Vaishnavii Paramashivam

Greater Chicago Area • 3129348621 • vaishnavii.pp@gmail.com
linkedin.com/in/vaishnavii-pp/ • github.com/vaishnavii28 • vaishnavii28.github.io/

SUMMARY

- Skilled resource with **4+ years** of Professional and Academic experience in Data Science seeking Data Scientist/ Machine Learning Engineer jobs
- Proficient in **Python, R, SQL, Power BI, Tableau, AWS, Data Mining, Machine Learning, Big Data, Statistics, Software Engineering & Consulting**

EDUCATION

The University of Illinois at Chicago	Master's - Business Analytics (MS)	GPA - 3.55/4	Exp. May '21

Courses: Statistics, Data Mining, Machine Learning, Data Visualization, Advanced DBMS, Big Data, Healthcare Analytics, Time Series & Forecasting, ML Deployment

Anna University, India	Bachelor's – Computer Science (BE)	GPA - 8.09/10	May '16

SKILLS

Programming Languages	: Python, R Studio, SAS, Java, C, PHP, Javascript, CSS, MATLAB
SQL / Big Data	: MS SQL, Oracle SQL, Hadoop, HIVE, Pig, Spark, MapReduce, MongoDB
Statistical Techniques	: Hypothesis, A/B testing, Chi-Square test, t-test, Anova, Confidence Intervals, Sampling
Packages	: numpy, pandas, scikit-learn, mlr, MLlib, NLTK, seaborn, TensorFlow, SparkML, caret, ggplot2, PyTorch
Data Viz./BI Tools	: Tableau, Power BI, Jaspersoft
Other Tools	: AWS, Docker, JIRA, Git, Clarity, Jenkins CI/CD

CORE SPECIALIZATIONS

- Data Science
- Data Structures & Algorithms
- Machine Learning
- Statistical Analysis
- Software Engineering
- Data Pipelining
- Big Data
- Deep Learning

PROFESSIONAL EXPERIENCE

University of Illinois College of Dentistry, Chicago, IL - Graduate Assistant (Data Analytics)   **June 2020 - Present**
Tools & Technology: Python, NLP, Tableau, Time Series Forecasting

- Successfully implemented analytical initiatives and roadmaps for the IT department - web-scraped large unstructured data from ticketing tool, performed data wrangling, feature engineering & formulated datasets. Applied NLP, LDA and ML models to derive operational insights, that effectively identified business process inefficiencies and reduced inventory costs by 28%
- Devised a Time Series Model to forecast the weekly, monthly & yearly demand of server usage which reduced server operating costs by 12%

Robert Bosch, India – Software Engineer   **Nov 2017 – June 2019**
Tools & Technology: Python, MathWorks Polyspace, Statistical Analysis, SQL, Tableau

- Developed statistical data models and machine learning algorithms from scratch to predict the severity of run-time errors in engine software
- Designed and developed the Software Preparatory Tool with Corporate Research team which aims to integrate all the interfaces to analyze engine ECU software for Gasoline and Diesel Automobiles
- Successfully scrapped manual data entry time by developing SQL Database modules to store engine run-time error data from Mathworks Polyspace
- Automated assembly language replacement in engine ECU software thereby reducing the operational cost by 34%
- Spearheaded data analytics team of 4 that performed run-time error analysis using descriptive statistics, exploratory data analysis and delivered detailed reports to customer engineering and leadership teams to derive key business decisions

Verizon Data Services, India - **Software Engineer**   **July 2016 – Nov 2017**
Tools & Technology: Clarity PPM, SQL, XML, Jenkins, Selenium, Jaspersoft, PowerBI

- Expertise in creating business workflows and streamlining processes based on business needs using Clarity PPM, SQL, Gel script & XML
- Automated manual processes using Selenium Web Driver for each business unit model in Verizon and integrated them with the Jenkins CI/CD pipeline, thus reducing the overall lead time of final delivery by 35%
- Designed periodic dashboards, data driven and ad-hoc reports in Jaspersoft and Power BI, to monitor KPIs and identify performance bottle necks
- Highly appreciated & recognized with '**Best Performer'** award in department for quickly learning skills & contributing to crucial tool migration

ACADEMIC PROJECTS   August 2019 - Present

Data Pipelining - Interactive Machine Learning Application
Tools & Technology: Machine Learning, Regression, Python, AWS ECR, Fargate, Docker, Flask, Dash, Gunicorn, Git

- Designed & deployed a screening tool in production, that self-serves a logistic regression model to identify patients at risk for Chronic Kidney Disease
- Created a data pipeline and hosted model as web endpoint in AWS Fargate to be consumed by targeted audience, saving cost & hours of manual work

Machine Learning - Accident Risk Prediction Model
Tools & Technology: Python, R, Tableau, Decision Trees, kNN, Multinomial Logistic Regression, SVM

- Built an "Accident Risk Prediction" model using algorithms like Decision Trees, kNN, Multinomial Logit Regression and SVM to predict the severity of accidents on traffic in the US, Hyperparameter tuning in SVM improved F1-Score to 80% at 100% precision
- Extracted 50k+ real-time traffic records from traffic streaming APIs, performed descriptive statistics, exploratory data analysis, missing value treatment, outlier detection, data cleaning, manipulation and normalization, identified hot-spot accident locations and key factors contributing to severe accidents

Data Mining – Propensity to Respond Model
Tools & Technology: Machine Learning, SMOTE, Python, R, Microsoft Excel, LASSO Regression

- "Propensity to Respond" Model was developed to predict the likelihood of consumer behavior at VWWare by applying ML algorithms like Random Forest, Gradient Boosting, Ridge Regression and Model Stacking, XGBoost model achieved a Micro Average F-Score of 100%
- Feature reduction was done from 707 to 90 using LASSO Regression, SMOTE technique was used to address class imbalance problem

CERTIFICATIONS

Tableau-Desktop Specialist || Tableau - Data Scientist || AWS Associate Developer – Udemy || HackerRank SQL Gold Badge || German Language - A2 || ISTQB Certified Agile Tester