

# Aashish Cheruvu

2508 Dienes Way, Warrington PA 18976 aashcheruvu@gmail.com (215) 500-5766

---

## EDUCATION

Central Bucks High School South, Warrington, PA – Cum. Unweighted GPA: 4.0 (Weighted: 4.60)

Class Rank: 1/577

Harvard Extension School Classes:

- Introduction to Statistical Modeling (2021) – Grade: A
- Web Applications for Data Science (2022) – Grade: A
- Computer Vision and Wearable Devices (2023)

Standardized Test Scores:

- American Mathematics Competitions (AMC 12): 91.5 (Top 10%, Qualified for American Invitational Mathematics Examination - AIME)
- AP Statistics - 5/5 (8<sup>th</sup> Grade); AP Calculus BC – 5/5 (9<sup>th</sup> Grade); AP Biology – 5/5 (10<sup>th</sup> Grade); AP Computer Science A (10<sup>th</sup> Grade); AP Chemistry - 5/5 (11<sup>th</sup> Grade); AP English Language and Composition - 5/5 (11<sup>th</sup> Grade)
- SAT score – 1570 (composite), Math-800/800, Reading/Writing 770/800 (11<sup>th</sup> Grade)
- PSAT score – 1520 (composite), Math-720/720, Reading/Writing 720/720 (11<sup>th</sup> Grade)
- ACT score – 34 (composite), Math-36/36, Science 31/36, English 36/36, Reading 34/36 (7<sup>th</sup> Grade)

## EXPERIENCE

**Yale School of Medicine** | Incoming Intern in Dr. Julius Chapiro: Radiology/Computer vision: June 2024 New Haven, CT

**Perelman School of Medicine, University of Pennsylvania** | Student Research Intern in Dr. Abbas Alavi: Radiology/Computer vision August 2023 – Present | Philadelphia, PA

**Lockheed Martin** | Machine Learning Engineering Internship: Deep Learning/NLP June 2022 – Present | Bethesda, Maryland

Applied deep learning techniques with text and sensor data using transformer/diffusion models; Analyzed large numbers of application log files to produce visualizations providing business insights using Splunk and Tableau

**Pennsylvania Biotechnology Center** | Research Internship: XAI-Based RNA Biomarker Discovery August 2022 – Present | Doylestown, Pennsylvania

Applied deep learning techniques for identifying RNA biomarkers; Performed in-lab validation of the biomarkers using molecular biology techniques (e.g., qRT-PCR)

**Food & Drug Administration** | Research Internship: Disease Progression Modeling June 2022 – August 2022 | Silver Spring, Maryland

Application of deep learning-based time series modeling using multimodal data in disease progression

## HONORS

2023: Coca-Cola Scholarship – Semifinalist; PennFTC State Championship Finalist; Qualified for AIME; HOSA State Leadership Conference- Health Informatics- First Place; Future Business Leaders of America (FBLA) State Database and Design – Top 5 in Nationals

2022: PennFTC State Championship Finalist; PennFTC Regional Think Award 1<sup>st</sup> Place; Lockheed Martin CodeQuest Senior Division – 1<sup>st</sup> Place; Future Business Leaders of America (FBLA) Regional Database and Design – 1<sup>st</sup> Place

2021: HOSA International Leadership Conference- Health Informatics-Silver Medalist; Bucks County Research Science Competition Computer Science Category- First Place; Delaware Valley Science Fairs – 3<sup>rd</sup> Place; Office of Naval Research- Naval Science Award.

2020: Delaware Valley Science Fairs (Gold Medal); BROADCASTMM Masters Entrant; Bucks County Research Science Competition (Mathematics Category) - CBCares Educational Foundation Research Award; You Bethe Chemist Competition National Competition Entrant (Final discontinued due to COVID)

2019: Scholars of Distinction- Duke University Talent Identification Program

## LEADERSHIP

Founder President, Data Science Club – High School: (2021 – present)

President, North American Telugu Society (NATS) – (2019 – Present)

Biotechnology Partnership Program Leadership Council (2022 – present)

HOSA (Future Health Professionals) Club Leadership Council (2021 – present)

Lead Programmer – FTC Team 16606 Tech-Knight Robotics (2021 – present)

FBLA (Future Business Leaders of America) Club Leadership Council (2022-present)

STEM Research Club Leadership Council (2020 – present)

Telugu Association of Greater Delaware Valley (TAGDV) Youth Committee (2019 – present)

## PROGRAMMING SKILLS

Python • R • SQL • Machine Learning • Deep Learning • Natural Language Processing • Computer Vision • rust • Java • unix • vim • Google Cloud

## PUBLICATION

Aashish Cheruvu, "Multimodal Recommender System in the Prediction of Disease Comorbidity," 2022 Fourth International Conference on Transdisciplinary AI (TransAI), 2022, pp. 79-82, doi: 10.1109/TransAI54797.2022.00020.

