ATHARVA CHOPADE

Pune, India

📞 +91-9689913435 ✉ atharvachopade90@gmail.com in linkedin.com/in/atharva7060 ⭘ github.com/Atharva07060

SUMMARY

Big Data enthusiast with hands-on experience in real-time data pipelines using Kafka, Spark, Hadoop, Hive, and Airflow. Skilled in data ingestion, processing, and analysis with Python. Strong problem-solver focused on building scalable data solutions that drive insights and business value.

EDUCATION

Government College Of Engineering   **July 2020 – June 2024**
Bachelor of Technology in Information Technology, CGPA: 8.04   *Amravati, Maharashtra*

Late Jyoti Janolkar Vidalaya   **April 2018 – March 2020**
Intermediate, Percentage: 75.50   *Akola, Maharashtra*

School Of Scholars   **March 2018**
Matriculation (CBSE), Percentage: 85.50   *Akola, Maharashtra*

Technical Skills

- **Big Data Ecosystem:** Hadoop, Hive, Apache Spark, Kafka, Airflow

- **Programming Languages:** Python, R, C

- **Databases:** MySQL, MongoDB

- **Python Libraries:** NumPy, Pandas, Seaborn, Matplotlib

- **Data Visualization Tools:** Tableau, Excel, Power BI

- **Soft Skills:** Problem Solving, Teamwork, Communication, Time Management

PROJECTS

Stock Market Data ETL Pipeline | *Big Data*

- Built a real-time data ingestion pipeline using Apache Kafka to stream stock market data from yFinance API.
- Cleaned and transformed incoming data using Python and Pandas before storing it in HDFS.
- Created and queried external Hive tables for efficient analysis of historical stock trends.
- Demonstrated practical knowledge of distributed systems, data lakes, and batch processing with Hadoop ecosystem.

Football Player Market Value Prediction | *Machine Learning*

- Built regression models to estimate market value of football players using attributes like age, club, nationality, and stats.
- Preprocessed the dataset using feature engineering and handled missing values, outliers, and encoding.
- Evaluated models (Linear Regression, Random Forest, XGBoost) using metrics such as RMSE and R^2.
- Enhanced model accuracy using ensemble methods and demonstrated applicability in sports analytics and player scouting.

CERTIFICATIONS

- Introduction to Cloud Computing – IBM Developer Skills Network

- The Complete Oracle SQL – Udemy

- Python (Basic) – HackerRank

CODING PROFILE

- **HackerRank:** hackerrank.com/profile/atharvachopade90