Gestión del inventario en una empresa del sector farmacéutico mediante algoritmos de Machine Learning

Jordi de Mas Jaumot
jde_mas@uoc.edu

Universitat Oberta de Catalunya
Barcelona - España

Máster Universitario en Ciencia de Datos – Presentación del Trabajo de Fin de Máster
Enero 2021

Índice

1. Objetivos
2. Metodología
3. Planificación
4. Estado de la cuestión
5. Diseño e implementación
6. Conclusiones
7. Líneas de trabajo futuro

Objetivo principal

Construir un modelo predictivo basado en algoritmos de Machine Learning que permita pronosticar el stock futuro óptimo de cada material en un sólo paso, con el fin de reducir los costes derivados del stock.

Metodología

Etapas:

- Carga de datos
- Preprocesamiento
- Clusterización
- Modelización
- Análisis de resultados
- Conclusiones

Herramientas software

python™

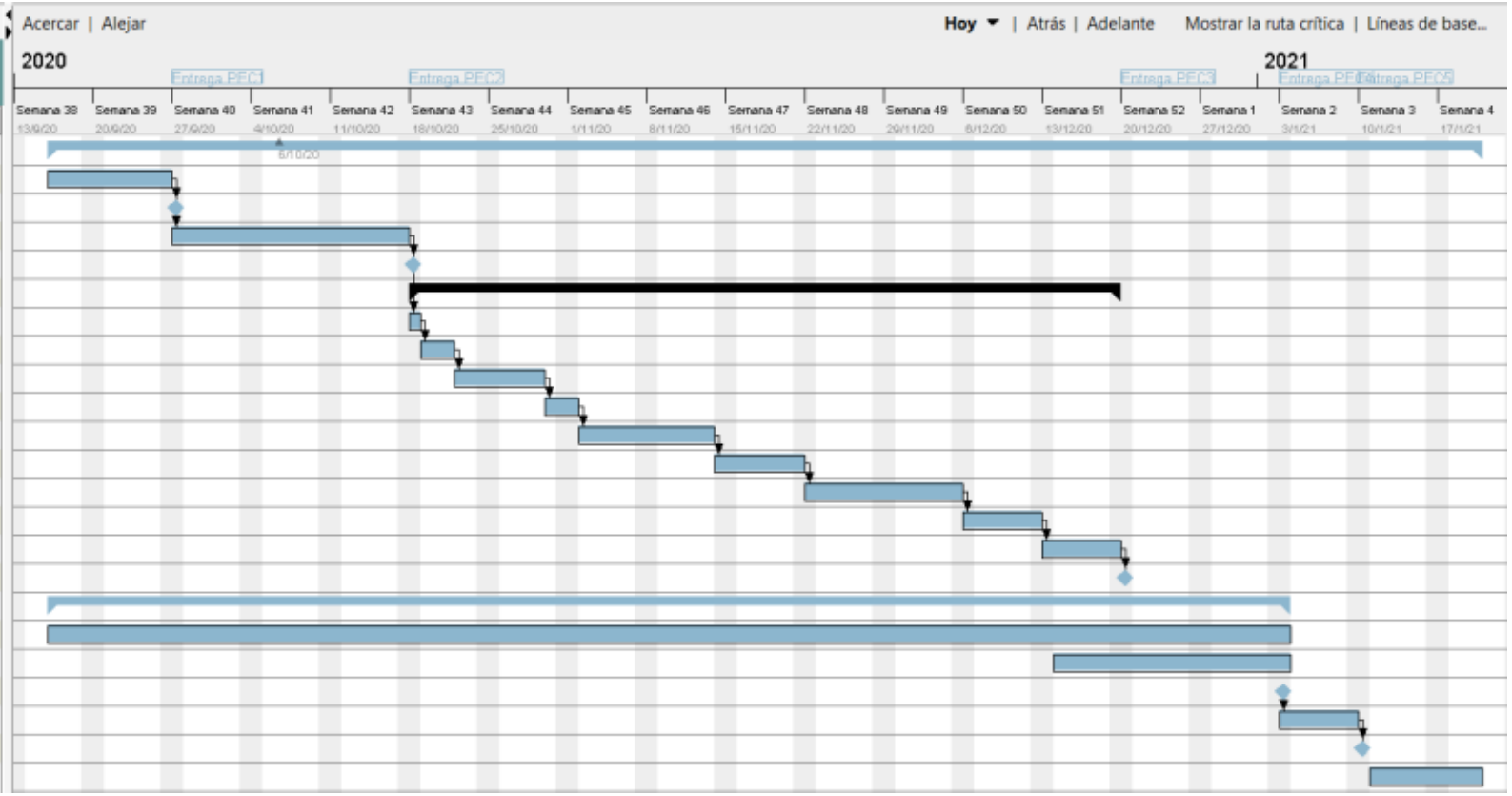
jupyter

ANACONDA.

