Daily Thoughts and Experience Diary

AutoML (formalising and automating the stages in training a Machine Learning algorithm)

Developing machine learning models to solve problems is often described as more of an art than a science. This is another way of saying that it is very hard to guess how well a given technique will work on a problem. AI developers therefore spend a lot of their time searching over different machine learning architectures and hyper parameters trying to find ones that work well for their problem. This can lead to messy code and create a feeling like you are building on sand, as soon as you decide that one part of the model is working well changing another part may invalidate that choice. This has led to the field of AutoML which is a process of automatically searching for good AI architecture and hyper parameters. This approach is a good way to approach AI problems where instead of writing a single algorithm that you change you develop a python script that can generate different AI experiments with different architectures and parameters. Using this approach you can automatically run lots of different experiments and automatically search for the best approach. As you work on the project you are adding different alternative architectures to your python script so that you can automatically search over more approaches. In this way you never take a step back, every choice you add is at worst making your system try unhelpful approaches but the optimal design can be easily produced from your configuration.

The goal of this project is to take one or more of the AI algorithms described as part of the FastAI course (the leading course on Deep learning which includes a number of state of the art models for solving common machine learning problems). For each problem express the choices available in how the algorithm is implemented using a csv format (like a row in excel).

Make the resulting training algorithms capable of running on google colab and as a docker gpu container so that the machine learning algorithms can be easily trained on a cloud server.

# Monday 18th January

* Having a first look at the fast.ai course and trying to get it set up.
* Spent around an hour trying to set up a jupyter notebook using pip and couldn’t get it working.
* Tried using anaconda to install the libraries and it seemed to work but when I got my notebook opened I couldn’t figure out how to import the files from the course into it.
* Decided to switch over to trying to use google colab to open the chapters of the course.
* Started reading the introduction chapter and it seems really interesting.
* Only some of the built in python scripts in colab seem to be working, a lot of them show errors of either a library not being found or a variable not being assigned.
* Because of this I'm not really able to run most of them.
* The first script for training a model worked but it took nearly an hour to train it, so I’m thinking that maybe i'm using the wrong GPU.
* Finished reading through the introduction chapter and enjoyed reading it. Looking forward to the next chapters, however I'm not sure if I would be able to write an algorithm myself yet or even how to go about that.

# Tuesday 19th January

* Just reading over the second chapter of the course today.
* Didn’t get much accomplished because I was lazy, only got part of the way through chapter two.
* Was sort of sidetracked by my Web Technologies module because I was watching the canvas videos and trying to do the lab practical.

# Wednesday 20th January

* Going to try to get stuck into chapter two of the course today.
* Decided I might switch over to using gradient instead of google colab. Trying to set up the notebook now.
* Switched over to gradient and it seems a lot better, things seem to be running correctly as of now.
* One of the scripts won’t run and so all of the scripts after won’t run. I’ve been looking around on the forums etc. and can’t find a solution. I will have to just not run the scripts. It is quite frustrating.
* Looked around more for a solution, none found.
* Watching youtube videos about the fastai library, I want to start experimenting with code ASAP to get my head around using the library and some simple model creation.

# Thursday 21st January

* Today my plan is to go over the video lectures for the fast.ai course while having the the notebooks open to experiment with the code
* I think I need to learn a bit more from the course before I try to start making a lot of code because otherwise I will pretty much just be copy pasting code without understanding the meaning behind it.
* Currently watching the fast.ai video lectures, noting down some important code and its meaning in a document so that I can go over it later and experiment.
* Starting to feel better about the practical side now as I’m getting some exposure into how the code will look.
* Trying to create a simple image recognition program for detecting different types of rodent. Running into a few roadblocks with syntax at the minute.
* I'm getting stuck with file handling and scraping the images from google images, I will go over it tomorrow and try to get it working.

# Friday 22nd January

* Had my first meeting with John today. It went well and he mentioned to me to contact him if I’m experiencing any roadblocks such as not being able to get certain code to run.
* We also determined what area of ML each person in our group will be focusing on. I will be focusing on Computer Vision which I am happy about. I find it interesting; it's very easy to see it working which I think is very useful and visually appealing.
* We discussed what each person would base their how-to guide on and I will be doing mine on how to get gradient and fast.ai up and running. This will be fun as I haven’t actually got it up and running yet. Although that is probably a good thing as I can list the problems I experience and that should be useful to anyone else who is trying to get gradient set up.
* Currently reading up a bit on python because I haven’t used it since GCSE Computer Science.

