

AGILAN SIVAKUMARAN

Toronto, ON | (289) 600-0719 | agilan.sivakumaran@gmail.com | www.agilan.online
linkedin.com/in/agilan-sivakumaran | github.com/agilancan

SUMMARY

Applied Data Scientist with a strong foundation in machine learning, generative AI, and business analytics, experienced in translating consumer and operational data into actionable insights. Proven ability to develop end-to-end ML solutions and collaborate with stakeholders to drive data-informed decisions across marketing, risk, and product domains in business-focused environments.

SKILLS

- **Programming & ML:** Python, SQL, Pandas, NumPy, scikit-learn, XGBoost
- **Modeling & Methods:** Classification, forecasting, predictive modeling, statistical experimentation, hypothesis testing, A/B testing, customer segmentation, risk modeling, model evaluation, SHAP
- **AI Tools & LLMs:** LangChain, OpenAI, Prompt Engineering, Transformers, Hugging Face, RAG
- **Data & Databases:** PostgreSQL, BigQuery, relational databases, data processing & validation
- **Cloud & Tools:** AWS, Docker, CI/CD, Tableau, Airflow, ETL, FastAPI

AI & DATA SCIENCE EXPERIENCE – More at github.com/agilancan

St. Michael's Hospital (University of Toronto, CAMH) [[Repository](#)] **Oct 2025 – Present**

AI Researcher (Volunteer) | Toronto, ON

- Developed an interpretable, survey-weighted ML model to identify past-year Major Depressive Episode (MDE) using NSDUH population data (n=45,000+), optimized for high-recall screening and risk stratification.
- Framed the problem as a high-recall classification task, prioritizing minimizing false negatives in real-world screening and triage contexts.
- Trained and evaluated logistic regression, Random Forest, and XGBoost models, achieving AUC ~0.93–0.94 and average precision ~0.61 on held-out data.
- Optimized models for screening use cases, achieving sensitivity ~0.91–0.92 and negative predictive value over 0.99 at selected thresholds.
- Performed model explainability, calibration, and subgroup analysis (SHAP, PDPs, reliability curves), identifying psychological distress and functional disability as dominant predictors.
- Developed and evaluated a hierarchical multi-agent LLM framework for suicide-risk classification using LangGraph and structured evaluation pipelines.
- Co-authored manuscript submitted to the Journal of the American Medical Informatics Association (JAMIA), currently under peer review.

Fake News Detection [[GitHub](#)] **Jul 2025 – Aug 2025**

- Fine-tuned RoBERTa for text classification and integrated a RAG pipeline with FAISS, structuring retrieval and inference workflows for context-aware, explainable fake news verification, achieving over 90% correct classifications in prototype evaluation on a predefined set of articles.

End-to-End ML Pipeline for Bankruptcy Prediction [\[GitHub\]](#) Jul 2025 – Aug 2025

- Developed an end-to-end ML pipeline for bankruptcy prediction using Logistic Regression, Random Forest, and XGBoost, achieving ~0.94 ROC-AUC on test data and incorporating SHAP-based explainability and PSI drift monitoring.

SafeRoomAI: Real-Time Security AI Platform [\[GitHub\]](#) May 2025 – Jul 2025

- Developed a real-time video intelligence system using OpenCV and YOLOv8, designing streaming inference workflows for continuous event detection and anomaly monitoring. Improved system reliability by handling edge cases (false positives, motion noise) in live video pipelines.
- Built backend APIs with FastAPI and PostgreSQL, integrating AWS RDS/S3 for scalable storage.

Greenearth Connect: Environmental Health AI Platform [\[GitHub\]](#) Jan 2025 – Apr 2025

- Developed AQI forecasting engine using XGBoost and LSTM (RMSE ↓ from 9.96 to 7.03).
- Designed Random Forest activity recommender based on health and environmental conditions, and built HuggingFace NLP pipelines for news summarization, overcoming limited datasets and integrating multi-pipeline AI models into a unified platform.

Pharmacy.ca May 2024 – Sep 2024

Director of Digital Marketing | Toronto, ON

- Designed and implemented end-to-end analytics pipelines using SQL, GA4, and HubSpot, incorporating data validation, quality checks, automated reporting, and compliance-aware practices.
- Led statistically grounded experimentation using A/B testing, hypothesis testing, funnel, cohort, and customer journey analysis to improve conversion efficiency, retention, and acquisition cost.
- Leveraged AI analytics & CRM segmentation to optimize targeting, messaging, and campaigns.
- Built KPI dashboards and automated reporting workflows to validate data quality and track customer behavior and marketing performance, enabling faster, data-informed decision-making.
- Translated analytical findings into actionable insights for non-technical stakeholders to support business and marketing decisions.

Doohickey Inc. Jul 2017 – Feb 2025

Senior Digital Strategist | Toronto, ON

- Led agile development of WordPress, React, and Unity projects, including ML-driven features and UX/UI, cutting time-to-market 30%.
- Leveraged GA4, SQL, data analysis, and A/B testing to optimize funnels, boosting retention 22% and conversions 15%.

EDUCATION & CERTIFICATIONS

Durham College Jan 2025 – Aug 2025

- Graduate Certificate in Artificial Intelligence Analysis, Design, and Implementation (Applied analytics, machine learning, and data science for business decision-making)

Trent University Sep 2007 – Apr 2012

- Honours Bachelor of Business Administration (Minor in Information Systems)

Project Management Association of Canada Sep 2012 – Oct 2012

- Certified Agile Project Manager (Cert.APM)