

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

# Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Auto: Jose Luis Casado Valero

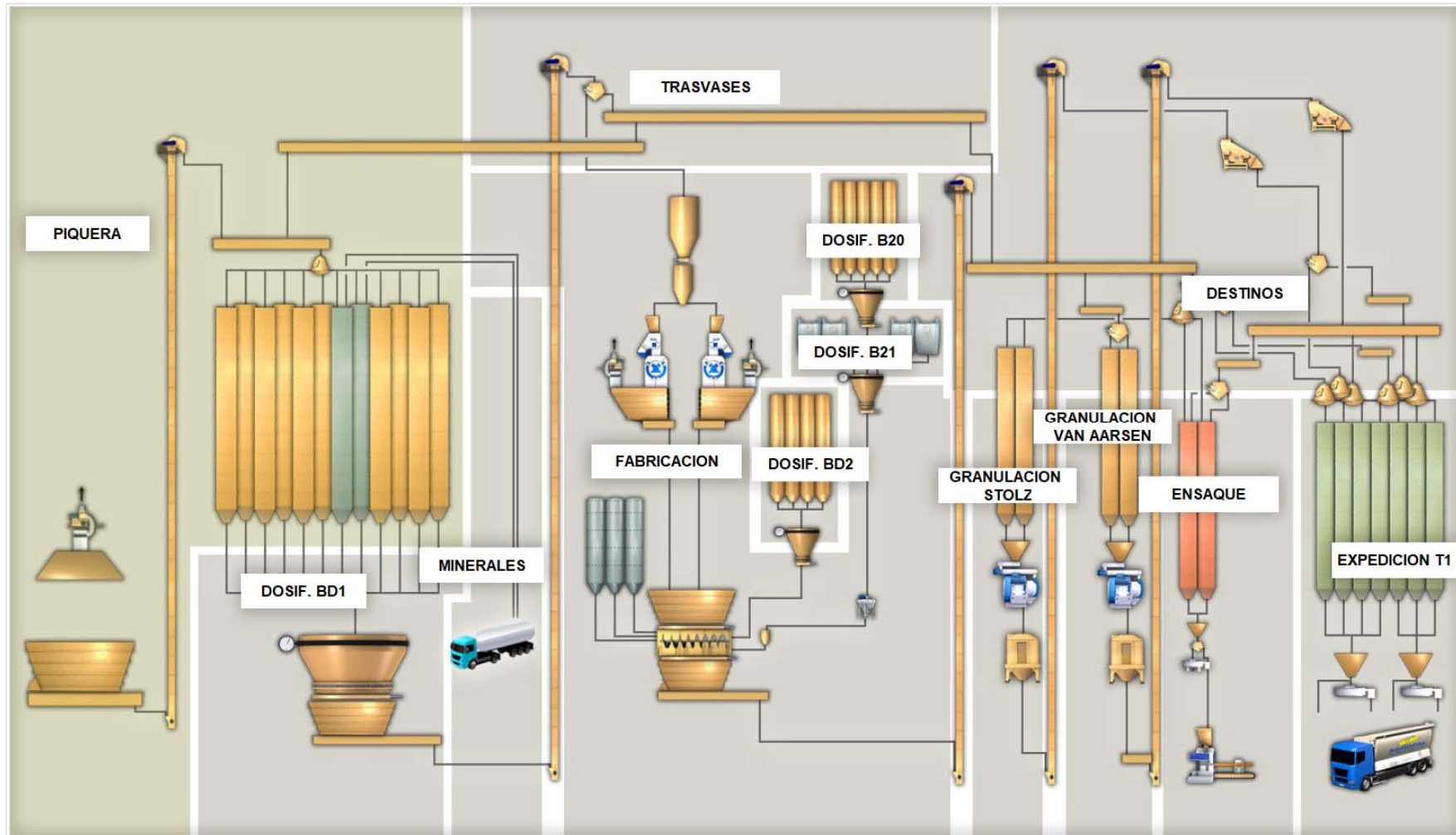
Tutor: Luis de la Ossa Jiménez

Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**

## Fabrica de piensos:

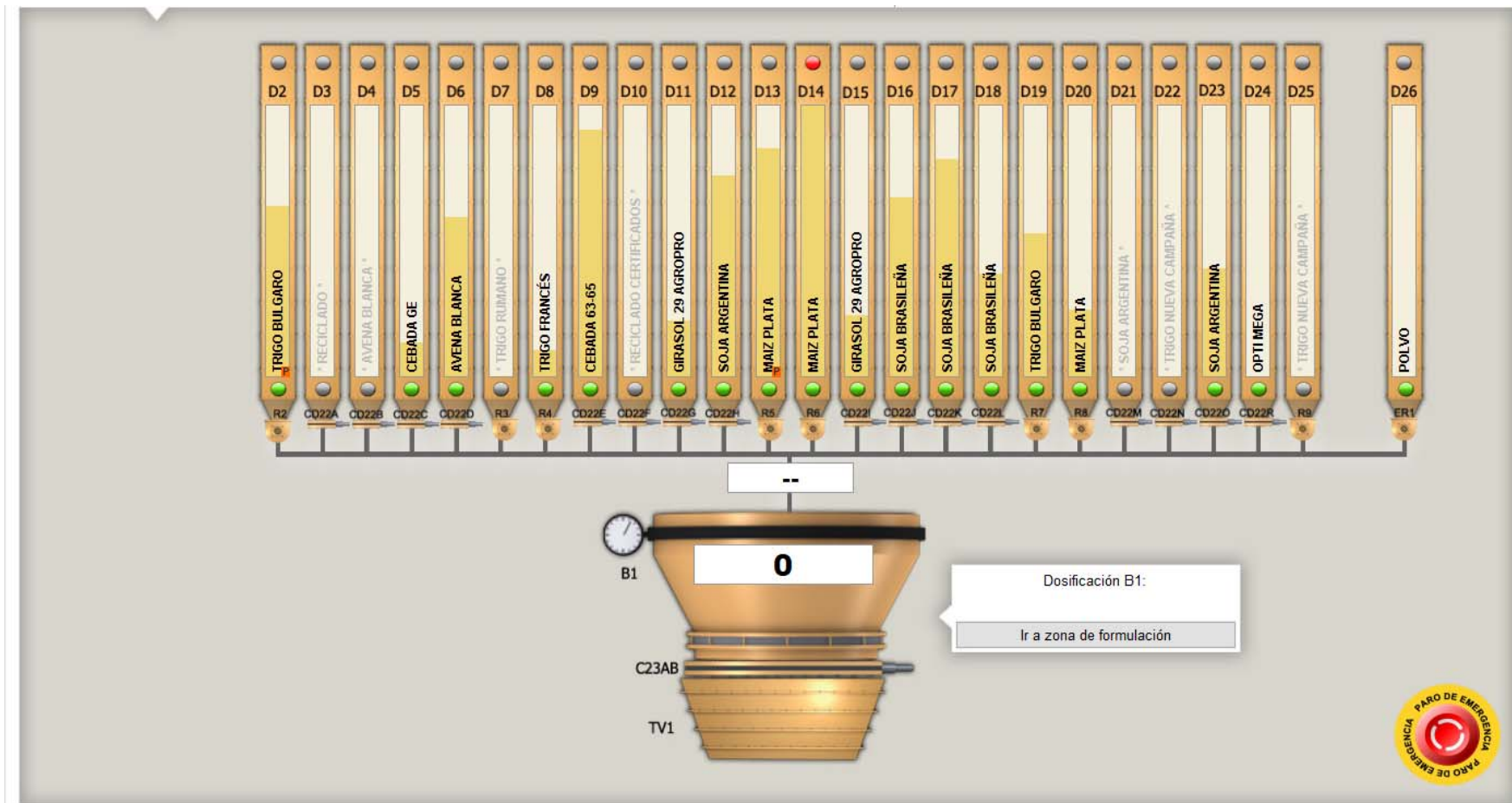


Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Dosificación en bascula:

