CS 3600 Project 4 Wrapper

CS 3600 - Fall 2022

Due December 7th 2022 at 11:59pm EST via Gradescope

Introduction

This Project Wrapper consists of a context paragraph, which identifies the topic of the wrapper, followed by four short-answer questions, each worth 1 point. Please limit your response to each question to **a maximum of 200 words**. The goal of this assignment is to train your ability to reason through the consequences and ethical implications of computational intelligence.

Context

You are working for a financial institution which is in the business of assessing risk of loaning money to potential clients. The CTO of the company has been an avid Twitter user swept up in the rumor that GPT-4 will have 100 trillion parameters, and has a dream to completely automate the loan risk assessment process. As a former CS3600 student and experienced engineer in the AI/ML team for the company the CTO comes to you to ask for a plan to implement this algorithm. You have historical data since the 1930s in this format (Note that all strings come from a fixed, enumerated set of values):

Pronouns	**Sex**	**Age**	**Zip Code**	**Race**	**Hobbies**	**Credit Score**	**Net Worth**	**Loan Amount**	**Accepted for Loan (target)**	**Name**
string	string	int	int	string	string	int	int	int	boolean	string

Table 1: Loan Dataset

Question 1

Fully engaged in the Twitter hype for GPT-4 and constantly up to date on the NeurIPS threads the CTO pitches a neural network architecture for predicting if the client will default on the loan. The CTO suggests the various columns of the table act as features to a deep neural network with 50 hidden layers. With your experience in the field you think that there could be a better model for such a sensitive task as this. Make an argument for why a decision tree is better for use in this particular situation (hint: you can mention runtime at inference, but there is an even stronger case to be made for a decision tree that we are looking for here! Especially consider that this ML model will be making highly sensitive decisions.)

Answer:

First, a decision tree would be much faster to make a decision compared to a ML model. It is also easier to understand how a decision tree may have arrived to a solution compared to why the ML model did. This is important in this application since many times the banks may need to explain why a loan is rejected. Additionally, it would be much faster to create a decision tree model compared to train a ML model. The ML model may also incorrectly weight random factors such as names or zip code which could lead to an incorrect decision. But more importantly the neural network accuracy may never converge to a high enough accuracy, and this could be due to errors made with historic data.

Question 2

The ethics review board at the company rejects the initial proposal from the CTO on the basis that algorithm could easily end up rejecting loan applicants based on race or sex. To fix this the CTO proposes that the sex and race columns be removed from the dataset. Will this completely prevent the machine learning model from discriminating based on race/sex? Why or why not?

Answer:

This probably won't completely prevent the problem since the ML algorithm may begin to correlate names or zip codes from a race or sex to rejection decisions. This is an example of data restriction but there may be secondary features that still create that bias. However, removing the direct link would drastically improve the algorithms' ability to not discriminate due to that factor specifically, it is a good first step!

Question 3

Algorithmic discrimination is a serious problem with the wide dissemination of machine learning models. (O'Neil, 2016) Unfortunately, these harms can go unnoticed due to the false assumption that math and algorithms are unbiased and objective. Machine learning models are only as good as the data used to train them. Research an instance where a machine learning model was used to make critical decisions, but was later found to be biased (excluding the example with bank loans). Summarize what the purpose of the model was, and how it ended up causing harm. Include a link to a news article or research paper discussing this issue.

Answer:

In 2015 Amazon had implemented an ML model to help with the hiring process for the company. However, they soon realized that the algorithm would prefer to hire men over women. This was due to the training data they used was the companies hires from the past 10 years which were predominantly men. This made the neural net link that the applicant being a certain gender was something that could influence if they should be hired or not which led the model to discriminate against woman.

Link: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

Question 4

What is the reason for having a separate train and test set when constructing a machine learning model? Why not use all the possible data for training and testing?

Answer:

This is so that when the model is being applied on the testing data, it is introduced to situations that it has never seen before. This will allow the programmer to see the true accuracy of the model and make sure that the model is not just overfitting the training data. If all the data was used for training and testing. The model would probably do very well on the testing data, but when introduced to completely new scenarios, we wouldn’t be able to tell if it would perform well or not.

References

Cathy O’Neil. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group, USA.