

Daniel Sakharuk

1.

The dataset that I will be using is a set of images of mathematical symbols. I want to create an algorithm that will be able to recognize handwritten mathematical symbols. This dataset is one of the only ones I could find on the topic.

The link is: <https://zenodo.org/record/50022#.XiUV88hKhPZ>

2.

I will approach this project by creating a classification algorithm that will have a dictionary of all possible math symbols. I will train the system on part of the dataset. In the end I hope to be able to write a math symbol on paper and after formatting the image I want to be able to have the computer recognize what I wrote.

3.

The dataset seems pretty good. For data preprocessing I don't think I will have to do much. I will have to convert the images into completely black and white pixels.

4.

For the machine learning model I will have to use a version of the classification algorithms. Maybe a CNN will work best. Any OCR algorithm is fair game because I don't know much about them yet.

5.

For the final project I think it will be cool if everyone who visits my station at the fair writes a math symbol on a paper and the machine will recognize it when I input the image. For the web app I can create a page where you upload an image of a math symbol and the symbol that the algorithm detects gets displayed in text format.