**Rahul Nair**

 rahulmnair1997@gmail.com • [www.linkedin.com/in/rahulmnair007](http://www.linkedin.com/in/rahulmnair007)• [rahulmnair7.github.io](https://rahulmnair7.github.io/)

**Education**

Aug 2019 – May 2021 **Illinois Institute of Technology** Chicago, IL

Master's in Science : Data Science

GPA:- 3.66

Aug 2015 – May 2019 **University of Petroleum and Energy Studies** Dehradun, India

Bachelor of Technology: Computer Science

GPA:- 3.47

**Work Experience**

Jan 2021 – May 2021 **LABELMASTER**  Chicago, IL

***Data Analyst (Logistics & Supply Chain)***

* Performed an Exploratory Data analysis on LabelMaster’s Books and Packaging production lines.
* Built a Sales Forecasting model for these departments with an accuracy greater than 85% using SARIMA.
* Integrated the model with an interactive UI using Flask, HTML and CSS.

Jan 2020 – Oct 2020 **Chicago Justice project** Chicago, IL

***Data Analyst***

* 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.

Jun 2018 – Jun 2019 **Epic Minds IT Pvt. Ltd.** Bengaluru, India

***Machine Learning Engineer***

* Built an end-to-end image classification model using python to predict the disease which the plant has with an accuracy of 86%.
* Trained Google Inception v3 model for classification and used DJANGO and FLASK for web interface.

**Projects**

Sep 2020 - Dec 2020 **Clustering and regression analysis of gerrymandering** Chicago, IL

* Performed an exploratory **data analysis** on 2012 and 2018 plans to find out the issues in 2012 redistricting plan.
* Implemented **weighted k-means** to develop a new redistricting plan for Pennsylvania 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.

Jan 2020 – May 2020 **Yelp Recommender System for Restaurants** Chicago, IL

* 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**.

Sep 2020 – Dec 2020 **StackOverflow Data Analysis** Chicago, IL

* 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.

**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