

**Sandeep Kumar Are**

**Mobile: +91 7989520772**

**E-mail: [sandeepkumar.are16@gmail.com](mailto:sandeepkumar.are16@gmail.com)**

**[LinkedIn](#)**

**[Personal website](#)**

---

### **CAREER OBJECTIVE**

A proactive and fast learning individual seeking an opportunity to work as a dynamic data analyst utilizing analytical & methodical skills and relevant expertise to help the company achieve business goals while sticking to vision, mission, and values.

### **PROFESSIONAL SUMMARY**

- Total 9+ years' experience in IT software Development field and in that 4+ years in Python, Machine Learning with large data sets of structured data, data validation, predictive modeling and data visualization.
- Conduct text-based analysis to derive insights and patterns from unstructured data sources.
- Perform data preprocessing tasks, including data cleaning, feature engineering, and transformation to prepare data for modeling.
- Collaborate with cross-functional teams to understand business requirements and translate them into data science solutions.
- Implement machine learning algorithms for classification and regression tasks, optimizing model performance and accuracy.
- Evaluate and select appropriate machine learning models and algorithms based on project requirements and data characteristics.
- Utilize Qlik Sense at a beginner level to develop interactive data visualizations and dashboards for stakeholders.
- Implement MLOps practices, including version control, continuous integration/continuous deployment (CI/CD), and model monitoring to ensure efficient and scalable machine learning workflows.
- Leverage AWS services such as EC2, S3, SageMaker, and Lambda for data processing, model training, and deployment in a cloud environment.
- Develop and maintain documentation for data science projects, including model architectures, workflows, and deployment procedures.
- Stay updated on emerging technologies and best practices in data science, machine learning, and cloud computing to drive innovation and enhance capabilities.

### **Technical Skills:**

- |                            |                                     |
|----------------------------|-------------------------------------|
| • Predictive Analytics     | • Natural Language Processing (NLP) |
| • Regression Analysis      | • ML & DL Techniques (RNN and LSTM) |
| • Transformer Based Models | • Time Series Analysis              |
| • Statistical Modeling     | • Data Visualization                |
| • Python/R Programming     | • SQL / PostgreSQL Sql              |
| • Qlik Sense               | • Deep Learning (CNN, RNN)          |
| • MLOPS                    | • AWS - S3                          |

### **EDUCATIONAL QUALIFICATION**

★ B. Tech Electronics & Communications engineering **JNTU University, India.** 2012

## **WORK EXPERIENCE**

### **Projects:**

#### **SENIOR SOFTWARE ENGINEER, 06/2022 - Current**

##### **ATMECS TECHNOLOGIES PVT LTD, Hyderabad, Telangana**

- Involved in multiple project based client requirements.
- Worked on Supervised and unsupervised, multi-class classification problems.
- Worked on Payroll data for predictive analysis and implementation of classification models
- Worked on Residual data to understand Anomalies detection.
- Implemented NLP NLTK , Spacy library and Transformed based models.
- Model development, Deployment and Monitoring model performance.
- Created dashboards through Qlik-Sense and Integrated the results.
- Embedded test-driven development practices into existing processes to elevate quality of code.

#### **SENIOR SOFTWARE ENGINEER, 02/2020 - 06/2022**

##### **Svobodha Infinity Private Limited (SAVART), Hyderabad, Telangana**

##### **Qualitative research (Iris)**

Iris uses textual, visual, and graphical data to generate insights and open threads, aspects that are critical to the investment but often elusive to spot and blindsiding analysts. For example, say a multi-national IT firm has 308 subsidiaries and is suspected to have shell companies evading taxes. The probability of spotting it is difficult and time-intensive for a group of research analysts but a matter of few milliseconds for Iris. It is obvious that the subsidiary would not be tagged as a 'shell company' by the firm itself. So, the approach undertaken by Iris entails capturing subsidiary data, compliance of the local jurisdictions, company supplied and third-party sourced data to pinpoint the non-compliances and then make assessments on the matter. However, scope of Iris includes not just legal jargon but around 850 topics including ethics, corporate governance, brand, moat and innovation.

##### **Roles and responsibilities:**

- I have involved scrap the data from various sites, API's and some more resources using Python Selenium.
- Involved in working with Python-MongoDB varioustasks.
- Involved in Topic Modeling with NLP.
- Worked with team on BERTModel.
- I have experience in working with Flask-Python for creation ofUI.

### **QUANTITATIVE ANALYSIS:**

The most popular form of fundamental research is to look at numerical data. Traditional analysts look at Profit & Loss statements, Balance Sheets and Cash Flow statements to make decisions, which are often based on (biased) assumptions and estimates. This led to hedge

funds building 'innovative' trading algorithms that analyze over millions of data points to make trading decisions. Savart has gone a step further and brought this technology for long term investments i.e., using billions of data points and not just back-test the strategies but build patterns and strategies from scratch using machine learning. While this does not sound like an enormous difference, the consequence is that the system generated strategy is far less biased and is thoroughly stress tested using not only historical data but monitored and tested in real time. Concisely, we input tons of raw data points and Quant outputs a shortlisted portfolios which is ready to be pushed through to Iris.

**Roles and responsibilities:**

- I have involved in what are approaches are gives the best results.
- Involved in Data cleansing, Data Cleaning, Data Analysis.
- Involved in Feature selection Models.
- Involved in Forecasting Models selections to predictions.

**SOFTWARE ENGINEER, 02/2015 - 01/2020**

**Netpeach Technologies Pvt Ltd, Hyderabad, Telangana**

**Project Name: Uncover****Autism Project****Description:**

The goal of the project is to help society mitigate the risk of autism in new born babies by identifying more reliable causes of autism. The project will be exploring existing research, and available information with feedback(surveys) from the patient population to define more reliable causes. The purpose of this study is to provide the healthcare industry with any findings discovered during analysis.

**Tools and Technologies:**

- Environments Anaconda Jupiter Notebook, NLTK tools.
- Numerical and Visualization libraries NumPy, Pandas, Matplotlib, Seaborn.
- Machine Learning libraries Scikit-learn.

**Projects:****Loan Predictions:**

Given historical data on loans given out with information on whether or not the borrower defaulted (charge-off), can we build a model that can predict whether or not a borrower will pay back their loan? This way in the future when we get a new potential customer, we can assess whether they are likely to pay back the loan.

\*This is which I have completed while doing TensorFlow 2.0 exercise from UDEMY Certifications. Worked with: Tensor Flow 2.0 , Feed Forward Neural Network(ANN), Keras with TF.

<https://colab.research.google.com/drive/1HbwYVE5erBLsqdwO1xqqYYRX1GqB5IFL?usp=sharing>

**My-Portfolio(Personal Website) -**

<https://sandeepkumar16nlp.github.io/sandeepmyportfolion.github.io/>