



TRABAJO DE GRADO
Opción Seminario-Diplomado.

Business Intelligence, Manejo Estratégico de los Datos Basados en IA

Corporación Universitaria Remington.

Facultad de Ciencias Empresariales

Marketing.

Autor: Kenny Chayanne Calderon Garcia

Docente: Jarrison Mosquera Rentería

Opción de Trabajo de grado Seminario-Diplomado

2025

Tabla de Contenidos

Resumen.....	3
Palabra clave.....	4
Pregunta orientadora de la búsqueda	5
Justificación	6
Metodología de búsqueda de la información.....	7
Sustentación teórica de la pregunta.....	9
Tablas.....	15
Conclusiones.....	16
Referencias.....	17

Resumen

La incorporación de inteligencia artificial (IA) en procesos organizacionales se ha vuelto clave para alcanzar mayor eficiencia operativa y optimizar recursos. Este documento examina cómo el adecuado procesamiento de datos permite entrenar modelos predictivos que potencian los resultados comerciales, especialmente a través de agentes virtuales aplicados en campañas de ventas.

El desarrollo de soluciones basadas en IA requiere datos limpios, estructurados y representativos del entorno real. En este sentido, la gestión de la información implica etapas como recopilación, filtrado, transformación y organización de los registros, lo que asegura que los modelos sean funcionales, precisos y relevantes. En el caso analizado, se parte de una base de 8.000 posibles clientes, con una tasa actual del 40% de contactabilidad y un 9% de efectividad en ventas.

Se diseñó un modelo predictivo con herramientas de la nube de Azure, como Machine Learning y AutoML, incorporando técnicas como balanceo de clases, clasificación supervisada y limpieza de datos. El objetivo es construir un voicebot capaz de priorizar contactos con mayor probabilidad de conversión y mejorar la gestión de campañas sin requerir una ampliación de personal.

El sistema predice cuáles registros tienen más posibilidades de cierre, lo que permite enfocar los esfuerzos en segmentos más receptivos, mejorando tanto la tasa de contacto como la conversión. Con una precisión del 70% en la predicción y un rediseño en la estrategia de acercamiento, se estima un aumento del 3% en la efectividad general, logrando aproximadamente 140 ventas adicionales con el mismo recurso humano.

Además del rendimiento técnico, se destaca la necesidad de aplicar principios éticos en el uso de IA, asegurando transparencia, equidad y protección de los datos personales. Plataformas como Azure ofrecen herramientas integradas para cumplir con estos estándares. El enfoque presentado permite a las empresas optimizar sus campañas y reducir costos operativos, aprovechando al máximo los beneficios de la automatización inteligente.

Palabras clave

Inteligencia Artificial:

Tecnología que permite a las máquinas simular capacidades humanas como el aprendizaje, la toma de decisiones y la resolución de problemas. Es el eje central en el desarrollo de agentes virtuales inteligentes.

Procesamiento de Datos:

Conjunto de técnicas utilizadas para recolectar, limpiar, transformar y estructurar datos, con el objetivo de que sean aptos para entrenar modelos de IA precisos y confiables.

Voicebot:

Asistente virtual automatizado que interactúa mediante voz con los usuarios. Su desempeño depende del entrenamiento previo a partir de datos históricos y patrones de comportamiento del cliente.

Automatización Comercial:

Aplicación de tecnologías para optimizar tareas dentro del proceso de ventas, reduciendo la intervención humana y mejorando la eficiencia operativa y la experiencia del cliente.

Modelos Predictivos:

Algoritmos de aprendizaje automático que utilizan datos históricos para anticipar comportamientos futuros, permitiendo tomar decisiones más acertadas en entornos comerciales.

Pregunta orientadora de la búsqueda

¿Cómo puede el procesamiento adecuado de datos mejorar el rendimiento de un agente virtual basado en inteligencia artificial en campañas comerciales del sector BPO?

Objetivo General

Explorar cómo se lleva a cabo la preparación y tratamiento de grandes cantidades de datos con el fin de entrenar modelos basados en inteligencia artificial. (voicebot)

Objetivos específicos:

Describir las fases fundamentales del tratamiento de datos necesarias para el desarrollo de modelos de IA aplicados a procesos comerciales.

Analizar cómo la calidad y el manejo adecuado de la información son importantes en el desempeño del agente virtual.

