Final Project Report:  
For Use in CS6460, CS6750, and CS7637

Winston Tan, Ryan Graddy  
wtan87@gatech.edu

***Abstract—***Welcome to Joyner Document Format (JDF) v2.2! JDF is primarily intended to standardize page lengths while ensuring readability. Note that you are required to use JDF for all written assignments, but we will not perform explicit formatting checks. So, while improper formatting may be subject to penalties, you should not worry too much about whether your submission conforms to every minute detail; the most important elements are margins, font, font sizes, and line spacing. Just make a copy of one of the provided templates and replace its contents with your own, using the built-in paragraph styles.[[1]](#footnote-2) If you do so, you do not need to verify that the style was followed.

# Dataset selection

## Dataset

After looking through many datasets in UCI’s machine learning repository, we both decided to work with the “National Poll on Healthy Aging (NPHA). The link to the dataset is as follows:

* <https://archive.ics.uci.edu/dataset/936/national+poll+on+healthy+aging+(npha)>

## Dataset Domain

This dataset’s domain would be in Public Accommodation (healthcare access).

## Number of Observations

This dataset has 714 observations, which satisfies the project requirements of at least 500 observations.

## Number of Variables

There are 15 variables within the dataset: Number of Doctors Visited, Age, Physical Health, Mental Health, Dental Health, Employment, Stress Keeps Patient from Sleeping, Medication Keeps Patient from Sleeping, Pain Keeps Patient from Sleeping, Bathroom Needs Keeps Patient from Sleeping, Uknown Keeps Patient from Sleeping, Trouble Sleeping, Prescription Sleep Medication, Race, Gender

## Dependent/Outcome Variables

The variables selected as the dependent/outcome variables are Mental Health and Physical Health.

## Protected Class Variables

Of the 15 variables in the dataset, there are three legally recognized protected classes:

* Age
* Race
* Gender

## Legal Precedence/Law for Protected Classes

* Age: Age Discrimination in Employment Act of 1967
* Race: Civil Rights Act of 1964, 1991
* Gender (Sex): Civil Rights Act of 1964, 1991

# Exploring the dataset

## Protected Classes and Subgroups (Gender, Race, Age)

|  |  |
| --- | --- |
| **Gender** | **Count** |
| 1 (Male) | 321 |
| 2 (Female) | 393 |

|  |  |
| --- | --- |
| **Race** | **Count** |
| 1 (White, non-Hispanic) | 578 |
| 2 (Black, non-Hispanic) | 52 |
| 3 (Other, non-Hispanic) | 20 |
| 4 (Hispanic) | 44 |
| 5 (2+ Races, Non-Hispanic) | 20 |

|  |  |
| --- | --- |
| **Age** | **Count** |
| 1 (50-64) | 0 |
| 2 (65-80) | 714 |

## Discretize Subgroups

There is only one age group within the dataset, so it’s already discretized.

## Protected Class Selection

For this project, we have decided to focus on race and gender for the rest of the analysis. Even though age is a protected class, all the observations for it are within the same age group so we feel it’s best to no include it.

## Protected Classes Vs Dependent Variables (Tables)

### Gender vs Number of Doctors Visited

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Number of Doctors Visted | 1 | 2 | 3 |
| Gender |  |  |  |  |
| 1 (male) |  | 56 | 173 | 92 |
| 2 (female) |  | 75 | 199 | 119 |

### Gender vs Prescription Sleep Medication

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Prescription Sleep Medication | -1 (refused) | 1 (use regularly) | 2 (use occasionally) | 3 (do not use) |
| Gender |  |  |  |  |
| 1 (male) | 0 | 21 | 15 | 285 |
| 2 (female) | 3 | 17 | 19 | 354 |

### Race vs Number of Doctors Visited

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Number of Doctors Visted | 1 | 2 | 3 |
| Gender |  |  |  |  |
| 1 (White, non-Hispanic) |  | 103 | 296 | 179 |
| 2 (Black, non-Hispanic) |  | 8 | 32 | 12 |
| 3 (Other, non-Hispanic) |  | 7 | 6 | 7 |
| 4 (Hispanic) |  | 13 | 24 | 7 |
| 5 (2+ Races, Non-Hispanic) |  | 0 | 14 | 6 |

### Race vs Prescription Sleep Medication

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Prescription Sleep Medication | -1 (refused) | 1 (use regularly) | 2 (use occasionally) | 3 (do not use) |
| Gender |  |  |  |  |
| 1 (White, non-Hispanic) | 2 | 31 | 31 | 514 |
| 2 (Black, non-Hispanic) | 1 | 3 | 0 | 48 |
| 3 (Other, non-Hispanic) | 0 | 1 | 0 | 19 |
| 4 (Hispanic) | 0 | 2 | 2 | 40 |
| 5 (2+ Races, Non-Hispanic) | 0 | 1 | 1 | 18 |

## Protected Class vs Dependent Variables (Bar Graphs)

A graph of a number of doctors visited by a number of doctors

AI-generated content may be incorrect.

A graph showing a number of patients

AI-generated content may be incorrect.

A graph of a sleep medication

AI-generated content may be incorrect.A graph of a number of doctors visited by race

AI-generated content may be incorrect.

