

Open Cities AI Challenge: Responsible AI Prize

II. Model documentation and write-up

1. Who are you and what do you do professionally?

Thomas Kavanagh shares DrivenData's hometown of Denver, CO, but now lives in Brooklyn, NY. He spent the first part of his professional career designing interactive media systems for live performance, which led him to rediscover a love of programming. He graduated from Duke University with a major in mathematics. Thomas is currently a data scientist at Alkymi, whose platform uses computer vision and deep learning to automate time-costly data extraction. His favorite quarantine coping strategies include attending virtual dance parties, teaching yoga, and cold-calling his entire contact list.

Alex Weston is a Brooklyn-based data scientist. After studying Mathematics at Vassar College, he spent the next six years working at a non-profit arthouse cinema in the greater Philadelphia area. The recent explosion in deep learning research inspired him to move to New York and study data science. He is currently a teaching assistant at Metis, a 12-week immersive data science program, where he and Thomas met.

2. What motivated you to compete in this challenge?

Write down any thoughts and I'll merge mine in. We can even write down who thinks what!

THOMAS: I wanted a means to develop my data skills, especially in computer vision, while also contributing to a great cause. DrivenData's competitions are the perfect place to do that. The stated humanitarian objectives behind the development of this algorithm were clear, and I could imagine how our contribution would serve an immediate need.

ALEX: My top priorities as I begin my data science career are 1) to learn and apply cutting-edge data science technologies, and 2) to use those technologies for the benefit of humanity. This challenge was a slam dunk for both, especially since I was looking to learn more about computer vision.

THOMAS: Our collaboration was very serendipitous. One day, we happened to be taking the same subway back home, and discovered we had both found interest in the Open Cities AI challenge. We worked together mainly in the month of February.

3. If you could continue working on this topic in the future, how would you add to your approach? What would you recommend for others that will continue to work in this domain?

From a technical perspective, we have a few lessons-learned and suggested next steps that should apply to most deep learning implementations. We joined the competition quite late, with a month left, and many of these recommendations we did not have time to fully explore.

Poor Label Quality? Address with Teacher-Student Semi-Supervised Approach

We didn't have time to fully explore this idea, but we believe it's a good lead. The poor quality of some mask labels no doubt contributes considerably to lower model performance if they are used as-is in model training. A workaround is to use a model trained on Tier 1 to predict better labels for Tier 2, then retrain a model on the actual Tier 1 labels and the predicted Tier 2 labels.

Upon manual inspection of the Tier 2 labels, it appears as if some of the label masks could be manually adjusted to better align the mask labels and the image data. But with semi-supervised learning, this is not necessary. In this paradigm, poorly labels are not used at all. Instead, models trained on the highest quality training data are used to predict outcomes for data with poor or no labels entirely. Subsequent models are trained on the entirety of both sets.

This cycle repeats, with each subsequent "student" model predicting better labels for the badly labeled set, and therefore becoming the "teacher" for the next generation model.

For Transfer Learning, and Encoder / Decoder Models, Try All "Backbone" Options

For our model – developed in PyTorch – we implemented the FastFCN architecture, which is outlined in [this paper](#) and implemented in this [repository](#). This model, which has origins in the original U-Net architecture for pixel-wise semantic segmentation, allows for transfer learning from models like ResNet, or another other pre-trained encoder. We used the Resnet, but it's possible that another model would be more appropriate for this task.

Additionally, we discovered (perhaps too late) that it was important for the model training to gradually "unfreeze" the pre-trained layers, as decoder layers began to converge. Although our latest model did unfreeze layers in ResNet, we neither had time to train this model as long as we'd like nor experiment with learning rate adjustments as these layers entered the training phase.

In terms of ethical recommendations going forward, we would want to continue a conversation with the end users of our model, as well as the data collection team. The privacy implications of this model and the data used to train it depend greatly on the needs of the disaster relief team. In what circumstance do predictions need to be made, and what are the tangible consequences of a low performance model? Do the predicted mask labels need to be stored in a database, or can the modelers store only the insights (total surface area of a region covered by buildings, for example), thereby affording residents one more degree of privacy?

- 4. This is the first time one of our competitions featured a Responsible AI track. What did you think worked well? What could have been better? Any feedback and ideas for future iterations are very welcome!**

THOMAS: The open ended submission format is a great idea. Although we ended up choosing a fairly traditional route of a policy briefing / essay, I'm curious to see if other submissions offered ideas in more creative formats.

ALEX: The fact that this submission track even existed is important. The data science community (myself included) is very excited about new advances in the field, but there's less enthusiasm to step back and consider the consequences of the technology we create. We must always consider the broader impact of our actions, but it's especially important when those actions have the potential to affect thousands of people worldwide.

THOMAS: Links to additional Ethical AI resources were excellent starting points for our write-up. Some we had time to look at. Others I'm only now discovering, such as the uaviators.org code of conduct, from the [Harvard Humanitarian Initiative](#) linked to in the Responsible AI tab of this competition. There's so much work that's been done on this space, any submission would ideally be iterative.

THOMAS: As for what could be done better... I am very curious to know more about GFDRR's most pressing ethical considerations in their unique domain. Are DRM interventions facing certain specific regulatory, compliance, or community-response challenges that tend not to come up in other fields?

ALEX: I would have liked to know more about the data ethics concerns considered by OpenCities and GFDRR during the data collection process. Were local residents and/or governments consulted? What steps were taken to mitigate risk of abuse? That information would have made it easier to address concerns specific to the project in our submission.

THOMAS: Overall, I think that the Responsible AI track was well laid out. I only wish we had more time to flesh out our thoughts on the issue. I look forward to hearing commentary before publication. I'd hope to integrate some additional references drawn from the links that the competition provided.

THOMAS: Here's an idea for a future iteration. What if the Responsible AI track asked posed a unique scenario or model deployment, and asked for submissions in the form of a strategic response? Example below.

Example Scenario:

"You've been contacted by department officials of _SOME_COUNTRY_ to assess disaster risk in certain cities, and are being asked to coordinate a surveying effort under an official government banner. The citizens in the affected area have a history of unrest with the government, and there has been resistance to government intervention in the past. How would you approach the surveying effort that maximizes DRM utility while minimizing harm to communities in the surveyed areas?"

Alex: An interesting realization we had when thinking about our Responsible AI submission was that many of the ethical concerns faced by machine learning practitioners come into play *before* the modeling process, during data collection. There are still many thorny questions in that domain that we didn't have time to fully address in our submission. What constitutes informed consent? How can we balance the need to collect more features to power our models with the individual's right to privacy? I'd like to see future Responsible AI challenges try to address those questions.