Aashish Cheruvu, " Application of Neural-Controlled Differential Equations in Disease Progression Modeling Using Irregularly Sampled Multimodal Data," 2023 Fifth International Conference on Transdisciplinary AI (TransAI), 2023.

Aashish Cheruvu, and Nathaniel Rigoni " ILDNet: A Neural Differential-Equation Based Disease Progression Model for Pulmonary Fibrosis," 2024 Society for Industrial and Applied Mathematics International Conference on Data Mining (SDM24), Houston, TX, US. (Under Review)

Aashish Cheruvu, " Bagged Fuzzy-Rough Nearest Neighbors (BFRNNs): A Novel Ensemble Learning Algorithm for Disease Diagnosis and Prognosis Prediction," Submitted to ArXiv. (Under Review)

## PRESENTATION

Invited presentation to ORISE fellows on "Multimodal Recommender System in the Prediction of Disease Comorbidity", Department of Pharmacometrics, Center for Drug Evaluate and Research, Food and Drug Administration (FDA), 2022.

Presented at 2022 Fourth International Conference on Transdisciplinary AI (TransAI) on Multimodal Recommender System in the Prediction of Disease Comorbidity

# PROJECTS

## **Deep Learning-based RNA Biomarker Discovery Using Transformer Models and Explainable AI (2023)**

Developed and evaluated multiple attention-based models and graph neural network-based models (graph attention networks, GATs; graph transformers) to identify biomarkers for early-stage liver diseases (non-alcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH)) using explainable AI. In-lab validation of identified was evaluated using reverse-transcription-quantitative polymerase chain reaction (RT-qPCR).

## **Bucks County Science Research Competition | A Novel Application of Neural Controlled Differential Equations in Disease Progression Modeling Using Multimodal Deep Learning (2022)**

Developed and evaluated multiple time series models (LSTM, Sequence-to-Sequence, and Neural Controlled Differential Equations) in combination with image models (EfficientNet) in predicting Alzheimer's and Pulmonary Fibrosis progression using both structured and image data.

## **Delaware Valley Science Fair | A Novel Application of Deep Learning Recommender Systems in Disease Diagnosis (2021)**

Developed and evaluated an implicit feedback-based hybrid content/collaborative filtering model in disease prognosis. Text features were extracted from clinical notes using SciSpacy, and feature combination was used as the hybridization technique.

## **Bucks County Science Research Competition | A Deep Learning Based Recommender System for Disease Prognosis (2021)**

Developed and evaluated an implicit feedback based neural collaborative filtering model in disease prognosis. Used the MIMIC-III database and performed well compared to the literature.

## **Using Medical Transcriptions to Accurately Predict Medical Specialties with tidymodels (2021)**

Extracted important text features from medical transcription text and used the tidymodels packages to train and evaluate the performance of various standard machine learning models in predicting medical specialty.

## **Use of Machine Learning Approaches in Predicting Chemical Reaction Rates of Compounds with OH and HO<sub>2</sub> in the Atmosphere (2020-2021)**

Extracted physico-chemical features of compounds and applied various machine learning techniques with the goal of predicting reaction rates of compounds with OH and HO<sub>2</sub> in the atmosphere.

## **Broadcomm Masters | A Novel Machine Learning Algorithm in Disease Diagnosis and Prognosis (2020)**

The performance of the Bagged Fuzzy Rough Nearest Neighbors Algorithm was evaluated on three datasets: the Wisconsin Breast Cancer Dataset, the Wisconsin Prognostic Breast Cancer Dataset, and the Pima Indian Diabetes Dataset. Various machine learning algorithms were used as baselines to evaluate relative performance.

## **Delaware Valley Science Fair | A Novel Machine Learning Algorithm in Aiding the Diagnosis of Breast Cancer (2020)**

Created a novel machine learning algorithm, Bagged Fuzzy Rough Nearest Neighbors that is an ensemble of the Fuzzy Rough Nearest Neighbors and uses a Borda count based voting metric. The algorithm was tested on the Wisconsin Breast Cancer Dataset and was evaluated alongside other baseline machine learning algorithms.

## **Bucks County Science Research Competition | A Comparison of Machine Learning Methods for Breast Cancer Diagnosis (2020)**

Compared the performance of several standard machine learning algorithms in breast cancer prediction using the Wisconsin Breast Cancer Dataset.