First Name: _____ Last Name: _____ Student ID: _____

BBM 406 – Introduction to Machine Learning (Fall 2016)

Final Exam

January 9th, 2016

Instructions

- Do not open this exam booklet until you are directed to do so. Read all the instructions first.
- When the exam begins, write your name on every page of this exam booklet.
- The exam contains **six multi-part problems**. You have **120 minutes** to earn 100 points.
- The exam booklet contains **8 pages** to the exam, including this one.
- This exam is a **closed book and notes exam**. But you are allowed to have an A4 sized *copy sheet* during this exam (you can use both sides of this sheet).
- Please write your answers in the space provided on the exam paper.
- Show all work, as partial credit will be given. You will be graded not only on the correctness and efficiency of your answers, but also on your clarity that you express it. Be neat.
- Good luck!

Problem	Points	Score
1	18	
2	20	
3	18	
4	8	
5	18	
6	18	
Total	100	

It is a violation of the Academic Integrity Code to look at any exam paper other than your own, any other reference material (books, lecture notes), or to give inappropriate help to someone or to receive unauthorized aid by someone. Please also not discuss this exam with the students who are scheduled to take a makeup exam.

Academic Integrity is expected of all students of Hacettepe University at all times, whether in the presence or absence of members of the faculty. Do NOT sign nor take this exam if you do not agree with the honor code.

Understanding this, I declare I shall not give, use or receive unauthorized aid in this examination.

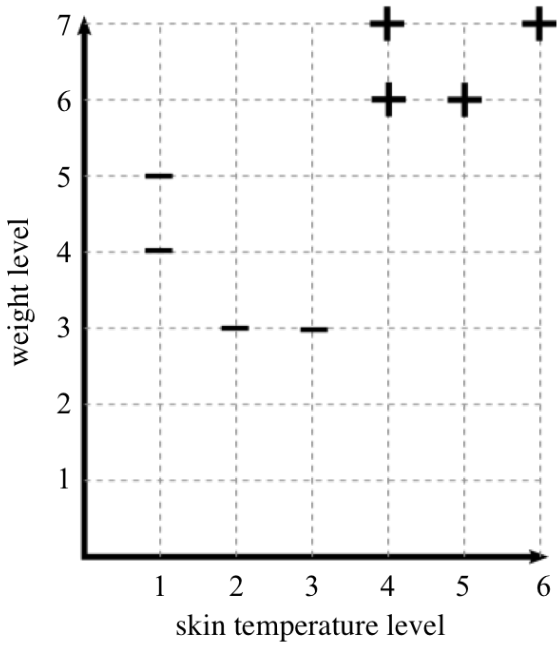
Signature: _____

First Name: _____ Last Name: _____ Student ID: _____

Q1.[18 points] **Support Vector Machines**

You have been hired by the Delos Corporation to investigate a recent incident at their Wild West-themed amusement park named Westworld, which is populated by androids. The park is technologically so advanced that one can not easily distinguish androids from ordinary humans, and your job is to train a classifier which looks at weight and skin temperature levels to decide whether the subject correspond to a human or an android. Some measurements together with their true labels are given below, where each "+" represents an android and each "–" represents an ordinary human.

© Photo from "Westworld" TV Series (2016)



(a) [6 points] You decide to use a linear SVM to classify humans vs. androids. Draw the resulting decision boundary in the graph above, and circle the support vectors.

