

Grading Criterion for Term Project Phase II

Advanced Data Structures and Algorithms

CS310, Fall 2010

Student Name: _____ TA: Binh Tran

Part 1: Classification using WekaJ48 and Feature Selection Using Randomized Algorithms (230 pts)	Points
<p>Write a Java program to do following tasks: classification using WekaJ48, feature selection by a randomized algorithm on the real-world Mars data (230 pts)</p> <ol style="list-style-type: none"> <li style="margin-bottom: 10px;"> 1. Compiles and runs correctly (20 pts) <ol style="list-style-type: none"> a. No warning or errors when compiling and running the program (10 pts) b. Output the best feature and the highest F1 score corresponding to that best feature (10 pts) <li style="margin-bottom: 10px;"> 2. Code Review (190 pts) <ol style="list-style-type: none"> a. Classification (50 pts) <ul style="list-style-type: none"> • Implement correctly the classification by using WekaJ48 to build a classifier using training data set train.csv b. Feature Selection & Randomized algorithm (140 pts) <ul style="list-style-type: none"> • Implement a correct a randomized algorithm to find the best feature that can achieve the highest F1 score on test set (100 pts) • Correct logic of find the best feature (40 pts) <li style="margin-bottom: 10px;"> 3. Programming Style (10 pts) Generating Javadoc, header comment, no stale comments proper, consistent indenting, well chosen variable names, etc... <li style="margin-bottom: 10px;"> 4. Testing (10 pts) Evidence of sufficient testing (examples of actual test results included, etc...) All functions in one run. 	I. _____ 1. _____ a. _____ b. _____ 2. _____ a. _____ b. _____ _____ _____ 3. _____ 4. _____
<p>Part 2: Report no less than 200 words (70 pts)</p> <ul style="list-style-type: none"> • Well organized, between 60 and 300 lines, spell-checked (5 pts) • Randomized Algorithm: discuss the time complexity of the algorithm, prove it, and draw a flowchart of the randomized algorithm (30 pts) • Data Structure: discuss your choice of data structures and algorithm design (15 pts) • Documents problems and solutions, IDE, UNIX problems, experiences in writing the programs (10 pts) • Analysis of the experimental results (10 pts) 	II. _____ _____ _____ _____ _____
Deductions:	

<ol style="list-style-type: none"> 1. Missing README.TXT, javadoc files(-5 pts) 2. Missing the program outputs (-5 pts) 3. Unbounded homework (-5 pts) 4. Incorrect submitted file name (-5 pts) 5. No hard copies or soft copies (-300 pts) 6. Missing citations/references (-300 pts) 	<ol style="list-style-type: none"> 1. _____ 2. _____ 3. _____ 4. _____ 5. _____ 6. _____
Total Points (300 pts max):	1. _____

Comment :