# Shraddha Anala

Data Scientist

shraddha.anala@outlook.com
https://github.com/shraddha-an
https://www.linkedin.com/in/shraddha-anala/

## SUMMARY

A data science professional proficient in Machine Learning, Deep Learning, NLP and Time-Series analysis looking for a Machine Learning/Data Scientist role.

Highly adept at analysis, applying ML techniques and developing algorithms to solve real-world business problems.

## KEY SKILLS

- **Machine Learning**
- **Data Analysis**
- **Deep Learning**
- **NLP**
- **Problem Solving**
- **Teamwork**
- **Communication**
- **Adaptability**

## TECHNICAL SKILLS

- **Tools:** Python, SQL, Git, Bash.
- **Libraries:** Keras, PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas, Statsmodels, HuggingFace Transformers, Gensim, fastText, Seaborn, Plotly
- **Models Built:** Regression, Classification, Artificial Neural Networks, Convolutional Neural Networks, LSTM, Word Embeddings, Topic Modelling.

## COURSES

- **Deep Learning A-Z: Hands-On Artificial Neural Networks** | Udemy
- **Machine Learning A-Z: Hands-on Python & R in Data Science** | Udemy
- **Statistical Inference** | Coursera

## BLOG

Data Science Blog:
**https://medium.com/@shraddha.anala**

## EXPERIENCE

### Freelance Data Scientist — Present

- Building a deep CNN + LSTM model to predict tick movements of Limit Order Books.
- In-depth Data Analysis and Data Cleaning project to extract useful insights.
- Churn Prediction model with EDA to identify factors leading to attrition of employees.

### Application Development Associate — 2019

**Accenture** — Bengaluru

- Database Admin/ Technical Support of ATCAT tool for over 30 clients.
- Identifying duplicate payments from ERP and incidents resolution.

### Control Systems Intern — 2017

**Schneider Electric** — Mumbai

- Industrial training on DCS and ESD Automation Systems.
- Learnt about how these systems contributed to plant-wide control operations and safety systems in oil refineries.

## PROJECTS

### 1) Classifying the quality of Stack Overflow Questions

In this project, I fine-tuned the BERT Base model to classify the quality of questions asked on Stack Overflow into 3 different categories. I chose this dataset as I wanted to explore the performance of an attention-based model on text belonging to a niche domain, in this case, questions related to programming.

### 2) Time-Series Forecasting of Stock Prices

Implemented several models to forecast Microsoft stock prices, such as the multivariate Vector Autoregression model and different univariate models such as the ARIMA, AR and Regression, along with neural models including an LSTM network and a deep CNN.

### 3) Pre-trained vs Trained Word Embeddings

A case study comparing the performance of the GloVe pre-trained word embedding against training own embedding on the text data. I wanted to view how a general embedding would fare against a locally-trained embedding and neural network for classifying the quality of Stack Overflow questions.

## EDUCATION

### B.Tech - Electronics and Communications Engineering — 2010

**Gandhi Institute of Technology and Management** — Visakhapatnam

- Graduated **First Class with Distinction** and a CGPA of 8.37.