

# Predictive Model Comparison Report

This report presents a comparative analysis of several predictive models trained to estimate production time.

## Machine Specifications

CPU: Intel Core i7-11800H @ 2.30GHz

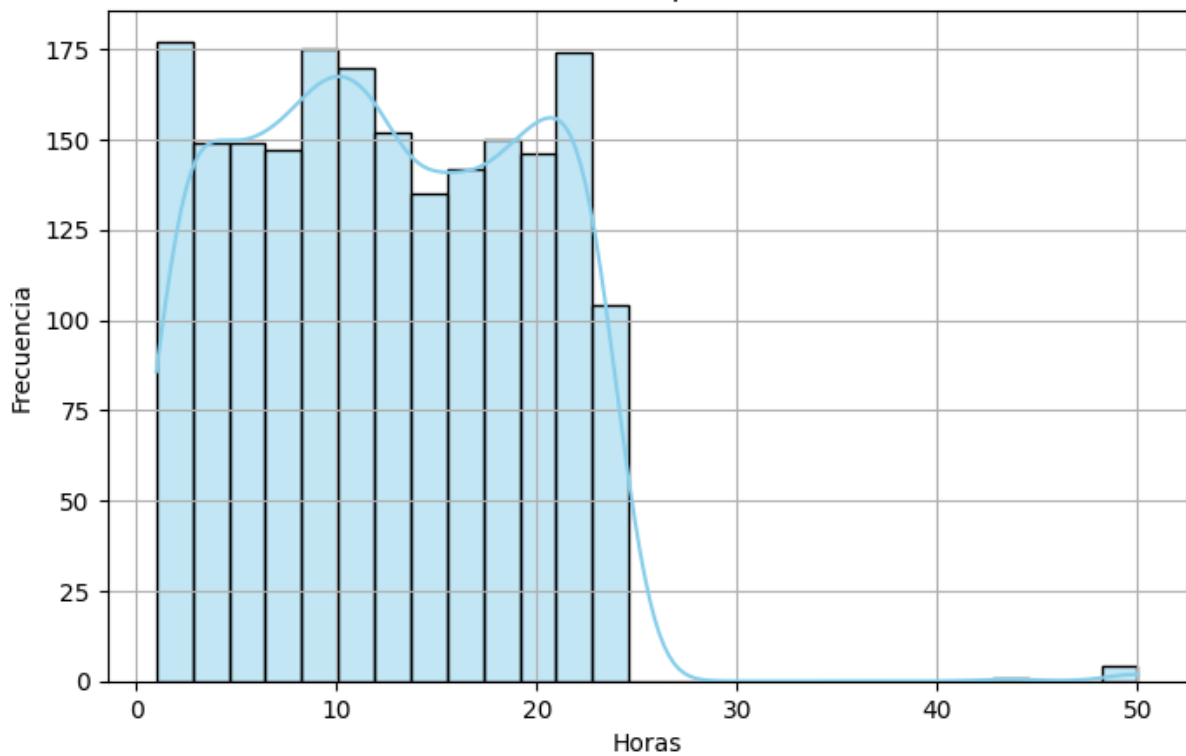
RAM: 16 GB DDR4

GPU: NVIDIA RTX 3060

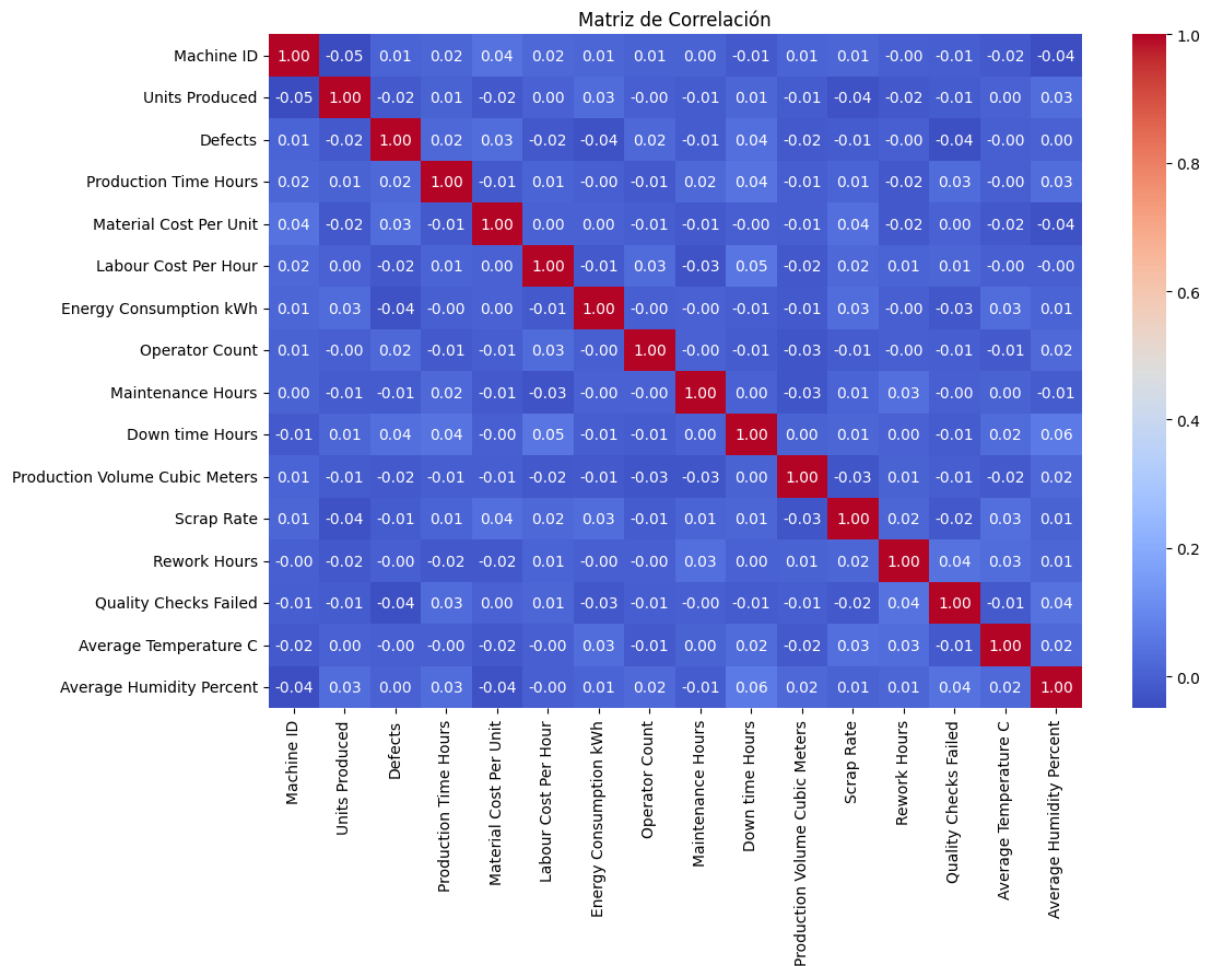
# Exploratory Data Analysis (EDA) Visualizations

