

# Noah Nguyen

(226) 899-7437 | [EMAIL](#) | [LINKEDIN](#) | [GITHUB](#) | [PORTFOLIO](#)

## Education

### University of Waterloo

Expected Dec 2025

*BMath (Hons) in Statistics, Minor in Economics*

*Waterloo, ON*

**Courses:** Statistics (Bayesian & Frequentist Methods), Probability, Calculus, Linear Algebra, DSA (Python)

**Community Involvement:** Data Science Club, Games & Sports Analytics Club, Actuarial Science Club

## Work Experience

### Data Scientist Intern – AI Trainer

Oct 2024 – Present

*Outlier*

*San Francisco, CA*

- Refining Gen AI models to enhance performance in NLP tasks by ranking domain-specific text generated by LLMs
- Developing mathematical queries to improve models' ability to answer questions with context awareness

### Actuarial Analyst Intern – Product Development

May 2024 – Aug 2024

*Bank of Montreal (BMO)*

*Toronto, ON*

- Conducted A/B testing across multiple platforms based on Gen Z's 8-second attention span and experimented with a gamification strategy for a travel insurance app, decreasing cost projections by 18%
- Collaborated with cross-functional stakeholders to leverage SQL and Python and utilize libraries to automate the Risk Integrity reporting process, reducing over 20 hours of work for the team
- Enhanced policyholder services and operational efficiency by building an NCPI calculator with VBA scripts and new mortality tables, reducing reserves by over \$500K, and decreasing pricing discrepancies by 16%

### Data Scientist Intern – Business Development

Jun 2023 – Sep 2023

*Mccredit - MB Shinsei Finance Limited*

*Vietnam*

- Utilized pre-trained NLP models to analyze customer intent toward the company's cross-selling strategies
- Implemented debt-to-income ratio (DTI) and credit utilization ratio metrics to remove 2,500+ high-income individuals, resulting in a 25% increase in loan approval accuracy and increasing revenue by over \$32,000
- Developed a time series anomaly detection model to flag suspicious income spikes such as borrowing from external sources, reducing high-risk approvals by 15%; utilized Tableau from model outputs to present insights to leadership

### Data Analyst Intern – Early Education

Jun 2022 – Sep 2022

*VUIHOC - K12 Online Education Platform*

*Vietnam*

- Applied SQL & Pandas to classify Math questions, boosting categorization accuracy by 34% across topics
- Collaborated with stakeholders to utilize OBS Studio to host pre-recorded educational videos by subject and broadcast content to over 5 million users nationwide, boosting quarterly revenue by \$7,500

## Projects

### Auto Car Claim Classification Challenge | [GitHub](#)

- Engineered Expected Claim Size as a proxy variable for CLV using Tweedie regression, boosting F1 score to 0.94
- Built a stacking model pipeline with LightGBM and CatBoost and optimized via grid search, achieving the best competition performance with a ROC-AUC of 0.96

### A/B Testing with Cookie Cats | [GitHub](#)

- Implemented Cohen's h and Mann Whitney U Test to assess retention differences between player groups

### Network Science Approach on Customer Churn | [GitHub](#)

- Engineered a Stochastic Block Model from 600,000 customers to identify influential groups for marketing strategies

### Customer App Review Analysis | [GitHub](#)

- Excluded over 190,000 stop words and developed TF-IDF to extract the top 20 topics among all customer reviews

## Skills & Interests

**Languages:** Python, SQL, R, VBA, C, Java, JavaScript, HTML/CSS, LaTeX, Racket

**Frameworks & Tools:** PostgreSQL, MySQL, AWS, Git, GitHub, TensorFlow, Excel, Tableau, GGY AXIS

**Expertise:** NLP, Bayesian Inference, Prompt Engineering, Product Sense, A/B Testing, Credit Risk

**Awards & Volunteering:** National Creative Contest Medalist, ABRSM Piano Highest Level, FMSC Volunteer

**Interests:** Entrepreneurship, Solo Traveling, Piano, Soccer, Billiards, Reading Books & Articles