**AKHIL THAKUR**
(720)-736-1161 | singhakhil33@gmail.com | https://www.github.com/thakurakhil

---

## KEY QUALIFICATION AND SKILLS

- **Languages** : Python, R, C, C++, Java, PHP, HTML, CSS, JavaScript, jQuery, SQL.
- **Frameworks and Tools** : Git, Scikit-Learn, Airflow, Hive, NLTK, AWS, TensorFlow, PyTorch, Keras, Hadoop, Spark, Kafka, Tableau, Docker, Kubernetes, Prometheus, Elastic Search, Causal Impact.
- **Skills :** Data extracting and cleaning, Experiment Design, Statistical Analysis, Optimization, Simulation, Forecasting, Feature engineering, A/B Testing, Defining Metrics.

## PROFESSIONAL EXPERIENCE

**VISIRIS** APR 2019 - MAY 2021

Data Scientist

- Performed Exploratory Data Analysis on observational data to estimate the impact of various campaign spendings on user's acquisition.
- Implemented a Time Series Forecasting Model using LSTM to forecast peak sales, reducing the RMSE by 12% which led to efficient campaign spending by the marketing team.
- Used XGBOOST to identify key levers for user engagement; Visualized the metrics and communicated with cross-functional teams to drive product decisions, leading to high user engagement.
- Automated the ETL process and optimized the data architecture for A/B testing metrics generation with ~50% reduction in the processing time.
- Created multiple dashboards and visualizations in Tableau to monitor the metrics and key performance indices.

**eBAY** JUN 2018 - DEC 2018

Software Engineering Intern

- Generated embeddings for the categorical attributes of product listings using the skip-gram model trained on the clickstream data of users.
- Improved the AUC by 3% for recommendation of "Similar Items" section using K-Nearest Neighbours on the generated embeddings.

**IIIT HYDERABAD**

Research Assistant DEC 2017 - MAY 2019

- Modeled and built a framework for congestion control using convex utility based optimization.
- Simulated the utilities in a network using TCP kernel modules and RYU network controller. The model resulted in 135% increase in bandwidth utilization.
- Published the work in IEEE International Conference on Advanced Networks and Telecommunications Systems. Title : "Utility Based Framework For Reactive And Proactive Congestion Control In SDN"

## EDUCATION

Masters in Data Science, University of Colorado Boulder AUG 2021 - Present
Masters By Research in CSE, IIIT Hyderabad JUL 2017 - JUN 2021
BTech in CSE, IIIT Hyderabad AUG 2013 - JUN 2017

## RELEVANT PROJECTS

Web Traffic Time Series Forecasting (Kaggle)

- Built a model based on LSTM to forecast future traffic for Wikipedia pages

Chemical Named Entity Recognition

- Implemented a NER model to extract chemical information and chemical-related entities from a list of bio and nano-informatics targets, incorporating POS tagging and molecular formula parsing. Applied various ML algorithms for comparative analysis using recall, precision and f1 scores as metrics.

Book recommendation engine

- Built a book recommendation engine using features such as POS tags, sentiment score, punctuation profile and TF-IDF scores (using Amazon EMR).