# MM AI Text Mining Assignment: Chapter 6 “A Closer Look at Machines that Learn”

1. TRUE/FALSE - The learning-from-data approach of deep neural networks has generally proved to be more successful than the “good old-fashioned AI” strategy, in which human programmers construct explicit rules for intelligent behavior. However, contrary to what some media have reported, the learning process of ConvNets is not very humanlike.
   True.
2. Why does your professor like the previous question?
   Even though the ConvNets and human thinking are similar in the way that it is hard to explain what exactly is happening in decision making, the two are still so far apart in how they learn.
3. TRUE/FALSE - As we’ve seen, the most successful ConvNets learn via a supervised-learning procedure: they gradually change their weights as they process the examples in the training set again and again, over many epochs (that is, many passes through the training set), learning to classify each input as one of a fixed set of possible output categories.
   True.
4. List some significant differences between the way that humans learn about objects and the way that ConvNets learn about objects.
   Humans do not need as many examples as ConvNets do and humans are inquisitive creatures. We don’t learn by just knowing what objects are called; we ask questions about them and make inferences about them and what they relate to.
5. Why is it inaccurate to say that today’s successful ConvNets “learn on their own?”
   ConvNets do not learn on their own because it takes a tremendous amount of human effort for ConvNets to be successful. Humans have to create the ConvNet itself and they have to label a large amount of data for ConvNets to be successful.
6. In answer to the rhetorical question “Where does all of the data come from to fuel big data applications?,” MM answers “You - and probably everyone you know.” Please elaborate on the answer.
   The data comes from people who use online services, anything that can get posted online can be used as data.
7. How do car companies acquire the big data (labelled images of pedestrians, cyclists and other obstacles) needed to train robo-cars?
   Car companies use dashcam footage to train self-driving cars.
8. What is the “long tail” phenomenon, and how does it relate to machines that learn (ConvNets)?
   The long tail phenomenon is a statistics term which refers to a large amount of events with a very low probability of occurring. This relates to machine learning because it is

impossible to teach a machine every possible unlikely event through the use of ConvNets.

9. TRUE/FALSE - A commonly proposed solution to the long tail problem in AI systems is to complement supervised learning with unsupervised learning.
   True.
10. What is “unsupervised learning?”
    Unsupervised learning is the use of unlabeled data to teach machines.
11. What colorful remark did Yann LeCun make about unsupervised learning?
    Unsupervised learning is the dark matter of AI.
12. TRUE/FALSE - For general AI, almost all learning will have to be unsupervised, but no one has yet come up with the kinds of algorithms needed to perform successful unsupervised learning.
    True.
13. TRUE/FALSE - Humans have a fundamental competence lacking in current AI sysgtems: common sense. We have vast background knowledge of the world, both its physical and social aspects. We have a good sense of how objects - both animate and living - are likely to behave, and we use this knowledge extensively in making decisions about how to act in any given situation.
    True.
14. TRUE/FALSE - Many people believe that until AI systems have common sense as humans do, we won’t be able to trust them to be fully autonomous in complex real-world situations.
    True.
15. TRUE/FALSE - Superficial changes to images, such as slightly blurring or speckling an image, changing some colors, or rotating objects in the scene, can cause ConvNets to make significant errors even when these perturbations don’t affect humans’ recognition of objects. This unexpected fragility of ConvNets – even those that have been said to “surpass humans at object recognition’” – indicates that they are overfitting to their training data and learning something different from what we are trying to teach them, a phenomenon that results in various manifestations of unreliability.
    True.
16. The unreliability of ConvNets can result in embarrassing – and potentially damaging – errors. Select a particularly embarrassing/damaging example of unreliability in ConvNets, and de- scribe it in just a sentence or two.
    One damaging example of unreliability of ConvNets relates to self-driving cars, obstructions to road signs, such as a sticker on a stop sign could fool the car.
17. At the end of the section on biased AI, MM observes that the problem of bias in applications of AI has been getting a lot of attention recently, with many articles, workshops, and even academic research institutes devoted to this topic. What questions does she raise in conjunction with this observation? What do you think are the appropriate answers to these questions?

The questions she raises are should the data sets being used to train AI mirror our own biased society or should they be changed to achieve social reform aims and who should be allowed to control it. I think we should always do what we can to achieve social reform to benefit everyone, but I do not know who should be in control of it.

18. TRUE/FALSE - You can often trust that people know what they are doing if they can explain to you how they arrived at an answer or a decision. However, "showing their work" is something that deep neural networks – the bedrock of AI systems – cannot easily do.
    True.
19. TRUE/FALSE - Recall that a convolutional neural network decides what object is contained in an input image by performing a sequence of mathematical operations (convolutions) propagated through many layers. For a reasonably sized network, these can amount to billions of arithmetic operations. While it would be easy to program the computer to print out a list of all the additions and multiplications performed by a network for a given input, such a list would give us humans zero insight into how the network arrived at its answer. A list of a billion operations is not an explanation that a human can understand.
    True.
20. What, according to MIT's Technology Review is the dark secret at the heart of AI?
    The dark secret at the heart of AI is explanations of the decisions that neural networks make.
21. What does the phrase "theory of mind" refer to, and how is it related to our interactions with AI systems such as deep networks?
    A theory of mind is what we humans make in relation to what others think and their goals in particular situations. This is related to deep networks because no one has an accurate theory of mind for how they make decisions.
22. One of the hottest new areas of AI is variously called "explainable AI," "transparent AI," or "interpretable machine learning." To what do these terms refer?
    These terms refer to developing AI that can explain the choices it makes.
23. The field of "adversarial learning" has emerged in response to the fact that AI systems can readily be fooled in dramatic fashion, like mixing up a guy in glasses with Milla Jovovich, or misclassifying a stop sign for a speed-limit sign. Briefly describe the field of adversarial learning.
    Adversarial learning is a field of AI research that's goal is to counteract the ways that can be used to fool AI.
24. Jeff Clune, an AI researcher at the University of Wyoming, made a very provocative analogy when he noted that there is "a lot of interest in whether Deep Learning is 'real intelligence' or a 'Clever Hans.'" Explain the essential question that underlies this analogy, being sure to incorporate a few words on the actual Clever Hans.
    The question that this analogy brings up is whether or not deep learning is actually using intelligence to operate, or they are like the horse Clever Hans, who was thought to be

intelligent at math but instead was just responding to unconscious cues given by the questioner.