

Tasha Gautam, Ph.D.

Toronto, ON, Canada • 16.tasha@gmail.com • 4374469172
linkedin.com/in/tasha-gautam-phd-80567885 • tgautam16.github.io

SUMMARY

– PhD in Astronomy with **5** years of experience leading data-driven projects using quantitative analysis, looking to use my skills in industry to help companies make data-informed decisions.

– Managed, processed, and analyzed large datasets using advanced statistics and ML techniques, proficient in Python and SQL. Reported work in **13** high impact publications, adept at communicating insights.

PROFESSIONAL EXPERIENCE

Data Scientist and Project Management expertise

gained as Research Associate at NRAO, VA, USA, Jan-Nov 2023

- Led a team of 8 to gain new insights from data on unknown stellar properties; acquired a **\$0.2M** NSF grant.
- Managed and processed complex TBs of data with Python, employing **time-series analysis, regression, and predictive algorithms**. Identified new patterns and discovered a General Relativistic effect in the data.
- Developed a python **pipeline** to process data and created new models for 68 star systems, achieved **150%** improvement in measurements. Implemented statistical tests for anomaly detection and performed regression analysis.

Data Scientist and Data Analyst expertise

gained as Research Assistant at MPIfR, Bonn, Germany, 2018-2022

- Led a project to find new insights from **10 years** of complex data using **regression and predictive** techniques. Discovered 3 new stellar properties and boosted accuracy by **400%**.
- Spearheaded data analysis of new stars from 2 surveys, **modeled and extracted new features** using hyperparameter optimisation, regression and regularization. Published findings in two scientific papers.
- Developed **ETL data pipeline** with **container support** for analyzing **100+** time-series datasets. Discovered **6** new neutron stars, marking the first such findings by the MeerKAT telescope.
- **Boosted computing efficiency** by **24x** via parallel computation and optimized CPU/GPU allocation, resulting in reduced survey costs.

Recent projects | Skills: Time Series Analysis, NLP, ML, Python, SQL, Analytics, Visualization
Toronto, Canada

Uncovered new relativistic effects in a neutron star

Performed time series analysis and used statistics and ML to analyze large-scale dataset and identified characteristic trends to discover new insights about the star system.

Discovered new stars in two large surveys

Analyzed complex data from more than 10 star clusters and discovered 6 new stars, determined new features for each system and created a data pipeline to efficiently look for rare signals in the data-set.

Sentiment Analysis of ChatGPT tweets

Analyzed 3 months of tweets utilizing NLTK and Vader tools to classify sentiments as positive or negative. Trained ML models, optimized model performance metrics, achieved 94% accuracy.

Time Series Forecasting of Stock prices

Identified high and low-risk investments using 7 years of stock price data from leading tech companies. Conducted advanced modeling and time-series forecasting of stock prices with 90% confidence.

Predictive modeling of customer reviews

Conducted predictive analysis on customer reviews of products from a clothing brand to determine sentiments. Trained decision trees and other classification models with 87% prediction accuracy.

EDUCATION

Ph.D., Astronomy

University of Bonn (MPIfR), Bonn, Germany, 2018-2022

MSc, Physics

Indian Institute of Science Education and Research, Mohali, India, 2015-2018

TECHNICAL SKILLS

Languages

Python (6 yrs) • SQL (3 yrs) • Git (5 yrs)

Data analytics

Pandas • Matplotlib • Tableau • PowerBI • PySpark

Data pipeline

Containers: Docker, Singularity • Kubernetes • Azure Databricks

Machine Learning

Scikit-learn • TensorFlow • Pytorch • NLP • Time Series Analysis

CERTIFICATIONS

Machine Learning A-Z: AI, Python & R + ChatGPT for data science

Udemy | Credential: UC-b32936e4-869b-4e5e-828b-67802b8f3c69

Azure Databricks & Spark For Data Engineers (PySpark / SQL)

Udemy | Credential: UC-c2e4fc01-81ed-47e4-891f-23a9038b96c3

The Complete SQL Bootcamp

Udemy | Credential: UC-9f65de70-2f9d-41ca-8e31-ca6831994fbb

CERTIFICATIONS

Leadership: Led multiple data science projects with cross-functional teams of **20** • Reported results in **4 lead author** papers in distinguished international journals • Supervised **2** interns • Representative at IMPRS.

Teamwork: Collaborated on 12+ data science projects with teams of up to **50** scientists • Organized and chaired weekly seminars and yearly conference at IMPRS, Germany.

Communication: Delivered 20+ public talks to technical and non-technical audience of up to **40** people in conferences • Supervised Master's course at University of Bonn, Germany.