

## **Objective**

I am a recently graduated physics Master's/former PhD student making the exciting leap into data science.

## **Education**

### **Central Connecticut State University, New Britain, CT 06053**

- Graduated winter 2019
- Bachelor of Science in Physics, Bachelor of Arts in Mathematics, Numerous Mechanical Engineering Credits.
- 3.69 GPA/4.0 Scale

### **University of California Santa Cruz, 1156 High St, Santa Cruz, CA 95064**

- Graduated with Master's from physics PhD program (Spring 2021)

## **Skills**

- Python (NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, Mlxtend), R (Ggplot, Dplyr, TidyR), SQL
- Exploratory analysis, Frequent patterns, Classification, Regression, Clustering, Knn, Neural networks, Parameter optimization, NLP, Web-Scraping (Selenium).
- Proficiency in all Microsoft Office applications, Entry level Matlab.
- Technical writing, oral presentations of research as poster or PowerPoint, team-oriented research, former university physics lab instructor and teacher's assistant.

## **Experience**

### **Exploratory Data Analysis of Professional (CS:GO) Gamer's Gear and Settings, and Modeling Player Accuracy Performance**

- Exploratory data analysis to describe trends in what gear/settings pro players prefer, and modeling whether a player has above average aim based only on their gear/settings.
  - Managed an increase in model classification accuracy of 17% (to 67%, without overfitting) over original data through feature creation (binning, frequent patterns, clustering, separation by player role), filter methods (chi2 tests with class target), wrapper methods (sequential feature selection) and parameter optimization (Gridsearch). (*Python, NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, mlxtend, exploratory analysis, Classification, regression, clustering, knn, gridsearch*).

### **Exploratory Data Analysis of Stroke Dataset in R**

- Exploratory data analysis of a Kaggle stroke dataset using Rstudio (R, Ggplot, Dplyr).

## **Relevant Coursework**

- CSE 242: Data Mining (Grad-level) – Theory, classification, and unsupervised learning.
- CSE 243: Machine Learning (Grad-level) – Theory, clustering, classification, and deep learning.
- ME 345: Engineering Statistical Analysis of Operations (Multiple regression, ANOVA, Statistics, Minitab).
- ENGR 240: Computational Methods for Engineers (Matlab and advanced functions of Excel).
- Math 218: Discrete Mathematics
- Math 377: Intro to Real Analysis
- Math 366: Intro to Abstract Algebra