
BABAN NAGARE, Ph.D.

902-412-2165 | [LinkedIn](#) | bnagare@gmail.com | Ontario, Canada

Experienced **Data Scientist** with a proven record in a fast-paced working environment of academic and corporate research. Proven record in experiment design, sampling, and data analysis and governance. Proficient in predictive modeling using various machine learning and time series algorithms. Managed and executed scientific data collection expeditions onboard research ships and university sites. Adept in conceptualizing and executing projects. Expertise in using tableau and Python libraries for data analysis and dashboard making.

Detail-Oriented | Result-Driven | Collaboration | Critical-Thinking | Business Sense | Effective Communication

KEY SKILLS

Programming Languages: Python (expert), SQL (expert), R (proficient)

Frameworks: TensorFlow, PyTorch, scikit-learn, Keras, Pandas, Matplotlib, NumPy, Open-CV, Xgboost

Environments: Azure ML, Azure Synapse, Databricks, AWS SageMaker, AWS S3, AWS EC2

**Machine Learning
Deep Learning
Natural Language Processing
Data Visualization**

**Large Language Models
Time Series analysis and forecasting
Mathematical and Statistical Modeling
Climate Science**

PROFESSIONAL EXPERIENCE

Data Scientist, *Fiddlehead Technology - Moncton, NB*

2022 – 2023

- Designed, developed, and deployed an in-house seasonality module for a time series package resulting in a 30% forecast improvement.
- Predicted bulk orders using ensemble average of ARIMA, SARIMA, Xgboost and Long Short-Term Memory (LSTM) for weekly predictions for retail stock-keeping units.
- Conceptualized and conducted comprehensive modeling to analyze the relationships between bulk orders and point-of-sale data at both province and country levels and used point of sale data as exogenous variable.
- Researched and calculated vegetation indices to input as features in the model to predict agricultural produce. Used multiple regression models including decision tree, random forest, Xgboost to predicted agricultural produce yields using ensemble averaging and best model performance.
- Clustered agriculture plots are on productivity, weather conditions, soil properties and irrigation facilities using unsupervised learning. Used clusters to classify plots using decision tree, support vector, random forest, Xgboost classification algorithms.
- Automate and update client weekly performance dashboards using PowerBI and tableau.
-

Research Associate, *Dalhousie University - Halifax, NS*

2018 – 2022

- Acquired, cleaned, and analyzed high-quality data during field experiments to support research initiatives.
- Lead data analysis workshops with collaborators resulting in multiple peer-reviewed publications in leading scientific journals.
- Implemented unsupervised machine learning clustering algorithms to air pollution data to uncover the physical relationship among the features, improving the understanding of environmental factors and their impact on fog formation.
- Developed a machine learning approach to automate data quality monitoring and detect events of interest using unsupervised machine learning.
- Managed two field experiments to collect air quality data over the Atlantic Ocean and coastal sites successfully.
- Taught subject expertise to graduate and undergraduate students and supervised two undergraduate students for their honors thesis.

Research Associate, Indian Institute of Tropical Meteorology - Pune, India 2016 – 2018

- Developed standardized data collection protocols for the cloud seeding experiment in India, ensuring data consistency and reliability.
- Parameterized the relationship between trace metal pollutants and other pollution features contributing to a deeper understanding of environmental processes.
- Parameterized the statistical relationship between RADAR signal and in situ measured weather data.

Credit risk advisor, Dhanlaxmi Bank, India 2016 – 2018

- Obtained data from stakeholders about client, history of client and assets.
- Evaluating asset value based on government regulation and local body norms including interacting with authorities to resource data and legal advisory.
- Checked authenticity and validity of client documents by cross checking on government websites.
- Advice credit managers on decision based on client existing assets and liabilities.

Research Assistant/Teaching Assistant/PhD Candidate, ETH Zurich - Switzerland 2012 – 2016

- Conceptualized, developed, and implemented a signal processing algorithm for backscattered light to identify the ice crystals and water droplets in an airstream, resulting in improved data accuracy for experiments.
- Wrote an extensive model for water droplet interactions with the surrounding environment.
- Developed a model to simulate the experimental situation and performed sensitivity analysis to understand causality.
- Developed a statistical parameterization based on electric charge on particle and water droplets enhancing the efficiency of estimate the collision estimate.
- Standardized data collection protocols for laboratory experiments, ensuring data consistency and reproducibility.
- Tutored undergraduate students in remote sensing principles demonstrating strong teaching and mentoring abilities.

EDUCATION

Doctor of Philosophy(Ph.D.) Atmospheric and Climate Science, ETH Zurich, Switzerland

Master of Science (M.Sc.) Physics and Atmospheric Science, Pune University, India

Bachelor of Science (B.Sc.) Physics, Pune University, India

TRAINING COURSES

Specialized Models: Time Series and Survival Analysis, Coursera – Online

Applied Machine Learning Specialization, 5 courses, Coursera – Online

Mathematics for Machine Learning, 3 Courses, Coursera – Online

Credit Risk Modeling using Machine Learning: Campusx-Online

Basics of MLOps: Campusx-Online

PROJECTS

- Predicting snowfall for city of Moncton using ARIMA, SARIMA, XGBOOST and multivariate LSTM. Collected open-source meteorological data, including windspeed, wind direction temperature, performed cleaning and preprocessing of data, training, and testing and evaluation of models. Presented the outcomes at group meeting at Dalhousie university.
- Enhancing customer engagement for Credit Cards: A segmentation approach for credit card users with model deployment on web. Performed cleaning and preprocessing and feature engineering on data. Performed K-means, Hierarchical clustering, DBSCAN clustering. Deployed the trained model on Streamlit.
- QuakeAI Fusion: Exploring time series dynamics, magnitude forecasting, clustering, anomaly detection, and generative AI for seismic events.

- SmartECG: A comprehensive approach to automated classification of Cardiac and COVID-19 Patients Using 12-Lead ECG Images – Integrating Machine Learning, Deep Learning, Transfer Learning, Explainable AI, and AutoML Techniques.
- EnergyForecast Pro: Harnessing the Power of Machine Learning and Deep Learning for Advanced Prosumer Energy Behavior Prediction
- EmoCare: An Empathetic AI Workshop on Fine-tuning of LLaMA 2 (Large Language Model) for Sentiment Analysis of Cancer Survivors Caregivers and Recommendation Systems in Healthcare