

UDIT GOEL

New Jersey, United States · goeludit7rma@gmail.com · (+1)7325589842

<https://www.linkedin.com/in/goeludit7rma/> · <https://github.com/goeludit/Data-Science>

EDUCATION

Rutgers University

Masters in Data Science and Statistics

Coursework: Regression and Time Series, Deep Learning, Database Systems, Probability,

New Brunswick, United States

August 2023 - Dec 2024 (E)

University of Mumbai (R.A. Podar College of Commerce and Economics)

Bachelor in Commerce

Coursework: Statistical Tech, Business Economics, Financial Mgmt, Marketing Research

Mumbai, Maharashtra

June 2016 - April 2019

PROFESSIONAL EXPERIENCE

Healofy

Machine Learning Product Manager

Bangalore, India

April 2021- July 2023

- Automated ETL marketing data pipeline using Branch and Google APIs for scheduled updates hosted on AWS S3.
- Implemented AutoML NLP-based transliteration and spell correction model for 8 vernacular languages.
- Built a hybrid filtering content recommendation system using matrix factorization for 7+ million daily users.
- Optimizing digital media compression algorithms decreasing cost by 10% and load times by 5% using AWS lambda.
- Spearheaded the implementation of the Healofy IOS App and the Shopify website, generating \$60k+ in revenue.
- Oversaw the end-to-end bug resolution process ensuring quicker resolution by 10% using Agile methodology.
- Established and optimized the Search (Levenshtein distance) algorithm, improving search recommendations by 30%.

Healofy

Data Analyst

Bangalore, India

Dec 2019 - April 2021

- Implemented and automated data collection methods on Google Cloud Platform and Metabase.
- Developed web scrapers for a competitive pricing strategy leading to growth in revenue by 15%.
- Reduced script's (data collection) runtime by 60% using multithreading utilizing Python's ThreadPool
- Integrated Google Analytics and analyzed data of over 20 million users from app and website to target revenue.
- Developed data visualization dashboards for daily KPIs using Google Data Studio, BigQuery connector, and Tableau.
- Devised data collection scripts using Clevertap, Mixpanel, Shopify, and other third-party APIs.
- Executed data wrangling, cleaning, transformation, and complex dataset preparation in Python.

TECHNICAL SKILLS

Computer and other Skills: Python, Java, R, SQL, Git, AWS, BigQuery, Metabase, Tableau, Agile, Azure, MongoDB, Jira

Packages: Pandas, NumPy, SciPy, Requests, Selenium, Matplotlib, Seaborn, TensorFlow, Tweepy, PyTorch, Airflow, SQLite

PROJECTS

- **Uber Price Prediction Model (Python Packages:** Statsmodel, Sklearn, Scipy, Seaborn, Matplotlib, Pandas, Numpy)
- Applied XGBoost, Random Forest, and bagging regressor to forecast Uber and Lyft ride prices using weather and ride features. Boosted accuracy to 97% R^2 score through bagging implementation.
- **Image Digit Classification Model using CNN (Python Packages:** Tensorflow, Keras, Matplotlib) - Built a TensorFlow-based Convolutional Neural Network (CNN) from scratch, achieving 98.15% accuracy on a 60k-image-trained dataset, enabling robust number recognition and output from images
- **Loan eligibility Model (Python Packages:** (Pandas, SKLearn, Matplotlib) - Implemented an SVM to assess loan eligibility, utilizing historical transaction data, credit ratings, previous loan records, etc achieving an accuracy of 85%.
- **Spotify ETL Pipeline through Airflow (Python Packages:** Airflow, SQLite3, Pandas, JSON) - Engineered a robust ETL pipeline with Python, leveraging Spotify's Open AI and SQLite3. Validated data integrity, and Automated the entire workflow ensuring daily updates to users' top artists.
- **Tweet Sentiment Analysis (NLP)(Python Packages:** Tweepy, Matplotlib, Textblob) - Using Tweepy, scraped trending tweets via the Twitter API, analyzed emotions with NLP: 50% neutral, remaining split evenly between positive and negative.