HW4

The ChatGPT states that it would not use race variables in the financial service, and I agree with that. The ChatGPT claims using race as a variable in financial service can lead to some ethical problems, I think this is very true. I also think using race as a variable in financial services can lead to biased data and increase the racial wealth gap. For instance, two people want to purchase a house, and looking to take out a loan for it. They have the same income and financial status, but one person is at a racial disadvantage because people of his race are more likely to default on a loan according to the financial service's past analysis. Therefore, the person who has a racial advantage would be more likely to get the loan even if they have the same financial status. Furthermore, even if the person who's at a racial disadvantage successfully gets the loan, they are more likely to have a higher interest rate, therefore, they have to spend more money on the loan than the other person. This can cause inequality in housing, same applies to higher education, people with racial disadvantages might hold back from attending college because they need to pay more for student loans. Therefore, I think using race as a factor in financial services can result in an increased racial wealth gap and discrimination in certain races which is unethical to do so. Moreover, the Chatbot suggests some other alternatives to avoid using race as a factor in financial service, such as using area code or neighborhood demographics as a factor in financial service instead of race. I also agree with this, since neighborhood demographics are closely related to race but not directly related to it, doing so can lower the risk of bias data. The Chatbot also includes the Equal Credit Opportunity Act and the Fair Housing Act. The chatbot uses the Social Contract Theory. It implies that the financial service should follow the rule that all people have agreed to obey for their mutual benefit.

Using race as a factor in financial services violates these rules therefore we should not use race as a factor. Even though the Chatgpt states its stance it does not use state ethical theory directly, I would directly state why we should use race as a factor in financial service. Additionally, I would add some examples to explain why using race as a factor is dangerous rather than saying it is problematic.

The ChatGPT replies with suggestions that can provide more insights into credit risk when offering loans to small businesses. I agree with the suggestion from the chatbot. I think using network traffic data to offer loans to small businesses without going deeper into the data is very dangerous. I would be collecting more information such as revenue statements, balance sheets, and cash flow statements, combined with the network traffic data to do a deeper analysis and make sure everything is looking good before offering a loan to a small business. And also be transparent about our data analysis model so our client can make the right decision without worrying about any unknown risk. The answers that the chatbot provides are not grounded in any particular ethical theory. I will be using professional ethics to elaborate on the importance of data practice ethics. As engineers we should be responsible for the product we build, especially collecting sensitive data from the user, and being transparent about how we make decisions is our priority. We should not be collecting data with a biased approach as this might harm certain groups of people. Lastly, as data engineers we should ensure that our product and related modifications meet the highest professional standards possible. Although the chatbot provides an excellent way to avoid credit risk and poor estimates for certain groups. It does not state whether or not to offer loans to small businesses when we have insufficient data. I would add some statements to clarify whether or not to offer loans to small businesses. I

would also add some examples of those suggestions and give some statistics on how well we can improve our data modeling. And discuss the trade-off when we implement that suggested method in our data model.

The ChatGPT states it would adhere to ethical guidelines and respect user privacy. Even though the chatbot did not say whether or not to consider the request, I still agree that this request might violate privacy rights and potentially lead to discrimination in the hiring process, therefore I would not agree on this request as a professorial data engineer. According to professional ethics, we should have good judgment to maintain integrity and independence in our professional judgment, so we can reject the client's request if we think the request might violate the privacy of others. The answers are not explicitly grounded in any particular ethical theory. I might use Utilitarianism to support my stance that we should not consider this request since it might harm the public interest and we should avoid any potential discrimination and bias in hiring and create a positive hiring experience for the public. We can also use virtue ethics, does not matter what profession we are in, we need to have good virtue to make decisions that respect individuals' rights. I might add some real instances of the hiring process that include bias and discrimination and show how dangerous that is and why we should avoid it. Because the chatbot did not explicitly state its stance on whether to consider the client's request, I will add this to eliminate some miscommunication.

The use of personal data such as social media activity and surveillance footage can violate individual privacy rights when the collecting data is leaked to some bad actor, and those data might threaten your safety. While the surveillance camera is installed for general security purposes, there is no way to stop recording an individual's activity when it shouldn't. For

instance, an individual is walking his dog in the park on the weekend, and there are surveillance cameras all over the park for general security purposes. This individual does not have any intent or suspicion of engaging in any criminal activity. The individual does not feel comfortable being watched as he is walking in the park. But there is no way to stop recording his activity even if there is no potential safety issue there. This can violate an individual's privacy rights. Furthermore, collecting data such as public profile photos and social media posts to predict an individual's behavior without his or her consent is also violating privacy rights.

