

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Posgrados

**Optimización del proceso de reedición de literatura clásica con LLMs:
Un estudio de caso de Publicaciones Educativas Ariel**

Jonathan Alexander Tayupanta Cárdenas

**Msc. Julián Maya Castro
Director de trabajo de titulación**

Trabajo de titulación de posgrado presentado como requisito
para la obtención del título de Máster en Gerencia de Datos y Negocios

Quito, 13 de julio de 2024

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ
COLEGIO DE POSGRADOS

HOJA DE APROBACIÓN DE TRABAJO DE TITULACIÓN

**Optimización del proceso de reedición de literatura clásica con LLMs:
Un estudio de caso de Publicaciones Educativas Ariel**

Jonathan Alexander Tayupanta Cárdenas

Nombre del director del programa: Julián Maya
Título académico: MSc
Director del programa de: Titulación

Nombre del decano del Colegio Académico: Ana María Novillo
Título académico: PhD
Decano del colegio: Decana CADE

Nombre del decano del Colegio de Posgrados: Hugo Burgos
Título académico: PhD

Quito, julio 2024

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombre del estudiante: Jonathan Alexander Tayupanta Cárdenas

Código de estudiante: 00338055

C.I.: 172179438-4

Lugar y fecha: Quito, 12 de julio de 2024

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017). Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETHeses>.

UNPUBLISHED DOCUMENT

Note: The following graduation project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017). Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETHeses>.

AGRADECIMIENTOS

A mi padre, quien, a través de años de arduo trabajo, ha podido sacar adelante una empresa editorial en un país donde los productos culturales son tan poco valorados. Él ha sido el más claro ejemplo que, a pesar de la adversidad y obstáculos que te presenta la vida, si hay constancia, esfuerzo, inteligencia y visión, es posible sobresalir en un mundo egoísta y competitivo.

A mi madre, quien, con su serenidad y prudencia ha sido un pilar fundamental en el negocio y la familia. De igual manera, por su continuo soporte en todos mis proyectos y decisiones; por sus enseñanzas; por su incondicionalidad, y el amor que propaga.

Finalmente, a mi hermano, mi instructor y gran influencia intelectual. Gracias a él, con sus ilustraciones y ejemplo, he podido formarme en varios campos de pensamiento como historia, política y filosofía.

RESUMEN

En este proyecto de titulación se explora la optimización del proceso de reedición de literatura clásica ecuatoriana y universal, utilizando grandes modelos de lenguaje natural (LLMs) basados en *transformers*, específicamente GTP-4o de OpenAI. La justificación radica en la necesidad de superar las limitaciones de los softwares tradicionales de reconocimiento óptico de caracteres (OCR) para realizar una transcripción inteligente basada en contextos y, a la vez, correcciones ortotipográficas profundas.

El libro que se ocupó en este estudio fue *Cosas de mi tierra* (1929) de José Antonio Campos, el cual presenta un desafío particular debido a los términos vernáculos y modismos ecuatorianos, así como a los errores ortográficos de la edición y aquellos intencionales que tratan de replicar la fonología costeña.

El uso de LLMs y la inteligencia artificial no solo mejora la calidad de las transcripciones, sino que también genera una ventaja competitiva de esta empresa en la industria editorial ecuatoriana. De la misma manera, la implementación de estos modelos permitirá que el patrimonio literario del país pueda preservarse y difundirse eficientemente desde una iniciativa privada.

Palabras clave: aprendizaje profundo, visión por computador, procesamiento de lenguaje natural, *transformers*, industria editorial, edición, OCR, Publicaciones Educativas Ariel

ABSTRACT

This work explores the optimization of the re-editing process of Ecuadorian and universal classic literature using large language models (LLMs) based on transformers, specifically GPT-4o from OpenAI. The justification lies in the need to overcome the limitations of traditional optical character recognition (OCR) software by performing intelligent transcription based on context and deep proofreading.

