

# Arth Akhouri

**in** arth-akhouri

✉ akhouriarth@gmail.com

🔊 FlintyTub49

☎ +44 7767 961 741

## EDUCATION

---

**The University of Edinburgh**

*MSc. in Data Science*

**Edinburgh, Scotland, UK**

*Sept 2023 - Ongoing*

**Mukesh Patel School of Technology Management and Engineering**

*B.Tech in Data Science*

**Mumbai, IN**

*Graduation: August 2022*

## EXPERIENCE

---

**eClerx**

*Associate Process Manager*

**Thane, Mumbai, IN**

*Jan 2022 – Jun 2023*

- Collected and processed Traffic data of app and website from different sources using APIs.
- Automated creation of Traffic Analytics reports based on client defined KPIs by using Python.
- Created an end-to-end pipeline in Databricks to generate a dashboard for Traffic Analytics.

**AVL**

*Data Science Intern*

**Thane, Mumbai, IN**

*Dec 2021 – May 2022*

- Analysed Time Series Electrical data generated by various machines in an IIOT system to find inefficiencies.
- Used Plotly to create visualizations of parameters of machines specifically focused on energy consumption.
- Utilised Statistical and Machine Learning techniques to find a threshold safety value of a particular machine.

**AkzoNobel**

*Data Science Intern*

**Thane, Mumbai, IN**

*Sep 2021 – Mar 2022*

- Collected large amounts of colour recipe data and linked it together to find different toners belonging to a colour.
- Built various Machine Learning and Deep Learning models to detect colours based on their reflection values.
- Used ML and DL models to generate new recipes for solid colours using fewer base colours.

**Amicus.AI**

*Founding Member/Data Scientist*

**Thane, Mumbai, IN**

*Feb 2021 – Nov 2021*

- Extracted and processed vast amounts of Legal textual data spanning over 100,000 case files using Regex.
- Implemented NER and other NLP techniques to obtain an accuracy of over 70% on unstructured case data.
- Created interactive visualizations and a web based dashboard for presenting them using Dash and Plotly.

**DKSH Smollan - In Partnership With Google**

*Data Science Intern*

**Thane, Mumbai, IN**

*May 2021 - July 2021*

- Responsible for scraping and filtering relevant Instagram data for gaining business intelligence.
- Building a cleaning pipeline to process the obtained data and derive meaningful insights from it.
- Helping in formulation of the Database structure and aiding the deployment of the application to the cloud.

**Vorfreude Community**

*Lead UI/UX Designer*

**Thane, Mumbai, IN**

*Apr 2020 - Nov 2020*

- Responsible for designing UI/UX for an app - WaitQ, created to help maintain social distancing at local shops.
- Helped in designing basic back-end architecture and data storage format for all customer and shops data.
- Participated in meetings with Government Officials to discuss the requirements of deployment of the app.

## ACHIEVEMENTS

---

**HackRx 2.0 - Hackathon, 2nd Position**

*July 2021*

- Created a crawler to scrape the Bajaj Finserv website. This data was used to train a model to improve the Search Recommendation Engine on the website using query extension.

**Paper Published - Review Of Data Privacy Laws and Case Studies**

*October 2020*

- Published a review paper on data privacy laws in the International Journal Of Creative Research Thoughts (IJCRT)

## ACM Mumbai 24 Hours - Hackathon, Winner

March 2020

- o Developed an App where investors, entrepreneurs and leading people from the industry could meet and grow. Designed the app in Adobe XD and coded in Flutter language.

## PROJECTS

---

### Document Scanner with OCR

4 Months

- o Developed a Flask app that can identify a document in an image and crop out the rest of the image. It uses line segment detection to find the edges of the document. Optical Character Recognition was also implemented to detect any text available in the scanned document. (Computer Vision, NLP)

### Music Genre Detection and Recommendation System

3 Months

- o Used a plethora of music files to analyse their audio pattern and train a model to predict a genre for the song based on it. This was deployed on a website where a user could upload a music file to a model and get similar genre song recommendations on the basis of it. (Deep Learning, NLP)

### COVID-19 Fake News Detector Bot

3 Months

- o Extracted data from X (formerly Twitter) and WhatsApp forwards regarding vaccine news and labelled it as fake or real. Used this data to create an NLP based WhatsApp bot, Telegram bot and a Website which takes news related to COVID-19 as input and verifies if the news is real or fake. (Deep Learning, NLP)

### Medium Article Recommendation System

2 Months

- o Scraped the Medium website for articles and created a model to classify each article in a particular topic. This was then extended to create a recommendation system where users could give a quantifiable preference to a particular topic desired in their article. The website was created using Streamlit. (Regex, Machine Learning, NLP)

### PyPi Python Package MAHA

3 Months

- o Created a python package which provided functionality to clean a given numeric dataset. The package offers to extrapolate missing values with appropriate methods. There are other small functions which can help with the task of data manipulation. (Data Cleaning, ETL, Machine Learning)

### Gesture Based Vault System

2 Months

- o A model is trained to look for particular sequence of hand gestures to unlock a vault. Hand gestures were manually captured and further augmented to create the training data. Principal Component Analysis was used to reduce the size of the model to make the program efficient. (Dimensionality Reduction, Computer Vision, Deep Learning)

### Automatic Library System

1 Month

- o A basic code trained on an image of a bookshelf which can tell when book/books have been added or removed from a certain place in it. (Dimensionality Reduction, Machine Learning, Computer Vision)

### Local Train Ticketing System

2 Months

- o Hypothesised a new ticketing system for Mumbai Local Train System which would enable more efficient data management about the local train ridership. Created the database design for this new system and tested it by running few sample queries. (Database Management, Oracle SQL)

### Analysis of Student Performance

1 Month

- o Used t-test to compare different aspects of student life and analyze how it affected their performance in college. (Statistical Analysis, R)

## ADDITIONAL INFORMATION

---

- o Participated in various Data Science Competitions and Hackathons
- o Online Certifications -
  - Machine Learning (Stanford Uni.)
  - Neural Networks and Deep Learning (Deeplearning.AI)
  - Business Analytics and Data Mining Championship (SAS)
- o Volunteered at the University Cultural Festival Sattva in the Management Department (January 2019)
- o Volunteered at Pet Care Centre and Raymond Rehabilitation Centre (December 2017 - March 2019)