National University of Computer and Emerging Sciences, Lahore Campus

Course:	Artificial Intelligence	Course Code:	CS401
Program:	BS(Computer Science)	Semester:	Fall 2016
Duration:	20 min	Total Marks:	10
Paper Date:	21-03-17	Weight	2%
Section:		Page(s):	
Exam:	Quiz 3	Reg. No	

Instruction/Notes:

Given the training data what attribute will be chosen as root node? Using Gini as impurity measure.

Training Data

Training Data				
Color	Size	Act	Age	Class(Inflated)
PURPLE	SMALL	DIP	CHILD	F
PURPLE	SMALL	DIP	CHILD	F
PURPLE	LARGE	DIP	CHILD	F
PURPLE	LARGE	DIP	CHILD	F
YELLOW	SMALL	STRETCH	ADULT	T
YELLOW	SMALL	STRETCH	CHILD	T
YELLOW	SMALL	DIP	ADULT	T
YELLOW	LARGE	STRETCH	ADULT	T
YELLOW	LARGE	STRETCH	CHILD	T

$$\text{Gini}(t) = 1 - \sum_{i=0}^{c-1} [p(i|t)]^2, \qquad (4.4)$$

Where C is number of classes, and P(i|t) is fraction of records that belong to class i at node t.

Question 2 Classify the following test instance using kNN and training data given in above table. Use K=3

PURPLE	LARG E	STRET CH	CHILD	T

All the attributes are symmetric (in case of binary)