

## Ejercicio 1

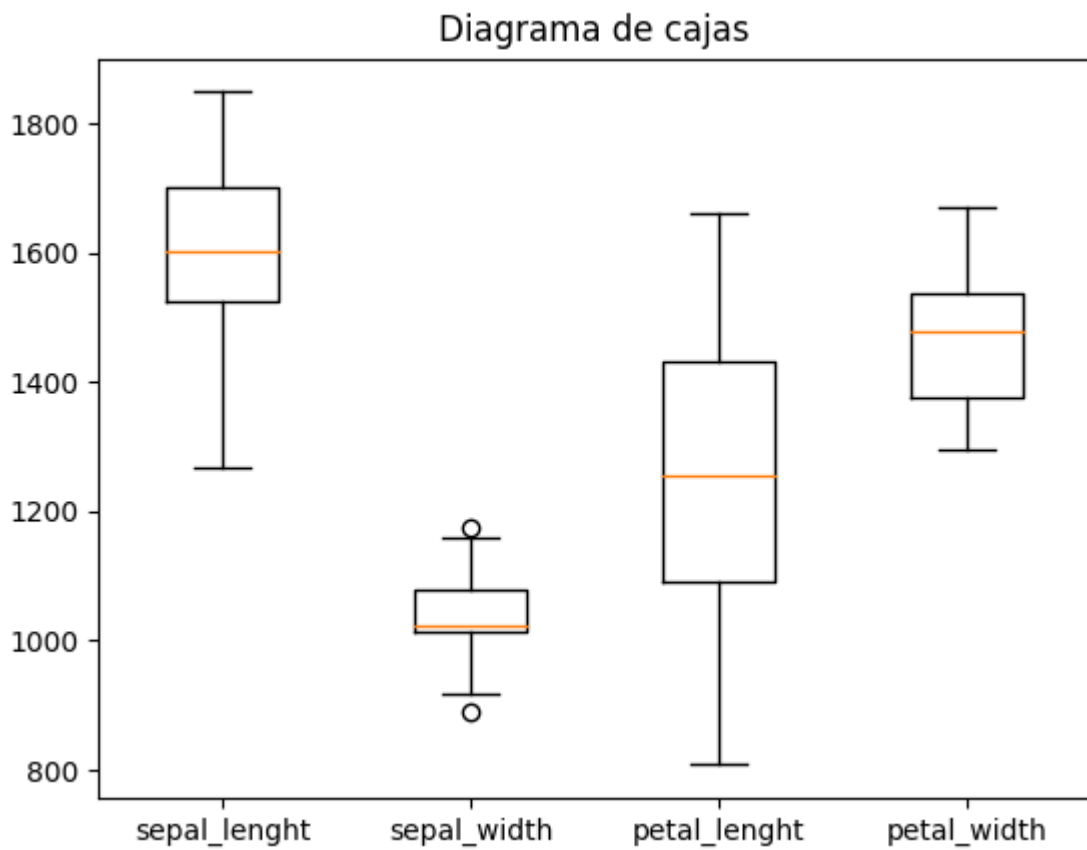


Figure 1: Diagrama de cajas de los resultados obtenidos de las 30 corridas

## Ejercicio 2

La ciencia de datos va más allá de las técnicas y algoritmos, requiriendo habilidades multidisciplinarias para traducir entre la tecnología y las preocupaciones comerciales. Por lo que hay un proceso para poder lograrlo, a continuación veremos estos pasos:

### ⇒ **Encuadre del problema.**

Lo más importante es definir de manera precisa un problema antes de resolverlo. Como científico de datos, se menciona que es común recibir aportaciones ambiguas, por lo que es necesario trabajar con los solicitantes y formular las preguntas adecuadas para comprender el problema en profundidad.

Se presentan una serie de preguntas clave que se deben plantear al inicio del proceso, con el objetivo de identificar a los clientes, comprender el proceso de venta actual, recopilar información sobre los clientes potenciales y conocer los niveles de servicio existentes. Además se enfatiza la importancia de adentrarse en la perspectiva del cliente, en este caso el vicepresidente de ventas, para comprender su visión del problema. Este conocimiento será invaluable para el análisis de datos y la presentación de puntos de vista.

Finalmente, es necesario realizar preguntas más precisas una vez que se tenga un conocimiento razonable del dominio, con el objetivo de comprender exactamente qué problema desea resolver el cliente. Se sugieren dos preguntas específicas relacionadas con la optimización del embudo de ventas y las áreas que el vicepresidente considera no optimizadas en el presente.

