



Segunda entrega proyecto final

Carrera de Data Scientist – Mayo de 2023

Autor: Emilse Bover

Tabla de contenido

| | |
|--|-----------|
| Abstract..... | 2 |
| Objetivo | 2 |
| Contexto comercial | 2 |
| Data acquisition..... | 2 |
| Data wrangling | 2 |
| Base de datos productivos..... | 2 |
| Base de datos climáticos de Concepción del Uruguay, Entre Ríos | 3 |
| Transformaciones de las bases de datos para el análisis | 3 |
| Creación de subset | 4 |
| Análisis exploratorio de datos | 4 |
| Histograma de ganancia media diaria de peso (A.D.) | 4 |
| A.D. por año..... | 4 |
| A.D. por estación del año | 5 |
| A.D. por mes del año | 5 |
| Histograma A.D. por tipo de ventilación | 6 |
| A.D. por sistema de ventilación | 7 |
| A.D. y temperatura máxima..... | 7 |
| A.D. y temperatura mínima | 8 |
| A.D. por zona de producción..... | 8 |
| A.D. por zona de producción dispersión..... | 9 |
| Machine learning aproximación | 9 |
| Modelo SVR | 10 |
| Conclusiones | 10 |
| Insights..... | 10 |
| Recomendaciones | 10 |

Abstract

En presente trabajo se analizaron los resultados productivos de una empresa productora de pollos parrilleros para detectar las variables que impactan sobre los resultados y realizar recomendaciones a partir de ello. En segunda instancia se construyó un modelo de predicción de resultados productivos en función de los datos que se suministren en el futuro.

Objetivo

El objetivo del presente trabajo es detectar el impacto de los factores que afectan al crecimiento de los pollos parrilleros para tomar decisiones que mejoren los resultados productivos.

Contexto comercial

El retorno económico de la empresa productora de pollos parrilleros depende de la obtención de mejores resultados productivos. A mayor ganancia de peso diaria, por ejemplo, el consumo total de alimento será menor disminuyendo el costo. Otro impacto positivo del aumento de la ganancia de peso diaria es que se reduce la edad a faena de los animales ya que se alcanza el peso deseado en menor cantidad de días. Esta reducción de días permite: liberación de superficie de producción (metros cuadrados de galpón, rotación), menor propensión a sufrir enfermedades y accidentes (por ejemplo, cortes de luz). También el dueño de la granja tiene menores costos de luz y gas y por lo tanto mayores ganancias finales también.

Data acquisition

La empresa ha provisto un **archivo Excel con información de resultados** y se utilizaron **datos climáticos históricos** obtenidos a través de una API de Open Meteo.

Data wrangling

Primero se eliminaron valores erróneos de la base de datos que se conocían antes de iniciar el análisis.

Base de datos productivos

Los resultados se agrupan por crianza (índice) de cada granja y se registran los diferentes resultados e indicadores de la producción.

| | granja | Nombre | Primer BB | Cantid. BB | A Faena | % a 7 Días | % Mortan. | % Fal. | Kg. Pollo | P. Prom. | ... | Índice |
|---|--------|-----------------------------------|------------|------------|---------|------------|-----------|--------|-----------|----------|-----|--------|
| 0 | 357 | MENESCARDI FABRICIO MEDARDO | 2014-11-11 | 23200 | 20699 | 1.42 | 10.78 | -1.09 | 58340 | 2.818494 | ... | 1 |
| 1 | 605 | COSSO CARLOS | 2014-11-11 | 23176 | 20678 | 1.51 | 10.78 | -1.88 | 61580 | 2.978044 | ... | 2 |
| 2 | 833 | CARDINAUX MARTIN ULISES | 2014-11-14 | 40000 | 36825 | 1.45 | 7.94 | -2.86 | 102480 | 2.782892 | ... | 3 |
| 3 | 1710 | STEGEMANN DANIELA CRISTINA | 2014-11-11 | 29000 | 26423 | 1.10 | 8.89 | 0.25 | 74800 | 2.830867 | ... | 4 |
| 4 | 2333 | OJEDA LIVIO | 2014-11-13 | 22000 | 19502 | 0.92 | 11.35 | -1.91 | 53020 | 2.718696 | ... | 5 |

La base de datos contiene resultados productivos de 15.571 crianzas desde el año 2015 hasta la primera quincena del mes de abril de 2023, no contiene valores duplicados y solo se hallaron valores en blanco para una columna que no será utilizada en este análisis.

