

Contents

1. Introduction	1
1.1 Motivation.....	2
1.2 Our contributions.....	2
1.3 Organization of the report.....	3
2. Background	4
2.1. Classification Algorithms.....	4
2.1.1. Naive Bayes.....	4
2.1.2. Support Vector Machine.....	5
2.1.3. Random Forest.....	6
2.2. Data mining in Weka.....	7
2.2.1. Weka data formats.....	7
2.2.2. Loading the data into Weka.....	7
2.2.3. Cross validation in Weka.....	8
2.2.4. Interpretation of the obtained result.....	8
2.3. Validation.....	8
2.4. Rough Set Theory.....	10
2.4.1. Basic Rough Set concept.....	11
2.4.2. Decision rules.....	12
2.4.3. Rough Set rule induction algorithms.....	13
3. Related Work	14
Existing Approaches.....	14
4. Work done	20
4.1. Feature Extraction and classification.....	20
4.2. Rule set generation.....	21
5. Results and Analysis	22
5.1. Classification	22
5.1.1. Naive Bayes classifier.....	22
5.1.2. Random Forest.....	24
5.1.3. Support Vector Machine.....	26
5.1.4. Ensemble of the classifiers.....	28
5.2. Rule set generation.....	30
5.2.1. Reduct.....	30
5.2.2. Core.....	31
5.2.3. Rule set.....	32
6. Conclusion and Future Work	33
References	34