

PROFESSIONAL EXPERIENCE

MONEY VIEW (Whizdm Innovations Pvt Ltd)	Senior Data Scientist	April 2021 – Present
<i>Built, productionized and maintained state-of-the-art machine learning (ML) models for credit underwriting; Leading a team of 3.</i>		
- Stack: Python, SQL, Pandas, NumPy, Scikit-learn, Flask, Plotly, AWS EC2, AWS S3, AWS Athena, Bitbucket, Jira, Metabase • Risk - SMS Data Pipeline, Risk Prediction ML Model and Business Intelligence (BI) Reporting - In Production: - Collaborated closely with product managers and credit-risk analysts to understand analytics and modeling needs - Optimized data pipeline by writing prioritization logic for RegEx template creation to process over 5 million SMS everyday - Wrangled 50TB of SMS data stored in AWS S3 file system using Athena and Python to create 330+ features, each highlighting a distinct consumer behavior, from lifestyle identifiers to financial literacy and cash flow management - Architected scalable target encoding and normalization process to reduce null values to 0% within each feature - Simulated target variable for rejected users to replicate production-like incoming population for model training - Utilized Python for training multiple linear and non-linear supervised ML models, picked the most stable model that stayed consistent over different time and segment cohorts, had an AUC of over 72% on test and out-of-time datasets - Worked with Java developers to implement the model in production, which reduced overall cohort risk by 16% - Developed offline feature store on MySQL using Linux Crons on EC2, processing and storing for over 1 lakh users everyday - Created BI dashboard using Metabase to track macroeconomic factors of financial stress from transactional SMS data • Risk - Bank Statement Data - Implementation: Developed ML model, expected risk reduction is over 26% in the cohort • Risk - Non-transactional Alternate Data - POC approved: Designing a third-dimensional credit risk ML model • Risk - Prediction Output Data from Base ML Models - POC approved: - Simplifying interpretation of predictions from multiple ML models, each of which is based on different base data, by creating a meta-model using tree-based classifiers and different ensembling techniques, for simplifying underwriting • Fraud - Mobile Device - In Production: Increased detection rate by >450% using fingerprinting techniques on SMS data • Fraud - Identity: Building deep learning (DL) model to identify image of photograph containing face from user selfie data • BI Dashboards and Data Summarization: - Used Flask, Plotly and Dash to develop web-based dashboard with complex analytical tools like inter-database connections, cross-tabs, comparisons, filters and more for credit, risk and collections teams - Summarized user journey and performance data-tables using MySQL and Cron, which saved 2 hours per analyst per day • Leadership and Training: - Leading a team of 3 data scientists who I trained right out of college, each of whom is responsible for building automated web-based dashboards for data pipeline maintenance and model performance monitoring respectively - Trained 15 fresh graduates on 'Hands-on Machine Learning' and 'Python for Data Science'; Received 5 star rating		

PALRIGRAPHY	Freelance Software Developer	December 2019 – April 2021
<i>Single-handedly planned, built, productionalized and managed end-to-end software development for corporate clients.</i>		
• Puranmal Sons, India, manufacturer and distributor of PVC Plumbing Systems in Central India: - Designed a Customer Loyalty Program and developed web and Android app, over 2500 DAU, rated 4.6 on play store Stack: Python (Flask), HTML, CSS, JavaScript, Java, Heroku, PostgreSQL, AWS S3, Werkzeug WSGI, TextLocal API - Automate BI reporting by building interactive GIS sales dashboard using Leaflet.js, Folium, GeoPandas, Google Sheets API - Reduced receivables by >20% by building a bot that sends regular outstanding reminders to debtors using CAPIWHA API • New Escape Room Designs, LLC, USA, developer and licensor of Escape Room Game Designs: - Reduced time from 4hr+ to 2 clicks by building GIS-enabled web application for new-license territory-discovery process Stack: Google Maps API, Google Sheets API, Python (Flask), HTML, CSS, Bootstrap, JavaScript, jQuery, Heroku, PostgreSQL		

DRONAMAPS PVT LTD	Co-Founder, Director	July 2017 – March 2019
<i>Created GIS 3D maps using drones and computer vision to provide analytics to clients like Reliance Industries, Water Works Departments, Industrial Development Authorities, among others, substituting the age-old surveying techniques by innovation</i>		
• Developed product to geolocate public assets using drones and 3D maps, reducing surveying time by 3-to-6 months • Headed product development for Municipal Corporation's Encroachment and Property Tax Assessment using computer-vision and machine learning models; Projected 4x increase in revenue earnings • Led ML team to develop AgriTech product for Reliance Foundation to help sustain crop health using multispectral sensors • Interviewed 350+ candidates, hired, trained and led team of 70+ professionals and terminated 18, across organization • Awarded "Most Innovative Startup" at NASSCOM Product Conclave, Bangalore – 2017		

INDEPENDENT PROJECTS

• Developed a Fake News Classifier using TfIdf Vectorizer and Passive Aggressive Classifier – ML, pandas, scikit-learn – 2021 • Created a Spam Email Classifier using Naïve Bayes and XGBoost Classifiers – ML, NumPy, seaborn – 2021 • Designed a Cat Image Classifier using multi-layer neural network – DL, NumPy, Matplotlib, h5py, PIL, SciPy – 2020 • Developed a web app to manage portfolio of stock with real-time prices using IEX API – SQLite, Flask, HTML, CSS, JS – 2020 • Built a terminal-run card game called Declare – Python – 2020		
--	--	--

EDUCATION

SELF-LEARNING, Computer Science and Data Science (Curriculum) <i>Key courses: MIT 6.0001, MIT 6.0002, Harvard CS50, Stanford CS229, Stanford CS229A</i>	2019-2021
INDIAN INSTITUTE OF MANAGEMENT, INDORE • Master of Business Administration - Trimester abroad at EGADE Business School, Tec de Monterrey, Mexico	2015-2017
• Bachelor of Arts - Founder, Photon, Photography Club Co-founder, Atharv, Management-Cultural Fest	2012-2015

SKILLS

• Programming Languages: - Pro - Python, SQL, HTML, CSS Know - JavaScript, jQuery, Java, Google Apps Script, Excel Macros • Data Science & ML: - Pro - Scikit-learn, NumPy, Pandas, Matplotlib, Flask, Plotly, Dash, APIs, Excel, Git, Metabase, AWS S3, EC2, Athena, data science pipeline (Cleansing, Wrangling, Visualization, Modeling, Interpretation) Know - PyTorch, fastai, TensorFlow • Languages: Spanish, German Certified with A1 level DELE (Diplomas of Spanish as a Foreign Language) – 2013		
--	--	--

HOBBIES & INTERESTS

• Photography, Driving, Backpacking, Neuroscience, Science-Fiction, Gadgets, Automobiles, Artificial Intelligence, Automation		
---	--	--