Determinar de qué manera un voicebot puede generar decisiones eficaces y flexibles al interactuar dentro de un entorno comercial.

Justificación

En un entorno empresarial cada vez más competitivo y orientado a la eficiencia, las organizaciones enfrentan el desafío de incrementar sus niveles de venta sin que esto implique un aumento proporcional en los costos operativos, especialmente en lo relacionado con el recurso humano. Ante este panorama, la inteligencia artificial (IA) se presenta como una herramienta estratégica capaz de automatizar procesos clave de gestión comercial y atención al cliente, lo cual permite mantener e incluso mejorar el rendimiento organizacional con una estructura de personal optimizada.

El desarrollo de agentes virtuales entrenados mediante modelos de IA representa una oportunidad para implementar soluciones que puedan atender clientes y muchas más

en el sector BPO donde la tecnología es clave, gestionar información y asistir en procesos de venta sin requerir una intervención humana constante. La implementación de soluciones basadas en inteligencia artificial permite disminuir significativamente los costos relacionados con la incorporación (atracción del talento, exámenes médicos) y formación de personal, al tiempo que ofrece atención continua, respuestas adecuadas y coherentes. Estos elementos son claves para potencializar la experiencia del cliente y favorecer el incremento en las tasas de efectividad de ventas con menores esfuerzos.

Este estudio genera valor al presentar una perspectiva tanto técnica como aplicada sobre el tratamiento de datos, considerado un paso esencial en el desarrollo de un voicebot de alto rendimiento y performance. Mediante la optimización de esta etapa, se pretende asegurar que la IA sea capaz de ajustarse a las exigencias del entorno comercial y contribuir directamente a la estrategia de crecimiento en ventas, maximizando la rentabilidad con una inversión reducida en talento humanos adicionales, lo cual también permite generar mejores condiciones salariales en la planta actual generando estabilidad y bajos niveles de rotación.

Metodología de búsqueda de la información

La elaboración de este trabajo se fundamentó en una estrategia de recopilación documental, enfocada en fuentes digitales y bibliotecas virtuales. Para ello, se llevó a cabo una revisión selectiva de contenidos especializados disponibles en internet, priorizando aquellos materiales con respaldo académico y técnico.

El proceso de búsqueda se desarrolló mediante el uso de herramientas como Google Académico, bases de datos de acceso universitario y plataformas oficiales de proveedores tecnológicos, entre ellas Microsoft Learn. Se establecieron criterios de pertinencia y actualidad, dando prioridad a publicaciones generadas entre 2012 y 2023 que abordaran temas relacionados con inteligencia artificial, gestión de datos, automatización en entornos comerciales y desarrollo de modelos predictivos.

La selección de contenido se guió por el uso de términos clave tanto en español como en inglés —por ejemplo: procesamiento de datos, automatización comercial, voicebot, machine learning, customer data modeling— así como por el uso de operadores de búsqueda que permitieron afinar los resultados obtenidos.

Una vez recopilada la información, se procedió a realizar una lectura crítica y análisis comparativo de las fuentes, extrayendo aquellos conceptos, metodologías y aplicaciones más relevantes para el contexto del estudio. Este enfoque permitió construir un marco conceptual sólido y al mismo tiempo contextualizado en el uso práctico de la IA dentro del sector BPO, asegurando que la información utilizada estuviera alineada con los objetivos del trabajo.

Sustentación teórica de la pregunta

1. Inteligencia Artificial y Conjunto de Datos Para Entrenamiento:

La inteligencia artificial (IA) corresponde a un campo de estudio dentro de la informática que busca diseñar sistemas capaces de ejecutar actividades que normalmente requieren habilidades humanas, tales como razonar, aprender y tomar decisiones informadas (Russell & Norvig, 2021). En el ámbito del aprendizaje automático, los algoritmos dependen del análisis de grandes volúmenes de datos para detectar patrones y tendencias y con esto construir modelos que permitan predecir o las mejores propensiones y la correcta forma de gestionar de acuerdo a las características de cada cliente final.

La eficacia de estos modelos está directamente relacionada con la calidad y estructuración del conjunto de datos utilizados durante su entrenamiento. es vital que los datos utilizados no tengan inconsistencias o sesgos, y que su volumen e integridad representen con precisión lo que se desea modelar. Solo de esta forma se logra que los modelos puedan adaptarse a las necesidades de cada momento.

