Project 11: Multi-stage Continuous-Flow Manufacturing Process

Company: Uniconverge Technologies Pvt. Ltd.

# Background

This project involved working with real-time process data from a high-speed continuous-flow manufacturing line. The goal was to predict factory output based on input parameters across multiple stages of production. Such models are essential for real-time control systems and anomaly detection.

# Problem Statement

To build a predictive model capable of estimating output properties from input parameters at both Stage 1 (Machines 1–3 and combiner) and Stage 2 (Machines 4–5). The model is expected to assist in simulations and real-time decision making.

# System Design

The manufacturing process was divided into two stages. The dataset contained 14,088 entries and 116 features. Stage 1 combined the outputs of Machines 1–3. Stage 2 processed those combined outputs using Machines 4 and 5. Target variables included 15 measurements from each stage.

# Implementation

XGBoost with MultiOutputRegressor was used for both stages. The dataset was cleaned, and numerical features were separated for Stage 1 and Stage 2 output predictions. Models were trained and evaluated using R² score and RMSE. Train-test split was applied with standard preprocessing.

# Results

Stage 1 model achieved an R² score of 0.704 and RMSE of 1.754.  
Stage 2 model achieved an R² score of 0.7326 and RMSE of 1.8303.

# Learnings

• XGBoost is highly effective for handling large-scale, multi-output regression tasks.  
• Preprocessing is critical when dealing with real-time sensor data.  
• Splitting process stages can improve model clarity and interpretability.  
• Evaluation metrics like R² and RMSE provide essential insights for model accuracy.