Sonia Tyburczy

Brooklyn, NY | soniatyburczy@gmail.com

EDUCATION

Baruch College, City University of New York *New York, NY*
B.A. in Mathematics — GPA: 3.88/4.00 *Expected December 2027*
Relevant Coursework: Linear Algebra, Calculus III, Introduction to Probability, Introduction to Machine Learning

RESEARCH & PROJECTS

MTH 4330 Final Project: Structural Predictors of Chronic Absenteeism in NYC High Schools
Baruch College | February 2026 – May 2026

- Investigated school-level chronic absenteeism as a supervised regression task using three merged NYC OpenData sources (attendance, demographics, socioeconomic indicators), spanning 2013–2018 with 7,607 observations and 50 features.
- Built and evaluated OLS and polynomial regression models (degrees 1–10) in R, applying bias-variance tradeoff analysis to identify optimal model complexity (degrees 3–4) under temporal train/test splitting to simulate real-world forecasting conditions.
- Discovered and documented an undisclosed 2017–18 methodological shift in NYC's Economic Need Index (ENI) that inflated values and created distributional mismatch between tune and test splits, substantively affecting model generalization.
- Implemented Random Forest and XGBoost with 5-fold cross-validation and hyperparameter tuning; conducted SHAP analysis to interpret feature contributions, identifying % Students with Disabilities as a structurally distinct predictor of absenteeism independent of economic need.
- GitHub: https://github.com/soniatyburczy/absenteeism-structural-predictors

Independent Study: QLoRA Supervised Fine-Tuning for LLaMA-2
Baruch College | September 2025 – December 2025

- Designed and implemented the full training and evaluation pipeline within a 5-person independent study, enabling adapter-based fine-tuning of LLaMA-2-7B-Chat on a single 16GB T4 GPU using QLoRA.
- Developed dataset cleaning heuristics and a structured prompt schema to reduce hallucinations across an 813-sample weakly supervised dataset.
- Achieved +43% relative ROUGE-L F1 improvement on long-form outputs (≥500 chars); evaluated using ROUGE-L, BERTScore, and repetition ratio with critical engagement with the limits of automated metrics for generative output quality.
- GitHub: github.com/soniatyburczy/llama2-qlora-sft-coverletter-project

EXPERIENCE

Analytics Intern *Incoming July 2026*
NYC Department of Youth & Community Development | New York, NY

Student Core Team Member, AI Leadership Initiative *May 2026 – Present*
Office of the President | Baruch College, New York, NY

- One of seven students selected college-wide to serve on the Campus Core Team for Baruch's AI Leadership Initiative, a presidential initiative advancing institutional AI capacity across teaching, learning, research, and administration.
- Working alongside faculty, staff, and administrators to develop AI governance frameworks, baseline AI curriculum and learning resources, and responsible AI policy for the institution.

Undergraduate Design Consultant, AI Discernment Project *April 2026 – Present*
Baruch College | New York, NY

- Member of a small student advisory group providing input on the design of a new first-year course on AI literacy.

Machine Learning Intern *January 2025*
Permanence | New York, NY

- Researched cross-language debugging of Common Weakness Enumerations (CWEs) using large language models, analyzing how code examples in one language influenced AI-generated fixes in another (C/C++ to Java).
- Designed an evaluation framework using cosine similarity, Jaccard similarity, and CodeBERTScore to assess semantic similarity of generated fixes across languages.
- Refined prompt structure and CWE examples to improve cross-language semantic similarity from 0.80 to 0.99; presented findings in a capstone report enabling pursuit of cross-language debugging features.

LEADERSHIP & SERVICE

Project Manager, Machine Learning & Data Science Club *February 2026 – May 2026*
Baruch College | New York, NY

- Led a 5-person team in designing and implementing a retrieval-augmented generation (RAG) pipeline for a CUNY-domain question answering system across the full stack: scraping, chunking, embedding, ChromaDB retrieval, and LLM integration.

Student Representative, Joint Committee for Student Evaluation of Faculty and Teaching
Baruch College | New York, NY | April 2026 – Present

- Serve as a student voice on an inter-constituency committee reviewing and redesigning Baruch's faculty evaluation instrument; member of the analysis subcommittee responsible for examining the current tool to identify strengths, limitations, and areas for improvement.

Vice Chair of Research and Policy, Academic Affairs Committee *August 2025 – May 2026*
Undergraduate Student Government (USG) | Baruch College, New York, NY

- Conducted higher education policy research to inform institutional initiatives including co-developing a scholarship proposal for international and first-generation students — researching program models, third-party administrators, and eligibility frameworks to support the initiative's design.

TEACHING

Peer Mentor *August 2025 – December 2025; returning Fall 2026*
Baruch College | New York, NY

- Co-led weekly recitation sessions for a cohort of 25 first-year students, designing and facilitating interactive lessons to support their transition to college-level work.
- Managed course logistics including grading, attendance tracking, and course communication via Brightspace.

SKILLS

Programming & Libraries: Python, R, PyTorch, Scikit-learn, NumPy, Pandas, Hugging Face Transformers, SHAP, Git

Methods: Data Preprocessing & Cleaning, Feature Engineering, Polynomial Regression, Ensemble Methods (Random Forest, XGBoost), Cross-Validation & Hyperparameter Tuning, Parameter-Efficient Fine-Tuning, Supervised Fine-Tuning, Prompt Engineering, Causal Language Modeling, Quantization, Dataset Curation & Quality Filtering, Retrieval-Augmented Generation (RAG)

Evaluation: Regression Metrics (RMSE, MSE, R^2, Adjusted R^2), NLP Metrics (ROUGE-L, BERTScore), SHAP Analysis, LLM Benchmarking, Hallucination & Repetition Analysis, Qualitative & Quantitative Analysis

Languages: English (native), Polish (fluent), French (intermediate)