Research Paper # 1: Machine Learning with Neural Networks

Due on May 1, 11:59PM.

Under Course Materials, click on Research Paper 1 and submit it there.

You can submit it early. It is now available for submission.

One of the most popular subfields of programming recently is machine learning(or statistical learning) using Neural Networks.

Machine learning is a set of tools and algorithms that can process and understand data. Linear regression that we learn from Chapter 10 and 12 is a machine learning algorithm. It finds the regression line that best describes the data and uses it to make new predictions.

Although, important machine learning algorithms were developed in the 1990’s, recent availability of data(e.g. Google images) and computing power make possible some recent breakthroughs in machine learning, particularly, image classification. Computers now can identify images better than humans. Neural Networks is the algorithm that is widely used for image classification.

I created six videos that introduces the Neural Networks. You are not required to watch all six videos. These videos can be found on my Youtube channel, see below. However, you will need to watch at least three to sufficiently answer some of the questions below. Links to them are posted in the Research Paper #1 Folder on Course Materials.

Watch at least the following three lectures. These three lectures contain some of the answers to the questions below. The videos contain all of the necessary information you need to write your paper. However,**feel free to also use other resources but cite all your resources!You don't need to cite my videos.**

Lecture 1: Introduction to Neural Networks

Lecture 4: What is a derivative? A gradient?

Lecture 5: Gradient Descent

Write a paper 3 pages or more (at least 1000 words) paper that explains how Neural Networks work.**Your paper must address ALL of following questions.These questions are answered in the videos above. Your paper should be a cohesive piece of writing and should not consist of simply a list of answers to the questions.**

1. Explain the MNIST dataset used. What is its 2D structure? Its flattened structure?
2. Explain the score function or the classifier. What are its parameters?
3. Explain the two basic functions from algebra that forms the basis for a score function. Explain the coding demonstration from the video(Lecture 1). What language is used?
4. Explain the process of training the computer to learn patterns found in images. Why is this process a data-driven approach? Discuss the terms: training set, the testing set, true labels or the ground truth.
5. What is a derivative? Explains and give a simple geometric interpretation. What is a gradient? Give a geometric interpretation of the gradient. Use the analogy from Lecture 4.
6. The main algorithm that makes Neural Networks capable of understanding data is the gradient descent algorithm. Explain this algorithm. How is the gradient used in the algorithm?