JOHN BURT
DATA SCIENTIST

Summary

Data Scientist with behavioral and bioacoustic research science and coding background. Experienced with machine learning, deep learning and analytics. Other areas of interest include natural language processing, behavioral analysis and sound analysis.

Recent Employment

Portland Data Science Group
Head Organizer — 2017 to Current

Co-organizer, elected Head Organizer in 2017. Managing the ongoing Applied Data Science meetup series. Secure meeting spaces, interface with sponsors, help prepare datasets, Data Science presentations, help participants.

- Data mining & munging, EDA, feature engineering, ML, NLP, regression, data visualization, deep learning, visualization, presentations, teaching, Python, SQL.

Encounternet LLC
Owner - CEO — 2012 to 2016

Co-founded company to design, manufacture and sell wireless sensing tags for animal research.

- Python Windows GUI, C microcontroller firmware dev, circuit design and manufacturing, sales, support, system deployment, presentations.

University of Washington
Research Scientist / Project Lead — Electrical Engineering Department, Wireless Sensing Lab, 2008 to 2012

Funded by NSF grant to develop my idea for Encounternet, a wireless smart-tag system for monitoring animal social behavior.

- Grant writing, talks, supervising graduate students, electronic and hardware design and prototyping, software dev, Python, C

University of Washington
Research Scientist / Engineer — Psychology Department, 2003 to 2008

Bird song learning research. Developed Matlab-based automated song stimulus and recording system to study captive songbird song 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
Post-doctoral Associate — Bioacoustics Research Program, Lab of Ornithology, 1999 to 2003

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 operations in Costa Rica, hardware system design, and software dev, C++ (Windows apps), Matlab.

Selected Projects

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

Oct. 2019 to Nov. 2019 — **Boardgame recommender system model and web app**
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.

Apr. 2018 to Mar. 2018 — **NOAA storm event visualization**
Interactive 4D geographic Bokeh visualization of severe weather events over time.

July 2019 to Sept. 2019 — **Reddit troll detector**
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).

July 2018 to Aug. 2018 — **ODOT crash statistics**
EDA and interactive Google map showing fatal crash density in the Portland OR area.

Contact

johnmburt65@gmail.com
206-214-8849
6022 SE Carlton St. Portland OR 97206
johnmburt

Education

University of Washington — 1990 to 1998
PhD Psychology 1999

The Evergreen State College — 1988 to 1988
BS Computer Science / Biology 1988

University of Oregon — 1987 to 1987
Marine Biology program

Southern Oregon University — 1984 to 1986
Major in Biology, minor in Computer Science

Skills

MACHINE LEARNING
classification
regression
clustering
feature engineering
data munging
neural networks (CNN, RNN)
natural language processing (NLP)
audio classification

DATA
web scraping
data munging
analytics
feature engineering
visualization
presentation

SOFTWARE AND PROGRAMMING LANGUAGES
Python (scikit-learn, numpy, scipy, pandas, gensim, nltk, hyperopt, keras)
SQL
C
C++
Matlab
Excel
Github
Jupyter Notebooks
embedded design

STATISTICAL METHODS
time series
regression models
hypothesis testing
multi-dimensional scaling (MDS)
dimensionality reduction (SVD, PCA)
matrix factorization
signal processing