Visualizing Genetic Influence on Autism Spectrum Disorder (ASD) Diagnosis: The Roles of Colors and High Dimensional Graphs

Researcher: Scott Neustadt (sfnzr6@mst.edu)
Advisor: Dr. Bih-Ru Lea (leabi@mst.edu)

Abstract

As new discoveries of Autism Spectrum Disorder (ASD) become prevalent, a need for faster and more efficient diagnosis is essential. Results from my 2015-2016 OURE project suggest that color and high dimensional graphs have a large impact on data retention of genetic data and helps in the discovery of important information hidden within the data. My proposed project will develop an interactive visualization software prototype for ASD genetic research data based on literature review and experiment design from my 2015-2016 OURE project.

My research will focus on educating people of different backgrounds (age, gender, race, etc.) on the genetic factors that have a direct influence on ASD through the roles of color and high dimensional graphs, particularly the chromosome locations of affected genes. I have collected and organized genetic data taken from SFARI Base and NDAR for my OURE research project that can be used to develop the proposed interactive visualization decision dashboard to collect data for the proposed hypotheses testing, analyze results and make recommendations.

The proposed research will develop multiple high dimensional visualization models to allow effective and efficient ASD diagnosis. From literature review I hypothesize that by utilizing interactive 3D visual models (such as a treemap) with a cold color scheme will help the general user to better understand genetic information by allowing better data retention and data discovery. It is also expected that results and prototypes from the proposed will contribute to the process that doctors and physicians diagnose autism, making for more efficient and earlier diagnosis.

Reflection on the Learning Experience

Research starts with an idea. You build on that idea by supporting it with literature review. Even though having new ideas is how great things happen, it is good to have a foundation that supports your idea and makes it worth pursuing. After gathering enough evidence to support your idea you should form a hypothesis based on your research. Having a hypothesis helps form your plan of action

and helps in the next step of constructing a model to test the hypothesis. You analyze your results and record your findings.

As my project for 2015-16 was mostly literature review I found the vast amount of information that the library at Missouri S&T provides through its databases and resources is very useful. I mostly used Scopus to find the articles that I referenced. I found that these resources were best utilized before writing my final report. I saved a lot of time by finding supporting literature before I started writing. I also used SAP HANA, Design Studio and Predictive Analysis to aid in the data analysis portion of my project. These tools will also help me construct my proposed visualization prototype in my continued project. The ASD (Autism Spectrum Disorder) data that I am using for this project is from two databases: The Simons Foundation Research Initiative (SFARI) and the National Database for Autism Research (NDAR). Both SFARI and NDAR made finding the information I needed easy, although formatting of the information will most likely prove to be difficult.

I learned that it is a tedious and time consuming process. You can't base your design on instinct. You have to describe your design and why it is the way it is with foundational support. I find it best to draw out your ideas on paper and start from there. The whole point of experimental design is to test a hypothesis. Take your time and make sure everything is done in accordance to your plan.

My current project results are what have helped me form a hypothesis for my future research design and survey prototyping. Multiple literatures suggest that there is a need for a better way to visualize genetic information in order to make for earlier and better autism diagnosis. Both color and high dimensional graphs have proven to make a large impact on people of different demographics. Utilizing these with genetic information is the next step for diagnosis.