

Kaustubh Daware

Academics			
Degree/Examination	Year	Board/University	%/CGPA
Post Graduate Diploma in Business Analytics	2016-2018	ISI Kolkata, IIM Calcutta, IIT Kgp	8.0
B.Tech. , Electrical Engineering	2011-2015	COEP, Pune	7.68
Class XII	2010-2011	Maharashtra State Board	90.17
Class X	2008-2009	Maharashtra State Board	92.61
Work Experience /Internships			
Data science manager	Jio Platforms Ltd.		May 2018-*
Marketing campaign based on consumer data usage	<ul style="list-style-type: none">To build targeted campaigns for Jio’s product offering based on consumer data consumptionTo create real time pipeline for data ingestion, aggregation and ML model for URL classificationBuilt pipeline using NiFi, spark streaming and Hive jobs for processing TBs of data in minutesDefined consumer brackets based on usage to aid campaign with reduced memory consumption		
Device management platform	<ul style="list-style-type: none">To build 360 view for devices connecting to Jio network and monitor their network experienceTo find out devices’ contribution in bad network experience and to explore newer business areasUsed NiFi, Shell, Hive, Sqoop, SFTP to create daily and weekly reports for Mobility PresidentIdentified models with poor experience, OEMs for enhanced collaboration, potential device buyer		
Jiophone insights	<ul style="list-style-type: none">To provide visibility of Jiophone users’ activity which constitutes roughly 25% of Jio networkTo monitor Jiophone users’ applications usage, device experience and network experienceDeveloped KPIs along with enabling business teams in Hadoop ecosystem for reduced timelinesIdentified data loss in pipeline and suggested schema changes to monetize Jiophone applications		
Real time devices ingestion(DC)	<ul style="list-style-type: none">To build real time ingestion pipeline for Jio devices capturing millions of users’ engagementTo handle frequent schema changes, corrupt records, data surge at minimum resource footprintStreaming pipeline using Spark, Kafka, NiFi with corrupt record logging, error specific treatmentsReduced timeline by identifying templates and improved data quality with inclusive process		
Internship			
Data science intern	Actify Data Labs(True North Managers LLP)		Oct 2017-April 2018
Customer support chat bot	<ul style="list-style-type: none">To develop customer support chat bot to answer potential and existing customer queriesTo identify underlying purpose and context of user queries and answer interactivelyBuilt chat bot using intent entity recognition, Python Flask with business specific workflow		
Improving customer experience for hospital chain	<ul style="list-style-type: none">To help leading maternity hospital chain improve overall degrading customer experienceConducted patient follow up exercise and built ML model to predict expected appointment timeEnhanced customer experience with reduced turnaround time and engagement suggestions		
Profit sharing simulation model	<ul style="list-style-type: none">To build interactive profit sharing simulation dashboard for PE firm to do scenario analysisTo calculate percent profit shared at each level with attrition, promotion and newcomer scenarioBuilt dynamic dashboard with R shiny to replace existing excel solution to reduce manual efforts		
Academic Projects			June 2016- May 2018
Compositional data regression	<ul style="list-style-type: none">To model effect of individual constituents’ proportion and interaction on mixture’s propertiesHandled singularity problem with compositional data with transformations and mixture modelProvided insight how each component and their interactions contributing to the yarn elongation		
Water pump classification	<ul style="list-style-type: none">To predict functional status of water pumps in Tanzania to improve maintenance of pumpsPerformed data preprocessing and exploration to derive insights and built classification modelsCorrectly categorized functional status with 82.47% accuracy of total pumps in test data		
Commodity price forecasting	<ul style="list-style-type: none">To enable agricultural decision makers by predicting prices of agricultural commoditiesCrawled Govt site to collect data and used time series and regression techniques to build modelPredicted near future prices with average error of 2.77% and long term future 5.05%		
GET	Fiat Chrysler Automobiles(FCA),R&D Chennai & Pune		July 2015-April 2016
<ul style="list-style-type: none">Ensured high quality vehicle by performing EOL(End-of-line), diagnostic and vehicle function testing and validationDeveloped comprehensive group vehicles benchmarking with current competitors' vehicles in terms of vehicle featureDocumented functions of electric component(sensors and actuators) in layman terms to help smoothen plant process			
Technical Skills and Relevant Coursework			
<ul style="list-style-type: none">Languages - SQL, shell script, R, PythonSoftwares/Tools- Hadoop, Hive, Spark, Sqoop, GG, NiFi, Kafka, MATLAB, Git, RTC, Git			