Proposal for QF Community Microgrant:

Algebraic nature of Machine Learning

The use of algebraic maps to algebraically characterize the behavior of machine learning models.

1 Motivation

Because of the recent advancement in Generative Models and in particular Large Language Models, it has come to the discussion that in order to reduce biases on the representations constructed by the Machine Learning models and more importantly place well-defined constrains on the capabilities of Artificial Intelligence systems, understanding what the machine learning models are doing is critical.

What is a machine learning model doing? Well, looking at the simple case of a supervised machine learning model, it appears to be implementing some sort of clustering or grouping of similar objects together and segregating clusters of dissimilar objects. A generative model, picks samples from the representative space. We know that. So one might ask, why do we still say that it is not well understood what these models do?

The clustering behavior of a supervised machine learning model, is a Topological understanding of the behavior; in that any change that happens in the behavior comes from moving around in the neighborhood of a point. Topology is about locality. However, most decisions made in human society, most things taught at school or most things achieved at a job are algebraic in nature. They are defined by generators and relations, which makes them specific. So, a machine learning network might give me lots of red pixels on an image if I say: give me a picture of a cat wearing a red hat, and give lots of white pixels on the image if I subsequently ask of it to give me a picture of a cat wearing a white hat. If I ask it how it achieved this understanding of my request, it would only be able to give me the mapping of the words red and white into the possible space colors of each pixel in an image. That is a topological answer. The state of the art machine learning models have difficulty in stating that the concept of color was used to make the decision. However, in the setting of human society, if a machine learning network were to distinguish between 2 candidates participating in a social event; like asking the bank for a loan, a distinction based on similarity is not satisfactory. A topological answer seems unsatisfactory.

A major part of human learning is repetition (experience) which is the tool we use, to form neighborhoods around concepts and ideas. Each time we come across an object or an experience, we tend to store (map) it near objects and experiences we believe to be similar. Lessons learned in life: like how long should you wait for something to work out before moving on, and virtually any other thing that classifies as ”subjective” in society, is learned topologically. But usually after some experience about certain objects and events, we tend to form algebraic representations of them. So it is not the case that we are unfamiliar with topological learning. But why are algebraic answers so important?

Let's say, I want to use tools provided by Khan Academy to teach a group of students. I can use machine learning models to customize the resources to their learning needs. Presumably, it affects a small group of students. If the model does not produce satisfactory suggestions, it can be discarded with ease. In this case, a resource that can be easily customized to a great extent affects a small group of people, and can be removed easily. But in society, machine learning models have begun affecting several million (or over a billion in a few cases) people, give little or no control to the users and; backed up by large corporations and markets ,can not be easily removed. So, to control the possibly inaccurate behavior or behavior misrepresentative of desired human values, laws are needed to govern the deployment of machine learning resources. Laws are by nature algebraic. They need to be specific and well defined by generators and relations. Laws also govern a large group of people; millions of citizens of a country and in some cases larger regions. The application of a law to various cases is not done based on variation by similarity (i.e. topologically) but rather by instantiation and specification by further laws; by adding generators and relations (i.e. algebraically).

The discussion of algebraic and topological learning suggests for the use of ideas from algebraic topology and algebraic geometry in the context of machine learning, with the aim of characterizing the topological behavior algebraically; in particular the algebraic understanding of constraints and decision making. This proposal follows in the footsteps of Geometric Deep Learning, Quantum Machine Learning, Dimensionality Reduction and Applied Algebraic Topology programs. Objects fundamental to algebraic topology and algebraic geometry like Vector Bundles will serve as inspiration.

2 Starting Point- Injecting operations between gradient descent steps

First we start with the simple idea: Apply gradient descent stepping only along one of the components of the input vector; the component along which the loss is maximum is a good choice. This is done so that various operations can be applied to each of the components indepdendently. The operations could be applied in between gradient descent steps; to change the representation prepared after each step, or in between the layers- equivalently as a layer. The former is expected to be simpler and more transparent. Operations could include projections and rotations which amplify or diminish the component of the representation vector along some direction vectors.

The commutation properties of such operations could also be used to understand the geometry of the loss space. A simple example could be to reflect the vector about an n-1 dimensional plane; i.e. reflect one of the components of the vector after one of the gradient descent steps, and then doing the reflection again (equivalently undoing it) after the next gradient descent step. Such operations when done for a few steps could be very difficult to track, so systematic strategies should be developed to apply the same set of operations over large number of steps. Properties like associativity of various operation applications and the composition behavior of various operations chained together could also be studied.

