David Pettersson
david.pettersson@code.berlin
Semester 04
SE_15 Advanced Machine Learning

Reflection Essay

Goals

The initial spark for taking this module lies in a fascination with everything generated by computers – and more precisely in the field of Machine Learning. I wasn't aware, or not really interested in this field until the experiments from Google popped up, and the following ones from a myriad of artists (Mario Klingeman – http://quasimondo.com/, Gene Kogan – https://genekogan.com/ or even Memo Atkens – http://www.memo.tv/works/). Having a machine dream up images of its own sounds like science-fiction.

After having visited the exhibition from Refik Anadol at Kraftwerk in late 2019, I knew I had the next subject of research coming up. While at the same time I was starting a collaboration with the artist Diane Drubay where the topic of synthesized images could be a major topic.

My overall goals for this whole endeavour was to get a deeper insight into the world of Machine Learning - from the basic regression models to the more complex GANs and other models used to generate images "on their own". I achieved to understand most of the concepts behind them, and can now read a research paper on Machine Learning with at least a good understanding of what is happening (even if certain mathematical details still seem blurry at the time of this writing). I can, with some help, identify the problem at hand and determine which kind of model (classification, regression, clustering, dimensionality reduction) would fit the best.

I was able to implement and use a model to generate images from a given dataset, create a latent space walk between two (or more) random points in space and generate the corresponding images as a sequence. So the major part of this project has been built, and is being used practically.

The less successful part of this project would be the implementation from scratch of a GAN entirely adapted to the situation. By lack of time (and resources) I chose to use a commonly used model (StyleGAN2 by Nvidia) to be able to generate results quickly (having access to a free training unit for a limited amount of steps in the form of RunwayML). I used the network weights to further train the model on Google Colab.

Another weakness would be the proper in-depth analysis through the provided metrics as the training time is quite lengthy, and the platform used for the initial training doesn't expose the

hyperparameters - hence the optimisation and hyperparameter adjustment is left to their default state. The further analysis and optimisation could enhance and optimize our model to the specific dataset provided. This could be done either through reduction of the initial dataset resolution (from 1024x1024 to 256x256) for a reduced training time from scratch, or investment in a cloud GPU provider, but that becomes quite costly rapidly.

The final thought is not relative to the module, but is more of an abstract nature: in this precise case, we are not looking at creating realistic wave pictures, but rather are looking for those synthesized pictures to use as a ground material for the installation. The metrics are bad but is it really that important from an artistic point of view?

Topics

Most of the topics that I learned about this semester were new for me. From data selection and cleansing, to the training methods, the python frameworks (or even python as programming language for that matter), the actual training and the metrics used for gauging the success or failure of a model.

I think the biggest fact that blew my mind is the inference of features through convolutional filters in the case of a deep learning algorithm. The concept that an algorithm can learn to recognise our external world and categorise everything (or almost everything) through a “simple” set of rules and arrays just pushed me forward to learn more and more about that topic. The semester is over, and this is the self-reflection due for this module, but the project and my learning about it will not stop here.

The generation / synthesis of images is now less of a mystery, less of a magical process but rather the result of a detailed and logical process (quite complicated, as I’m just touching the surface it seems - but still accessible due to the amount of documentation available in these times.).

Resources

- Codecademy - Get Started with Machine Learning - online course
    - Good introduction in the world of Machine Learning
    - Ground course over various ML models, and how to curate the initial data such as:
        - Regression

        - Classification
        - Unsupervised learning
        - Perceptrons and neural nets

From this course I have learned the basics of python and the use of easily available machine learning algorithms in various domains. Helped by exercises that are attached in the Jupyter Notebook, I gained ground knowledge in the field of Machine Learning to be able to understand more complex subjects that would be the focus in the next phase of this project.

- Udacity - Deep Learning nanodegree (in progress) - online course

This course has not been finished yet and is still a work in progress but it is going over the same concepts that the Codecademy course went over, adding a layer of complexity and going into depth into the fields of neural networks, convolutional networks, classification, RNNs and GANs. The current status of this course is the learning of GANs and their structure / functionalities.

- Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems - Aurelien Gueron - book
    - The same course path as the Codecademy but in much more depth

This book has been used to gain deeper knowledge about the coursework on Codecademy. It is also a work in progress, but I use it more as a in-depth knowledge holder than following precisely its content.

- Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play - David Foster - Book
    - Introduction to the various generative techniques allowing machine to “create”

This book has started me in the field of generative deep learning, allowing me to get an initial insight and warming up to the final phase of the project. It covers the building and use of variational autoencoders, GANs, and other models that are not in the scope of this project (such as RNNs for musical generation).

As a whole, following these sources from the Codecademy course, to the Udacity ones via the help of the two books listed above, allowed me to understand what is happening behind the scene, make a distinctions between the various ML models and how they can be used in real life situations (from prediction, to classification). The more artistic part came with the GAN part, and this part being more subjective, it is sometimes still hard to grasp the use of metrics (and their analysis) in this particular case. How do you link the metrics with the subjective view of generated images?

Level

The level expectancy for this project in this state is a Level 1.

From the basics of ML to the more intricate details of a GAN in the form of a StyleGAN2. It is true that because of performance and training issues, I have not been able to use various flavors of GANs and this is a weakness. In contrast, I have gained a deep insight into the creation of an appropriate dataset, into the training process and the shortcomings of my own computer hardware (and finding solutions to be able to still have a positive outcome to this project.) I have created a non-traditional (traditional as in creating the transformation pipeline from the initial dataset to the final result in one go) pipeline which can be used and reused for this specific goal outlined above.

I also have learned about the specifics of deep learning neural networks, and how they create images based on a dataset, with all the steps in between (from inference, to prediction). I also realise the importance of having a dataset that is as varied as possible (ideally a different image for each sample).

This model, learning the features of the incoming dataset, is one of the best I've seen for this task. In terms of learning, if I had to change the course of this semester, I would have done a deep dive into all the GANs, and would've done a comparison between all the different models - or even implementing a custom network. This out-of-the box solution which I have applied, limits my understanding in terms of optimisation, but still can be used as a springboard for further experimentation.