<p style="text-align:center">A Report on One Day Guest Lecture on
"Spark for Data Analysts: Transforming Big Data into Actionable Insights"
Organised by Department of Computer Science & Technology
In association with IIIC
19.03.2024 (Tuesday)</p>

Organized by: Mr. V. Naveen, Assistant Professor, Department of CST; Mr. D. Suresh, Assistant Professor, Department of CST
Submitted by: Mr. V. Naveen, Assistant Professor, Department of CST
Resource Person: Mr. Nethaji Muliki, Senior Consultant, Prolim Solutions India Pvt. Ltd, Bangalore.
Participants: II CST Department Students.
Attendance: 62 participants (Internal)
Mode of conduct: Offline
**Report Received on 04.04.2024**

Department of Computer Science & Technology, has organized One-Day industrial Lecture on **"Spark For Data Analysts: Transforming Big Data into Actionable Insights"** on 19.03.2024 (Tuesday) at MITS from 10:00 AM to 12:00 PM.

Objective:
The industrial Expert lecture on "Spark For Data Analysts: Transforming Big Data into Actionable Insights" was conducted with the aim of introducing participants to Apache Spark, a powerful framework for big data processing, and demonstrating its utility in deriving actionable insights from large datasets.

Welcome Address:
The event commenced promptly at 10:00 AM with a warm and engaging welcome address to all the audience by the **Mr. V. Naveen**, **Assistant Professor, Department of CST,** Madanapalle Institute of Technology & Science (MITS), Madanapalle. In his address, he highlighted the power of Spark for data analysis and decision-making in their respective domains.
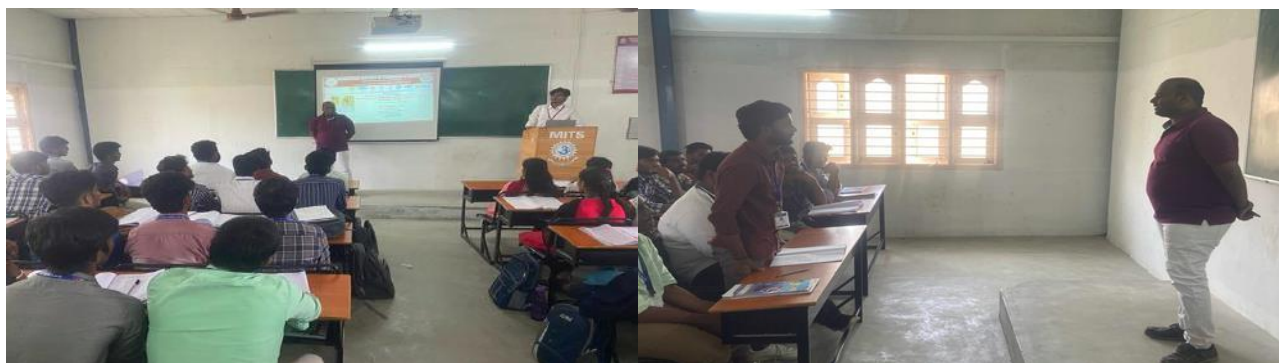
The resource person delivered lecture on the following topics:

Introduction to Apache Spark:
The lecture commenced with an overview of Apache Spark, highlighting its significance in the realm of big data analytics. The speaker elucidated on Spark's capability to handle large-scale data processing tasks efficiently through its in-memory computing capabilities, fault tolerance, and support for various programming languages such as Scala, Java, and Python.

Spark Architecture:
The lecture was divided into the architecture of Apache Spark, elucidating its components such as Spark Core, Spark SQL, Spark Streaming, MLlib, and GraphX. The speaker provided insights into how these components work together to enable distributed data processing across clusters.

Data Processing with Spark:
The practical aspects of data processing with Spark were demonstrated through examples. Participants were guided through the process of loading data into Spark, performing transformations and actions, and leveraging Spark's APIs for data manipulation and analysis. The speaker showcased the versatility of Spark in handling structured and unstructured data formats.

Spark for Data Analysis:
The lecture emphasized Spark's role in facilitating data analysis tasks for data analysts. Participants learned about Spark SQL and DataFrame APIs, which enable SQL-like queries and data manipulation operations on distributed datasets. The speaker highlighted the advantages of using Spark for exploratory data analysis, data cleaning, and feature engineering tasks.

Machine Learning with Spark:
The lecture also touched upon Spark's capabilities in the domain of machine learning. Participants were introduced to MLlib, Spark's machine learning library, which offers a rich set of algorithms and tools for building and deploying machine learning models at scale. The speaker illustrated the process of training machine learning models using Spark, showcasing its efficiency in handling large datasets and model training workflows.

Case Studies and Use Cases:
To provide a real-world perspective, the lecture included case studies and use cases demonstrating the application of Spark in various industries such as e-commerce, finance, healthcare, and telecommunications. Participants gained insights into how organizations leverage Spark for tasks such as personalized recommendations, fraud detection, predictive analytics, and sentiment analysis.

Vote of Thanks:
At 12:00 PM, the guest lecture formally concluded with a vote of thanks delivered by **Mr. D.Suresh, Assistant Professor**, **Department of CST**. In his address, he expressed sincere gratitude to resource person for taking the time to share their expertise on enriching learning experience, equipping participants with the knowledge and skills to harness the power of Spark for data analysis and decision-making in their respective domains.

Outcomes:
At the end of the guest lecture, participants gained insights into:

- Comprehensive understanding of Apache Spark and its capabilities in enabling data analysts to derive actionable insights from large datasets.
- Through practical demonstrations, case studies, and use cases, participants gained valuable insights into the potential applications of Spark across diverse domains.