

# Sarai Hoffberger

1021 Dulaney Valley Road, Towson, MD 21204



410-337-9191



[sarai.hoffberger@goucher.edu](mailto:sarai.hoffberger@goucher.edu)



[linkedin](#)



[github](#)

## Education

### Goucher College

Bachelor of Arts in Computer Science and Data Science

May 2026

Towson, MD

## Work Experience

### Institute for Research On Poverty

Data Analyst Intern

August 2024 – May 2025

Madison, WI

- Processed survey data (10,000+ entries) using Pandas and NumPy, ensuring 95%+ data accuracy.
- Built machine learning models (Logistic Regression, Decision Trees) with 85%+ accuracy to forecast poverty levels.
- Designed interactive dashboards with Matplotlib and Seaborn, enhancing insight delivery by 40% and achieving 90%+ satisfaction from policymakers.
- Validated models with Scikit-learn in Jupyter Notebooks, delivering actionable food security recommendations.
- Tools: Pandas, NumPy, Python (scikit-learn, machine learning), Matplotlib, Seaborn.

### Goucher College Biology Department

Undergraduate Research Assistant

May 2023 – May 2025

Towson, MD

- Analyzed 200+ student backgrounds using statistical techniques, uncovering trends to support diversity initiatives.
- Designed and launched a responsive website, improving accessibility for 1,000+ users.
- Collaborated with 5 faculty to propose actionable solutions promoting diversity in STEM majors.
- Streamlined operations by implementing a digital tool, reducing manual tasks by 30%.

### Goucher College Summer Science Research

Undergraduate Research Assistant

May 2024 – July 2024

Towson, MD

- Developed a predictive model using neural networks and random forest, achieving 95% accuracy in predicting UHI effects.
- Analyzed 10,000+ data points from satellite imagery and weather stations to identify trends.
- Identified 3 key factors contributing to urban heat islands and proposed 2 actionable mitigation strategies.
- Tools: Python (scikit-learn, machine learning), ArcGIS.

## Projects

### NCAA Data Analysis

January 2025

- Analyzed 100 years of NCAA data (1912-2012), uncovering athlete performance patterns and trends.
- Built a Streamlit app with 10+ interactive visualizations, delivering insights for 100+ universities.
- Discovered 5+ historical patterns, predicting future trends with 85%+ accuracy using data analysis techniques.
- Tools: Python, Streamlit, Plotly, Seaborn, Matplotlib

### Sentiment Analysis

October 2024

- Developed a sentiment analysis model for movie reviews with a 97% accuracy rate, leveraging TensorFlow for efficient training.
- Built an interactive Streamlit app to classify sentiments in real-time, processing 1,000+ reviews during testing.
- Improved preprocessing efficiency with NLTK, reducing model error rates by 15%.
- Tools: Python, Streamlit, Tensorflow, NLTK

## Technical Skills

- **Programming:** Python, Java, C, C++, Haskell
- **Data Science Tools:** Scikit-learn, TensorFlow, Pandas, NumPy, Matplotlib, Seaborn, SciPy
- **Web Development:** HTML/CSS, JavaScript, Streamlit
- **Developer Tools:** Git, Docker, SQL, AWS, GCP