For simplicity, the project would be started in tensorflow in jupyter notebooks and be done on simple datasets like MNIST digits recognition dataset for images.

3 Homotopy of the paths of gradient descent

Another idea that could help understand the features of the loss space is to look at various gradient descent trajectories (depending on the use of various meta parameters) and whether transformations between the various trajectories could help understand which variety of intermediate vectors (states) attained during the training converge to similar vectors at the end of the training.

4 Frame changing machine learning.

Following works on ideas in the field of using Gauge Transformations in Machine Learning and further the idea from section § 2, the operations applied to the representation vectors could be implemented in strategies that do operations on the basis of the vectors. For example, applying an orthonormal basis change transformation would effectively rotate the vector. Operations involving various components of the representation vector; like exterior algebra of doing cross products could also be done systematically.

5 Constructing relations from a vector to a representative copy

Whenever humans (or at least I) learn something new, we always associated it to concepts and objects that look similar to the new object or concept we have just encountered. Through a few encounters with the concept, we form a chain of ”experiences”, with each instance of the object mapping to other concepts we already know. This behavior of learning by repetition or imitative learning is also displayed by modern Machine Learning models. But after a few experiences, it seems to me that we humans form an algebraic entity that can continue to exist as an independent concept without a lot of references to other concepts. The idea is to do the same for representation vectors. Let us discuss the two parts to this idea.

The first part includes constructing a chain of objects and maps between an instance of an experience to the ideal representation. These maps could either: 1. generalize a concept, 2. make a concept specific, or stay on the same level of abstraction and, 3. use the comparison of equivalence or similarity. For example, if a discussion is about the road safety of a car, it may be implied that the situation applies to any car. In another example, if a discussion is about the legal regulations for ones company, it could be about how the legal regulations of a country could be adapted to a particular business. It could also be interesting to explore ideas of chain complexes, derived categories and homology in this context.

The second part includes constructing a system for injecting standard objects into a machine learning model. In games that include discovering a map of some sort, checkpoints are often employed. In the same manner, if I am training a model to recognize digits, I would like to take a representation vector of the digit 8 from the MNIST dataset and use that as a representative of the digit 8 (modulo some operations like dimension change, etc.) in my model, until a decision is made to replace the representative, with a better one. The easiest way to do this is to use this vector that I would like to use as one of the axes of my space (or change the coordinate system accordingly). The effectiveness of this approach is questionable. But this approach would allow reusability of resources between models.

6 Vector Bundles

Vectors in a latent space can be considered as equivalent to points in the space. So, it could be said that there is a 1-dimensional vector space at each point of the space. This leads to a question about the utility of having higher dimensional vector spaces at each point; that vary continuously over a space- a vector bundle. The simplest implementation of this could be to break a vector into 2; the simplest of which is to have two components whose direct sum gives the vector. Operations could be done with these pieces of a vector which could then be combined to form a vector for the next step. Operations on a vector could also be done based on another vector drawn from the space. For example, let us assume that by some strategy it can be identified that the first component (or some components) of the vector roughly represents the color of it. A first sample is drawn, then two samples are drawn from a neighborhood that is a ”large” distance from the first sample along the color component, but such that the two latter samples are within a ”small” distance from each other (in the same neighborhood).A map representing the rotation between the two nearby vectors could be applied; repeatedly, amplified in magnitude or otherwise, to the first sample drawn.

7 Training by sampling- a good mechanism

As hinted to in the previous section §[6], it seems like it could also be a good idea to use the operations on a generative model that has already been trained. Given that vectors can be sampled from a latent space, this approach would then produce a new space based on the samples drawn from the previous space. For example, given a model, another model where biases have been reduced or certain aspects have been exaggerated could be produced.

8 Symmetry breaking or deviation from a symmetry, in the context of combining features

Another possible idea that could be investigated is the operation of reordering or rearranging the pixels during the machine learning of an image data. The reordering could be done in an invariant way or an equivariant way with respect to coarse-graining (as implemented by convolution filters/masks).