Base de datos climáticos de Concepción del Uruguay, Entre Ríos

| | time | temperature_2m_max | temperature_2m_min | precipitation_sum |
|---|------------|--------------------|--------------------|-------------------|
| 0 | 2020-01-01 | 26.1 | 20.8 | 0.0 |
| 1 | 2020-01-02 | 26.2 | 17.9 | 0.0 |
| 2 | 2020-01-03 | 26.8 | 18.7 | 0.0 |
| 3 | 2020-01-04 | 27.8 | 18.1 | 0.0 |
| 4 | 2020-01-05 | 30.3 | 19.6 | 0.0 |

La base no posee valores en blanco ni fechas duplicadas.

Transformaciones de las bases de datos para el análisis

De la base de datos provista por la empresa se seleccionarán las siguientes columnas para el análisis:

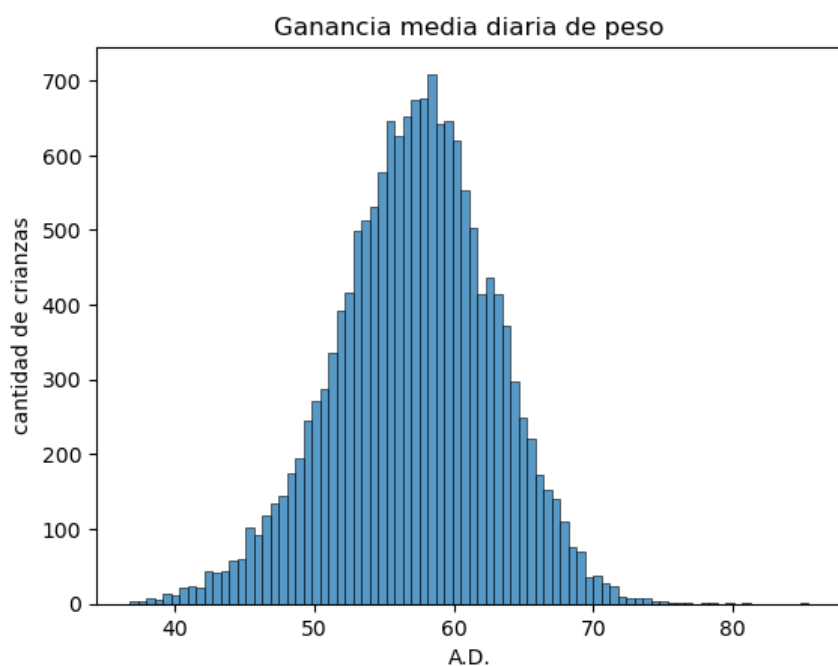
| Nombre | Descripción |
|--------------------------|--|
| granja | Unidad productiva compuesta por galpones |
| Primer BB | Fecha inicio crianza de aves |
| A Faena | Cantidad de aves al final de la crianza |
| % Mortan. | Mortalidad animales |
| P. Prom. | Peso promedio a faena |
| Ulto. Levan. | Fecha fin de crianza de aves |
| Edad | Edad al final de la crianza |
| A.D. | Ganancia media diaria de peso |
| año levante | año de fin de crianza |
| mes levante | mes de fin de crianza |
| Estación del año | Momento del año en que ocurrió el fin de la crianza |
| Consumo medio diario (g) | Consumo de alimento por ave |
| Destino | Destino de venta: mercado interno o exportación |
| Densidad | Cantidad de aves por metro cuadrado de galpón (de cada granja) |
| kg pollo/m2 | Cantidad de kilos logrados por unidad de superficie de granja |
| Ventilación forzada | Tipo de sistema de ventilación de granja |
| Ambiente controlado | Subtipo de sistema de ventilación |
| localidad | Localidad geográfica de las granjas |
| % desvío consumo std | Desvío del consumo de alimento respecto del estándar |
| Índice | Número de crianza (único) |
| zona climática | Zona de producción según clima imperante |
| Pes conv 2,7 | Indicador de eficiencia productivo |
| zona prod | Zona de producción de la granja |

Creación de subset

Se crearon diferentes subset para realizar los análisis de las bases de datos.