The use of predictive policing algorithms can exacerbate these historical biases in policing. For instance, there can be over-policing in a certain neighborhood more than others because it has a higher historical crime rate over the past 10 years, however, this neighborhood has had the lowest crime rate in the county in the past 2 years. This shows that historical data might not reflect the reality when algorithm does not account for some factors such as why we even need the data from 10 years ago where this neighborhood has the highest crime rate. Algorithmic bias can happen when the predictive policing platform targets or surveils certain demographic groups and exacerbates these biases even more. while one area has surveillance cameras on every corner of the street, it will detect or capture more crimes compared to the area that has no cameras at all. Thus, it will have more policing attention in that area. But does this imply that the area has lower reported crime rates is safer than those with higher crime rates? The answer is not necessarily. Therefore, predictive policing algorithms are biased and can exacerbate these biases.

The use of predictive policing platforms can lead to some injustice in certain demographic groups or people in certain areas. For instance, a kid walking on a street at night in

a historically high crime rate area is more likely to get wrongful accusations than a kid in an area that has a lower crime rate. the kid might have psychological impacts such as anxiety and fear when they approach the police or lead to mistrust toward law enforcement and authority. This also can lead to the entire neighborhood producing hate toward law enforcement agencies and exacerbating tension between different demographic groups.

From the perspective of law enforcement agencies, they think implementing a new predictive policing platform is promising. Not only it is a valuable tool for crime prevention but also leads to more efficiency in detecting crime. However, conflict might arise when community members and civil rights organizations are concerned about an individual's privacy right being violated and the potential biases of predictive policing algorithms can exacerbate biases toward certain demographic groups. From the perspective of a community member, they think implementing a new predictive policing platform might violate their privacy rights. However, community members are aligning intending to have a safer place with the implementation of a predictive policing platform. Moreover, conflict might arise if community members perceive the predictive policing algorithm can wrongly target innocent individuals or disproportionately target certain demographic groups. From the perspective of civil rights organizations, they think the implementation of predictive policing platforms can exacerbate the inequalities in the criminal justice system. They also suggested a more transparent platform and reduced the potential privacy violation by bad actors. Civil rights organizations might have conflict with law enforcement agencies in issues like racial profiling and privacy violations.

1A. You are working on a model for consumer access to a financial service. Race is a significant feature in your model. Would you use the race variable in your model? If you are disallowed from using it, what would you do?

Using race as a feature in a financial services model can be ethically problematic and may also raise legal concerns, particularly regarding discrimination and fairness. Discrimination based on race is not only unethical but also illegal in many jurisdictions, including the United States under laws such as the Equal Credit Opportunity Act (ECOA) and the Fair Housing Act.

If race is a significant factor in the dataset, it's crucial to address potential biases and ensure fairness in the model. However, directly including race as a variable in the model is generally not advisable due to the reasons mentioned above.

Instead, here are some alternative approaches that can be considered:

1. **Proxy Variables**: Look for proxy variables that may be correlated with race but are not directly related to it. For example, zip code or neighborhood demographics might correlate with race but are not explicitly racial variables. However, caution must be exercised here to ensure that the chosen proxy variables do not reinforce existing biases.
2. **Fairness Constraints**: Implement fairness constraints or adjustments to the model to ensure that decisions are not unfairly biased against certain racial groups. Techniques such as demographic parity or equalized odds can be used to mitigate bias in predictive models.
3. **Feature Engineering**: Focus on other relevant features that are not directly related to race but may still capture important information about individuals' financial behavior and creditworthiness. This could include factors such as income, employment status, education level, credit history, and so on.
4. **Transparency and Accountability**: Regardless of the approach chosen, it's essential to maintain transparency about how the model works and how decisions are made. This includes documenting the data used, the features considered, and any steps taken to address potential biases. Additionally, establishing mechanisms for ongoing monitoring and evaluation can help ensure that the model remains fair and unbiased over time.

Ultimately, the goal should be to develop a model that is both predictive and fair, ensuring equitable access to financial services for all individuals while avoiding discrimination based on race or any other protected characteristic.

