

# Martijn de Vries, Ph.D.

Data Scientist, Researcher

San Francisco Bay Area, CA | 415 533 5539 | [martijndevries91@gmail.com](mailto:martijndevries91@gmail.com) | [GitHub](#) | [Portfolio](#) | [LinkedIn](#)

## SUMMARY

---

As a committed data scientist with a background in astrophysics, I bring an academic and analytical perspective to the practice of modeling and interpreting extensive data sets. I have an innate curiosity and drive to find unique solutions to complex problems, and to translate my analyses into actionable insights that can make meaningful change and impact in the world.

## SKILLS

---

**CODING:** Python, Numpy, Scipy, Pandas, Scikit-learn, Tensorflow, Pytorch, Matplotlib, Seaborn, SQL, Bash, Git, R

**MACHINE LEARNING:** Classification, Clustering, Neural Networks, Recommender Systems, Regression, Time Series

**COMPETENCIES:** Analytics, Algorithms, API's, A/B Testing, Bayesian Inference, Critical Thinking, Communication Skills, Data Cleaning, Hypothesis Testing, Predictive Modeling, Problem solving, Public Speaking, Segmentation, Quantitative Research, Tableau, Web Scraping

## EXPERIENCE

---

### Data Science Immersive Fellow

General Assembly | Remote | Mar 2023 - Jun 2023

Honed critical data science techniques on data analysis, visualization, and machine learning across 480+ hours of expert-led instruction, finishing in the top 20% of the class. Individually and in small teams, completed 6 data science projects in two-week timeframes, succinctly communicating results to technical and non-technical audiences.

- **Computer Vision:** Designed a stand-alone tool that accepts handwritten math equations as input, then leverages a Convolutional Neural Network to render the equation digitally, correctly predicting at least 50% of the equation in 70% of cases.
- **Time Series Forecasting:** Predicted solar power available in Los Angeles using a Recurrent Neural Network and weather data from the National Solar Radiation Database, with an average error of 15% at solar noon.
- **NLP and classification:** Developed a binary classification model which correctly classifies Reddit posts and comments between r/politics and r/conservative with 88.6% accuracy, using logistic regression, random forest, and stacking.

### Postdoctoral Researcher

Kavli Institute for Particle Astrophysics and Cosmology | Stanford University | Oct 2019 - Dec 2022

Analyzed and modeled observational data of astrophysical objects in partnership with distinguished faculty researchers. Coordinated international research amongst cross-functional teams of 10+ scientists, and authored peer-reviewed papers in leading journals.

- **Communication:** Presented results of peer-reviewed studies at US conferences; collaborated with the *Chandra* press office to craft an accessible, scientifically meaningful, and visually appealing [press release](#) - in the top 10% most watched videos on the *Chandra X-ray Observatory* Youtube Channel.
- **Mentorship:** Supervised 3 new scientists/researchers in department operating procedures, instrumentation software, and best practices. Instilled data stewardship and rigorous application of statistical models.
- **Proposals:** Crafted compelling proposals with a success rate ~3x above the average, securing time at astrophysical observatories. As a member of the *Chandra* Time Allocation Committee in 2022, assessed and evaluated observing proposals based on merit, cost, risk, urgency, and suitability for the observatory.

## Select Peer-Reviewed Publications

- [Chandra Measurements of Gas Homogeneity and Turbulence in the Perseus Cluster](#)  
Developed a custom statistical method to measure galaxy cluster turbulence, using Bayesian Inference and Markov Chain Monte Carlo Sampling. Reduced error margins by 20%, and redefined benchmarks for simulations, theoretical work, and future observations.
- [The Long Filament of PSR J2030+4415](#)  
Discovered new pulsar filament, mapping its full extent with fresh observations, feeding new data-driven insights into the origin and physical nature of these filaments. Developed a new 'filament-finding' algorithm based on this discovery, which will be used to search for filaments in upcoming observations.
- [A Quarter Century of Guitar Nebula/Filament Evolution](#)  
Spearheaded first long-term study of a pulsar filament, comparing 4 epochs of observations (data from 25+ years). Developed generalizable and re-usable Python code to fit bow shock models to observational data.

## EDUCATION

---

**General Assembly Data Science Immersive** | Remote | Mar 2023 - Jun 2023

**Ph.D. in Astrophysics** | University of Amsterdam | Amsterdam | Sept 2015- Aug 2019

**M.Sc. in Astronomy and Astrophysics** | University of Amsterdam | Amsterdam | 2013-2015

- **Relevant Coursework:** Bayesian Statistics, Computational Astrophysics

**B.Sc. in Astronomy** | University of Groningen | Groningen | 2009-2013

- **Relevant Coursework:** Introduction to Python, Linear Algebra, Calculus, Statistics