Andrea Osika

602-214-8360 | andiosika@gmail.com | GitHub | Blog | Linked-In

Operations Analyst / Data Scientist with a background in both the private and non-profit sectors is passionate about wellness, data, and helping others and career-long dedication to process development/refinement and continuous learning. Experienced in back-end engineering (Python, SQL), research, KPI Development, A/B testing, statistical analysis, machine learning (sci-kit-learn, Pandas/NumPy), and deep learning (Keras). Highly skilled in synthesizing ambiguous or complex information to engage stakeholders in productive dialogue that drives sound business strategy and outcomes.

TECHNICAL PROJECTS

Hypothesis Testing, A/B Testing - GitHub
Drew upon SQL and logistic regression techniques to derive business intelligence
+ Queried SQL database to obtain relevant data for each test
+ Answered regression and classification questions of commercial interest using parametric and non-parametric testing.

Hyper-Imbalanced Classification and Analysis on large dataset - GitHub
Random Forest Classification to derive top ten feature importances for contracting COVID-19
+ Utilized open-source dataset containing 619K samples with 43 features with .13% reflecting the target class.
+ Achieved overall **accuracy average of 97% and weighted recall of 97%**

Deep Learning using NLP using Recurrent Neural Networks for multi-label classification - GitHub
Use of NLP to identify toxic or inappropriate language in an online open forum discussion
+ Trained a wide range of shallow- and deep-learning models, relating exploratory statistics to resulting feature importance
+ Use of Focal Loss to address classes that contained baseline rates of >10%, to **realize an increase of 29% recall**

Support Vector Classification & Clustering/ Segmentation using LDA - GitHub
Built an NLP app for word-prediction of app-reviews (85% & 92% recall, 41% & 31% baseline)
+ Used NLP and linear Support Vector Classification for classifying positive and negative sentiment from app reviews
+ Optimized subsets using Latent Dirichlet Allocation (LDA) to inform best practices and identify growth opportunity

TECHNICAL SKILLS

Data/Statistical Analysis, Regression, and Logistic Modeling, A/B Hypothesis Testing, Python, Pandas, NumPy, SkLearn, Keras, Random Forest, SQL, Data Visualization(eg.Matplotlib, Seaborn) Project Management, Presentation, Collaboration MS Office: Excel – samples upon request

EMPLOYMENT HISTORY

Cisco Systems, Corporate Social Responsibility, Phoenix, AZ
Program Specialist, Education Non-Profit Investments, Jan 2011 - Present
+ Grantmaking and assessment specializing in analysis, KPI/Metric development, research, reporting, and presentation skills using quarterly and historical data to inform portfolio performance and investment strategies to cross-functional teams. Working with non-profits for capacity building solutions in education technology. (portfolio of $2.6M)

Aries Technology, Inc, Tempe, AZ
Director of Operations, Contractor, Consultant, Jan 2001 - Jul 2017
+ Managed projects, contracts, SOW, RFP response, billing, accounts payable, HR support, vendor relations for 350 head workforce and $10M annual budget
+ Strengthened client relationships through clearly articulated, evolving advice, backed by data and regular reports

Cisco Systems, Phoenix, AZ
Project Manager, Networking Academy Jan 2001 - Jun 2002 Operations and budget management directly reporting to the Director and Co-Founder in infant stage to scale to serve over 7million students globally

EDUCATION

Flatiron School, New York, NY - 2018
+ Data Science Immersive Program

Colorado State University, Ft. Collins, CO - 1995
+ Bachelor of Arts in Soc. Science (Psychology, Sociology), Minor Spanish