**MOHIT RAJPUT**

ML | DL | AI | Data Collection | Data Engineering | ML Experiment Design

[LinkedIn](https://linkedin.com/in/mr901) | [GitHub](https://github.com/MR901) | [mohitrajput901@gmail.com](mailto:mohitrajput901@gmail.com) | +91-8979572630 | Bengaluru

|  |  |
| --- | --- |
| **Areas of Professional Interest** | * Reinforcement, Unsupervised Learning and Recommendation System |
| * EDGE Devices based ML |
| * Data Collection, Data Engineering and Experiment Design |
| * Data based Product Research and Development |
| * Deep Learning based solutions (RNN, CNN, LSTM, SOM, Auto Encoder) |
| * Natural Language Processing (NLP) and ChatBots |
| * Time Series analysis & Computer Vision |

|  |  |
| --- | --- |
| **Work**  **Experience** | **Cargill Digital Labs, Cargill**  Machine Learning Software Engineer Jun’19-Present |
| ML labs v2 setup   * Defining ways to Work with ML projects and other applications having ML related features * Brining sustainable practices for accelerated development while maintaining the agility * Planning phase of data governance, 3rd Party Data Annotators, Portal development for data tagging, selection and versioning, data collection experiment designs and pipelines. |
| Estimating Growth Indicator of Chicken (POC)   * Estimating the weight and age of the chicken based on video recording done using the mobile camera. * *Tags: Agile Experiment Setup & Design , Object Detection, Image Standardizer, Estimator, Ensembling* |
| Defective Nugget Detection and Accounting (POC then MVP)   * Defective Nugget Detection and Tracking on conveyor belt with output projected on display for manual removal. * This was to be accomplished by on-premise device and needed to be real-time. * *Tags: Jetson Nano, Raspberry Pi, Ardino Uno, Edge device, Object Detection, Tracking* |
| HR Chatbot (Myco) (MVP)   * Initial product version was launched for Singapore and Australia with capability for form filling, ticket creation, ticket checking, easy data integration from excel sheet * *Tags: Rasa, Redis, NLU, NDM, NLG, NLP, Chatbot, excel* |
| Financial Commentary Generation (POC)   * A utility for table-to-text based on template method was developed for the initial stage, b/c of the lack and too much fragmentation of data source this approach was deemed fit. |
| YOLO v5-v4 training and deployment integration on AWS suitable to multiple tailored needs.   * *Tags: Computer Vision, Sagemaker, ECR, S3, Lambda Function, API Gateway, IAM, CloudWatch* |
| Computer Vision Utility based on Streamlit   * Ability to fetch data from local, S3, and pcloud, parse data, label data, visualize annotation, and partial support of Yolo prediction. * Tags: Support for Local, S3, Pcloud, |
| MiApp feature involving weather data visualization using the GIS data. |
| Information Extraction from Ship Invoices (POC)   * Extracting multiple information through the information that was shared in old and current emails. * Tags: Text-to-Table, Analytics; Regex, Python Email to DB |
|  |
| **ShieldSquare (Radware)**  Data Scientist – R&D Apr’18-May’19  Associate Research Engineer – R&D Jun’17-Mar’18 |
| * ICLSSTA * Solely developed a framework, which develop and evaluate unsupervised learning solution in much depth and this from a single config file. * *Tags: Framework, Clustering, Anomaly, Outlier, Conceptual Drift Detection, Ensembling, Support all Dimension reduction, Anomaly, and Clustering Algorithm from Sklearn, Add. Custom Algorithm.* |
| * Adaptive Action Taking (AAT) * Developed a generalized module named which makes use of reinforcement learning to automatically take action on incoming traffic on the web property. * *Tags: Reinforcement Learning, Multi-Armed Bandit, Smart handling of Business Limits, Self-Adjusting* |
| * Deep Behaviour Analysis (IDBA) * Partially developed a module named Intent based, which utilizes LSTM to yield encoded features and scores that are used with supervised, anomaly and clustering. * *Tags: RNN, Auto-Encoder, Semi-Supervised Learning, Sequence Based Detection, Deep Learning* |
| * Threshold Online Reinforcement (ThOR) * Helped developed a module named, which makes utilizes online learning to develop a probability score distribution over of isotonic regression for action taking. * *Tags: Isotonic Regression, Online Learning, Self-Adjusting, Smart handling of Business Limits* |
| * Developed and implemented end to end machine and reinforce learning based solutions using ICLSSTA-AAT, IDBA-ICLSSTA-AAT, & IDBA-ThOR *Tags: Semi-Supervised, Supervised, unsupervised, reinforcement learning, behaviour analysis.* |
| * Titan Batch Analyser * Developed and Implemented, which utilizes multiple browser signatures, behaviour based rules, regular expressions and databases to generate a suspiciousness score to take action on traffic. * *Tags: SQL & Python Implementation, ETL, Advance SQL operations, Handling Multiple Datasets* |
| * Rule Scripts * Developed and Implemented multiple rules which were based on network, device and browser fingerprint, and behaviour of the visitors. These were adequately balanced between causal and correlated, yielding negligible FP. * *Tags: Domain knowledge, Behaviour Analysis, Rules development* |
| * Rule Mining * Created a Recommendation System for the highlighting possible bad signature coming in the traffic. * *Tags: Associate Rule Mining, Text Data Pre-processing* |
| * Developed Dashboard, Visualization and Analysis, spreadsheet to understand the whole behaviour of the traffic. Partially interactive visualization sheets were also developed in python. *Tags: SQL and Spreadsheet/Python Implementation, Visualization, Result Sharing* |
| * Developing Dynamic Moving Collective Intelligence which can evolve while working across multiple sid. |