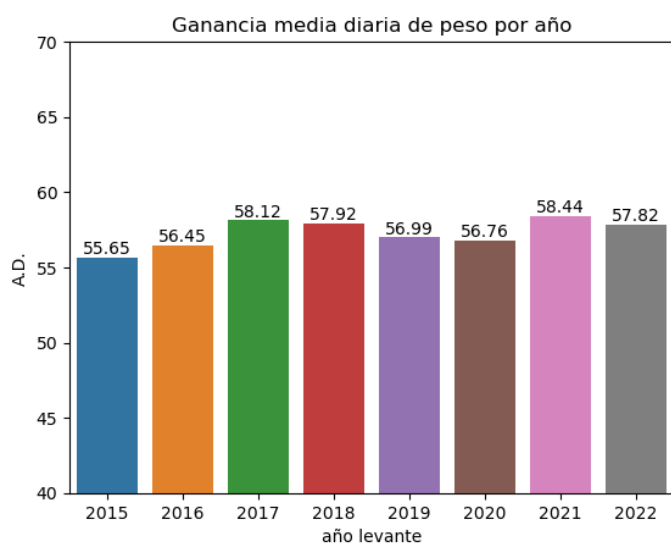
Análisis exploratorio de datos

Histograma de ganancia media diaria de peso (A.D.)



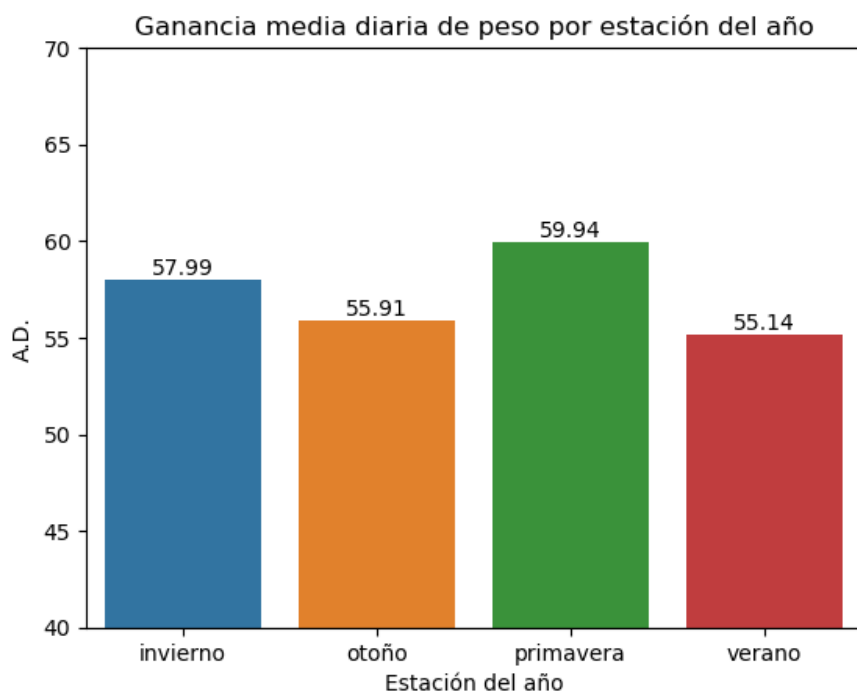
La **ganancia media diaria de peso** es una medida que permite comparar resultados entre granjas y además permite ver la evolución de los resultados globales de la empresa a lo largo del tiempo. A mayor valor, mejor será el resultado. El valor medio es de **57.18** gramos de peso ganados promedio por día.

