

nicolasfishman

njw.fish | (919) 995-1050 | njwfish@stanford.edu

interests

machine learning, bioinformatics, cryptography, privacy

education

- 2017 - 2021 **Stanford University, Stanford, CA**
- Pursuing a double major in Computer Science and Bioengineering
 - Relevant courses: Probability Theory, Quantum Chemistry, Genetics
 - Upcoming courses: Design and Analysis of Algorithms, Linear Algebra and Matrix Theory
- 2016 - 2017 **George Washington University, Washington, DC**
- Relevant courses: Machine Learning, Bioinformatics, and Cryptography
- 2013 - 2017 **Woodrow Wilson High School, Washington, DC**
- Weighted GPA: 4.5

experience

- 2017 **Technical Specialist** Star Lab Corporation
- automated the exploration of the strategy space of defensive and offensive cybersecurity, working on a small research team to:
 - develop a game theory computation library (**C**)
 - build a system to dynamically load and execute abstract Python modules on remote machines, with and without authorization credentials (**Python**)
 - developed project proposal for a system to detect shifts in training distributions in online learning algorithms to detect attacks using malicious training data
- 2015 - 2017 **Research Fellow** Comparative Genomics Branch, National Institutes of Health
- developed Database of Genomic Structural Variation (DOGSV) (**MySQL, Python**)
 - developed specialized random forest algorithm to identify candidate causal areas from single nucleotide permutation (SNP) data (**Matlab**)
 - developed script to identify inheritance model in SNP-SNP interaction (**C++**)
 - designed parallelized algorithm to identify quality of SNP tagging in a region (**Python**)
- 2013 - 2017 **Lead Programmer** Woodrow Wilson High School FIRST Robotics
- programmed FRC robot for 4 years, made it to World Championships twice (**Java**)
 - received Innovation and Control Award for vision targeting system (**Python**)

projects

- 2016 - 2017 **DOGSV** Comparative Genomics Branch, National Institutes of Health
- Built database for the storage and analysis of hundreds of millions of structural variants. Trained model to cluster variants and assess the likelihood of variants being false positives.
- 2017 **Predict Population Diversity from Unassembled Reads** George Washington University
- Developed model to estimate the genetic diversity of viruses with high rates of nucleotide substitution, using a nearest neighbor regression on k-mer analyses of the raw reads.
- 2017 **Predict Outcomes in Fantasy Baseball** George Washington University
- Developed model to predict outcomes in fantasy baseball by building a dataset of MLB player histories, and then proving the validity of using player histories to predict team outcomes.
- 2016 **Legal Aid Dashboard** Drupal Legal Aid
- Using Google web traffic data for various legal aid websites, classified the topics of the most popular web pages by location to facilitate allocation of legal aid resources to meet need.