George Archbold

CS35L

Lab 10 Report

<div align="center">AI Anxiety</div>

When I started my topic search for this assignment I knew I wanted to do a topic related to AI. I had just watched the movie Ex Machina, which explores the possibilities of AI-human coexistence and raises a lot of interesting points about technological and ethical questions surrounding the advancement of AI. The article I chose was published by The Washington Post, and titled "The AI Anxiety". Although the article did not delve very deep into the topics of AI, it introduced me to the leading scientists, Ray Kurzweil and Nick Bostrom, and the major talking points of their theories.

The article first introduces Nick Bostrom, a Swedish philosopher who has spoken out strongly against the reckless advancement of AI technology. His predictions tend to be very pessimistic and cynical, assuring that AI technology will lead to evil robots hell bent on the destruction of mankind, or a myriad of other technological doomsday scenarios. Bostrom is not the only one who is worried about rogue technology having unexpected adverse effects on the human race; Stephen Hawking, Elon Musk, Bill Gates and many other leading figures in the tech world have created a research fund committed to the responsible and ethical development of such technologies. Although the realm of AI is very exciting and already very integrated into our society today, with computers getting smarter and smarter, and the development of nanotechnology and genetic engineering, the fear is that this technology will compound on itself at a rate faster than we can predict or control.

The article then introduces Ray Kurzweil, who is the currently director of engineering at Google. He has been a major figure in the world of AI since its inception to the forefront of the technological spotlight. He is dubbed the most famous techno-utopian because "he believes that technological progress will culminate in a merger of human and machine intelligence" (Washington Post). I was more interested in the theories of Ray Kurzweil as I would like to think technology will help advance us as a society an not destroy us, and thus I found that he had written many widely acclaimed books, one being *The Singularity is Near.* The concept of the singularity is what intrigued me most about this whole study. It is the point at which AI is

perfected and parallels human intelligence. This level of intelligence and cognizance combined with the computing of power of a machine, which is unattainable to humans whose brains, in their natural states, are finitely limited, would result in an exponential explosion of advancement. After reading through parts of the book, I traced his projected timeline of advancement and tend to align my future predictions with his rather than the grimmer outlook of Nick Bostrom.

Kurzweil believes that AI will benefit humanity because he believes that by the time the human race reaches the singularity there will be no distinct binary between man and machine that could define two clear clashing sides. He predicts that machines will be a part of us, making humans a breed of cyborg. He predicts that engineering will get more precise and smaller computers will become more powerful eventually leading tot eh creation of nanotechnology that becomes distributed through the human body, aiding in the fight against disease, metabolism of food and energy, and general function of the body. Progressively the human body will adapt and advance through the introduction of such technology and these advancements will begin to compound and occur exponentially faster and faster. People often think of advancement linearly, so when they thinking about the future they look backwards at how much we have accomplished in the past. He describes this misconception in the image shown below.
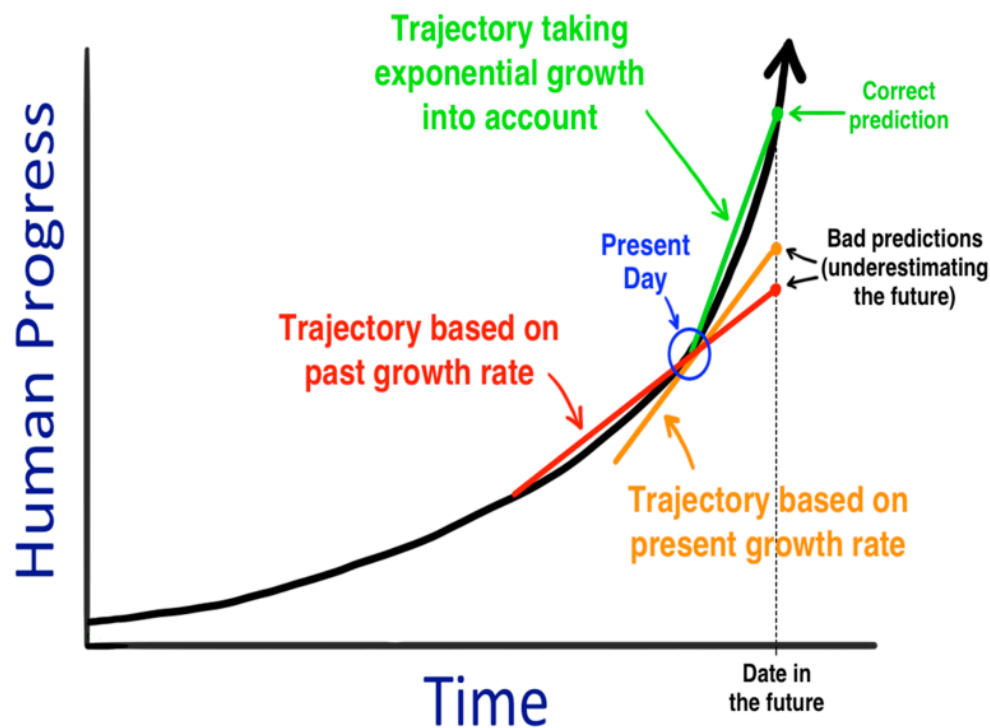
After the completion of the reverse engineering of the brain he predicts that humans will abandon their attachment or dependence on a biological body, and will live as an uploaded version of their brain, thus resulting in immortality in a sense. After this it all becomes quite scifi, and regardless of how accurate it may be, it all seems some what unfathomable. This idea of the singularity interested me the most because it was so far beyond the scope of what I had imagined for the foreseeable advancement of the tech world. Kurzweil explains the phenomena of exponential technological advancement, and how the more we discover the more it enables us to make the next great discovery even quicker. He gives an example of this, comparing his iPhone to his computer, which was one of the most powerful of its times, in his MIT lab forty years ago. His iPhone today is one millions times cheaper, one million times smaller, and one thousand times as powerful. He makes the point that in forty years, the technology that will exist, given the steeper slope of the exponential curve, is unfathomable.

The article then goes on to address the measures being taken to ensure the safe and responsible development of AI technology. It also discusses less prominent futurists who don't believe that the singularity is a likely occurrence in our life time or for centuries to come. With the emergence of a strong community of technical leaders organizing and collaborating to plan the best possible course of action for the development of AI technology, I cannot wait to see what kind of technology emerges and can only dream to be a part of its creation.