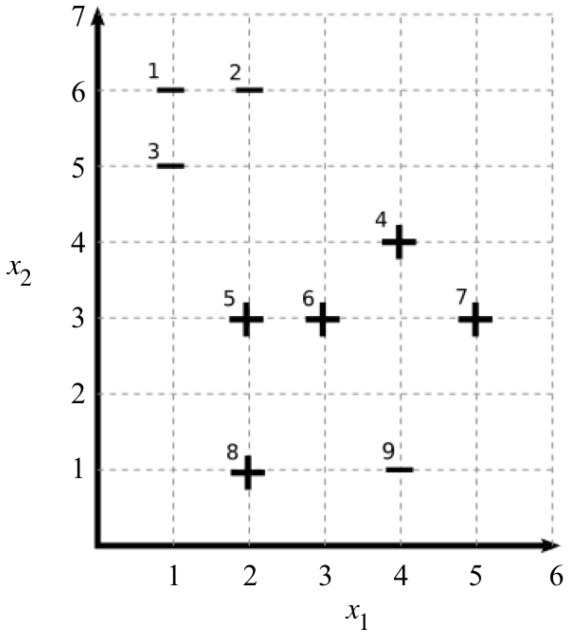
(b) [6 points] State the vector **w** and the bias b defining this linear SVM?

(c) [6 points] Suppose that Delos gunmen discover another subject whose weight level is 5 feet and skin temperature level is 6. What would the α value of this new data point be? Justify your answer.

First Name: _____ Last Name: _____ Student ID: _____

Q2. [20 points] **Kernels**

In this section, you will project the data below into a new space with $\varphi(\mathbf{u}) = \langle |x_1 - x_2| \rangle$. That is, you project the 2-dimensional vector $\mathbf{u} = (x_1, x_2)$ into a 1-dimensional vector in a 1-dimensional space.



(a) [6 points] What is the kernel function $k(\mathbf{u}, \mathbf{v})$ for this transformation?

(b) [8 points] In the transformed feature space, what is the vector $\mathbf{w}$ and the constant b?

(c) [6 points] What is the final classifier produced. Express your answer in the original feature space, not the transformed one.

First Name: _____ Last Name: _____ Student ID: _____

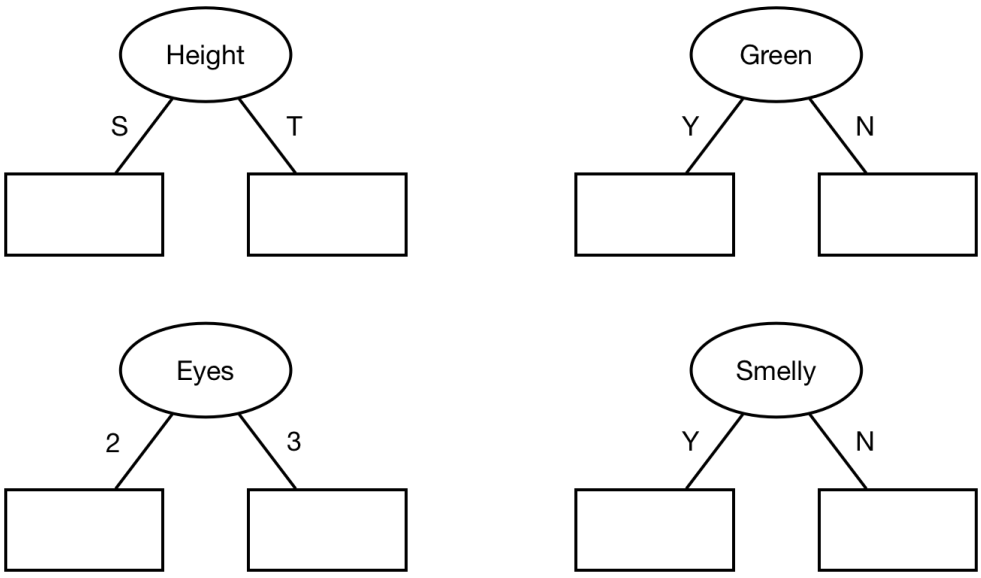
Q3.[18 points] **Decision Trees**

You want to learn a decision tree classifier to discriminate between Trolls (T) from Humans (H) based on the following four attributes: Green ∈ {Yes, No}, Eyes ∈ {2, 3}, Height ∈ {Short, Tall}, Smelly ∈ {Yes, No}

