Andrew Collin Stewart

San Francisco · California

✉ andrew.c.stewart@gmail.com   📞 +202 505 3633   ⌨ andrewcstewart   🌐 andrewcstewart.com

SUMMARY

Data scientist with a strong background in applied research and development in big data, machine learning, and statistics. Passionate about improving business processes through sound experimental design and analysis.

PROFESSIONAL EXPERIENCE

Twist Bioscience, Scientist (Bioinformatics)   2015-2018

Developed predictive models to optimize manufacturing processes involving cell colony picking and misassembly. Desgin and analysis of experiments for improving chemical manufacturing processes. Architect behind data analytics culture centered around data warehouse and ETL processes.

Veracyte, Scientist (R&D Data Analysis Team)   2014-2015

Developed high dimensional machine learning models for classification in cancer diagnostics using a combination of supervised and semisupervised techinques.

Smithsonian Institution, Data Science Specialist (Office of the CIO)   2013-2014

Provided analytical support on research projects in population genetics, phylogenmoics, and metagenomics. Developed ensemble machine learning approach for classification of signatures of natural selection in genetic populations using support vector machines.

Keygene, Bioinformatician (Applied Systems Biology)   2009-2012

Performed data mining and data visualization of gene regulatory networks using gene weighted correlation network analysis, R and graph visualizations. Performed exploratory data mining on hyperspectral imaging of field crop plots using dimensionality reduction techniques and 3D data visualization. Performed modeling and simulation of plant architecture using Lindenmayer systems.

Navy Medical Research Center, Research Assistant (Biodefense Research Directorate)   2006-2009

Applied Hidden Markov Models, clustering algorithms towards rapid identification of pathogens from genomic sequencing data. Designed and built data analytics system for collection, processing and analysis of microbial genomic data using perl, R, SQL, and HPC systems.

EDUCATION EXPERIENCE

Johns Hopkins University, M.S. Computer Science   2015

University of Maryland College Park, B.S. Biology   2006

TECHNICAL SKILLS

Expertise	Data Analysis, Analytics & Dashboards, Predictive Modeling, Research & Development, ETL, Big Data Queries, Data Mining, Machine Learning, AI
Programming	R, Python, Rust, C/C++, Java, Docker , AWS
Data Analysis	tidyverse, ggplot2, pandas, numpy, scipy, scikit-learn, Shiny, Tableau, Spark
Machine Learning	classification, regression, clustering, feature engineering

Statistics	bayesian inference, multilevel modeling, hypothesis testing, power analysis, dimensionality reduction

TEACHING EXPERIENCE

Johns Hopkins University, Graduate Teaching Assistant   2016-2017

Spring 2017: 605.448.81 - Data Science
Spring 2017: 605.445.81 - Artificial Intelligence
Fall 2016: 605.448.81 - Data Science
Fall 2016: 605.448.82 - Data Science

VOLUNTEER EXPERIENCE

Voteocracy, Data Science Volunteer   2016

Applied network analysis techniques to identify marketing targets from Twitter data.

Bayes Impact, Data Science Intern   2015

Developed multiclass predictive model of criminal activity based on geographic and census data.

PUBLICATIONS

Chen, Peter E., Christopher Cook, Andrew C. Stewart, Niranjan Nagarajan, Dan D. Sommer, Mihai Pop, Brendan Thomason, et al. 2010. "Genomic Characterization of the Yersinia Genus." *Genome Biology* 11 (January): R1. doi:10.1186/gb-2010-11-1-r1.

Chen, Peter E., Kristin M. Willner, Amy Butani, Shakia Dorsey, Matroner George, Andrew Stewart, Shannon M. Lentz, et al. 2010. "Rapid Identification of Genetic Modifications in Bacillus Anthracis Using Whole Genome Draft Sequences Generated by 454 Pyrosequencing." *PLOS ONE* 5 (8): e12397. doi:10.1371/journal.pone.0012397.

"Meetings - Paper Digital Library." 2017. Accessed December 6. https://dl.sciencesocieties.org/publications/meetings/2011am/8617/67897.

Mukhopadhyay, Sanghamitra, Arya Akmal, Andrew C. Stewart, Ru-ching Hsia, and Timothy D. Read. 2009. "Identification of Bacillus Anthracis Spore Component Antigens Conserved Across Diverse Bacillus Cereus Sensu Lato Strains." *Molecular & Cellular Proteomics* 8 (6): 1174–91. doi:10.1074/mcp.M800403-MCP200.

Pagan, Moraima, Richard T. Kloos, Chu-Fang Lin, Kevin J. Travers, Hajime Matsuzaki, Ed Y. Tom, Su Yeon Kim, et al. 2016. "The Diagnostic Application of RNA Sequencing in Patients with Thyroid Cancer: An Analysis of 851 Variants and 133 Fusions in 524 Genes." *BMC Bioinformatics* 17 (1): S6. doi:10.1186/s12859-015-0849-9.

Stewart, Andrew C., Brian Osborne, and Timothy D. Read. 2009. "DIYA: A Bacterial Annotation Pipeline for Any Genomics Lab." *Bioinformatics* 25 (7): 962–63. doi:10.1093/bioinformatics/btp097.

Zwick, Michael E., Sandeep J. Joseph, Xavier Didelot, Peter E. Chen, Kimberly A. Bishop-Lilly, Andrew C. Stewart, Kristin Willner, et al. 2012. "Genomic Characterization of the Bacillus Cereus Sensu Lato Species: Backdrop to the Evolution of Bacillus Anthracis." *Genome Research* 22 (8): 1512–24. doi:10.1101/gr.134437.111.

Zwick, Michael E., Maureen P. Kiley, Andrew C. Stewart, Alfred Mateczun, and Timothy D. Read. 2008. "Genotyping of Bacillus Cereus Strains by Microarray-Based Resequencing." *PLOS ONE* 3 (7): e2513. doi:10.1371/journal.pone.0002513.