

Steven Lio

Website: stevenlio88.github.io/Portfolio/

Linkedin: linkedin.com/in/steven-lio/

Github: github.com/stevenlio88

Summary

Result-driven professional with 6+ years of experience in providing engineered solution for data workflows, developing tools to extract actionable insights, and support cross-functional teams in strategic decision-making

Qualifications

- **Machine Learning Expertise:** Design efficient Data Pipelines, Features Engineering, and implement models to support data production in Python using NumPy, Pandas, scikit-learn, Scipy
 - **Advanced Analytics and Modeling:** Apply Machine Learning Algorithms for Classification, Clustering, Statistical Modeling, Regression Modeling in R, and A/B testing to provide comprehensive data-driven analyses
 - **Deep Learning Proficiency:** Develop Convolution Neural Networks (CNN) using Pytorch/Tensorflow for Image Segmentation and Classification, optimizing model accuracy and memory usage
 - **Data Processing Skills:** Hands-on experience in designing SQL queries to process, implement algorithms and aggregate large datasets in Database Management Systems such as MS SQL Server, Snowflake, SAS
 - **ETL Process Design:** Experienced in designing efficient ETL processes to handle structured and unstructured data types including CSV, HTML, JSON, XML, audio files, and images
 - **Software Development:** Proficient in backend web applications development with C# in the .NET framework using Visual Studio and experienced in utilizing version control systems such as Git and GitHub along with Jenkins and JIRA to drive production and deliver reliable CI/CD development pipelines
 - **Analytical Acumen:** Provide actionable insights to support research in marketing strategy including Customer Value Prediction and Segmentation, Sentiment Analysis, Market Potential and Shares Estimation, Optimizing Sales Channel and Competitor Analysis
 - **Visualization and Storytelling:** Skilled in creating compelling storytelling and dynamic data visualization tools using Tableau, PowerBI, ArcGIS, and libraries such as Plotly, Dash, matplotlib, ggplot and Shiny
 - **Detail-Oriented and Organized:** Highly organized, with the ability to manage multiple tasks simultaneously to meet ad-hoc requirements and intensive deadlines in a fast-paced working environment
-

Professional Experience

Blueprint Software System

Toronto, ON

Data Scientist

Sep 2022 – Jun 2024

- Optimized algorithm for converting Robotics Process Automation (RPA), achieving an **over 95% conversion rate** through the application of machine learning techniques similar to **Transformer Model**
- Prototyped intelligent **Recommender System** for improving efficiency in customer's RPA processes
- Increased conversion capability on major RPA platforms fourfold and optimized the operational workflow
- Developed tailored backend solutions using **.NET framework** and **C#** within an **Agile environment**
- Refined SQL queries, stored procedures and data storage architecture for scalability and robustness
- Identifying and engineered KPI and produced real-time performance reports that supported CTO in strategic decision-making, highlighting opportunities to improve efficiencies in application development

Quality Assurance Business Analyst

Apr 2021 – Aug 2022

- Developed data integration processes and **constructed probability models** for evaluating data quality
- Designed streamlined data pipeline and feature engineering process to consolidate usage data effectively
- Created testing and validation SQL scripts to analyze the performance of RPA processes conversion

Bell Mobility

Mississauga, ON

Data Scientist

Jan 2020 – Dec 2020

- Implemented **multivariate data models** to **predict customer lifetime value** and devised effective marketing campaign strategies to identify high-value market segments and enhancing customer acquisition quality
- Developed **multi-class classification models** for **churn prediction** and **enriched competitor analysis**
- Boosted customer analytics by **tripling available customer attributes** through engineered features and ML models
- Engineered a scalable process for aggregating Calling Detail Records and defining features to classify customers
- Analyzed sales and spatial data to **estimate market size** and **revenue**, and proposed optimal sales channel strategies

Steven Lio

Website: stevenlio88.github.io/Portfolio/

Linkedin: linkedin.com/in/steven-lio/

Github: github.com/stevenlio88

Bell Mobility

Market Insight Analyst

Mississauga, ON

Aug 2017 – Dec 2019

- Provided actionable insights to drive strategic decisions and business growth for marketing, sales, and product teams
- Created **compelling storytelling** dashboards using real-time data to **deliver comprehensive summaries**
- Leveraged customer activity and sales transaction data to highlight marketing focus to improve campaign efficiency
- Analyzed promotional response, assessed product category elasticity to identify cross-selling opportunities
- Supported Pre & Post campaign analyses, designing customized surveys to gain insights into customer behaviors

Environment and Climate Change Canada

Data QA Analyst

Burlington, ON

May 2016 – Apr 2017

- Designed data QA process and digitized 40+ years of Canadian Great Lakes freshwater quality data
 - Collaborated with project leaders and database managers to implement QA procedures in production
 - Assisted project leaders to identify suitable benchmarks for developing data quality guidelines
 - Identified patterns of anomalies and missing data in order to maintain data integrity
-

Projects

Segmentation of Coral baby

Personal

Toronto, ON

Jan 2023 – Jun 2023

- Developed an Image Segmentation CNN under the U-Net architecture for identifying coral baby from images
- Transferred learning from Resnet50 with early stopping and **achieved 90% IoU scores** on validation images
- Designed image processing pipeline for streamline training data into the model and optimize memory usage
- Image preprocessed to create cleaned masks using OpenCV for the training and validation images

Photo of Photo Detection

Trusting Pixels Inc.

Vancouver, BC

May 2022 – Jun 2022

- Collaborated with a team to build a proof-of-concept ML model for detecting recaptured images
 - Assessed the performance of classification models including SVM, Ridge and Logistic Regression, CNN model and the best model achieved **over 95% Cross-Validation accuracy** score in identifying recaptured images
 - Designed image preprocessed and assessed impact on model performance based on engineered features
 - Built an interactive Dash application in Python with the best model built-in for annotating recaptured images
-

Education

The University of British Columbia

Master of Data Science

Vancouver, BC

Aug 2022

McMaster University

Honors Bachelor of Science - Mathematics and Statistics Co-op

Hamilton, ON

Aug 2016