© Photo from "Trollhunters", A Netflix Original Animated Series by Guillermo del Toro (2016)

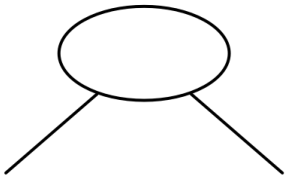
Instance	Height	Green	Eyes	Smelly	Species
x_1	S	Y	3	Y	T
x_2	T	Y	3	N	T
x_3	S	Y	3	N	T
x_4	T	Y	3	N	T
x_5	T	N	2	Y	T
x_6	T	Y	2	Y	H
x_7	S	N	2	N	H
x_8	T	N	3	N	H
x_9	S	N	3	N	H
x_{10}	T	N	3	N	H

(a) [8 points] For each possible choice of root feature (attribute) below, show the resulting species distribution. Give your answers as *m* M and *n* H.



(b) [4 points] Which root feature (attribute) would information gain select as the "best" root feature (i.e., the highest information gain)?

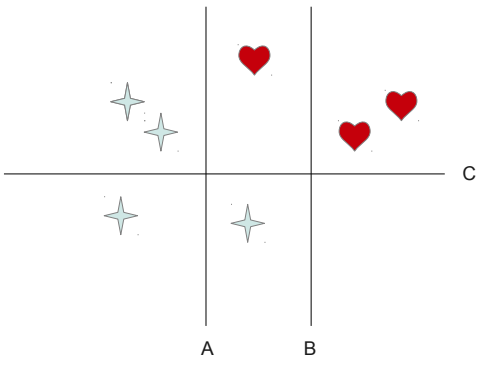
(c) [6 points] Draw the final decision tree that results from your choice of root feature (attribute) above.



Q4. [8 points] **AdaBoost**

The diagram shows training data for a binary concept where positive examples are denoted by a heart. Also shown are three decision stumps (A, B and C) each of which consists of a linear decision boundary.

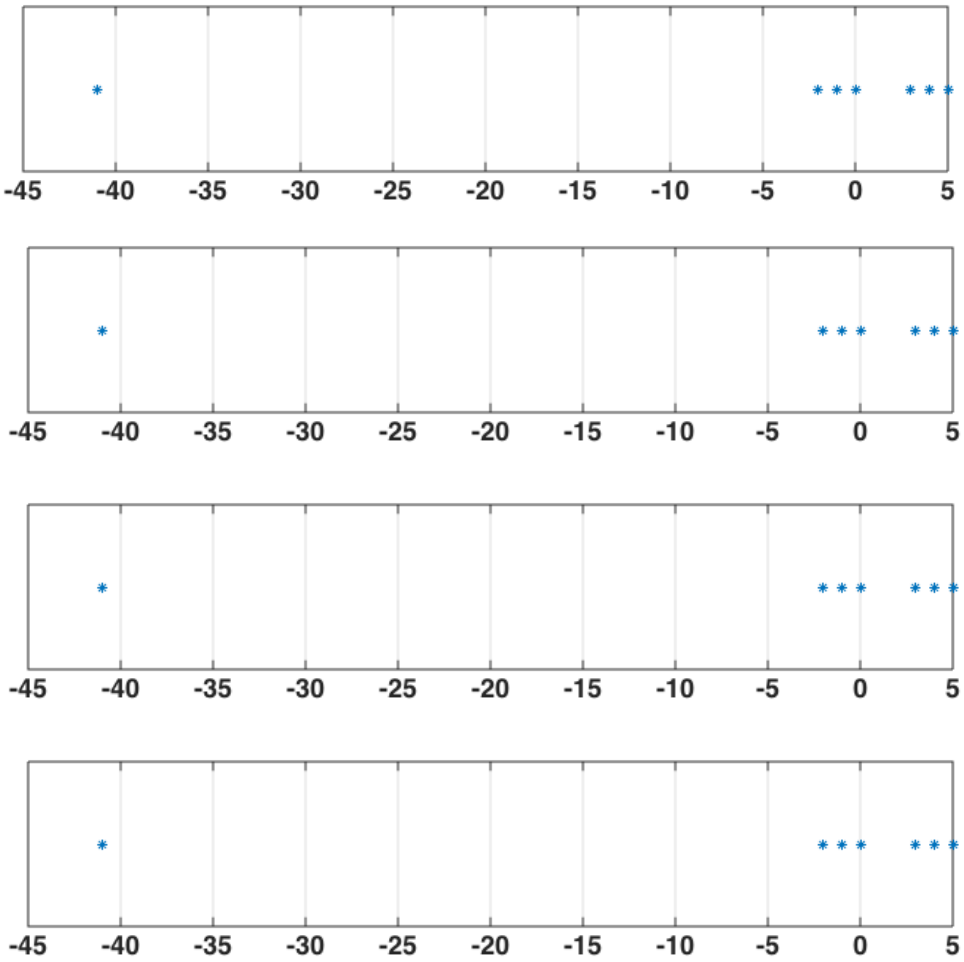
Suppose that AdaBoost chooses A as the first stump in an ensemble and it has to decide between B and C as the next stump. Which will it choose? Explain. What will be the ϵ and α values for the first iteration?