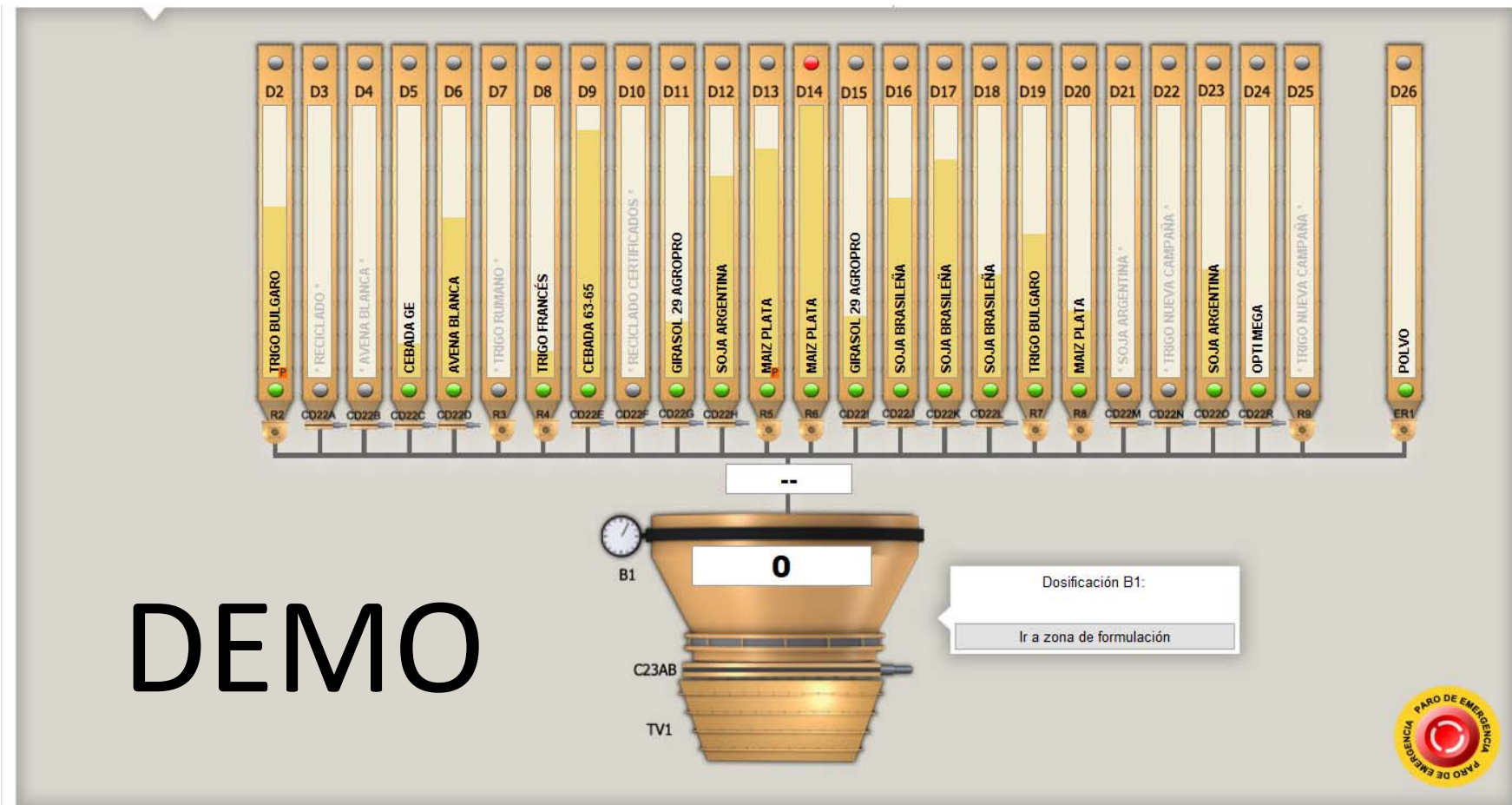


Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Dosificación en bascula:



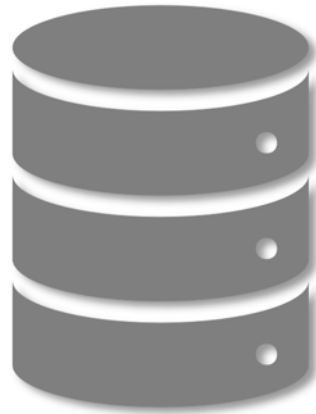
Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Fuente de datos:

- Registros de trazabilidad aplicación de automatización.
- 395.477 registros.
- No existen valores perdidos.



