Q3: Use google scholar to find one paper that uses functional data analysis methods to solve the research problem (this cannot be the paper we discussed in class).

Paper Discussed: Classification using functional data analysis for temporal gene expression data.

The research objective was to use machine learning techniques to split each gene expression into gene groups. The researchers proposed a new method which was using functional logistic regression with the implementation of functional principal component analysis. Functional data analysis is very helpful in this dataset since the data collected is sparse. The eigenfunctions of the functional logistic regression was chosen through leave-one-out cross validation. To compare the results this method was compared against the B-Spline implementation of functional data analysis, which is through my knowledge is the industry standard in this field.

After running simulations and the actual experiment the paper had some major contributions. Implementing the proposed methodology successfully and finding great results compared to the industry standards. One of the results was the ability to properly classify a gene sub-group. This has not been done before this paper and is considered a major success. Since the amount of eigenfunctions and base functions are also reduced it is less computationally expensive as well.

Though the comparative results to the industry standard are very good, and it is less computationally expensive as the industry standard there are still limitations to the proposed method. It is still computationally taxing and requires a long time to run. This is mostly due to the leave-one-out cross validation that selects the eigenfunctions. For possible future directions there are also a lot of room for growth since temporal gene expression analysis is a complicated dataset and has a lot of room for improvement in terms of high variation and low accuracy (90%).

Xiaoyan Leng, Hans-Georg Müller, Classification using functional data analysis for temporal gene expression data, Bioinformatics, Volume 22, Issue 1, 1 January 2006, Pages 68–76, https://doi.org/10.1093/bioinformatics/bti742