

# Shivan Kumar *Associate Data Scientist*

✉ shivan18july1997@gmail.com

☎ 6280963446, 7301694118

🔗 <https://github.com/Shivan118>

in <https://www.linkedin.com/in/shivan-kumar/>

📺 <https://www.youtube.com/@engineeringwalabhaiya>

kaggle <https://www.kaggle.com/shivan118>



## Profile

Data Scientist with 3 years of experience in collecting, examining, and interpreting data and statistical figures to drive growth for companies and clients. In-depth understanding of Machine learning & Natural language processing.

## Skills

**Programming Languages:** Python Programming Language

**Technical Skills:** Machine Learning(supervised & unsupervised learning), Statistics, Deep Learning(ANN, CNN,RNN,LSTM), TensorFlow, Pytorch, NLP(BERT, GPT, ROBERTA, ALBERT, T5),MySQL, MongoDB,Git & Github, MLOps, DVC, MLFLOW, Docker, AWS, Flask, FastAPI, Data Analytics (Excel, Power BI)

## Professional Experience

01/2020 – 11/2022  
Bengaluru, India

**Associate Data Scientist, iNeuron Intelligence** 

- **Requirement Gathering & Translation:** Gathering data from various sources, collecting our customer's data, and translating it into use cases to solve real-world problems.
- **Data Cleansing, Statistical & Data Analysis:** Clean, transform & extract insights and provide **sentiment analysis** using the feedback from the courses for providing data-driven recommendations.
- **Training & Implementation:** Developing content-based course recommendation systems and an automated chat system for questions and answers. Creating machine learning models as well for various use cases supplied by the clients.
- Managed projects for **CRM, Education** and **Natural Language Processing** with end-to-end deployment using **CI-CD pipelines**.
- **Delivered expert lectures** to data science and other industry professionals on Machine learning, Natural language processing & MLOps.
- **Lead a team** of 25+ **Jr. Data Scientist** in a various data science projects

11/2019 – 01/2020  
Chandigarh, India

**Machine Learning Intern, Aleph AI**

- Build a model for a client from **Healthcare Industry** Using Natural Language Processing.
- The main aim of this project to build a application to provide a medicine to the patient based on the medical issues.
- Extract the medicine prescription name from the medicine & segregate based upon the medical issues using Natural Language Processing.

## Projects

---

### **Chatbot for Automatic Ticket Creation, As Team Lead**

- The aim of this project is to make an automated CRM tools in a supply chain system of an organization where the queries raised by the customers are solved by raising a ticket automatically if the chatbot isn't able to resolve it.
- The model is trained on the custom data which is manually labeled.
- The model is implemented using CI-CD pipeline using MLFLOW and AWS as cloud platform.
- Tech Stack: AWS, Database-MongoDB, Core Model - Distill BERT, Data - In-House Data(2.5 Million), API - Language Translation

### **Course Recommendation System, As Data Scientist**

- Based on the customer's preferred skills and the skills they acquire from their registered courses, the system is designed to suggest courses they might be interested in.
- The user will submit certain information about himself and the system will filter courses depending on the user's preferences (content-based).
- Recommended courses are decided based on the **cosine similarity** with the preferred course.

### **Abnormal Event Detection**

- Abnormal Detection in the video is one of the challenging tasks in many sectors like Banking, Security.
- The main aim of this project is to establish a video system that can distinguish abnormally behaving people from normal ones.
- This model is trained on the banking Custom Data which is manually labeled based on people's behaviors.
- Classification method using ML and Deep learning(FCNN and R-CNN used for Detection of violence in Crowd scenes)
- Tech Stack: Model - FCNN, R-CNN, Database - MongoDB, Library - OpenCV, Data - Companies Data.

### **ML Automation, As Data Scientist**

- Design a web application to automate the various operations performed in machine learning projects to solve specific problem related to supervised or unsupervised use case.
- Our application must have the capabilities to perform some specific task like: Exploratory Data Analysis, Data Preparation, Feature Engineering and Model Building.
- Model is Implemented using CI-CD pipeline with automated testing and Dockerization.

## Personal Projects

---

### **EDA Automation**

- **Create a package** to perform Exploratory data analysis Operations and feature engineering steps.
- Maintain Dashboard based on user **EDA activities** also allow users to graphical interface for visualization & execution progress.
- Design a application based upon classification and regression problems

### **Placement Checker Web Application**

- Created a web app to classify whether the user will be placed in the **our Internal placement drive or not** based upon user tech skills, experience and projects.
- The user is also given a option to select which of the given algorithms will the classification be made. (Options to choose from are SVM or Support Vector Machine, Logistic Regression and AdaBoost Classifier)
- **Tech Stack:** Python, MongoDB, Flask, AdaBoost algorithm, Docker, MFLOW and AWS, Data - Companies Data.