MA124 Maths by Computer, Group Project Individual Submission

Dyson Dyson

Sophie and I worked on part B1, which was about researching techniques to improve neural networks and reduce overfitting. We both did the research; she wrote it up in markdown and on slides, and I used scikitlearn and matplotlib to create graphical examples of regularisation methods, and a contour plot to demonstrate the differences between L1 and L2. I also used git to coordinate and merge everyone's changes into a single notebook.

Originally I wanted to use a categorisation problem to show overfitting, but after a week or two of trying, I couldn't create a model that would actually overfit. All my examples either didn't fit the data at all, or they ignored the outliers and didn't overfit. Obviously that was unsuitable to demonstrate overfitting, so I changed my approach and decided to fit a polynomial to a set of data points instead.

In terms of improving group cohesion, I wish I'd been more involved with helping people. I've been programming for about 5 years and I did mention early on that I'd be happy to help if anyone needed it, but no-one ever asked. Everyone's code is fine, but I feel like I should've stepped up and offered to help more. But apart from that, we met often to coordinate who was doing what and we each kept the group up to date on our progress, so I think we worked well together overall.

My favourite part of the project was tinkering with various scikitlearn models. Most of them didn't make the cut and the quite frankly I think the graphs I ended up with are a bit boring, but they're the only thing I could get to work. It seems like a lot of the models in scikitlearn are designed for much larger, higher dimensional datasets, which are much more difficult to visualise. Finding the balance of a model that was complex enough to demonstrate overfitting while being simple enough to easily visualise, was surprisingly hard. But programming is one of my main hobbies, so delving into the scikitlearn documentation and experimenting with different models was really fun for me.

I've learned from this experience that I have no distinct passion for machine learning. I like it, but no more than I like programming in general, so I don't think I'll be focussing on machine learning in the future. I quite enjoyed the communication aspect of the project, so I'm considering doing some maths communication in the future, and I'm particularly looking forward to MA262 next year. My A Level Computer Science coursework was a Python program to visualise linear transformations, so I've been using software for maths communication for quite a while already.