First Name: _____ Last Name: _____ Student ID: _____

Q5. [18 points] **Clustering with K-Means**

(a) [10 points] Starting with initial cluster centers at -1 and 4, perform k-means clustering on the following data points {−41, −2,−1,0,3,4,5} (denoted by *). In each panel, clearly indicate the data assignment and the new cluster means. Stop when converged or after 4 steps whichever comes first.

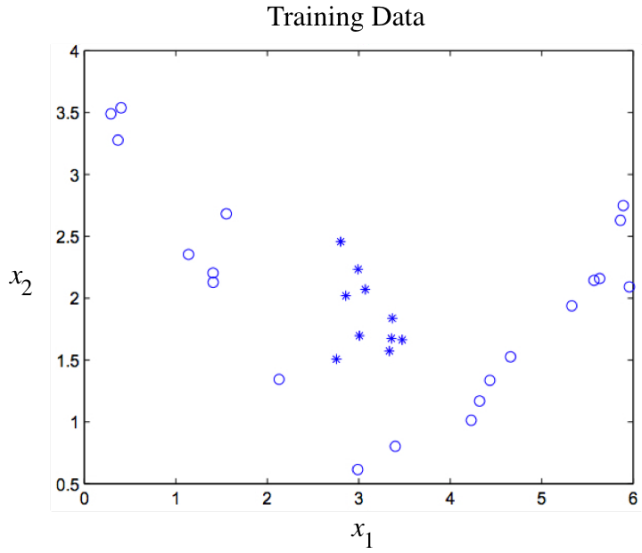


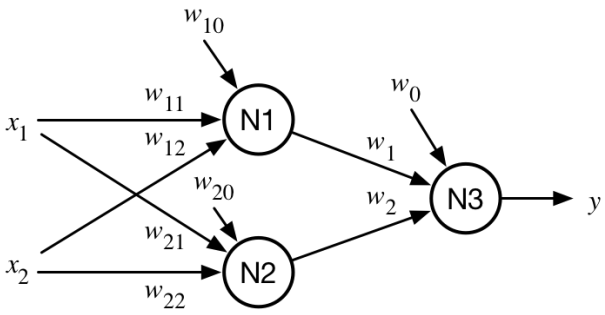
(b) [8 points] You now have reason to believe that your data has outliers. How will you change your k-means algorithm to be robust towards outliers? Throwing away data-points is not an option.

First Name: _____ Last Name: _____ Student ID: _____

Q6.[18 points] **Neural Networks**

Suppose that the following two-layer neural network (shown on the right), which uses sigmoid activation functions, is trained on the binary classification data given below and has suitable weights (w's) so that it is able to correctly classify this data set.



