

MANORANJAN KUMAR

☎ : +91-9321014955
✉ : softmanoranjank@gmail.com
in : in.linkedin.com/in/TheCodeCache
🌐 : equity-updates.com
🔥 : stackoverflow.com/users/6842300/TheCodeCache
🐱 : github.com/TheCodeCache

Date of Birth:	22 nd May 1987
Address:	Room No-09, Sahu House, Kalina, Old CST Road, Santacruz (E), Mumbai-400029
Comm. Skills:	English, Hindi
Location Pref:	Mumbai, Bengaluru, Pune

PROFILE – SUMMARY

- ❖ **11+ Years** of Versatile & Exciting Work Experience
In – Analysis, Design, Development of **Data Pipeline**, Testing, **Troubleshooting**, **Performance**
on – **Big Data** | AWS | **Data Science** | Concurrency | **Python** | **Scala** | Java | Spark | TensorFlow
at –
 - **Impetus Technologies** Module Lead May – 2015 till – Present
 - **Sonata Software Ltd** Sr. System Analyst Apr – 2013 till Apr – 2015
 - **Eka Software Solutions** Software Engineer Sep – 2010 till Mar – 2013

TECH – SKILL

- ❖ **Design**
 - **4+ Years** of work-ex in Software **Design**
 - Hands-on:
 - **Data Structure**, Algorithm, **Design Patterns**, **System Design**
 - Multi-Threading, Cache, I/O, SOLID, UML, Normalization
- ❖ **Full Stack**
 - **3+ Years** of work-ex in Web **Development**
 - Hands-on:
 - JavaScript, **AJAX**, HTML, JSP, JSON, XML, CSV, **Web Scrapping**
 - **SQL**, Oracle, **MySQL**, PL/SQL, Struts, **Spring**, Hibernate, JDBC
- ❖ **Big Data**
 - **5+ Years** of work-ex as a Big **Data Engineer**
 - Hands-on:
 - **Python**, **Scala**, Core **Java**, **SQL**, Shell Script, AWK, Linux, Hadoop
 - **Spark**, **PySpark**, Hive, **YARN**, **S3**, **HDFS**, ORC, **Batch**, **Streaming**
- ❖ **Cloud**
 - **4+ Years** of work-ex on Amazon Web Services
 - Hands-on:
 - **Lambda**, **Step Functions**, CloudWatch, EC2, **EMR**, SQS, SNS, **S3**
 - **Cost Optimization**, Disaster Recovery, Operational Efficiency
- ❖ **Data Science**
 - **2+ Years** of work-ex in **Deep Learning**
 - Hands-on:
 - **TensorFlow**, GPU, **NumPy**, **Pandas**, Jupyter Lab, **CNN Model**
 - Data Preprocessing, Predictive Modeling, Hypothesis Testing
- ❖ **Leadership**
 - **2+ Years** of work-ex in **leading a team** of 3 colleagues
 - Hands-on:
 - Proper Delegation, Teamwork, Resolved Team Conflicts etc.
- ❖ **Tools**
 - **Project Tools** – Jira, **Confluence**, Agile, Scrum, **BitBucket**, Tortoise **Git**
 - **Dev Tools** – **Maven**, **PyCharm**, **Eclipse**, **Putty**, **CI/CD**, Deployment
 - **Clean Code** – Sonar, **Code Coverage**, **Code Quality**, **Best Practices**

ACHIEVEMENT

❖ At Work

- > **30% Improvement** in unified pipeline's total execution time
- > **25% additional improvement** in unified pipeline's run time
- **Open-Source Activities** with OpenCSV, Hue & **Catboost** package
- **Solved** OOM Error, **Tuned** Spark jobs, **Fixed** Performance issue
- **Better Productivity** through Code Quality, **Refactoring**, Utility
- > **35% Cost Saving** on objects stored on S3 – using **S3 Versioning**
- > **60% CPU Utilization** of Node – for a good **50%** of job's runtime
- **Wrote** [Utility](#) for generating SQL Views – used by > **8** teams
- **Automated Tasks on AWS**, e.g., stopping idle EMR, Alerts
- **0 bug** in data-ingestion code, running in production for **3** years
- > **20** Teams uses data-ingestion to process > **1000 feeds** daily
- > **12** Teams uses **Aggregator** – a utility that combine small files
- Analyzed **Heap Memory**, **Thread Contention** using Dump files
- > **80% Memory** Utilization of EMR Cluster – for spark jobs

❖ At Home

- <https://equity-updates.com> – this is my website, secured, basic, built from scratch
- As low as INR 72 /- per month costing to keep this website up and running
- **Tech Stack** Used – Python, Nginx, **Flask**, WinSCP, **Route 53**, Deployment on **EC2**
- Developed **Games** (based on older Nokia mobile phones) using Graphics and **C++**
- Built **Tic-Tac-Toe** game using just **C** language
- **Open-Source** activity with Apache **Spark**, [Set up](#) the source code in local Eclipse IDE
- Qualified **GATE-2021** in CS&IT paper, helped me to clear fundamental concepts
- **Automated** profile update activity on <https://www.naukri.com> for job searching

PROJECT

❖ **Marsh AWS Big Data Lake**®, June 2018 – Present

Environment: Spark, Scala, Python, Java, Hive, Hue, Git, MongoDB, AWS, Jenkins
Description: A big-datalake platform on AWS, supports data-catalog, ingestion

❖ Key Results:

- Developed Data-Ingestion module from scratch
- Played a Key role in Going Live MVP (Sep-2018)

❖ **Stream Analytix (Image Classification)**, June 2017 – May 2018

Environment: Python 3, TensorFlow, NumPy, Pandas, GPU, Dask, Linux (CentOS)
Description: A multi-engine stream processing and machine learning platform

❖ Key Results:

- Wrote **Dask** Scheduler for launching **training** processes
- Developed a **classification training** program for images

EDUCATION

❖ 2010	B.Tech.	CS & IT, IET MJPRU, Bareilly, UP	78.10%
❖ 2004	XIIth	BIEC, Patna	65.22%
❖ 2002	Xth	BSEB, Patna	69.59%

CERTIFICATE

❖ **MapR** certified Spark Developer – 2018 [click here](#)

❖ Successfully qualified:

- **GATE**–2010 in CS&IT with a score of 468 i.e. (96.25%tile)
- **GATE**–2021 in CS&IT with a score of 423 i.e. (AIR – 8317)
- **G – Level of CST – 2010** conducted by **C – DAC, Mumbai**

DECLARATION:-

I would like to hereby inform that all the above information is true to the best of my knowledge.

Date: -

Manoranjan Kumar