|  |  |
| --- | --- |
| **Python Package Developed** | **Algorithmic Machine Learning Exploration & Exploitation Tool (AMLEET)** |
| * A python package for multiple purpose code for logger wrapper, configuration, notification, cloud storage integration, local db, Redis Big query integration * Other application involve numerous computer vision functionalities ranging from stream creation to prediction wrapper for face detection, recognition, pose estimation and such. * Then there are functionalities for NLP related functions for data cleaning pipeline and even visualization. * Additional support for Supervised and unsupervised ML integration is also available |
| **COCO Transformation Utility (CTU)** |
| * A python package to perform the same transformation to coco-annotation as performed on the image. * v0.5 release will have the capability to create synthetic dataset. |

|  |  |
| --- | --- |
| **Patent & Research Paper** | * First Author. “System and method for detecting bots based on iterative clustering and feedback-driven adaptive learning techniques”; us US20200099713A1. * Second Author. “System and method for detecting bots using semi-supervised deep learning techniques”; us US20200099714A1. * First Author. “Electric field and current assisted alignment of CNT inside polymer matrix and its effect on electrical and mechanical properties”, in International Journal for the Science and Technology of Polymers 2016. |

|  |  |
| --- | --- |
| **Education** | **Indian Institute of Technology, Roorkee** Jul’12-Jul’17  B.Tech. + M.Tech. (Dual Degree Course) in Metallurgy and Material Engineering, CGPA: 7.152/10 |
| Ministry of Human Resource Development (MHRD) Scholarship for M.Tech. as for qualifying GATE. |
|  |
| **Udacity Nano Degrees**   * Data Product Manager, 2021 * Computer Vision, 2020 * Artificial Intelligence, 2018 |

|  |  |
| --- | --- |
| **Non-Credit Courses Taken** | |
| * Deep Learning and Computer Vision A-Z * Neural Networks and Deep Learning, DeepLearning.AI * Deep Learning: Advanced NLP and RNNs * Deep Learning with TensorFlow 2.0 * Business Fundamentals: Corporate Strategy * Real-time OCR and Text Detection * Deep Learning A-Z: Hands-On Artificial Neural Network * Statistics for Business Analytics A-Z * Introduction to Big Data, UC San Diego * Big Data Integration and Processing, UC San Diego * Customer Analytics, University of Pennsylvania * Linear Regression and Modelling, Duke University * A crash course in Data Science, Johns Hopkins * Managing Data Analysis, Johns Hopkins University * Introduction to R Programming, Microsoft Corporation * Mastering Data Analysis in Excel, Duke University | * Machine Learning A-Z: Hands-On Python & R in Data Science * Deep Learning and NLP A-Z: How to Create a ChatBot * Artificial Intelligence A-Z: Learn How To Build An AI * Introduction to Data Analytics for Business, University of Colorado * Big Data Modeling and Management System, UC San Diego * Data Management and Visualization, Wesleyan University * Practical Machine Learning, Johns Hopkins University * Building a Data Science Team, Johns Hopkins University * Data Science in Real Life, Johns Hopkins University * R Programming, Johns Hopkins University * Programming Foundations with JavaScript, HTML & CSS, Duke * Setting expectations & Assessing Performance Issues, UC Davis * Coaching Conversations, University of California, Davis * Java Programming: Solving Problems with Software, Duke Univ. * Managing as a Coach, University of California, Davis * Coaching Practices, University of California, Davis |
| (completed these certified courses either via udemy or coursera or edx) | |

|  |  |
| --- | --- |
| **Skills and Tools** | **ML Languages/Framework :** Python , scikit-learn , Tensorflow, PyTorch  **Other Language:** R, SAS, JAVA, C++  **Other:** Google Big Query, Kafka, Elasticsearch, Docker, Selenium, Redis , MongoDB, Grafana, Kibana, HTML, CSS  **Databases:** SQL, No-SQL  **Software:** Tableau, Excel, XLMiner, Spyder, Jupyter Notebook, RStudio, Gimp, SolidWorks  **Cloud Computing Services:** Google Cloud Platform, Amazon Web Services  **Cloud Data lakes**: S3, Cloud Storage, Drive, Pcloud  **Hardware:** Jetson Nano, Raspberry Pi, Arduino Uno, IP Cameras, Sensors, basic electrical devices and circuits  **Languages:** English(SRW), Hindi(SRW) |

|  |  |
| --- | --- |
| **Projects** | IP camera (RTSP protocol) to consistent HTML Live Stream with offline Data Sync |
| Edge Device based CV application for counting of sheets processed in a workshop |
| Text Detection and OCR |
| Face Detection, Face Landmark Detection and Face Recognition Wrappers |
| Human Pose Estimation and Activity Recognition |
| WhatsApp based Chatbot (Hack around way) |
| Selenium based web scrapping and Web Application Control of Tinder & WhatsApp |
| Streamlit based threshold adjustment for object detector and annotation visualizer |
| Unsupervised Learning Framework with easy selection of basic data pre-processing, transformation, dimension reduction/clustering/anomaly algorithm selection or pipeline creation, hyperparameters setting, ensemble, and EDA. |
|  |
| **Other projects:** |
| * Identifying customer segment, Udacity * Tagging part of speech, Udacity * Isolation game playing agent, Udacity * Cargo route planning, Udacity * Sudoku solver, Udacity * Twitter Sentiment Analysis * Black Friday: Understanding the customer purchase behaviour, Analytics Vidhya * Estimation of the audience score of movies, Coursera * Predicting the way in which exercise were done, Coursera * Forecasting bike rental demand in the Capital Bike Sharing Program, Kaggle * Predicting Survival on the Titanic, Kaggle |

Full Resume can be found at: <https://mr901.github.io/resume/>