AAKASH KUMAR HARIT

Delhi, India

+91 9310236672 aakashharit17@gmail.com LinkedIn-Profile GitHub-Portfolio

TECHNICAL SKILLS

Languages: Python, SQL (MySQL Server, PostgreSQL), R, C++
Tools: MS Excel, PowerBI, VS Code, Jupyter notebook, Google Data Studio, SQL Server Management Studio
Python Libraries: NumPy, Pandas, PySpark, Matplotlib, Seaborn, Scikit-learn, Pytorch, TensorFlow
Data Analysis Skills: Data Analysis, Data Mining, Data Visualization, Predictive Analytics, Time-Series forecasting
Machine Learning: Goggle Cloud AI, Azure, Data Bricks, AWS
Non-technical skills: Data Storytelling, Analytical, Leadership, Time management, Research paper writing

PROFESSIONAL EXPERIENCE

AnaAr Solutions *Remote*
Data Analyst *January 2024 – May 2025*

- Utilized PivotTables, VLOOKUPs, and conditional formatting to efficiently analyze large datasets and extract key insights.
- Automated repetitive tasks using Excel macros, improving efficiency and reducing manual effort.
- Performed data extraction, cleaning, and transformation with SQL, ensuring accuracy and consistency across large datasets.
- Enhanced query performance by optimizing complex joins and aggregations, reducing data retrieval time.
- Built interactive Power BI dashboards with dynamic filters and DAX calculations, providing real-time business insights.
- Created visual reports with Power BI and Excel, enabling clear and effective data-driven decision-making.

EDUCATION

University of Exeter *Sep 2022 - Jan 2024*
Masters in Data Science, Merit *Exeter, UK*

Guru Gobind Singh Indraprastha University *Aug 2018 - Aug 2022*
Bachelors in Computer Science and engineering, 8.0/10 *Delhi, India*

PERSONAL PROJECTS

Top UK Youtubers 2024 Dashboard using Excel, SQL and PowerBI — *Github*

- Designed and implemented an interactive dashboard to identify the top-performing UK YouTubers based on subscriber count, videos uploaded, and total views, enabling data-driven marketing decisions.
- Analyzed reach, engagement, and potential revenue to recommend YouTube channels best suited for different campaign types (e.g., product placement, influencer marketing), optimizing ROI
- Data Quality Assurance: Conducted thorough data quality checks, including row/column validation, data type verification, and duplicate removal, ensuring the accuracy and reliability of insights.
- Strategic Marketing Insights: Delivered actionable recommendations backed by data-driven justifications, helping the marketing team prioritize high-performing YouTube channels for future collaborations.
- Created detailed documentation covering data sources, transformation processes, and analysis conclusions, ensuring the solution's reproducibility and maintainability for future updates.

Power Consumption Forecasting using LSTM — *Github*

- Built a time series forecasting model using LSTM neural networks to predict household energy consumption from historical smart meter data.
- Processed and cleaned raw .txt data with custom delimiter handling, and engineered datetime features for temporal analysis.
- Performed exploratory data analysis (EDA) and normalization to uncover consumption trends, voltage fluctuations, and sub-metering patterns.
- Designed and trained an LSTM model in Keras/TensorFlow, reshaping time-based sequences and optimizing for low RMSE.
- Forecasted short-term power usage with visualizations comparing actual vs predicted results, supporting smart grid decision-making.

Spotify HIT Prediction, Dissertation project(Machine Learning) — *Github*

- Imported and preprocessed Spotify track data, including various audio features like tempo, danceability, and more
- Performed EDA to understand the distribution and relationships of features within the dataset, including visualization of audio metrics.
- Implemented machine learning models to predict target variables based on the audio features, likely involving splitting the data into training and testing sets.
- Evaluated model performance using metrics such as accuracy, precision, recall, possibly involving cross- validation
- Created visualizations to present the findings and model performance, ensuring the results are actionabl