
Contents

Abstract	ix
Notations	xi
Introduction	xiii
Chapter 1. State of the Art of Statistical Learning Theory	1
1.1. Preliminaries	2
1.1.1. Definitions	2
1.1.2. No-free lunch theorem	5
1.1.3. Risk minimizing strategies	6
1.2. Generalization bounds	8
1.2.1. Vapnik–Chervonenkis bounds	8
1.2.2. Rademacher bounds	11
1.2.3. PAC–Bayesian bounds	13
1.2.4. Uniform stability	15
1.2.5. Algorithmic robustness	17
1.3. Summary	18
Chapter 2. Domain Adaptation Problem	21
2.1. Notations and preliminaries	22
2.2. Overview of the adaptation scenarios	24
2.2.1. Covariate shift	25
2.2.2. Target shift	28
2.2.3. Concept shift	30
2.2.4. Sample-selection bias	32
2.2.5. Other adaptation scenarios	33

2.3. Necessity of a divergence between domains	34
2.4. Summary	35
Chapter 3. Seminal Divergence-based Generalization Bounds . . .	37
3.1. The first theoretical analysis of domain adaptation	37
3.2. One seminal result	38
3.3. A generalization bound for domain adaptation	41
3.4. Generalization bounds for the combined error	42
3.5. Discrepancy distance	45
3.5.1. Definition	46
3.5.2. Generalization bounds with discrepancy distance	47
3.5.3. Domain adaptation in regression	49
3.6. Summary	56
Chapter 4. Impossibility Theorems for Domain Adaptation	59
4.1. Problem set-up	60
4.1.1. Definitions	60
4.1.2. Common assumptions in domain adaptation	61
4.2. Impossibility theorems	62
4.2.1. Theoretical results	62
4.2.2. Illustrative examples	63
4.3. Hardness results	66
4.4. Usefulness of unlabeled data	68
4.5. Proper domain adaptation	71
4.6. Summary	73
Chapter 5. Generalization Bounds with Integral Probability Metrics	75
5.1. Definition	75
5.2. General setting	76
5.3. Wasserstein distance	79
5.3.1. Generalization bounds with the Wasserstein distance	81
5.3.2. Combined error and semi-supervised case	84
5.4. Maximum mean discrepancy distance	85
5.4.1. Generalization bound with kernel embeddings	86
5.4.2. Combined error and semi-supervised case	88

5.5. Relationship between the MMD and the Wasserstein distances	89
5.6. Summary	92
Chapter 6. PAC–Bayesian Theory for Domain Adaptation	93
6.1. The PAC–Bayesian framework	93
6.2. Two PAC–Bayesian analyses of domain adaptation	94
6.2.1. In the spirit of Ben-David <i>et al.</i> and Mansour <i>et al.</i>	94
6.2.2. A different philosophy	96
6.2.3. Comparison of the two domain adaptation bounds	98
6.3. Specialization to linear classifiers	99
6.3.1. Setting and notations	99
6.3.2. Domain disagreement, expected disagreement and joint error of linear classifiers	100
6.3.3. Domain adaptation bounds for linear classifiers	102
6.3.4. Generalization bounds	103
6.4. Summary	104
Chapter 7. Domain Adaptation Theory Based on Algorithmic Properties	105
7.1. Introduction	105
7.2. Robust domain adaptation	105
7.2.1. λ -shift a measure of similarity between domains	105
7.2.2. Domain adaptation bounds using robustness	106
7.3. Hypothesis transfer learning	107
7.3.1. Problem set-up	107
7.3.2. A biased regularized least squares algorithm for HTL	108
7.3.3. HTL for regularized ERM with smooth and convex loss functions .	111
7.3.4. Comparison with standard theory of domain adaptation	113
7.3.5. HTL for metric learning	114
7.4. Summary	119
Chapter 8. Iterative Domain Adaptation Methods	121
8.1. Iterative self-labeling domain adaptation	122
8.2. Boosting-based frameworks	127
8.2.1. TRADABOOST with both labeled source and target data	128
8.2.2. SLDAB for unsupervised domain adaptation	129
8.3. Summary	132
Conclusions and Discussions	135
Appendices	137
Appendix 1	139

Appendix 2	153
Appendix 3	165
Appendix 4	171
Appendix 5	175
Bibliography	179
Index	189