

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**



**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS**

**DEPARTAMENTO DE POSTGRADO**

**PROYECTO DE TITULACIÓN**

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

**“MAGÍSTER EN ESTADÍSTICA CON MENCIÓN EN CALIDAD Y  
PRODUCTIVIDAD”**

**TEMA:**

IMPLEMENTACIÓN DE UN CUADRO DE MANDO INTEGRAL PARA  
UNA EMPRESA DE CARGUE Y DESCARGUE DE MERCADERÍA.

**AUTOR:**

OLVERA ORELLANA ANDREA PAMELA

Guayaquil - Ecuador

2022

## **RESUMEN**

A medida que los mercados crecen y los servicios evolucionan, las empresas se ven cada vez más necesitadas de generar una visión clara que les permita aportar valor a sus clientes, pero a la vez rentabilizar su operación. En este trabajo se encuentran condensados los pasos seguidos por el operador logístico líder del país para el potenciamiento de los resultados en su empresa de cargue y descargue de mercadería, TORRESTIBAS S.A. El trabajo considera la incorporación de la nueva visión de negocio para convertirse en un referente de excelencia en servicio y asesoría logística integral, desde la identificación de la mejor forma de crecer y obtener un rendimiento superior basado en el planteamiento realizado desde su Plan Estratégico con metodologías PMS (Performance Management System) hasta la incorporación de nueva tecnología para lograr plasmar en cifras reales los objetivos principales de la compañía, esto para beneficio de la toma de decisiones de sus directivos y logro de la mencionada nueva visión empresarial.

En una etapa previa de Alineamiento Estratégico la organización evaluó su contexto empresarial, expectativas y requisitos de las consideradas como sus partes interesadas (stakeholders) y las acciones identificadas como claves a gestionar. Además de las definiciones estratégicas que la dirección ha determinado como marco de la estrategia empresarial. Esto se conjuga en la

ejecución de los pilares del PMS en donde se busca que la brecha de alineamiento entre la estrategia (dirección) y la ejecución (operaciones) sea mínima en función de lograr los objetivos planteados.

Como parte de esta implementación también se genera la identificación de los factores claves de éxito que deben ser evaluados con la consecución de OKRs y KPIs, estos indicadores serán planteados desde el entendimiento de las variables críticas del negocio evaluadas a través de la analítica de datos, con un enfoque en el análisis exploratorio en aquellos niveles operativos que aplique para identificar patrones y relaciones, así como la aplicación del análisis confirmatorio a través de la inferencia estadística, con modelos predictivos que aporten a la planificación de las operaciones y recursos, además del uso e incorporación de herramientas de Business Intelligent como Power Bi que aporta para dinamizar la toma de acciones, poniendo al alcance de la mayor cantidad de directivos y puestos claves dentro de la compañía las cifras más relevantes para su accionar.

Todo esto en el marco de conseguir la generación de un modelo exitoso que permita la capitalización de este conocimiento en la experiencia de la implementación de sus recursos y gestionar un modelo que sea replicable en el resto de las empresas del grupo al que pertenece Torrestibas.

## **ABSTRACT**

As markets grow and services evolve, companies are increasingly in need of generating a clear vision that allows them to add value to their customers, but at the same time make their operation profitable. In this work, the steps followed by the country's leading logistics operator to enhance the results in its cargo loading and unloading company, TORRESTIBAS SA are condensed. The work considers the incorporation of the new business vision to become a benchmark of excellence in service and comprehensive logistics consulting. The scopes includes the identification of the best way to grow and obtain superior performance based on the approach made from its Strategic Plan with PMS (Performance Management System) methodologies along with the incorporation of new technology to achieve translate into real figures the main objectives of the company. All this in order to obtain benefit of the decision-making of its managers and achievement of the aforementioned new business vision.

In a previous stage of Strategic Alignment, the organization evaluated its business context, expectations and requirements of those considered as its interested parties (stakeholders) and the actions identified as key to manage. In addition to the strategic definitions that the management has determined as a framework for the business strategy. This is combined in the execution of the PMS pillars where it is sought that the alignment gap between the strategy

(direction) and the execution (operations) is minimal in terms of achieving the objectives set.

As part of this implementation, the identification of the key success factors that must be evaluated with the achievement of OKRs and KPIs is also generated, these indicators will be raised from the understanding of the critical variables of the business evaluated through data analytics, with a focus on exploratory analysis at those operational levels that apply to identify patterns and relationships, as well as the application of confirmatory analysis through statistical inference, with predictive models that help planning of operations and resources, as well as the use and incorporation of Business Intelligent tools such as Power Bi, which contributes to stimulate the taking of actions, making the most relevant figures for their actions available to the greatest number of executives and key positions within the company.

All this within the framework of generating a successful model that allows the capitalization of this knowledge in the experience of implementing its resources and managing a model that is replicable in the other companies of the group to which Torrestibas belongs.

## **DEDICATORIA**

Dedicado a los que trabajan para hacer que en este mundo cosas grandiosas pasen, a mis hermanos, un recordatorio de lo que podemos lograr si no descansamos por lo que soñamos.

## **AGRADECIMIENTO**

A mis padres, Patricia y Leonardo que no pintaron barreras en mi mapa mental. A mi equipo de refuerzo, Yahir y los amigos que han sido un apoyo en este camino.

## **DECLARACIÓN EXPRESA**

La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación, me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo, corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.



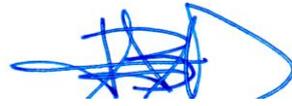
Ing. Andrea Olivera Orellana

# TRIBUNAL DE GRADUACIÓN



---

Francisco Vera Alcívar, Ph.D.  
PRESIDENTE



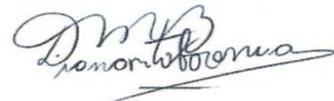
---

Jorge Fernando Abad Morán, PhD  
DIRECTOR



---

María Nela Pastuzaca Fernández, Ph.D.  
VOCAL 1



---

Mgtr. Diana Montalvo Barrera  
VOCAL 2

## TABLA DE CONTENIDO

RESUMEN.....	i
ABSTRACT.....	iii
DEDICATORIA .....	v
AGRADECIMIENTO .....	vi
DECLARACIÓN EXPRESA .....	vii
TRIBUNAL DE GRADUACIÓN .....	viii
TABLA DE CONTENIDO.....	ix
LISTADO DE FIGURAS .....	xi
INTRODUCCIÓN .....	xii
CAPÍTULO 1: Antecedentes .....	1
1.1 Descripción del problema.....	1
1.2 Objetivo General:.....	3
1.3 Objetivos Específicos:.....	3
1.4 Alcance.....	3
CAPÍTULO 2: Marco teórico .....	5
2.1 Metodología.....	5
2.2 Fase de Gestión Estratégica.....	7
2.3 Datos e información disponible para la investigación.....	8
2.4 Información del negocio y sus procesos a evaluar. ....	9
CAPÍTULO 3: Conocimiento del Negocio.....	11
3.1 Direccionamiento Estratégico 2022 – 2025. ....	11
CAPÍTULO 4: Desarrollo y Elaboración del Balance Scorecard.....	14
4.1 Evaluación de oportunidad de Venta cruzada. ....	20
CAPÍTULO 5: Desarrollo y resultados del modelo.....	22
5.1 Limpieza y estructuración de la información. ....	22
5.2 Estructura de la información de Ciateite – Estibas.....	23
5.3 Limpieza y estructuración Tabla principal Estibas. ....	25
5.4 Limpieza y estructuración Tabla CLIENTE_OPERACIONES.....	26
5.5 Creación de nuestra master table.....	27
5.6 Creación de las variables objetivo.....	27
5.7 Creación de variable objetivo de Estibas.....	28

5.8	Selección de Técnica Machine Learning.....	29
5.9	Entrenamiento del modelo de contratación de servicio de Estibas. ....	30
5.10	Validación de los resultados. ....	33
5.11	Resultados.....	36
CAPÍTULO 6: Conclusiones y Recomendaciones. ....		38

## LISTADO DE FIGURAS

Figura 2.1 5 The Knowledge discovery in databases process.....	9
Figura 2.2 Dimensiones claves para la exitosa implementación de estrategia en un negocio. ....	10
Figura 3.1 Pilares Estratégicos y de negocio para gestionar la estrategia del Grupo T&T.....	11
Figura 3.2 Atributos de la Propuesta de Valor.....	13
Figura 4.1 Ejemplo mapa estratégico con líneas estratégica y relación causa - efecto. ....	15
Figura 4.2 Mapa estratégico Torrestibas 2022.....	16
Figura 4.3 Cuadro de Mando Torrestibas 2022. ....	18
Figura 4.4 Matriz de Metas por Pilares Estratégicos y UNES: Crecimiento Rentable. ....	19
Figura 4.1 .....	24
Figura 4.2 Variables de facturación a considerar en la conformación del modelo de aprendizaje. ....	25
Figura 4.3 .....	25
Figura 4.4 Variables operativas a considerar en la conformación del modelo de aprendizaje. ....	26
Figura 4.5 .....	26
Figura 4.6 Variables Cliente – Operaciones a considerar en la conformación del modelo de aprendizaje.....	27
Figura 4.7 .....	27
Figura 4.8 .....	29
Figura 4.9 .....	31
Figura 4.10 .....	32
Figura 5.1 .....	36

## **INTRODUCCIÓN**

En este nuevo entorno creado a raíz de la crisis sanitaria por COVID – 19, los problemas de globalización que aplican a la logística y la lucha de los clientes por disminuir sus costos operativos, se crea un entorno en donde es emergente la priorización de las empresas por rentabilizar su operación pero que a su vez le permita incluir formas de identificar oportunidades en donde pueda ofertar nuevos servicios a sus actuales clientes o generar información que le permita fidelizarlo o cuidar su experiencia a través de un servicio de excelencia. Este es también el reto planteado para Torrestibas que es una empresa dedicada al manejo de Cargue y descargue de mercaderías, así como del acondicionamiento de las mismas y trabajos internos de bodegas, siendo un entorno distinto para cada una de sus sucursales en Quito y Guayaquil

En este trabajo se busca mostrar los resultados de una correcta implementación de las mediciones de desempeño en la compañía frente al negocio y los empleados o conocido en el mundo de la gestión como el Performance utilizando la metodología PMS®, con lo que se busca conectar la estrategia organizacional con lo que hacen las personas en su día a día para traducirlo en resultados tangibles y reportables en modo de OKRs (Objectives and Key Results u objetivos y resultados clave) y KPIs (Key Performance

Indicators o Indicadores Claves de Desempeño), estos indicadores serán planteados desde el entendimiento de las variables críticas del negocio evaluadas a través de la analítica de datos, con un enfoque en el análisis exploratorio en aquellos niveles operativos que aplique para identificar patrones y relaciones, así como la aplicación del análisis confirmatorio a través de la inferencia estadística, con modelos predictivos que aporten a la planificación de las operaciones y recursos.

