

Setting Up and Running the Student Performance Spark Analysis Project

Prerequisites:

Apache Spark: Ensure Apache Spark is installed on your machine. Download it from the official Apache Spark website.

MySQL Database: Have a MySQL database installed and running. Create a database named "Student" and execute the provided SQL script to set up required tables.

Project Setup:

Download Code: Obtain the Student Performance Spark code files.

Set Configuration:

Open the test.scala and test1.scala files in a Scala and Spark-compatible text editor or IDE. Update the user and pwd variables in the test object with your MySQL username and password.

Configure Spark:

Ensure proper Spark configuration in the SparkSession creation within the test object. Adjust master and appName parameters as needed.

Load CSV Files:

Place the CSV files (e.g., student.csv) in a directory accessible to your MySQL database. Update file paths in the SQL script to match the location of your CSV files.

Execute SQL Script:

Run the SQL script provided in a MySQL client or command-line interface to create the "Student" database, define tables, and load data from CSV files.

Run Spark Application:

Execute the Student Performance Spark analysis code using an IDE or spark-submit command. Adjust the main method call as needed.

Run Machine Learning Code:

For machine learning, execute the test1.scala file. This contains the machine learning code for predicting student grades.

Run the Student Performance:

Execute the main method in the test object to initiate analyses and methodologies on the loaded data.

Observations:

Examine the console output for results and insights generated by the Student Performance Spark application.

Notes:

Ensure all dependencies are satisfied, including Spark, MySQL, and correct file paths.

Review the Spark documentation for any additional configuration requirements specific to your environment.

By following these instructions, you should be able to seamlessly set up, configure, and run the Student Performance Spark analysis, enabling you to derive meaningful insights from the provided student data. The inclusion of the test1.scala file allows for the execution of machine learning code to enhance the analytical capabilities of the project.