Tyler Prehl
CSC 481 - W3
5/1/22

Task 2:

The exciting aspects of “The Guardian view on bridging human and machine learning: it’s all in the game” lay with the potential of AI beating another human-mastered game, rationalizing its decisions, and becoming more “insight” than computation based. The skeptical aspects of the article include its loose usage of certain words (such as “wining” or “beating”), general position on how “black box” AI cannot be understood, and then calling out recent shortcomings in the AI industry as clear evidence of the limitations of machine learning but trying to simultaneously differentiate it from machine learning that can rationalize its decisions.

presents a discussion about breaking out of the “black box” AI issue - that an AI may arrive at a given solution that cannot be rationalized after the fact, but was found through a machine learning process. It is always exciting to hear when AI makes advancements in typically human activities, especially when it comes to mastering human games. In this case, the game in assessment is the card game “Bridge.” However, when the author mentions that the AI NooK won by formulating rules instead of brute force computation, I became skeptical of their understanding of AI, since much of AI is rule creation/revision such as in reinforcement learning. Upon further inspection of a link to a different article (“Artificial intelligence beats eight world champions at bridge” by Laura Spinney of The Guardian), I found that NooK is far from actually “beating” Bridge. NooK was not even required to play the full game, “it did not involve the initial bidding component of the game during which players arrive at a contract that they must then meet by playing their cards,” nor did it play against real players. Its performance against other Bridge AI was simply compared, and it performed better given certain hands than the human players. Because the author of “...it’s all in the game” claims that NooK “beat” Bridge, the reader should be wary of information to come considering the very loose definition of “beat” that the author is using.

The key aspect of Bridge that the author focuses on is the information sharing piece. If a player wants to share information with their partner, they must also share that information with their opponents. Because of this, the AI NooK was required to be able to rationalize its decisions in language that humans can understand, which is described by the author - “This feature of bridge meant NooK could explain how its playing decisions were made, and why it represents a leap forward for AI.” Although this is certainly excellent and exciting, tracking what rules an AI is using to make its decisions does not seem wildly difficult considering our in-class experiences with reinforcement learning. What makes me most skeptical is when the author claims that a black box AI spends millions of iterations to improve its performance, but that it’s a mystery how it arrived at its conclusions. Although people who never saw the code may not understand how the AI came to its conclusions, the developers crafted algorithms that weigh certain positive and negative outcomes and provide direction for the AI to learn through them. As Dr. Bender alludes in her Twitter post, this type of writing by the author makes it seem as though the AI is fully independent of humans, but in reality, there is still a code base to point to and say “look at the various outcomes that cause the AI to ‘learn’.”

Despite the drawbacks in this article, the author does refer to the famous British AI specialist Donald Michie about his statement regarding strong vs “weak” AI. More specifically, “NooK nods to the work of British AI pioneer Donald Michie, who reasoned that AI’s highest state would be to develop new insights and teach these to humans, whose performance would

be consequently increased to a level beyond that of a human studying by themselves. Michie considered ‘weak’ machine learning to be just improving AI performance by increasing the amount of data ingested.” Unfortunately, the author then goes on to tear apart all of the lack of AI progress that has been made in the past year, attributing it to “weak” AI’s limited ability to learn with statements such as “[Mitchie’s] insight has been vindicated as deep learning’s limits have been exposed. Self-driving cars remain a distant dream.” While it may be true that increasing data ingestion may have a diminishing return, the author’s solution is for AI’s decisions to become understandable to humans. Dr. Bender would likely consider - although it would certainly be helpful for analysis, would this really change the AI’s decisions? The overall quality of this article must be seriously doubted because of the several points of general terminology and gaps of logic that the author assumes as it refers to AI.