



UPENDRA GOSAVI

Data Scientist

About Me

I'm a passionate Data Scientist and Analyst with strong analytical skills. I'm always exploring and implementing cutting-edge technologies like **Machine Learning, Python, Power BI, Excel, NLP, Deep Learning** and **MySQL**. I thrive on unleashing my talent and using it to its full potential. My focus is always on practical solutions that can improve technology and management.

As a Computer Science expert, I'm results-oriented and highly motivated to gather, analyze, and interpret data sets. I have a keen understanding of business needs and excel at gathering client requirements. Let me help you harness the power of data to drive your business forward!

MY CONTACT

✉ upendrasgosavi@gmail.com

☎ 8689809107

📍 Thane, Mumbai, Maharashtra

GITHUB



<https://github.com/007kakashi>

LINKEDIN



www.linkedin.com/in/upendra-gosavi-4068a2195

SKILLS

- Programming languages: **Python, SQL**, and other languages relevant to **data science**.
- Data manipulation and analysis: Experience with **data cleaning, transformation, and analysis**.
- **Statistical modeling**: Familiarity with **statistical modeling techniques**, such as **regression analysis** and time series forecasting.
- **Machine learning**: Experience with **supervised and unsupervised machine learning** techniques.
- Data visualization: Proficiency with data visualization tools like **Power BI, Seaborn** or **Matplotlib**.

Education Background

- **EXCEL'R SOLUTION**
- July 2022 - Feb 2023 Course of Data Science & Data Analyst
- **NAVJEEVAN EDUCATION SOCIETY POLYTECHNIC**
- 2014 -2019 Diploma In Information Technology

Professional Experience

ISF Analytica and Informatica (Data Scientist)

April 2023 - Present

Key responsibilities:

- Utilized Looker Studio and Dashboard to develop interactive dashboards for EasyLearn platform.
- Created SQL views and queries for data extraction and manipulation.
- Conducted data analysis using statistical techniques to gain insights for EasyLearn platform.
- Collaborated with data science team and received mentorship.
- Pre-processed and cleaned data for analysis and reporting.
- Developed skills in Looker Studio, Dashboard, SQL, and data analysis.

Ai-Variant (Data Scientist As Intern)

Sep 2022 - feb 2023 (6 Months)

Key responsibilities:

- Pre-processed the data to prepare it for analysis.
- Analyzed the data using statistical techniques to gain insights.
- Gained practical experience working on other data science projects throughout the internship.
- Developed valuable skills that will help prepare for a career in data science.
- Collaborated with a team of data scientists and received mentorship from experienced professionals.

Wipro (IT-Support)

oct 2019 - aug 2021

Key responsibilities:

- Troubleshooting technical issues related to hardware, software, and networking.
- Providing excellent customer service to end-users.
- Documenting technical support activities for accurate record-keeping.
- Providing remote support using tools to assist users remotely.
- Effective time management to handle multiple requests.
- Adaptable and flexible to changing priorities and evolving technologies.
- Basic understanding of IT infrastructure to troubleshoot network-related issues.

PROJECTS

Customer Churn

- **Data collection and preprocessing:** Collect relevant data from various sources, clean and preprocess the data. Handle missing values and encode categorical variables.
- **Exploratory data analysis (EDA):** Perform EDA to understand the data and relationships between features and the target variable.
- **Feature engineering:** Create new features that can be relevant to the problem statement.
- **Model selection:** Choose the appropriate algorithm(s) for the problem. This may include supervised learning algorithms such as logistic regression, decision trees, random forest, or neural networks.
- **Model training:** Split the data into training and validation sets, train the model using the training set, and tune the model parameters to achieve optimal performance.
- **Model evaluation:** Evaluate the performance of the model on the validation set. Use appropriate metrics such as accuracy, precision, recall, or F1 score.
- **Model deployment:** Deploy the trained model in a production environment for real-world predictions.

Sentiment Analysis

- **Data collection:** Collect the necessary data from the e-commerce website by web scraping. This includes extracting the product name, customer reviews, and ratings.
- **Data preprocessing:** Preprocess the data by cleaning it, removing unwanted characters, and transforming the data into a format suitable for NLP algorithms.
- **Exploratory data analysis (EDA):** Perform EDA to get insights into the data, such as identifying the most frequent words in the reviews and their sentiment.
- **Feature engineering:** Create new features from the existing ones that may be relevant to the problem. This may include creating a bag of words or TF-IDF matrix.
- **Model selection:** Choose the appropriate NLP algorithm(s) for the problem. This may include sentiment analysis algorithms such as Naive Bayes, Logistic Regression, or Support Vector Machines.
- **Model training:** Split the data into training and testing sets and train the model using the training set. Tune the model parameters to achieve optimal performance.
- **Model evaluation:** Evaluate the performance of the model on the testing set. Use appropriate metrics such as accuracy, precision, recall, or F1 score.
- **Model deployment:** Deploy the trained model to predict the sentiment of new reviews.

Leaf Disease Detector

- **Data collection:** Gathered a comprehensive dataset comprising images of both healthy and diseased leaves from various sources, including field photographs.
- **Data preprocessing:** Preprocess the data by resizing the images, normalizing pixel values, and splitting the data into training, validation, and testing sets.
- **Model architecture selection:** Choose the appropriate CNN architecture for the problem, such as VGG, ResNet, or Inception.
- **Model training:** Train the model using the training set, using techniques such as data augmentation, transfer learning, or fine-tuning.
- **Model evaluation:** Evaluate the performance of the model on the validation and testing sets using metrics such as accuracy, precision, recall, or F1 score.
- **Model deployment:** Deploy the trained model in a production environment to detect leaf diseases automatically.