



## RISHI SHAH

### ASPIRING DATA SCIENTIST/MACHINE LEARNING ENGINEER

Data Science and Machine Learning enthusiast, eager to learn and make a growth in this domain. I am highly motivated, curious and detail oriented analytic individual with over 3+ years of experience in Data Preprocessing, Data Analyzing, Building Model, SQL, Oracle ERP & PL/SQL. I have in-depth knowledge of building machine learning, deep learning models and their deployment. This makes me a perfect blend for data science role in your organization. I am waiting for the right opportunity to prove my skills and grow with the establishment.

## CONTACT



Golghar, Maidagin, Varanasi  
UP, INDIA



+91 7701871214  
+91 7030981582



<https://www.linkedin.com/in/rishi-shah-54b96013a>



<https://github.com/RishiShah25>



[rishi.gujrati@gmail.com](mailto:rishi.gujrati@gmail.com)

## WORK

**Designation:** Associate Consultant

**Company:** Capgemini India Private Limited

### Role & Responsibility:

1. Part of the team, which created Machine Learning model, based on the employee data, whether they will attend the seminar or not.
2. Access and compile data sets from various sources for exploratory and pre-defined analysis.
3. Audit data sets for completeness, validity, and other measures
4. Consult with stakeholders or client on data availability and quality and outcome of analysis.
5. Consulted and worked with development team to determine, execute and deliver relevant solution.
6. Independently designed and executed reports (RDF, RTF) for infrastructure support and development. Support to all PO, AP, AR, HRMS, GL and FA activities.
7. Advocated for end-users to perform testing and problem analysis for server, desktop, and IT infrastructure work.

## PROJECTS

### Project 1: Road Traffic Severity Classification

**Description** - This data set is collected from Addis Ababa Sub-city police departments for master's research work. The target feature is Accident severity, which is a multi-class variable. The task was to classify this variable based on the other 31 features step-by-step by going through each day's task. The metric for evaluation is f1-score.

### Project 2: Patient Survival Prediction

**Description** - Knowledge about chronic conditions can inform clinical decisions about patient care and ultimately improve patient's survival outcomes. The target feature is hospital death, which is a binary variable. The task was to classify this variable based on the other 84 features step-by-step by going through each day's task. The scoring metric is Accuracy/Area under ROC curve.

### Project 3: Cyberbullying Detection

**Description:** This dataset is a collection of datasets from different sources related to the automatic detection of cyber-bullying. The data contains different types of cyber-bullying like hate speech, aggression, insults, and toxicity. We are provided with twitter\_parsed\_tweets: you have to perform a systematic NLP approach to identify the toxicity of the tweet and classify the tweet in a binary value.

## SKILLS

- Machine learning
- Deep Learning
- Data Analysis & Visualization
- Natural Language Processing
- Data Preprocessing
- SQL
- PL/SQL
- Python
- Feature Engineering
- OS - Ubuntu, Windows 10
- Google Collaborator
- VS Code
- AWS (Basic)
- Streamlit
- Numpy
- Pandas
- Scikit-Learn
- Matplotlib
- Tensor flow
- Jupyter Notebook

## EDUCATION

- Bachelor of Engineering – (2014-2018) Electronics Engineering
- G.H Rasoni College of Engineering, Nagpur
- CGPA=8.7

## HOBBIE

- Trekking
- Swimming
- Dancing
- Cooking
- Reading

## LANGUAGE KNOWN

- Hindi
- English
- Gujarati

## PERSONAL DETAIL

- Father's Name: Navin Shah
- Mother's Name: Manju Shah
- DOB: 25th-June-1996
- Gender: Male
- Marital Status: Unmarried
- Nationality: Indian

## INTERNSHIPS

---

- **Company** - The Machine learning Company  
**Duration** – 3 months  
**Roles & Responsibility** – The Company provides the real-world problem statement every week and we need to perform task to simplify the data and provide an insightful model.
- **Company** - iNeuron  
**Duration** – 1 months  
**Roles & Responsibility** – Perform a task and develop a model on given problem statement by the company to help them improve their performance in the educational field.
- **Company** - Regional Remote Sensing Centre (ISRO)  
**Duration** – 6 months  
**Roles & Responsibility** – The basic objective of project is to map water surfaces and monitor these water surfaces across India using Sentinel-1 ESA satellite. The objective of analysis part of the project is that it will exactly tell whether water bodies are being extracted properly or is there any deformation in the shape, size of the water body.
- **Company** - INFOCEPTS  
**Duration** – 4 months  
**Roles & Responsibility** – Fuel Analysis dashboard analyses indicators, such as production, consumption, price etc. that provides valuable insights to oil and gas sectors, effective policy making for government, private industries, customers, and administrators. It presents the most current and accurate data available and includes national and regional estimates which helps in providing predictive analysis.

## CERTIFICATES AND ACHIEVEMENTS

---

1. Project Star Award (June-2021)
2. Best Project (September -2017)
3. The Machine learning Company experience certificate
4. iNeuron internship completion certificate
5. What is Data Science? (Aug -2021)
6. The Data Science Course 2021
7. Agile Software Development