Data Scientist

Senthil Kumar S

Trichy, India
ssenth14@gmail.com
Ph: +91-9962463755

Summary

I have 23+ years of experience in information technology solutions across diverse industries and more than 5+ years working as a Data Scientist; Knowledgeable in artificial intelligence, machine learning, and Deep Learning with extensive technical knowledge able to serve as a trusted and valuable consultant; Collaborates with all levels of internal and external developers and project stakeholders; working in the development of new products; building machine Learning/data, and driving integrated solutions that encourage long-term business growth.

Education

Thiagarajar College of Engineering, Madurai,
Kamaraj University, Madurai, India
Bachelor of Engineering in Computer Science (1995-1995)

Professional Skills:

Libraries Frameworks: Python, pandas, numpy, Scikit-Learn, Natural Language Toolkit (NLTK), Natural Language Understanding (NLU), Natural Language generation (NLG), SQL, Seaborn, Tensor flow, Keras, Db2,Tableau
ML/Neural Networks: Regression techniques, Linear regression, Logistics, Decision Tree Learning, Neural Networks, Random Forest, SVM, Naïve Bayes, KNN, RNN, CNN.
RDBMS **:** DB2/Oracle 12c

Employment History

Freelancer, India
Data Scientist (Feb 2020- till date)

Perform Data profiling, preliminary data analysis and handle anomalies such as missing, duplicates, outliers, and imputed irrelevant data.
Remove outliers using Proximity Distance and Density based techniques.
Involved in Analysis, Design and Implementation/translation of Business User requirements.
Experienced in using supervised, unsupervised and regression techniques in building models.
Performed Market Basket Analysis to identify the groups of assets moving together and recommended the client their risks
Experience in determine trends and significant data relationships using advanced Statistical Methods.

Implemented techniques like forward selection, backward elimination and step wise approach for selection of most significant independent variables.

Data Scientist

Performed Feature selection and Feature extraction dimensionality reduction methods to figure out significant variables.

Used RMSE score, Confusion matrix, ROC, Cross validation and A/B testing to evaluate model performance in both simulated environment and real world.

Performed Exploratory Data Analysis using Python. Also involved in generating various graphs and charts for analyzing the data using Python Libraries.
Involved in the execution of multiple business plans and projects Ensures business needs are being met Interpret data to identify trends to go across future data sets.

Developed interactive dashboards, created various AdHoc reports for users in Tableau by connecting various data sources.

OCR Project: Using OCR to archive client-related paperwork, like on-boarding material, to easily create a client repository and significantly reduces the on-boarding time and thereby improves the user experience. Banks use OCR to extract information like account number, amount, cheque number from cheques for faster processing.

Freelance, India
Data Scientist Jan 2016 – Jan 2020

1. Project: Chabot – NLP (Python TF/IDF, NLTK)
Objective:
Worked on Natural language processing and build a Q&A/Chabot service for banking client. This was basically a Q&A system where user can ask questions and service will respond based on trained data. Technique applied - Used TF/IDF approach (NLTK, Pandas) and deep learning algorithm Word2Vec for creating word vectors and to get sense of question data & respond with correct answer.

2. Project: Twitter Sentiment Analysis on manufacturing domain - NLP (Python NLTK)
Objective:
Conducted Sentiment analysis for twitter data to monitor and analyse social phenomena on Goods and Services Tax - GST. Sentiment Analysis is the process of 'computationally' determining whether the GST implementation is positive, negative or neutral. It's also known as opinion mining, deriving the opinion or attitude of the public. So that we can know that the public opinion on GST implementation. In our Study we identified that 75% people are positive about GST. We have reported this classification with bar charts.

Responsibilities:

- Conducting sentiment analysis with NLTK package.
- Data analysis (end to end) – data import, data preparation, Identification of model data, model building, model evaluation, model deployment and presentation of results.
- Quality check (QC) of all the intermediate and final outputs. Presentation of results, helping project owner in interpreting outputs and Clarifying Clients' technical queries through mails and calls

3. Project: The purpose of this project was to predict the sales of retail store.

Responsibilities:
• Creation of Hypothesis using Mind-maps to understand the problem in minute details.
• Visualize the data through pie charts and histogram and did Uni-variate and Bi-variate Analysis.
• Worked on missing value and feature Engineering.
• Model building with Linear Regression, Random Forest and XGBoost.

Data Scientist

4. Project: Prediction Loan Repayment

Objective: Bank offers a wide variety of credit cards for customers, and it evaluates every new profile before approving the card to check if the customer is a potential defaulter or not. Using historic data available, predictive analytic models are created to predict the future behaviour of a new customer.

Responsibilities: Developed the predictive model for the customers that are most likely to churn to predict weather to issue a credit card to the new customer or not.

Tad Solutions Inc, USA
System Analyst Jan 2014 - Dec 2016
Oracle EBS /ADF /SOA – Techno Functional consultant Nov 2013 - Dec 2015.

Tata Consultancy Services, USA
Oracle EBS – Techno Functional consultant April 2010 - Nov 2013.
Payment Systems and Credit Card Systems.

DXC Technology, India
Mainframe Technology – Solution Architect Sep- 2005 – Apr 2010

ORACLE FINANCIAL SERVICES SOFTWARE LIMITED, India
Mainframe Developer – Jan 2004 – Aug -2005

United Systems Inc, India
Systems Engineer – Jun 1999 – Dec 2003

Professional Certifications:

- Neural Networks and Deep Learning - Coursera
- Elements of AI – University of Helsinki
- Applied Data Science with Python - IBM
- Data Analysis with Python – IBM
- Data Science Methodology – IBM
- Data Visualization with Python -IBM