GanttProject

Planificación

Nombre	Fecha de inicio	Fecha de fin
Trabajo Final de Máster	16/9/20	20/1/21
Definición y planificación del trabajo final	16/9/20	26/9/20
Entrega PEC1	27/9/20	27/9/20
Estado del arte del proyecto	27/9/20	17/10/20
Entrega PEC2	18/10/20	18/10/20
Diseño e implementación del trabajo	18/10/20	19/12/20
Obtención de datos	18/10/20	18/10/20
Carga de datos	19/10/20	21/10/20
Preprocesamiento de datos	22/10/20	29/10/20
Clustering de referencias	30/10/20	1/11/20
Construcción de modelos alternativos	2/11/20	13/11/20
Análisis de resultados obtenidos	14/11/20	21/11/20
Revisión de los modelos	22/11/20	5/12/20
Análisis de nuevos resultados	6/12/20	12/12/20
Conclusiones	13/12/20	19/12/20
Entrega PEC3	20/12/20	20/12/20
Redacción de la memoria	16/9/20	3/1/21
Memoria	16/9/20	3/1/21
Redacción presentación	14/12/20	3/1/21
Entrega PEC4	3/1/21	3/1/21
Presentación y defensa del proyecto	3/1/21	9/1/21
Entrega PEC5	10/1/21	10/1/21
Defensa pública	11/1/21	20/1/21



Estado de la cuestión

- Introducción conceptual
  - Entorno VUCA
  - Equilibrio: Inversión vs. servicio
  - Costes
    - Almacenamiento
    - Emisión de pedidos
    - Adquisición
    - Ruptura de stock
  - Plazo de reposición

Estado de la cuestión

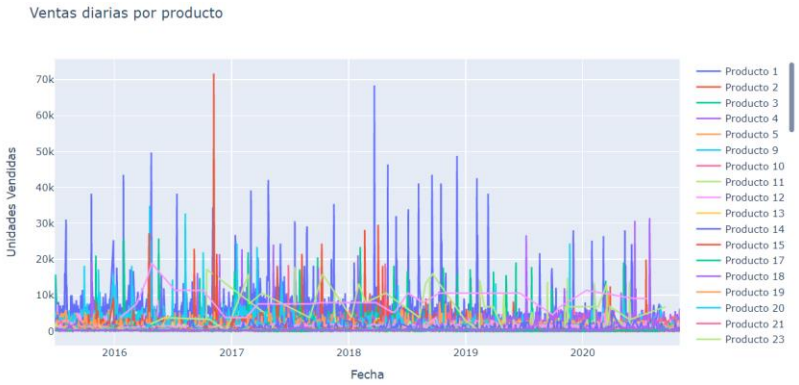
- Revisión bibliográfica
  - Métodos clásicos
  - Métodos Machine Learning

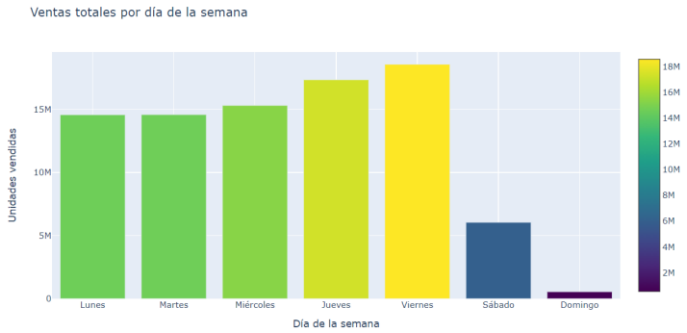
Diseño e implementación

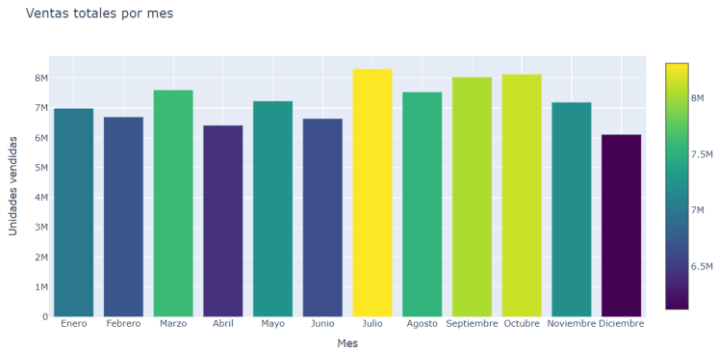
- Recogida de datos
- Estrategia abordaje de proyecto
  - Análisis comportamiento de venta
  - Trabajo con datos de stock
- Preparación, limpieza y preprocesado
  - Juntar datos y añadir variables
  - Tratamiento nulos
  - Eliminación outliers

