



ING. MECATRÓNICA

**Tesis previa a la obtención del título de
Ingeniero en Mecatrónica.**

Autor: Pástor González, José Alejandro

Tutor: Quito Carrión, Angélica Verónica

Modelado computacional de la predicción del tamaño de grano bajo deformación plástica severa mediante conjuntos de datos híbridos de aprendizaje automático.

Computational Modeling of Grain Size Prediction Under Severe
Plastic Deformation Using Hybrid Machine Learning Datasets

Quito-2026

Resumen

El presente Trabajo de Integración Curricular desarrolla un modelo computacional para la predicción del tamaño de grano en materiales sometidos a deformación plástica severa (SPD), a partir de variables de proceso como la velocidad, el ángulo de ataque, la deformación equivalente, la tasa de deformación, la temperatura, el parámetro de Zener–Hollomon y una razón geométrica. El conjunto de datos original estuvo conformado por 7249 registros experimentales; sin embargo, tras un proceso riguroso de limpieza y filtrado físico, se obtuvo un conjunto reducido de 410 muestras confiables. Debido a esta limitación, se evaluaron técnicas de generación de datos sintéticos mediante Gaussian Copula y CTGAN, proponiéndose una estrategia híbrida que combina datos reales, muestreo aleatorio controlado con ruido gaussiano y datos generados por CTGAN restringidos a rangos físicamente consistentes. Se entrenaron y compararon modelos de regresión tipo ensemble, incluyendo Random Forest, Gradient Boosting y XGBoost, utilizando una partición 80–20 que garantiza que el conjunto de prueba contenga únicamente datos reales. Los resultados muestran que el conjunto híbrido no saturante con CTGAN restringido al rango de 0.17–7.87 μm alcanza el mejor desempeño, con coeficientes de determinación cercanos a 0.94 y errores concentrados alrededor de cero. La metodología propuesta constituye una alternativa reproducible y eficiente para reducir la dependencia de ensayos experimentales costosos.

Palabras clave: deformación plástica severa, tamaño de grano, aprendizaje automático.

Abstract

This thesis presents the development of a computational model for grain size prediction in materials subjected to severe plastic deformation (SPD), based on key process variables such as speed, rake angle, equivalent strain, strain rate, temperature, the Zener–Hollomon parameter, and a geometric ratio. The original experimental dataset consisted of 7,249 records; however, after an extensive data cleaning and physical filtering process, only 410 reliable samples remained. To overcome this limitation, synthetic data generation techniques using Gaussian Copula and CTGAN were evaluated, and a hybrid strategy was proposed by combining real data, noise-controlled random sampling, and CTGAN-generated data constrained to physically meaningful ranges. Ensemble regression models, including Random Forest, Gradient Boosting, and XGBoost, were trained using an 80–20 split, ensuring that the test set contained only real samples. The results indicate that the non-saturating hybrid dataset with CTGAN restricted to the 0.17–7.87 μm range achieved the best performance, with determination coefficients close to 0.94 and error distributions centered around zero. The proposed approach offers a reproducible and computationally efficient alternative to extensive experimental characterization.

Keywords: severe plastic deformation, grain size, machine learning.