Portfolio Website: https://sachinmishraportfolio.streamlit.app/

Sachin Mishra

Phone:+91 8928450584 | Github | sachin19566@gmail.com | LinkedIn

Analytical and solution-oriented background, with great people skills and structured problem-solving capabilities & more than 2.5 years of experience in IT and 2 years experience in Data Science.Organized with the ability to manage multiple workloads and consistently meet deadlines.

Skills

Programming	:	Python- (Scripting / Automation), C, R/Rstudio, SQL
Visualization	:	Tableau, Power BI, Seborn, Plottly, Matplotlib, Tableau, Prepflow-builder, Excel
Database	:	MySQL, MongoDB
Skills	:	**Python, Machine Learning**, **Web Scrapping**, **Deep Learning -** *(ArtificialNeuralNetwork),* **Computer Vision -** *(ConvolutionalNeuralNetwork |Image processing | Object Classification | Object Detection)* **Natural Language Processing -** *(Transformers, Bert, GPT)* | **Git** | **GitHub, MLOps -** *(DVC | MLflow | Sagemaker| Docker| GitHub Actions)* | **APIs** *(FastAPI| Flask)* | **Cloud** *(AWS / GCP)* ***Snowflake , LLM***
Frameworks	:	Streamlit, Flask, Langchain,FastAPI
Competency	:	Communication – YouTube Channel (Loc Data - YouTube), Teaching/Mentoring

Experience

Data Society - Data Scientist (Jul 2022 – Current)

- Creation of **supervised and Unsupervised ML models** and deploying it on cloud based upon requirements.
- Day to day activity includes cleaning and analyzing marketing data for better insight.
- Creation of ML courses for corporate employees in Data science field.
- Worked on CNN project to classify fruits images using transfer learning. Written a production grade code for building a **Convolutional neural network(CNN)** using transfer learning approach on pretrained **VGG-16 model, and Mobile-Net model** to classify fruits images.
- Worked on **NLP** project for doing sentiment analysis on companies policy documents.
- Worked on Project of **Tableau Dashboards** for North-carolina government which I used **tableau prep flow builder** for creating dataflows and finally build the dashboard for the same.
  Link : https://dashboards.ncdhhs.gov/#/site/DCDEE/projects/125

Digikull - Data Analyst/Scientist Mentor (June 2022 – Current- Part-Time)

- Facilitated learning and growth as a mentor and instructor, delivering engaging lessons and practical examples to students on **Python programming, machine learning, statistics, and Tableau.**
- **Developed and implemented a comprehensive curriculum for Python programming, machine learning, statistics, and Tableau,** catering to students with diverse backgrounds and skill levels.
- **Mentored and coached junior data analysts**, providing guidance on best practices, troubleshooting techniques, and code review to facilitate their p**rofessional growth and development**.
- **Mentored and guided students in their learning journey,** providing individualized support and feedback to help them grasp complex concepts and apply them effectively.
- Designed and conducted h**ands-on coding exercises, projects, and assessments to assess students' understanding and proficiency in Python programming, machine learning, statistics, and Tablea**u.
- Introduced and implemented **MLOps methodology for the first time in the organization using MLFlow**

Ineuron.ai – *Data Science Intern* (Feb 2022 - Oct 2022)

- Completed generic training in **MySQL, Python, Statistics, Tableau, and Machine Learning.**
- Cleaned and formatted Big Mart Sales data with over 8524 rows and 12 columns and made it ready for analysis.
- Developed interactive dashboards to visualize Key Performance Indicators (KPIs) and provided business recommendations.
- Build and Compare different Machine Learning Models such as **Linear Regression, Lasso Regression,SVM, and Random Forest** to Predict Sales of the different stores of Big Mart

Tata Consultancy Services Ltd - *SAP Analyst* (Jul 2021 - Jul 2022)

- Day to day creating and maintaining Clients data in SAP hana database in Production, Development and Quality System.
- Prepared BI dashboard in **Tableau** for tracking of monthly incidents, tasks and change request reported. This work improved the tracking of different task assigned to the different teams and reduced the SLA.
- Worked closely with the engineering and business team using **scrum/agile** methodology
- Creating SAP BI objects, Providing necessary roles and authorizations to clients and monitoring process chains.

