

LAKSHMI VENKATASUBRAMANIAN

(626) 487-8857 | lakshmi2688@gmail.com | [Linkedin](#) | [Portfolio](#) | [Github](#) | Redmond, WA

EDUCATION

University of Washington, Seattle, WA **Master of Science in Data Science** (GPA: 3.9 / 4.0) Anticipated Grad date: **Mar 2022**
Coursework: Machine Learning, Data Visualization, Human-Centered Data Science, Probability and Statistics, Applied Statistics and Experimentation Design, Scalable Data Systems

Anna University, Chennai, TamilNadu **Bachelor of Engineering in Electrical and Electronics**(GPA: 3.7 / 4.0) **Jun 2005 – May 2009**

EXPERIENCE

Meta, Redmond, WA **Graduate Student Researcher (Capstone)** **Nov 2021 – Present**
● Evaluate change point detection in Time Series using inhouse packages on open-sourced mobility datasets and propose recommendations to improve the performance of algorithms, if any.

Tesla, Redmond, WA **Analytics Engineer Intern** **Sep 2021 – Present**
● 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

Intuit, Redmond, WA **Data Scientist/Machine Learning Intern** **Jun 2021 – Sep 2021**
● 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
Worked on building model and deploying the model generating an API endpoint.

Puget Sound Energy, Bothell, WA **Senior Analytics Engineer** **Jun 2019 – Mar 2021**
● Led Data Analytics projects by validating Business 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, Cassandra & SAP HANA studio to verify data correctness and completeness.

Microsoft (Consultant), Redmond, WA **Senior Analyst/Program Manager** **Dec 2017 - Aug 2018**
● Spearheaded a cross-functional Partner Rebates project and built a team of 8 from scratch, using Program Management, leadership, that uncovered major gaps in the Rebates Eligibility program saving millions 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 writing technical specs and handled adhoc data investigations.

Liberty Mutual (Consultant) , Seattle, WA **Software Developer in Test** **Jul 2016 - Nov 2017**
● 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, Automation coverage to aid in product release.

Kaiser Permanente (Consultant), Pasadena, CA **Lead Analyst** **Mar 2010 – Jun 2016**
● Developed code for HR Payroll and Health care benefits projects using Java, SQL, DB2, Cobol, JCL, Proc
● Built tools using Rexx to automate redundant tasks & optimized SQL queries to meet the SQL standards and improved performance by 80%.

ACADEMIC PROJECTS

Analysis of COVID impact on US Households (Skills: *Regression, K-means Clustering, PCA, Statistical analysis, Hypothesis Testing*) **Dec 2020**
● Developed Logistic, Ordinal Logistic regression models and performed statistical analysis to evaluate the relationships between variables, developed K-means clustering in PCA space to identify patterns based on certain characteristics pertaining to COVID.

Sentiment Analysis and Bias in Data (Skills: *Exploratory Data Analysis, Predictive Modeling, NLP toolkit and Naïve Bayes*) **Jan 2021**
● Developed a Naïve Bayes model to predict the sentiment of Wikipedia discussion comments and to identify any sources of bias that may exist in the data using Exploratory Data Analysis and visualization.

Illicit transactions detection in the Bitcoin blockchain (Skills: *Supervised and Unsupervised techniques*) **Jul 2021**
● Compare the performance of Supervised (Logistic regression, Random Forest and XGBoost) and Unsupervised techniques(PCA,K-means clustering, DBSCAN, Kernel PCA) to detect illicit bitcoin transactions on Elliptic dataset.

Streaming Platform comparison (Skills: *Exploratory Data Analysis, Data Cleaning, Tableau Visualization*) **Oct 2020**
● Created 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.

AWARDS AND HONORS

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

TECHNICAL SKILLS

Languages:	Python (NumPy, Pandas, Scikit-learn, matplotlib, seaborn, Spark MLlib), SQL, R, Java, HTML, CSS
Frameworks and Tools:	Spark, Git, AWS, Selenium, Docker, Databricks, Snowflake, Hive
Machine Learning:	Regression, Classification, Clustering, PCA, SVM, Decision Trees, Data Mining, Data Analysis, Modeling
Statistics:	Probability distributions, AB testing, Hypothesis Testing, Experimentation
Data Visualization:	Tableau, Power BI, Excel