Arnold Kalmbach

Senior Data Scientist.
Expert in machine learning, computer vision, and signal processing algorithms.
Experienced at all levels of the data science dev-path, from exploratory analysis to deplyment. Specialist in the scientific python stack.

Email: akalmbach@gmail.com | Web: akalmbach.gitlab.io | Phone: (250)-885-8155

Languages & Skills

Scientific Python	pytorch, numpy, scipy, scikit-learn, pandas / polars
Scalable Inference	ONNX, TorchScript, Triton Server, Argo
GIS Software	geopandas, gdal, QGIS, PostGIS
Audio + Linux Audio	JUCE framework, torchaudio, librosa, jack, Bitwig Studio
Spoken Languages	First language, English. Basic proficiency: French and Spanish.

Professional Experience

Senior Data Scientist — Kinsol Research Inc. Jul 2022 — Present

Team lead for multiple client projects. Worked with clients and Kinsol data science teams to align development path and business goals in weekly update meetings.

Managed teams of 2-3 junior developers, guiding their technical development and coaching on client interactions.

Example Projects:

Oversaw the technical plan and helped implement a project to build an inventory of all electrical substations in the USA, using a combination of existing databases and an ML model applied to satellite data. Scaled the ML pipeline to perform inference on high-res satellite data from millions of candidate points, covering 100 metro areas in the USA.

Oversaw the technical plan and helped implement a project to detect clinically significant events in microphone and PPG data on a medical wearable device. Kept data science (Kinsol), and clinical, DSP (client) teams in sync.

Developed experimental features for a domain registration and site generation project using NLP techniques. Responded to evolving business constraints by adapting the project's scope and priorities, resulting in a 3x increase in the contract.

Data Scientist — Kinsol Research Inc. May 2017 — Jul 2022

Implemented machine learning, computer vision, and signal processing algorithms for a broad range of client projects.

Provided value to clients by filling gaps from science to production; As scientist performing literature review and implementing rapid prototypes; As machine learning expert generalizing existing hand-tuned models based on data; As engineer optimizing and productionizing models too dense for engineering teams.

Communicated with clients several times per week, delivering quantitative progress reports and coordinating development plans to meet business milestones.

Example Projects:

Implemented specialized signal processing algorithms for motion tolerant heartbeat detection and significantly improved accuracy while also increasing yield more than 3x.

Productionized an existing flood model, reducing runtime 2x from hydrologists original version and facilitating scaling of the model from city to national level.

Applied black-box optimization techniques to tune parameters of real-time vocal pitch detection algorithm, leading to best-in-class pitch detection accuracy.

Implemented several variants of real-time object detection using convolutional neural nets, each tailored to minimize data collection effort for their particular application.

Teaching Assistant, Introduction to Robotics and Intelligent Systems - Jan —
McGill University, School of Computer Science, COMP 417 May 2017

Automated Video Analysis Intern — Ocean Networks Canada Jun — Nov 2015

Implemented an algorithm to automatically produce sea-floor classification maps from dive-log videos (Kalmbach, WACV 2016 in Publications below).
Implemented algorithms for automated quality control of live deep-sea video, and deployed these tools in the automated pre-processing workflow for live video streams from a network of underwater observatories.

Test Technician & Autonomy Software Co-Op — Clearpath Robotics Inc. Jan — May 2015 (Contract)

Developed proof of concept Gazebo (robot simulation) test suite to be deployed in continuous integration for Husky robot autonomy stack.
Automated a CI test suite for robot autonomy behaviors such as localization, mapping, and path following simulated in a variety of environments.
Increased CI testing led to improved reliability of the Husky simulation and robot autonomy packages. May — Aug 2014 (Co-Op)
Deployed ROS tools for a robot with an integrated industrial arm, including arm control through the CAN protocol and integration with high level planning through MoveIt, MoveBase and Gazebo.

Education

McGill University — MSc. Computer Science Jan 2016 — Sep 2018

Thesis - *Unsupervised Learning of Interpretable Models for Sparse, Smooth Data* (Sep. 2018).
Contributed a variety of Bayesian non-parametric methods to learn temporally varying semantic maps.
Significant work on machine learning, computer vision, and algorithmic robotics problems as well as robot deployments and field experiments.
Coursework focus on the fundamentals of machine learning, computer vision, digital signal processing, and optimization.
Supervisor Greg Dudek, Mobile Robotics Lab
Achieved 3.88/4.0 GPA

McGill University — BSc. Comp. Sci. Sep 2009 — May 2013

Emphasis on mathematical foundations of Comp. Sci.

Began involvement with robotics research, Mobile Robotics Lab (2012).

Honors and Awards

2nd place, Student Poster Competition, MTS Oceans 2017
2017 ICRA IEEE RAS Travel Grant Recipient
Guest Student Appointment, Summer 2016, Woods Hole Oceanographic Institution
3rd place, McGill CS Undergraduate Research Symposium 2012

Selected Publications & Research Experience

Learning Seasonal Phytoplankton Communities with Topic Models. **MTS OCEANS 2017**
A Kalmbach, H M Sosik, G Dudek, and Y Girdhar

An unsupervised method to learn probabilistic associations between species (communities) with a topic model, incorporating priors for sparseness and spatio-temporal smoothness.

The method finds an interpretable community model by ensuring community distributions can be predicted from simple environment data.

Phytoplankton Hotspot Prediction with an Unsupervised Spatial Community Model. **A Kalmbach**, Y Girdhar, H M Sosik and G Dudek. **ICRA 2017**

A method to predict the presence of sparse spatio-temporal phenomena by their denser associates.

The method predicts the 'hotspot locations' of phytoplankton species by the abundance of other species which may be easier to observe.

Learning Deep-Sea Substrate Types with Visual Topic Models. **WACV 2016**

A Kalmbach, M Hoeberechts and A Branzan Albu

Unsupervised learning of distribution-of-visual-features representations for sea-floor types from ROV dive logs.

Sea-floor classifications were used in further bio-geophysical studies of the study sight.

Unsupervised environment recognition and modeling using sound sensing. **ICRA 2013**

***A Kalmbach**, Y Girdhar, and G Dudek*

Unsupervised learning of distribution-of-audio-features representations for ambient noise from walking through urban environments.
Similarity in the learned topic space indicates intuitively similar environments and can be used for soft loop-closure.