# Fairness Metric Selection and Mitigating Bias

## Privileged and Unprivileged Groups

The privileged subgroups for the protected classes in the dataset are:

* Male
* White

The unprivileged subgroups for the protected classes in the dataset are:

* Female
* Black
* Hispanic

The two dependent outcomes variables selected previously are:

* Mental health
* Physical health

## Fairness Metric Selection and Calculation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mental Health (Race) | Mental Health (Gender) | Physical Health (Race) | Physical Health (Gender) |
| Demographic Parity | 2.61 | 1.10 | 1.58 | 1.12 |
| Equal accuracy | 2.61 | 1.04 | 5.10 | 3.88 |

* 1. **Bias Mitigation Algorithm**

Transformed the data using Disparate Impact Remover.

## Bias Mitigation Algorithm

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mental Health (Race) | Mental Health (Gender) | Physical Health (Race) | Physical Health (Gender) |
| Demographic Parity | 2.61 | 1.10 | 1.14 | 1.14 |
| Equal accuracy | 2.61 | 1.04 | 5.10 | 1.39 |

# Mitigating Bias Using Classifier

Below are the results from training and testing using a classifier.

* 1. **Original Dataset Fairness Metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mental Health (Race) | Mental Health (Gender) | Physical Health (Race) | Physical Health (Gender) |
| Demographic Parity | 1.05 | 1 | 1.04 | 1.23 |
| Equal accuracy | 1.12 | 1.06 | 1.08 | 1.41 |

* 1. **Transformed Dataset Fairness Metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mental Health (Race) | Mental Health (Gender) | Physical Health (Race) | Physical Health (Gender) |
| Demographic Parity | 1 | 1 | 1 | 1 |
| Equal accuracy | 1 | 1 | 1 | 1 |

* 1. **Positive or Negative Change to Bias**

Calculated fairness metrics for each step are shown above.

|  |  |
| --- | --- |
|  | Change Compared to previous |
| Original Dataset | N/A |
| After Transforming Dataset | No change / very minimal change both negative and positive |
| After Training Classifier on Original Dataset | Positive change |
| After Training Classifier on Transformed Dataset | Positive change |

# Fairness Metric Selection and Mitigating Bias

## Group Members

* Ryan Graddy
* Winston Tan

## Graph of Two Fairness Metrics

## Explain Which Fairness Metric is Better

Neither fairness metric is better in this case. Both showed similar results in the original dataset, transformed dataset, and after applying a classifier. We chose to use Equal Accuracy and Demographic Parity for our fairness metrics using the What-If Tool provided. We modified the code to take NHPA dataset as an input and used features and labels discussed previously to train the classifier. Additionally, we used AI Fairness 360 toolkit to transform the dataset using Disparate Impact Remover. The results were consistent across all fairness metrics and did not show one being better than the other.

## Ryan Graddy Response to Questions

The DNN classifier trained using the transformed dataset seemed to work best. It provided the best fairness result, and I think this is a result of the dataset having characteristics that are somewhat predictable. No specific group was positively or negatively affected by these approaches. Rather they increased fairness and level the playing field among the groups. This could increase the chances of adding more opposite bias to the data. I believe the best way to mitigate bias is to conduct a larger poll for the dataset, but that can be costly and unnecessary if the bias/fairness mitigation techniques perform adequately.

## Winston Tan Response to Questions

# References

1. Joyner, D. A., Ashby, W., Irish, L., Lam, Y., Langston, J., Lupiani, I., Lustig, M., Pettoruto, P., Sheahen, D., Smiley, A., Bruckman, A., & Goel, A. (2016). Graders as Meta-Reviewers: Simultaneously Scaling and Improving Expert Evaluation for Large Online Classrooms. In *Proceedings of the Third Annual ACM Conference on Learning at Scale*. Edinburgh, Scotland.
2. Joyner, D. A. (2017). Scaling Expert Feedback: Two Case Studies. In *Proceedings of the Fourth Annual ACM Conference on Learning at Scale*. Cambridge, Massachusetts.
3. Joyner, D. A. (2018a). Intelligent Evaluation and Feedback in Support of a Credit-Bearing MOOC. In *Proceedings of the 19th International Conference on Artificial Intelligence in Education*. London, United Kingdom. Springer.
4. Joyner, D. A. (2018b). Toward CS1 at Scale: Building and Testing a MOOC-for-Credit Candidate. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale*. London, United Kingdom. ACM Press.
5. Newman, H. & Joyner, D. A. (2018). Sentiment Analysis of Student Evaluations of Teaching. In *Proceedings of the 19th International Conference on Artificial Intelligence in Education*. London, United Kingdom. Springer.

# Appendices

You may optionally move certain information to appendices at the end of your paper, after the reference list. If you have multiple appendices, you should create a section with a *Heading 1* of “Appendices.” Each appendix should begin with a descriptive *Heading 2;* appendices can thus be referenced in the body text using their heading number and description, e.g. “Appendix 5.1: Survey responses.” If you have only one appendix, you can label it with the word “Appendix” followed by a descriptive title, e.g., “Appendix: Survey responses.”

These appendices do not count against the page limit, but they should not contain any information *required* to answer the question in full. The body text should be sufficient to answer the question, and the appendices should be included only for you to reference or to give additional context. If you decide to move content to an appendix, be sure to summarize the content and note it in relevant place in the body text, e.g., “The raw data can be viewed in *Appendix 5.1: Survey responses.*”

1. [↑](#footnote-ref-2)