

Yashwant Singh

Data Scientist

 [homepage](#)  [yashwant-63](#)  [ysingh63](#)  ysingh64@gatech.edu  [+1 4708151406](tel:+14708151406)

EDUCATION

GEORGIA INSTITUTE OF TECHNOLOGY

MS IN ANALYTICS (MSA)
Aug 2021 - Dec 2022 | Atlanta, Georgia
Cum. GPA: 3.9 / 4.0

IIT KANPUR

M.TECH IN MATERIALS SCIENCE
AND ENGINEERING
July 2013 - June 2018 | Kanpur, U.P
Cum. GPA: 9.1 / 10.0

IIT KANPUR

B.TECH IN MATERIALS SCIENCE
AND ENGINEERING
July 2013 - June 2018 | Kanpur, U.P
Cum. GPA: 7.8 / 10.0

SKILLS

PROGRAMMING

- Python: Pandas, Numpy, SciPy, Scikit-learn, Matplotlib, Streamlit, Stellargraph, UMAP, HDBSCAN, tsfresh
- SQL, Git, Github, R, Spark, Hadoop, Looker, Tableau, D3

ANALYTICAL METHODS

Regression, Principal component analysis, Cross-Validation, K-means clustering, K-nearest neighbor, Time-series analysis, XGBoost

COURSEWORK

GRADUATE

- Machine Learning
- Web Search & Text Mining
- Data Mining & Stat. Learning
- Regression Analysis
- Data & Visual Analytics
- Analytical Modelling
- Data Analytics for Business
- Principles of Management
- Computing for Data Analytics
- Constraint Programming

EXPERIENCE

EQUIFAX | DATA SCIENCE INTERN

May 2022 - Aug 2022 | Atlanta, Georgia

- Set up an automated RCA process to identify top features causing credit score changes on Google Cloud Platform (GCP) to help preserve accuracy of data/credit scores for over 300 million people
- Used logistic regression and XGBoost to identify the most important attributes for over 600 credit models which contribute to \$3.5Billion of Equifax's revenue
- Achieved KS score, Gini coefficient of 0.93, 0.92 for the best XGBoost model and 0.87, 0.84 for the best logistic regression model respectively

SHARECHAT | ASSOCIATE PRODUCT MANAGER/ ANALYST

Sept 2019 - June 2021 | Bangalore, India

- Built a full fledged grievance redressal and ticketing system for ShareChat keeping an average resolution time to less than 8 hours
- Led a team of 5 to migrate CleverTap (a tool used for analytics and notifications) to Looker: Built Looker models from scratch while retaining all CleverTap functionalities, thereby reducing cost by 90 %, translated to savings worth \$ 27000/month

MYNTRA | DATA ANALYST

July 2018 - Aug 2019 | Bangalore, India

- Used Python to build a tool to improve line balancing and produce proactive alerts to reduce leakage in customer order consolidation, resulting in 2% decrease in leakage and annual savings worth \$250,000
- Assisted respective warehouse teams with implementing changes in the order-fulfillment process, resulting in shrinking order to pack (O2P) time from 4.0 hours to 3.0 hours

PROJECTS

PERSONALISED ITEMS RECOMMENDATION | [LINK- COURSE PROJECT](#)

- Implemented Continous Time Dynamic Network Embeddings (CTDNE) and Joint Dynamic User-Item Embeddings (JODIE) to recommend items based on 7 million reviews on Amazon
- Achieved a mean reciprocal rank (MRR) of 0.29 for CTDNE and 0.92 for JODIE

FANTASY PREMIER LEAGUE POINT PREDICTION | [LINK- HOBBY PROJECT](#)

- Scraped player level data from fbref.com/ for the English Premier League 2021-22 season using Beautiful Soup
- Built an ML based app on streamlit allowing users to select players and predict points using linear regression and XGBoost models, and compare players on various metrics
- Achieved accuracy of 50% for linear regression model and 59% for XGBoost

FULL BAND CAPTURE UTILISATION | [LINK- COX COMMUNICATIONS HACKATHON](#)

- Extracted relevant features using tsfresh, further reduced dimensions using UMAP, used HDBSCAN to cluster the FBC data to identify common impairment patterns and evaluated models using DBCV score to find the optimum model
- Achieved DBCV score of 0.78 for HDBSCAN as compared to 0.08 for Kmeans
- Awarded 3rd position among 18 teams