Machine Learning, Homework 1

Ben Senator

October 14, 2024

Question 1. *Will the cumulative number of people who have gone to college converge or diverge?*

Under the assumption that the human population will continue to exist into the distant future, and constraining our thinking to that future, the cumulative number of people who have gone to college will diverge.

Question 2. *The release of chlorofluorocarbons (CFCs) was banned by treaty and CFCs decay in the atmosphere over time. Would we expect the quantity of CFCs to converge after an arbitrarily long period of time without new emissions? Explain.*

We would indeed expect the quantity of CFCs to converge after an arbitrarily long period of time without new emissions. Specifically, we would expect the CFCs to decay to a count of 0. Consider that each CFC emitted affects the overall count of CFCs by $+1$ and each CFC that decays affects the overall count -1. Given a limited number of CFCs and that all *must* decay eventually, this series converges to 0.

Question 3. *Which series converges the fastest:* $\frac{i^2-3}{i^3}$, $\frac{i^3-3}{i^5}$, *or* $\frac{100}{i^3}$*? Explain.*

The fastest-converging series is $\frac{100}{i^3}$ as this series has the largest ratio of denominator-heavy powers.

Question 4. *Consider two algorithms: (A) has quadratic convergence and requires* $\mathcal{O}(n^3)$ *operations and (B) has linear convergence and requires* $\mathcal{O}(n^2)$ *operations. What would be the characteristics of a data science project when algorithm (B) would be preferable?*

As there are an exponentially increasing number of required operations in both algorithms but algorithm (B) has a lower rate of exponential growth than algorithm (A), we would desire algorithm (B) when n is large. This would save us time and computational energy.

Question 5. *Describe whether and how the probability of rain on a given day could be described as a series with the Markov property.*

This series would likely fail to be considered to hold the Markov property, because we know that meteourologically, weather systems cause rain to occur non-randomly in terms of time and space. That is, whether or not it rained on the previous day would influence the probability of it raining today. As such, we should know the past history of the system to make predictions of the system in the present, which is a contradiction of the Markov property.

Question 6. *Describe how representation bias could affect a facial recognition algorithm?*

A facial recognition algorithm would need to be trained or a representative sample of the population so that it can accurately and reliabily function across people of different race, gender, and other relevant characteristics that influence facial appearance. If we did not have a representative sample to train our algorithm, our algorithm would likely struggle to perform on underrepresented populations.

Question 7. *Describe how algorithmic practices can lead to poor outcomes and one possible approach to mitigating algorithmic bias.*

Algorithmic practices can lead to poor outcomes by ignoring the potential social inequities that can result from societal progress through practice alteration resulting from algorithmic deployment. For example, an algorithm may be developed and deployed in a corporate setting that all but removes labor demand for low-skilled workers, raising significant concerns for such workers to gain employment. We would attempt to mitigate such biases by working closely with those who define the problem to be tackled, to understand not just how our algorithm could be developed and deployed, but whether it should be developed in the first place given its societal implications if developed.