# **CAPÍTULO 1: Antecedentes**

## **1.1 Descripción del problema**

La creciente competencia entre compañías latinoamericanas, tanto a nivel local como internacional, lleva a un mayor interés en el establecimiento de procesos tendientes a revisar sus estrategias de negocios (Govindaran y Gupta, 2001).

El Grupo Torres y Torres cuenta con más de 30 años en el mercado, con 4 empresas alineadas a la logística se ha convertido en referente en el manejo de carga; comercio exterior y transporte de carga pesada. Dentro de su cadena logística se encuentra TORRESTIBAS S.A., empresa dedicada al manejo de Cargue y descargue de mercaderías, así como del acondicionamiento de las mismas y trabajo internos de bodegas, esta es la última de las 4 empresas que integra su cadena logística, cuenta actualmente con un Administrador de Estibas encargado de dos sucursales Quito y Guayaquil, quien ha visto como necesario priorizar el manejo integral de la información de la compañía, además se suma la necesidad de reevaluación de la estrategia y metas de la compañía para dar concordancia al nuevo contexto creado en este 2020 y 2021.

La empresa ha trabajado con el acompañamiento de una herramienta de asignación de personal que ayuda a distribuir la carga de trabajo y gestiones del día a día, sin embargo, no maneja indicadores de apoyo a la estrategia,

tácticos ni operativos, su información se ha utilizado principalmente para evaluar períodos de información muy anteriores y sólo para el reporte de las cifras primarias que la compañía reporta a su directorio. Se ha evaluado la expectativa gerencial acerca de esta necesidad de implementación y uno de los puntos más relevantes es la disponibilidad y organización de la información, es decir, la herramienta a implementarse debe cumplir con lograr que la información y cifras relevantes del negocio se encuentren siempre actualizadas y disponibles, además de que la organización de la misma sea de ayuda y se muestre con claridad para el manejo del usuario con el fin de incrementar la eficiencia de los procesos y mejorar las cifras críticas del negocio. Por estos motivos la herramienta seleccionada para la implementación es Power BI.

En el planteamiento de este trabajo además se ha definido el reto de lograr usar los datos de la compañía para obtener no sólo descriptivos, sino que además con un enfoque en el análisis exploratorio en aquellos niveles operativos que aplique para identificar patrones y relaciones, así como la aplicación del análisis confirmatorio a través de la inferencia estadística, con modelos predictivos que aporten a la planificación de las operaciones y recursos u colocación de ventas cruzadas entre las empresas del grupo.

## **1.2 Objetivo General:**

Construir un Cuadro de Mando Integral en Power BI para la empresa de Carga y Descarga de mercadería, a través de la construcción de OKRs y KPIs claves para la toma de decisiones y gestión de la estrategia.

## **1.3 Objetivos Específicos:**

- a) Definir los Indicadores Claves que inducen al logro de las metas estratégicas dentro del Modelo de Gestión por Indicadores.
- b) Describir los indicadores de forma que sean medibles y aporten valor a las decisiones de la organización mediante una Matriz de Indicadores.
- c) Generar para cada Indicador Clave un tablero de gestión dinámico que permita tener la información actualizada en línea y disponible para todo el equipo de trabajo.

## **1.4 Alcance.**

Este proyecto será implementado en una empresa ecuatoriana de cargue y descargue, con operaciones a nivel nacional.

Los procesos de los cuales se utilizará la información de la gestión estratégica serán los de negocio o estratégicos, claves y de apoyo, además de la data para la generación de modelos con información operativa y los maestros de clientes del grupo logístico en general para utilizar variables registradas a nivel

general de los clientes, se usará la información recabada en sus sistemas tecnológicos actuales, tanto DOBRA operaciones como sus aplicaciones Web de Planificación y de Asignación de personal.

El período de tiempo en el que esta información ha sido recolectada es de 3 años, desde el 2018 a 2020 completos, más los meses del 2021 que han transcurrido y las consideraciones y políticas que operativamente se han considerado hasta el día de hoy (octubre 2021).

Los insumos requeridos por parte del estudio, sean de tipo decisiones o información del direccionamiento del negocio, serán proporcionados por la Gerencia General del Grupo TYT y la administración de Torrestibas a quienes se suman el directorio o junta mayor, este grupo de ejecutivos será el definidor y aprobador de lo implementado y detallado en el estudio.

## **CAPÍTULO 2: Marco teórico**

### **2.1 Metodología**

La implementación de un Cuadro de Mando comprende el uso de los Pilares del PMS (Performance Management System) que consiste en alinear la visión de negocio con la ejecución de los planes estratégicos en todos los niveles organizacionales a través de factores claves de éxito que deben ser evaluados con la consecución de OKRs y KPIs, estos indicadores han sido planteados desde el entendimiento de las variables críticas del negocio evaluadas a través de la analítica de datos, con un enfoque en el análisis exploratorio en aquellos niveles operativos que aplican para identificar patrones y relaciones, así como la aplicación del análisis confirmatorio a través de la inferencia estadística, con modelos predictivos que aporten a la planificación de las operaciones y recursos.

Estas herramientas permiten desarrollar una cultura empresarial Data Driven mediante la gobernanza de datos y DataOps para Analítica instrumentada a través de una poderosa plataforma de integración de datos que los consume y transforma en tiempo real. La simplicidad de la herramienta permite generar modelos de relaciones entre small data, datos sucios, no estructurados, o de otra manera, trabajamos con todos ellos utilizando las propiedades de end-to-end. Una vez estructurado el modelo y validado se obtendrá como resultado final la implementación de

tableros de información que cubran las necesidades para la toma de decisiones ágiles y con gobernanza.

Además, se utilizan técnicas que ayudan con la creación de modelos predictivos, utilizando la información con la que cuenta la empresa, como datos de sus clientes, transacciones, etc., en la que permite obtener un grado de propensión de estos clientes para adquirir un nuevo producto o a mejorar alguno que ya se ha comprado con anterioridad o se tiene en la cartera de la organización. (Llorente, 2018)

Las opciones más tradicionales de Cross-Selling se enfocaban en segmentar una base de consumidores basándose en unas cuantas variables, como por ejemplo se hubiera usado la definición de quien compró uno de los servicios del Grupo T&T para presumir quién compraría el de Estibas. En la práctica se sabe que no es la única variable que implica oportunidad de compra. Los datos que hoy en día almacenan las empresas son mucho más amplios y permiten entender a los clientes en muchos otros aspectos y dimensiones, así como segmentarlos o categorizarlos. Dado la cantidad de datos que las compañías tienen a disposición sería imposible o nada óptimo que una persona cree las reglas que identifiquen estos segmentos, ya que debería tener en cuenta múltiples variables y sus relaciones. Este es el campo en donde son valiosos los algoritmos de aprendizaje automático y aporta valor el uso de inteligencia artificial.

Estas técnicas permiten crear los llamados modelos predictivos que, en basados en los datos proporcionados, dan como resultado el grado de propensión que tiene el cliente para adquirir un nuevo servicio como en este caso o producto final.

Las gerencias y mandos medios de la organización, ven facilitados el acceso a la información, la toma de decisiones y planteamiento de acciones con el análisis de esta información se enfocan estrategias para la gestión de las métricas identificadas como necesarias a controlar y evaluar.

## **2.2 Fase de Gestión Estratégica.**

Se define una metodología de trabajo que permita asegurar la continuidad de lo planteado ya que además de la implementación se busca manejar los resultados y acciones que deben mantenerse actualizadas al contexto del negocio. El presente proyecto aplica PMS como parte de su aseguramiento. Lo que implica 3 fases iniciales para la descripción y aterrizaje de los planes de negocio, se detalla:

### **i. Pensamiento y Diagnóstico Estratégico.**

En esta etapa es en donde se generan las sesiones con los líderes de las UENS (Unidades Estratégicas de Negocios), que son el conjunto de procesos o actividades más pequeñas en las que se puede dividir un negocio y que usualmente se rige por una estrategia común y diferente al

resto de unidades, para crear el comité estratégico y confirmar los equipos que participarán del proceso. Seguido de una sesión de pensamiento estratégico Think Tank con expertos sobre los escenarios futuros empresariales y nuevas tecnologías. Además, es relevante ya que contempla la visión de los accionistas para establecer la dirección en el largo plazo, los grandes desafíos y metas de crecimiento, mismos que son plasmados en la matriz de crecimiento del negocio.

ii. Construcción del Plan Estratégico.

