Well, I keep going round in circles trying to find a way to do it. Perhaps the easiest way to do it is that we look for genes that are present in both the 47 chromosome database and the autism database. We can then train a deep learning network on these genes only for autism, 47 chromosome syndrome and normal. So the stages would be - look for the identical genes in both datasets. Make a dataset with normal, autistic and 47 chromosome individuals. Label the 47 chromosome individuals (since autism and normal are already labelled). finally train a deep learning model on this data. There will be a few stages to this process. First you should see if there are any genes that appear in both databases

Then we will see how much data we have from all the sources. Once this is done, we can see how much we have got. If that looks OK, I'll transform all the data to turn it into a tagged dataset for training. This will take approx 1/2 day. Once that is done, I'll build the model and train it. That will take another half day in estimation. I'll then be able to provide you with the completed dataset and the trained model. We can discuss any tweaks once the model is up and running and it looks like it will work. Males with 47,XYY syndrome have an increased risk of behavioral, social, and emotional difficulties compared with their unaffected peers. These problems include attention-deficit/hyperactivity disorder (ADHD); depression; anxiety; and autism spectrum disorder, which is a group of developmental conditions that affect communication and social interaction.

Physical features related to 47,XYY syndrome can include increased belly fat, a large head (macrocephaly), unusually large teeth (macrodontia), flat feet (pes planus), fifth fingers that curve inward (clinodactyly), widely spaced eyes (ocular hypertelorism), and abnormal side-to-side curvature of the spine (scoliosis). These characteristics vary widely among affected boys and men (Genetics Home Reference). So, we can include data from related syndromes like ADHD also. OK.. What we are looking for is genes that occur in both the 47 chromosome syndrome data And the autism dataset. but they will be a handful

how are we going to build a deep learning model?

The crux of it is that you need to find a datasets where the data for the same genes exists for 27 Chromosome syndrome individuals/normal individual/autistic individuals. This will be a dataset rather than research papers. This is exactly what they did for the S2 autism/normal dataset. We need data like that but that includes 27 chromosome syndrome patients as well. They actually used a support vector machine which is what I would also go for. You could cluster the autism data to see whether there is a group of higher expressing genes

OK, so what I'll do is look for all of the common genes and create a common dataset ready for training. I can import the datasets into a python scrip and write the code that will do that. So you will end up with a dataset that looks a but like that training set in S3 but with labels for autism/normal and 27 chromosome syndrome

Thats stage one

Stage two will be to train the model

well, create and train the model

Is it possible to do the same model building using

TensorFlow, PyTorch or Spark

If you think it’s possible, we can talk about the budget also because I just realized that the project requires that one of these data tools is used.

I don't know if our model used any activation functions

It would have an activation function for each node in the network

what activation function did it use?

That's how it collates activation from the nodes attached to it and sends it on up the layers

Our is a classical neural net Deep learning model

If its just reinstating the model in tensorflow (I'll use keras on top)