Projects

Sports Celebrity and Data Scientist's Image Classification **- Github Link , Docker Image**

- Developed a Classifier using **Haar-cascade and Wavelet Transform** to classify different sports personssuch as Maria Sarapova , Virat Kohli , Serena Williams, Roger Federer and Lionel Messi.
- Extracted face region from the images and labeled them for feature extraction using OpenCV in ScikitLearn.
- Compared SVM, Random Forest and Logistic Regression to compare the results. Finally I choose **SVM**for better results and predicted results with **84.13% accuracy**.

Credit Score Classification **- Docker Image**

- Analysed & cleaned Credit dataset over 50,000 records
- Built Classifier using Random Forest and achieved prediction accuracy of 72%
- Built docker image and deployed it using Kubenetetes cluster in local environment.

Pet-Image Classification using CNN **- Docker Image , Github**

- Cleaned Images for pets (Dog/Cat) and prepared data for building a CNN model.
- Written a production grade code for building a **Convolutional neural network(CNN)** using transfer learning approach on pretrained **VGG-16 model**.
- Achieved accuracy of 94%, containerized the service and created model version tracking using **mlflow**.

Python Pypi Package **- Package-Link**

- Created well tested **python pypi package and published on pypi** official page.
- This package is available for **MAC-OS** and for **Windows & Linux** still it is in developing phase.
- This package tries to record time spent on each & every platform & then creates analytics based upon those platform and time spent on those platforms.

NLP Emotion Detection **- API Link**

- Trained an NLP Emotion detection model using BERT, which can detect what kind of emotion is present in a sentence.
- Achieved accuracy of 84%, containerized the service and created model version tracking using **mlflow**.
- Deployed (Hosted) on **AWS,Azure and Streamlit-community cloud**.

Atliq Hardware Sales/Profit Dashboard **- Github Link**

- Developed Sales and Profit dashboard for Atliq hardware in tableau
- Cleaned 10,000 rows of the sales transaction table and performed some basic data munging.
- Connected tableau to mysql real-time database so that new records can be updated in the dashboard automatically.

Aadhar Card Masking & Information Retrivel **- Github Link**

- Written a python Module which takes Aadhar card images as input and gives all the neccesaary info (Name, Age, Gender, Aadhar Card number, Address) and masked version of Aadhar card.
- For Masking and extracting all the info from from Aadhar card is done by **easyOCR** and **Open-CV** library.

Flipkart review scraping using selenium **- Github Link**

- Built Flipkart review scrapper bot using selenium.
- Using cron job scheduled it via github actions on github to get daily scrapping results.

Community Sessions:

- Train ML/DL model in just 2 mins using **Teachable Machine -** **Build Your ML/DL Model in just 2 mins using Teachable Machine with Streamlit web app ...**
- **Tableau** Project based learning - Tableau Project Based Learning | Chapter-2 | Part-1
- Python Automation - Pdf_tables_to_Excel_tables in Python | Loc Data

Education

Thakur College of Engineering & Technology, Mumbai, Maharashtra (Jul 2017 - Jun 2021)

- BACHELOR OF ENGINEERING, Electronics and Telecommunication - CGPA - 9.04/10
- Related Courses: Signals & Systems, Discrete Time Signal Processing, Image Processing and Machine Vision, Database Management System.

T.P Bhatia College of Science, Mumbai, Maharashtra (Jul 2015 - May 2017)

- H.S.C,Science; Physics,Math,Chemistry and Electronics - 75.07%
- Related Courses: Electronics,Statistics&Maths,Microprocessors and Controllers