

SKILLS

- AI/ML,
- GenAI,
- Data Science & Data analytics
- MLOps
- Big Data Technologies
- Application Development,
- Excellent communication,
- Project management
- Work Ethics and Ownerships

CERTIFICATIONS

- Introduction to Machine Learning in Production-Certification from Coursera- Grade 94.20%
- Introduction to Tensorflow for Artificial Intelligence, Machine Learning, and Deep Learning Coursera-Grade 97.06.
- Foundations of Project Management- Coursera-Grade Achievement: 91.70%
- Udarity Nanodegree Graduation Certificate Certification from Udacity
- Domain Masterclass on Gen AI and Its Impact on Insurance (TCS)
- Google-Gen AI intermediate level (E2) (TCS)
- Machine Learning-E3 (TCS)

TOOLS & DBS

Eclipse, Anaconda, Tableau, Power-BI, QlikSense AWS Services, GCP, RPA Tools MySQL, Oracle, MongoDB, Elastic Search, RPA, Web Technologies, Power BI, QlikSense, Excel, Streamlit. Big Data Technologies, Pyspark. InfluxDB.

DOMAINS

EDUCATION

- M.Tech-CSE (2016-2018)- 8.3
- B.Tech -CSE (2011-2015) - 6.9

LANGUAGE

- Hindi (Native)
- English (Professional proficiency)
- Deutsch (A1)



+91-8076154115



aditya0025@yahoo.com



www.linkedin.com/in/aditya0025



Delhi, India

PROFESSIONAL PROFILE

Analytical, energetic, and task-oriented with experience in Data Science and Data analytics, including AI/ML. GenAI Automation tools, Web Application & Cloud Services.

PROFESSIONAL SUMMARY

- 6.4 years of experience in leading high-quality tasks for Professional Services, Banking, Transportation and Logistics, Pharmacy, IoT, B2B, B2C sectors from different geographic locations.
- Comprehensive working experience in Data Science and Machine Learning models.
- Working Experience in Predictive and Agile software development models.
- Experience in Predictive model building from end- to-end process, Model architecture to production.
- Well versed in Planning, designing & executing Functional/System/Regression/Integration/E2E testing Test scenarios/ Test Case
- Design/Execution/ Defect Management.
- Worked with different Web scraping and crawling tools and techniques for Data collection purposes.
- Worked with IoT device integration with system (Lidar, GPS, Onboard SDK)
- Directed a skilled team of three, fostering innovation and teamwork.
- End to End Application Development with different python framework (Django and Flask) and MEAN stack

WORK EXPERIENCE

Tata Consultancy Services (TCS) <https://www.tcs.com/>
AWSBU-LLM Developer (June 2023 - till date (Delhi, India))

PoC Implementation for Japanese client (Takenaka) - AWS Environment

- Japanese Text Document - Color, Images and Tables
- Data Extraction, Pre-processing and Creation
- Fine-tuning Llama 2 70B with instruction fine-tuning and domain adaption fine-tuning
- Data parallelism with 8 gpus instance

Document Summarization with FLAN-T5

- Create S3 bucket for text files
- Deploy FLAN-T5 in Sagemaker jumpstart
- Create Lambda function with logic
- API with API Gateway - API will accept s3 path for the file as input.
- Integrated API with UI
- Test the functionality

RLHF with PPO on FLAN-T5 for Insurance bot (AWS Environment)

- Define fine tuning strategy
- Lambda Creation for querying fine-tuned model

Tata Consultancy Services (TCS) <https://www.tcs.com/>
CIGNA (USA-Pharmacy) - Python Developer
September 2022 May 2023 (Delhi, India)

- Worked with Data Engineering team for Application Maintenance. Create multi-threading processes for increasing system efficiency.
- Initiated the Code review system to refine the junk code and do proper documentation
- Follow Agile methodology for Application development
- Integrated CI/CD workflow for various backend services
- Git Implementation for code versioning.
- QnA testing

Genpact Digital India <https://www.genpact.com/>

CITI Bank-AL/ML Consultant

September 2021-September 2022 (Delhi, India)

ENR (Ending Net Receivables) and NCL prediction for predictive reporting for loss management.

