Thoughts	131
A	Closed-form Solution for Computing Expected Misclassification Costs	132
B	Scatter Plots of the Artificial Data Sets	134
C	Additional Experimental Results	135
	Bibliography	141