DIVYA PATTISAPU

divyapattisapu.vercel.app | divyapattisapu@gmail.com | +1 773-219-4003 | linkedin.com/in/divyapattisapu

SUMMARY

Experienced Engineer from IIT Bombay and University of Chicago, specialized in Distributed Data Processing, Backend Development and Cloud Engineering

EDUCATION

The University of Chicago *Chicago, IL*
Master's Program in Computer Science *Mar 2024*
Coursework: Cloud Computing, Algorithms, Machine Learning, Distributed Systems

Indian Institute of Technology, Bombay *Mumbai, India*
Bachelor and Master of Technology in Mechanical Engineering *Aug 2021*
Teaching Assistant - Engineering Data Mining & Applications: Instructed 200+ students, Designed & evaluated their assignments

SKILLS

Languages	: Python, SQL, C/C++, Go
Cloud Services	: Amazon Web Services (EC2, S3, Glacier, SQS, SNS, ELB, Lambda - Serverless, CloudFormation, DynamoDB)
Data Frameworks	: Apache Spark, Hadoop/HDFS, Apache Hive, PostgreSQL, MS SQL Server
Web frameworks	: HTML, CSS, JavaScript, React.js, Next.js, Bootstrap
Libraries	: FastAPI, Flask, Pytest, Tensorflow
Machine Learning Models	: Classification, Clustering, Natural Language Processing

EXPERIENCE

MasterCard *New Delhi, India*
Big Data Engineer *Jul 2021 - Jul 2022*
- Optimized the performance of production forecasting models running on Spark and Hive in the inference pipeline
- Built automated data validation and quality check pipelines for continuously observing and validating production runs
- Performed brand analytics to select relevant customer groups for client advertisement campaigns in a customer facing Digital Marketing team
- Developed a customer segmentation model using KMeans model to fuel the strategic debit portfolio enhancements and targeted campaigns
- Built automated pipelines for competitor analysis & market research for campaign targeting
- Developed a PySpark application to carry out correlation analysis between audience segments to identify cross-selling opportunities
- Core Team Member, Girls4Tech: Hosted collaborative learning workshops to motivate young girls to build STEM careers

University of Chicago Professional Education *Chicago, IL*
Data Analytics Intern *Feb 2023 - Mar 2024*
- Developed KModes model to create target segments with observed student admission decline patterns for retention campaigns by performing segmentation models; Used these outcomes to identify key contributors for decline in each segment using Decision Tree Classifiers
- Developed an in-house pipeline for campaign reports on LinkedIn Ads and Google Analytics using their APIs saving dollar cost
- Created an ETL pipeline for student grades, enrollment and admission, leveraging dbt, Trino and MS SQL Server
- Built and automated dashboards on Tableau to present key insights to partners, facilitating data visualization and quantitative analysis
- Identified factors contributing to accepted, denied or declined applications using Decision Tree Classifier and Logistic Classifier models

PROJECTS

Genomics Annotation Service – AWS Cloud Computing
- Developed a SaaS application for file upload, processing and retrieval with job status view and free/premium user file storage features
- Implemented a tiered storage serverless archival process to transfer free users' result files from S3 to Glacier to reduce the storage cost incurred
- Integrated the application with a Stripe payment system and included a notification system to inform the users of their job completion

Microservices-Powered Auction Website
- Designed an eBay-modeled auction website in a team of 5 using Agile methodology ensuring robust integration via API testing
- Developed a full-stack solution using Postgres for relational database modeling, React.js for frontend and FastAPI for the backend
- Integrated the services using RabbitMQ for communication, and ensured robust integration via API testing

Parallel Image Processing Package using GoLang
- Designed and implemented a versatile image processing package in Golang, with a command-line interface to input images
- Implemented a map-reduce algorithm for parallel image processing and a partitioning mechanism based on a specified attribute
- Integrated a work-stealing algorithm to opportunistically steal work from busy threads, improving the speedup by 10%

Distributed Messaging Queue
- Implemented an object-oriented distributed message queueing system like RabbitMQ using Flask with quorum-based consensus for replication
- Conducted rigorous testing to validate fault-tolerance mechanisms, ensuring 100% code coverage at each development iteration

Analyzing Customer Behavior towards Electric Vehicles
- **Publication objective:** To identify the contributing factors to EV buyer readiness across cities in India to improve its adoption (link)
- Curated a 35-feature survey dataset, applying a variety of classifiers including Logistic Regression and Deep Learning
- Achieved 72% accuracy using Logistic Classifier and analyzed the contributing factors using Shapley Value & Relative Weight Analysis

Other Projects
- Developed an AI Checklist **website** which creates customized checklists with support for subtask creation using Next.js, Figma and GPT API
- Built and compared two real-time messaging apps, one using Restful API (React.JS & Flask) and the other using ZeroMQ