EECS 1001 Machine Learning Multiple Choice

Instructor: Jeff Edmonds

1. In classical computer algorithms, the software that instructions what and how to compute is painstakingly written by humans. In a neural net,
   Indicate which of A-D is wrong or state all correct.

   (a) In a neural net, the instructions are 100% dictated by the weights on the edges.

   (b) The weights are painstaking set. This is why there are so many jobs in machine learning.

   (c) These weights are learned automatically using gradient descent.

   (d) We dont understand the resulting neural network.

   (e) All of these are correct.

   - Answer: b

2. Are there any parallels between machine learning and evolution?
   Indicate which of A-D is wrong or state all correct.

   (a) Both start with simple structures, some low level building blocks.

   (b) Both use just a little random change that is encouraged just a little by some feed back to go in the right direction.

   (c) Both show emergent complexity: And what arises is awe inspiring, producing machines that are amazingly powerful.

   (d) Jeff feels like this a glimpse into how he, himself, came to be. He finds it oddly spiritual.

   (e) All of these are correct.

   - Answer: e

Extra The following are characteristics of a problem that indicate that machine learning may succeed where traditional coding methods failed.
Indicate which of A-E is wrong.

(a) Software too complex to write by hand

(b) Human expertise does not exist (navigating on Mars),

(c) Learn to calculate payroll and taxes

(d) Humans are unable to explain their expertise (speech recognition)

(e) Solution changes over time

- Answer: c

3. Positive things that machine learning will likely bring the average person in the next 20 years.
   Indicate which of A-E is wrong.

   (a) Speech recognition, Natural language processing

   (b) Self driving cars,

   (c) Robots

   (d) There will be lots of machine learning jobs.

   (e) Medical: Give machine symptoms and get diagnoses. AI is getting better than doctors at reading an x-ray

   - Answer: d

4. Negative effects that machine learning may have on the average person in the next 20 years.
Indicate which of A-E is wrong.

   (a) In 10-15 years, half current jobs will be obsolete.

   (b) In 20 years, machines will be much smarter than humans in every way.

   (c) The authorities being able to find a criminal in a crowd might be scary for those being found.

   (d) Is it fair? Are the criteria for the decisions understood?

   (e) War: Should humans build AIs that decide when to kill humans?

   - Answer: b

5. Suppose we want our machine to recognize whether a given image is a face or a bicycle.
Indicate which of A-E is wrong.

   (a) The first step in recognizing a face is finding the nose.

   (b) $\vec{x}_d$ is the d^{th} training image

   (c) y_d is the answer given to us by the supervisor of whether it is a face or a bicycle.

   (d) $\vec{w}$ is the vector of weights/parameters defining the machine $M_{\vec{w}}$. It is basically the software. It is what is learned.

   (e) $M_{\vec{w}}(\vec{x})$ is the answer given by this machine as to whether the image $\vec{x}_d$ is a face or a bicycle.

   - Answer: a

6. The error $E(\vec{w})$ of the machine is the sum over all the training data of the square of the distance between the answer given by the supervisor and that given by the machine. The following are reasons we square this distance.
Indicate which of A-E is wrong.

   (a) To make it positive so that errors in the negative and positive direction do not cancel each other.

   (b) To punish big errors even more.

   (c) Absolute values make the math near impossible.

   (d) Quadratics were learned in high school.

   (e) The error is a function of $\vec{w}$ because the training data is fixed and our goal is to find the weights $\vec{w}$ that minimize this error, namely $\vec{w}_{opt} = ArgMin_{\vec{w}} E(\vec{w})$.

   - Answer: d

7. After a neural net has been trained on data labeled by a supervisor, the hope is that it will also get the right answer to input instances that it has never seen before. The following are conditions when this does or does not work.
Indicate which of A-D is wrong or state all correct.

   (a) Underfitting occurs when the machine just isnt complex enough to model the training data.

   (b) If the machine summarize the material, then hopefully it understands it and do well when a new question comes.

   (c) Occam's Razor is the problem-solving principle that the most simple solution tends to be the right one.

   (d) Overfitting occurs when the machine is specified by so many parameters/weights that it can memorize all D of the answers in the training data. It then learns nothing and hence does not know the answer for a new input.

   (e) All of these are correct.

- Answer: e

8. There is a theorem that says that the trained neural net will work well on never seen before instances as long as the following are all true.
   Indicate which of A-D is wrong or state all correct.

   (a) The machine found must do well on the training data.

   (b) The training data must be randomly chosen so that it "represents" the general data.

   (c) The machine does not have enough resources to simply memorizing the answers to the training data.

   (d) Then the learned machine will always work on never seen before instances.

   (e) All of these are correct.

   - Answer: d

9. In steepest gradient descent, assume you know your current $\vec{w} = \langle w_1, ..., w_m \rangle$.
   Indicate which of A-D is wrong or state all correct.

   (a) Our goal is to finding weights $\vec{w}_{min}$ that minimize the error $E(\vec{w}_{min})$.

   (b) Forward propagation calculates our height $E(\langle w_1, ..., w_m \rangle)$.
   Backward propagation calculates the direction of steepest decent.
   We change $\langle w_1, ..., w_m \rangle$ by a small step in the direction that the hill is steepest down.

   (c) Weights w_k that influence correct answers are increased and those that influence wrong answers are decreased. This is determined by $\frac{\delta E}{\delta w_k}$.

   (d) Stop when you are at the bottom of a valley, i.e. every direction is up. This will be the globally minimum location.

   (e) All of these are correct.

   - Answer: d