

# Results Evaluation (Primera iteración)

La etapa de evaluación compara modelos en conjunto según las métricas definidas y analiza el rendimiento en unidades operativas. El informe de evaluación debe incluir ranking de modelos, análisis de errores, sensibilidad a hiper parámetros y una recomendación fundamentada que vincule métricas técnicas con criterios de negocio.

## 4.1. BenchMark de modelos

Los resultados obtenidos son:

Ranking	Modelo	R <sup>2</sup> Mean	R <sup>2</sup> Std	RMSE Mean	RMSE Val
1	Random Forest	0.7046	0.0004	2.8656	0.0078
2	Ridge	0.4865	0.0053	3.7782	0.0119
3	Linear Regression	0.4865	0.0053	3.7782	0.0119
4	Lasso	0.2386	0.0031	4.6009	0.0208

Debemos que tener en cuenta que a pesar de que el error estándar de Ridge o Linear Regression sea mayor solo por un poco, se debe considerar que su error cuadrado (RMSE Mean) es considerablemente mas alto, lo que nos indica que el modelo es variable a datos nuevos

Por eso mismo por y por el error estándar muy bajo es que nos decantamos mejor por nuestro **Random Forest**, este mismo puede refinarse mediante GridSearchCV, ajustando sus hiper parámetros para mejorar la precisión y estabilidad.

## 4.2. Ajuste de Hiper Parámetros al mejor modelo

Dado el desempeño superior de Random Forest, se aplicó **GridSearchCV** para optimización de hiper parámetros:

```
Python
param_grid_rf = {
    "model__n_estimators": [100, 300],          *# Número de
    árboles*
    "model__max_depth": [None, 10, 20],        *# Profundidad
    máxima*
```

```
"model__min_samples_split": [2, 5]          *# Muestras  
mínimas para dividir nodo*  
}
```

#### Resultados:

- **R<sup>2</sup> óptimo:** 0.7106
- **Mejora:** 0.7033 → 0.7106 (+0.73 puntos porcentuales)
- **Mejores hiper parámetros:** [identificados por GridSearchCV]

Aunque la mejora es modesta, representa un incremento en capacidad predictiva manteniendo la estabilidad del modelo. En general: El Random Forest demostró ser un modelo bastante robusto para este primer acercamiento a la problemática.

### 4.3. Validación Visual del Modelo Final

Se generó una scatter plot de **predicciones vs valores reales** usando `cross_val_predict` con el mejor estimador:

#### Características del análisis:

- Predicciones obtenidas mediante validación cruzada (no en conjunto de entrenamiento)
- Línea de referencia  $y=x$  para identificar desviaciones
- Ejes con escala 1:1 para interpretación proporcional
- Métricas finales:  $R^2 = 0.711$ , RMSE = 1.44 kg

#### Observaciones:

- Los puntos se concentran cerca de la línea diagonal, indicando buena calibración
- Hay varios puntos dispersos que podrían considerarse como Outliers

### 4.4. Reproducibilidad y Control de Calidad

#### Configuraciones implementadas:

- Seeds fijas (`random_state=42`) en todos los componentes aleatorios
- Métricas estandarizadas ( $R^2$ , RMSE) para comparación objetiva
- Documentación de transformaciones y decisiones metodológicas

Con estas configuraciones pretendemos dar claridad a nuestros hallazgos y descubrimientos

### 4.5. Conclusiones y Recomendaciones

El Random Forest con sus hiper parámetros optimizados es el candidato final, justificado por:

- Mayor  $R^2$  y menor RMSE en validación cruzada
- Menor variabilidad entre folds (mayor estabilidad)
- Capacidad de capturar relaciones no lineales relevantes en el sistema de ordeña

Este enfoque sistemático garantiza que la decisión de selección de modelo está basada en evidencia cuantitativa y es reproducible por terceros. Sin embargo, la naturaleza de hacer un problema de regresión no fue lo suficientemente robusta para ajustarse bien a los datos. Un error de 0.71% sigue siendo alto considerando que ya optimizamos de la mejor manera a nuestro RF.

Es necesario abordar la problemática de otra manera con la idea de cumplir nuestros objetivos de negocios y de minería de datos. Por ende, tenemos que retroceder y reevaluar otro enfoque a los datos y probar con otras arquitecturas o datos de entrada.