



PAULAMI RAY

Data Science and Big Data Analytics Engineer

Experienced data enthusiast having the skills to use extremely large, complex data sets to draw conclusions and effectively communicate findings to stakeholders. Motivated to constantly develop new skills and grow professionally. I follow three main values at work – Respecting individuals, innovative thinking and sharing knowledge with each other.

Looking for opportunities in Analytics field in Sweden

Phone: +1 (773) 727-5540 || **Email:** paulamilahiri@gmail.com

LinkedIn: <https://www.linkedin.com/in/paulami-lahiri-rav-b157a736/>



Education

MS in Information Management - Specialization in Data Science and Analytics

GPA: 3.8/4.0

University of Illinois, Urbana-Champaign, IL, USA

Dec 2018

Coursework: Data Statistics, Data Mining, Text Information Systems, Data Visualization, Machine Learning, Big Data Analytics, Data Science Storytelling, IoT & Business Analytics, Socio-Technical Information Systems

Literature Review – Text Mining approaches in Security domain

Bachelor of Technology (B.Tech) in Computer Science & Engineering

GPA: 3.8/4.0

West Bengal University of Technology, India

Aug 2009



Technical Skills & Certifications

- **Computer Languages:** Python, R, SQL, PLSQL, Java, HTML, Shell Scripting
- **Data Science Models & Libraries:** Numpy, Pandas, Scikit-learn, Matplotlib, Plotly, BeautifulSoup, nltk, dplyr, Tidyverse, Linear regression, Logistic regression, SVM, LDA, KNN, Random Forest, XGBoost, Bagging, Randomized Control Trials, Causal Analysis.
- **Big Data Stack:** HDFS, Hive, Spark, AWS - EC2, EMR, S3, RedShift, DynamoDB, Lambda, Teradata, NoSQL databases – MarkLogic.
- **Frameworks/Tools:** Jupyter, RStudio, Qlik Sense, QlikView, MicroStrategy, Power BI, SAS Viya, Tableau, Toad
- **Certifications:** Oracle Applications Certified Specialist in R12 - AR, AP and GL modules, Coursera Machine Learning – Stanford University, Algorithms & Data structure in Python by Udemy.



Professional Experience

Senior Business Technology Analyst, ZS Associates, Evanston, IL, USA

Apr 2019 – Till Date

- Analyse physician/account data and define KPIs, metrics, charts, complex expressions in Qlik Sense dashboards.
- Designing a custom-built Targeting Analytics platform using AWS data lake for call plan and incentive calculations.
- Optimizing existing QVDs, data models and load scripts to improve performance of the dashboard.
- Coordinating with multiple teams across globe for R&D new product launches and drive business requirements and deliveries.
- Leading the automation for quality checks of the dashboards across franchises to optimize time and effort.
- Presenting weekly status reports and project updates to key stakeholders and senior management on a regular basis.
- Building prototypes to enable self-service and customized reporting within the Qlik Sense application dashboards.
- Working closely with industry leadership on building strategies, assessments and roadmaps for Big Data modelling.

Data Architect, Intern, CCC Information Services, Chicago, IL, USA

May 2018 – Dec 2018

- Built Serverless Architecture using AWS features - CloudFormation, Data Pipeline, S3, Lambda, Athena, using JSON/Parquet files.
- Predicted notice of loss for damaged vehicles through ML algorithms and visualized those using Python libraries.
- Performed Proof of Concepts (POC) on Big Data use cases and experimented with big data tools and technologies.

Research Assistant, NCPRE, Coordinated Science Lab, UIUC, IL, USA

Sep 2017 – May 2018

- Optimised data collection procedures and generated reports weekly using Advanced Excel, Tableau and Google Analytics.
- Used HTML5/CSS scripts to migrate online resources from existing website to WordPress site and managed the dashboard.

IT Analyst, Tata Consultancy Services Limited, Mumbai, India

Feb 2010 - Jun 2017

- Developed PLSQL packages, interfaces, stored procedures, functions, triggers and Oracle Application EBS Forms & Reports.
- Created SQL scripts for extracts, adhoc-requests, reported and analysed large data sets and modified tables/views.
- Developed performance improvement scripts using Gather Stats, TK Prof, Explain Plan improving 70% of the query run time.
- End to end implementation of Oracle Application EBS R12 projects in Finance and Supply Chain modules.
- Team Lead for global support & development projects and worked across 16 APAC and 5 EMEA locations.



Academic Projects

<https://github.com/Paulamiray>

- **Green Taxi Tip Predictive Analysis** Nov 2018 – Dec 2018
Predicted whether a commuter will tip or not while travelling in the green taxis and estimated the percentage of tip.
Language/Tools: Python ML libraries, Tableau
- **Course Explorer using text mining features** Feb 2018 - April 2018
Created Course explorer for the university to predict the subject choices by web scraping and text mining algorithms like LDA. **Language/Tools:** Python ML libraries
- **New York City Bike Visualization** Feb 2018 - April 2018
Visualised the popular stations, days of the week having heavy traffic, start & end stations of city bikers in New York.
Language/Tools: Python libraries like matplotlib, bokeh, folium maps, bqplot
- **Eco-Bins in Smart Cities Using IoT application** Nov 2017 – Dec 2017
We handled the issue of waste management through the applications of IoT device which would efficiently notify about the fill level of the bins through sensors. **Language/Tools:** Raspberry Pi3, IR Sensors, Python, Tableau, Microsoft Office PowerPoint, Word, Excel.
- **Missing Value Prediction in a Gene Set by Parabolic Curve Fitting Algorithm** Dec 2008 – Jun 2009
We investigated automated methods for estimating missing data in gene expression data sets by using parabolic curve-fitting algorithm. **Language/Tools:** C++, UNIX, Microsoft Office PowerPoint, Word.

Poster Presentation: TED Talks–Predictive Analysis Using Classification Algorithm at the Research Showcase, University of Illinois, Urbana Champaign May 2018



Leadership Activities & Accomplishments

- Winner in “Go Green in the City-2018” Innovation Competition from North America region organized by Schneider Electric.
- Won the “Datathon” in graduate student category organized by the Illini Statistics Club at University of Illinois.
- Participated in Hackathons and Kaggle 5-day Data challenge on Statistical modelling using R and Python.
- Won the 2nd position at the “Student Data Visualization Competition”, iSchool, University of Illinois.
- Presented the innovative startup concept to the “Founders” platform at University of Illinois.
- “Best Team Player” and “Service Commitment Award” through “TCS GEMS” from Tata Consultancy Services on several occasions.
- Student Representative at iSchool, University of Illinois and given prep talks on campus recruitments, internships.
- Member of the “ZS Cares” club and volunteered for a number of activities like free software education to children, awareness spread for “Pediatric Cancer”, brown bag lunches to vulnerable neighbors in Chicago.