### ⇒ **Recopilar los datos brutos**

En este paso se resalta la importancia de considerar los datos necesarios en un proyecto de ciencia de datos. Se plantean varias preguntas clave relacionadas con los datos de la base de datos CRM, como cuáles son necesarios, cómo extraerlos y en qué formato deben almacenarse para facilitar su análisis.

Es importante ser un científico de datos ético, con un enfoque tanto en la seguridad como en la privacidad de los datos. Además de tener cuidado de no extraer información de identificación personal de la base de datos y asegurar que toda la información en el archivo sea anónima y no se pueda rastrear hasta clientes específicos. La mayoría de los proyectos de ciencia de datos se utilizan datos existentes que se están recopilando en tiempo real. Sin embargo, ocasionalmente puede ser necesario liderar esfuerzos para recopilar nuevos datos, lo cual implica trabajo de ingeniería y tiempo para obtener resultados satisfactorios.

### ⇒ **Procesar y explorar los datos**

Se revisan los datos extraídos para comprender el significado de cada columna y se verifica la presencia de valores faltantes, intervalos adecuados y el formato correcto. Se decide qué hacer con los valores perdidos o corruptos, ya sea eliminarlos o utilizar valores predeterminados.

Una vez completada la limpieza de datos, se inicia la exploración para obtener conclusiones. Se busca descubrir qué información contienen los datos y qué partes son relevantes para responder a las preguntas planteadas. Se pueden examinar patrones en los datos dividiéndolos en grupos según si los clientes se han convertido o no.

### ⇒ **Realizar análisis**

Después de recopilar suficiente información, es momento de crear un modelo que responda a la pregunta inicial. Para ello, se utilizan técnicas de Aprendizaje Automático (Machine Learning) que se basan en conjuntos de datos y vectores de características. Para desarrollar el modelo, se elige la Regresión Logística, un algoritmo simple de clasificación dentro del Aprendizaje Automático Supervisado. Este algoritmo aprende a partir de ejemplos etiquetados y proporciona una predicción binaria y una probabilidad de conversión.

Después de aplicar la Regresión Logística y ajustar los parámetros, se obtienen resultados emocionantes con una precisión del 95%. Sin embargo, es importante comunicar estos resultados de manera convincente y comprensible para el cliente.

### ⇒ **Comunicar los resultados**

En este paso, se aborda la importancia de comunicar los resultados del modelo de Machine Learning de manera efectiva al cliente. La habilidad de comunicación es fundamental en el campo de la ciencia de datos, ya que implica traducir el trabajo realizado en información comprensible y práctica. Es decir, es la necesidad de contar una historia basada en los resultados obtenidos del análisis exploratorio y del modelo predictivo. La historia debe incluir conclusiones relevantes que respondan a las preguntas más importantes para el cliente.

En la siguiente etapa, se reúnen con el cliente y se explican las conclusiones. El cliente se muestra entusiasmado con los resultados y plantea la pregunta sobre cómo pueden utilizar mejor estos hallazgos. Como científico de datos, se espera que no solo se analicen los datos, sino que también se brinden recomendaciones sobre cómo utilizar los resultados.

## Ejercicio 3

Al hablar sobre el surgimiento de la “nigromancia digital” es importante reconocer que, además del debate sobre la moralidad de utilizar el rostro de personas fallecidas o vivas sin su consentimiento, este fenómeno plantea también preocupaciones éticas más amplias, cómo se puede percibir en todo lo relacionado con la tecnología de los Deepfakes. Los Deepfakes comprenden las tecnologías que generan imágenes o videos por medio de el uso de redes neuronales conocidas como Autoencoders. Estas redes neuronales, se adiestran utilizando grandes conjuntos de datos de rostros, buscando aprender los patrones subyacentes que codifican; y decodificando las características faciales. El Autoencoder, al recibir una imagen de entrada, la comprime en un cuello de botella donde la información se reduce a una representación compacta, para que luego, la red intente reconstruir la imagen original a partir de esta representación comprimida.

Una de las herramientas más populares, sino es que la más utilizada en lo concerniente a la nigromancia digital, es DeepFaceLab, un software que aprovecha las capacidades de los Autoencoders para generar sustituciones faciales realistas. Esta herramienta permite ejercitar redes neuronales con una amplia variedad de datos, incluyendo fotografías y videos de personas fallecidas, para luego aplicar esta información a nuevas imágenes o secuencias de video, teniendo como resultado lo que bien podría parecer la capacidad de “revivir” digitalmente a alguien, mostrándolo con expresiones o en situaciones en las que nunca estuvo presente.

