

# Naruto uzumaki

## Data Scientist

✉ hello.ineuron@gmail.com ☎ 9923312341 📍 Telangana 🌐 <https://github.com/>

---

### PROFILE

---

I have 4+ years of work experience in Data Science with a proven ability and history of developing full-stack data science projects. Hands-on experience leveraging machine learning, deep learning, and transfer learning models to solve challenging business problems.

---

### PROFESSIONAL EXPERIENCE

---

#### **Data Scientist, MU SIGMA**

Jan 2022 – present

##### **DATA QUALITY AND VALUE ASSESSMENT**

Bangalore, India

- Designed and developed a comprehensive and versatile Data quality and Value assessment module for a US based asset management company.
- The module developed on Jupyter notebook with a Python kernel brought down the time of data quality assessment of more than 500 million rows of data from 60hours to 50mins, was compatible with different types of datasets (transaction, health-care, geolocation etc.) and generated 2 reports.
- It allowed flexibility to the user to tailor the conditions and checks according to the business and assess the quality of data

#### **THE ENERGY INSIGHTS APP**

- Developed a one stop application hosted on Predix platform for a fortune 100 energy firm which helps the business development managers throughout customer engagement journey by leveraging the data available from multiple sources to have richer conversations with customers and help them in taking better decisions to achieve their energy efficiency goals (by generating realistic project roadmap using mixed integer optimization and linear programming) .
- Handled the entire backend of the application using python.

#### **Software Engineer, iNeuron Intelligence**

Jul 2020 – May 2022

##### **CUSTOMER SEGMENTS**

Bangalore, India

- In this project, I analysed a dataset containing data on various customers' annual spending amounts of diverse product categories for internal structure using PCA and Gaussian Mixture Model clustering algorithm to best describe the variation in the different types of customers that a wholesale distributor interacts with to provide insight into how to best structure their delivery service to meet the needs of each customer. (GitHub)

#### **Software Engineer, Thales**

Aug 2020 – May 2021

Hyderabad, India

- One checkout Platform – Real time streaming application for Financial data
- Tech Stack – Flink, Python, Kafka, Oracle, Kafka Connect, Docker, AWS, Airflow
- Building one checkout platform for different business domains of Expedia i.e; Vrbo, Hotels, Car Rentals, Lodging
- Real time streaming platform captures & process financial data for Accounting and helps suppliers to track total amount to be paid via Auto Pay or Request Pay model
- Crafted a scalable streaming solution using Flink StateFun to handle ~800k booking streams each day.

#### **Project Engineer, Wipro Technologies**

Jul 2015 – Aug 2016

Worked on investment banking project for Barclays capital as client.

Bangalore, India

---

## EDUCATION

---

### **Bachelor Of Technology in Computer Science**, *Jaypee Institute Of Information Technology*

Jan 2013 – Aug 2017

Noida, Uttar Pradesh

- Cumulative GPA : 8.1/10.0
- Relevant Coursework : Algorithms and Data Structure, Database Management System, Operating System, Computer Network, Linear Algebra, Discrete Mathematics, Fuzzy Logic and Neural Network, Big Data

### **Information Technology**, *SASTRA University Thanjavur*

May 2005 – Apr 2009

Thanjavur, India

- Cumulative GPA : 9.05/10.0
- Relevant Coursework : Machine Learning, Deep Learning, Probability and Applied Statistics, Advance Data Structure and Algorithms, Natural Language Processing, Medical Image Processing

---

## CERTIFICATES

---

- Elastic Search Silver contributor (06/2020 - Present)

---

## SKILLS

---

**Apache Spark** | **Microservices** | **Kafka** | **Data Warehousing** | **Event Streaming** | **Machine Learning**

**Deep Learning** | **Azure** | **Data Ingestion** | **Distributed Computing**

**Data Science Pipeline** ((*collection, cleansing, visualization, modeling, interpretation*), *Statistics, Hypothesis Testing, CI/CD Pipeline*)

---

## INTERESTS

---

Playing Badminton, Tennis, Football | Writing Tech Blogs | Traveling | Movies