Do Hyeong Kim   Software Engineer

Sunnyvale,CA   718-687-3315   dhk371@nyu.edu   linkedin.com/in/do-hyeong-kim | brightdo.github.io | github.com/brightdo

PROFESSIONAL EXPERIENCE

Big Data Software Engineer (System Title: Software Engineer 2) | Paypal   **August 2022 – May 2024**

- Pioneered development and maintenance of Paypal's model variable monitoring tool. Employed Python and spark to extract crucial metrics, and facilitated improved data science solutions and actionable insights through early identification and investigation of problematic variables.
- Orchestrated strategic change in backfill job execution, reducing costs by 12% through optimization of cluster instances and centralized event time processing, leading to more efficient data handling.
- Designed and Spearheaded variable categorization etl spark pipeline improvement project, achieving remarkable 20% reduction in incorrect categorization rates by flagging and storing high chance wrong categorization variable and re categorizing it in future when more detail of variable is accumulated.
- Located and investigated into increase in model variables' missing values by utilizing monitoring tool. Isolated problem by going through variable logic and communicated with responsible team to fix issue based on findings.
- Reduced cost of creating and maintaining backup of Database server by taking advantage of InfluxDB's backup and restore API features.
- Proposed and implemented technique to resolve front-end rendering issue due to large data volume by aggregating multiple metrics to decrease data requirements without compromising monitoring efficiency.

SOFTWARE ENGINEER INTERN | Paypal   **June 2021 - July 2022**

- Produced a data pipeline with Spark, and Big Query to process millions of rows from various sources to provide merchants' funding instruments and transaction approval rates for Smart Wallet project.
- Investigated and resolved a critical bug in audience insight project, ensuring all merchants receive accurate web visitor insights. Identified root cause from source table and implemented a solution to ensure data integrity.
- Engineered a crucial data pipeline pivotal in calculating metrics to showcase effectiveness of a PayPalproject within team.

ENGINEERING INTERN | IpserLab   **February 2021 - May 2021**

- Developed a secure user authentication system/backend for client using NodeJS and MySQL. Ensured security through hashingand salting with Bcrypt library.
- Proposed and executed revamped signup/login page, collaborating with other interns to engineer cleaner UI.

EDUCATION

Master of Science in Software Engineering, San Jose State University, May 2022
Bachelor of Arts in Computer Science, New York University, May 2019

PROJECTS

Fake Tweet Classifier (Python)

- Preprocessed dataset and trained variety of classification models(multinomial Naïve-Bayes, Random Forest, Decision Tree, SGD, KNN, and CNN) for performance comparison of fake tweet detection.

Distributed Key-Value Data store with Consistent Hashing (ZeroMQ, Python, Consul)

- Composed a distributed key-value data store with Consistent and HRW hashing adopting ZeroMQ as transport protocol.
- Rebalanced data when node is add or removed and dynamically controlled node membership with Consul.

Starbucks App (JAVA)

- Constructed a text based Starbucks Mobile App Simulator leveraging design patterns such as State, Composite, Chain of Responsibility, Command, Singleton and Observer. Generated module with minimal coupling, making code easier to understand and maintain.

CORE COMPETENCIES/TECHNICAL SKILLS

Core Expertise: Algorithm & Data Structures, Software Architecture, Design Pattern, machine learning

DevOps: Testing & Debugging, Continuous Integration/Continuous Deployment (CI/CD), Agile, Data Mining

Soft Skills: Problem-Solving, Communication, Team Collaboration, Time Management, Leadership, Attention to Detail, Creativity

Programming Languages: Java, Python, C, Objective C, Assembly, Pyspark, Scala

Web Development: HTML, CSS, NodeJS, JavaScript, Express, AJAX, JSON, Jira, Mongoose, PHP, jQuery

Database Management: MongoDB, MySQL, SQL, InfluxDB, Redshift,PostgreSQL, NoSQL,

Big Data & Analytics: Spark, Hive, Hadoop, Tableau, Data Mining, ML Recommendation Model, AWS EMR,AWS RDS, Data Warehouse, AWS Glue

Operating Systems: Windows, macOS, Linux

Development Tools: Vim, UML, GitHub, Airflow, Docker, Kubernetes