A.D. por año



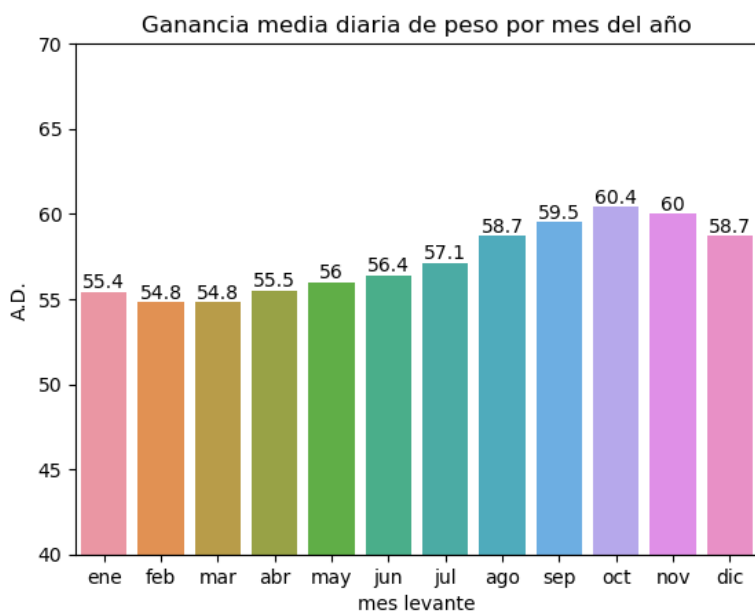
Se observa que las ganancias de peso de los años 2017 y 2021 fueron superiores al resto de los años analizados. El año 2023 fue excluido por estar incompleto.

A.D. por estación del año



Se observa que la ganancia de peso es variable según la estación del año en la que se desarrolle la crianza, las mayores ganancias de peso ocurren en primavera y las peores en verano dando indicios de que el ambiente puede estar afectando el resultado productivo.

A.D. por mes del año



Se observan también diferencias en los valores de ganancia media diaria de peso en los meses del año. Debido a ello a continuación se realizará un análisis según los tipos de ventilación que poseen las granjas.

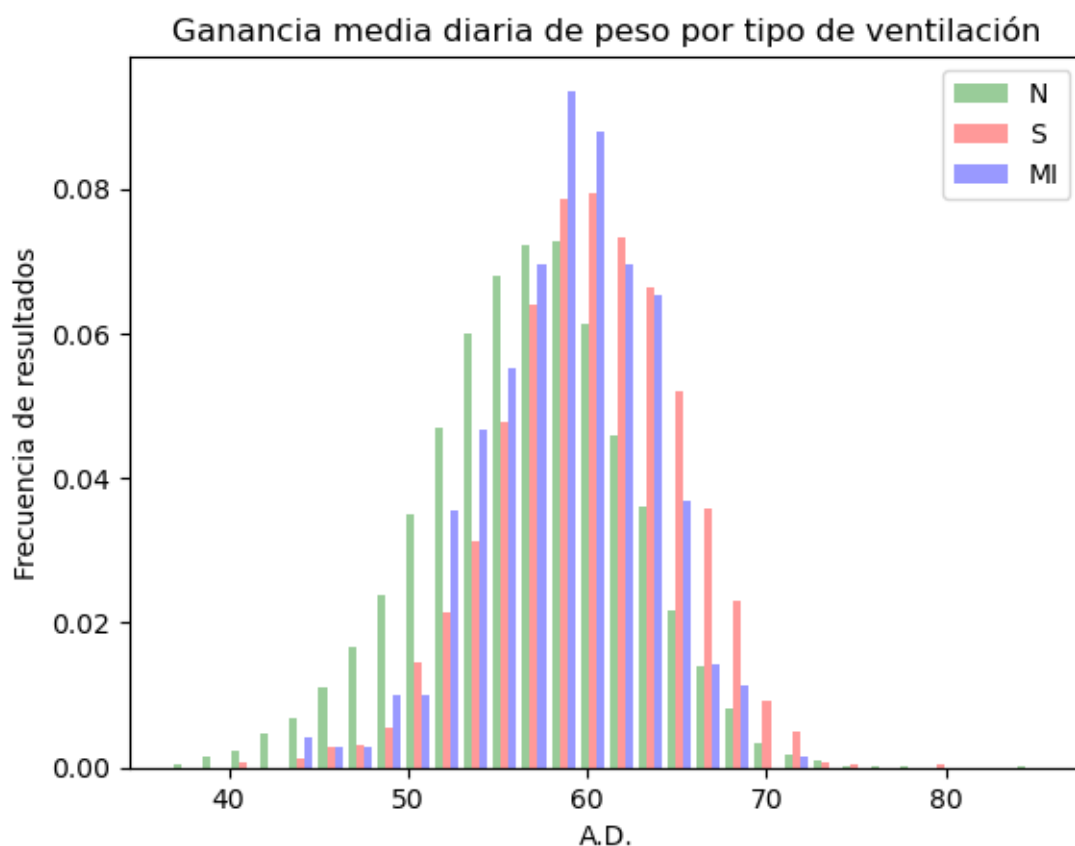
Para ello se han clasificado las granjas en tres categorías:

