Prabhu Subramanian

Union City, CA | 617-982-4335 | prabhus165@gmail.com | www.prabhusubramanian.com | github.com/prabhuSub | linkedin.com/prabhu-subramanian

EDUCATION

Northeastern University, Boston, MA   **September 2018 - December 2020**
Master of Science in Information Systems

TECHNICAL SKILLS

Programming languages: Python, SQL, HTML, CSS, Java, JavaScript
Database: MySQL, Microsoft SQL Server, PostgreSQL, Oracle, Snowflake, Arora, Redshift, DynamoDB, Teradata
ETL & Tools: Alteryx, Talend, Athena, Putty, Visual Studio Code, Docker, Kubernetes, Terraform, Jenkins, Atlantis
Cloud Engineering: Google Cloud Platform, Amazon Web Services, Microsoft Azure, Databricks
Business Integration: Tableau, Microsoft Power BI, Excel, SAP, BI HANA
Version Control: Git, GitHub, Bitbucket, SVN Tortoise

EXPERIENCES

Data Analyst, Levi Strauss & Co., San Francisco, CA   **April 2021 – Present**
- Provided guidance as a Subject Matter Expert (SME) for critical reports, and aligning with business objectives for over 500 users.
- Innovatively designed and executed multiple Proof of Concepts (POCs) using Python, Alteryx, and ETL techniques, enhancing operational efficiency.
- Developed an API to bridge the gap between Alteryx and SharePoint, enabling comprehensive reporting analysis and saving approximately 15 hours of manual data transfer per week.
- Leveraged PySpark on Databricks to automate reporting solutions and visualize data on Tableau/PowerBI for business users, leading to a 30% increase in data accessibility and insights. Led the development of GDO and Databricks ETL pipelines, ensuring seamless operation of reporting solutions, resulting in a 40% reduction in data processing time.

Graduate Programmer Analyst, SquarkAI, Boston, MA   **January 2019 – December 2020**
- Enhanced model accuracy by an impressive 70% through the application of advanced NLP techniques such as word2vec & BERT, resulting in more precise predictions and decision-making processes.
- Developed production code integrating the MinIO framework with AWS S3, significantly streamlining data storage and retrieval processes, leading to a 50% reduction in data access time.
- Formalized subprocesses to execute the JVM H2O engine, effectively optimizing computational performance and resource utilization, resulting in a 40% increase in processing speed. Developed and devised a multiprocessing and multithreading solution to enhance efficiency further.
- Implemented an automation script using Python and CLI, dramatically reducing developers' testing time by an outstanding 90%, thereby accelerating the development cycle and time-to-market.

Data Engineer Intern, Peterbilt Motors Company (PACCAR), Denton, TX   **January 2020 – June 2020**
- Implemented scripts to preprocess and optimize the speed by an impressive 50% for cleaning customer warranty data using Python & SQL, resulting in a significant reduction in data processing time and operational costs.
- Enhanced NLP model accuracy by implementing custom spelling dictionaries and integrating a specialized manufacturing plant failure modes word corpus, resulting in a 20% increase in prediction accuracy.
- Administered data warehouse pipelines from various sources including Snowflake, Teradata, and Excel, deploying CI/CD pipelines using Jenkins and CloudFormation.
- Incorporated Flask application for the internal team's interaction with the data warehouse.
- Deployed pipeline using Docker, Kubernetes (EKS), and AWS Services - SNS, SQS, SES, S3, EC2, Lambda.

Software Engineer, Tata Consultancy Services - Hewlett Packard Indigo, Pune, India   **November 2015 – July 2018**
- Designed and implemented Enterprise Resource Planning (ERP) dashboards catering to Supply Chain, Manufacturing, and Finance departments, resulting in a 20% increase in supply chain visibility, a 15% reduction in manufacturing downtime, and a 25% improvement in financial forecasting accuracy.
- Achieved a remarkable 68% improvement in application query optimization, leading to a 30% reduction in report generation time and a 25% increase in overall system performance, resulting in significant cost savings and enhanced productivity.
- Improved Customer Self-Service operations by 30%, leading to a 25% rise in satisfaction ratings and a 20% drop-in ticket resolution time, showcasing leadership in operational excellence and customer focus.

PROJECTS

WORD2VEC – Model Interpretability (H2O, NLP, Visualization)
- Orchestrated an in-depth comparison of findings against BERT model; skillfully merged TensorFlow and TensorBoard.
- Pioneered writing a GitBook on model interpretability based on the above analysis - subramanian-pr.gitbook.io/research-paper.

H2O Hyperparameter Recommendation System (Text Analytics, Feature extraction, Feature Selection, Feature Engineering)
- Trained H2O AutoML on various datasets. Scrapped and stored the hyperparameters of the best model from the leaderboard.
- Developed hyperparameters recommendation system using the data to train the model.

Kaggle Competition - Kobe Bryant Shot Selection (H2O, AutoML, Kaggle, Visualization, XGBoost)
- Performed EDA on the data of Kobe's shot on the field and built a predictive model to predict successful shots given the position.
- Decreased model Log Loss from 0.60 to 0.34 by hyperparameter tuning. Secured 1st rank on the leaderboard.

SERVICES AND LEADERSHIP AWARDS

- Awarded company's leadership value holder certificate and badge at TATA Consultancy Services training department (2015).
- President, and Co-Founder of NEU AI Skunkworks (Jan 2019 – Dec 2020).
- Graduate Student Ambassador & accoladed as a best TA for Information Systems at Northeastern University (2019).
- Awarded as Global MPIM "Innovation Champion" at Levi Strauss & Co among 10 other nominees (2023).