# Monday 25th January

* Not much progress as I was catching up with videos and work for other modules.
* Plan for this week is to try to get gradient up and working properly and start on a bit of code to get my head around fast.ai.
* Got an email from John with a pdf attached which went over the importance of AutoML and how it is used in Enterprise which was an interesting read. Took some notes from it such as the goal of AutoML and some key areas that should be focused on: usability, stability etc.

# Tuesday 26th January

* Went through one of my group members’ test projects in a notebook and everything ran perfectly, with what seemed to be identical code to the fast.ai notebook. Starting to think that my problems originated from imports not working correctly or me not running them for some reason.
* Originally I thought that the problem was my Bing search API key, however, even after swapping my group member’s key out with my own, the notebook ran perfectly. I’m now possibly over this roadblock in the code, but the only way to know for sure is to try to set up my own application.
* Also today, going over the lesson 3 lecture on the fast.ai course, which by now is getting more in depth in the code. Learning about pytorch syntax and about tensors in pytorch.

# Wednesday 27th January

* Rewatching part of yesterday’s fast.ai lecture to remind myself. Learning about gradient descent now and pytorch code.
* Today I think I will try to create the image recognition app.
* I'm trying to set up my app but for some reason now the pip installs and the imports aren’t working. It's confusing because the exact same code was running yesterday.
* I think the problem was that I was trying to create a new notebook from google colab that wasn’t connected to the main fast.ai repository meaning I couldn’t import properly.
* Started a new notebook from gradient and the code is now running.
* Ran into a problem when downloading images where I created a path but then an error occurred because of a spelling mistake in a method name. This led to the if statement not being activated because the condition was “if not path.exists()” meaning that the downloading wouldn’t happen.
* Fix: I just had to delete the empty directory from the main notebook repository so that the code would run and it works now.
* Running into an error when trying to train my model.
* The notebook kernel keeps dying and I’m getting memory allocation errors.
* I think I will finish for tonight and try to get the errors sorted tomorrow.
* I think possibly the problem with memory could be due to the host GPU, I just noticed the host was a CPU so I will need to try to change the host to a GPU and see if that fixes it.

# Thursday 28th January

* I have restarted my notebook and switched the machine type over to a free GPU after it had automatically switched to a CPU. It seems that the training is now working, so I’m assuming the CPU host doesn’t have enough memory or just isn’t powerful enough compared to the GPU host.
* I was able to train my model and clean my dataset. I incorporated voila to make into an actual usable application within a jupyter notebook. It classifies an image of rodents between 3 types of rodent: Rat, hamster and capybara.

# Friday 29th January

Meeting with John

* Todays meeting:
  + Discussed what we should be doing currently to progress.
    - What I need to focus on is trying to create an object detection program that will detect objects within an image and put a box around them with a label of what the object is.
    - The program will use the YOLO algorithm and John suggested using YOLOv4.
    - I will need to look into the YOLO algorithm and create a YOLO style application in a Jupyter notebook.
  + We also discussed the overall goal of our projects
    - To be able to take a labelled dataset and input it to our program
    - It will then compare different algorithms and find the most efficient for the problem and write that code in a jupyter notebook and run it.
* Read some of this paper: “YOLO Nano: a Highly Compact You Only Look Once Convolutional Neural Network for Object Detection”. Just trying to read a bit about YOLO and also learn a bit of the academic vocabulary.

# Monday 1st February

* I was reminded today by email that I have to complete and hand in my how-to guide this Friday and I wasn’t aware of that, and so today I have been getting started on it by writing the basic introduction and researching other how-to guides that were linked on canvas.
* My how-to guide is called “A guide to set up fast.ai + Gradient for Machine Learning”
* I am currently researching fast.ai to figure out how I am going to explain it and describe its usefulness to the audience of the guide.

