# PRAHALD SIWAKOTI

Physicist    Data Scientist

📞 225-210(0199)

📍 Harrisburg, PA

⌨ github.com/siwa-p

🔗 prahlad-siwakoti

## SUMMARY

I am an experimental physicist turned Data Scientist. In my career as a researcher, I have used data from a wide range of measurements sources to analyze and investigate phases of matter and presented my findings through publications in academic journals and research conferences. As a data scientist, I have experience with machine learning, AI, and statistical models as well as deriving insights from complex datasets using Python, R, SQL and other open-source tools. I am looking forward to integrate my educational experience, research background and programming tools to solve complex problems.

## SKILLS

| | |
|---|---|
| **Languages:** | R, SQL, Python, SAS, Excel |
| **Version Control:** | Git/Github |
| **Visualization:** | Power BI, ggplot, Matplotlib, R-Shiny |
| **Machine Learning:** | Sci-kit Learn, Tensorflow, Pytorch |
| **Query:** | SQL |
| **Additional Skills:** | Object Oriented Programming Debugging, AWS, Docker, Time-Series Analysis, Deep Learning |

## EDUCATION

**1/2024 – 5/2025**  **Masters in Data Science**  <span align="right">**University of Texas at Austin**</span>
Courses taken: Machine Learning, Probability and Statistical Inference, Data visualization, Algorithms, Advanced Predictive Models, Deep Learning, Natural Language Processing, ReInforced Learning, Data Science for Health.
Graduation date: May 2025

**09/2023 - 7/2024**  **Data Science Apprenticeship**  **Nashiville Software School, Nashville, TN**
Intensive part-time boot-camp focusing on data science fundamentals and problem solving
Graduation date : June, 2024

**08/2015 - 11/2021**  **PhD in Physics**  **Louisiana State University, Baton Rouge, LA**
Dissertation:
Effects of Structure, Crystallographic Orientation, and Dimensionality on Emergent Properties of Transition Metal Oxide Thin Films

**08/2011- 03/2014**  **Masters of Science in Physics**  **Tribhuvan University, Kathmandu, Nepal**
Dissertation:
*First-Principles Study of Neutral $((N_2)_n)$ and Singly Cationic $((N_2)_m^+)$ Molecular Nitrogen Clusters; (n = 1, 2 and m = 1, 2, 3, 4, 5 and 6)*

## DATA SCIENCE EXPERIENCE

**9/2023 – 7/2024**  **Data Scientist Apprenticeship**  **Nashville Software School**

- Wrangled data and performed exploratory data analysis using Python's pandas library and R's tidyverse packages
- Created data visualizations using matplotlib, seaborn, and ggplot2
- Performed geospatial analysis using geopandas and folium
- Gathered data through APIs and webscraping
- Retrieved and analyzed data using PostgreSQL and sqlalchemy
- Built and evaluated statistical and machine learning models using the scikit-learn and statsmodels libraries
- Developed and evaluated machine learning models for classification and clustering tasks, with hands-on experience interpreting confusion matrices, ROC curves, and precision-recall metrics.
- Applied natural language processing using the nltk and spaCy libraries
- Performed network analysis on graph data using Neo4j
- Built and deployed interactive data visualizations using the R Shiny library
- Source code version control with Git/GitHub
- Project management/tracking with GitHub project boards and issue tracking
- Interacted with AWS using the CLI and ssh

| 8/2024 – 10/2024 | **Advanced SQL Professional** | <div align="right">**Nashville Software School**</div> |
|---|---|---|

- Gained hands-on experience with PostgreSQL.
- Dived deeper into each topic through additional resources and exercises.
- Learned the fundamentals of database design, including creating Entity-Relationship Diagrams (ERDs) and normalizing data.
- Mastered SQL scripting to modify database structures using CREATE and ALTER statements.
- Practiced data manipulation through INSERT, UPDATE, and DELETE operations.
- Explored advanced querying techniques, including Sub-Queries, Common Table Expressions (CTEs), and Window Functions.
- Automated processes with VIEWs, Stored Procedures, Triggers, and Transactions.

## PROFESSIONAL EXPERIENCE

| 11/2021 – 12/2023 | **PostDoctoral Researcher** | <div align="right">**University of Tennessee at Knoxville**</div> |
|---|---|---|

- Designed and executed experiments in the field of Quantum thin films, creating novel functional properties in artificial structures. Successfully completed projects resulted in publications in reputable academic journals.
- Wrote successful synchrotron beamline proposals and performed experiments on site in Advanced Photon Source, Argonne National Lab
- Developed and maintained a data analysis pipeline for synchrotron data using Python and R
- Wrote python scripts to simulate observed data and to perform statistical analysis
- Collaborated with researchers from various disciplines to analyze, interpret data and deduce conclusions
- Provided mentorship and training to graduate students with research, instrumentation, and troubleshooting

| 01/2018 – 11/2021 | **Graduate Research Assistant** | <div align="right">**Louisiana State University, Baton Rouge**</div> |
|---|---|---|

- Explored non-trivial physics of transition metal oxide perovskite thin films with respect to their symmetry and growth orientation and studied various two-dimensional defects.
- Provided training to undergraduate students with research, instrumentation, and troubleshooting

## PROJECTS

| Project 1 | **Time-Series Forecasting : A python Implementation** | **LINK** |
|---|---|---|

Exploration of various time-series forecasting methods using Python. Various statistical and machine learning models were implemented to predict the future values of a time-series data and compared with the actual values. The models include ARIMA, XgBoost, and LSTM.

**Skills:** Time-Series Analysis, Data Wrangling, Data Cleaning, Data Visualization, LSTM, XgBoost, ARIMA, Pytorch

| Final-Capstone | **Air Quality: Machine learning models applied to air quality data** | **LINK** |
|---|---|---|

Constructed a predictive model for air-quality monitoring from data obtained from inexpensive air-sensors by PurpleAir and various meteorological data.I have utilized various tree-based spatio-temporal models as well as neural networks to predict the air quality.

**Skills:** Time-Series Analysis, Spatial regression, Kriging interpolation, Machine Learning, Deep Learning, Data Visualization

| Midcourse-Capstone | **Wildland fires and their effects on visitation data in US National Parks** | **LINK** |
|---|---|---|

Created an interactive R Shiny app of various National Parks in the US featuring wildfire events in the past to visualize the effect of these events in the park visitation statistics.

**Skills:** ARIMA forecasting, R Shiny, Data Wrangling, Data Cleaning, webscraping

## PEER REVIEWED PUBLICATIONS

Google Scholar :                                                                                     ***Prahlad Siwakoti***