

RESUME

SOURABH KUMAR PRAJAPATI

PERMANENT ADDRESS

J-79 New Colony, Churk- Sonebhadra(U.P.)

Mobile : (+91) 87070-86797

Mail Id : prajapatisourabh22@gmail.com

Linkedin Id :

<https://www.linkedin.com/in/sourabh-prajapati-51129918a/>

Github Id:

<https://github.com/sourabhprajapati22>

CARRIER OBJECTIVE

As a fresh Data Scientist, I have a strong foundation in statistics, machine learning, and programming. I have experience in collecting, cleaning, and preprocessing data to derive meaningful insights and draw data-driven conclusions. I am proficient in programming languages such as Python and have hands-on experience with popular data analysis tools and libraries such as pandas, NumPy, and scikit-learn, TensorFlow. I am always eager to learn and continuously improve my skills, making me a quick learner and a valuable asset to any team.

PROFILE SYNOPSIS

Highly motivated, strong work ethics, fast learner, strong analytic & communication skills with positive approach & innovate ideas and have an invaluable strength essential to qualify performance on job.

PROJECTS

- ❖ [ICR - Identifying Age-Related Conditions](#) **Nov 28, 2023**
Technologies:- Catboost
Methods:- StandardScaler, CatboostClassifier
Zindi Rank- 10 out of 6431
score- 0.35326 (Private score)
- ❖ [9th-place-GEO-AI-Challenge-for-Cropland-Mapping](#) **Oct 14, 2023**
Technologies:- XGboost, LightGbm, Catboost
Methods:- ensemble techniques use to modify the accuracy
Zindi Rank- 9 out of 72
score- 0.9266 (Private score)
- ❖ [HR Analytics PowerBI](#) **May 16, 2023**
Technologies:- PowerBI (Microsoft)
- ❖ [ML Project CI/CD Pipeline](#) **March 18, 2023**
Technologies:- Python-3.8.0, Flask, Github, VsCode, Kubernetes, Docker (Google Cloud) ,Anaconda and ML libraries
- ❖ [AgriFieldNet India Challenge](#) **Oct 31, 2022**
Technologies:- XGboost, LightGbm
Methods:- ensemble techniques use to modify the accuracy
Zindi Rank- 22 out of 179
score- 1.378947408 (Private score)

- ❖ [UmojaHack India Income Prediction Challenge by UmojaHack Africa](#) **Oct 22, 2022**
 Train accuracy:- 97.04%
 Test accuracy:- 95.83%
 Technologies:- CatboostClassifier Methods:- iterations and learning rate modify
 Zindi Rank- 28 out of 128
 F1 score- 0.615538512 (Private score)
 F1 score- 0.627316857 (Private score after competition close)
- ❖ [Wheat Kernel Classification](#)
 Train accuracy:- 97.62%
 Test accuracy:- 95.24%
 Technologies:- StandardScaler, LogisticRegressor, RidgeRegressor, PCA, LDA
 Methods:- Without scaling and with scaling data analysis.
- ❖ [Email Classification](#)
 Train accuracy:- 99.10%
 Test accuracy:- 98.47%
 Technologies:- NLTK, MultinomialNaiveBayes, Pandas. Methods:- Stopwords, Stemming, Vectorization
- ❖ Malbourne house price prediction (Kaggle Project).
 Tools:- Jupyter notebook, Github Methods:- RandomForestRegressor
- ❖ Digit recognizer using **LogisticRegression** and also done using Artificial Neural Network(**ANN**).
Developd these POC for the Internship purpose (fresher).

EDUCATION

B. Tech (Mechanical Engineering) 69.4% Institute of Engineering and Technology, Lucknow (U.P.)	2016-19
Diploma (Mechanical Production) 70.5% Govt. Polytechnic, Lucknow (U.P.)	2013-16
Intermediate 76.6% U.P. Board	2012
High School 71% U.P. Board	2010

ACHIVEMENTS

- ❖ I found-out another **Low Severity** bug in Infosys. **Feb 22, 2022**
- ❖ I found-out the **High Severity** bug in one of the Maharatna company **IOCL**. **Aug 8, 2021**

CERTIFICATIONS

- ❖ [Intro to SQL \(BigQuery\)](#) Kaggle **March 20, 2023**
- ❖ [MySQL Essential Training](#) LinkedIn **Dec 9, 2022**
- ❖ [Intermediate Machine Learning](#) Kaggle **July 2, 2022**
- ❖ [Intro to Machine Learning](#) Kaggle **June 1, 2022**
- ❖ [Python For Machine Learning](#) Great Learning **June 2022**

- ❖ [End-to-End Machine Learning with TensorFlow on GCP](#) Google Cloud (Coursera) **May 4, 2022**
- ❖ [My SQL Basics](#) Great Learning **May, 2022**
- ❖ [Data Analytics on AWS](#) Amazon Web Services (Coursera) **Apr 15, 2022**
- ❖ [Machine Learning](#) 11week course Stanford University (Coursera) **Feb 8, 2022**

WORK EXPERIENCE

Current Company Name- PPS International, Greater Noida **Jun2021-July2022**

Graduate Engineer Trainee Supervised the work of other members, and maintain the quality of products.

TECHNICAL SKIL

Python: Pandas, Numpy, Scikitlearn, Tensorflow, Keras.

MachineLearning: Supervised (LinearRegression, LogisticRegression, Regularisation) Unsupervised (K clustering), Data Compression (PCA), xgboost, catboost, lightgbm.

DeepLearning: ANN, CNN, RNN, Xception, VGG , YoLo, U-net, Image classification, Image detection.

NLP: Tokenization, StopWords, Stemming, Lemmatization, NLTK, Bag of words (DictVectorizer, CountVecitorizer, Tf-IdfVectorizer).

Database: MySQL, BigQuery.

CI/CD Pipeline, PowerBI

HyperParameterTuning: GridSearchCv, Pipeline.
Some basic knowledge of **Cyber Security**.

ADDITIONAL ACTIVITIES

- ❖ My final year B.Tech. project was **3D-PRINTER** completed in 2019.
- ❖ Passed the **NCC 'C' Certificate** examination held in 2012.
- ❖ Passed the **NCC 'B' Certificate** examination held in 2011.

PERSONAL PROFILE

- ❖ **DATE OF BIRTH** : September 22nd , 1995
- ❖ **MARITAL STATUS** : Unmarried
- ❖ **GENDER** : Male
- ❖ **FATHER** : Mr. Vijay Kumar Prajapati ~ Teacher
- ❖ **MOTHER** : Mrs. Kiran Prajapati ~ Housemaker

DECLARATION

I hereby declare that the foregoing information is correct to the best of my knowledge and belief. I have not suppressed any material fact of factual information in the above statement. References can be provided upon request.

Oct 2023

SOURABH KUMAR PRAJAPATI