

Passionate data science enthusiast skilled in programming, statistics, data analysis and machine learning. Committed to leveraging these skills to uncover valuable insights and solve real-world challenges. Actively seeking opportunities to apply expertise and stay current with the latest advancements in the field.

Work History

2022-05 -
2022-08

Machine Learning Data Scientist Intern

Mitchell International, San Diego, CA

- Executed the development and deployment of XGBoost-based claims automation pipeline on Amazon SageMaker
- Integrated preprocessing of million rows of data and error handling in AWS Lambda
- Built a Node.js script with Newman as library and expedited API testing by 80%.
- Created data visualization graphics, translating complex data sets into comprehensive visual representations.

2020-11 -
2021-08

Application Development Associate

Accenture, India

- Developed and managed change requests for 'Employee Portfolio' web application components, employing ASP.NET MVC Framework
- Assisted a team of 30 software architects to maintain and improve in house software modules and technical support
- Accelerated development time by 40% by leveraging Bootstrap's pre-designed UI components and CSS classes, streamlining prototyping and feature development.
- Collaborated with multidisciplinary teams to design and implement new technology features.

2020-08 -
2020-11

Machine Learning Intern

Imaginea, Accenture, India

- Formulated a Conditional Random Field (CRF) based NLP model in python to automate and escalate employee's reimbursement process by 35%

Good2go Recruiter

Data Scientist

Contact

Address

Buffalo, NY 14223

Phone

000000000

E-mail

prajvaln@buffalo.edu

LinkedIn

<https://www.linkedin.com/in/prajval-nataraj-a0b119>

Technical Skills

Python

R

SQL

MATLAB

C

Data Structures

NumPy

Pandas

Scikit learn

TensorFlow

Jupyter

Notebook

RStudio

Visual Studio IDE

- Conducted data extraction and annotation of large dataset of employee bills, utilizing POS tagging and sentiment analysis to categorize bills into relevant expenditures
- Evaluated the effectiveness of Tesseract OCR and Python libraries such as PDFMiner and PyPDF2 in automating data extraction from scanned images and PDFs, achieving a 50% reduction in manual data entry.

2019-07 -
2020-05

Machine Learning Intern

Tekatron IT Solutions, India

- Constructed end-to-end machine learning pipelines to address complex business problems, spanning data preprocessing, feature engineering, algorithm selection and hyperparameter tuning
- Employed a range of algorithms, including Linear Regression, Logistic Regression, Decision Tree, SVM, Naive Bayes, KNN, and K-Means, to produce effective solutions
- Spearheaded research efforts on cutting-edge machine learning techniques, with a focus on neural networks and GANs (Generative Adversarial Networks), resulting in a 20% improvement in model performance and scalability
- Created and presented detailed reports and visualizations, utilizing data visualization tools such as Tableau, Matplotlib, and Seaborn, to effectively convey the results of machine learning experiments to technical and non-technical stakeholders.

Education

2021-08 -
2022-12

Master of Science: Data Science

University of Buffalo - Buffalo, NY

GPA: 3.592/4

2016-08 -
2020-08

Bachelor of Engineering: Computer Science

JSS Academy of Technical Education - Bengaluru, India

Tableau

Snowflake

AWS (SageMaker, Lambda)

MS Office

Data Analysis & Modeling

Data Visualization

Machine learning

Deep Learning

Regression

Clustering and Classification

Natural Language
Processing (NLP)

Neural Networks

Time Series

Analysis/Forecasting

A/B Tests

Accomplishments

1. Covid Prediction and Analysis

- Charted a time series forecasting model to analyze future trend in Covid deaths and performed statistical analysis with R.
- Employed ggplot2, Plotly, and lattice libraries to introduce data visualization of vaccines administered, impact, and coverage of completely immunized US residents

2. ATP (Association of Tennis Professional) Analysis

- Pioneered an analysis technique that provides real-time statistics data on players and events
- Performed data injection and exploratory data analysis using Pandas, Plotly, and SQL.