

Rahul Nair

rahulmnair1997@gmail.com • www.linkedin.com/in/rahulmnair007 • rahulmnair7.github.io

EDUCATION

2019 - 2021	ILLINOIS INSTITUTE OF TECHNOLOGY Master's in Science : Data Science GPA:- 3.55	Chicago, IL
2015 - 2019	UNIVERSITY OF PETROLEUM AND ENERGY STUDIES Bachelor of Technology: Computer Science GPA:- 3.47	Dehradun, India

SKILLS

Languages: Python , R, Scala, Java, SQL, C, C++, Pyspark
Databases: MapReduce, MongoDB, Pig, Hive, Cassandra, MySQL, PostgreSQL, Neo4j, BigQuery
Tools: Google Analytics, FLASK, Microsoft Excel (Advanced), AWS, Docker, TensorFlow, GitHub, Tableau, DJANGO, Power BI, SAS, Alteryx, Microsoft PowerPoint, Microsoft Suite, React, Looker
Soft Skills: Problem solving, analytical, team building, team player, leadership, communication, multitasking, organizational

WORK EXPERIENCE

2021 – Present	LABELMASTER <i>Data Scientist (Logistics & Supply Chain)</i> <ul style="list-style-type: none">Perform an Exploratory Data analysis on LabelMaster's production lines, especially Books and Packaging.Build a Sales Forecasting model for these departments with an accuracy greater than 85% using VAR and LSTM.Integrate the model with an interactive UI using Flask.	Chicago, IL
2020 - 2020	CHICAGO JUSTICE PROJECT <i>Data Analyst</i> <ul style="list-style-type: none">Performed ad-hoc analyses on Chicago Crime Data provided by Chicago Justice Project from 1996 – 2020.Structured the data and performed an exhaustive data analysis leveraging python to find out certain patterns.Built a dashboard using Tableau for effective visualization for stakeholders.	Chicago, IL
2018 - 2019	EPIC MINDS IT PVT. LTD. <i>Machine Learning Engineer</i> <ul style="list-style-type: none">Built an end-to-end image classification model using python to predict the disease which the plant has with an accuracy of 86%.Utilized Google Inception v3 model for training along with DJANGO and FLASK to integrate the model with an interface.	Bengaluru, India

PROJECTS

2020 – 2020	CLUSTERING AND REGRESSION ANALYSIS OF GERRYMANDERING <ul style="list-style-type: none">Performed an exploratory data analysis on 2012 and 2018 plans to find out the issues in 2012 redistricting plan of PA.Implemented weighted k-means to develop a new redistricting plan for PA making sure the population distribution remains proportionate across districts. This new plan has improved the fairness by about 60%.Experimented with certain regression analyses such as Best Subset, Ridge and Lasso to find which factors influence elections the most.	Chicago, IL
2020 - 2020	YELP RECOMMENDER SYSTEM FOR RESTAURANTS <ul style="list-style-type: none">Developed an end-to-end recommender system leveraging python to suggest restaurants to users utilizing hybrid matrix factorization method with an accuracy of 97%.Analyzed other algorithms such as content-based, collaborative, Approximate nearest neighbor, etc. and tuned hyperparameters using Bayesian optimization.Deployed final model as a web-app employing Angular JS and Flask.	Chicago, IL
2019 - 2019	STACKOVERFLOW DATA ANALYSIS <ul style="list-style-type: none">Ran queries on Hive and Pig on dataset to perform some exhaustive data analysis.Created a Tag Predictor operating Pyspark to anticipate tags for any StackOverflow post with an accuracy of 83%.Extracted Stackoverflow Data operating Bigquery using SQL queries.	Chicago, IL