

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Significance of the study	2
1.3	Objectives	3
1.3.1	Main Objective	3
1.3.2	Specific Objectives	3
1.4	Expected Results	3
1.5	Thesis Structure	3
2	Related Work	5
2.1	Supervised Learning	5
2.2	Gradient Descent	6
2.2.1	Stochastic gradient descent	6
2.2.2	Adam	6
2.3	Interpretability	7
2.4	Feature Selection	8
2.5	Dempster-Shafer theory	9
2.6	Regression Methods	10
2.6.1	Seasonal Auto Regressive Integrated Moving Average	10
2.6.2	K-Nearest Neighbors Regression	10
2.6.3	Support Vector Regression	11
2.6.4	Gaussian Process Regression	11
2.6.5	Regression Tree	12
2.6.6	Random Forest	13
2.6.7	Gradient Boosting	13
2.6.8	Artificial Neural Network	14
2.6.9	Long Term Short Memory Recurrent Neural Network	15
3	Model proposal	17
3.1	Evidential Regression	18
3.2	Improving Computing Time	21
3.3	Weighted Evidential Regression	24
3.4	Weight Learning	26
4	Model Evaluation	30
4.1	Synthetic data	30

4.2	Well-known datasets	36
4.2.1	Boston housing	38
4.2.2	Medical expenses	40
4.2.3	Red wine quality	42
4.2.4	Abalone	43
4.3	Results	45
5	Real case 1: Store Traffic	49
5.1	Related Work	50
5.2	Data and problem description	51
5.3	Feature vector	52
5.4	Experiments and results	55
5.4.1	Prediction performance	56
5.4.2	Feature selection performance	57
5.4.3	Confidence interval	58
6	Real Case 2: Health Care Costs	61
6.1	Related Work	62
6.2	Data and Problem Description	63
6.3	Feature Vector	65
6.4	Experiments and Results	67
7	Conclusions and future work	72
	Bibliography	75
	Appendices	85
A	Methods Configuration	86
A.1	Model evaluation	86
A.1.1	Synthetic data	86
A.1.2	Well-known data	87
A.2	Store Traffic	88
A.3	Health Care Costs	88