**Duke** **MIDS Capstone Project Proposal for Academic Year 2023/2024**

**Title: Fairness in ML for Financial Services**

2nd Order Solutions (2OS) is a leading credit advisory firm that operates at the intersection of deep credit expertise and advanced data science. We work with top banks and leading fintechs. Our data scientists build statistical models to predict consumer behavior, such as who to lend to, in what amount and at what terms – one might think this would be core capabilities for major lenders, but our expertise has led them to rely on us to solve the most challenging data science problems – machine learning models are key to driving this performance increase. One of the great advantages of machine learning over traditional regressions is its ability to capture non-linear effects. The downside of the method is they can be hard to interpret, and as such may lead to biased decision-making, which causes issues in fairness.

Fairness in machine learning is the process of detecting, understanding, and rectifying algorithmic bias in a machine learning model or system. Algorithmic bias includes gender, race and ethnicity, disability, etc.[[1]](#footnote-1) Since many of our models deal with who receives what credit product and in what capacity (e.g., credit cards, loans, mortgages, etc.), it is incredibly important to manage fairness. The financial services market is highly regulated and has certain fairness requirements for models used in lending. US law, through the Equal Credit Opportunity Act (ECOA),[[2]](#footnote-2) protects consumers by prohibiting unfair and discriminatory practices, which includes lending unfairly.

Given this high burden, 2OS is always looking for methodologies or approaches that improve what is being done currently. As such, we are interested in the comparison of available tools for machine learning fairness. While we list several options below, we encourage the students to research and find other potential tools.

|  |  |
| --- | --- |
| **Package Name** | **GitHub Repository** |
| AI Fairness 360 | <https://github.com/Trusted-AI/AIF360> |
| DALEX | <https://github.com/ModelOriented/DALEX> |
| Deon | <https://github.com/drivendataorg/deon> |
| Fairlearn | <https://github.com/fairlearn/fairlearn> |
| fairness-comparison | <https://github.com/algofairness/fairness-comparison> |
| fairness-in-ml | <https://github.com/equialgo/fairness-in-ml> |
| FairSight | <https://github.com/ayong8/FairSight> |
| Responsible AI toolbox | <https://github.com/microsoft/responsible-ai-toolbox> |
| Smclarify | <https://github.com/aws/amazon-sagemaker-clarify> |
| Themis-ML | <https://github.com/cosmicBboy/themis-ml> |

The project would include the following:

* Use publicly available datasets for comparison purposes
* Research, understand, and explain what fairness is, why do we care about fairness, how fairness is determined and measured, etc.
* Research what fairness Python packages are most popular, most promising, most used, etc.
* Choose five of the most promising Python packages and compare based on the models you built:
  + **Is your model fair?**
    - How does the package define being unfair?
    - If it isn’t, what’s unfair about it?
    - Do the packages agree that your model is fair?
  + **Useability**
    - How easy is it to use?
    - What methodologies are used?
    - While none of the datasets have missing values, how would the package deal with missing values?
  + **Interpretation**
    - What kinds of conclusions does this package lead to?
    - How transparent and easy to understand are the package interpretations?
  + **Mitigation**
    - Does the package include options to mitigate identified biases?
    - If it does, how much does the method impact model performance (e.g., predictive power)?
    - What kind of tradeoff do you observe between mitigating the identified biases and model performance?
  + **Generalizability**
    - How widely applicable are the techniques included in the package?
    - For what kinds of data would this package work?
  + **Adherence to regulations**
    - Are the output and conclusions sufficient to satisfy fairness regulations?
* Provide a package recommendation and back up your reasons
* (EXTRA) If you could build a fairness in ML package from scratch, what would it look like?
  + Create your own fairness repository
  + How does your version compare to existing repos?

2OS has proposed a few potential datasets below but is open to others if the team would like to use additional datasets. 2OS can help determine if the other dataset is like the dataset used for typical model builds. Given we are a consulting firm, our clients do not want us to share data, so publicly available datasets are necessary.

There are two publicly available datasets, as well as a synthetic dataset. For the project, the team will need to use the provided datasets and build out baseline machine learning models for each dataset. First data set is the default of credit card clients in Taiwan (<https://doi.org/10.24432/C55S3H>), and the second data set is bank marketing data from a Portuguese bank (<https://doi.org/10.24432/C5K306>). The synthetic data set simulates mortgage default probability for credit decisioning (<https://github.com/SelfExplainML/PiML-Toolbox/blob/main/examples/Example_SimuCredit.ipynb>).

2OS can support this project through bi-weekly meetings with the project team. In those meetings, 2OS can provide guidance on how to address the problem along with feedback on the results. The expected output is a write-up and presentation describing fairness in ML, the fairness packages that were investigated, results based on the models built, and recommendations. If the “extra” goal is met, then the expected output is a fairness GitHub repository with adequate documentation and usage examples.

Contact information:

Alice Liu – Senior Data Scientist – [alice.liu@2os.com](mailto:alice.liu@2os.com)

Joe Hammond – Senior Director, Data Science – [joe.hammond@2os.com](mailto:joe.hammond@2os.com)

Scott Barton – Managing Partner and Founder – [scott.barton@2os.com](mailto:scott.barton@2os.com)

1. Starting point references: <https://arxiv.org/pdf/2206.04101.pdf> and <https://arxiv.org/pdf/2210.02516.pdf> [↑](#footnote-ref-1)
2. See additional details from the Office of the Comptroller of the Currency (OCC) here: <https://www.occ.treas.gov/topics/consumers-and-communities/consumer-protection/fair-lending/index-fair-lending.html>. [↑](#footnote-ref-2)