

# Luan Nguyen

778-668-3009 | [luan.nguyen.3@sfu.ca](mailto:luan.nguyen.3@sfu.ca) | [Portfolio](#) | [LinkedIn](#) | [GitHub](#) |

## TECHNICAL SKILLS

**Languages & Tools:** Python, SQL, MongoDB, R, Jupyter, Tableau, Snowflake, dbt, Salesforce, MS SQL Server  
**DS & ML:** EDA, feature engineering, prompt engineering, time-series analysis, regression, classification, clustering, A/B testing  
**Deep Learning:** NLP(tokenization, Word2Vec, sentiment analysis), CNN, RNN, RAG-based LLM, Fine-tuning LLM  
**Automation & Scripting:** Tableau API, Encryption, Task Scheduling, Web Scraping, Data Report

## EXPERIENCE

- Data Scientist Co-op** | ETL, Visualization, NLP, Prediction Model Sep. 2024 – Current  
*Workstream BC, Canada*
- Developed turn-over prediction system using XGBoost and ensemble methods, analyzing applicant applying patterns and key features to identify turn-over candidates, resulting in a improved turn-over rate and reduced hiring cost
  - Applied NLP techniques (Word2Vec, TF-IDF) and clustering algorithms to analyze 10,000+ sales team notes, uncovering key themes in customer downsell that led to a 20% reduction in customer churn through targeted intervention strategies
  - Implemented end-to-end data pipelines from backend data sources to production data warehouse using Hevo, dbt, and Snowflake, reducing data processing time by 40% and improving data quality through optimized SQL queries and automated validation checks and error handling
  - Created interactive Tableau dashboards tracking key business metrics (CLV, churn, product metrics), enabling data-driven decisions that help sale and marketing teams to improve customer acquisition and retention efficiency
- Data Analyst Co-op** | *Databrick, Azure, MS SQL Server, Tableau* Jan. 2024 – May. 2024  
*Vancouver Coastal Health, Medical Quality Practice Dept. BC, Canada*
- Developed ETL pipelines using PyODBC and PySpark to integrate MS SQL Server and Denodo, automating data extraction and transformation for real-time physician performance evaluation system
  - Designed and implemented annual performance metrics dashboards for the Radiology department using Tableau, tracking key indicators such as patient wait times and scan volumes, enabling data-driven resource allocation decisions
  - Automated quarterly performance report generation and distribution for 500+ physicians using Python and Tableau API, reducing manual effort by 80% and improving reporting efficiency
  - Contributed to predictive modeling for patient outcomes (readmission, length of stay, mortality) and conducted A/B testing on quality improvement projects, providing data-driven insights for strategic initiatives
- Personal Banking Associate** Jul. 2022 – Jan. 2023  
*TD Bank BC, Canada*
- Excelled in clear communication by translating complex financial and technical concepts into understandable terms for diverse clients, resulting in 50% higher than expected sale revenue during Q4-2022
  - Developed strong leadership skills by mentoring new team members and coordinating with multiple stakeholders including financial planners, branch managers, and back-office teams to deliver comprehensive financial solutions
  - Demonstrated exceptional multitasking abilities by managing multiple client portfolios while maintaining high service standards and meeting strict regulatory compliance requirements

## HIGHLIGHTED PROJECTS

- S&P500 Analysis** | *NLP, Data Handling, ML, Statistical Analysis, Feature Engineering* | [Github](#) Sep. 2024
- Conducted sentiment analysis on 7000+ financial news articles using TextBlob, Vader, and HuggingFace, linking them with S&P500 stock data to predict next-day prices
  - Built and optimized regression models with Python, achieving a \$12.6 mean prediction difference by integrating sentiment scores and stock parameters (open, close, volume,...)
  - Applied statistical T-tests and data visualization to confirm significant impact of major events on stock prices
  - Addressed data challenges, including weekend gaps and missing sentiment values, through interpolation and feature engineering
- Probabilistic Models** | *TensorFlow, Probabilistic Model, Git* | [Github](#) Aug. 2023
- Performed feature extraction and hypothesis tests using MANOVA, Chi-square test, identifying relevant attributes
  - Conducted regression and time series analysis, investigate trends and patterns
  - Developed binary classification models (KNN, Naïve Bayes, Logis. Regress., SVM, Neural Net.) using TensorFlow
  - Evaluated models using multiple techniques (train-test split, cross-valid., bootstrapping), achieved 87% accuracy
- Other Projects** | *Scikit-Learn, PyTorch, R, HTML/CSS/JS, C++, Java* | [Github](#) 2018 – 2023
- Web scraping, features extraction/analysis, prediction model, report generation, web/game development

## EDUCATION/PROFESSIONAL DEVELOPMENT

- BASc, 4th Year, Computer Science Major, Applied Statistics Minor** Sep. 2025  
*Simon Fraser University Burnaby, BC*
- GPA: 3.5, Dean's Honour Roll 2022-2023
  - Concentration: Machine Learning
  - Relevant coursework: ML, DS, KDD (3rd year); Probabilistic Analysis (3rd year); DSA (3rd year)
  - Data Analytic Certificate: Python/R data wrangling; Tableau data visualization
  - Social Activities: SFU Canadian Cancer Society Chair (2022-2023)