

# Pratik Geoffrey Saxena

91-9890957202

saxenageoffrey@gmail.com

January 24, 2021

geoffreysaxena.github.io

A budding Analyst & ML Engineer

## EDUCATION

### University of Petroleum and Energy Studies

*B.Tech in Computer Science and Engineering; CGPA: 7.68*

Dehradun, IN

*Jun. 2018 – Current*

## TECHNICAL SKILLS

**Languages:** C, C++, Java, Python, SAS, MySQL, SQLite, TensorFlow.js

**Data Science:** Statistical modeling and inference, Pipeline construction, ETL, Data Wrangling

**Machine Learning:** TensorFlow, Sklearn, SciPy, Pandas, NumPy, Matplotlib

**Analytics:** Tableau, Microsoft Power BI, Microsoft Excel

**Developer Tools:** Git, VS Code, Visual Studio, PyCharm, Anaconda

**Others:** MS Office (Complete tool set), Scrapy, Streamlit, Figma

## EXPERIENCE

### STEM-Away

Santa Clara, CA

*Artificial Intelligence Intern*

*Jun. 2020 - Jul. 2020*

- Leveraged **Natural Language Processing** techniques to deploy a 7 layered Neural Network.
- Utilized **Python & Scrappy** built pipeline to automate scraping & pre-processing from over 120 websites.
- Devised an **automatic web scraper** top pull and analyze data from over **8000** forum posts.
- Constructed a **pipeline** of over **30** custom data science utilities in Python, to facilitate data cleaning, visualization, modeling, cross-validation, hyper-parameter tuning, and model ensembling.

### Government of Uttarakhand

Dehradun, IN

*Data Analyst & Machine Learning Intern*

*May. 2020 - Jun. 2020*

- **Spearheaded** the development of a registry that stored sensitive data of over **17,000** villages.
- Over **55 Insights** pivotal to deploy a skill registry initiative were **extracted** using data cleaning methods.
- A **classification** model was delivered, that was further deployed over an app within a span of **60** days.

## PROJECTS

### Doodle Classification Neural Net | *Python, TensorFlow, TensorFlow.js, & p5.js*

- Constructed **Image Classifier** that classifies **345** different doodled items with about 89% accuracy.
- Trained model on **1,72,50,000** doodles with TensorFlow on Google Colaboratory.
- Converted **trained** model and **deployed** over the browser with **TensorFlow.js** and **p5.js** libraries.

### Analysis on Criminalization of Politics | *C Programming, k-means clustering, & Tableau*

- Addressed **unawareness** among general public by **analyzing** criminal data of over 300 politicians.
- Formulated a **visualization** that colored attributes of criminals and politicians using Tableau.
- Visualized GitHub data to show **collaboration** over the span of 150 days.

### Prediction of Used Car Price | *Python, SAS, & SQL*

- Wrangled, Analysed data-set with **1.75 million** records and **25** different features.
- Analysed each of the **25 features** to determine the impact on the output variable.
- Utilized **SQL** and **PROC SQL** queries to subset and clean data.
- Formulated **regression models** to predict the value of the used cars.

## POSITIONS OF RESPONSIBILITY

### Joint Secretary

Dehradun, IN

*Computer Society of India*

*Jun. 2020 - Present*

- **Counsel** the functioning of finance, logistics, event management & public relations.
- **Mentor** and **Supervise** 18 other decision makers to ensure smooth function of the organization.