Sanjay Yadav

Data Science/Data Analyst

I am Submitting herewith my resume for your perusal and consideration in your organization. to describe myself in nutshell, I am systematic, organized & hard working, ready to take up any challenge of life & a Team Player with Good communication skills. I am confident of making visible contribution for the growth of the organization. I would appreciate the chance to meet with you in person to discuss as to how I could be a vital part of your organization.

✉ sanjaykumaryadavcool@gmail.com

📍 New Delhi, India

🐦 @uniqueDisability-UK0620619880002991

📱 9041176459

in https://www.linkedin.com/in/sanjay-yadav-

📄 www.kaggle.com/sanjay_19/notebooks

ML ALGORITHMS

Linear Regression, Logistic Regression, SVM, Decision Tree Random Forest, Ada Boosting KNNGBMXG Boosting, Naïve Bayes, K-Means Clustering, Hierarchical Clustering, Principal Component Analysis

TIME SERIES FORECASTING

Exponential Smoothing Moving Average ARIMA Facebook Prophet

NATURAL LANGUAGE PROCESSING

Named Entity Recognition Text Classification Sentiment Analysis Pattern Matching

SKILLS

Python, OOPs, Flask, Scikit learn, NumPy, Pandas, Seaborn, Statistical models, Matplotlib, MySQL, HTML, CSS, Advance Excel, Power BI, Scikit Learn, Model Technique, Data Analysis, Predictive Analysis, Web Scaping, Data Cleaning, Data Visualization, NLP, Mongo DB, MS Office, Power query.

LIBRARY & TOOLS KNOWLEDGE

Anaconda, Jupyter notebook, Power query, Advance Excel, Spyder, PyCharm.

PROFESSIONAL SYNOPSIS

- Overall 3+ years of experience as an IT Trainer and their all operation also work data analytics in the areas of Risk Modelling, Predictive Modelling, Forecasting, Customer analytics and Marketing and Campaign analytics, able to tech technical skill to visually impaired person *etc.*
- 9 months of experience as Data Science practitioner with expertise in data exploration (EDA), predictive model development, model validation and model deployment.
- Good Knowledge of EDA and visualization using NumPy, pandas, matplotlib and seaborn in python
- Implemented Machine Learning algorithms using scikit-learn, XG Boost and states models in python.
- Good Knowledge on NLP – Sentiment Analysis, Entity Recognition using Bag-of-words, TF-IDF and Spacy
- Good knowledge of advance MS Excel and Power query.

TRAINING & EMPLOYMENT SUMMARY

Technospecs	Python	Dec- 2017 – may- 2018
Talent Shiksha (weekend Virtual session)	Data Scientist	Mar- 2021 – Nov- 2021
Samarthanam	IT Trainer	May- 2018 – Till Now

ACADEMIC CREDENTIALS

Qualification	Board/University	Year	Percentage
B. Tech (CSE)	Punjab Technical University	2014-2018	70%
Intermediate	Bihar School Examination Board, Patna	2014	62.4%
High School	Bihar School Examination Board, Patna	2011	64%

PROJECT EXPERIENCE

Following projects have been completed as part of the course, have done the complete end-to-end coding in python. As part of the project have done – data cleaning and pre-processing, development of multiple models, hyper- parameter tuning, model selection.

FLIGHT PRICE PREDICTION

This dataset is about predicting the ever-varying prices of tickets. The dataset consists of data collected from various sources and includes the following features.

TITANIC: MACHINE LEARNING FROM DISASTER

Objective: Use machine learning to create a model that predicts which passengers survived the Titanic shipwreck. Have performed the following tasks on the data set – data cleaning and visualization, data transformation and feature engineering, created following models – Logistic Regression, Decision Tree, Random Forest and XG-Boost. Did hyper parameter tuning, and algorithm was selected based on precision, recall and F1 score.

HOUSING PRICE PREDICTION

Developed machine learning model that predicts the housing price. With 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa. Performed feature engineering and advanced regression techniques like random forest and gradient boosting for accurate prediction housing price.

PREDICT THE STOCK PRICE – TIME SERIES FORECASTING

Built time series forecasting models to predict the daily price of the stocks listed in Nifty 50. Following techniques are used Exponential Smoothing, ARIMA, Auto Arima and Facebook Prophet. Performed univariate time series analysis. Used MAPE to compare the model._____Model produced the least forecasting error.

PREDICT LANGUAGE TAG FROM QUESTION TEXT – NLP

Used Spacy and Scikit-Learn to build a model to identify tags from the question text. Dataset had 10% of questions and answers from the Stack Overflow programming Q&A website which had approx. 1.2M text and about 15 tags. Performed text pre-processing like tokenization, removal of stop words, punctuations, lemmatization and created vector representations used spacy matcher to create a rules to generate labelled examples for training. Employed Count Vectorizer (n-gram) and TF-IDF along with advanced classification algorithms like Random Forest, XGBoost for tag identification.
Count Vectorizer (n-gram) and TF-IDF along with advanced classification algorithms like Random Forest, XGBoost for tag identification.

ZOMATO DATA ANALYSIS

Restaurants from all over the world can be found here in Bengaluru. From United States to Japan, Russia to Antarctica, you get all type of cuisines here. Delivery, Dine-out, Pubs, Bars, Drinks, Buffet, Desserts you name it and Bengaluru has it. Bengaluru is best place for foodies. The number of restaurants is increasing day by day. Currently which stands at approximately 12,000 restaurants. With such a high number of restaurants. This industry hasn't been saturated yet. And new restaurants are opening every day. However, it has become difficult for them to compete with already established restaurants

INTEREST

- Learn something everyday
- Travelling
- Social Services
- Reading Book

STRENGTH

- Multitasking
- Teamwork
- Analytical Skill
- Time management

CERTIFICATION

- Data Science
- Python

ADDITIONAL ACTIVITY

- Able to Tech computer visual impaired person

PERSONAL DETAILS

Father Name – Late Ramchandra Rai

Permanent Address – Gopal Ganj, Bihar

Date of Birth – 12/05/1995

Language Known – English & Hindi

Marital Status – Single

Nationality / Religion – Indian / Hindu