

# Satyam Tripathi

## Aspiring Data Scientist

8 months+ experienced in Data Science intern, looking for a transition to the Data Scientist Role. I have experience in data analysis to get insights, data visualization, data preprocessing i also have knowledge of machine learning models and NLP and Deep Learning models. I am waiting for the right opportunity to prove my skills and grow with the organization



<https://github.com/Satyamwork43>



7895084559



[satyamwork43@gmail.com](mailto:satyamwork43@gmail.com)



<https://www.linkedin.com/in/satyam-tripathi-1141111a4/>



<https://medium.com/@satyamwork43>

## WORK EXPERIENCE

### Data Science Intern | Flip Robo Technologies | May 2021-PRESENT

#### Project: MICRO-CREDIT DEFAULTER PROJECT (END TO END)

**Problem Statement:** The client wants to know whether their users are paying the loaned amount within the time frame or not, if not what are the criteria.

**Solution:** Did EDA, Treated missing values, removed outliers, removed skewness, feature engineering, feature transformation, data was imbalanced so i balanced the data with SMOTE, did hyperparameter tuning and finally deploy the model using streamlit on heroku.

### Deep Learning Intern | IIT Jammu | March 2021-May 2021

#### Project: Stone Pelting (END TO END)

**Problem Statement:** As in Kashmir people used to do stone-pelting so to stop it make a model that can predict whether stone pelting is happening or not

**Solution:** First i have collected data using google and downloaded videos from youtube and extracted framed from that video as i have to give two classes for my model i also collected different images of all types after that i build a CNN model which was able to understand whether the stone-pelting is happening or not.

## Personal Projects

### Project: Chatbot

**Problem Statement:** The client wants to build a chatbot for Scaling Businesses Globally.

**Solution:** created a custom data as per client questions and answers using json file and did all the preprocessing and build a model using RNN LSTM.

### Project: Census Income Prediction

**Problem Statement:** The prediction task is to determine whether a person makes over \$50K a year or not.

**Solution:** First i have collected the data and did all the preprocessing then did EDA on data to understand the problem and to understand the effect of every feature on my target column and finally i build a model and deployed it on heroku

### Project: Video Content Summarization To Text

**Problem Statement:** Text summarization refers to the technique of shortening long pieces of text. The intention is to create a coherent and fluent summary having only the main points outlined in the document.

**Solution:** I have used transformer here and YouTubeTranscriptApi and i used a video from youtube then i summarized the video into text.

### Project: Sign Language Detection

**Problem Statement:** Conversion of sign language into words by an algorithm or a model can help bridge the gap between people with hearing or speaking impairment and the rest of the world.

**Solution:** First i have collected the data of different sign languages and then did data augmentation to increase the size of my data then i used labelling for labeling of my data then finally i have build a model using yolov5l model.

### Project: Movie Recommendations System

**Objective:** To provide recommendations based on recorded information on the users' preferences.

**Solution:** Collected data using web scraping from imdb and build my recommendation system using cosine similarity.

## EDUCATION

### B.Tech[Mechanical Engineering]

A.K.T University, Lucknow  
2017-2020

## KEY SKILLS

**Machine Learning | Data Analysis & Visualization | Statistics | Deep Learning | NLP | Object Detection | Feature Engineering | Feature Selection | Data Pre-Processing | Python | Computer Vision | Time Series**

## TECHNICAL SKILLS

**Scripting Language:** Python

**Database:** SQL

**Machine Learning:** Supervised & Unsupervised Learning

**Deep Learning:** ANN | CNN | RNN | Object Detection | YOLO | TFOD | Fast RCNN | Faster RCNN

**NLP:** Text Normalization (Stemming & Lemmatization) | NLTK | Word Embedding | Transformers | BERT | ALBERT | DistilBERT

**Cloud Deployment:** Heroku

**Web Framework:** Streamlit

## CERTIFICATION

**PG Program in Data Science, Machine Learning & Neural Networks in collaboration with IBM**

**DATATRAINED | Jan 2021 - Nov 2021**

## ACHIEVEMENTS

**HACKATHONS** -Predict The Flight Ticket Price Hackathon.

**RANK** - 250 Out of 3000

**HACKATHONS** -Predict The Price Of Books

**RANK** - 350 Out of 3500

**HACKATHONS** - Predicting House Prices In Bengaluru

**RANK** - 270 Out of 2500

## Software Related Tools

**Libraries**-Pandas, Numpy, Matplotlib, Seaborn, Scipy, Scikit-Learn, Keras, Tensorflow, Transformers, Spacy, NLTK, Textblob

## PERSONAL TRAITS

**Enthusiastic and self learner**

**Good Communications**

**Adaptability**