

DIPALI MEHTA

Data Scientist

Email : dipalivinodmehta@gmail.com

Contact : 9284306783

Linkedin : www.linkedin.com/in/dipalivmehta

SUMMARY

Data Scientist with a track record of driving data-driven solutions and delivering actionable insights. Proficient in SQL, Excel, and Python, with a strong background in data analysis and statistical modeling. Expertise in exploring large datasets, designing methodologies, and translating data into business results. Dedicated to revolutionizing health decisions through real-time data.

EXPERIENCE

Associate Data Scientist

Blackstraw.ai Chennai,India 09/2021 - Present

- Developed and implemented machine learning models for automated data extraction, resulting in increased efficiency and reduced manual effort.
- Applied data analysis techniques to deliver valuable insights, contributing to data-driven decision-making and business strategies.

Machine Learning Intern

Avishkar Tech Solution Hyderabad,India 06/2020 - 01/2021

- Applied Machine Learning algorithms to solve real-world problems and developed models using Python libraries like Sci-kit learn, TensorFlow, and Keras.
- Utilized data exploration, visualization, and preprocessing techniques for improved model performance and assisted in deploying models on production platforms.

PROJECTS

Invoice Data Extraction

Aim is to provide end to end capability on automated extraction of fields in invoices/receipts and an efficient invoice extraction solution.

- Developed and implemented an OCR-based machine learning model for automated invoice extraction, significantly reducing manual effort.
- Utilized LayoutLMv2 transfer learning to train the model, incorporating image, layout, text, and location information.
- Integrated the model into the organization's workflow, resulting in improved invoice processing times and accuracy.

Crowd Sourcing Estimation

Aim is to build a scalable asset estimation system by gathering and processing data

- Leveraged SQL queries to extract insights from large datasets and design effective methodologies for measurement products and user identification.
- Collected data from multiple sources and implemented barcode reading to extract relevant fields. Identified and resolved system errors and bugs, ensuring accurate estimation results.
- Conducted gap analysis to identify areas for improvement and enhance
- Contributed to end-to-end testing and successfully deployed the system from a country-specific version to a generalized version, improving scalability and usability.

INSIGHTS

PROGRAMMING

Python, R, SQL

PACKAGES

NumPy, Pandas, Matplotlib, Seaborn, NLTK, OpenCV, TensorFlow, Keras, SciKit Learn, PyTorch

PLATFORM & TOOLS

PyCharm, Visual Studio Code, Jupyter Notebook Anaconda, DBeaver, Databricks, R Studio, Excel, PowerBI/Tableau, Git

TECHNIQUES

Machine Learning, Deep Learning, NLP, Computer Vision

ALGORITHMS

Linear Regression, SVM, Logistic Regression, KNN, Decision Trees, PCA, Layout LM, K-Means Clustering, Random Forest, Naive Bayes, Neural Networks, Time Series, Random Forest, Support Vector Machines (SVM), Principal Component Analysis (PCA)

MATHEMATICS

Statistics, Probability, Linear Algebra, Calculus

SKILLS

Data Analysis, Statistical Modeling, Data Visualization, Data Exploration, Data Wrangling, Data Preprocessing, Data Cleaning, Data Transformation, Feature Engineering, Hypothesis Testing, Exploratory Data Analysis (EDA), Data Mining

EDUCATION

Bachelors Degree in IT

Savitribai Phule Pune University

06/2017 - 06/2021

CERTIFICATIONS

Training

- Python Programming from Ethan's Tech
- C++ & OS Certification from NPTEL
- Data Science from Great Learning
- Deep Learning from Coursera