When faced with the challenge of using network traffic data that does not sufficiently inform credit risk, especially with poorer estimates for certain groups, there are several steps you can take to address the issue:

1. **Data Augmentation and Enrichment**: Consider augmenting the network traffic data with additional data sources that provide more insights into credit risk. This could include traditional financial data such as income statements, balance sheets, and cash flow statements, as well as alternative data sources like transaction histories, social media activity, or third-party credit scores. By enriching the dataset, you may gain a more comprehensive understanding of borrowers' creditworthiness.
2. **Feature Engineering**: Focus on identifying and extracting meaningful features from the network traffic data that may indirectly correlate with credit risk. For example, patterns in transaction frequency, volume, or timing could provide insights into the financial health and behavior of small businesses. Additionally, consider incorporating features related to the industry, geographical location, or size of the business, which can also influence credit risk.
3. **Modeling Techniques**: Explore advanced modeling techniques that are robust to noisy or incomplete data and can effectively handle imbalanced datasets. This may include ensemble methods, deep learning models, or machine learning algorithms specifically designed for handling sparse or unstructured data. Additionally, consider techniques for dealing with biased data, such as re-sampling methods or algorithmic fairness techniques.
4. **Fair Lending Practices**: Ensure that lending decisions are made in accordance with fair lending laws and regulations, such as the Equal Credit Opportunity Act (ECOA) in the United States. Avoid using features that directly or indirectly discriminate against certain groups, and implement fairness constraints or adjustments to mitigate any biases in the model's predictions.
5. **Validation and Monitoring**: Thoroughly validate the model's performance across different demographic groups to identify and address any disparities in predictive accuracy. Implement mechanisms for ongoing monitoring and evaluation to ensure that the model remains fair and equitable over time, and be prepared to update the model as new data becomes available or as business conditions change.
6. **Transparency and Accountability**: Maintain transparency about how lending decisions are made and communicate clearly with borrowers about the factors considered in the decision-making process. Provide avenues for recourse or

appeal if borrowers believe they have been unfairly denied credit, and establish governance structures to oversee the use of predictive models in lending.

1B: You are asked to use network traffic data to offer loans to small businesses. It turns out that the available data does not rigorously inform credit risk and estimates are relatively poorer for certain groups. What do you do?

By taking a holistic approach that combines data augmentation, feature engineering, advanced modeling techniques, and adherence to fair lending practices, you can improve the accuracy and fairness of lending decisions based on network traffic data for small businesses.

1C You are talking to a new potential client who wants you to use a public profile photos on the web to get age and race for “inclusion” in their recruiting. Would you consider this request? Justify your reasoning.

As an AI language model, I don't access external data or perform actions, but I can certainly provide guidance on this topic.

Using public profile photos to infer age and race for recruiting purposes raises significant ethical concerns and potential legal issues. Here are some reasons why this request should be carefully considered:

1. **Privacy Concerns**: Using publicly available photos without consent for purposes such as age and race determination can infringe upon individuals' privacy rights. Even though the photos are publicly accessible, individuals may not have intended for their images to be used in this manner.
2. **Bias and Discrimination**: Inferring age and race from photos can introduce bias and perpetuate discrimination in the recruiting process. Age and race should not be factors in hiring decisions, and using this method could lead to unfair treatment of certain groups.
3. **Accuracy Issues**: Attempting to determine age and race from photos can be highly inaccurate and unreliable. People may appear younger or older than their actual age in photos, and racial identification based on physical features can be subjective and prone to error.
4. **Legal Implications**: Using age and race as factors in recruiting decisions can potentially violate anti-discrimination laws, such as the Age Discrimination in Employment Act (ADEA) and Title VII of the Civil Rights Act of 1964. Employers should focus on qualifications and merit rather than demographic characteristics.
5. **Reputation Damage**: Adopting such practices can damage the company's reputation and brand image, leading to public backlash and negative publicity.

Trust and integrity are crucial in recruiting, and using questionable methods can erode stakeholder trust.

Instead of relying on superficial characteristics like age and race, it's advisable for companies to focus on merit-based criteria, skills, qualifications, and experience when making recruiting decisions. Employers should implement fair and unbiased hiring practices that prioritize diversity, equity, and inclusion. Additionally, utilizing anonymized application processes and blind recruitment techniques can help mitigate bias and promote a more inclusive workforce.