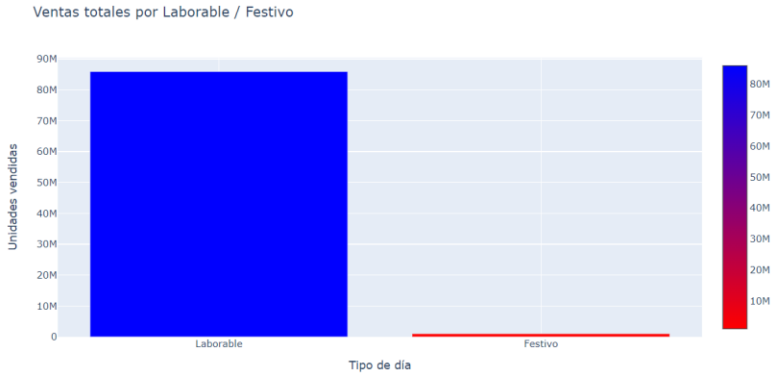
Diseño e implementación

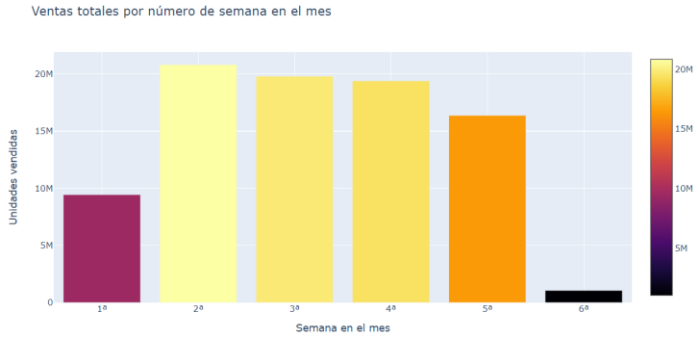
- Análisis de ventas





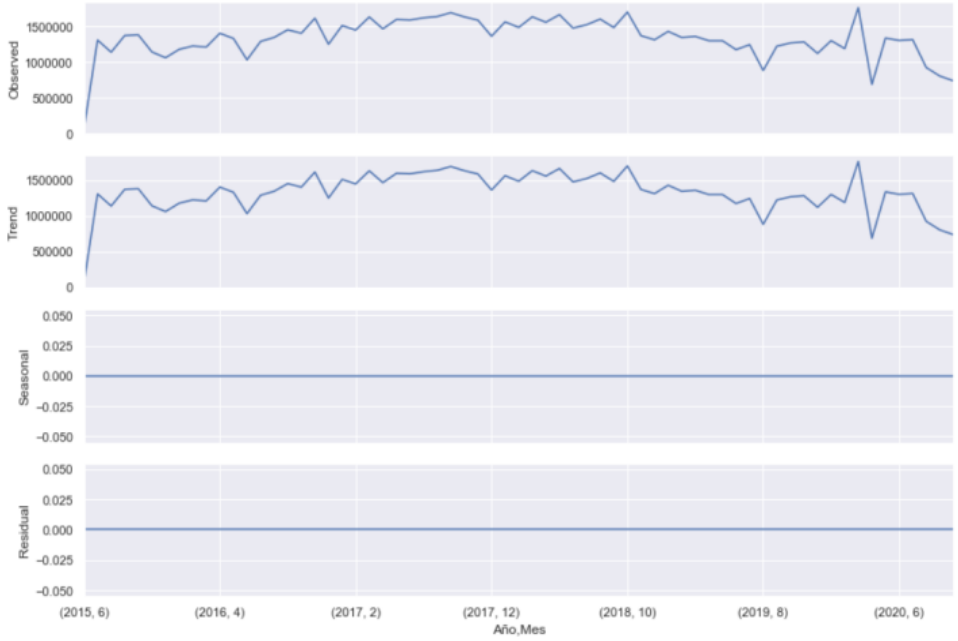


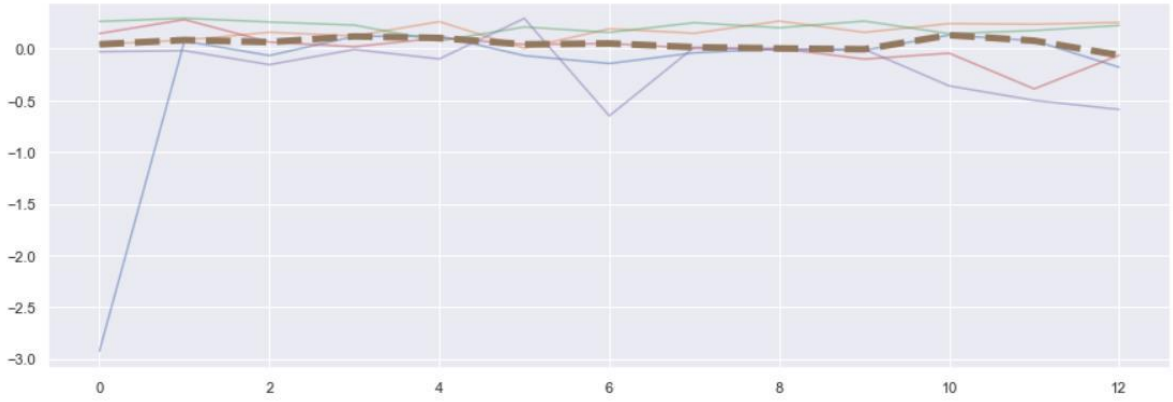




Diseño e implementación

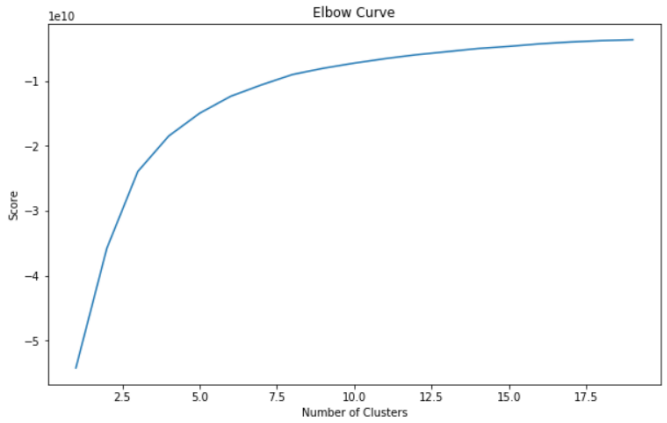