Histograma de Producción

Distribución de Tiempo de Producción



Matriz de Correlación



## ⚙ Data Preprocessing

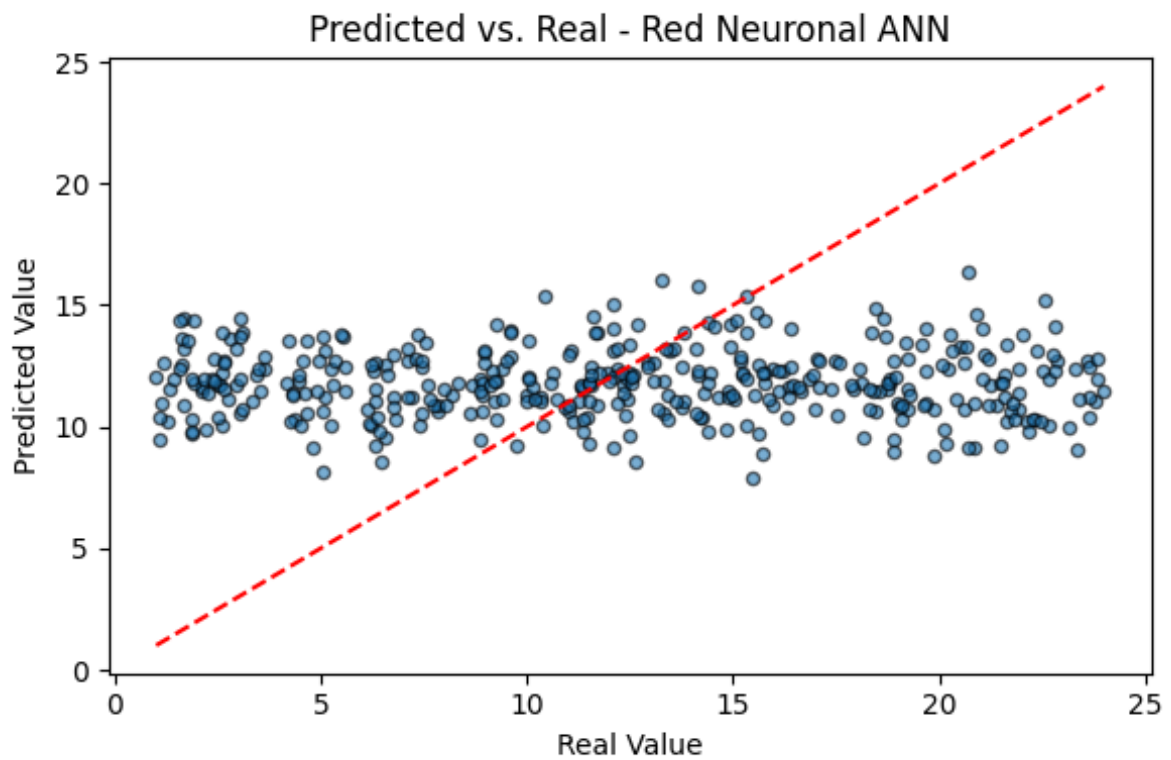
Cleaning, categorical variable encoding (One-Hot Encoding), and feature scaling (StandardScaler) techniques were applied to prepare the data before model training.