En esta etapa a través de un Bootcamp de Estrategia, las unidades de negocio construyen la Planificación Estratégica bajo la metodología Strategy Model Canvas® e identifican las sinergias del modelo de negocio bajo la metodología Business Model Canvas.

iii. Despliegue de Proyectos Estratégico.

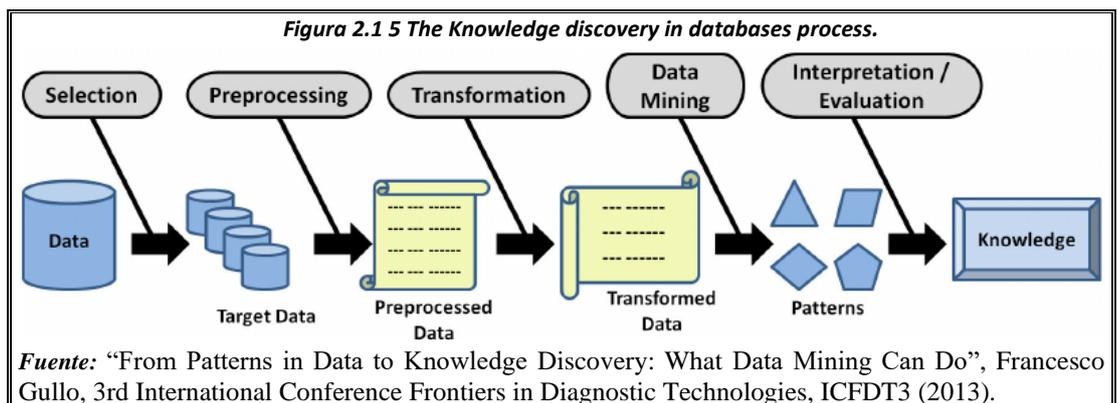
En la etapa de despliegue los talleres de Matrices de Proyectos Estratégicos contienen los despliegues siguiendo la metodología de Kaplan y Norton.

### **2.3 Datos e información disponible para la investigación.**

Para el diseño de los OKRs y KPIs además de contar con las definiciones estratégicas del negocio esta investigación toma en cuenta datos almacenados por la compañía en el lapso de alrededor de 3 años, proveniente del sistema DOBRA que es la herramienta transaccional de

operaciones, su Bandeja de Planificación de Operaciones y los datos maestros de Clientes.

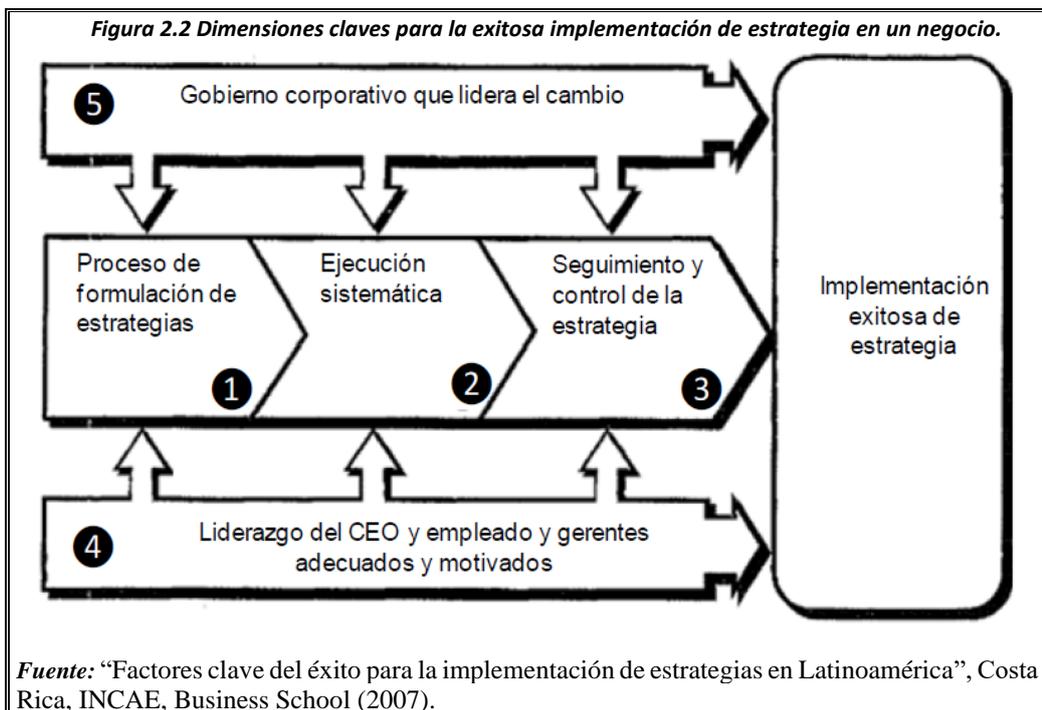
A través del proceso de Knowledge Discovery in Databases, que en español significa “Descubrimiento de Conocimiento en Bases de Datos” o conocido por sus siglas KDD (Piatetsky-Shapiro, 1989), tal como lo explica la figura 2.1, que implica un trabajo extenso en minería de datos para el tratamiento de los mismos antes de aplicar las diferentes técnicas de análisis en la búsqueda de patrones que nos lleve a brindar nueva información y conocimiento en este caso para el negocio.



## 2.4 Información del negocio y sus procesos a evaluar.

A fin de explorar los procesos de la compañía el estudio se concentra en cinco dimensiones de análisis que tienen un impacto en la implementación de la estrategia de negocios. Éstas incluyen el proceso de formulación de estrategias, la ejecución sistemática, el control y seguimiento de la implementación, el liderazgo del CEO y empleados y gerentes adecuados y motivados, y finalmente, gobierno corporativo (consejo directivo y

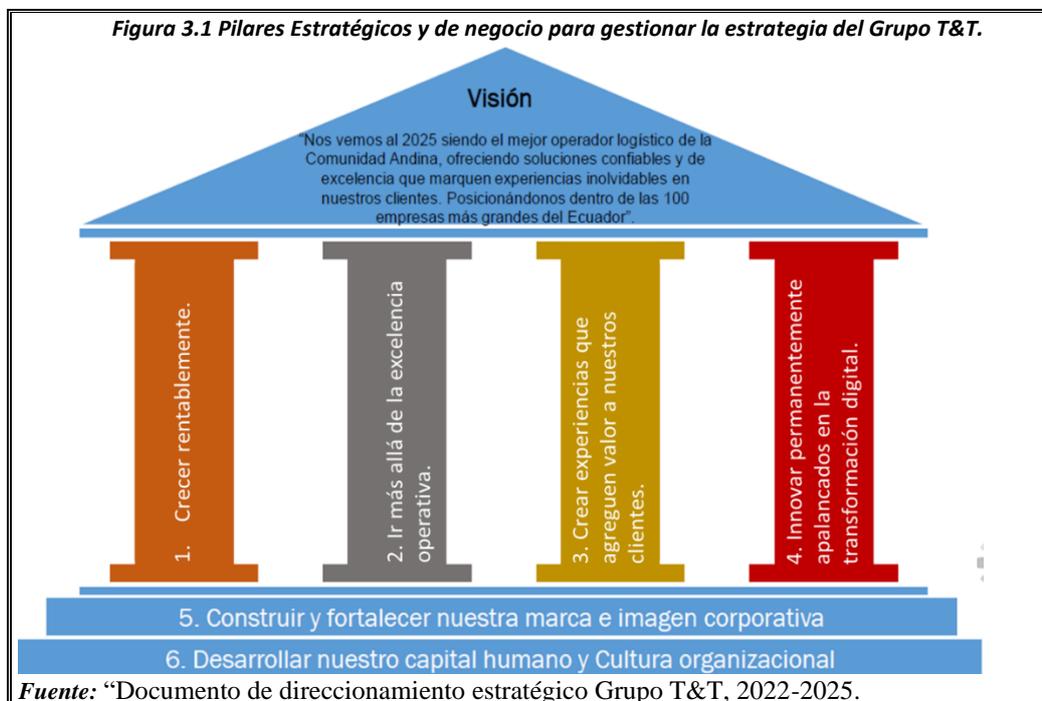
accionistas) que lideran el cambio, ver Figura 2.2. Además de la exploración preliminar de los datos del periodo base para 3 iniciativas que se identificaron en un diagnóstico elaborado en conjunto con el equipo de procesos y excelencia del grupo y la empresa consultora London Consulting.



## CAPÍTULO 3: Conocimiento del Negocio.

### 3.1 Direccionamiento Estratégico 2022 – 2025.

En el direccionamiento estratégico la formulación de las finalidades y propósitos de una organización se ven planteadas desde la construcción de la nueva visión de negocio, fase en la que la empresa expresa su sueño futuro. De esta manera se fijan además los pilares estratégicos que darán soporte y forma a la consecución de la visión planteada. En la figura 3.1 se identifican la visión del grupo empresarial en donde conecta Torrestibas S.A., y cada uno de los pilares que serán transcritos luego a metas estratégicas específicas, estas son las que dan como origen a la interpretación del Mapa Estratégico.



