

Machine Learning-based Genome Wide Association Studies of Rheumatoid Arthritis

Allie Burton

Advisor: Prof. Sandra Batista

Introduction

Rheumatoid arthritis (RA) is an autoimmune disease affecting the membrane between the joints causing pain, joint damage, and eventually severe disability. Although RA has been identified as a multifactorial disease (meaning that there is more than one factor causing it)[1], this paper will specifically focus on the genetic components.

To look for these genetic components, researchers perform what is known as a genome-wide association study, or GWAS, which the National Human Genome Research Institute defines as “an approach that involves rapidly scanning markers across the complete sets of DNA, or genomes, of many people to find genetic variations associated with a particular disease.” The results of these studies, however, have been generally inconclusive collectively. For example, Takahashi et. al.’s genome-wide association study (GWAS) studying approximately 1,400 Japanese females shared data on 87 of the top 100 single nucleotide polymorphisms (SNPs) found in Sharma et. al.’s [2] racially diverse GWAS of 419 families, however, only one of those SNPs showed significant association in Takahashi et. al.’s GWAS [3]. The goal of my project is to study the usage and efficacy of machine learning for GWAS of scoliosis.

Background

Related Work

Although there have not been any studies done to date using machine learning for GWAS of idiopathic scoliosis, there have been many studies using machine learning for other phenotypes including IgM and rheumatoid arthritis, as mentioned above, in addition to myocardial infarction, coronary artery calcification, and anti-cyclic citrullinated peptide[4]. These studies will provide the basis for my methodology, specifically D'Angelo et. al.'s [5] and Tang et. al.'s [6] GWASs of rheumatoid arthritis and Stassen et. al.'s GWASs of IgM. Since there are no prior machine learning-based GWASs of AIS, I plan replicate their respective methodologies as best as possible, adapting where necessary for the specifics of scoliosis and the data sets I use.

References

- [1] Y. Alamanos and A. A. Drosos, "Epidemiology of adult rheumatoid arthritis," *Autoimmunity Reviews*, vol. 4, no. 3, pp. 130–136, 2005.