Manaal Saxena

Hyderabad, TG | LinkedIn | +91 9666760611 | saxenamanaal@gmail.com | Porfolio

PROFESSIONAL SUMMARY

As a professional Python developer, I excel in data science and web development technologies. I have experience in building complex ETL processes, implementing machine learning, and automating tasks. Proficient in Git and working with relational and non-relational databases such as MySQL, Cassandra, and Redis. Proven track record in developing applications with Flask and Django, and expertise in cloud services and data visualization tools. Experienced in working with Unix-based operating systems. Committed to staying current and delivering robust solutions as a dedicated Python developer.

SKILLS:

Data Analytics: *Pandas, Numpy, Scipy, Pyspark, Dask, Power Bi, Tableau, Matplotlib, Bokeh, PLotly, Excel, Gsheets*
Data Engineering: *ETL's, SQLAlchemy, Spark, Airflow, Elasticsearch*
Machine Learning: *Sci-kit Learn, NLTK, Spacy, Tensorflow, Keras, PyTorch, OpenCV, ML flow*
Databases: *Mysql, Postgresql, MariaDB, Apache Cassandra, Redis*
Web Development: *Flask, Django, Streamlit, Rest API, OMR's, Nginx, Apache*
Miscellaneous: *GIT, Shell, Digital Ocean, AWS, Vim, Jupyter, Linux*

EDUCATION

B.Sc Computer Science — **Hyderabad, TG**
Osmania University — *October 2021*

WORK EXPERIENCE

Python Developer — **Hyderabad, TG**
Muzigal — *May 2022 – Present*

- Delivered meaningful insights to the company through expertise in Python and data visualization tools, utilizing powerful libraries such as **Pandas** and **Pyspark** to process large datasets and extract valuable insights that informed business decisions and optimized operations.
- Implemented efficient and scalable data workflows through the design and development of complex **ETL** processes, reducing manual workload and **streamlining** data management.
- Contributed to the modernization of the company's infrastructure by integrating various databases and systems, utilizing both relational and non-relational databases like **MySQL**, **Postgresql** and **Cassandra**.
- Utilized extensive knowledge of machine learning algorithms to deliver impactful results for the company. Developed **predictive models** and applied **time-series forecasting** techniques, as well as worked on **NLP** projects, Utilizing technologies like **Sci-kit Learn**, **Tensorflow** and **MLFlow**. Provided data-driven insights to inform company decisions.
- I have played a significant role in delivering high-quality web applications for my company, leveraging my expertise in web frameworks such as **Flask** and **Django**. My ability to create scalable applications has greatly contributed to improving the **Data Latency** which resulted in better user experience.

Research Fellow — **Hyderabad, TG**
Fellowship.ai — *September 2022 – January 2023*

- Contributed to the 'Competitor Analysis' research project, aimed at analyzing e-commerce competition to enhance sales performance.
- Led the web scraping aspect of the project, built a scrapper engine utilizing technologies such as **Beautiful Soup**, **Selenium**, and **Scrapy** to gather information on similar products from competitors.
- Achieved the goal of providing insight into the competitiveness of the products compared to our own and advising pricing strategies for increased profitability through the following techniques:
    - Scraped similar products based on various factors such as brand, category, size, color, etc.
    - Assessed the similarity of the products using **text embeddings**, **image embeddings**, price, and SKU matching.
    - Offered information on price changes and positioning over time, as well as the competitiveness of the products compared to our own to inform pricing strategies.
- The project utilized advanced models for data analysis and processing, including the **CLIP ViT-B/32** Transformer Architecture and **VisionTextDualEncoderModel**.
- The project employed powerful deep learning models such as **CLIP-ViT-32** and **Roberta-Base** to extract and analyze visual and textual data respectively.
- The **KNN** algorithm was used to classify and categorize data, providing a comprehensive and in-depth analysis of the market.

Machine Learning Engineer (Internship**)** **Bangalore, KA**
inueron.ai *January 2022 – April 2022*

- Proactively contributed to the development of Machine Learning Science projects, product development & PoCs.
- Constructed end-to-end Data Science projects, performing tasks **Data Gathering, Data Cleaning**, **Model Building**, **Model Deployment**.
- Worked on building multiple complex machine learning and NLP models, which utilized technologies like **Text Classification, Keyword Extraction, Word Embedding, Entity Recognition**.
- Built integrations with **Apache Cassandra** and **Maria DB** to store date and finetune the model. **ML Flow** was used to maintain the project workflow.

Associate Developer (Internship**)** **Hyderabad, TG**
Roy Projects Pvt Ltd *March 2021 – August 2021*

- Managed, supported and monitored **MySQL** Server database servers, proactively **troubleshooting** and **maintaining** servers.
- Executed **MySQL** database installations and **upgrades**, performance monitoring and tuning, **backup** and **recovery**, space management and resource utilization.
- Created scripts and processes to **integrate** and **maintain** data.

PROJECT EXPERIENCE

Joblix | Code | App
- A powerful tool that will find you all the jobs listed on the web with a single search.
- Built a complex **web scraper engine** that can scrape **40,000+ data points** in no time, that is followed by a dedicated pipeline which sends the data to the database.
- **MongoDB** is implemented to store and manage data, data is pulled from the DB using **SQLAlchemy** to avoid data latency and is displayed to the user using **Django OMR.**
- It is built on **Django** framework, maintained using **Docker** containers and deployed using **Kubernetes**.

RE-VEU | Code | App
- RE-VEU is a powerful web app built streamlit framework. It provides a seamless experience to analyze any review data, transformers any any review into clean text and produces analyzed results.
- It runs on a powerful algorithm that is built with open-source packages such as **NLTK**, **TextBlob**, **pandas**.
- **Jupyter Notebook** and **PyCharm** for scripting, **Streamlit** cloud for Deployment and **Github** for code maintenance.
- Technologies in implementation for this project are **Python**, **Pandas**, **NLP**, **Streamlit**, **Jupyter**, **vader-sentiment**, **Git** etc.

Domanify | Code
- Domainify is a web tool that is able to predict if a URL domain is real or malicious.
- The ML model used in Domainify is trained on data which is initially collected from Mendeley data repository, then cleaned to eliminate errors and unwanted features.
- Built using following models - **SVM**, **Random Forests, Logistic Regression, AdaBoost, KNN** Hyperparameter Tuned models.
- **Streamlit** is used for **model deployment**, performing **feature engineering** so URLs are structured to feed to the model requirements. **Jupyter Notebook** and **PyCharm** for scripting, **Github** for code maintenance.
- Technologies in implementation for this project are **Python**, **Pandas**, **Sci-kit Learn**, **Streamlit**, **Jupyter**, **PyCharm**, **Git** etc.

HwDR GUI | Code
- Developed a **Deep Learning** model that achieves high performance on the handwritten digit recognition task using the **MNIST** dataset available in the **Keras** library.
- Built a **GUI** based on **Tkinter** where you can draw the digits and recognize it straight away by a bounding box surrounding each digit.
- Performed **Robust Feature Extraction** to improve the performance of machines as the challenge in handwritten digit recognition is mainly the variation of writing styles of various individuals.
- Technologies used in the project include **Python**, **Keras**, **Tensorflow**, **OpenCV**, **Tkinter**.

ACHIEVEMENTS

- Attained Gold badge for SQL Scripting in Hackerrank competitive coding.
- Achieved Silver Badge for Python programming in Hackerrank competitive coding.
- Selected as Campus Ambassador by Internshala Student Program. To represent Internshala at college campus.
- Placed in top 7% India wide in SoloLearn coding challenges.