**Pratibha Rana**

Shimla-171009, Himachal Pradesh, India |pratibharana08@gmail.com | +91 8219637232 | |<https://www.linkedin.com/in/pratibha-rana/>

**EDUCATION**

**University of Huddersfield** **England, United Kingdom**

Master of Science (MS) Data Analyst September 2021

**Panjab University Punjab, India**

Bachelor of Science (B.Sc.) Information technology May 2018

Introduction to Analytic Computing | **COURSEWORK**

Data Statistics| Big Data Analytics| Data Mining, Data Visualization | Data Warehousing | Database Design and Database Development | Mathematical Foundations of Analytics | Structures and Algorithms Business | Effective research in analytics| Case study in Artificial intelligence| Machine learning

**TECHNICAL SKILLS**

**Programming Languages:** R, Python, MySQL, NoSQL

**Tool/Software:** Tableau, SPSS, Microsoft Office, Jupiter notebook, Power Bi, Apache, Apache Spark

**Analytical Skills:** Hypothesis testing, Data collection, Data Modeling, Data Manipulation, Data Cleaning, Data Processing, Data infrastructure, Machine Learning, Statistical analysis (statistics), API, Predictive Analysis, Regression Analysis, Quantitative Analysis, TensorFlow, Keras

**EXPERIENCE**

**HOYER PETROLOG UK LTD Huddersfield, United Kingdom**

**Operations Analyst 21st December- 5th May 2019**

* Analyzed the operations by tracking and reporting documentation to achieve efficient logistic supply.
* Maintained and updated the MS Access database to feed in the detailed information about the operations as well as if any discrepancies in the planning route occur.
* Operated various internal software to analyze the client base and the product requirement.
* Generated MS Excel reports and manipulated information using pivots, slicers, and formulas.
* Reported the documentation for a smooth transition between shifts.
* Managed client relations by communicating the changing information and keeping records of any requirements.

**SMOHAR for a project under Cerner Corp. Himachal Pradesh, India**

**Project Analyst November-December 2018**

* Analyzed and managed the healthcare project run by the state government designed for enhancing healthcare facilities for Children in rural and urban areas.
* Performed an in-depth exploratory analysis to delve into useful findings for the project which can further be helpful in strategy building.
* Cleaned and Processed the available data using Python and SQL.
* Developed data-driven visualization using python (Matplotlib, Seaborn, ggplot), Tableau for senior management to clearly define the success and suggesting further implementations for the pilot project.
* Co-ordinated with the CERNER team to adopt new technologies to make the project more efficient.
* Collaborated with a team of 5 in analyzing the overall operations and ground effects of the project.

**PROJECTS**

**UBER REVIEWS (TEXT AND SENTIMENT ANALYSIS) September 2019 - December 2019**

* Web scraped the Uber reviews, performed text mining classification on that complex dataset, converted “Words” to “Numbers” using NLP Bag of Words model, Trained few supervised learning models (Naive Bayes, Logistic Regression). Evaluated the models using Accuracy and AUC-ROC.
* Deployed the final model using Flask a micro web framework in python. It predicts the sentiment of a customer new text reviews input as “Positive” (1) OR “Negative” (0)

**NYC PARKING DATA September 2018 - December 2018**

* Data is cleaned and further analyzed using libraries like pandas, matplotlib and spark for exploration and finding graphical relationships.
* Conclusions regarding violations and cars parked are made to present further insights for improvements.

**MOVIE RECCOMMENDATION SYSTEM May 2019- June 2018**

* The data is cleaned and processed in two phases working towards creating a recommendation system considering the content and the ratings. The python code uses pandas, NumPy for exploration and representation and sklearn for implementing text mining through cosine similarity scores for popularity and content-based recommendation system.
* The spark file implements the collaborative filtering using PySpark implementing ALS (Alternate Least squares) and regression calculator and further evaluating the root mean square error.