- Análisis de ventas (componentes serie temporal)

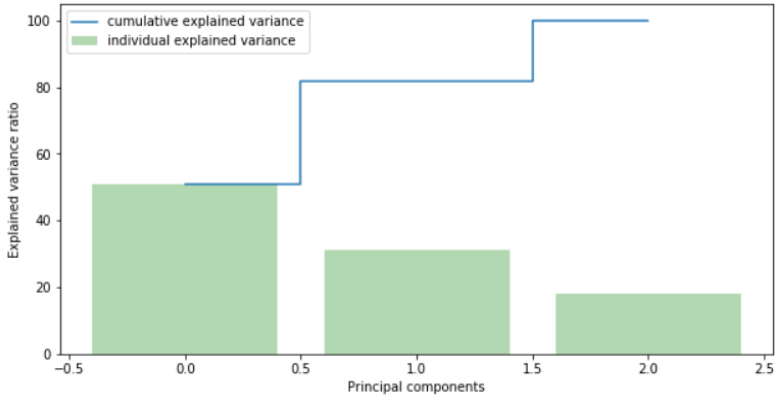


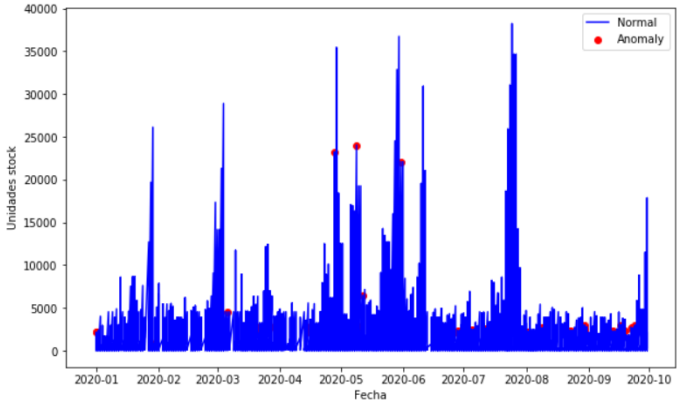


Diseño e implementación

- Tratamiento datos de stock
  - Preprocesado
  - Clusterización y detección automática de outliers

