

## SUMMARY

New Graduate Software Engineer with a strong foundation in Big Data, specializing in optimizing Spark jobs for large datasets, and aiming to leverage my skills in distributed computing to drive business growth and efficiency.

---

## EXPERIENCE

No direct industry experience, but actively seeking opportunities to apply my skills and knowledge in real-world settings.

---

## PROJECTS

### Big Data Analytics Platform *Personal Project, 2022*

- Designed and developed a scalable data processing pipeline using Apache Spark and Hadoop, achieving a 30% reduction in processing time for large datasets.
- Implemented data visualization tools using Tableau to provide insights into dataset trends and patterns, resulting in a 25% increase in data-driven decision making.
- Collaborated with peers to integrate the platform with various data sources, including MongoDB and Cassandra, and ensured seamless data ingestion and processing.

### Distributed Computing Research *Academic Project, 2021*

- Researched and implemented various distributed computing algorithms using Apache Spark and Hadoop, focusing on optimizing job performance and reducing latency.
  - Developed a comparative analysis of different distributed computing frameworks, including Apache Spark, Hadoop, and Apache Flink, and presented findings in a research paper.
- 

## TECHNICAL SKILLS

TECHNICAL SKILLS Languages: Java, Python, Scala Frameworks: Apache Spark, Hadoop, Apache Flink Cloud: AWS, Azure, GCP Tools: Git, Docker, Jenkins Databases: MongoDB, Cassandra, HBase OS: Linux, Windows, macOS

---

## EDUCATION

B.S. in Computer Science, XYZ University, 2022 GPA: 3.5/4.0 Relevant Courses: Distributed Systems, Big Data Analytics, Data Mining, and Machine Learning.