Visualization of Loss Landscapes in Neural Nets to identify minima of loss functions

CS 529: Visual Data Science

Varunya Yanamadala

([vyanam3@uic.edu](mailto:vyanam3@uic.edu))

Department of Computer Science

University of Illinois at Chicago

**Abstract:**

The efficiency of the neural network depends on its ability to employ the best loss function, which are highly complex. This has been done theoretically until a paper was released to study through visualizations the loss functions of a network. The goal of the proposed work is to discover new methodologies to incorporate better understanding of these functions so that the network could train and generalize better. The solution obtained from the visualization would help the users of the system to decide on the hyperparameters, which otherwise would require theoretical analysis often not leading to better decisions.

Client: Dr. Sathya Narayanan Ravi, Assistant Professor, Department of Computer Science, University of Illinois at Chicago

Participant: Varunya Yanamadala, Graduate Student, Department of Computer Science, University of Illinois at Chicago

**Introduction:**

Neural networks are models simulated from human cell neuron. The mathematical models represent the synaptic energies that flow from one cell to another and are designed in a way to perform complex computations. These computations involve non-linear functions and the loss function is used to optimize the parameters that are in the neural network. As the parameters increase it becomes highly challenging to minimize a high dimensional non-convex function. Visualization in this scenario would help the users understand how a highly non-convex function could be minimized and consequently, how does it help in generalization.

**Aim of the project:**

The aim of the project is to understand the geometry of the loss function landscape in terms of sharpness or flatness. Since any change in the parameters in a network architecture turns the convex nature of the visualization to chaotic behavior, several observations could be made from the visualization as an effect of change in the parameters. Since, the loss functions are high dimensional, representing them over a 1D or 2D space is highly challenging. Representations could be made in parameter space by selecting two sets of parameters but cannot be used to perceive the non-convexities which are important to identify generalization in network architectures. So, a novel approach is employed to select the axis on which the loss values are plotted, and implementation is to be made to identify the parameters of the model at a particular local minimum.

**Significance:**

The significance of this project completely lies in the novel approach of visualizing the loss function values obtained from a network architecture to convey the users the optimal choice of parameters for the dataset considered for training. It will further help the network designers or architects to eliminate choices that would not lead to expected results. The approach corresponding to this project could be related to the following research topics:

* Representation of data and knowledge which comes under VAST
* Consideration of multivariate data, InfoVis
* Novel Interaction Techniques, InfoVis and VAST
* Novel encodings for algorithms, Methods at scale, LDAV
* For employing Geometry based techniques and developing scalable system, SciVis

**Requirement Analysis:**

1. Humans:

* Machine Learning Researchers
* Machine Learning Engineers
* Machine Learning Modelers
* Architects or Designers of Neural Networks
* The usage frequency ranges from at least a couple of times a day and should be up until all observations can be studied.
* Access is available 24x7
* User could use anywhere.

1. Tasks:

High Level:

Select a particular algorithm and different parameters of that algorithm are displayed. A visualization is displayed to understand the local minima values that can be attained at

different parameter values.

Questions Prioritized:

* Can by any means a user compare one or more algorithms of interest?
* Can a representation of one algorithm be overlaid on the other for comparison?
* Could we have an indicator to show the global minimum for the model?
* Can the stability of predictions in parameter space be studied?
* Can the user get to any conclusion the complexity of the model selected?
* Could few parameters be changed to see its effect on the overall loss landscape?
* Can we compare the parameter set values for the similar loss function value in a network?

1. Data:

* Network models information:

1. This multivariate data contains information related to the network architecture parameters - number of layers, width of each layer, edges – connectivity, biases, skip connections, etc.,

* Loss function values:

1. This data contains the loss function values for parameter set of a model.
2. Flow:

The flow of the system is as follows:

* The user lands on the main page where he could see the types of algorithms.
* The user could select an algorithm, the visualization depicting the loss function values is shown.
* The user could hover over a point of interest to see the values.
* Since the visualization is a representation of non-convex loss function, the user could select the wireframe option to observe the descent in any local space.
* A representation of how the architecture looks is also displayed to the right to help the architectures understand the neural network model.

1. Non-functional Requirements:

Performance:

The system should have faster response times

Scalability:

The developed system should have great adjustability from laptops, desktops to larger displays.

Accessibility:

Cross-platform and preferably should be able to run using an executable file. This is not hosted on any server and is not viewed on browser.

Reliability:

There is no server downtime, as it’s a standalone application. It should be accessible at all times for the users.

Security:

No need of data security.

Privacy:

No restriction on data accessibility.

User Experience:

No training should be required, should be understandable by the user how to operate through and the user should be able to navigate through the flows easily.

1. Probes:

* Rationale for visualization:

Visualization wasn’t used in this field of identifying the loss function values in neural networks. It is a novel approach to help the researchers to identifying the minimizers and global minimum for an algorithm of interest.

* Alternative tools:

There is no tool present which is publicly available. A base version of the selected paper was released by Telesens. We intend to improvise the visualization so that the above objectives are achieved.

* Other possibilities:

Could be integrated with data residing on a server increasing the scope for integrating more up-to-date algorithms, making the visualization system dynamic.

**Workplan of the project:**

*Week1 – Week2:* Understand the data background on how its generated and check for the ways of converting the data as required for visualization. Prepare functional requirements document.

*Week3 – Week4:* Produce prototypes using Five Design Sheet methodology and take approval from the client. Develop alpha phase of the visualization and present to the client.

*Week5 – Week7:* Corrections (if any) of the existing system and continue with other requirements of the project; be ready for beta release.

*Week8:* Get a final approval of the client. Prepare the final abstract, and for presentation and demonstration of the project.

**Challenges:**

* The data is not in ready to be used for visualization. Data has to be transformed and the data corresponding to the relevant attributes has to be considered.
* Requires basic understanding of the hdf5 file format (format used for high dimensional scientific data).
* Dataset has values differing with values differing in the order of 10-3. Care should be taken to visualize these values correctly in an understandable way to the users.

**Conclusion:**

This project is the state-of-the-art in identifying the loss function minima values (in the field of machine learning), an area which is studied theoretically. Visualization is chosen relevant for this to ease the procedure of constructing an efficient neural network.

**References:**

<https://arxiv.org/abs/1712.09913>

<http://www.telesens.co/2019/01/16/neural-network-loss-visualization/>

<http://ieeevis.org/year/2019/info/call-participation/vast-paper-types>

<http://ieeevis.org/year/2019/info/call-participation/infovis-paper-types>

<http://ieeevis.org/year/2019/info/call-participation/scivis-paper-types>

<https://ldav.org/call-for-papers.html>

**Curriculum Vitae:**

Varunya Yanamadala

Masters Student (Aug 2019 – May 2021)

Department of Computer Science, University of Illinois at Chicago.

Skills:

HTML, CSS, Knockout JS, SQL, SSRS, .NET, Spark

Experience:

* Worked as a developer in ASP.NET applications
* Backend developer (SQL Server)
* Advanced Analytics projects involving python, spark and hive.