

# Pavan Borole

pborole2@illinois.edu | +1-217-693-8978 | LinkedIn

## EDUCATION

---

### University of Illinois at Urbana Champaign

*Master of Science in Information Management; GPA: 4.0*

Champaign, USA

January 2021 - May 2022

### University of Mumbai

*Bachelor of Engineering in Electronics and Telecommunication; GPA: 3.5*

Mumbai, India

August 2015 - July 2019

## EXPERIENCE

---

### Business Intelligence Group

*Consultant*

Champaign, USA

January 2022 - present

- Working with stakeholders of a renowned educational institution on a data science project involving statistical analysis and data visualizations.

### Synchrony Financial

*Emerging Technology Center Intern - Strategic Automation*

Champaign, USA

January 2022 - present

- Developing process automation workflows for the FinOps team using SQL to reduce process costs and time.

### AGCO Corporation

*Market Analyst Intern*

Champaign, USA

October 2021 - January 2022

- Collaborated with internal stakeholders to identify key factors and data sources required to build a market intelligence model that could provide consistent and uniform global and regional market reports.
- Developed a data pipeline by aggregating the data sources and analysing data using AWS S3 and Athena.

### University of Illinois System-AITS

*Data Science Analyst Intern*

Champaign, USA

April 2021 - October 2021

- Developed ETL pipeline connecting SQL DBs and KNIME to generate BI reports that delivered strategic insights.
- Designed a workflow to extract text from research proposal PDFs and identify keywords from the abstracts present in research proposals using LDA and RAKE algorithms. Automated the process of identifying the scope and domain of research projects and made it 60 times faster.
- Developed a machine learning model that uses Latent Dirichlet Allocation(LDA) to create a body of tweets(from data pulled through the Twitter API) and extract the most relevant trending topics based on domain specified.

### Altizon Inc.

*Solutions Analyst*

Pune, India

August 2019 - December 2020

- Collaborated with the client's operations staff to understand business challenges and design solutions to streamline processes, improving manufacturing efficiency and reducing downtime by 13%.
- Analyzed and interpreted large datasets and visualized data on dashboards with Apache Superset using SQL which helped for driving data driven business solutions and enhance production by 8%.
- Developed Python scripts to extract, clean and transform network data from the cloud and automate data acquisition and collection processes. This reduced the time required for process completion by 50%.

## PROJECTS

---

### Hybrid Movie Recommendation Engine

- Combined content and collaborative filtering models to create a hybrid movie recommendation engine by using ML and NLP libraries. Tested the engine by using a dataset created by scraping data from TMDB.
- The engine offers recommendations that are more personalized and tailored towards each user.

### NYC Taxi and Uber Trips Analysis

- Analyzed over 133 million trips of publicly available taxi and Uber data through exploratory data analysis. Used the vaex library that allowed dataframe processing and operations to be upto 1000 times faster.
- Visualized the trends that defined the city's habits before and after the COVID-19 pandemic using Tableau.

## TECHNICAL SKILLS

---

**Languages and Databases:** Python(pandas, NumPy, sklearn, keras, tensorflow, nltk, transformers, vaex, seaborn, plotly), R, C++, MySQL, PostgreSQL, neo4j, MongoDB, Greenplum, HTML

**Tools:** Jupyter Notebook, Tableau, PowerBI, Apache Superset(BI), AWS, Azure, KNIME, Talend, Git, Docker, Linux, Oracle SQL Developer, SAP BusinessObjects, MS Excel

**Skills:** Data Cleaning, Data Visualization, Exploratory Data Analysis, Data Mining, Hypothesis Testing.