This leads to another important idea that can be studied- the breaking of symmetries. An implementation of this idea could be for example, in a model that combines two input images and based on the features of those images, generates an output image. For example, given an image of a cat and another image of a person, an image of a person in a reasonable setting with a cat would be a good output. The training of such a model could be done for example with images with a cat and a person, or generally, with images of any two objects as the input and images containing both the objects as the known output. This could serve as a starting point for the study of what two pieces of data can be combined in a structured and fairly defined manner. This idea is almost reminiscent of Galois Group in Algebra; here studying the algebraic features associated with the isomorphisms between features of two images. Each time during the process of combining smaller features to form larger features; like in an image, a symmetry of the ordering of the smaller features could be said to have been broken.

Based on the idea of constructing standard representatives in section §[5], the breaking of symmetry could also be implemented as deviations from the standard.

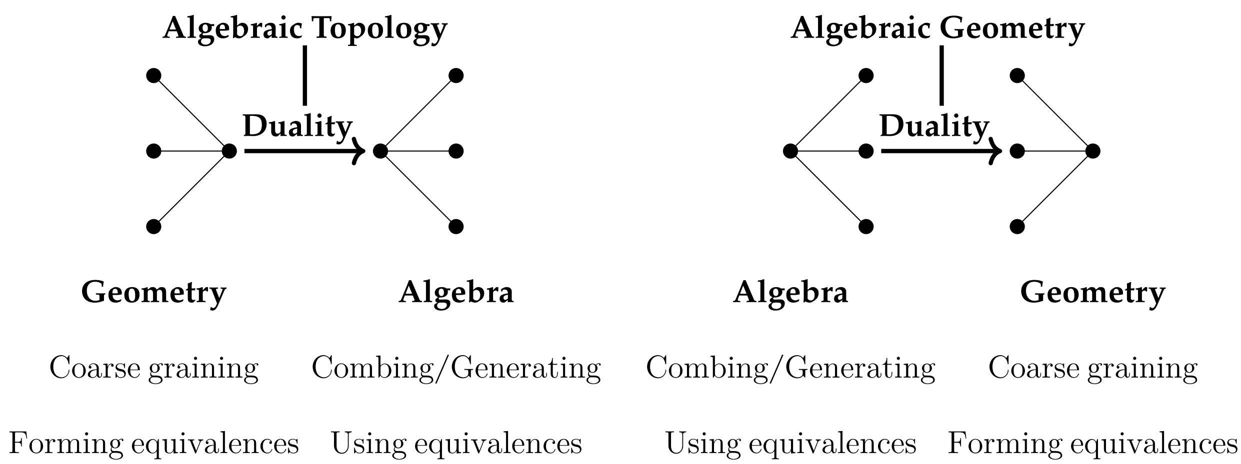
9 Plan and Logistics

The main starting point of the project is to use algebraic operations in the machine learning process. Some ideas have been suggested in this proposal but the implementations of these ideas are still vague. Without a systematic approach, most things done might yield nonsensical results. So, based on the study of papers in Geometric Deep Learning, Quantum Machine Learning, Dimensionality Reduction and themes of similar endeavor, implementations of various ideas could be tried. Ideas from dimensionality reduction techniques would be interesting in understanding the implementation of projection maps. Rotation maps can be studied from Gauge Transformation techniques. Symmetry and Invariance from Quantum Machine Learning and Geometric Deep Learning would serve as initial guidelines for implementing the algebraic operations. This would define the progress of the project. The idea of studying, using and applying symmetry transformations is likely to be a major component of the project.

10 Duality between Geometry and Algebra; coarse graining and combining

The main theme of the project is the algebraic characterization of the machine learning process. This relates to a perspective on the duality between algebra and geometry. Geometry arises in the world because of us trying to understand more than we can. We create equivalences in the coarse-graining process. We use algebra to understand more than we can, also by generating equivalences. So geometry is also about combining things and algebra is also about combining things. This is the duality between algebra and geometry. It's the same thing. Or at least dual (opposite directions of the same structure), with the same fundamental defining structure: Generators and Equivalences (Relations).

Generative models seem to be able to sample an object similar to some fixed object by drawing a sample from its neighborhood. This is similar to the coarse-graining process. Classifier models seem to generate a vector by combining the input features in some sense.



11 Personal plan

For a grant of \$8000-\$10000; starting in August or September, I would like to work on the project for 12 months; 40-60 hours a month (about 500-700 hours in total). After working on this project, I would feel more confident to apply for other opportunities in Mathematics and Artificial Intelligence in the future.

Kaushal Timilsina