Parth Kohli

parthkohli2k3@gmail.com | linkedin.com/in/psk28

EDUCATION

VIT Bhopal University Bhopal, India
Bachelor of Technology in Computer Science and Engineering *Aug. 2022 – Ongoing*
- Current CGPA: **7.81** (as of July 2025)

Delhi Public School (CBSE) Bulandshahr, India
All India Senior School Certificate Examination (Class 12th) *May 2022*
- Percentage: **93%**

Delhi Public School (CBSE) Bulandshahr, India
All India Secondary School Examination (Class 10th) *May 2020*
- Percentage: **92%**

EXPERIENCE

Data Science Intern (Python) February 2025 – April 2025
Main Flow Services and Technologies Pvt. Ltd. *Remote Internship*
- Completed a Data Science with Python internship under the Ministry of Corporate Affairs and MSME
- Developed 8 data-driven projects over 2 months, applying statistical analysis, data visualization, and machine learning techniques
- Demonstrated strong dedication, enthusiasm, and willingness to learn, contributing effectively to assigned tasks
- Gained hands-on experience with Python libraries such as NumPy, Pandas, Matplotlib, and Scikit-learn

PROJECTS

Book Recommender | *Python, Scikit-learn, Pandas, NumPy, K-Means Clustering*
- Built a content-based recommender using K-Means clustering on a dataset of 8000+ books
- Clustered books based on genre, author, and metadata to display 6–10 similar titles
- Engineered features and optimized clustering to ensure meaningful recommendations

Nimbus ETL | *Apache Airflow, Astro CLI, Python, REST APIs, PostgreSQL*
- Designed and deployed a scalable ETL pipeline to fetch and store daily weather data from a public API
- Configured DAGs in Airflow to automate data collection for 5+ countries over 7+ days with flexible date and location support
- Transformed raw weather data (temperature, windspeed, etc.) and loaded into a structured PostgreSQL table
- Enabled real-time weather tracking and future expandability for 50+ locations with minimal configuration

Tumour Detection Model | *Inception Model, TensorFlow, Keras, Data Augmentation, Early Stopping*
- Developed a CNN-based tumour detection model using the Inception architecture on a dataset of 20,000+ medical images
- Augmented 15,000+ images to improve model generalization and robustness
- Implemented early stopping to prevent overfitting and optimize training efficiency
- Achieved validation accuracy and test accuracy of 96%, demonstrating strong predictive performance

TECHNICAL SKILLS

Programming Languages: Python C++, SQL, PostgreSQL,HTML/CSS
Frameworks & Libraries:TensorFlow, Keras, Scikit-learn, pandas, NumPy, Matplotlib
Developer Tools & Platforms: Git, Docker, Apache Airflow, Apache, Kafka, AWS, Google Cloud Platform, VS Code, PyCharm, Visual Studio
Other Technologies: REST APIs, WordPress