

# Dan Nguyen

xxx-xxx-xxxx | [sidan.nguyen@gmail.com](mailto:sidan.nguyen@gmail.com) | [linkedin.com/in/dannguyen111](https://www.linkedin.com/in/dannguyen111) | [github.com/dannguyen111](https://github.com/dannguyen111)

## EDUCATION

### Gettysburg College

*Bachelor of Science in Computer Science and Mathematics. Cum. GPA: 4.11.*

Gettysburg, PA

*Jan. 2023 – May 2026*

## EXPERIENCE

### AI Engineer Intern

KPMG

June 2025 – Aug. 2025

Berlin, DE

- Pioneered a project to train a specialized LLM for Global Audit, building an ETL + GraphRAG pipeline with **Python** and **Neo4j** that improved QA processing efficiency by 30%.
- Engineered prompt-based LLM solutions using **LangChain** and **Jupyter Notebooks** for NLP tasks like markdown extraction, enhancing document parsing reliability by 30%.
- Developed QA dataset generators and an "LLM-as-a-judge" evaluation framework in **Python**, enabling scalable data augmentation and model validation.

### Data Scientist Intern

KPMG

May 2024 – Aug. 2024

Berlin, DE

- Leveraged **Python**, **Jupyter Notebooks**, and **MS Excel** to conduct data analysis on public bank datasets, developing a deterministic rule-based baseline model for peer group clustering.
- Engineered a **stochastic unsupervised algorithm** for the **Clara 2.0** AI, a system for **anomaly detection** in accounting data to identify potential failure risks, achieving a **15% improvement** in peer group cohesion.
- Designed and evaluated various peer group quality measures by comparing different methodologies to validate and select the most effective models for the risk assessment tool.

## PROJECTS

### Saint Petersburg AI Player | *Java, Python* | [View on GitHub](#)

Aug 2025 – Dec 2025

- Engineered an intelligent agent utilizing **Monte-Carlo Tree Search** and heuristic evaluation functions trained via **logistic regression** and **random forest** models, achieving a **first-place finish** in a class tournament.
- Analyzed data from Monte Carlo simulations between players to identify card imbalances and generated a **balance change report** to optimize game mechanics.

### Mancala AI Player | *Java, Python* | [View on GitHub](#)

Jan 2025 – Feb 2025

- Implemented a competitive Mancala AI using minimax search with alpha-beta pruning, transposition tables, and heuristic evaluation for efficient state-space exploration.
- Designed a dynamic time management module to adapt search depth and node expansion based on real-time computational budget and board complexity, achieving top-tier class performance.

### Quantitative Analysis of Healthcare Disparities in Pennsylvania | [View Paper](#)

Spring 2024

- Analyzed county-level healthcare access disparities in response to a **Deloitte-issued prompt**, using the Gini coefficient, Atkinson index, and a weighted Utilitarian welfare function.
- Performed regression analysis linking income, rurality, and uninsured rates to provider distribution, delivering data-driven insights to **contribute to Deloitte's decision-making** process.

## TECHNICAL SKILLS

**Languages & Frameworks:** Java, Python, JavaScript, HTML/CSS, php; LangChain, LangGraph, scikit-learn

**Data Stack:** Neo4j, SQL, mySQL; pandas, NumPy, Matplotlib, Microsoft Excel

**Developer Tools:** Git, Azure DevOps

## ACTIVITIES

### President & Founder

*Gettysburg College Fashion Initiative*

January 2024 – December 2024

Gettysburg, PA

- Founded the organization and managed a 13-member executive board across five departments.
- Served as Creative Director for our fashion show, managing a \$10,000 budget and drawing 300 attendees.
- Established and ran a semiannual Vintage Market on college campus, raising \$9,000 for charity.