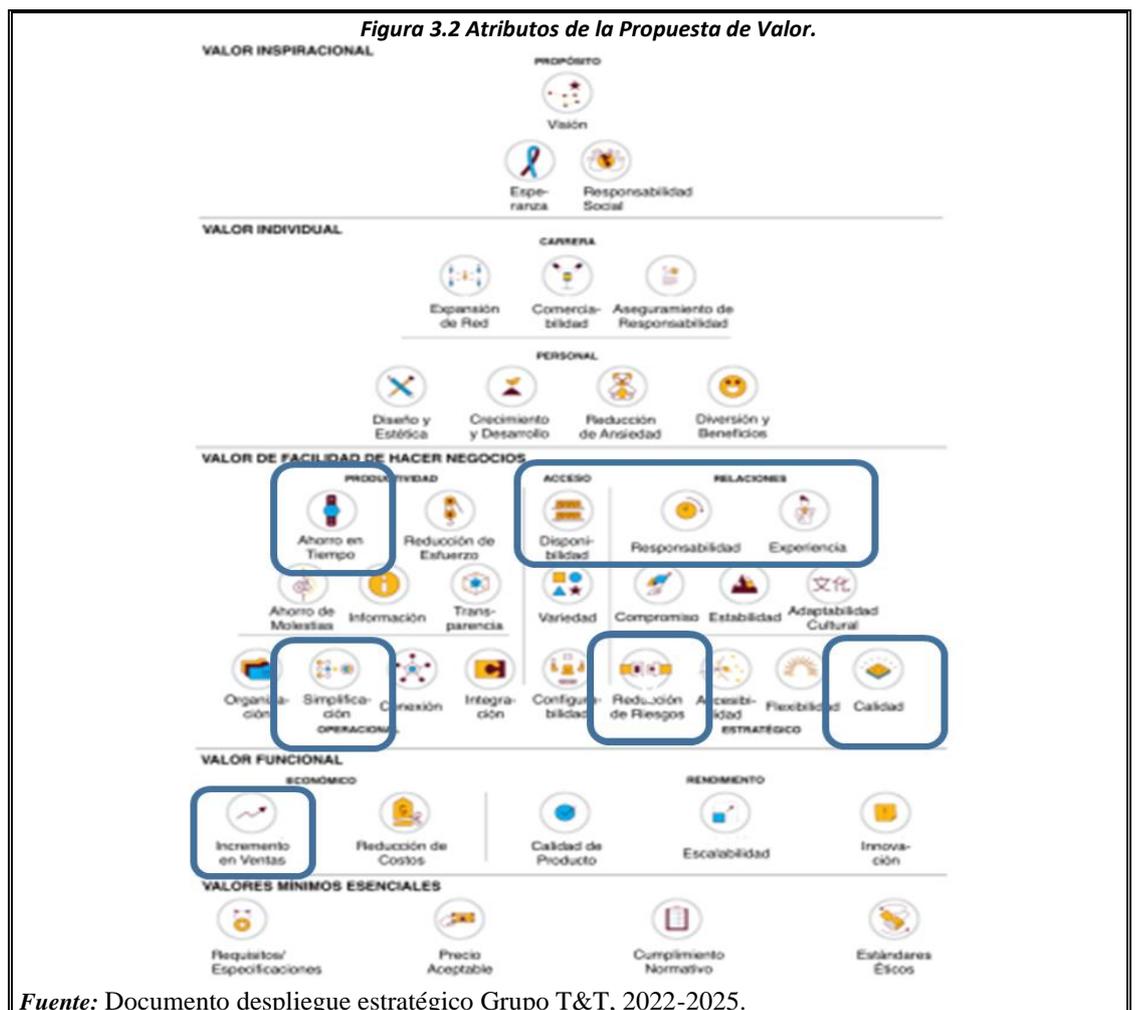
Al definir los pilares que sustentarán el grupo y su visión organizacional es necesario se defina además cuál es la misión de la unidad de negocios de Estibas, detallando el problema que se ama resolver del cliente y su correspondiente propuesta de valor.

La empresa ha definido en el documento de Direccionamiento estratégico que la misión de esta unidad de negocio es: “Garantizar una eficiente y correcta manipulación de carga a través de técnicas seguras y confiables que permitan disminuir tiempos operativos y riesgos laborales con personal capacitado dentro de sus bodegas”. Esta misión es la que responde a solucionar el problema de “Manipulación riesgosa e insegura de las cargas en las bodegas del cliente con el personal informal que pueda ocasionar accidentes laborales” que la dirección identifica como lo que aman resolver hacía el cliente y le da sentido a su propuesta de valor.

Los elementos de valor son las palancas con las que el negocio busca tener una relación establecida con su consumidor, en el caso de Torrestibas la relación es de tipo B2B, los atributos identificados como aplicables para apalancar este servicio son los mostrados en la figura 3.2.

En “Los elementos de valor”, HBR, septiembre de 2016, se explica el modelo B2B que clasifica los elementos descritos en niveles que forman parte de una pirámide, aquellos elementos que brindan un valor más fundamental o básico en la base y ubica aquellos que ofrecen un valor más subjetivo y de complejidad en la cima, por ende se entiende que mientras

más alto en la pirámide se asocian los elementos de valor generará mayor fidelidad y apego ya que se enfocan particularmente en vincular emociones de disminución de ansiedad o preocupaciones generales del consumidor.



**Fuente:** Documento despliegue estratégico Grupo T&T, 2022-2025.

Además de la matriz de crecimiento de nuevas unidades de negocio, las cuales no son parte del alcance de este proyecto, la organización ha detallado hasta la etapa de Pensamiento y construcción estratégica. Estos insumos son parte de lo que dan forma al mapa estratégico.

## **CAPÍTULO 4: Desarrollo y Elaboración del Balance Scorecard.**

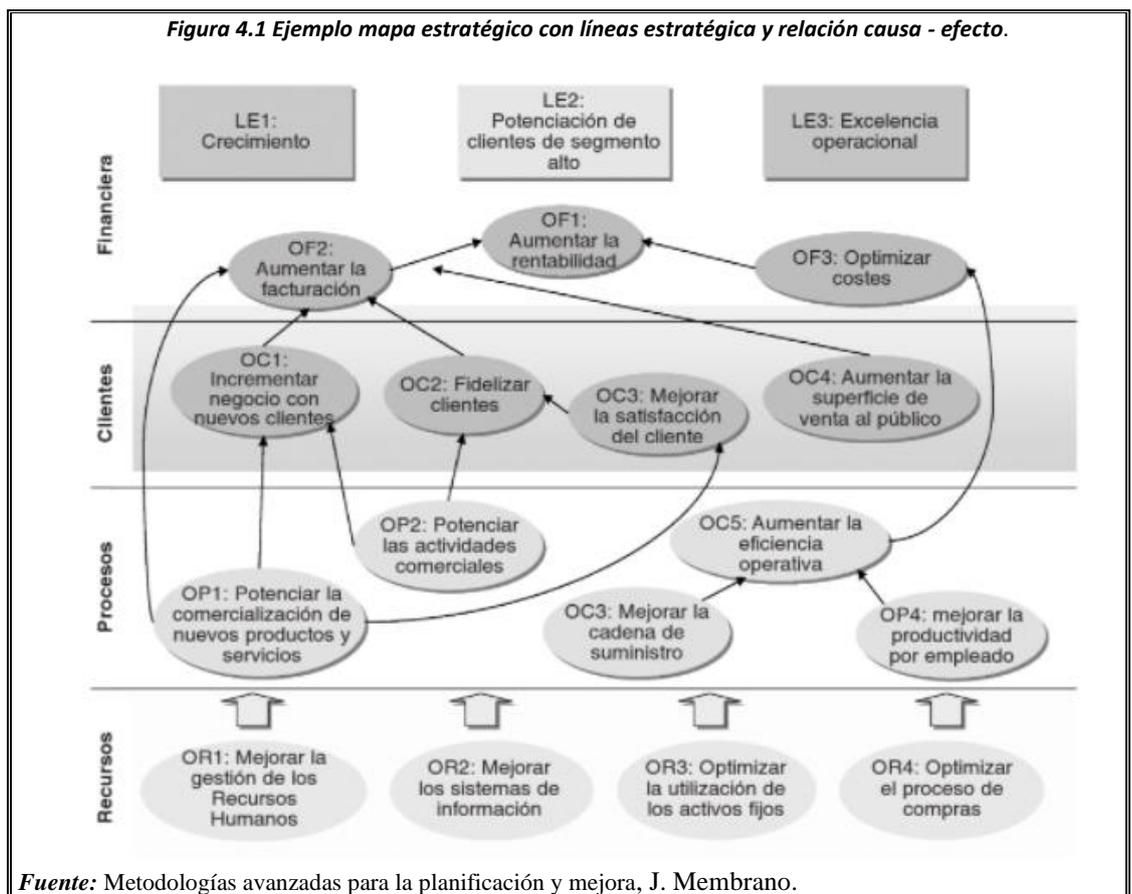
Sainz de Vicuña, J. (2017) afirma que “si un buen plan estratégico quieres tener debes elaborar uno que sea realista, hecho a la medida y redactado con el usuario y deseado por la dirección”. Esta práctica se ve reflejada en el capítulo previo; en este capítulo se detalla la información que lo vuelve realista y además específico, ya que el mapa estratégico traducirá las acciones o deseos que la dirección necesita para hacer realidad la visión y, a su vez, el cuadro de mando integral mostrará específicamente la meta a cumplir para solventar el deseo de la dirección.

En esta implementación se toman como base los pilares definidos y además se divide los mismos en 4 perspectivas.

