Daniel Ryan Furman

(310) 993 9757 • daniel_furman@berkeley.edu • github.com/daniel-furman

PROJECTS

Multi-Platform Mixing for Social Media NLP - Python (NLTK, Hugging Face). [ACL22, In Review]
- Developed a data-centric domain adoption method that improves the portability of NLP models by mixing platforms at training (Instagram & Twitter captions), which significantly boosted F1 performance by 17+% on average for celebrity gender profiling & authorship attribution tasks (SVMs and LRs) under domain shift (Facebook caption test-set).
- Exploration of over 43,000 celebrity Twitter, Instagram & Facebook post linguistics (Rshiny). [DataViz]

SurfCrowds.ai - Python (NumPy, Pandas, Scikit-learn), AWS (Lambda, S3), Dash, Heroku. [Web-App]
- Built a crowd size predictor for surfers in Malibu, CA with Random Forest & Gradient Boosting ensemble regressors deployed as a serverless AWS Lambda function. Embedded models within a Dash web-application that enables my hometown surf community to explore new types of informative AI forecasts.

Kaggle BirdCLEF21 Competition - Python (PyTorch, Pandas, Scikit-Learn), Neptune.ai. [Blog]
- Won solo bronze out of 816 teams for Sound Event Detection with blends of multi-label, split-attention audio RNNs & CatBoost metadata classifiers with non-competition geospatial climate features (397 species classes, ~63k obs.), overcoming a large train to test domain shift & weak labeling (F1=0.61).