The book used in this study was “Cosas de mi tierra” (1929) by José Antonio Campos, which shows a particular challenge due to vernacular terms and Ecuadorian idioms, as well as orthographic errors in the edition and intentional ones that replicate the coastal phonology.

The use of LLMs and artificial intelligence not only improves the quality of transcriptions but also generates a competitive advantage for this company in the Ecuadorian publishing industry. Likewise, the implementation of these models will allow the country's literary heritage to be efficiently preserved and disseminated through a private initiative.

Palabras clave: Deep Learning, Computer Vision, Natural Language Processing, Transformers, Publishing, OCR, Publicaciones Educativas Ariel

Agradecimientos	5
Resumen	6
Abstract	7
Introducción	12
CAPÍTULO 1. CONTEXTUALIZACIÓN	14
1.1. Antecedentes y contexto	14
Aporte de Publicaciones Educativas Ariel al desarrollo de la industria editorial en Ecuador	14
Publicaciones Educativas Ariel en la actualidad.....	16
1.2. Justificación del proyecto	18
1.3. Definición del problema	18
1.4. Objetivos	19
Objetivo general.....	19
Objetivos específicos.....	19
CAPÍTULO 2. ANÁLISIS ESTRATÉGICO DE LA EMPRESA	20
2.1. Misión	20
2.2. Visión	20
2.3. Valores	20
2.4. Objetivos estratégicos de la empresa	21
2.5. Diagnóstico y análisis de la estrategia general	22
2.6. Diagnóstico y análisis de la estrategia de analítica de datos y transformación digital	25
.....	25
CAPÍTULO 3. REVISIÓN DE LA LITERATURA	27
3.1. Evolución del reconocimiento óptico de caracteres (Optical Character Recognition)	27
.....	27
3.2. Inteligencia artificial	28
3.3. Aprendizaje automático y profundo	30
3.4. Redes neuronales recurrentes (RNN) y convolucionales (CNN)	31
3.5. Procesamiento de lenguaje natural (NLP)	32
3.6. Redes neuronales de atención y transformers	34
3.7. Funcionamiento de <i>transformers</i>	34
3.8. Uso de una interfaz de programación de aplicaciones (API)	36
CAPÍTULO 4. IMPLEMENTACIÓN DEL MODELO	37
4.1. Datos	37
4.2. Modelo GPT-4o de OpenAI	38
4.3. Tiempo de procesamiento	39
4.4. Resultados y limitaciones del modelo	40
4.4. Comparación de tiempos de edición	41
CAPÍTULO 5. ANÁLISIS FINANCIERO	44
5.1. Estimaciones de egresos e ingresos	44
Costos operacionales	44
Costos no operacionales	46
Ingresos operacionales	46
Ingresos no operacionales.....	47
5.2. Métodos de evaluación: VAN (valor actual neto) y TIR (tasa interna de retorno) ...	47
5.3. Punto de equilibrio y análisis de sensibilidad	49
CAPÍTULO 6. CONCLUSIONES Y RECOMENDACIONES	51
6.1. Conclusiones	51

6.2. Recomendaciones	53
6.3. Impacto del proyecto en la empresa	54
Referencias.....	55
Anexos.....	57
Anexo A: Errores de transcripción comunes de un OCR.....	57
Anexo B: Comparación entre una imagen y el texto producido por GPT 4o	59

ÍNDICE DE TABLAS

Tabla 1. Análisis FODA.....	23
Tabla 2. Costos fijos para la reedición 100 títulos (8 meses).....	45
Tabla 3. Ingresos mensuales estimados a partir del noveno mes	47
Tabla 4. Flujo de caja proyectado para dos años.....	47
Tabla 5. Análisis de sensibilidad.....	50

ÍNDICE DE FIGURAS

Figura 1. Los seis niveles de la lingüística.....	33
Figura 2. Inferencia de un token de un modelo de GPT de OpenAI.....	35
Figura 3. Fluctuación del tiempo de procesamiento por página	40

INTRODUCCIÓN

La edición de libros escaneados —de los que no se dispone de archivos editables— representa un desafío para las editoriales, especialmente cuando estos son antiguos. Las manchas, páginas deterioradas y fallas de impresión dificultan el reconocimiento óptico de caracteres (OCR), resultando en transcripciones inexactas. Incluso si el software OCR no es conservador, puede realizar inferencias de palabras, pero, cuando se trabaja con literatura costumbrista y realista, que utilizan muchos términos vernáculos, modismos y tratan de representar de manera escrita el habla de regiones particulares, los resultados no son óptimos. En pocas palabras, el tiempo de edición es largo y cada libro debe ser revisado múltiples veces por más de un editor.