No obstante, si bien el desarrollo de tecnologías de nigromancia digital representa un avance importante en cuestión de inteligencia artificial, estas tecnologías también han generado suficiente alerta entre la comunidad técnica, ya que existen desafíos en la creación de Deepfakes de alta calidad y que arrojen resultados realmente convincentes y, a su vez, que sean difíciles de identificar, ya que, en la medida en que los algoritmos mejoran y se tornan más sofisticados, se necesita también, un enfoque paralelo en el desarrollo de métodos de detección de Deepfakes igualmente avanzados.

Los desarrolladores e investigadores en este campo están trabajando en nuevas técnicas de detección, más avanzadas, que buscan utilizar análisis de patrones y aprendizaje automático para identificar inconsistencias sutiles en los videos e imágenes generados por Deepfakes. Al analizar características faciales, movimientos oculares y parpadeos, y otras señales naturalmente sutiles, se intentan identificar y desarrollar algoritmos que puedan distinguir entre el contenido auténtico y el que es producto de la manipulación digital.

El “traer de vuelta a la vida” a personas que ya no están presentes, incluso de manera virtual, sugiere la apertura de un debate sobre la moralidad, la integridad y el uso responsable de su imagen. En primer lugar, la posibilidad de “revivir” digitalmente a personas fallecidas plantea, además, fuertes interrogantes éticos fundamentales que obedecen a la cuestión central de si utilizar los rostros de los difuntos en contenidos audiovisuales, es una forma respetuosa de honrar su memoria o si, por el contrario, es una violación de su privacidad y dignidad e incluso un atentado hacia la propia dignidad de los seres queridos. Ahora bien, es necesario también considerar otro aspecto ético que cuestiona la moralidad del uso de las Deepfakes para engañar, manipular o influenciar negativamente a terceras personas.

Si bien al ser capaces de generar contenido audiovisual falso que ha sido editado hasta llegar a ser altamente convincente, estas tecnologías pueden prestar las condiciones óptimas para que se lleven a cabo prácticas o situaciones de extorsión, difamación o engaño. Partiendo desde el entendido de que este contenido audiovisual posee la capacidad de mostrar videos o imágenes en los que las personas fallecidas parecen estar diciendo o haciendo cosas que nunca antes han hecho, podría también, ser explotado con fines maliciosos, trayendo consigo consecuencias devastadoras para la reputación y la vida personal de los individuos afectados, finados o no.

Lo anterior plantea interrogantes considerables sobre la responsabilidad y el uso adecuado de la tecnología, ya que se corre el riesgo de socavar la confianza en la información, al mismo tiempo que el de aumentar la desinformación en un mundo ya de por si afectado por la información y noticias falsas.

Por si esto fuera poco, y en un ámbito más humano, existe también la preocupación de que este tipo de prácticas, que normalizan la “nigromancia digital”, puedan generar emociones perturbadoras o sumamente negativas en los conocidos y seres queridos de los fallecidos: el disponer de material audiovisual de un ser querido fallecido, especialmente si éstas le muestran en situaciones inventadas o fuera de contexto, puede ser un catalizador emocionalmente traumático que llegue a causar desconcierto, angustia, daño moral e incluso hasta depresión a las personas que están de luto. Es en ese sentido que la falta de consentimiento previo por parte de los difuntos agrava aún más este dilema ético, ya que se toma una decisión sobre su imagen sin su aprobación e incluso muchas veces sin siquiera la aprobación de los afectados más cercanos al fallecido.

En ese sentido, el uso o aparición de los rostros de personas vivas en contenido digital no autorizado y los debates que han surgido a razón de dichos cuestionamientos, han intentado arrojar algunas medidas tanto técnicas como legales en aras de reglamentar su uso y distribución. En primer lugar, se requiere de una mayor investigación y desarrollo de métodos efectivos para la detección de

Deepfakes que permitan identificar y prevenir su difusión, de esta manera, se buscaría ayudar a mitigar los riesgos asociados con la manipulación no autorizada de la imagen y la desinformación.

En resumen, la "nigromancia digital" plantea un debate ético complejo referente al uso de los rostros de personas fallecidas y vivas en contenidos audiovisuales generados por la inteligencia artificial o Deepfakes. La privacidad, la dignidad y la manipulación son cuestiones éticas clave que deben abordarse a través de medidas técnicas, legales y educativas que se busquen implementar en el marco regulatorio de estas nuevas tecnologías y será, solamente mediante un enfoque equilibrado y ético, que la comunidad involucrada podrá aprovechar al máximo, el uso y las capacidades de la tecnología de los Deepfakes de una manera ética, responsable y respetuosa.