

Ramesh Kumar

Data Scientist

Delhi | Contact number – 8607409748

GitHub - [rameshgupta41 \(Ramesh Kumar\) · GitHub](#)

E-mail - ramesh41gupta@gmail.com

LinkedIn - [Ramesh kumar - Data Scientist - Infosys | LinkedIn](#)

Objective

My career objective is to succeed in an environment of growth and excellence and earn a job which provides me job satisfaction and self-development and help me achieve personal as well as organization goals.

Summary

I have 1.5 years' experience as Data Scientist, Machine Learning, NLP and Time Series. Working experience and extensive knowledge in python, gradio, transformers, fastApi, flask and some libraries such as sklearn, NumPy, Pandas and Matplotlib Platform and misc.: Anaconda Enterprise Edition, Jupyter Notebook, Spyder IDE, SQL Server Management Studio, Visual studio

Work experience

Data Scientist at Infosys	06/2021 - Till Date
Machine Learning training at Ministry of Electronics and Information's Technology Delhi	01/2020 - 05/2020

Education

Master of Computer Application, Computer Science Central University of Haryana, Mahendragrah	-2020
Bachelor of Computer Application, Computer Science Magadh University, Patna	-2017
Intermediate, C.B.S.E Board, DELHI	-2014
Matriculation, JNV Madhepura, C.B.S.E, Bihar	-2011

Skills

- **Python**
- **Data Analysis& Pre-processing**
- **Machine Learning**
- **NLP/ NLTK**
- **Time Series**
- **Database Management**
- **Hugging Face - Transformers**
- **Gradio**
- **Project Management**

Programming Skills

Python



NumPy



Pandas



Matplotlib



SCIKIT Learn



NPL/NLTK



Hugging Face - Transformers



Gradio



SQL Server



Projects

Infosys Projects

- 1) **BE Forecasting:** The idea of this project is forecast the employee Estimate based on past data of employee that contain Recommendation, Revenue, RTBR and other many features. This is fully Regression based Model that used for forecasting the Infosys Employee Estimate.
- 2) **Text Classification:** The idea of this model is to predict class of the Text based on AHD data, the model is based on NLP preprocessing and Word2Vec embedding techniques used for better accuracy and provide better result in comparison to TF-IDF.
- 3) **Generic Regression:** The idea of this project is creating a model, where we have passed any type of regression-dataset into model, and they predict better result based on low variance of model. Internally many models created but it generates one pickle file of model.
- 4) **Generic Time Series:** This Generic model is based on ARIMA Model. The idea behind it, we can pass any data in the form of JSON, that contains Datetime and its Values and no. of forecast period. Model predict Future data based on it.
- 5) **Sentiment Analysis:** In this model, I have used pretrained models from Hugging Face to determine the sentiments from given text like Reviews/ Feedbacks/ Comments (in bulk also) as Positive, Negative or Neutral label with accuracy.
- 6) **Word Cloud:** This project is hosted on Gradio, where i have used NLP for preprocessing on CCD data, that provides n-gram and its frequency to determine which query on high demand.

Ministry of Electronics and Information's Technology

1) Intrusion Detection Techniques using Machine Learning: -

- Analyze Format data using Machine Learning algorithm by Python Scikit-Learn
- Generating various capacity planning reports (graphical) using Python packages like NumPy, matplotlib
- Clean data and processed third part spending data into maneuverable deliverable within specific format with Excel macros and python libraries such NumPy, SQLAlchemy and matplotlib
- Retrieving data from the database through SQL as per business requirements.
- Used Python Flask framework to build modular and maintainable application.

2) Android Malware Detection Techniques using Machine Learning: -

Description: In this approach using Machine learning algorithm to detect malicious and benign apps with the help of Static features available in Android dataset. In this approach, we used different algorithms and compared which one algorithm gives a better model to classified malicious apps. This approach further developed into software and application.

Tools and Languages: Python, Pandas, matplotlib, Seaborn, SK-learn

Responsibilities: - Guide the team member to understand the requirement and allocate the work accordingly. - To analyze the requirement and prepare best suited database structure design for the project. - Create database, tables, function and stored procedure for the project. - Work on the code and develop the application and provide support to the system team for go live process.

Personal Details: -

- Date of Birth: 03/03/1995
- Nationality: India
- Languages: Hindi, English