S: Ventilación forzada: La ventilación del galpón se realiza a través de la entrada del aire por inlets o aberturas en la zona superior de los galpones y su posterior expulsión por extractores.

N: Granjas con sistema de ventilación convencional con encendido manual de ventiladores y/o apertura y cierre de las cortinas laterales del galpón.

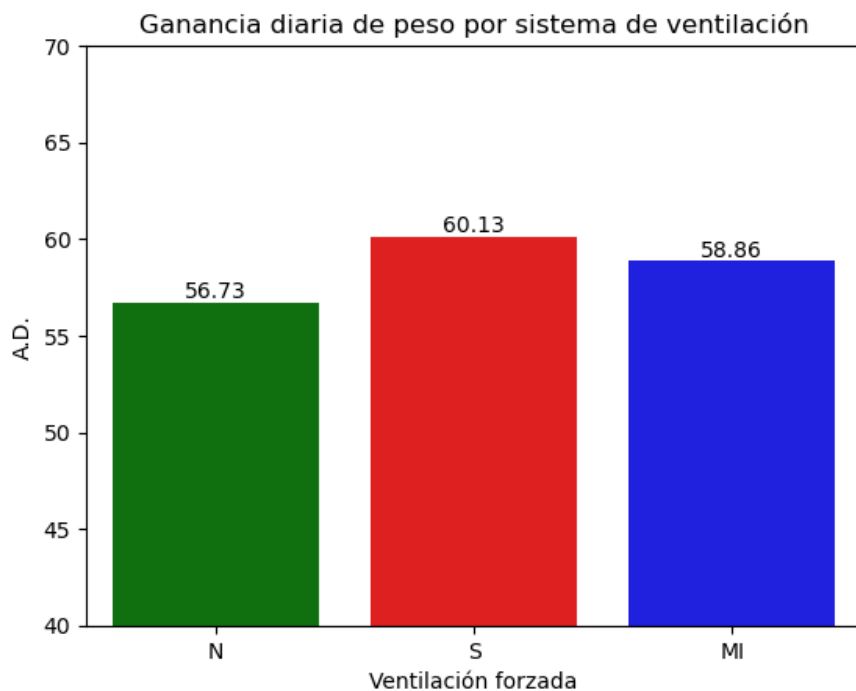
MI: Mixtas: Granjas que tienen ambos tipos de galpones.

Histograma A.D. por tipo de ventilación



En el gráfico se observa que los valores de ganancia de peso de los tres sistemas tienen una distribución normal y que el sistema convencional logra menores ganancias de peso que los sistemas de ventilación forzada y mixto. La distribución de los datos de las granjas convencionales se encuentra desplazada hacia la izquierda, es decir hacia valores menores de ganancia. Su promedio se ubica alrededor de los 56 gramos/días mientras que el promedio de las granjas de ventilación forzada está alrededor de los 60 gramos/día.