2. Preparación y Diseño de Datos en Proyectos de IA

La preparación de datos para su uso en modelos de IA involucra un proceso estructurado de actividades técnicas que garantizan que la información esté lista para ser procesada por los algoritmos. Este flujo contempla etapas como la selección de fuentes de

datos confiables, la limpieza de los datos, la formulación de variables, el etiquetado de casuísticas y la división de los datos en subconjuntos. (Zhang et al., 2020).

Un diseño errado en esta etapa puede provocar desviaciones importantes en la calidad del modelo final, tales como reprocesos y la introducción de sesgos no intencionados en los análisis. Por ello, es vital emplear estrategias como el reequilibrio de clases, el uso de técnicas de aumento de datos y la revisión de las variables utilizadas en el modelo (Domingos, 2012).

3. Consideraciones Éticas y de Seguridad

Azure incorpora funcionalidades como Azure Purview para la gobernanza de datos, y marcos de trabajo como Responsible AI para promover el desarrollo ético, inclusivo, transparente y seguro de sistemas inteligentes (Microsoft, 2023d).

Asimismo, la plataforma garantiza el cumplimiento de normativas internacionales de seguridad, como GDPR, mediante el cifrado de datos en reposo y en tránsito, autenticación multifactor y control de acceso basado en roles.

Contexto de la Campaña piloto aplicando el procesamiento de datos

La campaña en estudio corresponde a un proceso de ventas de seguros realizado por un equipo comercial que dispone de una base de datos de 8,000 registros de clientes potenciales. Esta base incluye información como nombre, edad, historial de contacto, producto de interés, ubicación, entre otros. En su ejecución actual, la campaña presenta una tasa de contactabilidad del 40% (3,200 personas) y una tasa de efectividad de cierre del 9% (288 ventas), lo que indica márgenes de mejora significativos tanto en la priorización de contactos como en la estrategia de abordaje.

Ante esta situación, se plantea desarrollar un agente virtual inteligente que apoye en la gestión de contactos y priorice a los clientes con mayor probabilidad de conversión. Para lograrlo, es necesario entrenar un modelo de inteligencia artificial basado en los registros históricos de la campaña, utilizando herramientas de Azure Machine Learning y aplicando las mejores prácticas en el diseño y procesamiento de datos.

Diseño de Datos para el Entrenamiento del Modelo

Selección de variables relevantes: De los registros disponibles, se deben identificar las variables que más influyen en la conversión. Por ejemplo: número de intentos de contacto previos, edad del cliente, tipo de seguro consultado, fecha del último contacto, región, entre otras.

Limpieza de datos: Se eliminarán registros duplicados, se imputarán valores faltantes y se corregirán inconsistencias en los formatos. Esto asegura que los datos no introduzcan ruido en el aprendizaje del modelo (Goodfellow et al., 2016).

Etiquetado supervisado: Cada registro debe tener una etiqueta binaria ("venta realizada" vs. "venta no realizada"), lo cual es necesario para modelos de clasificación supervisada.

Balaneo de clases: Dado que solo el 9% de los casos resultaron en ventas, el dataset está desbalanceado. Para evitar que el modelo aprenda a predecir mayoritariamente "no venta", se aplicarán técnicas como sobre muestreo de la clase minoritaria (SMOTE) o submuestreo de la clase mayoritaria (Domingos, 2012).

Partición de datos: El conjunto se dividirá en subconjuntos de entrenamiento (70%), validación (15%) y prueba (15%) para asegurar una evaluación justa y objetiva del rendimiento del modelo.

Uso de Azure Machine Learning y Herramientas Complementarias

El entrenamiento del modelo se realizará en Azure Machine Learning, aprovechando su capacidad de gestionar experimentos, registrar métricas, automatizar pipelines y reproducir resultados. A continuación, se describe la arquitectura aplicada:

Azure Data Lake Storage: Usado para almacenar los 8,000 registros en un formato estructurado y accesible para procesos automatizados.

Azure AutoML: Utilizado para probar múltiples algoritmos con foco en árboles de decisión y redes neuronales para seleccionar automáticamente el modelo con mejor precisión y capacidad de efectividad.

Toma de Decisiones Basadas en el Modelo Entrenado de análisis de datos

Una vez entrenado el voicebot, este será capaz de predecir la probabilidad de ser efectivo en cada contacto, permitiendo priorizar los registros con mayor propensión de conversión. De esta forma en lugar de gestionar los contactos de forma aleatoria sin estrategia, el sistema indicara cuáles son los clientes con mayor capacidad e intensidad de compra para enfocar la gestión ahí primero.