De esta manera, el siguiente trabajo de titulación propone utilizar un modelo de aprendizaje profundo basado en *transformers* —que actualmente se encuentra disponible en el mercado, especialmente para desarrolladores— para implementar transcripciones inteligentes en Publicaciones Educativas Ariel, una editorial ecuatoriana con una rica historia en la publicación de literatura ecuatoriana y universal.

Para ello, se propone realizar el modelo para el libro de literatura costumbrista *Cosas de mi tierra* de José Antonio Campos, perteneciente a una insigne colección de los años setenta de la editorial: la Biblioteca de Autores Ecuatorianos de Clásicos Ariel. Con estos resultados, se planteará un proyecto financiero para implementar la digitalización de alrededor de 100 títulos la colección mencionada y de otra de clásicos universales.

En este sentido, con la finalidad de desarrollar el objetivo principal de esta investigación —realizar una transcripción eficiente y de calidad con inteligencia artificial—, se estructuró este proyecto en seis secciones: contextualización; análisis estratégico de la empresa; revisión de la

literatura; implementación del modelo; análisis financiero, y, finalmente, conclusiones y recomendaciones.

En el primer capítulo, se contextualizará sobre la historia de Publicaciones Educativas Ariel y Radmandí, y su relevancia en el mercado editorial ecuatoriano; de la misma manera, se justificará por qué la implementación de este proyecto aportará en el proceso productivo editorial y se desarrollarán explícitamente el objetivo principal y los específicos. En la segunda parte, bajo esta misma línea, se realizó un análisis estratégico de la empresa, que permitirá conocer su razón de ser y realidad interna.

En los tres siguientes capítulos, se aborda la parte técnica. Se revisó la literatura que está detrás de los grandes modelos de lenguaje natural: el amplio espectro de la inteligencia artificial, el aprendizaje automático y profundo, y la revolución de las redes neuronales de atención y *transformers*. En este sentido, se presenta la implementación del modelo GPT-4o a través de su interfaz de programación de aplicaciones (API), en donde se muestra la naturaleza de los datos y se analizan los resultados, comparando los tiempos de edición tradicionales con esta propuesta de transcripción inteligente. Los resultados son sobresalientes. También, se realiza un análisis financiero para evaluar su viabilidad económica.

Finalmente, en el capítulo seis, se presentan las principales conclusiones y recomendaciones.

CAPÍTULO 1. CONTEXTUALIZACIÓN

1.1. Antecedentes y contexto

Publicaciones Educativas Ariel es una marca editorial ecuatoriana fundada, en 1963, por el guayaquileño Tomás Rivas Mariscal, quien, en un principio, publicó textos escolares y, a finales de los sesenta, literatura. Las tres colecciones más conocidas de esta editorial fueron la Biblioteca de Autores Ecuatorianos, cien títulos rigurosamente escogidos por Hernán Rodríguez Castelo de la literatura más trascendente del Ecuador desde la Colonia hasta mediados del siglo XX; Ariel Universal, cien títulos de clásicos universales, con dirección de otro gran literato, Rafael Díaz Ycaza, y, finalmente Ariel Juvenil Ilustrada, a finales de los años setenta, a cargo del mismo poeta, que la componen otros cien títulos adaptados e ilustrados de literatura universal. En el año 2006, Lucas Marcelo Tayupanta adquirió la marca y los derechos intelectuales de estas tres colecciones. Por otro lado, Radmandí Proyectos Editoriales, fundada por este último, nació en 1993 y se dedicó a la publicación de textos escolares y literatura contemporánea. Actualmente, las dos marcas forman parte de la misma persona natural.