Nombre columna	Ejemplo
Fecha_inicio	2013-08-14 16:51:48
Fecha_fin	2013-08-14 16:52:39
Cantidad_solicitada	235.026
Cantidad_dosificada	236.5
Mezcla	1
Pesada	0
Peso_inicial	0.0
Manual	0
Id_lote_destino	232527
Materia_origen	13
Materia_destino	66
Id_silo	90
Tipo_materia	Prima
Tipo_destino	Premezcla
Densidad	1.0
Tam_mezcla	500
Desviacion	1.47

Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**

## Objetivo:

- Esgrimir los datos en busca de información para entender las desviaciones.
- Prevenir las desviaciones en las dosificaciones.
- Conseguir un método complementario al cálculo de colas actual.

## Planteamientos:

- Predicción del error. (Difícil de caracterizar)
- Predicción de duración.
  - o Corrección de cola condicionada a la duración.

## Duración:

- Variable a predecir.
- Segundo que está activo el elemento dosificador.
- Dataset. Segundos que transcurren desde fecha inicio a fecha fin.
- Ejemplo: 51 seg.

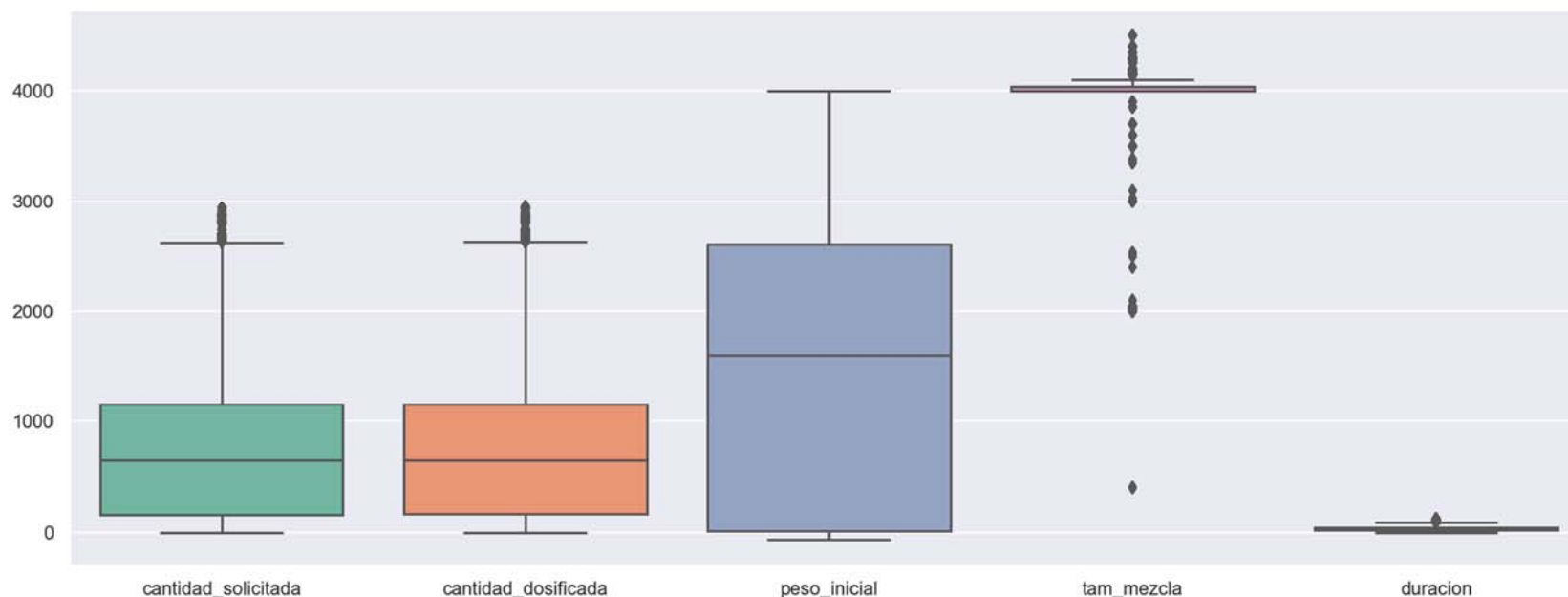
Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**