Problem Statement: ENR and NCL are some of the major areas where a bank can detect their future losses and predict that will help the bank to manage the future scenarios.

Technologies: Python, TensorFlow, Pyspark, Tableau

Key Points: A hybrid model with embedding layer and deep neural network. Dealing with 20-30 GB of datasets (Collection-Analysis - Model training Reporting) Tableau reporting Business acceptance Working knowledge of test management software

Transport Corporation India Limitedes (www.tcil.com)

Data Scientist

December 2018 - September 2021 (Gurgaon, India)

Route Optimization with TomTom and Google maps

Problem Statement:

Creating Route calculation for a Truck (with Multiple Axle) with alternative route with their en-routes calculation of tolls, petrol pumps, RTO etc. Display a dynamic dashboard which provide all routes calculation.

Technologies:

Python, Flask, SQL, MongoDB, Rest APIs, JavaScript, HTML, JQuery, Bootstrap 4

Prediction of Transit time of a truck(ETA)

Problem statement:-

To predict the Estimate time of arrival of a truck for the Plan management.

Key Points:-

It's one of the concerned and key problem of logistics.

Data cleaning and processing of past 3 years of gps data.

Pyspark for Big data Analytics.

Features engineering and Feature Selection (Parameters) for the model input.

Designing and Tuning Deep neural network (LSTM).

Saving predicted model In HDFS File format(Hadoop distributed File system).

Regular Presentation directly to M.D (TCI Ltd)

Analytical mind and problem-solving aptitude

Technologies: Python, SQL, No-SQL, Deep Neural Network.

Multiple Dashboards using QlikSense and Python

Problem statement:-

1. Revenue prediction and Create a Dashboard for the CEO's.

2. Freight-Solution Dashboard.

Key Points:-

Create Revenue Prediction Dashboard by connecting QlikSense with Python.

Freight-solution Dashboard (Multi-modal Freight Dashboard) for overall Group Business Development Head.

Data Collection and wrangling.

Technologies:- Python, SQL, Flask, Big data Analytics, QlikSense, Excel

Last Mile Delivery/Pickup

Problem statement:-

For the optimize distribution of the consignment in inter or intra city network.

Key Points:-

An optimize route planning for the last miles (HUB Or Branch Distribution)

Clustering last mile customers and planning Best route for each cluster with minimum cost Function.

Technologies: Python, Graph-Search, Numpy, Pandas, matplotlib, Scikit-learn, SQL, No-SQL, Machine learning Libraries.

Search-Engine for location Search(TCI Place API)

Problem statement:-

The Purpose of the project is to develop a search engine which used as and alternative to Google Place API which decrease the Dependency on the Google to track live location of a Truck.

Key Points:-

Deal with huge amount of GPS location data which is GB's.

Data Pre-processing and cleaning for exact Geo-code of the location.

Search Engine for the Fast and accurate search of the location with auto-complete facility and Geo-location.

Worked with Structured Data and semi-Structured Data.

Technologies:- Python, SQL, Elastic-Search, MongoDB, Flask, Big data Analytic, web-development Technologies.

RPA TOOL and Python for web-scraping

Problem statement:-

Using RPA tool (Tru-bot) for web-scraping and saving it to database.

Key Points:-

Web scraping from different Travel Websites for Travelling Rate Purposes using RPA Tool.

Used Python where RPA tool is not working.

API for data insert to database.

Technologies:- Python, SQL, Flask, Big data Analytic , Beautiful-Soup, Crappy, web-development Technologies.

OCR (Object Character Recognition)

Problem statement:-

Text-detection from Cheque data (Account No, IFSC, MICR etc.) and other bank related data.

Text-detection from Vahaan Details (Lorry details) from RC pics and PDF provided from the field.

Key Points:-

Data Collection from Cloud-Platform (AWS) services.

Data Pre-processing and cleaning for exact Text.

Image Processing using TensorFlow.

Data Collection and Insert relevant data to the database.

Worked with Semi-Structured Data and Unstructured Data.

