

II. Model documentation and write-up

You can respond to these questions in an e-mail or as an attached file (any common document format is acceptable such as plain text, DOCX, etc.).

1. Who are you (mini-bio) and what do you do professionally?

Originally from South Africa, I've been lucky to enough over the last six years to live in Sweden, the US, and now Spain. I love working with data, especially when latitude and longitude are involved, and get a kick when I can apply this to my personal interests (mountains, oceans) and my profession (energy, climate). I consult for international donors like the World Bank, and more recently have started working with smaller startups and research groups using remote sensing to tackle interesting problems.

2. What motivated you to compete in this challenge?

I'd been looking for a good excuse to get to grips with neural networks, but in my day job traditional machine learning and data analysis are normally enough. So I jumped at the chance to work on something relevant to my work, with well-defined inputs and outputs. Cloud instances get expensive quickly, so I only trained and predicted three times. But I'd learned enough for my purposes, so decided I could share my thoughts for the Responsible AI part of the competition.

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?

There seems to have been lots of progress on deep learning techniques that are less compute-intensive, and also on techniques that go further with CPUs. I'd definitely like to figure out if I could train more models on more data using these approaches, because otherwise this type of challenge can be a bit unaffordable to hobbyists like me. I'd also suggest to data scientists in general to spend more time thinking about what problem they're actually trying to solve, and for communities like DrivenData to continue to foster a discussion of this.

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!

It seems like there weren't very many entries on the Responsible AI track? I'd assume this is because of the public leaderboard, so only those in the top three were strongly incentivised to submit. If there had somehow been a private leaderboard or final test

dataset released after all submissions... then maybe much more ML model submissions could have been joined by Responsible submissions. I also think the debrief needs to come quicker at the end of the competition! Who won, why how etc, before I forget how it relates to my failed attempts. Although I imagine more time was needed for the subjective judging, and probably other issues caused by the current situation...