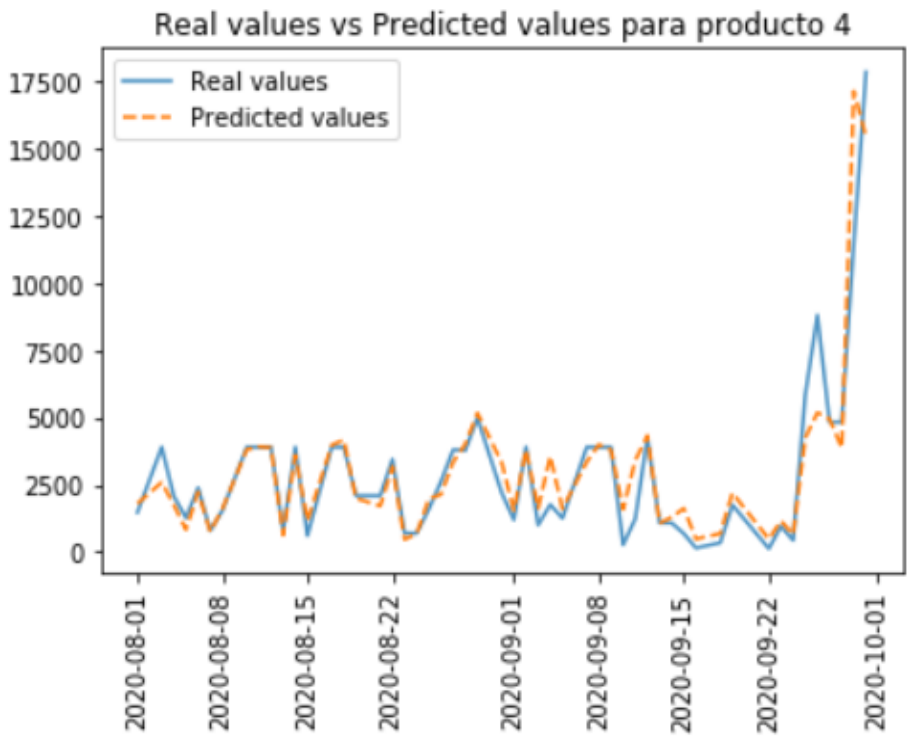






Diseño e implementación

- Tratamiento datos de stock
  - Algoritmos para análisis predictivo
    - Red Neuronal Artificial multicapa con Embeddings
    - Gradient Boosting
    - Random Forest
    - Decision Tree
    - Light Gradient Boosting Machine
    - Extreme Gradient Boosting



Diseño e implementación

Comparativa resultados algoritmos ML

Método	Explained variance score	Var. uds. (abs)	% Var. uds. (abs)	Var. valor (€) (abs)	% Var. valor (abs)
Gradient Boosting	0,86	50.909	2,17	234.863,80	2,32
Extreme GB	0,80	160.093	6,82	677.444,80	6,68
Red Neuronal	0,82	161.043	6,86	840.021,72	8,29
Random Forest	0,85	162.732	6,93	658.840,70	6,50
Light GB	0,75	181.819	7,75	733.691,90	7,24
Decision Tree	0,72	193.398	8,24	788.288,10	7,78

Conclusiones

- Lecciones aprendidas
  - Caso real
  - Importancia fase preparación, limpieza y preprocesado
  - Importancia revisión de la literatura
- Logro de objetivos planteados
- Referentes a planificación y metodología
  - Importancia disponibilidad de tiempo
    - Revisión de literatura
    - Preparación, limpieza y preprocesamiento
    - Afinación de hiperparámetros

Líneas de trabajo futuro

- Conseguir mayor precisión en el modelo de redes neuronales
  - Afinar hiperparámetros incluida la estructura de la red neuronal

- En el resto de algoritmos basados en árboles se debe profundizar en el conocimiento de hiperparámetros para conseguir un mejor ajuste a la serie histórica.

Gracias

¿Alguna pregunta?

Jordi de Mas Jaumot
jde_mas@uoc.edu

Universitat Oberta de Catalunya
Barcelona - España

Máster Universitario en Ciencia de Datos – Presentación del Trabajo de Fin de Máster
Enero 2021