

Rishabh Agarwal

rishabh5530@gmail.com | +919096074621
rishabh's_linkedin

SKILLS

Languages/Frameworks:

- PySpark • SQL
- Python • Databricks

Softwares/Tools:

- AWS Services • Git • Snowflake

SDLC Methodologies:

- Agile • Waterfall

Soft Skills:

- Leadership • Communication
- Problem Solving • Teamwork
- Adaptability • Presentation

EDUCATION

BITS PILANI WILP

M.Tech in Data Science & Engineering

Nov 2020 - Oct 2022

CGPA: 8.49 / 10

PICT, PUNE

B.E. E&TC

Jul 2015 - Jun 2019

CGPA: 8.67 / 10

THE BISHOP'S SCHOOL

2015 | 2013 | Pune, India

(12th Sci.) ISC: 94%

(10th) ICSE: 93.4%

CERTIFICATIONS

- CCA SPARK AND HADOOP DEVELOPER (CCA175) - CLUDERA
- MICROSOFT CERTIFIED: AZURE DATA FUNDAMENTALS
- MICROSOFT CERTIFIED: AZURE FUNDAMENTALS

ACHIEVEMENTS

- **STAR Award: Quick on the Block** at Fractal for making strong contributions to work in team from week 1.
- **Invited to talk** by PICT about Career in Data Engineering with the current students, staff and alumni.
- **Best Team** and **On the Spot awards** for successful on-time releases.
- **Applause award** for actively leading the AnI Kosh'19 activities.
- **2nd/80** at InC'19 (National Level project competition) for BE project

WORK EXPERIENCE

- Strong analytical, data modeling, and organizational skills; able to manage competing demands with large data sets in high-paced environments.
- Passionate about data and solving complex problems efficiently.

MCKINSEY & COMPANY DATA ENGINEER 2 | 1.9 YEARS

November 2022 - Present

• Sustainability Data Marketplace:

- Led development of robust ingestion pipelines for 2 new data sources and the design and implementation of 2 new API endpoints

• Automated Mapping:

- Spearheaded the design and development of a tool for mapping client data to standard taxonomy leveraging LLMs and Databricks - leading to reduced manual efforts and cost, and increasing overall efficiency

• Carbon Markets:

- Migrated 5 major pipelines to Kedro framework - beating deadlines, and achieving robust and scalable codebase for future enhancements
- Modelling various voluntary carbon market scenarios for 30 year horizon

FRACTAL ANALYTICS DATA ENGINEER | 1.3 YEARS

September 2021 - November 2022

• License Compliance - Piracy Analytics:

- Collaborated with the client software team, developing data pipelines to convert pirated to licensed users. Utilized PySpark on AWS EMR for intricate data processing, orchestrated with Airflow.

• License Compliance - Internal Product:

- Enhanced the product by optimizing PySpark code, refining logic based on user feedback, helping users achieve 95+% accuracy in identifying piracy. Maintained product uptime consistently.

TATA CONSULTANCY SERVICES DATA ENGINEER | 2.3 YEARS

June 2019 - September 2021 | Pune, India

• AML Insights - CounterParty Search:

- Developed PySpark scripts for client-specific data transformation pipelines, spanning *6 years of daily data* for a web search portal. Utilized by users for monitoring suspicious transaction activities.

• AML Insights - C&IB (Corporate & Institutional Banking):

- Designed and optimized final Hive tables for Tableau visualization. Transitioned client from static Excel to dynamic Tableau dashboards, lauded for flexibility. System *adopted by 1000s of RMs*.

• Portfolio Command Module - Domestic Correspondent Bank:

- Collaborated in an agile team of 4 to analyze source data and construct PySpark data pipelines for Tableau visualization. Implemented intricate requirements with custom transformations and aggregations on dataframes.

VOLUNTEERING

- **Certified Mentor** for *The Mentoring Club*, helping graduates and experienced folks explore data engineering as a career by providing insights based on my experience.
- **Invited to Judge** projects at *InC'22, InC'23, InC'24* - National Level BE Project Competition @ PICT.
- **Conducted Spark and Hadoop training** for *over 20 freshers at TCS*.
- **Invited to Mentor** at *GSSoC'21* GirlScript Summer of Code - 3 month long open source contribution program
- **Head of Partnerships at TEDxPICT'17** - Got **9 sponsors** for a successful event.
- **Organized virtual team activities** for **45+ members** including onsite folks while WFH.