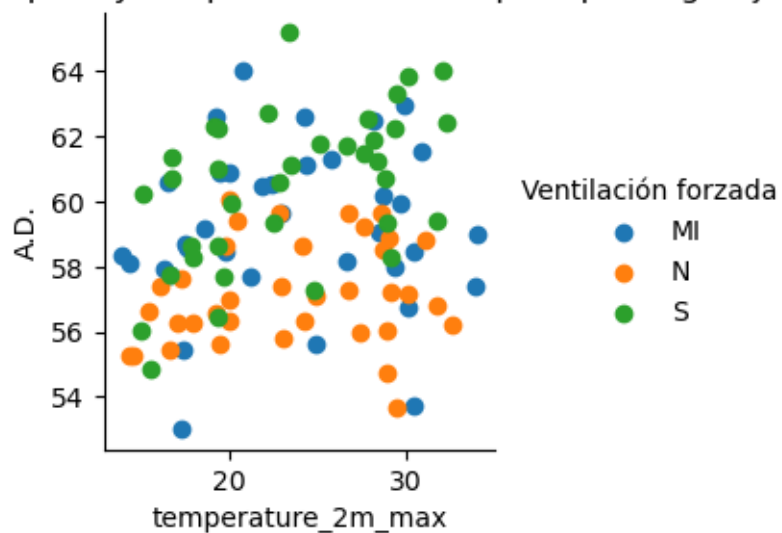
A.D. por sistema de ventilación



Habría indicios de que el sistema de ventilación tiene impacto sobre la ganancia de peso de los animales.

