

HIMANSHU BOBADE

P: +91 705-700-8846 | Email: himanshubobade007@gmail.com
linkedin.com/in/himanshu-bobade | github.com/himanshubobade

EDUCATION

UNIVERSITY OF MUMBAI - Bachelor of Engineering in Computer Science
Bharati Vidyapeeth College of Engineering (CGPA : 8.25/10)

Mumbai, India
Aug 2018 - May 2022

PROFESSIONAL EXPERIENCE

JIO PLATFORMS LIMITED

Data Engineer (Assistant Manager), Data Platforms

Mumbai, India
Jul 2022 – Apr 2024

Business Use Case Engineering:

- **Optimized JioFibre Service Efficiency:** Streamlined a 52-step tracking process by integrating real-time monitoring for technician service status. Engineered a data pipeline with Spark, Scala, Kafka, Airflow, APIs, Hive, and MySQL, ensuring seamless synchronization across 5 disparate systems with CI/CD testing. Enhanced service resolution efficiency, impacting 4.3 million users with real-time tracking and proactive alerts.
- **Live Device Theft Prevention:** Built a real-time Geofencing alert system to detect unauthorized device movement across geographical regions. Leveraged Spark, Hive, Kafka, Scala, and Airflow to process live location updates, reducing device theft incidents and enhancing security monitoring.
- **IMSI Location-Based Customer Segmentation:** Developed a customer profiling solution by tracking day and night locations using public Kafka streams. Integrated NiFi, Kafka, Spark, Scala, Airflow, and Hive to process and analyze location data, enabling targeted premium service offerings based on user affordability and behavior insights.
- **Enterprise Mobility Data Engineering:** Engineered scalable data pipelines for Network Tower Optimization, GIS, JPW, OrderCare, and WorkOrder, capturing KPIs to support enterprise-wide decision-making. Employed Spark, NiFi, Scala, Hive, MySQL, Kafka, and Airflow to aggregate and process critical mobility data. Additionally, directed migration efforts from HDP to CDP, ensuring seamless system transition and improved performance. Implemented metadata tracking and data lineage frameworks to boost data governance and compliance.
- **User Engagement Tracking for MyJio:** Developed an analytical framework to trace user login/logout activity on the MyJio app. Automated data ingestion and processing using Shell scripts, NiFi, Hive, and Airflow, providing actionable insights into user retention and engagement patterns. Combined master data validation checks to improve data consistency and quality.

Cloud-Based Platform Engineering:

- **Azure-Based Data Quality Assurance:** Designed a generic Data Quality solution applicable to Azure pipelines utilizing Dataflows and Copy Activity. Automated anomaly detection by implementing Z-score calculations on SQL Pool data, improving data reliability through daily validation processes. Integrated data governance policies to ensure master data consistency and regulatory compliance.
- **GCP Data Quality & Monitoring Framework:** Led a proof of concept (POC) to replace Azure and on-premise systems with GCP-based solutions. Developed data quality monitoring solutions for Dataproc (using Spark Listeners), DataFusion (via CDAP API endpoints), and DataFlow (leveraging Cloud Monitoring Metrics). Identified and reported a critical Apache Spark listener bug, resulting in an open-source fix by GCP. Centralized monitoring across multiple projects and visualized insights using Looker Studio.
- **Hybrid Cloud Data Pipeline:** Engineered a high-performance Apache Flink-based real-time data pipeline to bridge on-prem Hive and Oracle MySQL via Azure Synapse. Designed a 24x7 running Azure Notebook to process and sync Kafka-based data, achieving a peak processing capability of 500 million records daily. Executed robust checkpointing mechanisms to ensure fault tolerance and data integrity.
- **Data Completeness:** Developed a solution to assess data completeness for kafka producer and consumer to ensure no data leakage in the pipelines.

GOOGLE DEVELOPER STUDENTS' CLUBS – GOOGLE

Technical Lead

Mumbai, India
Aug 2021 – Jun 2022

- Organized, facilitated and orchestrated 12+ technical seminars, workshops and hands-on sessions on multiple domains involving web and app development, machine learning and cyber security, managing a team of 6 professionals along with creating projects.

- Led a 3500+ student community, fostering collaboration and engagement, growing community across social media platforms. Organized sessions with entrepreneurs, providing students industry exposure and career opportunities, resulting in a staggering 330% increase in community strength.

UNRUFFLED FEATHERS ENT. PVT. LTD

Remote

Data Science Intern

May 2021 – Feb 2022

- Applied TensorFlow and ML models (YOLOv3), custom similarity matrix (87% accuracy), CNN, object detection, eye heat map visualization, context (text/audio) summarization, & brand-logo detection models, image captioning.
- Led 20+ interns delegating tasks and deploying ML workloads with a revolutionary context-based adTech recommendation.

THE TANN MANN GAADI FOUNDATION

Remote

Flutter Developer Intern

Jul 2021 – Jan 2022

- Constructed Frontend and Backend Development with Flutter Bloc MVC framework with Google maps APIs integration for live location tracking.
- Accomplished unique face recognition and user token model along with Blockchain wallet integration.

MADRAS SCIENTIFIC RESEARCH FOUNDATION

Remote

Research, Data Science & Content Authoring Intern

Feb 2021 – May 2021

- Researched, designed algorithms and authored technical articles on Virtual Voice Assistant deploying Deep Learning and NLP, Roles of AI and ML in GNSS Positioning, Image Captioning, Diabetes Prediction using six different classification algorithms.

RISING YOUTH FOUNDATION

Remote

PHP Backend Development Intern

Oct 2020 – Nov 2020

- Coded PHP MVC framework with MySQL to visualize student ambassador progress data dashboards for a non-profit organization.

ACHIEVEMENTS

- **Coding Competitions:** Google Hashcode 2022 (Rank- 1773), Google Hashcode 2021 (Rank- 4912)
- **Programs/Fests:** Google Cloud Program 2021(Completed 11 quests and 9 skill badges from Quiklabs), DigitalOcean's Hacktober Fest 2020
- **Projects:** Flutter Voice Assistant With Realtime Object Detection, Twitter Sentiment Analysis with AWS Lambda & Comprehend, completed 50+ AI/Data Science projects.
- **Certifications/Courses:** Softtype, Inc- Oracle Netsuite - Centre of Excellence; Udacity- AWS Machine Learning Foundation, Salesforce- Einstein AI Academy Course; Coursera - Natural Language Processing with Classification and Vector Spaces, Neural Networks and Deep Learning

SKILLS

Languages/Web	Python, Scala, Shell, R, PHP, SQL, HQL, HTML, CSS, HQL
Databases/Warehouse	BigQuery, Hive, Firebase, Oracle MySQL, Oracle, Azure SQL pool, Azure serverless SQL
Tools/ Frameworks	Spark, NiFi, Hadoop, Flink, Keras, Flask, SBT, Sqoop, MVC, MVT, Kafka
Platforms	CDP, HDP, Azure Synapse, GCP (Google Cloud Platform), Heroku, AWS, Azure Synapse
Other	REST APIs, CI/CD, Airflow