NIRMITK TRIPATHII

Gurugram, India
nirmitktripathii@yahoo.com / 91-9528006968
Date of Birth: 10/03/1995

Github Profile: https://github.com/nirmitktripathii
Linkedin Profile: https://www.linkedin.com/in/nirmitk-tripathii-553a92158/

CAREER OBJECTIVE

I have gained good exposure to understand various aspects of **Data Science** through my Masters in Data Modeling and Simulation which has developed a great sense of confidence at a very personal level. Hence, I seek a challenging position in the area of **Data Science, Machine Learning, Data Analysis & related fields** where I can share my skills and expand my capabilities further in the pursuit of progressive career advancement.

CERTIFICATION

1. Masters Program -Data Scientist : Simplilearn Certified in collaboration with IBM - 2023

- Data Science with Python, R
- Advance Machine Learning
- Tableau 10
- Big Data Hadoop & Spark Developer
- SQL Training
- Deep Learning
- Natural Language Processing
- Data Science Capstone

2. Other Certifications

- Completed Football Data Analytics certification course from Mad About Sports
- Completed IBM certification course from Purdue University on Data Science, Tableau. Currently undergoing courses for NLP and Deep Learning.
- Undertaking online course on Hadoop Ecosystem.

EXPERIENCE

AI/ML Programmer / HRBOTICS - Noida, India *04/2022 - 03/2023*

Raising human capitals by SaaS-Talent Discovery & Mgmt, HCM, Resource as a Service, Recruiter as a Service B2B & Adv Sal

- Acquiring data of all the registered commercial outlets in NCR from 2010 to 2021 using data scraping with multi-threading.
- Stored and fetched the acquired data in both SQL (SQLite) and noSQL(MongoDB) databases to and from a Django based backend system.
- Automating the entire workflow of job posting and sharing as well as applicant hiring in Linkedin using headless Selenium an BeautifulSoup, creating an ATS protoype.
- Design reusable libraries and application programming interfaces (APIs) for streamlining and automating ATS.
- Creating React UI/UX components for designing web pages in React-Native as part of web and android based applications.

Football Data Analyst / Boomer 11 - Dubai, United Arab Emirates *11/2021 - 03/2022*

Boomer11 is a digital engagement and sports technology platform created to connect and engage sports driven communities across all GCC nations.

- Regular acquisition of real time statistics from matches. Application of the acquired data into performance analysis of individual players and teams.
- Prediction and analysis of key battles (a defender vs a striker, midfielder vs midfielder and winger vs full back) between the players from both the sides.
- Analysis and prediction of important players to look out for in the next match based on the percentile and raw stats acquired for the players from previous matches.
- Video podcasts discussing the latest happenings in the footballing world and analysing important matches, transfer strategies, and playing tactics.

.

PROJECTS

IMAGE CLASSIFIER USING IMAGENET CNN MODEL:

- Created an Image classification Deep Learning Model using an ImageNet CNN network architecture (VGG16) trained on pictures of different unhealthy plants, leaves and fruits and classify them according to their different diseases.
- Achieved an accuracy of 84% on unknown test data.
- Created a web application using Flask which takes input images and classify the species, health status and disease of the plant, if it classified as not healthy.

TWITTER SENTIMENT ANALYSIS:

- Built a Twitter sentiment analysis tool that classifies a tweet into positive or negative sentiment.
- Achieved an overall accuracy of 87% on unknown test data using TF-IDF token vectorizer.
- Developed a model app on Flask with a frontend which takes a tweet as input, sends it to the Twitter Sentiment model in the backend and displays the result back in the frontend as output.

MODELING SHOT QUALITY USING EXPECTED GOALS METRIC - Pune, India - 6/2021 - 01/2022

- Developed an ultra fast data scraping code to web scrape shots data from Understat.com website that reduced the time for webscraping shots data from the site from 8000+ seconds to 250 seconds.
- Developed a Random Forest (RF) and Logistic Regression (LR) model to predict and classify a shot taken by a player as a goal or no goal depending on the distance and angle made by the shot location from the goal mouth as well as whether the shot was taken
  by strong foot/weak foot/header.
- Got 89.77 % accuracy with RF Classifier model without hyperparameter optimization and 90.55% accuracy with hyperparameter optimization. xGBoost Classifier achieved 90.55% accuracy as well but had a poor f1- score in classification of goals as compared to RF classifier.

EXPLORATORY DATA ANALYSIS, & DEVELOPING A VAEP MODEL FOR PLAYERS & PREDICTIVE MODEL FOR TOP 5 EUROPEAN LEAGUES IN FOOTBALL - Pune, India - 10/2020 - 01/2021

- EDA aimed at studying the football dataset, to analyse, extract information from it and make important conclusions based on the data, primarily on the different styles of plays in different teams, finding weaknesses and strengths of the teams and assess the ways of measurement and improvement of the team performance.
- For predictions, used multiple output regressor with different regressors such as Random forest regressor (RF), Support Vector Regressor (SVR), to predict the points table. Got 67% accuracy in Top 6 prediction and 50% accuracy in Bottom 4 prediction for 2019-20 season.
- Created a GUI wherein we display the goal plots in which all the previous actions leading to a goal are displayed and top 10 players, of each league, according to VAEP (Valuing Actions based on Estimated Probability) rankings are also displayed.

EDUCATION

M.Tech: Modeling And Simulation *2022*
Savitribai Phule Pune University - Pune | CGPA : 7.4/10

B.Tech: Electrical Engineering *2018*

Pandit Deendayal Petroleum University - Gandhinagar, IN

Intermediate: PCM+Computer Science *2013*

Hartmann College (ISC) – Bareilly, IN

High-School: *2011*

Hartmann College (ISC) – Bareilly, IN

TECHNICAL SKILLS

- Python
- R
- SQL
- Selenium
- Scrapy
- Azure
- Postman
- React.js
- React-Native
- Tableau
- Django
- MATLAB
- VMware
- Linux
- Javascript
- Hadoop
- MongoDB
- SQLite3
- Pandas
- Scikit-Learn
- OpenCV
- NLTK
- Machine Learning
- NLP
- Deep Learning
- Data Analysis
- Django
- Backend Development
- ReactJS
- Frontend Development
- Statistics and Probability
- Statistical Mechanics
- Quantum Physics
- Football. data analysis
- Computational Physics
- Numpy
- Sentiment Analysis
- Image Classification

PASSION

- Loves numbers and patterns and likes exploring new and fascinating aspects of mathematics everyday.
- Currently reading Soccermatics by David Sumpter
- Likes reading Stephen Hawking, Richard Dawkins, Erwin Schrodinger, Dan Brown and Jeffrey Archer.
- History buff- European, History and Ancient and Classical Indian History
- Physics- Astrophysics and Quantum Mechanics

ACHIEVEMENTS

- Created models to quantify ratings of football players based on estimated probabilities of their in game actions.
- Created an Expected Goals model that gives probability of a shot taken in a football match being a goal or not.
- Created an automated ATS system for HR recruitment using Selenium and BeautifulSoup.

PUBLICATIONS

- Modeling Shot Quality using Expected Goals Metric