Technologies:- Python, SQL, Flask, Big data Analytics, TensorFlow, Neural-Networks

TCI-Browser

Problem statement:-

Developed a Web-browser for TCI ERP system to secure the URL-link for Security Purposes.

Key Points:-

Chromium Based Web-Search engine.

Integrating with the ERP system.

Pre-defined management requirement for Security.

An exe to run on Field systems (Linux Based).

Technologies:- Python, Chromium, PYQT5.

TCI-Directory

Problem statement:-

Search engine to detect the Caller details at Real-time and display to mobile and web(IVR). For Customer, Suppliers, Vendors and Employees.

Key Points:-

Data collection, Creating Own data-set with number details and validating the Data.

Android app for display at real-time.

Web-API for the customer Delight locations.

Worked with Structured Data and Un-structured Dataset.

Technologies:- Python, SQL, Flask, Big data Analytics, MongoDB.

Sugal and Damani Group (www.sugaldamani.com)

Skill lotto solution (Group Company)

Python Developer

July 2018 - Decmeber 2018 (Gurgaon, India)

API Development

Blockchain PoC

Payworld (Group Company)

Developer and Analytics consultant

Churn customers analysis and Financial Data Analysis for Payworld (division of group):

Problem statement:-

Churn customer analysis and Credit analysis.

Key Points:-

Detection the churn or about to churn customers of the Payworld.

Data collection and Pre-processing.

Financial Data Analytic.

Technologies: - Python, Matplotlib, Pandas, Numpy.

Indian Institute of Technology (IIT Delhi)

Project Associate

June 2015 - June 2016 (Delhi, India)

Online Teaching Platform:

Problem statement:-

Developing an Online Platform for the professor of IIT Delhi.

Key Points:-

Application development using Java.

Contributing in all phases of the development life-cycle.

Writing well-designed, efficient, and testable code.

Managing Java and Java EE application development.

Maintaining UI of project.

Technologies:-Java, web-development Technologies.

Ø Languages/Technologies / Databases:

Python, JavaScript, MySQL, Oracle, MongoDB, Elastic Search, RPA, Web Technologies, Power BI, QlikSense, Excel, Stramlit, Big Data Technologies, Pyspark, InfluxDB, Grafana

Ø Tools / Software / IDEs/ Services:

Eclipse, Anaconda, Tableau, Power-BI, QlikSense, AWS services(S3, EC2, SageMaker), GCP, RPA Tools.

Ø Operating Systems:

Windows (8, 8.1, 10), Linux (Ubuntu, Kali, Fedora), Android.

Ø Libraries/ frameworks/ Clouds:

TensorFlow, keras, SciPy, NumPy, Pandas, matplotlib, Hibernate, Flask/Django, AWS, GoogleCloud

RESEARCH PUBLICATIONS

Currently having 3 research paper published in the field of - Machine Learning, Deep Neural Network, Cloud Computing and Cryptography.

- "Cloud and Biometric: The Future of Authentication" In Volume 8, No. 2, March 2017 (Special Issue) International Journal of Advanced Research in Computer Science
- "Rainfall Prediction with Machine Learning Algorithms and Comparison" in Volume 7, Issue 4(July, 2018) of our journal IJSRCSAMS.
- "Estimated Time of Arrival for Sustainable Transport using Deep Neural Network":-
https://link.springer.com/chapter/10.1007/978-981-19-4052-1_9

COURSES

- > Introduction to Machine Learning in Production- Certification from Coursera - Grade Achieved: 94.20%
- > Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning - Certification from Coursera - Grade Achieved: 97.08%
- > Foundations of Project Management
 - Certification from Coursera - Grade Achieved: 91.70%
 - Agile Expert Certificate of Proficiency - Certification from TCS internal training.
- > Udacity Nanodegree Graduation Certificate Certification from Udacity.
- > Microsoft Azure Essentials Certification from Great-learning.

EXTRACURRICULAR

ACTIVITIES:

Social Work: Worked with social organization Trishakti charitable trust, which provides primary education to street children in Delhi.

Hobbies: Gym and Travelling to different locations and explore different cultures.
(Travel to different states of India and Countries of EU like Germany, Luxemburg, Belgium, Netherlands, France, Italy))