

Blima Arguellez

Big Data Engineer

I am a Big Data Engineer with over 3 years of experience in the field. I have a strong background in developing and managing Hadoop clusters for large-scale data processing, ETL pipelines, and data mining. I have experience in using languages such as Java, Python, and Pig, as well as various Big Data technologies such as Hadoop, Hive, Spark, and Kafka. I have successfully designed and implemented data pipelines for various types of data sources. I am also well-versed in developing data models for data analysis and machine learning. I have a good understanding of cloud technologies such as AWS and Google Cloud Platform, and I am comfortable with deploying applications and services in the cloud. I have a passion for data engineering and always strive to stay up-to-date with the latest technologies and trends.

blima.arguellez@gmail.com 

(142) 871-9186 

Hartford, CT 

Education

**Bachelor of Science in
Big Data Engineering at
University of Connecticut,
Storrs, CT**

Aug 2015 - May 2020

Relevant Coursework:
Algorithms and Data
Structures, Database Design
and Administration, Machine
Learning, Data Mining, Data
Visualization, and Cloud
Computing.

Links

[linkedin.com/in/blimaarguellez](https://www.linkedin.com/in/blimaarguellez)

Skills

Hadoop



Apache Spark



Data Warehousing



ETL (Extract, Transform and
Load)



NoSQL databases



Machine Learning



Python/R Programming



Employment History

Lead Big Data Engineer at Big Data Corp., CT

Jan 2023 - Present

- Developed a Big Data infrastructure for Big Data Corp., CT, which enabled the company to process over 30 terabytes of data daily, resulting in a 15% increase in overall productivity.
- Automated the data ingestion process using Apache Spark, reducing manual labor by 80%.
- Designed and implemented a distributed NoSQL database with scalability and high availability, which increased customer satisfaction by 20%.

Senior Big Data Engineer at DataAnalytics Solutions LLC, CT

Aug 2020 - Nov 2022

- Developed a distributed data processing platform using Apache Spark to analyze large volumes of data from various sources, resulting in a 25% reduction in processing times.
- Created a data pipeline for streaming data from various sources into the data lake, increasing data availability by 30%.
- Built an AI-based system to detect anomalies in the data and identify potential risks, resulting in a 90% increase in accuracy of predictions.

Certificates

Cloudera Certified Professional: Data Engineer

Aug 2021

Hortonworks HDP Certified Apache Hadoop Developer

Mar 2020

Memberships

Institute of Electrical and Electronics Engineers (IEEE)

Association for Computing Machinery (ACM)