

SUMMARY

Motivated, teamwork-oriented data analyst with significant experience of 3+ years in solving complex business problems using SQL, Python, and MS Excel for fast-paced product and services teams. Seeking full-time and summer internship opportunities in Data Science and analytics.

COMPETENCIES

- **Programming** – SQL, Python, R, Java, SAS • **Databases** – MySQL, T-SQL, BigQuery, DataBricks, Teradata, Oracle, SQLite
- **Skills** – Machine Learning, K Means Clustering, Bayesian Inference, Hypothesis Testing, NLP, Web-Crawling, A/B Testing, Data Mining, ETL, Deep Learning, A/B Testing, Anova, Classification
- **Libraries** – pandas, scikit-learn, matplotlib, NumPy, seaborn, TensorFlow, pyspark, NLTK • **Platforms** – Hadoop, Spark, Linux, Microsoft Azure, AWS
- **Tools** – Visual Studio Code, SQL Server, Power BI, Microsoft Office Suite, Stata, Jupyter Notebook, MS Excel, Tableau, MS PowerPoint

EDUCATION**J. Mack Robinson College of Business, Georgia State University, Atlanta**

- **Master of Science in Data Science and Analytics** (August 2021 – April 2023)
- **Master of Science in Quantitative Risk Analysis and Management** (August 2021 – April 2023)
 1. **Web Scraped data of Fortune 500 for better Earnings Per Share Estimations**
 - Scraped the data using **selenium**(python) and created a database from scraped data with non-redundant and normalized tables for better query execution time.
 - Executed multiple regression algorithms(**Linear and Logistic**) to predict the Earning per Share for fortune 500 companies.
 2. **Movie/Sitcoms Recommendation System**
 - Performed **exploratory data analysis** and data cleaning on Netflix and IMDB data to draw meaningful patterns and insights.
 - Built the enhanced recommendation model on Netflix data applying **K-Nearest Neighbour** manually to predict the top 10 most similar movies/sitcoms.

Bundelkhand Institute of Engineering and Technology, Jhansi, Up

- **Bachelor of Technology in Computer Science and Engineering** (August 2013 – June 2017)
 1. **Email Spam Classifier**
 - Cleaned the data(Enron Emails) using **NLTK** library to help remove poor quality results from our machine learning model.
 - Designed a spam classifier model to predict whether the email is spam or not using a **Naïve Bayesian algorithm** with an accuracy of 93%.
 2. **Monte Carlo Simulations on Betting**
 - Build a simulator to help one visualize most or all of the potential outcomes involved in betting decision-making.
 - Through the simulations of **randomized** betting trials, we concluded that "House Always Wins" even with a 49% probability of winning.

EXPERIENCE**Georgia State University**

(Atlanta, Georgia, USA)

Graduate Research Assistant

(January 2022- Current)

- Working in a mutual fund lab, to mine insights between mutual fund disclosures and inflation using **Text Analytics**(topic modeling)

Starr Insurance

(Atlanta, Georgia, USA)

Graduate Research Assistant

(August 2021 – December 2021)

- Developed a model to predict the likelihood of an existing customer if they will buy the recommended add-ons insurance using **basic statistics and Random Forest Classifier**.

Bipp Consulting

(Agra, Uttar Pradesh, India)

Business Intelligence Analyst

(September 2020 – July 2021)

- **Streamlined, backtracked, and automated** 4 KPI's for Johnson & Johnson, which reduced the person-hours by 30%.
- Decreased database **complexity** by **normalizing** and **optimizing** tables and queries in Johnson & Johnson SQL Databases, reducing production time by approximately 20%.

ParallelDots

(Gurgaon, Haryana, India)

Senior Business Analyst

(July 2019 – August 2020)

- Created processes and solutions for seamless execution of multiple projects, which involved image processing of 3 million images per month in helping Hindustan Unilever identify critical aspects of **advertising**, thereby increasing sales by 3%-4%.
- **Fostered** resources of 27 data annotation employees and multiple business analysts according to skill, timeline, relevance, and revenue.
- Performed **sentiment analysis** using **logistic regression** on survey results for Indian tobacco company that helped them launch **2 low-cost** products in the Indian market, increasing their market share by 2%.

Decision Point Analytics

(Gurgaon, Haryana, India)

Business Analyst

(December 2018 – June 2019)

- **Collected, cleansed, and transformed** 10TB of structured and unstructured Coke Data on Big Query every month.
- Created PowerBI reports based on sales to identify key marketing strategies for Coke in Latam Region.
- Drove the improvements in database processes resulting in a decrease in time and daily data usage in Big Query, which reduces the overall server cost by 25% and data processing time by 10%.

MAQ Software

(Hyderabad, Telangana, India)

Software Engineer

(March 2018 – December 2018)

- Developed a workflow in SSIS to improvise user experience by streamlining incoming survey data from Adobe Omniture for Microsoft.
- Created **interactive** dashboards using PowerBI to analyze the hit ratio of web pages, thus identifying pages with more traffic.
- Implemented funnel analytics and statistics on web traffic data of Microsoft which led to 10% increase in user signups.

CERTIFICATES AND ACCOMPLISHMENTS

- **761-QUERYING DATA WITH TRANSACT-SQL** by Microsoft.
- **778-ANALYZING AND VISUALIZING DATA WITH MICROSOFT POWER BI** by Microsoft.
- **MICROSOFT CERTIFIED PROFESSIONAL: MICROSOFT CERTIFIED PROFESSIONAL** by Microsoft.
- **RANK 1** and **GOLDEN BADGE** in **SQL MODULE** on **HACKERANK**. (USERNAME- BrownBatman)