

# PRADYUMNA BADA

pbada2@illinois.edu | (217) 200-3976 | [LinkedIn](#) | [Github](#) | Urbana, Illinois - 61801

## EDUCATION

**University of Illinois, Urbana - Champaign | GPA 3.8/4**

**May 2025**

Master of Science, **Advanced Analytics**

Coursework: Big Data Analytics, Data Mining, ML Cloud (**AWS**), Deep Learning, Machine Learning.

**JSS Science and Technology University | GPA 8.5/10**

**Nov 2020**

Bachelor of Engineering, Industrial Engineering

Coursework: Advanced Calculus, Linear Algebra, C Programming, Operations Research, Statistical Quality Control.

## SKILLS

**Coding:** Python, SQL (Oracle, MySQL, PostgreSQL), NoSQL (MongoDB), C Programming, COBOL, Jenkins, GITLAB, Azure, Docker.

**Data Analytics:** Pandas, Numpy, Excel, Matplotlib, Power BI, Tableau, PySpark, Prophet, Scikit-Learn, FAISS, Synapse Analytics.

**Other Skills:** JIRA, Insurtech Experience, Azure OpenAI Studio, Web Scraping (Beautiful Soup), Power Query.

## PROFESSIONAL EXPERIENCE

**Cline Center for Advanced Social Research | Research Assistant**

**Sep 2023 - Present**

- Leveraged LLMs to extract structured data from unstructured text (news articles), specifically tracking police use of lethal force.
- Specialized in data cleaning, analysis, schema design, **DDL**, **DML**. Built a national dashboard ( [Link](#) ) using **MySQL** & **Power BI**.
- Built retrieval models using embeddings for ranking data and clustering articles to automate the data extraction process.

**Seagate Technology | Data Analyst Intern**

**May 2024 – Aug 2024**

- Analyzed Volume Data using statistical tools and data visualization to help characterize emerging memories.
- Developed **Python**-based automation & reporting systems, utilizing **Jenkins**, **GitLab**, & other technologies to streamline operations.
- Calculated failure metrics, including Bit Error Rates across address levels, and modeled CFR to predict potential failures.

**RailTEC at Illinois | Summer Research Intern**

**May 2023 - Aug 2023**

- Implemented **Big Data Analytics** to ensure safe railway operations, by using **LRAIL**, an AI-based system for Visual Track Inspection.
- Analyzed periodic data to track & forecast the health of rail components by leveraging **PySpark** and **MS Azure Synapse Analytics**.
- Visualized data with **Power BI** and **Python** to provide actionable insights for efficient decision-making.

**Sapiens International Corporation | SQL Developer**

**Mar 2021 - Nov 2022**

- Enhanced our Insurtech product using **Oracle SQL**, and **ETL**, which helped companies administer legacy **Insurance** products.
- Expertly managed datasets encompassing policy, claims, customer, and fund prices, and crafted complex SQL data fix scripts
- Developed reporting interfaces using SQL and **Shell Scripting**, including 'ECASS', a cache facility designed for client's actuarial team.

## ACADEMIC PROJECTS

**Stock Price Forecasting Using News Sentiment Analysis | Python, Streamlit, Prophet.**

- **Goal:** Develop an interactive web app to forecast stock prices by integrating sentiment analysis of financial news.
- Used a pretrained model for sentiment analysis and Prophet for forecasting, delivering a dynamic dashboard with real-time plots.

**How Successive Wars Crippled Indian Economy | Python, Jupyter Notebook.**

- **Goal:** To understand the impact of war on Indian economy using data visualization.
- Used Pandas, NumPy for data analysis and Matplotlib for data visualization. Visually presented the results using line chart.

**Hypothesis Test | Python, SciPy**

- **Goal:** To assess the contrasting effects of a recession on university towns as compared to non-university towns in the USA.
- Used pandas for preprocessing and SciPy for hypothesis testing.

**Predictive Modeling for Failure Prevention in Scania Truck Air Pressure System (APS) | Python, GitHub, MICE.**

- **Goal:** Develop a cost-based model to predict and minimize failures, considering the costs of unnecessary checks and breakdowns.
- Performed EDA, & data imputation to develop various ML models, including LR, KNN, DT, RF & used Grid Search CV to compare it.

## LEADERSHIP & INVOLVEMENT

- Ensured timely meal service for **hundreds** of students daily through food servicing, dish cleaning, and table sanitation.
- Led the development team in resolving high-priority production issue in the absence of senior team members.
- Completed certification courses like Applied Machine Learning, Applied Plotting & Data Representation in Python, and more.
- Organized multiple treks to the Great Himalayan Ranges for groups of more than 10.
- Directed and acted in short films, cover songs, and created **travel vlogs**. [Video1](#), [Video2](#), [Video3](#), [Video4](#), [Video5](#)