Aporte de Publicaciones Educativas Ariel al desarrollo de la industria editorial en Ecuador

Durante la década de los sesenta, Publicaciones Educativas Ariel comienza a ganar reconocimiento en el mercado por sus llamativas publicaciones y «su didáctico y valioso contenido» (El Universo, 2008). Entre sus primeros lanzamientos, se encuentran *Así nació Ecuador* y *Conozca Ecuador*, elaboradas por el historiador Francisco Huerta Rendón, y la colección «Compañerito», textos integrales de Educación Básica. A finales de aquella década, Tomás Rivas Mariscal, influenciado con el éxito de la venta de literatura clásica por colecciones, como en España y en otros países de Latinoamérica, lanza la Biblioteca de Autores Ecuatorianos de Clásicos Ariel.

Los primeros libros de la primera colección se comenzaron a vender a 12 sucres, un precio muy competitivo y accesible para la familia promedio si se compara con las ediciones de literatura ecuatoriana disponibles en aquellos años (El Universo, 2012).¹ El proyecto fue un éxito total. Por primera vez, en el siglo XX, una verdadera industria editorial se instauraba en el país, pues, anteriormente, la oferta de literatura se limitaba a pocas imprentas, universidades y al joven legado de Benjamín Carrión: la Casa de la Cultura Ecuatoriana, fundada en 1944.² De la misma manera, la escasa demanda se condicionaba a selectos grupos literarios, intelectuales y académicos.

De esta forma, la persona común, con el propósito de armar su primera biblioteca en su hogar, comenzó a adquirir periódicamente los libros de Clásicos Ariel. La clandestina literatura ecuatoriana salió a la luz y despertó interés general. El historiador Rodolfo Pérez Pimentel (1988) afirma que los primeros volúmenes aparecieron primero en la Sierra y posteriormente en la Costa, «tímidamente al principio y luego en forma triunfal, pues el país entero comprendió que se trataba del esfuerzo editorial más importante del siglo». Se hicieron tirajes iniciales de veinte mil por libro. Al poco tiempo, varias reimpressiones de igual o mayor tiraje fueron necesarias.

Al alcanzar el título número cien de la colección, por el aporte cultural de este emprendimiento, Hernán Rodríguez Castelo (n.d.-b) menciona que el mismo presidente Velasco Ibarra expresó: «Ustedes han llevado la cultura al pueblo». De igual manera, en 1973, el Ministerio de Educación otorgó a la editorial la Orden al Mérito Educacional de primera clase. También, en la Comisión Internacional del Año del Libro del mismo año, se reconoció

¹ El salario mínimo vital entre 1971 y 1972 era de 750 sucres (Pazmiño, 2010).

² Hernán Rodríguez Castelo afirma que «... se hacían tirajes de cuatrocientos o quinientos y no se vendía nada» (El Universo, 2012).

... que la empresa editorial Ariel ha cumplido de modo ejemplar con los ideales propuestos para este año, lanzando en millares de ejemplares, al alcance de las clases populares, las obras mayores de la cultura y la literatura nacionales, lo cual ha significado el comienzo de una nueva etapa en la historia del libro ecuatoriano» (Hernán Rodríguez Castelo, n.d.-b).

El impacto no se limitaría únicamente a Ecuador. En 1973, la editorial publicaría la colección Ariel Universal, lo que le llevaría a conquistar el mercado regional, especialmente, en Colombia y Perú. Estos libros tuvieron gran acogida internacionalmente, y algunas generaciones se formaron y se identificaron con ellos. Ya en los años ochenta, nacería otra colección más popular: Ariel Juvenil Ilustrada.

Con el éxito de Publicaciones Educativas Ariel, la industria editorial nacional comenzó a desarrollarse tanto para literatura y textos escolares. Nuevos competidores emergieron y, lastimosamente, en los noventa, debido a la falta de innovación, la marca evidenció su decaimiento.

