

# Steven Lio

Website: [stevenlio88.github.io/Portfolio/](https://stevenlio88.github.io/Portfolio/)

Linkedin: [linkedin.com/in/steven-lio/](https://linkedin.com/in/steven-lio/)

Github: [github.com/stevenlio88](https://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, 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 analytics supporting 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, and ggplot
  - **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*

- Enhanced algorithm for converting robotics process automation (RPA), achieving over 95% conversion rate
- Increased conversion capability on major RPA platforms fourfold through enhancements on algorithm and optimizing the operational workflow, including decoding RPA processes in XML, JSON, XAML formats
- Developed tailored backend solutions using .NET framework and C# within an Agile environment
- Optimized SQL queries, stored procedures and data storage architecture for scalability and robustness
- Identifying and engineered key metrics and produced real-time performance reports that supported CTO in strategic decision-making and, 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 for identifying high-value market segments and enhancing customer acquisition quality
- Developed multi-class classification models for churn prediction and improved and enriched competitor analysis
- Enhanced customer analytics by tripling available customer attributes through engineered features and ML models
- Designed process for aggregating calling detail records and defined features to classify customer based on usage
- Analyzed sales and spatial data to estimate market size and revenue, and proposed optimal sales channel strategies

# Steven Lio

Website: [stevenlio88.github.io/Portfolio/](https://stevenlio88.github.io/Portfolio/)

Linkedin: [linkedin.com/in/steven-lio/](https://linkedin.com/in/steven-lio/)

Github: [github.com/stevenlio88](https://github.com/stevenlio88)

---

## Bell Mobility

Mississauga, ON

*Market Insight Analyst*

*Aug 2017 – Dec 2019*

- Analyzed customer activity data to refine marketing focus and improve marketing campaign efficiency
- Leveraged sales transaction data to study promotional response and product category elasticity
- Forecasted sales trends, identified market share gap within specific customer and product segment
- Supported Pre & Post campaign analyzes and designed customized surveys to understand customer behaviors
- Design interactive dashboard with real-time data that provide comprehensive summaries for the analytics team

## Environment and Climate Change Canada

Burlington, ON

*Data QA Analyst*

*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 design and implement QA procedures
  - 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

Toronto, ON

*Personal*

*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 set
- 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

Vancouver, BC

*Trusting Pixels Inc.*

*May 2022 – Jun 2022*

- Collaborated with a team to build a proof-of-concept ML model for detecting recaptured image
  - Assessed the performance of classification models including SVM, Logistic Regression, CNN model and best model achieved over 95% accuracy on validation 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 a dynamic image verification
- 

## Education

### The University of British Columbia

Vancouver, BC

*Master of Data Science*

*Aug 2022*

### McMaster University

Hamilton, ON

*Honors Bachelor of Science - Mathematics and Statistics Co-op*

*Aug 2016*