CS 188 Section Handout: Classification

1 Introduction

In this set of exercises we will experiment with a binary classification problem. The data comprise a training set of *feature* vectors with corresponding *class* labels, and a test set of unlabeled feature vectors which we will attempt to classify with a decision tree, a naive Bayes model, and a perceptron.

Scenario: You are a geek who hates sports. Trying to look cool at a party, you join a lively discussion on professional football and basketball. You have no idea who plays what, but fortunately you have brought your CS188 notes along, and will build some classifiers to determine which sport is being discussed.

Training data: Somehow you come across a pamphlet from the Atlantic Coast Conference Basketball Hall of Fame, as well as an Oakland Raiders team roster. You study these to create the following table:

Sport	*Position*	*Name*	*Height*	*Weight*	*Age*	*College*
Basketball	Guard	Michael Jordan	6'06"	195	43	North Carolina
Basketball	Guard	Vince Carter	6'06"	215	29	North Carolina
Basketball	Guard	Muggsy Bogues	5'03"	135	41	Wake Forest
Basketball	Center	Tim Duncan	6'11"	260	29	Wake Forest
Football	Center	Vince Carter	6'02"	295	23	Oklahoma
Football	Kicker	Tim Duncan	6'00"	215	27	Oklahoma
Football	Kicker	Sebastian Janikowski	6'02"	250	27	Florida State
Football	Guard	Langston Walker	6'08"	345	27	California

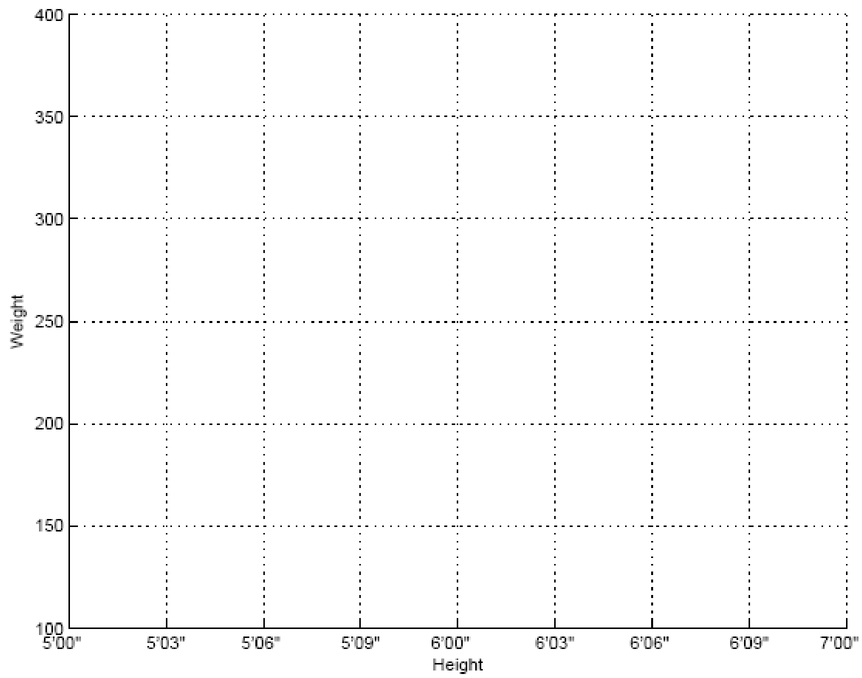
Test data: You wish to determine which sport is played by these two subjects of discussion:

Sport	*Position*	*Name*	*Height*	*Weight*	*Age*	*College*
?	Guard	Charlie Ward	6'02"	185	35	Florida State
?	Defensive End	Julius Peppers	6'07"	283	26	North Carolina

2 Perceptron (and MIRA)

A perceptron can be applied to input features with continuous values. The weights of the perceptron define hyperplanes that represent decision boundaries for classification. For these problems, consider only the input features $f_1 = Height$ and $f_2 = Weight$, along with a bias $f_0 = 1$.

Linear separability: A binary or two-class perceptron can only represent linearly separable functions. Is either feature f_1 or f_2 alone sufficient to separate the training data? Plot the training data features below to determine if a linear separator exists.



Binary vs. multi-class updates Write down the multi-class updates for perceptron and MIRA and adapt them for binary classification (you only need a single weight vector w).

Training: Use the perceptron and MIRA updates to fill in the table below. Since this is a binary classification problem, you only need to keep track of a single weight vector, w, which is initialized to $(0, 0, 0)$, corresponding to (f_0, f_1, f_2).

Training Data		**Perceptron Update**		**MIRA Update**		
$f(x)$	y^*	y'	w	y'	τ	w
–	–	–	$(0,0,0)$	–	–	$(0,0,0)$
$(1, 6\frac{6}{12}, 195)$	B					
$(1, 6\frac{2}{12}, 295)$	F					
$(1, 6\frac{6}{12}, 215)$	B					
$(1, 6\frac{0}{12}, 215)$	F					
$(1, 6\frac{8}{12}, 345)$	F					
$(1, 5\frac{3}{12}, 135)$	B					