

Contents

1	Introduction	1
1.1	Multi-relational learning	1
1.1.1	Factorization models for Multi-relational data	4
1.1.2	Applications of Multi Relational Learning	5
1.2	Contribution	6
1.3	Submitted and Published Work	7
1.4	Chapter Overview	9
2	The Multi-Relational Factorization Problem	11
2.1	Problem Formulation	12
2.2	State of the Art	14
2.2.1	Parametrization of Multi-relational factorization models	15
2.2.2	Optimization objectives for multi-relational learning	18
2.2.3	Summary of the presented models	21
2.3	Evaluating Multi-Relational Learning Models	21
2.4	Open Problems in Multi-Relational Factorization	23
3	Loss functions for multi-relational learning tasks	26
3.1	Application scenario: Mining RDF Knowledge Bases	27
3.1.1	RDF inference and Related Work	28
3.2	Predicting RDF Triples	30
3.3	Predicting RDF triples by Tensor Factorization	31
3.3.1	Factorization Models	31
3.3.1.1	Three-way Interaction Model	31
3.3.1.2	Pairwise Interaction Model	32
3.3.2	The open world assumption and the loss function	33
3.3.3	Dealing with the open world assumption through the BPR Framework	35
3.3.4	Filtering the Results of Tensor Factorization Models	36

3.4	Evaluation	39
3.4.1	Datasets	39
3.4.2	Methods	40
3.4.3	Evaluation Methodology	41
3.4.4	Results	42
3.4.5	Comparing BPR against RMSE	44
3.4.6	Evaluation of Type information through Post Filter	45
3.5	Conclusions	50
4	Target-Specific Parametrization of Multi-Relational Models	51
4.1	Learning with multiple target relations	52
4.2	Multi-Target Factorization	53
4.2.1	Optimizing models for Multiple Target Relations	54
4.2.2	Coupled Auxiliary and Target Specific Factorization	57
4.2.3	Setting up CATSMF	60
4.3	Evaluation	62
4.3.1	Comparison against the state-of-the-art	62
4.3.2	Evaluation Protocol and Metrics	64
4.3.3	Experiment I: Benchmark Datasets	66
4.3.4	Experiment II: Web Datasets	67
4.4	Conclusion	74
5	Factorization models for Semi-Supervised Classification	76
5.1	Semi-Supervised Classification	77
5.2	Related Work	80
5.2.1	Semi-Supervised learning	80
5.2.2	Multi-Relational Factorization for Semi-Supervised Classification	82
5.3	Semi-supervised classification problem formulation	83
5.4	Factorization models for Semi-supervised Classification	85
5.4.1	Neighborhood Based Feature Extraction	86
5.4.2	Semi-Supervised Learning of PNT-CMF	88
5.4.3	Learning Inductive Factorization models for Classification	90
5.5	Evaluation	91
5.5.1	Datasets	92
5.5.2	Setup	92
5.5.3	Baselines	93
5.5.4	Model Selection	94
5.5.5	Results and discussion	94

5.6 Conclusion	98
6 Conclusion	100
6.1 Summary	100
6.2 Discussion	101
6.3 Future Direction	102
Index	106
References	114