## Estadísticos variables numéricas:

	C. Solicitada	C. Dosificada	P. Inicial	T. mezcla	Duración	Desviación
Count	395477	395477	395477	395477	395477	395477
Mean	801.251	800.430	1342.961	4000.817	37.613	-0.8203
Std	744.934	745.126	1249.587	211.905	37.362	30.4486
Min	0.000	0.000	-64.000	400.000	0.000	-2623.09
25%	167.984	168.000	11.000	4000.000	16.000	-3.000
50%	668.040	664.000	1574.000	4000.000	25.000	0.0040
75%	1163.096	1168.000	2600.000	4040.000	49.000	3.0198
Max	3121.984	3127.000	3995.000	4500.000	7786.000	1179.891



Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**



## Filtrado de datos:

- Cantidad solicitada mayor de 0.
- Cantidad solicitada menor de 1300.
- Cantidad dosificada mayor de 1.
- Cantidad dosificada menor de 1400.
- Desviación  $> -100$ .
- Desviación  $< 100$ .
- Peso inicial de la báscula menor de 2700.
- Duración menor de 120 segundos.

Como resultado nos queda una base de datos con **235.045** registros.

Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**



### Filtrado de columnas:

- Fecha\_fin.
- Cantidad\_dosificada.
- Mezcla.
- Pesada.
- Manual.
- Densidad.
- Tipo\_materia.
- Tipo\_destino.
- Desviación.

Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

### Creación de nuevas características:

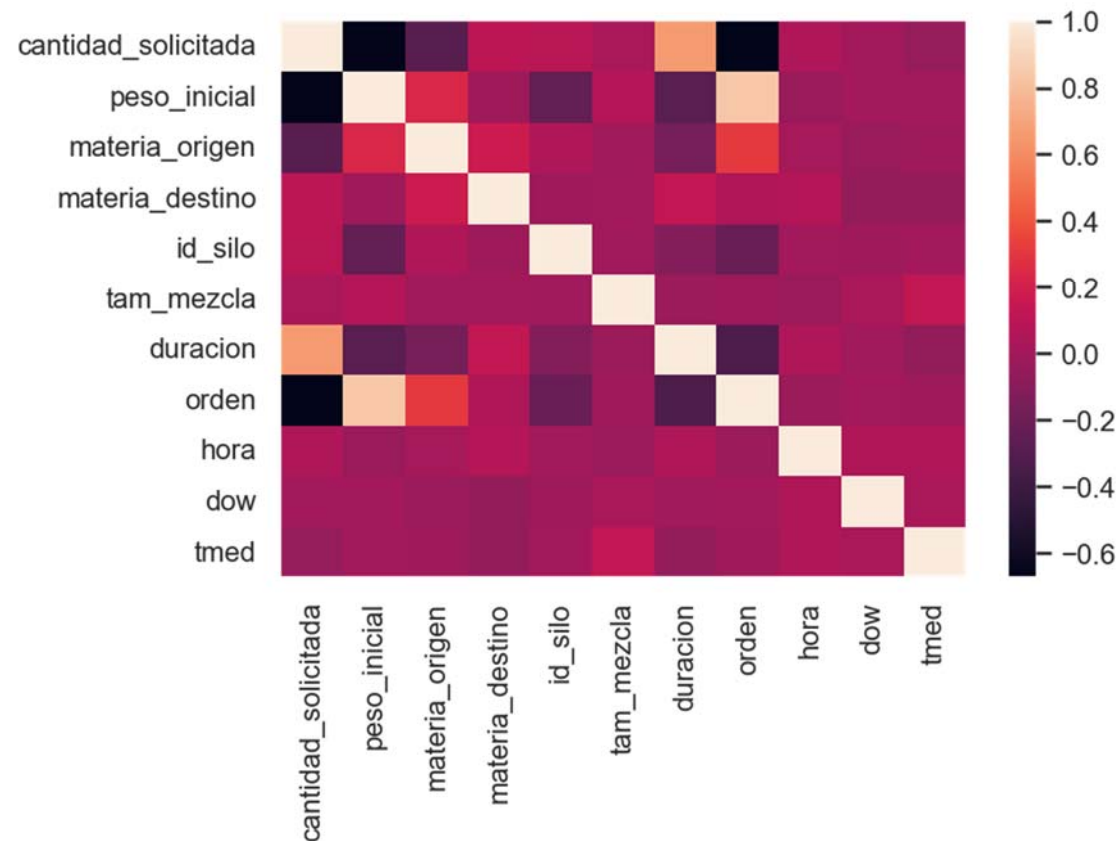
- Orden. Orden de cada dosificación dentro de cada mezcla.
- Hora. Hora a la que se produce la dosificación.
- DOW. Día de la semana en el que se produce la dosificación.
- Tmed. Temperatura media del día (OpenData AEMET).

Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**

Correlación entre variables:

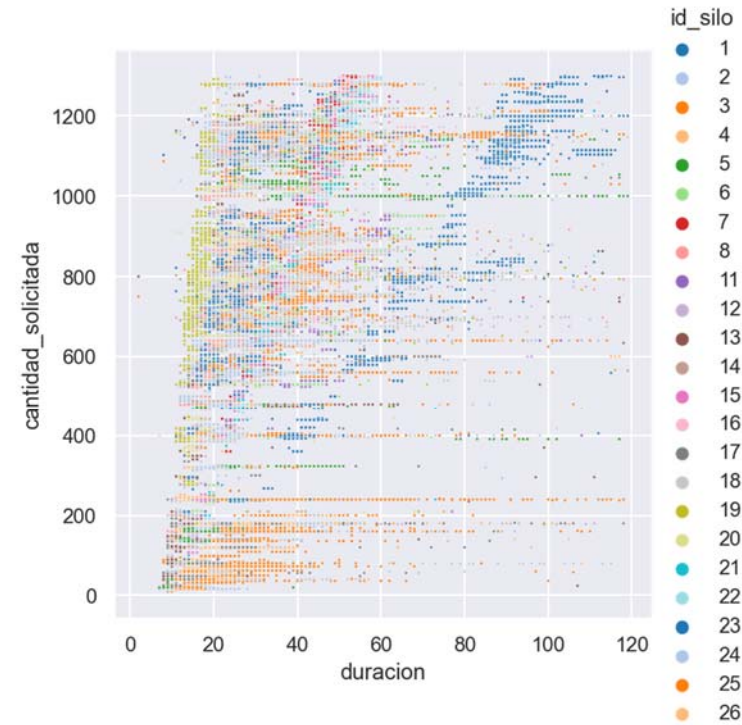
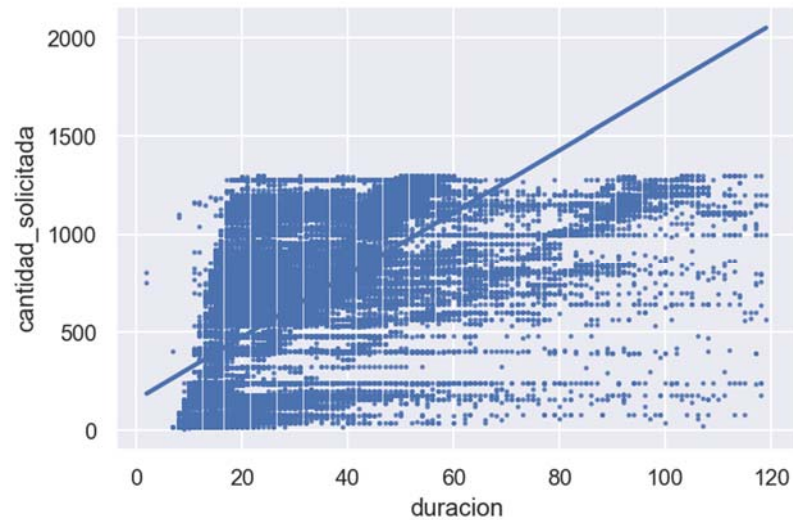


Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

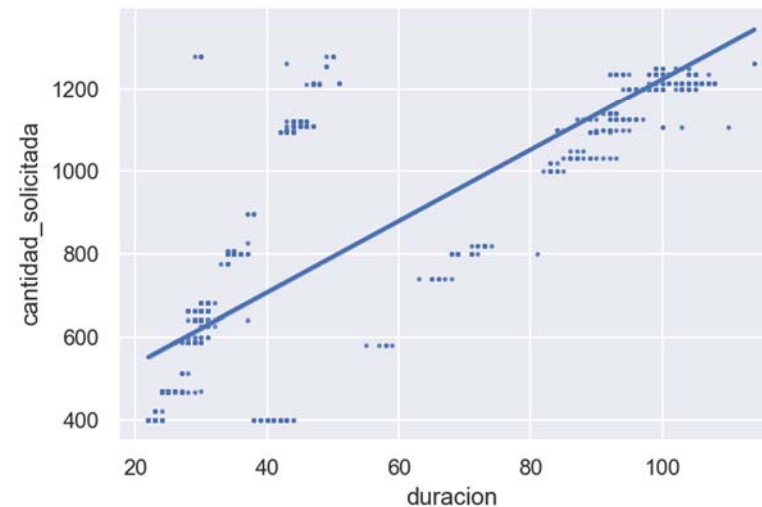
Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Cantidad solicitada – duración:



Por silo



Por silo y  
Materia

Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Preprocesamiento:

- Numéricas: Cantidad solicitada, peso inicial, tamaño mezcla y tmed.
- Categóricas: Materia origen, Materia destino, silo, orden, hora, dow.



## Train y test:

- Aleatorio. 66% y 33%
- Curiosidad: Sobreajuste al hacerlo cronológicamente.

Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Entrenamiento de modelos:

- Utilizamos le error absoluto medio.
  - o Cuantificar y comparar modelos
  - o Medida de valoración desde el punto de vista operativo.

Modelo	C.V búsqueda de Hiperparametros
R. Lineal	
Ridge	Alpha = {1, 0.1, 0.001, 0.0001}
Random Forest	Estimadores = {50, 100, 200, 1000}
Random Forest 1 silo	Max Atributos = 'auto', 3, 5, 7, 10
Random Forest 1 silo 1 materia	Max Profundidad = None, 3, 5, 10, 20

Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Resultados:

Modelo	Error Train	Error Test	Hiperparametros
R. Lineal	4.18 Seg.	4.15 Seg.	
Ridge	4.18 Seg.	4.15 Seg.	Alpha = 1
Random Forest	1.31 Seg.	1.55 Seg.	Estimadores = 200
Random Forest 1 silo	0.39 Seg.	0.87 Seg.	Max Atributos = 'auto'
Random Forest 1 silo 1 materia	0.35 Seg.	0.88 Seg.	Max Profundidad = 20

