HARSHIT GAUR

gaurharshit04@gmail.com | +16173960129 | **Portfolio**: harshitgaur6155.github.io/
linkedin.com/in/harshitgaur-/ | public.tableau.com/app/profile/harshitgaur | github.com/harshitgaur6155

WORK EXPERIENCE

Greatpoint HR Inc. Boston, MA
Lead Data Analyst *May 2023 - Present*
- Architected and managed high-volume quality **ETL** data pipeline for multi-source ingestion, transformation, processing over 2.3 billion records with **Snowflake SQL**, and **Snowpark Python**, facilitating **data analysis** and **machine learning model** development.
- Analyzed compensation data to identify trends and anomalies, delivering actionable insights that guided executive leadership in strategic planning for salary budgeting, achieving 15% improved employee productivity and strategic business outcomes.
- Developed interactive **Tableau** reporting dashboards using job postings data to analyze compensation trends and benchmarks, effectively communicating insights and empowering clients to optimize job pricing strategies and metrics performance.
- Implemented problem-solving machine learning model with **Large Language Models (LLMs)** as **BERT** and **Sentence Transformers** to automate job category assignments, achieving 88% accuracy, enhancing HR analytical precision via complex analyses.
- Led and collaborated with cross-functional teams to develop **Recommendation System** that analyzes job posting, market trends and geographical factors to deliver decision-making salary pricing, with 20% improved key performance indicators (KPIs) accuracy.
- Communicated and delivered data-driven insights using salary recommendations, enriching clients' understanding of market dynamics and contextual factors, leading to more informed compensation strategies and an 18% increase in client satisfaction.

Mobikasa Delhi, India
Data Scientist (Software Engineer) *December 2018 - August 2021*
- Utilized **Python, Azure Cognitive Service** to implement **unsupervised** and **deep learning** techniques on 4TB unstructured data, achieving 87% accuracy in Speech-to-Text Translation of complex surgical debrief and timeout audio sessions.
- Designed and created databases using **Structured Query Language (SQL)** and utilized **Python, PHP, Shopify** to develop applications: ***1-800-Flowers.com*, AnytownUSA, DigiNurse (BSSNY Hospital)**, driving 40% growth in annual business revenue.
- Forecasted a 30% increase in sales and enhanced customer engagement through insightful data analysis using **Tableau** and **machine learning libraries (Pandas, Seaborn, Scikit-Learn)**, improving website traction through focused marketing strategies.
- Led a team of 4, enhancing technical collaboration and teamwork, driving project success to deliver data-driven solutions that align with strategic business objectives and optimize operational efficiency.

Appster Gurugram, India
Software Engineer *July 2017 - December 2018*
- Created REST APIs in Python, PHP, and Java for applications, utilizing MySQL and maintaining version control with GIT.
- Analyzed user stories via **JIRA** and **Excel**, optimizing project roadmap, reducing deadlines by 7% with data-informed planning.
- Developed mobile applications (MyHero, Bazar, COEY) catering clients' business requirements with 350,000 plus downloads.
- Incorporated cryptography for 2-way security in applications. Integrated payment gateways to streamline online transactions.

PROJECTS

Airbnb Price Estimator (Recommendation System)
- Developed **Predictive Model** on Streamlit with **Random Forest, XGBoost, PCA**, reaching 80% accurate dynamic price estimation.

EDUCATION

Master of Science in Computer Science, *Machine Intelligence and Data Analytics* *September 2021 - April 2023*
Northeastern University, Boston **GPA: 4.0**

Bachelor of Technology, *Information Technology* *August 2013 - July 2017*
Guru Gobind Singh Indraprastha University **GPA: 3.71**

SKILLS

- **Programming Languages & Databases**: Python, SQL, Snowflake, R, PHP, MySQL, NoSQL, PostgreSQL, MongoDB
- **Data Science and Machine Learning**: Pandas, NumPy, Scikit-learn, PyTorch, Matplotlib, Regression, Classification, XGBoost, Clustering, PCA, Deep Learning (CNN, RNN, Transformers), Natural Language Processing (NLP), Large Language Models (LLMs)
- **Data Visualization & Business Intelligence Tools**: Tableau, Microsoft Power BI, Microsoft Excel
- **Data Engineering & Statistical Analysis**: Data Pipeline, ETL (Extract, Transform, Load), Exploratory Data Analysis (EDA), Root Cause Analysis (RCA), Causal Inference, Feature Engineering, Hypothesis Testing, A/B Testing
- **Cloud Technologies and Services**: Microsoft Azure, Amazon Web Services (AWS), GCP, Spark, Hadoop, GIT, Agile, JIRA, Docker
- **Other Skills**: Problem-solving abilities, Attention to detail, Presentation and Communication skills for engaging stakeholders