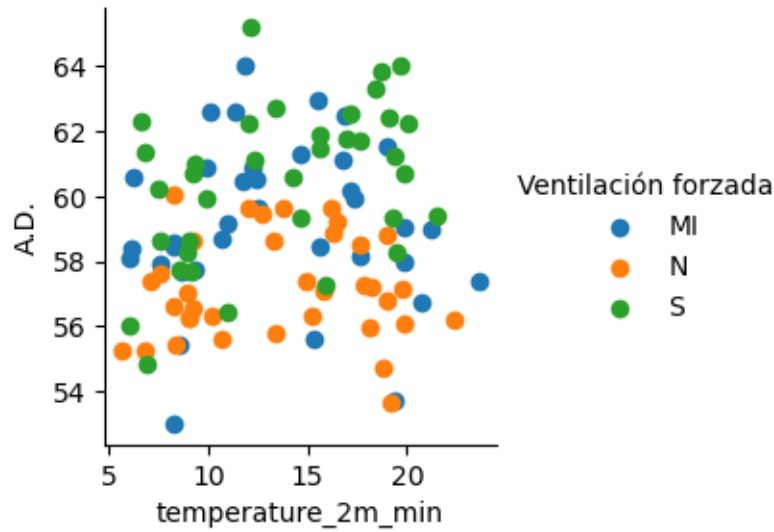
A.D. y temperatura máxima

Ganancia de peso y temperatura máxima por tipo de granja



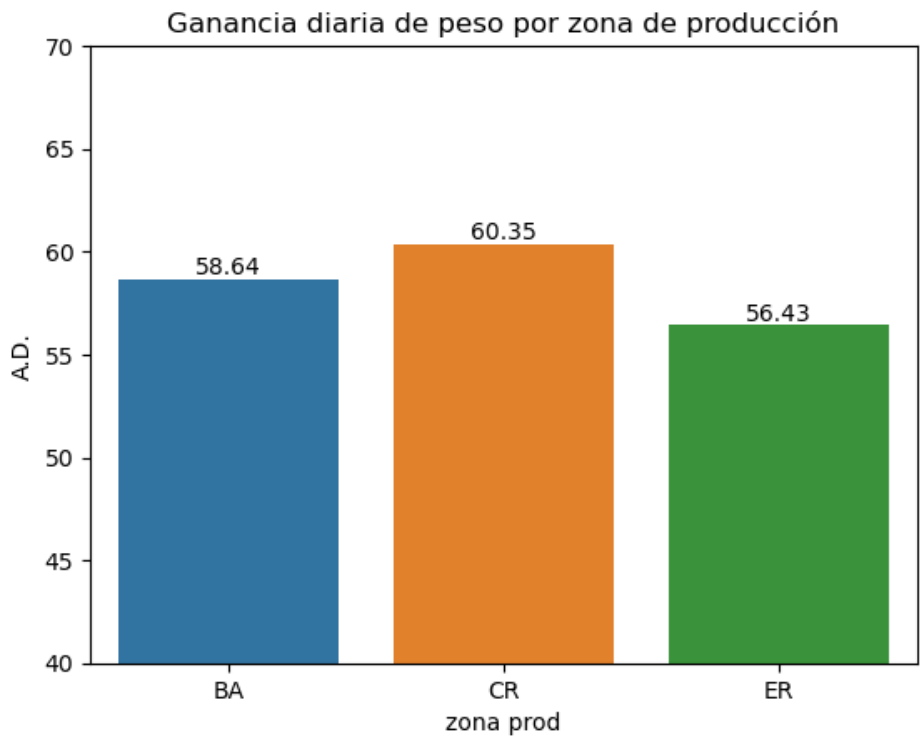
A.D. y temperatura mínima

Ganancia de peso y temperatura mínima por tipo de granja



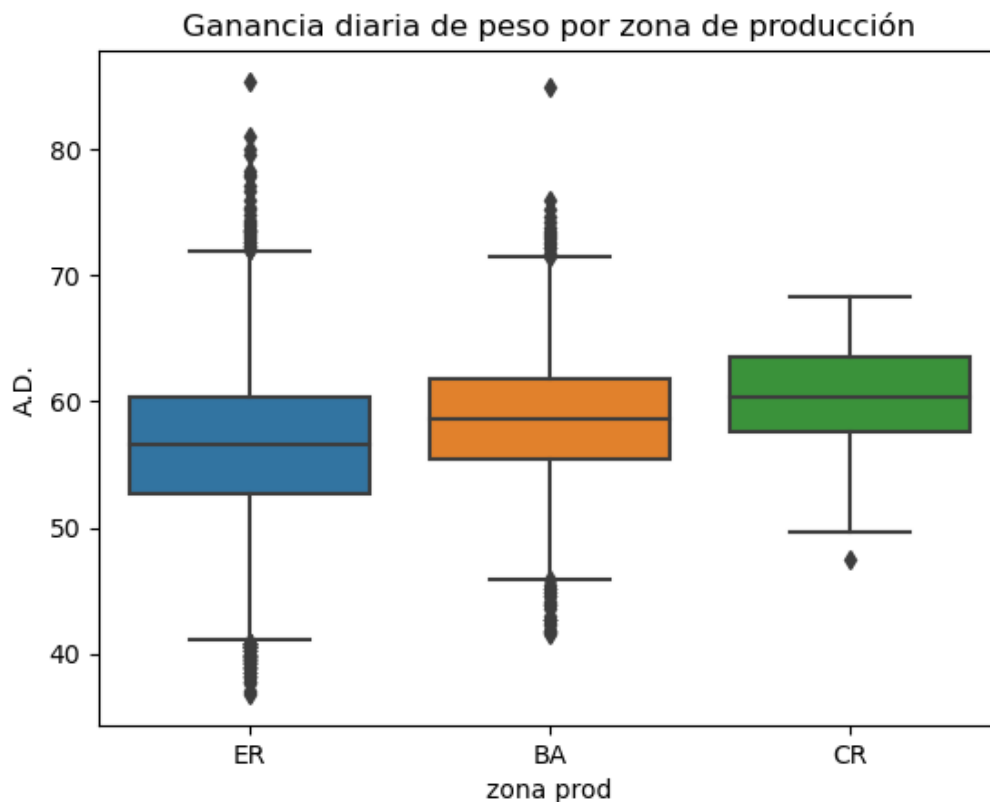
Los gráficos de correlación entre la ganancia de peso y las temperaturas máximas y mínimas diarias no muestran ninguna relación entre estas variables y tampoco con relación al sistema de ventilación.