Publicaciones Educativas Ariel en la actualidad

En el año 2006, Tomás Rivas Mariscal vendió la marca editorial con sus tres colecciones insignes a Lucas Marcelo Tayupanta, quien fue uno de sus principales distribuidores en los noventa y en los dos mil a través de la compañía limitada CODISLISI (Comercial Distribuidora de Libros el Siglo, que en su mejor momento tuvo alrededor de 80 empleados). Es menester mencionar que Lucas Marcelo Tayupanta pudo capitalizar y comprar la marca con los beneficios económicos de esta compañía y de su otra empresa Radmandí Proyectos Editoriales, que se dedicaba principalmente a la edición de textos escolares para Educación General Básica.

Sin embargo, en aquellos años, con el gobierno de Alfredo Palacios, se institucionalizó la entrega gratuita de textos escolares a nivel nacional. Esta política lapidó a las editoriales pequeñas en el país e instauró a otras que pudieron ganar millonarias licitaciones para ofertar

los nuevos libros a todo el sistema educativo fiscal del país. De esta manera, CODISLISI no pudo sobrevivir a este *shock* y se la liquidó en el año 2007. Es decir, Lucas Marcelo Tayupanta se quedó únicamente con dos marcas editoriales: Radmandí y Publicaciones Educativas Ariel, disminuyendo drásticamente el tamaño de la actividad económica en la que se desenvolvía.

Esta grave crisis ocasionó que la reedición de todos los libros de la marca recién comprada no se haya podido materializar en su totalidad. Igualmente, por la limitación de recursos y por ciega confianza, se contrató a un director editorial que se encargó de la publicación de algunas series de textos escolares y varios libros de literatura de Publicaciones Educativas Ariel. Algunos de estos productos no fueron de la calidad esperada. Empero, desde el 2018, con la salida de esta persona, la marca ha hecho grandes esfuerzos por mejorar sus ediciones y sacar productos de calidad.

Por otro lado, la dinámica del mercado no ha sido la misma y la factibilidad económica para relanzar los libros de la antigua marca ha sido determinada principalmente por la demanda de escuelas y colegios, mas no por el consumo del mercado no educativo.

Actualmente, se ha reeditado la totalidad de la colección Ariel Juvenil Ilustrada (101 títulos), 21 libros de Ariel Clásicos Universales y 20 de Ariel Clásicos Ecuatorianos. En este sentido, considerando ciertos libros que no se pueden publicar por no estar en dominio público, faltan de reeditar alrededor de 100.

Actualmente, Lucas Marcelo Tayupanta es una persona natural y tiene como actividades principales la edición, diseño, impresión, distribución y comercialización de libros. La pequeña empresa cuenta con tres personas de planta en el área editorial: dos editores y una diseñadora. Sin embargo, cuando se presentan nuevos proyectos en los que no puede trabajar su personal de planta, se externaliza la edición y diseño. Las otras personas de la empresa están distribuidas en las áreas administrativa, ventas, bodega e imprenta. En total, la empresa tiene 14 empleados.

1.2. Justificación del proyecto

La reedición de literatura clásica es relevante cultural y educativamente, especialmente en un país como Ecuador, donde la difusión y preservación del patrimonio literario enfrenta desafíos considerables.

Publicaciones Educativas Ariel, con más de seis décadas en este mercado, ha jugado un papel crucial en la promoción de autores ecuatorianos y en la formación de lectores nacionales e internacionales. Sin embargo, la digitalización y posterior reedición, que incluye alrededor de 100 títulos, se ha visto obstaculizada por problemas técnicos y económicos.

1.3. Definición del problema

Publicaciones Educativas Ariel enfrenta problemas significativos con la digitalización de su colección de clásicos ecuatorianos y universales debido a la calidad física deteriorada de los libros originales. Las manchas, páginas amarillentas —con sombras en escala de grises— y fallas de impresión dificultan el reconocimiento óptico de caracteres (OCR), resultando en transcripciones inexactas, especialmente cuando se encuentra lenguaje vernáculo y ecuatorianismos que no están en diccionarios estándar.³ Igualmente, por sus recursos económicos limitados, no existe la cantidad suficiente de editores para relanzar estos libros.

