Kyle Willett

ML/Applied Scientist

Seattle, WA, USA
651.278.2643
willettk@gmail.com
willettk.github.io
Github: @willettk

Experience

2016-present **Applied scientist**, *Amazon*, Seattle, WA

- Developed and launched ML solutions for Amazon's third-party Pricing and Fees Science teams. Applications include pricing estimates, search ranking, multi-class categorization for billions of items, causal inference, LLM agentic tools, and risk-aware prediction models.
- Responsibilities include code review and production deployment, statistical analysis, A/B testing, and integration at industry-scale.
- Supervisor and manager for three graduate student science interns at Amazon. Team lead for causal inference group for 2.5 years.

2016 **Fellow**, *Insight Data Science*, Palo Alto, CA (remote)

- Developed an application that used image processing and real-time ML to identify and recognize humans in online videos. Packaged as a standalone web-service backed by AWS and delivered as a consulting project for **Muse AI**.

2015-2016 **Lead data scientist**, *Galaxy Zoo*, Minneapolis, MN/Oxford, UK

- Oversaw day-to-day maintenance, backend processing, and dataset generation for Galaxy Zoo, a citizen science project with $> 250,000$ participants worldwide.
- Created online data archive for both internal and public access to project publications and scientific results.
- Designed and administered a Kaggle competition (with 326 teams participating) to replicate human classification of galaxy images using convolutional neural networks.
- Lead author for two major catalog releases. Author/co-author on 48 project publications.

2011-2016 **Research associate**, *University of Minnesota*, Minneapolis, MN

- Researched structure and morphology of galaxies, including the relative influences of galactic bars, disks, and mergers on supermassive black holes and star formation.
- Developed new algorithms to measure and correct the effects of image quality on volunteer classification accuracy.

2006-2011 **Graduate research assistant**, *University of Colorado*, Boulder, CO

- Ph.D. thesis research on hydroxyl megamasers in external galaxies, including discovery of two new megamasers and numerical modeling of physical conditions at maser sites.
- Extensive observing experience with radio, infrared, and optical facilities, including the Very Large Array, Green Bank Telescope, and Spitzer Space Telescope.

Education

2005–2011 **Ph.D.**, *University of Colorado*, Boulder, CO
Astrophysical and Planetary Sciences

M.S., *University of Colorado*, Boulder, CO
Astrophysical, Planetary, and Atmospheric Sciences

2001–2005 **B.A.**, *Carleton College*, Northfield, MN
Physics (*magna cum laude*)

Skills

Languages **Python, SQL, PySpark, Stata, MongoDB, IDL, LaTeX, Mathematica**

Tools AWS, numpy, pandas, scikit-learn, matplotlib, Jupyter, Git, MySQL, scikit-image, OpenCV, Spark, Docker, HTML, Flask