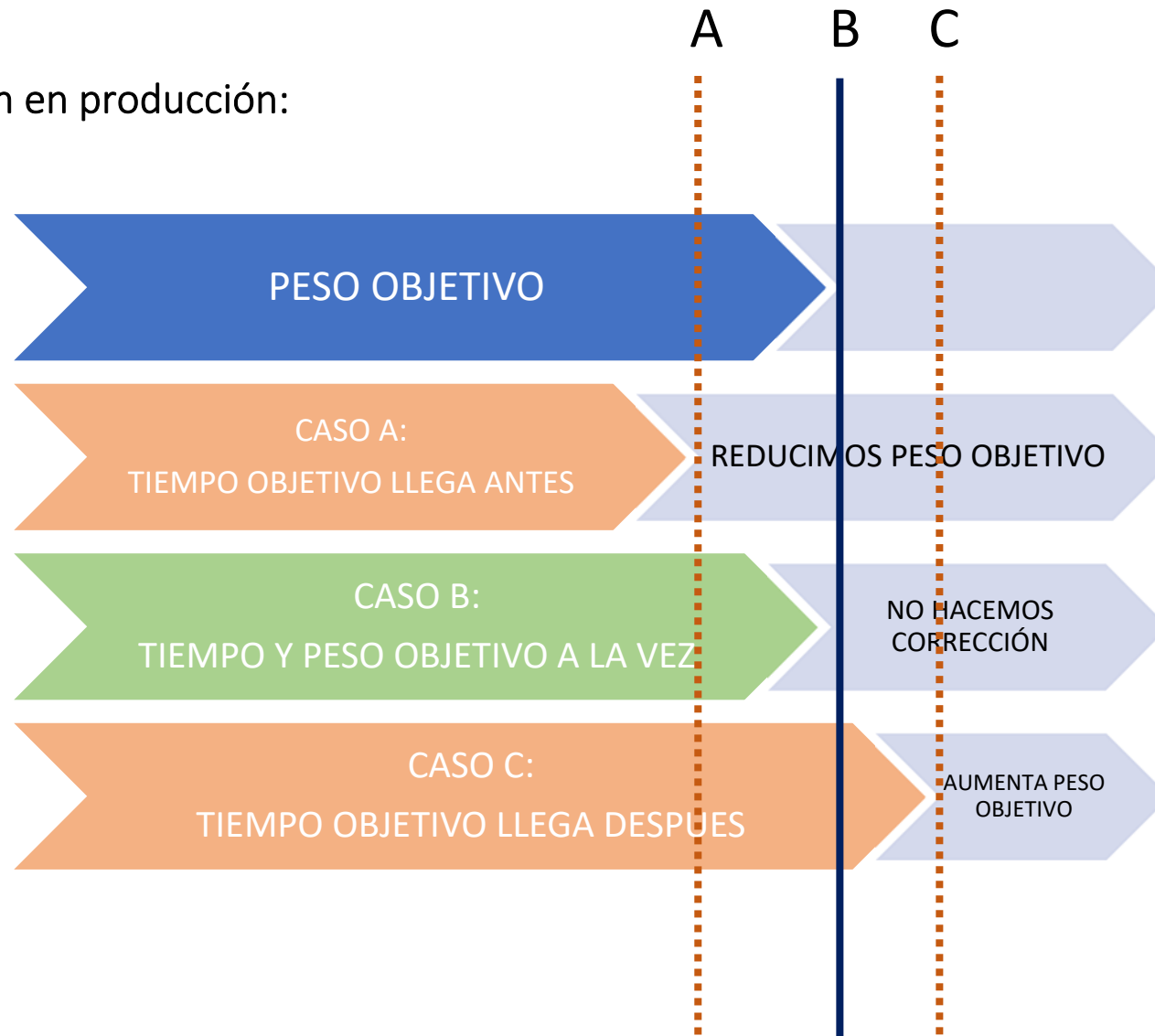
Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**



Implantación en producción:



Predicción de duración de dosificación para corrección de colas en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

Máster en Ciencia de Datos e Ingeniería de Datos en la Nube

## Conclusiones:

- Importancia de la fase de análisis.
- Importancia Generación de características.
- Random Forest funciona perfectamente con gran cantidad de datos y variables.
- En nuestro problema, la dosificación está muy caracterizada por el diseño y la disposición del elemento mecánico.

## Trabajos futuros:

- Modelo de clasificación de materia en silos en función de menor desviación en la dosificación.

Predicción de duración de dosificación para corrección de colas  
en procesos automáticos de fabricación de pienso.

Trabajo Fin de Máster

**Máster en Ciencia de Datos e Ingeniería de Datos en la Nube**

# GRACIAS POR SU ATENCIÓN