John Burt

johnmburt65@gmail.com
206-214-8849
Portland OR 97206
johnmburt.github.io

Summary

- Data Scientist / Analyst with hands-on experience working with healthcare data and large production databases. Background in scientific research, product development. Years of coding, problem solving and solution building.
- Experienced with data mining, munging and cleaning, statistics, EDA, experimental design, validation testing, machine learning, deep learning, and analytics.
- Great presentation and visualization skills: have presented seminars and workshops, written technical documentation, conducted hands-on training for my own products.

Keywords

Machine Learning: classification, regression, feature engineering, deep learning (CNN, RNN), natural language processing (NLP), AWS SageMaker, random forest, decision trees, SVM, dimensionality reduction, statistical modeling, A/B testing, simulation models.

Data: web scraping, data wrangling, exploratory data analysis (EDA), analytics, feature engineering, visualization, presentation. SQL coding for production databases.

Software and Programming Languages: Python (Data Science stack), SQL, Golang, C, Power BI, Excel

Statistics: time series analysis and forecasting, regression models, classifier models, hypothesis testing, experimental design.

Education / Accreditation

- **Certification, Springboard Data Science Career Bootcamp** - 2020
  The Springboard Data Science Career program involves over 600 hours of coursework covering the entire Data Scientist's toolbox: data collection and preparation, EDA and statistics, Python programming, most popular ML methods, SQL coding, and two substantial capstone projects.
- **PhD Psychology 1999**, University of Washington - 1989 – 1999
- **BS Computer Science / Biology** 1988, The Evergreen State College - 1988 – 1989
- **Marine Biology program**, University of Oregon - 1987 – 1988
- **Biology / Computer Science**, Southern Oregon University - 1984 – 1987

Experience

6 Degrees Healthcare   **Jul 2020 – April 2021**
Junior Data Analyst
- Software engineering and ETL in Golang and SQL for website products and large PostgreSQL / MongoDB databases. ETL, data cleaning, data munging and data research in SQL, Golang, Python and Excel. Machine Learning modeling. Power BI report design.

Portland Data Science Group   **Jul 2017 - Jul 2020**
Head Organizer
- Managing the ongoing Applied Data Science meetup series. Interface with sponsors, collect and prepare datasets, Data Science presentations, help participants.
- Data mining & munging, EDA, feature engineering, machine learning, NLP, regression, classification, data visualization, deep learning, presentations, teaching, Python, SQL

Encounternet LLC   **Apr 2012 - Sep 2017**
Owner – CEO
- Co-founded company Encounternet LLC to design, manufacture and sell wireless sensing tags for animal social network research. I fulfilled all roles: sales, hardware circuit design, software and microcontroller firmware coding, testing, manufacturing, customer training and deployment.
- Python Windows GUI, C microcontroller firmware dev, circuit design and manufacturing, sales, support, system deployment, presentations.

University of Washington   **Aug 2008 - Jul 2012**
Assistant Professor / Project Lead, Electrical Engineering Department, Wireless Sensing Lab
- Funded by NSF grant to develop my idea for Encounternet, a wireless smart-tag system for monitoring animal social behavior. Worked with electrical engineers to design and prototype our system, which included custom smart tags and wireless base stations.
- Grant writing, talks, supervising graduate students, circuit and hardware design and prototyping, coding in Python and C.

University of Washington   **Nov 2003 - Aug 2008**
Research Scientist / Engineer, Psychology Department
- Bird song learning research. Developed Matlab-based automated audio stimulus and recording system to study bird vocal learning. Co-wrote NSF grant the project was funded on.
- Experimental design, management lead role at lab, scientific articles and grant proposals, taught upper division Animal Communication course, Matlab, C++.

Cornell University   **Sep 1999 - Nov 2003**
Post-doctoral Associate, Bioacoustics Research Program, Lab of Ornithology
- Developed and deployed a 16 channel radio-microphone array hardware and acquisition software system for recording bird song in Costa Rica. Developed automated analysis software to locate sounds in array recordings. Developed free windows app for sound analysis (SyrinxPC, with 1000s of downloads),
- Field research, managed field research operations in Costa Rica with teams of up to 10 biologists, hardware system design, and software development, C++ (Windows apps), Matlab.

Selected projects

Hosted at johnmburt.github.io

A/B testing simulation - Jun 2020
- I created an A/B testing simulation model that emulates users interacting and responding to a web site. A tester can run experiments on the sim-user population to determine their preferences.

Training a SageMaker Deep Learning model and deploying it as a microservice - May 2020
- Trained an AWS SageMaker image classifier (ResNet) to identify 120 dog breeds from images, then deployed the model as a microservice via the AWS API Gateway

Building a whale song detector based on a deep learning model - Feb 2020
- I developed a classifier to detect humpback whale songs in a year long hydrophone recording. Built a CNN based classifier model, trained it, ran validation testing, and then ran the model on a year's worth of hydrophone recordings. I then analyzed the resulting detection data.

Boardgame recommender system model and web app - Oct 2019 - Nov 2019
- Scraped 15M boardgame ratings, conducted EDA and developed models to recommend boardgames based on input of several games a user likes. Developed a fully functional demo recommender web app (Bokeh server), deployed to Heroku.

Reddit toxic comment detector - Jul 2019 - Sep 2019
- Conducted EDA and developed a classifier model using NLP to detect toxic comments on Reddit. Scraped Reddit comments, developed a toxicity label, and trained and cross-validation tested several classifier models (Multinomial Naive Bayes, Random Forest, XGBoost, RNN).

Music Genre Classification - Aug 2019 - Sep 2019
- Built a classifier to identify the music genre of a song clip. Conducted EDA to identify useful features, generated features, trained and cross-validated an XGBoost classifier, examined feature importances.