A.D. por zona de producción



A nivel de promedios, hay diferencias entre las zonas del país donde se encuentran ubicadas las granjas.

A.D. por zona de producción dispersión



Los mejores resultados se encuentran en la zona de Crespo (CR), no solo porque tiene mayores valores de ganancia de peso sino también porque los resultados tienen menor dispersión y prácticamente no posee resultados atípicos.

Machine learning aproximación

Se seleccionó la variable ganancia media diaria de peso como valor a predecir en función de los valores de:

- Año de levante
- Mes de levante
- Estación del año
- Ventilación forzada
- Zona de producción

Teniendo en cuenta el posible impacto del clima (representado por las tres primeras variables), del sistema de ventilación de los galpones y de la zona geográfica en la que se encuentran las granjas.

Se dividió el conjunto de datos en subconjuntos de entrenamiento y prueba y se aplicó un modelo SVR y sus correspondientes pruebas.

Modelo SVR

Se presentan algunos resultados de la prueba del modelo como ejemplo:

| | Actual | Predicho | Sesgo | Error_porc |
|-------|--------|-----------|-----------|------------|
| 11404 | 54.89 | 54.999957 | -0.109957 | -0.200322 |
| 4819 | 55.00 | 61.720292 | -6.720292 | -12.218713 |
| 9353 | 59.25 | 54.289041 | 4.960959 | 8.372926 |
| 14717 | 55.76 | 61.080591 | -5.320591 | -9.541950 |
| 3948 | 57.08 | 52.641326 | 4.438674 | 7.776234 |

El modelo aplicado arrojó un valor promedio de ganancia de peso predicho de 57.29 frente a un valor promedio real de 57.17. El sesgo promedio entre los valores fue de -0.12 en valor absoluto y -0.98 en porcentaje.

Conclusiones

El objetivo del presente trabajo fue detectar el impacto de los factores que afectan al crecimiento de los pollos parrilleros para tomar decisiones que mejoren los resultados productivos, se observó que los resultados varían en las distintas épocas del año, así como en cada uno de los meses. Las diferencias entre los resultados anuales también son un indicador del impacto que tienen las condiciones ambientales sobre la ganancia de peso. Otros factores que influyen en el resultado son el sistema de ventilación de los galpones y la ubicación geográfica de las granjas (zonas de producción). Se desarrolló un modelo que permitirá estimar valores de ganancia media diaria de peso en función de valores conocidos de año de levante, mes de levante, estación del año, sistema de ventilación y zona de producción.

Insights

- El año 2021 fue el que tuvo mejores resultados de ganancia de peso promedio.
- En todos los años los meses de la primavera son los que tienen los mayores valores de ganancia de peso.
- Las granjas con sistema de ventilación forzada presentan mejores resultados productivos y menor dispersión que las granjas convencionales.
- Las granjas de la zona de producción de Crespo presentan mejores resultados productivos y menor dispersión que las granjas de las otras zonas.
- Los valores de temperatura máxima y mínima diarios tienen una baja correlación con la ganancia diaria de peso.

Recomendaciones

- Incrementar la cantidad de granjas en la zona de Crespo (CR).
- Si es posible, dejar de producir en las granjas de peores resultados de la zona de Entre Ríos (ER).
- Buscar acciones para mejorar el ambiente interno de los galpones para obtener mejores resultados, haciendo hincapié en las granjas con sistema de ventilación convencional.
- Se recomienda tomar medidas y adoptar nuevas estrategias de manejo de ambiente durante los meses de verano, por ejemplo, agregar más ventiladores en los galpones o realizar capacitaciones al personal de granja para el manejo de las cortinas en esta época del año.