

Lakshmi Venkatasubramanian

Greater Seattle Area | lakshmi2688@gmail.com | [Linkedin](#) | [Portfolio](#) | [Github](#) | 626-487-8857

SUMMARY

Overall experience spanning Data Science, Test Engineering and Data Analytics space, managing, supporting and driving Customer Insights, Business Decisions, User Experience and Data projects end to end

EDUCATION

- **Master's in Data Science**, University of Washington, USA (GPA 3.91/4) (2019-2022)
- **Bachelor of Engineering in Electrical and Electronics**, Anna University, India (GPA 3.7/4) (2005-2009)

SKILLS

- **Programming/tools:** SQL, Java, Python, PySpark, R, JavaScript, HTML, CSS, Git, Linux terminal, Extract Transform Load(ETL), AWS
- **Machine Learning:** Regression, Classification, Clustering, Principal Component Analysis, Sentiment analysis, Anomaly detection, hands-on using Scikit-learn and Spark ML package, NumPy, Pandas, PyTorch, Modelling, Predictive Analytics, Statistical Data Analysis
- **Databases:** Oracle SQL, HANA Studio, AWS Athena, MongoDB, SQL, SAP HANA Studio, SAP Data Objects, Cassandra, Data Warehousing
- **Experimentation & Statistics:** Probability distributions, A/B testing, Hypothesis Testing, ANOVA
- **Data Visualization :** Tableau, Power BI, Excel

WORK EXPERIENCE

Analytics Engineer Intern Sep 2021 – Current
Tesla Redmond, WA

- Automate metrics calculations based on contractual obligations using **Python** and **PySpark** and troubleshoot data inconsistencies.
- Create tables in production, schedule and monitor workflows through **Airflow** DAGs to analyze performance of energy storage products.

Machine Learning Engineer Intern Jun 2021 – Sep 2021
Intuit Redmond, WA

- Built Customer Churn **Binary Classification model** using **Spark MLlib**, **Databricks MLflow** and evaluated the best model based on F1-scores.
- Evaluated end to end **AutoML** solution using **Databricks** and **H2O** to support propensity models to end customers leveraging **AWS** infrastructure from building model to deploying in production and generating an API endpoint

Senior Analytics Engineer in Test Jun 2019 – Mar 2021
Puget Sound Energy Bothell, WA

- Led testing efforts of Data Analytics projects by validating Extract Transform Load(ETL) workflows and **Data Science** algorithms using **Python** and **SQL** which improved data-driven decision-making capabilities for End Customers by 80%.
- Achieved 90% improvement in the overall test approach and test strategy by adopting **Agile** best practices, improving User Experience and mitigating project timeline constraints.
- Validated **Power BI** dashboards and performed data validations in **Athena** and **SAP HANA** studio.

Senior Analyst/ Program Manager Dec 2017 - Aug 2018
Microsoft (via Tata Consultancy Services) Redmond, WA

- Spearheaded a cross-functional Partner Rebates project and built a team of 8 from scratch, using **Program Management**, **Leadership** skills, that uncovered major gaps in the Rebates Eligibility program saving thousands of dollars in revenue.
- Transformed ambiguous requirements into well-defined actionable tasks by debugging **Stored Procedures**, analyzing data using **SQL** and **Python**, meeting with Business experts, and reduced the team's effort by 50% and increased the productivity by 90%.
- Supported by writing **technical specs** and handled **adhoc data investigations**

Software Development Engineer in Test Jul 2016 - Nov 2017
Liberty Mutual (via Tata Consultancy Services) Seattle, WA

- Maintained development and supported Agile teams for User Interface and Web applications using **SOAP UI**, **Java**, **Groovy**, and **Selenium** framework that resulted in reduction of execution time and manual efforts by 60%.
- Provided visibility on the overall **Product Quality** to Stakeholders by timely reporting of metrics such as **Entrance/Exit criteria**, **User Experience**, **Defects**, **Downtime**, **Automation coverage** to aid in product release.

Lead Analyst Mar 2010 – Jun 2016
Kaiser Permanente (via Tata Consultancy Services) Pasadena, CA

- Developed code for HR Payroll and Health care benefits projects using **Java**, **SQL**, **DB2**, **Cobol**, **JCL**, **Proc**
- Built tools using **Rexx** and **Java** Automation framework to automate redundant tasks.
- Established and configured testing environments for System, Parallel, UAT testing and dress rehearsal

PROJECTS (Selected projects on [Github](#))

- **[Streaming Platforms Comparison](#)**: The goal of the [project](#) is to create an interactive visualization using Tableau that can be perceived as a go-to marketplace for comparing popular streaming platforms on various metrics such as Genre, content type, average IMDb rating.
Skills used: Exploratory Data Analysis, Data Cleaning, Tableau Visualization
- **[Analysis of COVID impact on US Households](#)**: The goal of this [analysis](#) is to gauge the impact of the pandemic on overall household and answer the research questions using **Regression Analysis**, **K-means Clustering**, **PCA**, **Statistical analysis**, **Hypothesis Testing**.
- **[Sentiment Analysis and Bias in Data](#)**: The goal of this [project](#) is to predict the sentiment of Wikipedia discussion comments and identify any sources of bias that may exist in the datasets using **Exploratory Data Analysis**, **Predictive Modelling**, **NLP toolkit** and **Naïve Bayes**.
- **[Illicit transactions detection in the Bitcoin blockchain](#)**: The goal of this [project](#) is to compare the performance of **Supervised** and **Unsupervised techniques** to detect illicit bitcoin transactions on Elliptic dataset

ACHIEVEMENTS

Second place in Husky AI Hackathon 2020 out of the 12 teams and contributed towards the design of the User Interface.