- i. Perspectiva Financiera: que describe los resultados que en términos financieros o rubros se desean obtener. En donde figuran los más tradicionales como rentabilidad, inversión, ventas, crecimiento o costes.
- ii. Perspectiva Cliente: que detalla las metas en función de la estrategia comercial, valor del cliente y sus segmentaciones de mercado.
- iii. Perspectiva Procesos Internos: identifica los procesos con mayor impacto, centrándose en el núcleo de la gestión del servicio o elaboración del producto.

- iv. Perspectiva Recursos: Aborda las capacidades de la compañía, usualmente relacionado a capacidades de desarrollo del capital humano, capital de la información y organizativo, estos deben estar alineados con los procesos internos.

La representación de un mapa estratégico se basa en la relación causa – efecto. A esto también se lo conoce como inductores ya que cada definición en nivel se influye del anterior e influye a uno siguiente. En la figura 4.1 se observa un ejemplo de mapa estratégico con líneas estratégica y relación causa – efecto.



Cada uno de los pilares se aterriza en deseos del negocio, en donde se expresa claramente la relación que existen y la inducción del deseo de logro hacia la visión. En la figura 4.2 se presenta la recopilación de deseos de la dirección y cómo se han organizado cara a los ejes y perspectivas.



Este mapa será traducido a un Cuadro específico de Metas y objetivos que ayudarán a que el negocio comprenda qué medir y controlar.

El cuadro está definido para un periodo de tiempo de un año, es decir, aplica para el 2022, así mismo en función de la visión organizacional se actualizará en cada período que la organización así lo crea necesario o al cabo del cierre de gestión del periodo definido, a 2023 nuevas métricas se generarán en función del logro de los objetivos y su mejor aplicabilidad bajo el contexto reevaluado.

Las principales iniciativas que la dirección ha logrado identificar para el negocio en este 2022 corresponden a los que se vieron también reflejados en su misión, visión de negocio y propuesta de valor, en un correcto ejercicio de pensamiento estratégico. La organización busca lograr estas sinergias, ya que éstas son importantes para hacer realidad la visión o el sueño futuro.

La figura 4.3 muestra cómo la figura anterior pasa a ser más específica y detallada. En una siguiente etapa estas mismas metas deben tener responsables para el correcto proceso de seguimiento y medición de la estrategia. Un mismo deseo de la dirección podría traducirse en uno o más objetivos estratégicos y así mismo estar implícitos dos deseos de la dirección en un mismo objetivo. Esto se verá impulsado a través del modelo de gobierno de la compañía en donde, apoyados de la tecnología y el modelo de retroalimentación, se generarán las tomas de decisiones que alinearán el cumplimiento y mejora continua de los procesos.

El cuadro de mando integral será retroalimentado en juntas diversas con un foro y espacio exclusivo para cada componente o equipo de gestión.

Figura 4.3 Cuadro de Mando Torrestibas 2022.



En el caso de Torrestibas, se toma como principal eje de estrategia el de Crecimiento rentable, ya que, en la constitución del Estado de Resultados de la compañía, uno de los puntos más críticos es la gestión del personal que representa más del 40% de los costos fijos de la compañía, por lo cual es uno de los factores más relevantes o que preocupan a la dirección. Se busca establecer el cómo solventar mayor cantidad de operaciones con la base del mismo personal e invirtiendo la menor cantidad de recursos posibles en esta captación.

Se ha detallado en la Matriz de metas relacionadas, que consta en la Figura 4.4, que este primer factor estratégico de crecimiento rentable, además de estar potenciado por controles financieros, se ve impulsado por las ventas cruzadas con respecto al portafolio actual del grupo, lo que

significa para Estibas no tener que hacer mayor trabajo de captación sino usar información del inventario de datos de la compañía CTT para la colocación de su servicio, ya que Ciateite es proveedora de operaciones y Clientes potenciales al ser el paso previo en la cadena logística.

**Figura 4.4** Matriz de Metas por Pilares Estratégicos y UNES: Crecimiento Rentable.

	META ASOCIADA
<b>1. CRECER RENTABLEMENTE</b>	1. Incrementar las ventas cruzadas en el servicio de Estibas con respecto a nuestro portafolio de clientes actuales.
	2. Alcanzar al menos el 18% de rentabilidad sobre las ventas.
	3. Aperturar nuevas líneas de negocio con respecto a la diversificación de la propuesta de valor de Estibas (servicios complementarios).
	4. Lograr que al menos el 70% de nuestros clientes soliciten Operaciones Integrales.
	5. Incrementar en un 30% el volumen de las Operaciones Integrales a nivel de Grupo.
	6. Alcanzar un nivel de control en la meta de costos directos e indirectos por unidad de negocio.
	7. Incrementar el nivel de cumplimiento en los acuerdos de ventas establecidos con el cliente por cada línea de negocio.

*Fuente:* Documento despliegue estratégico Grupo T&T, 2022-2025.

Se considera relevante para la determinación de la posibilidad de que un cliente de CIATEITE requiera el servicio de TORRESTIBAS considerar las variables de tipo de carga, cliente, peso, ruta, entre otros factores, esto se buscará interpretar como etiquetado para la determinación de ser más propenso a la compra del servicio. Por lo cual, se hace necesario la composición de un data set integrado entre las compañías que recabe información de la facturación, y conocimiento de la operación con variables que se pueden homologar y ser necesarias para afinar en entendimiento del tipo de clientes que tienen en realidad potencial con Estibas.

#### **4.1 Evaluación de oportunidad de Venta cruzada.**

Durante un diagnóstico realizado con información perteneciente a más de 3 años, se identificó que la empresa cuenta con datos que reflejan que se tiene una participación de mercado dentro de la cartera de Ciateite que es su principal empresa generadora de demanda con respecto a la operación logística, se revisó los datos más relevantes para el entendimiento de la Operación de Ciateite y Torrestibas, sin dejar de lado las empresas primarias como Torres aduana y Tolepu internacional y se identificaron oportunidades de cruce en la data operacional. En el caso de Torrestibas y de la venta cruzada, además, se ha identificado dos niveles de acción o gestión:

- i. Los clientes que pertenecen al portafolio de la empresa CTT con poca o nula participación, es decir, tienen una tarifa pactada pero no son cuentas significativas para el servicio; y,
- ii. Clientes que tienen el 0% de participación con Torrestibas, pero son clientes de una o más empresas del grupo.

Estas definiciones forman parte de la exploración del proyecto, el análisis de esta oportunidad se ha limitado hasta ahora en revisar la gestión realizada de los clientes y no en identificar las características o el aporte de las variables operativas o del cliente que hacen que estos sean un prospecto para la venta cruzada por lo cual se propone generar un modelo matemático que aporte en la identificación del potencial acorde a los datos

actualmente recopilados, con la intención de dar un primer paso para el uso de algoritmos de aprendizaje.

## CAPÍTULO 5: Desarrollo y resultados del modelo.

### 5.1 Limpieza y estructuración de la información.

Según lo publicado en el artículo de xeridia(<https://www.xeridia.com/blog/la-importancia-del-preprocesamiento-de-datos-en-inteligencia-artificial-limpieza-de-datos>) la limpieza de datos forma una parte fundamental al momento de construir un modelo de machine learning debido que permite arreglar errores puntuales que se dan en el conjunto de datos y ayuda a mejorar el rendimiento del mismo. A continuación, se detalla los defectos más comunes que se pueden encontrar al momento de estructurar la información:

**Falta de valores:** La falta de valores es un error muy común cuando se trabaja con bases de datos grandes, la ausencia de valores debe tratarse de manera correcta debido que puede impedir que el sistema de aprendizaje automático pueda entrenarse de manera correcta. Se puede solucionar de diferentes maneras las cuales se detallan a continuación:

- a) **Interpolación:** Si los datos que tenemos en nuestro DataSet son datos temporales, la política que se puede llevar a cabo es interpolar los datos ausentes tomando en consideración los datos próximos
- b) **Completar con valores fijos:** Se puede rellenar los valores faltantes con la media, la moda o colocando un valor fijo que puede ser 0.

- c) **Rellenar con técnicas:** Existen técnicas de regresión que permiten rellenar el valor perdido utilizando el resto de las variables que compone nuestro conjunto de datos.
- d) **Eliminar el registro completo:** Si ninguno de los descriptos antes funciona, en ocasiones se puede optar por la eliminación del registro completo.

**Inconsistencia de los datos:** En muchas ocasiones cuando se está analizando y estructurando la información se detectan errores en el formato o en el tipo de datos que se está analizando.

**Valores duplicados:** En ocasiones en el conjunto de datos se encuentran valores duplicados, es importante poder detectarlos de manera correcta y poder eliminarlos. Si no se realiza esto, el elemento duplicado se tenga en el aprendizaje automático, por lo tanto, se entrenaría de manera sesgada.

**Outliers:** Debido a errores de inserción de datos o almacenamiento pueden aparecer datos atípicos, estos valores pueden afectar la distribución de los datos haciendo que el aprendizaje automático se vea afectado.

## **5.2 Estructura de la información de Ciateite – Estibas.**

Con el fin que el modelo se ajuste a las necesidades solicitadas por el área comercial en donde se requiere una lista de prospecto de manera mensual, se procedió a estructurar la información por meses.

La limpieza y estructuración, así como la creación y selección de las variables más importantes de la tabla. En donde se encuentra toda la información relacionada a los valores que se paga por facturas y el servicio que se contrató. Para el tratamiento de esta tabla se realizó la limpieza de información, y el tratamiento de los datos utilizando las técnicas descritas anteriormente.

