

Sagar Patel Data Scientist

9913758770 | patel.sagar939@gmail.com | <https://github.com/ITianssj> | <https://www.linkedin.com/in/sagar-patel-2b2aa6167>

Profile

As a data analyst with 3 years of experience, I possess robust data analysis, visualization, and process improvement skills. Proficient in Python and SQL, I have successfully led end-to-end data analysis projects, utilizing statistical techniques to derive insights from large datasets. I am skilled in creating visually appealing dashboards using tools such as Power BI and have a track record of optimizing processes and achieving measurable results. With meticulous attention to detail and a results-oriented approach, I am proficient in Python, SQL, Power BI, and Excel and am familiar with machine learning techniques.

Experience

PREDICTIVE RESEARCH sr. Data Analyst

Bangalore [Feb2021-Aug2023]

LEAD SCORING ENGINE

- Objective of the project is to give the leads to the OEM about the propensity of the customers to buy the product. Customer: Auto data Solutions/Hyundai Motor America.
- Tech Stack: Python, scikit-learn, MySQL Server, TensorFlow, flask, docker.
- Role: Data collection, data cleaning, feature selection, model building, continuous testing and deployment, report submission, model accuracy improvisation.

PREDICT-OPS

- Predictops is a product for log monitoring under development by Predictive Research.
- Tech Stack: Python, Django Rest, Elasticsearch, Prometheus, EC2, MongoDB, PostgreSQL.
- Data ingestion from Elasticsearch to predict-ops. Pattern recognition, micro-service design and development. Elast-alert configuration with the product.

CUSTOMER LIFETIME VALUE

- Objective of the project is to give the who is your best customer
- Role: Data collection, data cleaning, feature selection, model building, continuous testing and deployment, report submission, model accuracy improvisation.

TOPIA LIFE SCIENCE Data Scientist

AHMADABAD [Jan2024-Present]

Retrosynthesis with GNN

- Implemented Graph Neural Networks (GNN) for retrosynthesis, automating organic compound pathway prediction
- GNN Model: Developed and deployed a GNN model for molecular structure analysis.
- Data Processing: Preprocessed chemical data, engineered features, and optimized model performance.
- Integration: Collaborated with cross-functional teams to seamlessly integrate GNN into retrosynthetic analysis.
- Results: Accelerated synthetic route design, improving efficiency, and reducing time-to-market for new compounds.

Projects

Care Predict

- Developed a machine learning and deep learning model to diagnose and predict seven common diseases: Alzheimer's, brain tumors, breast cancer, COVID-19, diabetes, heart disease, and pneumonia.
- Utilized various ML algorithms: fine-tuned CNNs for image recognition, and Random Forests for tabular data.
- Built and deployed a Flask application: allows healthcare professionals to input patient data and receive diagnostic predictions.

"Buy Till You Die" - Customer Lifetime Value Prediction

- Leveraged the lifetime package in Python to predict customer lifetime value (CLTV) and gain valuable insights into customer behavior.
- Built and trained a model using the Beta-Gamma-NBD (BG/NBD) approach. This method effectively captured both the frequency and monetary value of customer purchases, providing a comprehensive CLTV estimate.

Movie Recommender System with Streamlit Integration

- Developed a movie recommender system that leverages cosine similarity and Streamlit for an interactive user experience.
- Utilized cosine similarity: Analyzed movie metadata (genres, directors, actors, etc.) to calculate the similarity between movies, recommending those closest to the user's preferred choices.
- Integrated Streamlit: Built a user-friendly interface where users can input their favorite movies and receive personalized recommendations.
- Technical Skills: Python, Cosine Similarity, Streamlit, Data Visualization

Dell Image Generation with LLM APIs

- Leveraged cutting-edge Large Language Models (LLMs) and APIs to generate unique and creative images showcasing Dell's technology and vision.
- Utilized OpenAI's DALL-E 2 APIs: Experimented with different prompts and phrasing to craft compelling narratives that guided the LLMs towards generating Dell-specific imagery.

Education

B.E.IN INFORMATION TECHNOLOGY

GANDHINAGAR [2017-2020]

GOVERNMENT ENGINEERING COLLAGE GANDHINAGAR 7.23 CGPA

DIPLOMA IN INFORMATION TECHNOLOGY SURAT

SURAT [2014-2017]

DR. S&S. S GANDHI COLLAGE SURAT 6.98 CGPA

Skills & Abilities

- **Programming Language:** Python
- **Machine Learning:** TensorFlow, Keras, Scikit-Learn
- **Data Analysis:** NumPy, Pandas, Matplotlib, Seaborn
- **Cloud Computing:** AWS, Azure, GCP Web
- **Tools:** Power BI, Excell
- **Databases:** PostgreSQL, MySQL, MongoDB
- **Development:** Django, Flask