

CSCI 446 Artificial Intelligence Project 2 Final Report

ROY SMART

NEVIN LEH

BRIAN MARSH

November 20, 2016

Abstract

1 INTRODUCTION

2 DATASETS

2.1 Discretization

2.2 Stratified Cross-Validation

2.3 Missing Values

3 k -NEAREST NEIGHBORS

3.1 Training

3.1.1 Constructing Probability Table

3.2 Validation

3.2.1 Value Distance Metric

3.2.2 Determining k and p

4 NAIVE BAYES

4.1 Training

4.1.1 Constructing Probability Table

4.2 Validation

4.2.1 Determining Class Probability Distribution

5 TAN

5.1 Training

5.1.1 Constructing Probability Table

5.1.2 Constructing Augmented Tree

5.2 Validation

5.2.1 Determining Class Probability Distribution

6 ID3

6.1 Training

6.2 Validation

7 RESULTS

7.1 Algorithm Convergence

7.2 Algorithm Precision

8 CONCLUSION