Reflections on Machine Learning Module.

Module Code: ML\_PCOM73

Written by: Ganapavarapu Lakshmi Narayana Bharadwaj

Student Code: GANAP01509

E-Portfolio URL:

**Reflections:**

The Concept of Machine Learning (ML) has been around since the 1950s. Its usage has become immense since its inception, in present world. We now have a significant number of powerful and insightful algorithms, which, we humans can make use of and generate models with immaculate performance and possibly near perfect results. We are not however there, but as per recent social events, OpenAI (2022) has been a fascinating example of how far Artificial Intelligence (AI) and Machine Learning have come.

This module from University of Essex Online (ML\_PCOM7E) has extensively covered some really good ML algorithms using which the students were asked to tackle certain problems and offer ethical, analytical, and sustaining solutions. The beginning of the module was with the introduction of Machine Learning, covering its past, present, and future, and the professional and ethical challenges posed by it.

There were a few interesting lecture casts along the road, which helped us understand the concepts and the algorithms better. The Correlation and Regression (2022) and Clustering (2022) were a couple of interesting ones. At this point we had to submit an assignment that the New York AirBnB data ([access here](https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data)), from Kaggle (2019). This is where students learned to apply ML techniques like regression analysis, clustering analysis, and many more algorithms.

As we progressed further, we had dived into even more powerful ML algorithms that would handle complex form of data like images. These concepts were termed as neural networks and there’s different types of neural networks. All this time, we have also been performing small data related activities, which strengthened our application of our concepts. More information on this is available on the e-portfolio.

The first major concept after the regression and clustering analysis was an Artificial Neural Network (ANN). To put this briefly, ANN was a concept that closely was related to how a human brain work. The lecture on ANN by Qazi (2022), covers the basics of this concept starting from how a single brain cell, called a neuron, connects to multiple of them in order to make a decision, to how powerful it actually is, in the present day. Along with Dr. Qazi’s lecture, another lecture cast, Artificial Neural Network (ANN) (2022), has also extensively explained how the neurons in the brain work and based off of it, how can we design a neural network, that is capable of accomplishing complex and time-consuming tasks within minutes. After learning about the neural network, we had then looked at how to improve the model’s performance with something called as back propagation.

We then took this learning further by comprehending another powerful concept called Convolutional Neural Networks (CNN). CNNs are basically high level ANNs with very densely packed layers, which are also, huge in number. CNNs are what were used to get the final task in the module done. The students were instructed to perform the analysis on CIFAR10 data. CIFAR10 is a unique dataset ([more info here](https://www.kaggle.com/c/cifar-10)) that consists of 60,000 images of 10 different categories. The principal task to be performed on this dataset is to recognize each image and categorize it into its respective class.

We had to perform ML algorithms like the ANNs and CNNs on it and observe how the model was doing and also talk about how we can improve the model’s performance. The results were later composed into a presentation, which each team had to present, as a part of the submission.

During the second half of the module, we still had some e-portfolio related tasks that were to be completed, like, hyper-parameter tuning, choosing the right activation function for the neural net, what weights to update in order to get the desired output and many more.

On the flip side to all of this, there were underlying challenges to this powerful concept, that had to be ethically and professionally, taken care of. For instance, irrespective of where, how, how much and why the data was collected, the threat of privacy breach, misuse, and many other ethical issues prevail, indirectly hindering human rights. Rodrigues (2020) points out in her article that the debate regarding the ethical and legal human rights issues, how are they being tackled and what principles and protocols need to be followed to ensure good practice of the concept of AI. Similarly, Naik et al. (2022) highlighted the importance of ethical issues in the healthcare domain and explains the vulnerability and the consequences that comes with the misuse of a patient’s data.

If this module were to be compared with Rolfe et al.’s (2001) approach of what concepts were covered in the module, what could be done with those concepts, like, analytically, ethically, and professionally, and what benefits could be reaped out of it, the module would unfold in the following manner:

* The ‘what’ part would be the concepts like regression analysis, clustering analysis and all the neural network concepts.
* The ‘so what’ part would be all the data analysis techniques performed on various datasets, which were a part of the module, to make the data make sense, and applying various ML techniques on the data to get insights.
* The ‘now what’ part would be how can we improve the ML model and make better decisions by keeping the ethical and professional challenges in mind.

Brown (2015) has given a short and crisp detail about Rolfe et al.’s approach on her blog where she explains each part concisely.

As a whole, this module has certainly helped us gain very fascinating and powerful ML concepts that can be applied in almost our daily lives to get various tasks done. Working alongside a teammate in a group has also made me learn and improve on things like effective communication, be it about the challenges that occurred or explanation of concepts, division and sharing of work, writing effective reports, and many more. This kind of teamwork has made a positive impact on me, and it will definitely be helpful in the future to accomplish any team-oriented tasks. I am extremely grateful to have taken up this module, which has immensely helped me enrich my data science skills and has raised my hunger for knowledge in the world of data science.

**References:**

OpenAI (2022) *CHATGPT: Optimizing language models for dialogue*, *OpenAI*. OpenAI. Available at: https://openai.com/blog/chatgpt/ (Accessed: December 12, 2022).

University of Essex Online (2022) “Correlation and Regression,” November.

University of Essex Online (2022) “Clustering,” November.

Dgomonov (2019) *New York City airbnb open data*, *Kaggle*. Available at: https://www.kaggle.com/datasets/dgomonov/new-york-city-airbnb-open-data (Accessed: December 12, 2022).

Qazi, N. (2022) “Emerging Research in ANN,” 15 November.

University of Essex Online (2022) “Artificial Neural Network (ANN),” November.

Cukierski, W. (2013) *CIFAR-10 - object recognition in images*, *Kaggle*. Available at: https://www.kaggle.com/c/cifar-10 (Accessed: December 12, 2022).

Rolfe, G., Freshwater, D. and Jasper, M. (2001) *Critical reflection for nursing and the helping professions: A user's guide*. Basingstoke: Palgrave.

Brown, N. (2015) “Reflective model according to Rolfe et al.,” *Dr Nicole Brown Social Research & Practice and Education Ltd.* Available at: https://www.nicole-brown.co.uk/reflective-model-according-to-rolfe/ (Accessed: December 12, 2022).

Rodrigues, R. (2020) “Legal and human rights issues of AI: Gaps, challenges and vulnerabilities,” *Journal of Responsible Technology*, 4, p. 100005. Available at: https://doi.org/10.1016/j.jrt.2020.100005.

Naik, N. *et al.* (2022) “Legal and ethical consideration in artificial intelligence in Healthcare: Who takes responsibility?,” *Frontiers in Surgery*, 9. Available at: https://doi.org/10.3389/fsurg.2022.862322.