

# CS-534 Machine Learning

## Implementation Assignment 3

---

Mohammand Velayati (33.3%)    Lucas Wells (33.3%)    Chirag Shah (33.3%)

### Introduction

#### Part I

(a) **Decision tree learning algorithm**

(b) **Information gain**

Table 1: Information gain for each feature and threshold at the root node

$x$	$\theta$	$I(Y, X)$
$x_1$	5.15	0.38
	5.35	0.48
	5.55	0.59
	5.95	0.48
	6.15	0.4
	6.25	0.37
	6.35	0.3
	6.55	0.18
	6.75	0.16
$x_2$	2.1	0.01
	2.35	0.03
	2.55	0.08
	2.65	0.1
	2.85	0.2
	3.15	0.34
	3.25	0.32
	3.35	0.3
	3.85	0.07
$x_3$	2.45	0.92
	4.85	0.7
$x_4$	0.8	0.92
	1.35	0.67
	1.45	0.64
	1.55	0.64
	1.75	0.68

- (c) Training and testing error vs  $k$
- (d) Effect of  $k$  on testing accuracy

## Part II

- (a) Feature bagging
- (b) Bootstrap aggregation
- (c) L decision trees
- (d) Training and testing errors vs  $k$  for each  $L$
- (e)