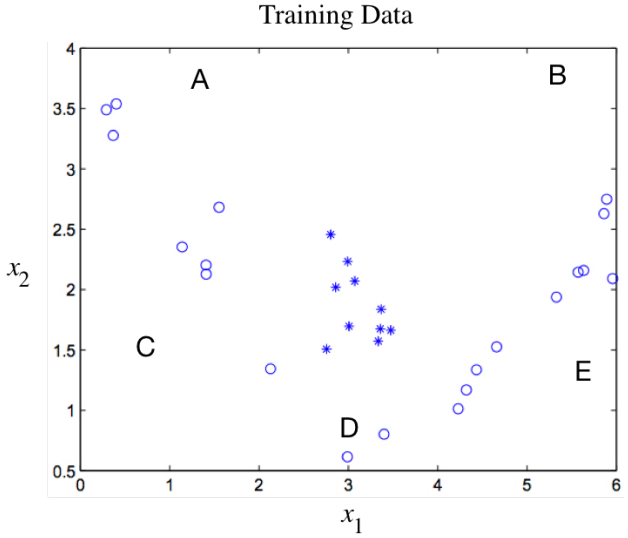


In this problem, you will plot the decision boundaries learned by each of the nodes in the network.

(a) [6 Points] The N1 node takes the feature vector (x_1, x_2) as input so you can visualize the decision boundary learned by this node in the original feature space. Plot N1's decision boundary in the figure below. That is for neuron N1, plot the line where $w_{10} + w_{11}\,x_1 + w_{12}\,x_2 = 0$.

When you draw the decision boundary in the figure below, you should notice that it passes through or close by two or more of the letters A–E in the figure. State your answer by circling in the bubbles for the corresponding letters closest to the decision boundary. For example, if your decision boundary is a curve that connects D to E to B, then circle those three letters.

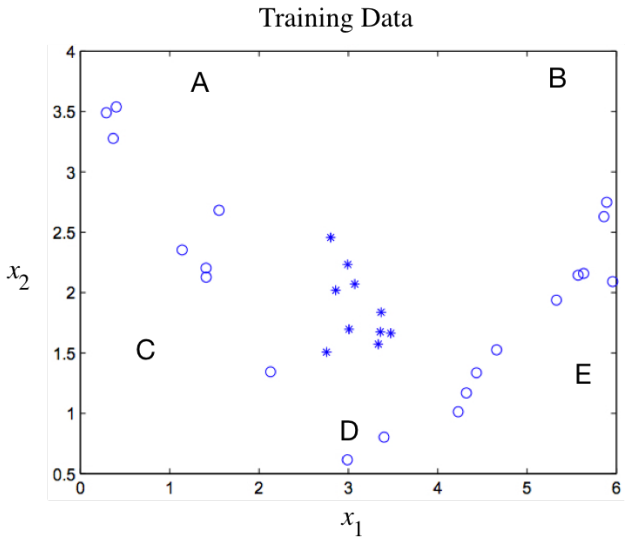
Note: There may be more than one correct answer!



First Name: _____ Last Name: _____ Student ID: _____

(b) [6 Points] Now, plot the decision boundary for N2 in the figure above, and circle the letters closest to N2's decision boundary in the figure.

Note: Your answer depends on your answer in part (a).



(c) [6 Points] Plotting the decision boundary for N3 is a bit different. N3 takes as input the output of N1 and N2, and so the graph below has axes v_1 and v_2 to correspond to the outputs of those nodes. Start by plotting $\big(v_1(x_1, x_2), v_2(x_1, x_2)\big)$ for a few training points. Once you have a good idea of the input space to N3, draw a decision boundary such that the neural net will correctly classify the data. Once again, specify your answer by circling the letters closest to that decision boundary in the figure.