Deepak Kumar Singh

Data Science | Machine Learning (India)

www.linkedin.com/in/dks916 , ds9207260@gmail.com +91-9167622916

OBJECTIVE

- Seeking an employment opportunity that challenges and enhances my technical abilities as a Data Scientist, Machine Learning engineer.

PROFESSIONAL SUMMARY

- **Over 3+ years of IT** experience in Analysis, Design, Development, Implementation with python, machine learning and in data science profile.
- Experience in developing applications using technologies: **python**, **Machine learning algorithms.**
- Worked in developing application with Machine learning cases like **classification, regression, clustering.**
- Worked on **Supervised and Unsupervised problems.**
- Hands on Visualization, Missing value imputation, Encoding, Feature scaling, PCA, Regularization, Feature selection, Outlier detection, Feature transform, Cross validation.
- Worked on **Linear Regression, Logistic Regression, SVM, Naïve Bayes, Decision Tree, KNN, Random Forest, K-Means**, Data Analysis, Feature Engineering, Hyper Parameter Optimization, Statistical and Predictive Modelling, **Artificial Neural Network, Convolutional Neural Network, Recurrent Neural Network**.
- Worked on Python, Sklearn, Numpy, Pandas, Matplotlib, Seaborn, Stats, Pickle, Tensorflow-Keras, SQL queries.
- Hands on PySpark, Data Bricks | Azure-ML, **AWS – EC2** | Spyder, Jupyter Notebook, Google Colab.
- **Conference paper submission** in Fig Congress Warsaw, Poland for MTech thesis.
- Have knowledge on Python data types, Use Case, OOPs.
- Exposure of application development in different environments: **Windows** and **Ubuntu**.
- Experience in IDE tools: **Jupyter notebook, PyCharm, Spyder, Python IDLE, Edit Plus** and **Google colab**.
- Team player with ability to work under minimal supervision possessing excellent communication, interpersonal, analytical and problem-solving skills.
- Flexible and versatile to adapt to any new environment and work on any project and able to meet deadlines.

ACCOMPLISHEMENTS

Project #1	**Insurance Verification | Random Forest & NLP (Vassar Labs)**	**Duration**	**Jan 2022 to July 2022**
Description	To provide verification report on crop yield by farmers over a period of time against provided insurance from state government of Haryana. Worked on data acquisition and text data cleaning to get farmers details from Haryana govt site. Performed various text cleaning task such as text cleaning, HTML tag removal, handling punctuation etc. **Dataset: -** Data acquisition task is done through the provided access to Haryana State Govt. website.		
Responsibilities	➢ Client interaction and requirement analysis. ➢ Designed and Implemented Random Forest model along with tuning. ➢ Implemented on python and google earth engine platform. ➢ Performed data pre-processing. ➢ Involved in presentation and documentation of the project POC. ➢ Given technical knowledge transfer sessions to team on Google earth engine and Sentinel data. ➢ Conducting Code reviews.		
Client	**Reliance, Bajaj, Haryana state government.**		
Role	**Team member**		
Environment	**Python.**		

Project #2	Detection and classification | Random Forest (Vassar Labs)	Duration	Jan 2022 to July 2022
Description	Design and analyzed different machine learning algorithms for detection of water pixel and classification of vegetation from buildup area, similarly water detection needs to be implemented after classification task for better understanding of water stability over area of Telangana state. **Dataset: -** More than 2 plus year of Sentinel 1 VV-VH data is collected from Google earth engine platform. **Modelling: -** Random Forest model is selected for training purpose which result in accuracy of more 90%. Implementation result able to show the correct availability of water and healthy vegetation which ultimately helps in Telangana Govt to launch future scheme.		
Responsibilities	➢ Client interaction and requirement analysis. ➢ Designed and Implemented Random Forest model along with tuning. ➢ Implemented on python and google earth engine platform. ➢ Performed data pre-processing. ➢ Involved in presentation and documentation of the project POC. ➢ Given technical knowledge transfer sessions to team on Google earth engine and Sentinel data. ➢ Conducting Code reviews.		
Client	**Telangana Govt**		
Role	**Team member**		
Environment	**Python, Google earth engine, Random Forest, QGIS.**		

PROJECTS

Projects #3	Customer Churn Prediction
Business Problem	This project aims towards: - ➢ Understanding the key indicators/features responsible for customer churn. ➢ Estimating the likelihood of an active customer leaving an organization. ➢ Implementing retentive strategies to diminish prospective customer churn. Dataset: - Telecom customer churn dataset (Kaggle). Workflow: - Data analysis, Feature engineering, Model selection, Model tuning, Predictions, Feature importance.
Algorithm	Logistic Regression, KNN, Decision Tree, Random Forest.
Environment	Pandas, NumPy, Matplotlib, Seaborn, Jupyter.

Projects #4	Fraud Detection
Business Problem	Fenix (financial service provider) facing a huge revenue and profitability crisis due to significantly large number of unauthorized transactions being made through their credit/debit cards. This project aims towards: - ➢ Building proactive monitoring and fraud prevention mechanisms in place to reduce the unauthorized transactions. ➢ Develop a machine learning model to detect fraudulent transactions based on the historical transactional data of customers with a pool of merchants. Dataset: - Fenix fraudulent transactions dataset (Kaggle). Workflow: - Data analysis, Feature engineering, Model selection, Model tuning, Predictions, Feature importance.
Algorithm	KNN Classifier, Random Forest Classifier.
Environment	Pandas, NumPy, Matplotlib, Seaborn, Jupyter.

Project #5	Recommendation System	Duration	July 2021 to Aug 2021
Description	Build Movie Recommender System using TMDB dataset, preprocessing, stop word, stemming and deployment as web application On Heroku.		
Responsibilities	➢ Designed and Implemented NLP. ➢ Implemented on python and Google colab. ➢ Performed data pre-processing. ➢ Involved in presentation and documentation of the project.		
Environment	**Python, Kaggle, Windows.**		

Project #6	MTech Research and Conference Paper (IIT Kanpur)	Duration	Aug 2021 to Aug 2022
Description	**Title: -** Zero Velocity Detection Using Foot Mounted IMU. Determination of zero velocity instance using IMU sensor during human localization task. We generated dataset and also used **pyshoe dataset**. Implemented ML/DL techniques to build a zero-velocity detector.		
Responsibilities	➢ Involved in Literature review, understanding system requirements for data generation. ➢ Explored Neural Network for processing pyshoe dataset. ➢ Designed and Implemented ML/DL algorithms.		
Organization	**IIT Kanpur**		
Paper Link	*Google scholar*		
Role	**Research team member**		
Environment	**Python, Google colab, windows**		

PROFESSIONAL EXPERIENCE

Duration	Designation	Organization
Jan 2022 to July 2022	Data Scientist	Vassar Labs IT Solution PVT LTD, India
July 2021 to Aug 2021	Machine Learning Intern	GRIP, India
Jun 2021 to July 2021	Natural Language Processing Internship	TCS iON, India
Mar 2016 to Jan 2019	Software Engineer	Techwob IT services & solution, India

EDUCATION

Course	Organization	Percentage/CGPA
MTech **(**Geoinformatics**)-2022**	IIT Kanpur	7.12

CERTIFICATIONS/COURSE

- IIT KANPUR Machine Processing of Remotely Sensed Data (Machine Learning).
- NASSCOM Certified Full Stack Data Science & AI (Naresh IT).
- NASSCOM Certified Python Developer (Naresh IT).