**Figura 4.1**  
*Limpieza de la data correspondiente a valores facturados por clientes.*

```

auxFactura <- Valor_Factura[PERIODO_CORTO == PERIODOS]

auxEmpresa <- dcast.data.table(data = auxFactura,
                              formula = RucOrden + ClienteOrden + PERIODO_CORTO ~ Empresa ,
                              fun.aggregate = sum,
                              value.var = c("Total"))

setnames(auxEmpresa, "CIA", "VALOR_TOTAL_CIA")
setnames(auxEmpresa, "EST", "VALOR_TOTAL_EST")
setnames(auxEmpresa, "TOL", "VALOR_TOTAL_TOL")
setnames(auxEmpresa, "TYT", "VALOR_TOTAL_TYT")

auxDocumento <- dcast.data.table(data = auxFactura,
                                 formula = RucOrden + ClienteOrden + PERIODO_CORTO ~ Documento,
                                 fun.aggregate = sum,
                                 value.var = c("Total"))

setnames(auxDocumento, "AVISO COBRO", "AVISO_COBRO")

auxFinal <- merge(x = auxEmpresa,
                  y = auxDocumento,
                  by.x = c("RucOrden", "ClienteOrden", "PERIODO_CORTO"),
                  by.y = c("RucOrden", "ClienteOrden", "PERIODO_CORTO"),
                  all.x = T,
                  all.y = F)

```

De dicho Script se obtuvieron 10 variables, las cuales están conformadas por variables artificiales o compuestas y variables que ya se tenían en los datos de la empresa, las cuales se detallan a continuación:

**Figura 4.2** Variables de facturación a considerar en la conformación del modelo de aprendizaje.

RucOrden	ClienteOrden	PERIODO_CORTO
VALOR_TOTAL_EST	VALOR_TOTAL_TOL	VALOR_TOTAL_TYT
VALOR_TOTAL_AVISO_COBRO	VALOR_TOTAL_HONORARIOS	VALOR_TOTAL_REEMBOLSOS
VALOR_TOTAL_CIA		

### 5.3 Limpieza y estructuración Tabla principal Estibas.

Se realizará la limpieza, se generará la creación, estructura y la selección de variables importantes a contemplar para la tabla principal de Estibas se encuentra toda la información relacionada a lo que es el servicio de que brinda el Grupo Torres – Estibas el cual se encarga de la carga y descarga de todo tipo de mercadería, ensacado de productos en granel, etc.

Para el tratamiento de esta tabla se realizó la limpieza de información, y el tratamiento de los datos utilizando las técnicas descritas en el capítulo anterior.

**Figura 4.3**

**Limpieza de la data correspondiente a variables de operaciones Estibas.**

```
auxCargaCat <- dcast.data.table(data = auxGrabar,
  formula = ruc + cliente + PERIODO_CORTO ~ Carga_cat ,
  fun.aggregate = length,
  value.var = c("Carga_cat"))

auxCarga_tipo <- dcast.data.table(data = auxGrabar,
  formula = ruc + cliente + PERIODO_CORTO ~ Carga_tipo ,
  fun.aggregate = length,
  value.var = c("Carga_tipo"))

auxTipoEstibas_cat <- dcast.data.table(data = auxGrabar,
  formula = ruc + cliente + PERIODO_CORTO ~ TipoEstibas_cat ,
  fun.aggregate = length,
  value.var = c("TipoEstibas_cat"))

auxRiesgo_cat <- dcast.data.table(data = auxGrabar,
  formula = ruc + cliente + PERIODO_CORTO ~ Riesgo_cat ,
  fun.aggregate = length,
  value.var = c("Riesgo_cat"))

auxFinal <- merge(x = auxCargaCat,
  y = auxCarga_tipo,
  by.x = c("ruc", "cliente", "PERIODO_CORTO"),
  by.y = c("ruc", "cliente", "PERIODO_CORTO"),
  all.x = T,
  all.y = F)

auxFinal <- merge(x = auxFinal,
  y = auxTipoEstibas_cat,
  by.x = c("ruc", "cliente", "PERIODO_CORTO"),
  by.y = c("ruc", "cliente", "PERIODO_CORTO"),
  all.x = T,
  all.y = F)
```

De dicho Script se obtuvieron 17 variables, las cuales también están conformadas por variables artificiales o compuestas y variables que ya existían orgánicamente en los registros de la compañía, se detallan a continuación:

**Figura 4.4** Variables operativas a considerar en la conformación del modelo de aprendizaje.

RUC	Cliente	Periodo corto
CARGA_20	CARGA_40	CARGA_CS
CARGA_OTRO	CARGA_tipo_DIU	CARGA_tipo_NOC
CARGA_tipo_OTRO	TIPO_STIBAS_OTROS	TIPO_STIBAS_SUELTA
RIESGO_AMARILLO	RIESGO_AZUL	RIESGO_NARANJA
RIESGO_ROJO	RIESGO_OTROS	

#### 5.4 Limpieza y estructuración Tabla CLIENTE\_OPERACIONES.

En la tabla CLIENTE\_OPERACIONES se encuentra toda la información relacionada a la cantidad de transacciones que realiza un cliente del Grupo Torres y Torres por empresa. Para el tratamiento de esta tabla se realizó la limpieza de información, y el tratamiento de los datos utilizando las técnicas descritas.

**Figura 4.5**  
*Limpieza de la data correspondiente a variables de Cliente – Operaciones.*

```

auxOperaciones2 <- auxOperaciones[PERIODO_CORTO == PERIODOS, .(ruc,cliente,Empresa,PERIODO_CORTO)]
auxOperaciones2 <- dcast.data.table(data = auxOperaciones2,
                                   formula = ruc + cliente + PERIODO_CORTO ~ Empresa,
                                   fun.aggregate = length,
                                   value.var = c("Empresa"))

setnames(auxOperaciones2,"CTT","CANTIDAD_OPERACIONES_CTT")
setnames(auxOperaciones2,"EST","CANTIDAD_OPERACIONES_EST")
setnames(auxOperaciones2,"TYT","CANTIDAD_OPERACIONES_TYT")
#setnames(auxOperaciones,"TOL","CANTIDAD_OPERACIONES_TOL")
print(PERIODOS)

```

De dicho Script se obtienen 6 variables, las cuales están al igual que las demás tablas están conformadas por variables compuestas y variables de registro generado por la empresa, las cuales se detallan a continuación:

**Figura 4.6** Variables Cliente – Operaciones a considerar en la conformación del modelo de aprendizaje.

RUC	Cliente
PERIODO_CORTO	CANTIDAD_OPERACIONES_CTT
CANTIDAD_OPERACIONES_EST	CANTIDAD_OPERACIONES_TYT

## 5.5 Creación de nuestra master table

La master table será nuestro set de datos en donde se reúne y recopila toda la información de las diferentes tablas con el fin de tener una estructurada y limpia que posteriormente va alimentar al modelo.

**Figura 4.7**  
*Creación de una master data.*

```
masterDf <- merge(x = DATA_OPERACIONES_X_EMPRESA,
  y = DATA_ESTIBAS_OPERACIONES,
  by.x = c("ruc", "cliente", "PERIODO_CORTO"),
  by.y = c("ruc", "cliente", "PERIODO_CORTO"),
  all.x = T,
  all.y = F)

for (var in names(masterDf)){
  set(x = masterDf, i = which(is.na(masterDf[[var]])), j = var, value = 0)
}

DATA_VALORES_FACTURA <- as.data.table(read_csv(paste0("C:/ProyectoAndrea/data/DATA_ESTRUCTURADA/DATA_VALORES_FACTURA
  col_types = cols(VALOR_TOTAL_CIA = col_number(),
  VALOR_TOTAL_EST = col_number(), VALOR_TOTAL_TOL = col_numb
  VALOR_TOTAL_TYT = col_number(), VALOR_TOTAL_AVISO_COBRO =
  VALOR_TOTAL_REEMBOLSOS = col_number())))
```

## 5.6 Creación de las variables objetivo.

La variable objetivo o de destino es la variable cuya explicación se busca a través de un conjunto de datos, en este caso es por la cual se desea tener un juicio más profundo de lo que se está analizando. Un algoritmo de aprendizaje automático supervisado utiliza los datos históricos en el cual aprender y descubre patrones de su conjunto de datos y de su objetivo en particular. En este caso planteado la variable objetivo será la identificación del cliente prospecto o potencial.

Es de suma importancia tener el objetivo claramente definido ya que lo que genera el algoritmo es un mapeo de las relaciones que se tiene entre los datos de entrada y este objetivo. En conclusión, si la variable objetivo no tiene sentido su objetivo tampoco lo tendrá, si deseáramos obtener una explicación de una variable que quizá es dada por la operación y no influida a través de otras, la construcción de un modelo no tendría sentido.

Dicho esto, en la función dcast en su versión simple necesitamos entregar:

- a) la tabla sobre la que operará.
- b) qué columnas disponer en filas.
- c) qué columnas disponer en columnas.

Debemos entender que la función dcast utiliza una serie de reglas para determinar cuál es la columna de la cual se extraen los datos para incluirlos en la tabla resultante.

### **5.7 Creación de variable objetivo de Estibas.**

Para nuestra creación de la variable objetivo de estibas, obtenemos los datos de los clientes que teniendo el servicio CTT han activado el servicio de estibas, la información se guarda por períodos mensuales. Debido a las directrices del área comercial se va a entregar una lista de prospecto cada mes.

