Kyle A. Chezik

(206) 832-5883 | Seattle, Washington | kchezik@gmail.com
github.com/kchezik | linkedin.com/in/kyle-chezik | kchezik.github.io

Technical skills

Machine Learning	**Deep Learning**: *Convolutional-NN* **Clustering**: *K-means, Mixture Models* **Classification**: *Neural Networks, Decision Trees, Random Forest, Naïve Bayes*	Machine Learning	**Feature Engineering**: *PCA, AIC, BIC, Spike-and-Slab, L1/L2 Regularization* **Regression**: *Linear, Logistic, Ridge, Lasso*
Analysis	**Time Series**: *ARIMA, Bayesian Structural & Hidden Markov Models* **Simulation**: *Parametric Bootstrapping* **A/B Testing**: χ^2, *ANOVA*	Tools	**Python**: *pandas, NumPy, scikit-learn* **R**: *tidyverse, R-Studio, caret, lme4, rstan* **SQL**, Unix, Git, HTML, CSS, Regex, Stan **VIS**: *ggplot2, Matplotlib, Seaborn, R-Shiny*

Professional experience

2019 **Insight Data Science Fellow**, Seattle WA, USA

- Developed *Native Garden Builder*, an **interactive recommender** that leverages **deep learning** in **computer vision** to provide conservation minded gardeners aesthetically desirable native plants.
- Decomposed plant images into **512 features** using the **convolutional neural network** ResNet18 within the **PyTorch** framework and compared images using cosine similarity.
- Scraped **1400**$^+$ web images and collated meta-data from multiple databases using **BeautifulSoup** in **Python**.

2013-19 **Doctoral Research Assistant**, Simon Fraser University, Burnaby BC, Canada

- Automated error identification in temperature time series using a novel **Bayesian Hidden Markov model** that demonstrated **84% accuracy** across **1 Million**$^+$ records, significantly **reducing human work hours** and earning the SFU KEY Big Data Graduate Scholarship.
- Identified novel river-network properties using **linear regression**, overcoming limited replication by simulating **ARIMA** processes using **parametric bootstrapping**, leading to **98% certainty**.
- Combined time series and spatial network models to assess salmon migratory heat risk, **reducing regulatory uncertainty** of a multi-million dollar fishery.
- **Feature engineered** gridded data using GIS tools (e.g.,*GDAL, OSGEO*), leveraging **parallel computing** within Python (*WhiteboxGAT*) and on the command line (*GNU parallel*).

2016 **Data Engineer Contractor**, ESSA Technologies Ltd., Vancouver BC, Canada

- Constructed and quality controlled a relational stream temperature database of ~**4 million records**, foundational to a province wide river network model for British Columbia Canada.
- Collated data from multiple data-custodians and packaged functions in R for **end-to-end reproducibility** of QAQC and continued maintenance.
- Improved model accuracy through feature engineering of river flow, temperature and landscape data.

2009-11 **Medical Device Reporting Specialist (MDR)**, Medtronic Inc. Mounds View MN, USA

- Monitored clinical trials and addressed challenges identified by principle investigators at multiple hospitals while maintaining **experimental design** efficacy.
- Systematically recorded and reported medical device failures to the FDA to improve detection of device related negative health outcomes.

Education

2019 **Ph.D.** Biological Sciences, Simon Fraser University, Burnaby BC, Canada

2013 **M.Sc.** Conservation Biology, University of Minnesota, St. Paul MN, USA

2009 **B.A.** Biology, St. Olaf College, Northfield MN, USA