



# SAHIL JOSAN

## Aspiring Data Scientist | Data Analyst

8 years of Industry Experience with 1 year of internship as a Data Scientist in ineuron.ai, looking for transition to the Data Science Role. With interest in analyzing raw data and statistics, brings excellent communication abilities with hands-on experience in building machine learning models with deployment and creating dashboards with a story to tell. Working Experience and knowledge in python with libraries such as sklearn, Numpy, Pandas, Matplotlib, seaborn. Worked on tools like- Pycharm, Visual Studio, Jupyter Notebook, Power BI

- 
- 📞 +91 98250 15084
  - ✉️ sahijosan50@gmail.com
  - 🔗 <https://www.linkedin.com/in/sahil-josan-a1256622b/>
  - 🔗 <https://github.com/Sahiljosan>
  - 🌐 <https://datawithsahil.github.io/>
  - 📍 Ahmedabad, Gujarat, India

## Key Skills

Data Analysis & Visualization | Predictive Analysis | Statistical Modeling | Web Scraping  
Data Preprocessing | Clustering  
Machine Learning Algorithms | Model Deployment

## Technical Skills

### Tools:

Python | MySQL | Pycharm | Jupyter Notebook  
MongoDB | Google Colab | VS Code

### Packages:

- ✓ Scikit-Learn
- ✓ Tensorflow
- ✓ Numpy
- ✓ Pandas
- ✓ Seaborn
- ✓ Matplotlib
- ✓ Autoscraper ,  
Beautiful soup

### Web Framework:

- ✓ Flask

### Machine Learning Algorithms:

Linear Regression | Logistic Regression  
K Nearest Neighbors | Support Vector Machine | Naive Bayes | Decision Tree  
Random Forest | Boosting Algorithms | many more..

### Hackathons:

Data Science Hackathon, organised by Deets Digital Pvt Ltd at IIM Ahmedabad. The problem statement was to predict the future defaulter based on bank credit card records.

[[Certicate](#)]

## Education

B.Tech in Electronics & Communication Eng.  
**GITS, Gujarat | 2011 - 2014 | CGPA 8.14**

Diploma in Electronics & Communication Eng.  
**LE College, Gujarat | 2008 - 2011 | CGPA 7.82**

## Work Experience

### Data Science Intern | Ineuron.ai | May 2022 - Present

I have learned Python, Statistics, MongoDB, Mysql, in ineuron, learned the mathematical intuition behind supervised, unsupervised machine learning and clustering algorithms.

- **Project:** Content Based Movie Recommendation System [Source Code](#)
  - Gathered **4803 Movies** dataset from TMBD. Extracted genres, keywords, cast (Top 3 actors) and crew and include all these to Tags Feature.
  - For Text Vectorisation we have used **CountVectorizer**. Perform stemming using **PorterStemmer**, so that the words will not repeat in tags.
  - Find the distance between all vectors of all movies using **cosine\_similarity**. and Top 5 movies having lower distance are recommended.
  - Used **Streamlit** to create web app and deployment
- **Project:** Predict APS Failure at Scania Trucks using ML Algorithms [Source Code](#)
  - Used dataset having readings of 171 sensors.
  - After **EDA**, **Preprocessing**, **Feature Selection** and **Handling the outliers**, We seperated dependent and independent features.
  - For **Model Building** i have used different classifiers as this was a classification problem statement and finds that **XG Boost** gives highest accuracy of 99%
  - To be absolutely sure and to avoid model from over-fitting, i have done **Hyperparameter tunning** using **GridSearchCV** and again the accuracy of **XGBoost** Classifier was maximum (99%)
- **Project:** Predict the car price and build Flask Web App for Deployment [Source Code](#)
  - Used dataset from kaggle having 892 records and 6 features.
  - The data has many categorical features so we have used **OneHotEncoder** for encoding all features.
  - For model building we have used **linear regression** but at first our **r2\_score** was very less 57%, so we run a for loop of 1-1000 to check r2\_score for different values of **random\_state**. and finds that at random\_state 433 r2\_score id max 84%
  - We store the model in pickle file and made **Flask web app** for deployment in **Heroku**.
- **Project:** Web Scrapping Using Python [Source Code](#)
  - Scrapped <https://github.com/topics>
  - Used **requests** to download the page
  - Used **BeautifulSoup** from **BS4** to parse and extract information
  - For each topics, extracted Topic Title, Topic Page URL, and Topic Description
  - For each topics, extracted Top 20 repositories.
  - For each repository, graped the Repo name, username, stars, and repo URL
  - Converted the extracted information into a pandas dataframe and for each topic, created a CSV file.

### Former Business Owner | MG Engineering | July 2014 - Jan 2022

- Create business plans, arrange financing, hire staff, review sales, develop marketing strategies, oversee daily activity, and identify business opportunitites.
- Do continous research of new materials that can be used to manufacture Tiles.
- Do line trails of existing products in different companies (part of B2B marketing).
- Satisfactorily handling a team of 6-7 workers and 4-5 marketing technicians.
- Ensure that equipments and machinery are adequately maintained and promptly repaired by assigned workers.



# SAHIL JOSAN

Aspiring Data Scientist | Data Analyst

## Work Experience

### Freelance Editor | May 2020 - Present

- 3 years of experience as a video editor with a proven track record of creating high-quality video content for various clients.
- Worked with clients to understand their video needs and created content that met their goals and exceeded their expectations.
- Collaborated with other team members to ensure that videos were completed on time and within budget
- Editing online freelancing projects from upwork and other freelancing websites. Maintained 90% Job Completion rate on upwork.
- Managing youtube channel of INNOFIED, a kolkata based startup.

### Video Editor & Motion Graphics Designer | V3 Production | 2019 - 2020

- Performed post-production work, including, assembling raw footage, Editing, sound and sequences and providing final film within time and budget constraints.
- Have done many outdoor projects including aarambh event, held at tent city, statue of unity where our honourable PM Shri Narendra Modi was Chief Guest.
- Edited Interviews, Logo Animations, Infographics Videos and Corporate Films.

### Video Editor | Ansh Production | Jan 2019 - Apr 2020

- Editing educational videos for Gujarat University.

	+91 98250 15084
	sahiljosan50@gmail.com
	<a href="https://www.linkedin.com/in/sahil-josan-a1256622b/">https://www.linkedin.com/in/sahil-josan-a1256622b/</a>
	<a href="https://github.com/Sahiljosan">https://github.com/Sahiljosan</a>
	<a href="https://datawithsahil.github.io/">https://datawithsahil.github.io/</a>
	Ahmedabad, Gujarat, India

## Language:

English	
Hindi	
Gujarati	
Punjabi	

## Creative Skills

ADOBE PHOTOSHOP	
ADOBE ILLUSTRATOR	
ADOBE PREMIERE PRO	
ADOBE AFTER EFFECTS	
ADOBE AUDITION	
CORAL DRAW	
CINEMA 4D	