

# SAMMY MUSTAFA

BOSTON, MA | 201-957-4754 | [sammymustafa@hms.harvard.edu](mailto:sammymustafa@hms.harvard.edu)

**AI Engineer**

[LinkedIn](#) | [Website](#) | [GitHub](#)

## EDUCATION

**Harvard Medical School** | Boston, MA

*Expected Dec 2024*

*MS in Biomedical Informatics* | GPA: 4.0

- **Relevant Coursework:** ML for Computational Biology (MIT), Computational Biomedical Statistics, Computational Psychiatry (ETH Zurich)
- **Activities:** Harvard Biotech Club, Harvard Innovation Labs, Harvard Graduate Consulting Club, Camp Kesem

**Northwestern University** | Evanston, IL

*Jun 2023*

*BA in Computational & Systems Biology, Data Science, and Linguistics* | GPA: 3.93 (Dean's List 8/9)

- Honors Program in Medical Education (**HPME**): 7-year BA/MD program
- **Relevant Coursework:** Data Science I-III, Data Visualization, Text Processing, Bioinformatics, Biostatistics

## DATA SCIENCE EXPERIENCE

**ML Researcher** | *Center for Precision Psychiatry, Mass General Hospital* | Boston, MA

*Oct 2023 – Present*

- Identify and validate brain-based depression biotypes through clustering of 4,000 fMRI neuroimaging features
- Develop 50+ predictor regression models on large-scale genetic data to dissect the polygenic nature of MDD
- Design novel methodologies to preserve spatial integrity information from imaging data using 3D deep learning

**Data Scientist** | *Behavior* | Pittsburgh, PA

*Sep 2023 – Dec 2023*

- Implement supervised and unsupervised ML models on 100 parameter, 200k record datasets to predict cravings
- Leverage label propagation algorithms for refined, improved scikit-learn clustering (15% gain in model precision)
- Improve Cypher algorithm efficiency for accelerated real-time anomaly detection in patient wearable devices

**NLP Researcher** | *Behavioral Research Lab, Yale University* | New Haven, CT

*Mar 2022 – Jul 2022*

- Conducted a comparative analysis of bargaining methods in verbal and non-verbal communication experiments
- Analyzed over 6,000 minutes of negotiation data to quantify and discern key linguistic behaviors via Python
- Employed sentiment analyses and word embeddings to identify effective buyer and seller negotiation strategies

## LEADERSHIP EXPERIENCE

**Associate Consultant** | *H2Ok Innovations* | Boston, MA

*Oct 2023 – Nov 2023*

- Scaled application of artificial IoT-enabled industrial liquid optimization across top-tier Fortune 500 enterprises
- *Project manager:* conducted strategic market and industry sensitivity analyses for a \$14 billion market entry
- Enhanced revenue generation by 25% through tailored, client-specific pricing models from competitive analyses

**Strategy Consultant** | *Mercurial-AI* | Chicago, IL

*Mar 2023 – Aug 2023*

- Developed and implemented strategic operational changes that successfully secured \$500k in seed capital
- *Project manager:* overlooked deep ensemble/casual framework personalizing breast cancer treatments

## PUBLICATIONS & PROJECTS

**Predicting Future Depressive Biotypes from rs-fMRI with Bayesian Network Models** – Harvard Medical School

- Pioneer depressive biotype prediction with ensemble clustering of UK Biobank neuroimaging and genomic data
- Leverage generative embedding to forecast/predict initial depressive episodes in never-depressed individuals

**Optimized Pharmacophore Modeling: Enhancing Drug-Ligand Dynamics with DL-Conformational Analyses** – MIT

- Identified caveats in deep learning molecular docking techniques impacting prediction accuracy and feasibility
- Developed novel harmonic prior models utilizing 3D structural consensus distances from multiple ligand conformers with thresholds on long-range pairwise distances for enhanced predictions (10% lower RMSD)

**SERBP1: Exploiting RNA-Binding Protein-Mediated PAI-1 Inhibition** – Feinberg School of Medicine

- Developed a transgenerational epigenetic method of prion-mediated PAI-1 destabilization for cardiac fibrosis
- Compared gene expression patterns with and without SERBP1 treatment *in vitro* via PCA on RNA-Seq data in R

## AWARDS

Top 20 in 2023 Penn Healthcare Case Competition  
Selected Venture @ Harvard Innovation Lab

HPME (2021) and Weinberg (2022) Summer Research Grants  
Regeneron Science Talent Search Scholar

## SKILLS

**Languages:** Python, R, SQL, C++, MATLAB, HTML, JavaScript

**Technologies:** PyTorch, TensorFlow, Git, Linux, RL, NLP, API

**Additional Skills:** Spanish (Intermediate), Arabic (Intermediate)