My objective with this how-to guide

* My goal for this guide is to help anybody who is trying to engage with Machine Learning by getting them started with the tools to begin their journey(fast.ai + Gradient).
* A lot of the facilities that are used in the field of Machine Learning development and Research are made for specialists. Therefore they are hard to use and hard to set up for non-specialists.
* During my start-up process, I encountered many bugs and errors with both fast.ai and also with Gradient. It was frustrating when I was trying to figure out how to fix them, but now that I have, I want to make it easier for those who are trying to accomplish the same thing in the future.

How do I plan to do this?

* I want to make this guide in a similar way to existing guides because they were clearly successful and helpful.
* One guide I had a look at was on the fast.ai website [here](https://course.fast.ai/start_gradient).
* I think that laying it out in clear steps would be effective at making the audience keep reading and be able to follow easily.
* However, I think I will try to make the steps as simple as I can, to increase accessibility further because more accessibility is generally better for more people to engage with content.
* Finally, based on the previously linked successful guide, I will need to include an attention-grabbing piece of text to engage the viewers and make them more inclined to read on and investigate my guide.

# Tuesday 2nd February

* Today I will again be focusing on writing up my how-to guide.
* I am now going through the instructions for creating a notebook.
* I plan to finish the project creation section and then go through the problems I faced during my journey and the solutions I found for them in order to help future users.
* ISSUE: Now when trying to set up a notebook it isn’t working so I’ll have to figure out how to fix it.
* FIX: I was trying to create a project rather than a Notebook.
  + This issue was an easy fix in terms of simplicity (I just had to create a Notebook instead) but it was an issue that didn’t show any errors, meaning it could be hard to figure out what the problem is.
  + Including small fixes like this will allow the viewers to get around the hurdles and get started faster than if they were trying to do this by themselves.
* Now I’m trying to figure out how I’m going to lay out this guide in terms of a website.
  + It needs to be clear and visually appealing to the viewer. (Otherwise they will just click away).
  + To do this I think I will have multiple named pages with clear purposes to allow for a concrete structure throughout.

# Wednesday 3rd February

* Downloaded one of the free templates from [Free responsive website templates | CSS, HTML5 web templates](http://www.html5webtemplates.co.uk/).
* I’m now trying to add all of my information from my how-to guide into the template.
* Firstly, I need to figure out the best layout for my material for any users.
* It is likely that my website will resemble a common website layout:
  + Home/../../Contact Info
  + This is because it is both used and works well on many other websites, so it is clearly a successful way to organise a website.
* It will probably consist of 3-4 tabs/pages.
  + I don’t think I want it to be a very large collection of pages as that may encourage viewers to quit if they think it will take too long or be too complicated.
* I'm obviously going to need a page for the tutorial itself- I’ll call it “Getting Started” or something similar.
  + Again, this is a standard name for such a webpage so it is likely going to be effective and also will be familiar to the viewer.
* Finished the “Home Tab”, it is just a basic overview/introduction of the tutorial including what both fast.ai and Gradient are.
  + I might need to add an attention grabbing piece of text as I mentioned in my plan.
* Started work on my “Getting Started” tab.
  + This will contain the tutorial content.

# Thursday 4th February

* Was at work all day today so my progress was impeded a bit.
* However, i’m going to try to finish up my how-to guide website tonight.
* Currently just finishing off the “Getting Started” tab by adding in the images and finishing off the text.
* I’m also noting down some ideas for information to include on a “My Goal” tab.
  + It will be a page about my reasons for creating this tutorial and what I hope to achieve with it (mainly to help students of the fast.ai course by allowing them to get started with the chapter notebooks and work alongside the lecturer).
* ISSUE: Now when I open the Notebook, none of the fast.ai material is in it for some reason.
* FIX: I just had to create a new Notebook and it seemed to work fine.
  + This is yet another bug that takes time to figure out how to solve by yourself. It took me at least 30 minutes and I was familiar with the process.
  + So hopefully, including this simple fix will help any readers of the guide if the error occurs for them.
* I added in the aforementioned attention grabbing statement at the top of the home page.
* I have finished the website now-
  + Layout is Home/Getting Started/My Goal/Contact Me
* Tomorrow I will be able to get back on track with trying to create a YOLO implementation for object detection. I’m eager to get a YOLO notebook up and running.