## Model Evaluation Metrics

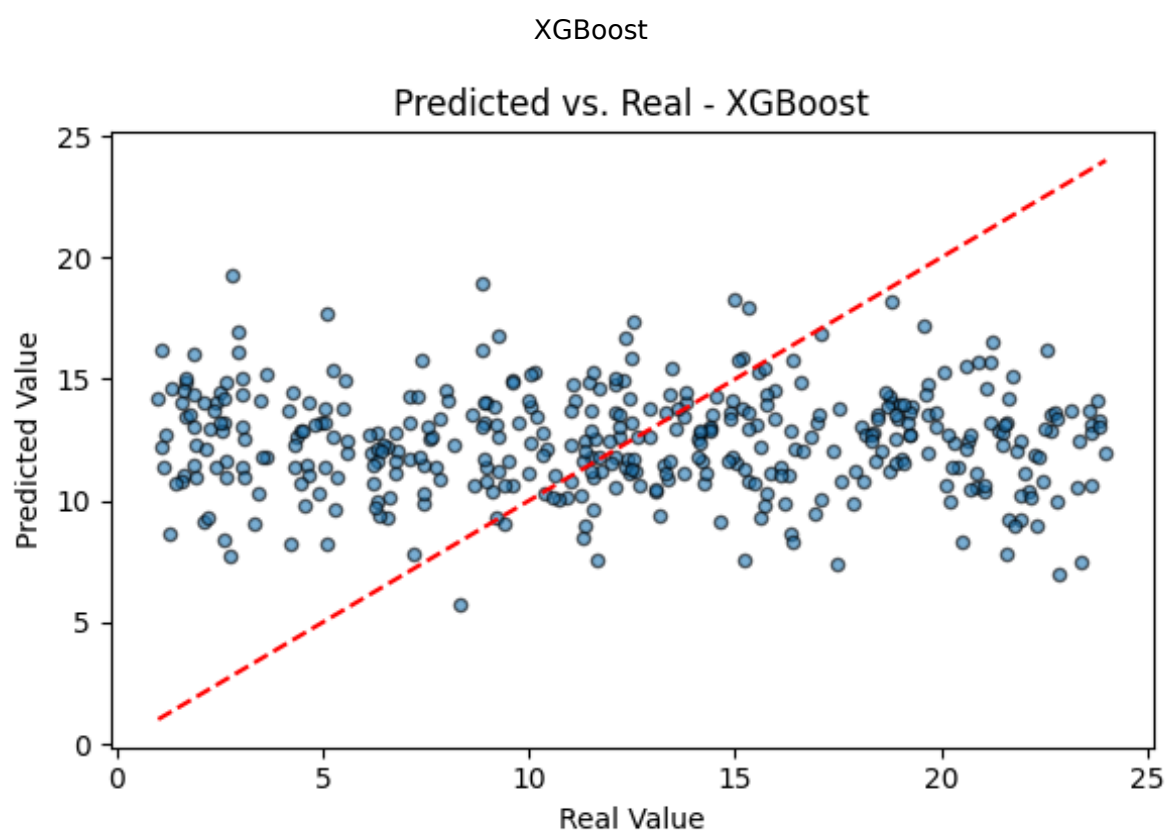
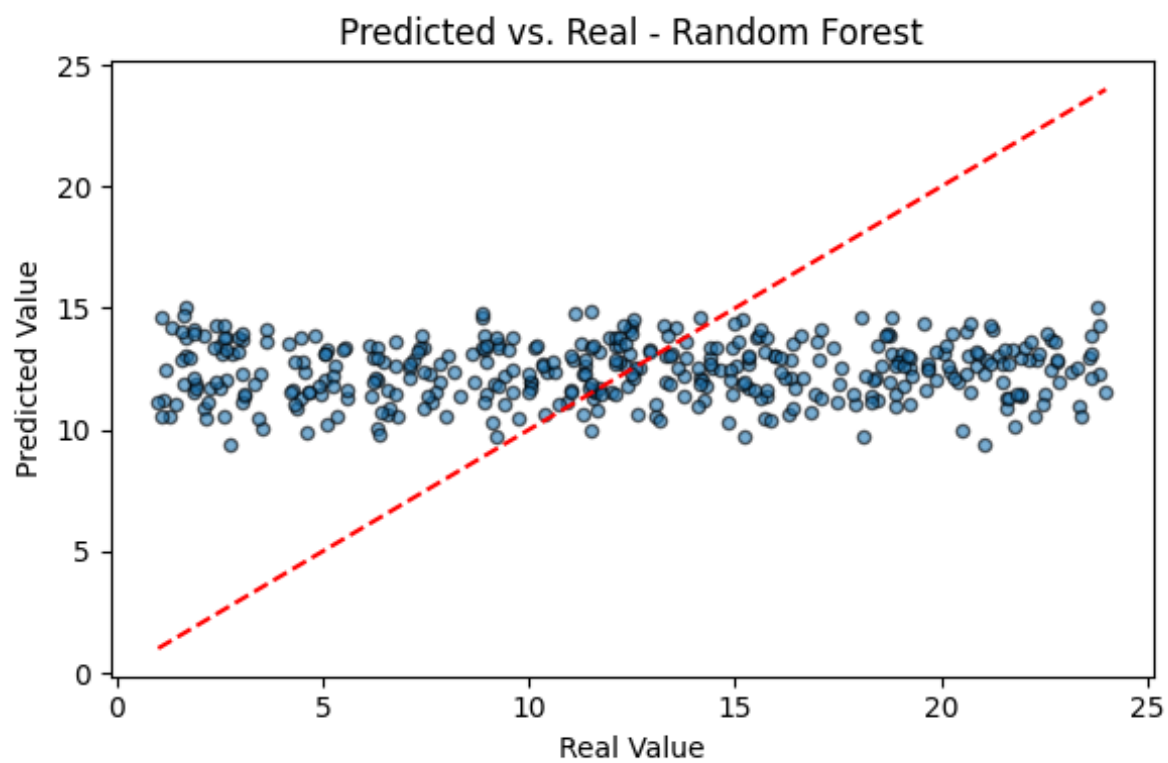
Model	MAE	MSE	R <sup>2</sup>	Time (s)
Red Neuronal ANN	5.625	44.995	-0.052	15.20
Random Forest	5.561	43.346	-0.013	3.10
XGBoost	5.761	47.924	-0.120	2.50

## Predictions vs. Real Values

Red Neuronal ANN



Random Forest



## Theil's U Coefficient

Red Neuronal ANN:  $U = 0.2601$

Random Forest:  $U = 0.2496$

XGBoost:  $U = 0.2610$

## Diebold-Mariano Test

Comparison	DM Statistic	p-value
ANN vs RF	-1.441	0.150
ANN vs XGB	1.720	0.086
RF vs XGB	3.682	0.000

## Conclusion

After analyzing performance metrics, statistical tests, and training times, the model with the best overall performance, considering the balance between accuracy and efficiency, was: Random Forest.