*Figura 4.8*  
*Generación de la variable objetivo.*

```
EstibasOperaciones[, PERIODO_CORTO := substr(FechaOperacion ,1,7)]  
  
Empresas[, PERIODO_CORTO := substr(FECHA_PRIMER_TRAMITE ,1,7)]  
  
INFORMACION_CTT <- Empresas[empresa == 'CTT']  
INFORMACION_EST <- Empresas[empresa == 'EST']  
  
setnames(INFORMACION_CTT, 'FECHA_PRIMER_TRAMITE', 'FECHA_PRIMER_TRAMITE_CTT')  
setnames(INFORMACION_CTT, 'PERIODO_CORTO', 'PERIODO_CORTO_CTT')  
  
setnames(INFORMACION_EST, 'FECHA_PRIMER_TRAMITE', 'FECHA_PRIMER_TRAMITE_EST')  
setnames(INFORMACION_EST, 'PERIODO_CORTO', 'PERIODO_CORTO_EST')  
  
INFORMACION_CTT$empresa <- NULL  
INFORMACION_EST$empresa <- NULL  
  
TarjetDF <- merge(x = INFORMACION_CTT,  
                  y = INFORMACION_EST,  
                  by.x = c("ruc", "cliente"),  
                  by.y = c("ruc", "cliente"),  
                  all.x = T,  
                  all.y = F)
```

## 5.8 Selección de Técnica Machine Learning.

Existen varias técnicas que se pueden utilizar en el aprendizaje automático y a su vez permite solventar diversos tipos de problemas en los negocios, uno de los más utilizados en estos años por las empresas es el de la venta cruzada. Según (Nielsen, 2016), la efectividad en el uso de estas técnicas depende, de los datos que se tienen disponible y su calidad de registros.

Según el artículo expuesto de (Chen & Guestrin, 2016), el uso de la técnica Extreme Gradient Boosting (XGboost) es altamente utilizada debido a su nivel de efectividad. A continuación, detallaremos la razón por la que se utilizó dicha técnica.

- a) Debido que contamos con un DataSet de no más de 10.000 registros y no más de 30 variables la técnica es muy buena para la escasez de datos.
- b) La técnica de árboles de decisiones ha resultado bastante útil al momento de realizar predicciones en diversas aplicaciones; es por eso que la técnica XGBOOST presenta ventajas, como la inclusión de parámetros diversos que ayudan con la regularización de los tamaños en los árboles resultantes y los pesos en cada uno de los nodos o terminales, lo que representa un beneficio ya que se genera reducción en la varianza resultante en cada árbol.
- c) Estos algoritmos tienen una buena relación con el sesgo - varianza en el ajuste del modelo, lo que ayuda a aprender de forma eficiente la estructura de los árboles y además aplicar penalización de manera asertiva, lo que ayuda que el modelo tenga óptimo performance al momento de ajustarlo.

### **5.9 Entrenamiento del modelo de contratación de servicio de Estibas.**

Para la realización de este modelo se utilizó el algoritmo de árboles de decisiones XGBoost, en el cual se probaron 2 configuraciones de parámetros diferentes con el fin de poder obtener los mejores resultados para el modelo.

Para trabajar con el algoritmo XGBoost, todos los datos que entren al modelo tienen que ser numéricos y se deberá tener instalado la librería en R llamada XGBoost.

El fin del modelo de contratación del servicio de estibas es poder predecir cuál es la propensión que de un cliente CTT me compre el servicio de estibas. Para ello se realizaron los siguientes pasos:

- a) Se carga la información de la master table de los períodos que se va a realizar el entrenamiento; una vez cargada la información se procede a añadir la variable objetivo o predictora.
- b) Similar a lo descrito en el punto de arriba, tenemos que escoger cuál será el conjunto de datos de prueba, el cual ayudará a identificar qué tan preciso es el modelo.
- c) Como tercer punto cargamos la información del cual modelo va a realizar el test, y así mismo con su respectiva variable predictora.
- d) XGBoost requiere que los datos que vayan a ingresar al modelo sean de tipo de datos `xgb.DMatrix`, se deberá realizar la conversión al tipo de datos que pide el modelo y especificarle cuál es la variable objetivo:

**Figura 4.9**  
**Conversión para el tipo de datos que necesita el modelo.**

```
modelCols <- setdiff(names(masterTrainDf), c("ADQUIERE_SERVICIOS_ESTIBAS", "ruc", "PERIODO_CORTO", "cliente"))

masterTrainDmatrix <- xgb.DMatrix(data = as.matrix(masterTrainDf[, modelCols, with = F]),
                                label = masterTrainDf[["ADQUIERE_SERVICIOS_ESTIBAS"]])
gc()
masterDevDmatrix <- xgb.DMatrix(data = as.matrix(masterDevDf[, modelCols, with = F]),
                                label = masterDevDf[["ADQUIERE_SERVICIOS_ESTIBAS"]])
gc()
masterTestDf$TIPO_STIBAS_OTROS <- 0
masterTestDmatrix <- xgb.DMatrix(data = as.matrix(masterTestDf[, modelCols, with = F]),
                                label = masterTestDf[["ADQUIERE_SERVICIOS_ESTIBAS"]])
gc()
```

- e) Antes de ingresar los datos al modelo se deberá realizar la parametrización de los datos de la siguiente manera:

**Figura 4.10**  
*Conversión para el tipo de datos que necesita el modelo.*

```
xgbParams <- list(booster = "gbtree",
  objective = "binary:logistic",
  max.depth = 6,
  eval_metric = "auc",
  missing = NA,
  colsample_bytree = 0.7,
  subsample = 0.7,
  alpha = 0,
  lambda = 1,
  eta = 0.1,
  early_stopping_rounds = 10,
  nrounds = 400) |
```

A continuación, se detalla la parametrización de cada uno de los datos (xgboost, s.f.):

- a) Booster = "gbtree": Por lo general este parámetro define la versión del árbol de regresión o tipo.
- b) Objective = "binary: logistic": Debido que nuestros de contratación de Estibas es modelo que me ayuda a predecir qué tan probable es que adquiera el servicio, es le coloca de tipo binary.
- c) Max.depth = 6: El valor por defecto es 6. Este parámetro indica el valor de profundidad máximo y es un número entero del 1 al 100. Indica el número mínimo de muestras que debe tener cada rama después de dividir un nodo. Incrementar este valor hará que el modelo sea más complejo y probable que se sobreajuste. Tenga en cuenta que XGBoost consume memoria de forma agresiva al entrenar un árbol profundo.

- d) Eval\_metric = "auc": Se escoge de tipo "auc" para generar el cálculo de un área bajo la curva ROC (AUC). AUC puede variar entre 0 y 1.
- e) Colsample\_bytree = 0.7: este parámetro define qué porcentaje de características o columnas se usará como submuestra para la construcción de cada árbol. Usualmente se logran que el conjunto de características sean distintas para cada árbol construido.
- f) Subsample: Proporción de la submuestra de las instancias de entrenamiento se utilizará el 70% y este muestreo ocurrirá una vez en cada interacción.
- g) eta = 0.1: Logra la reducción de las ponderaciones de características y logra un proceso más conservador, evita el sobreajuste.
- h) Early\_stopping\_rounds = 10: es un parámetro que controla la cantidad de iteraciones que esperaremos para la disminución del valor de pérdida, es necesario ya que los valores disminuyen aleatoriamente en cada iteración.
- i) Nrounds = 400: Representa el máximo de iteraciones que tendremos.

## **5.10 Validación de los resultados.**

Según Ratnakar, 2019 existen 3 métricas en la que podemos medir el modelo utilizado en este estudio, estas son la curva de Gini, KS o ROC. Las más utilizadas en investigaciones son el KS y el Gini que es en las métricas que se va a centrar el modelo propuesto.

### **a) Curva Loren e Índice Gini:**

Se basa en mostrar qué tan bueno es un modelo en comparación con otro modelo aleatorio (modelo ingenuo), llamada también medida de desigualdad en una distribución.

*El coeficiente de Gini es un número que va de 0 a 1:*

En donde 0 se interpreta como la igualdad perfecta (es decir, todos tienen los mismos ingresos). Y en donde 1 es el valor que se otorga a la perfecta desigualdad (es decir, una persona acapara los ingresos y los demás tienen ninguno).

El estadístico está dado por la siguiente ecuación:

$$Gini = 1 - \sum_{i=L}^{H'} (P_b(i+1) - P_b(i))(P_g(i+1) + P_g(i))$$

Donde:

$i$ : es el valor de score, en el rango  $L - H$ , que es,  $L \leq i \leq H$ .

$P_g(i)$ ,  $P_b(i)$ : Proporción de buenos y malos con score menor o igual a  $i$ , en la población, respectivamente.

Se considera que para modelos de originación, un coeficiente de Gini menor a

35% es sospechoso y mayor o igual a 50% es más que satisfactorio (Anderson 2007, 205).

Una interpretación sugiere que logra mostrar el poder discriminatorio del modelo al lograr la separación de lo “bueno” de lo “malo” frente a una selección que es aleatoria o lo “válido” de lo “no válido”. El índice de Gini se interpreta como la proporción entre las áreas marcadas bajo  $[A / (A + B)]$ . Esto mide el desempeño del modelo en comparación con la selección aleatoria. El resultado de un Gini de 0% indicará que el modelo no es mejor que el aleatorio o, en otras palabras, que este no posee poder de predicción. En su lugar, un resultante de Gini del 100% indicará un el modelo es perfecto o predecirá con precisión absoluta (100%) bueno y malo. Los valores de resultantes de Gini por encima del 60% son considerados buenos, es decir, el modelo emplea un buen desempeño o ajuste en su predicción.