Esto optimiza dos aspectos importantes:

- Contactabilidad de la base: al enfocarse en clientes con mayor disponibilidad o interés, aumenta la probabilidad de contacto con cliente (alo)

- Efectividad: el voicebot al haber aprendido de los patrones históricos, va a indicar la mejor estrategia (canal, speech) para aumentar la probabilidad de venta.

Además, mediante Azure, este modelo puede hacer mach con canales automatizados (correo, WhatsApp, llamadas) para aumentar los recorridos de la base y el número de interacciones con cliente sin aumentar los costos.

Resultados Proyectados

Si se implementa un sistema basado en IA de priorización con una efectividad del 70% para identificar prospectos efectivos (propensión alta), se podrían mejorar las tasas actuales de la siguiente forma:

De los 8,000 contactos, el modelo podría identificar los 1,800 más propensos.

Con una nueva contactabilidad efectiva del 60% mejorando la estrategia de contacto y los toques a las propensiones más altas y una tasa de efectividad mejorada al 12%, se lograrían aproximadamente 140 ventas adicionales, mejorando los resultados actuales con menos esfuerzos.

Tabla 1

Proyección de resultados tras implementación de modelo de IA en campaña

comercial

Indicador	Situación actual	Proyección con IA	Variación estimada
Total de registros en la base	8.000	8.000	–
Contactos efectivos	3.200 (40%)	4.800 (60%)	+1.600 contactos
Tasa de efectividad	9%	12%	+3 puntos porcentuales
Ventas logradas	288	432	+144 ventas
Costo de personal adicional	Alto (si se escala)	Nulo (automatización)	Reducción significativa
Intervención humana	Alta	Baja	Reducción operativa

Conclusiones.

La etapa de procesamiento de datos representa la parte más esencial en el diseño de soluciones basadas en inteligencia artificial, particularmente en entornos comerciales donde las decisiones deben ser ágiles, precisas y dinámicas. En el análisis realizado, la adecuada organización de un conjunto de 8.000 registros facilitó la estructuración de la información necesaria para entrenar un modelo con focos predictivos.

El modo en que los datos son preparados influye directamente en el rendimiento del voicebot. Procesos como la depuración, estandarización y elección adecuada de variables contribuyen de forma importante a mejorar la identificación de patrones y tendencias relacionados con la efectividad de ventas. Esto permite enfocar los esfuerzos de contacto hacia los perfiles más propensos a responder, elevando así la eficiencia y rentabilidad de las campañas.

La adopción de herramientas disponibles en la nube de Microsoft, como Azure constituye una plataforma integral y flexible para gestionar el ciclo de vida completo del modelo que se requiere.

Incorporar un sistema predictivo en campañas de ventas representa un cambio notable en los resultados. A través de una priorización inteligente de contactos, se abre la posibilidad de duplicar las ventas alcanzadas sin necesidad de incrementar la carga operativa (ahorro en costos de nómina) ni el número de interacciones.

Por último, resulta indispensable promover un uso ético y responsable de la inteligencia artificial. Esto incluye velar por la equidad en las decisiones, prevenir sesgos y proteger la privacidad de los datos (habeas data). Las funcionalidades de cumplimiento normativo incluidas en la infraestructura de Azure brindan respaldo para operar bajo estos principios fundamentales.

Referencias

- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87.
<https://doi.org/10.1145/2347736.2347755>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Microsoft. (2023a). What is Azure Machine Learning? Microsoft Learn.
<https://learn.microsoft.com/en-us/azure/machine-learning/overview-what-is-azure-ml>
- Microsoft. (2023b). Introduction to Azure Data Lake Storage Gen2. Microsoft Learn. <https://learn.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-introduction>

- Microsoft. (2023c). What is Azure Databricks? Microsoft Learn.
<https://learn.microsoft.com/en-us/azure/databricks/what-is-azure-databricks>
- Microsoft. (2023d). Responsible AI principles from Microsoft. Microsoft AI. <https://www.microsoft.com/en-us/ai/responsible-ai>
- Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
- Zhang, Y., Chen, X., & Qiu, M. (2020). Data preparation for machine learning: A hands-on approach. Springer.