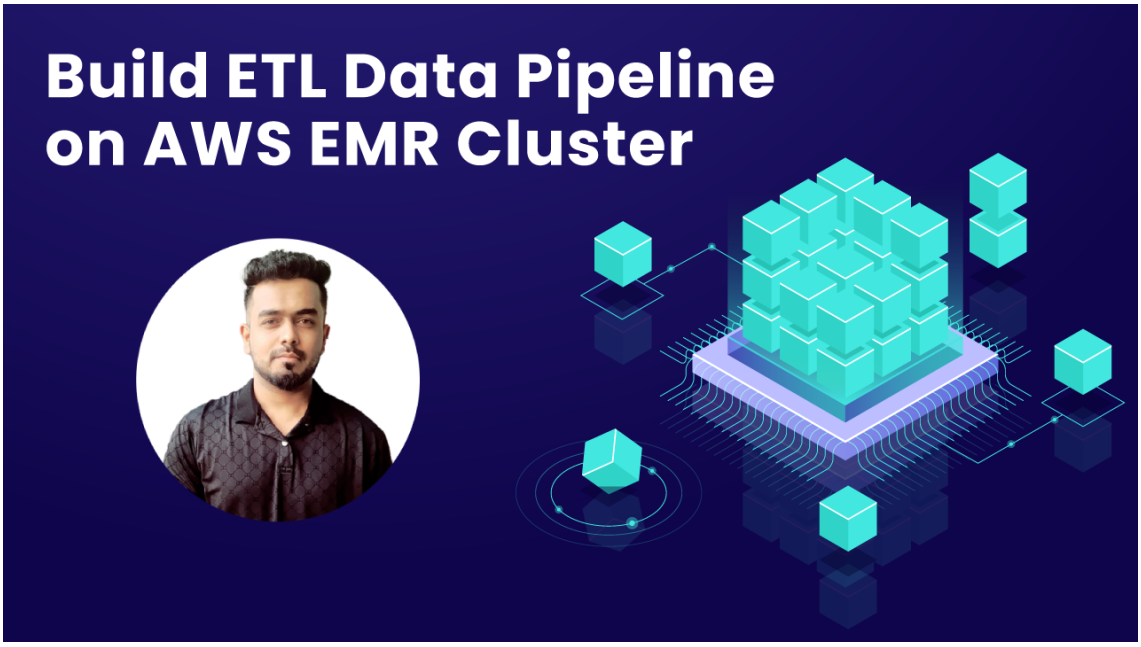
Welcome to ineuron.ai



Build ETL Data Pipeline on AWS EMR Cluster

Description:

With the advent of powerful data warehouses like SnowFlake, BigQuery, redshift spectrum, etc that allow separation of storage and execution, it has become very economical to store data in the data warehouse and then transform them as required. This Project goes over how to design such a ELT system using AWS EMR and Hive. The main objective is to keep the code complexity and server management low, while automating as much as possible

Start Date: 3rd Jan'23

Doubt Clear Time:

Course Time:

Features:

Do Everything In Industry Grade Lab

Learn As Per Your Timeline

Hands-On Industry Real-Time Projects.

Self Paced Learning

Dashboard Access

What we learn:

Real Time Projects

Build ETL Data Pipeline on AWS EMR Cluster

Components of a Data Engineering Platform

Building ETL Pipeline

Store data in the data warehouse

Build Dashboard using Tableau

Hive

Requirements:

System with minimum i3 processor or better

At least 4 GB of RAM

Working internet connection

Dedication to learn

Instructor:

Name:

MD Imran

Description:

Working as Data Scientist with experience in solving real world business problems across different domains.

>Welcome to the Course:

>>Course Overview

>>Dashboard Introduction

>Project :- Build ETL Data Pipeline on AWS EMR Cluster:

>>Introduction of Instructor

>>Introduction to ETL

>>Project Overview

>>End Notes

>>Problem Description

>>Understand the application scope

>>Tour to existing solution

>>End Notes

>>Data Infrastructure: Components used

>>Aws services

>>Data Visualization Tools

>>End Notes

>>Solution Description

>>Data Architecture

>>Tour to Architecture diagram

>>Cost Involved

>>End Notes

>>Exploration of the dataset

>>Creating EMR Cluster

>>Login into EMR hive Project

>>Upload Data into Amazon S3

>>using HIve as ETL Tool

>>Hive Data Insertion

>>CXonnect Tableau to Amazon EMR Hive

>>Plot Charts

>>Plot Dual Combination Charts

>>Other Carts

>>Building Dashboard

>>End Notes

>>Conclude the project

>>Assignments & External Resources