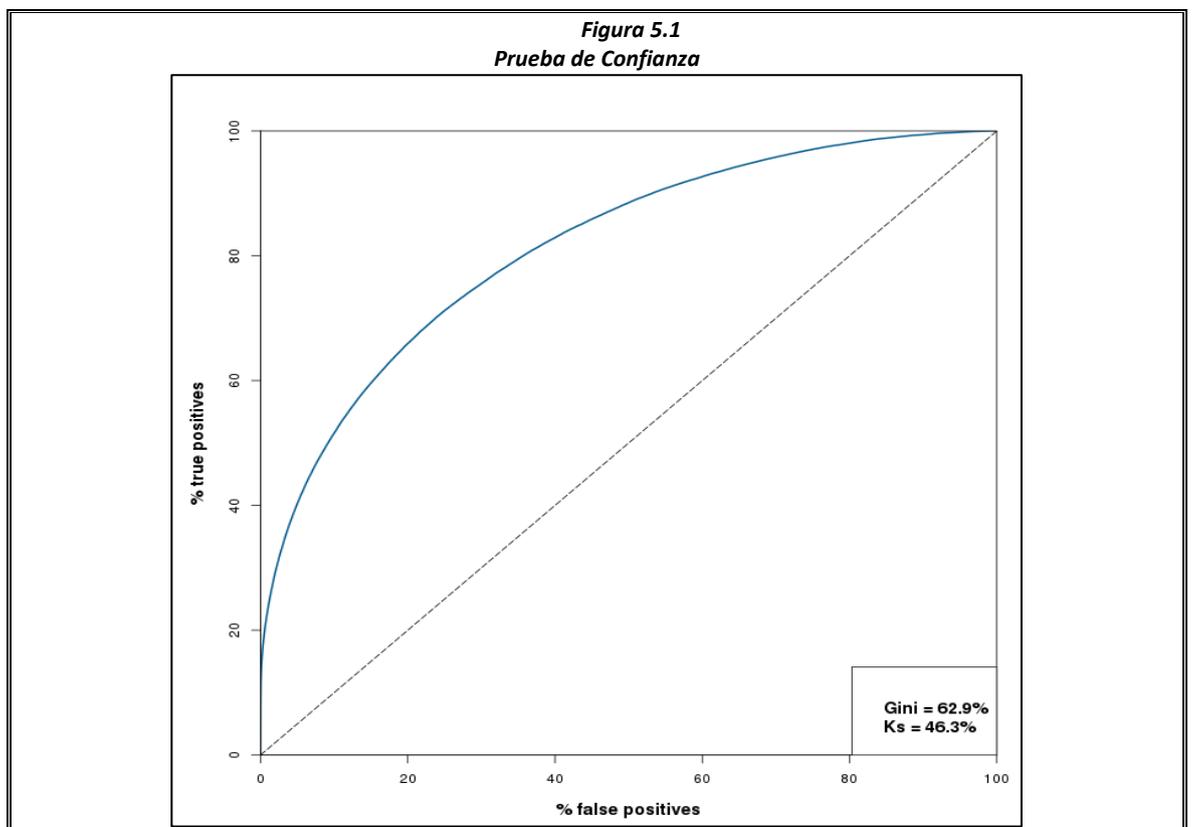
#### **b) Kolmogorov-Smirnov statistics (KS)**

En este caso al igual que Gini, la prueba de Kolmogorov-Smirnov también tiene el poder de discriminación del modelo al separar "Bueno" de lo "Malo", la conocida también como prueba de bondad de ajuste se explica como la separación más alta entre la Tasa acumulativa buena y la Tasa acumulada mala. Cuanto mayor sea este valor de KS, mejor es el modelo (mayor separación entre lo bueno y lo malo). KS es un parámetro que puede ir de 0 a 100%, se considera que un valor superior a 20% es aceptables para la valoración de un modelo.

## 5.11 Resultados

Los resultados del entrenamiento del modelo aplicado sobre la data de esta investigación se han testeado para poder verificar la capacidad predictiva lograda a través de la construcción de este modelo basado en la primera toma de variables que la organización tiene a disposición y logra identificar.

Basado en el trabajo realizado por Sohrabi de Multiclass Cross-selling Model en el año 2018 para el Banco Danske de Dinamarca, en donde explica la evolución de las técnicas aplicada para el Cros-Selling, notamos que el uso de estos dos test nos ayuda a determinar si el ajuste logrado es el deseado o nos permitirá el modelo concluir al respecto con confianza.



Para medir la capacidad predictiva del modelo se calcula los estadísticos y se puede observar que para el modelo y testing el estadístico KS se encuentra en el rango de satisfactorio de acuerdo con los estándares empleados en TransUnion. De igual forma, el coeficiente de Gini es más que satisfactorio al alcanzar un valor superior al 50%.

## **CAPÍTULO 6: Conclusiones y Recomendaciones.**

En este trabajo se abordaron varias técnicas y metodologías en donde se busca aportar valor al negocio, no sólo ubicando un caso de investigación por la cantidad de datos que se tengan registrados sino también enfocado en solventar la verdadera necesidad o propósito principal cara a estrategia.

Por lo cual las principales conclusiones del trabajo son:

1. El efecto de combinar metodologías ágiles y de performance en una compañía con estadística, potencia el efecto de las futuras decisiones.
2. Los equipos de trabajo pueden potenciarse desde sus acciones del día a día, hasta sus resultados con una correcta aplicación de la estadística en su enfoque de acción, como las áreas de comercialización y fidelización quienes podrán tener claridad del potencial de venta cruzada que realmente tienen los clientes antes de gastar esfuerzos en un barrido total en su gestión de captación de cuentas.
3. Se tiene mayor claridad y entendimiento de los posibles casos de éxito o respuesta que se buscan dar con una investigación estadística a partir de entender el contexto empresarial, para qué existe la compañía y sus pilares estratégicos, lo cual facilita la construcción de los modelos y dar respuesta a las interrogantes que como profesionales de la estadística podemos solventar.

4. Una etapa de mucho enfoque y que necesita trabajo de campo y entendimiento es la limpieza y estructuración de la información, las empresas de hoy en día tienen gran cantidad de datos almacenados, sin embargo, la mayoría desconoce si significan algo o el tipo de decisiones que pueden lograr o accionar con el buen uso y manejo de los mismos.
5. Los algoritmos de aprendizaje son valiosos no solamente en la gestión de datos masivos y transaccionales tipo Big Data sino también en el marco de acción que una PIME puede tener.

Así mismo hubo limitantes o lecciones aprendidas en el logro de este resultado presentado, por lo que se tienen las recomendaciones siguientes:

1. Con un modelo enfocado en explorar inicialmente las variables que los negocios hoy en día gestionan se puede aprender para siguientes pasos de ajuste que vendrán acompañados de mejores decisiones y acciones más precisas, por lo cual este modelo generado debe ser constantemente monitoreado para pulir o mejorar su ajuste de predicción.
2. En el grupo empresarial TYT existen otros puntos de acción para las estadísticas inferenciales o predictivas, que pueden hacer que otros resultados como la planificación correcta de las operaciones tengan mucho más tiempo para organizar sus recursos y responder de mejor

forma a los clientes, se recomienda que se exploren estos puntos para dar inicio a un modelo que aprenda de esta necesidad y ayude en el logro de este resultado.

3. El equipo de Excelencia del grupo T&T maneja estadística relevante para el negocio y ha logrado manejar las cifras claves del negocio en su plataforma de BI, sin embargo, se necesitan la incorporación de más profesionales enfocados en otro tipo de herramientas y aplicaciones para lograr que se tomen en cuenta otro tipo de escenarios, menos reactivos y más proactivos.
4. En las juntas mayores de revisión de resultados además de los indicadores habituales se pueden incluir los resultados de las iteraciones o aprendizajes de los modelos para mantener constantemente retroalimentados a los ejecutivos de estas nuevas aplicaciones y promover que se descubran basados en las necesidades de la dirección, nuevas oportunidades.

## **CAPÍTULO 6: Referencias Bibliograficas.**

- <sup>1</sup> Govindarajan, V. & Gupta, A. (2001), "Building an Effective Global Business Team" MIT Sloan Management Review; Cambridge Vol. 42, Iss. 4.
- <sup>2</sup> INCAE Business School (2007), "Factores clave del éxito para la implementación de estrategias en Latinoamérica", Costa Rica.
- <sup>3</sup> Llorente, A. (20 de Noviembre de 2018). PiperLab. Obtenido de <https://piperlab.es/2018/11/20/como-puede-la-inteligencia-artificial-ayudarte-a-mejorar-tu-venta-cruzada/>
- <sup>4</sup> R. J. BRACHMAN y T. ANAND. "The Process of Knowledge Discovery in Databases". En Advances in Knowledge Discovery and Data Mining, Fayyad, Piatetsy-Shapiro, Smyth y Uthurusamy Eds., AAAI Press, Menlo Park, California, 1996. Págs. 37-57.
- <sup>5</sup> "From Patterns in Data to Knowledge Discovery: What Data Mining Can Do", Francesco Gullo, 3rd International Conference Frontiers in Diagnostic Technologies, ICFDT3 (2013).
- <sup>6</sup> Nielsen, D. (2016). Tree Boosting With XGBoost.
- <sup>7</sup> Ratnakar. (2019). Obtenido de <https://www.quora.com/Which-one-is-better-to-evaluate-a-logistic-regression-Gini-KS-or-ROC>.
- <sup>8</sup> Sohrabi, A. (2018). Lund University Publications. Obtenido de Lund University Publications Web.