El problema de digitalización afecta la capacidad de la editorial para reeditar libros con precisión y eficiencia; mantener la calidad y la integridad del contenido original; aumentar la oferta de libros digitales en plataformas como KDP y iBooks, y satisfacer la demanda del mercado, especialmente en sectores educativos que dependen de estos textos. Adicionalmente, dada la inversión que se hizo en la compra de las marcas y de los derechos de edición de los libros, muchos de estos activos son ociosos.

³ En el anexo A, se podrá encontrar una muestra de los errores de transcripción más comunes utilizando un software OCR.

1.4. Objetivos

Objetivo general

Realizar una transcripción inteligente del libro *Cosas de mi tierra* (1929) de José Antonio Campos, utilizando el último modelo de procesamiento de lenguaje natural de OpenAI.

Objetivos específicos

1. Comparar los tiempos de reedición tradicional con la nueva propuesta de aplicación de modelos de aprendizaje profundo.
2. Realizar un análisis de viabilidad económica para implementar este proceso para la reedición de los libros de Ariel Clásicos Ecuatorianos y Ariel Clásicos Universales.

CAPÍTULO 2. ANÁLISIS ESTRATÉGICO DE LA EMPRESA

Con la finalidad de analizar la razón de ser, los principios, los objetivos y el entorno de esta actividad económica, en esta sección, se presentará un resumen del análisis estratégico de la empresa.

2.1. Misión

Poner al alcance del consumidor libros educativos y literarios asequibles y de alta calidad que contribuyan al desarrollo intelectual y cultural de estudiantes y lectores en Ecuador y América Latina. Buscamos innovar y adaptarnos a las nuevas tecnologías para ofrecer productos tanto en formato físico como digital.

2.2. Visión

Ser una editorial líder en la producción y distribución de materiales educativos y literarios en América Latina, reconocida por la excelencia de contenidos, el compromiso con la educación y cultura, y la capacidad de adaptarse a los cambios tecnológicos y del mercado.

2.3. Valores

- **Excelencia:** compromiso con la calidad y mejora continua en productos educativos.
- **Innovación:** adaptación constante a las nuevas tecnologías y metodologías educativas.
- **Sostenibilidad:** compromiso con el medio ambiente mediante la implementación de prácticas sostenibles en la producción y adquisición de insumos.
- **Colaboración:** fomento de un ambiente laboral colaborativo que promueva el desarrollo profesional y personal de los empleados.

- **Integridad:** prácticas éticas y transparentes en todas las operaciones.

2.4. Objetivos estratégicos de la empresa

1. Reposicionar las colecciones clásicas.

Meta: Atraer nuevas generaciones de lectores.

Estrategia: Actualizar el diseño y contenido de las colecciones clásicas.

2. Incrementar la oferta de libros digitales.

Meta: Aumentar la presencia digital en plataformas como Kindle Direct Publishing (KDP), iBooks y Blink Learning

Estrategia: Cada libro reeditado deberá publicarse en no más de dos semanas después de su diagramación en las plataformas digitales antes mencionadas.

3. Optimizar los costos de producción.

Meta: Reducir costos mediante la innovación en la edición e impresión bajo demanda.

Estrategia: Optimizar el proceso de reedición con un buen software OCR e implementar un sistema de impresión bajo demanda para libros de baja rotación, reduciendo inventarios y costos de almacenamiento. Es decir, evitar la producción offset para productos con bajo movimiento comercial.

4. Expandir la presencia internacional.

Meta: Aumentar las ventas en mercados internacionales.

Estrategia: Fortalecer las alianzas con los distribuidores internacionales actuales y generar nuevas, participando en ferias del libro.

5. Capacitación y desarrollo del personal.

Meta: Mejorar la competencia y eficiencia del equipo.

Estrategia: Implementar programas de capacitación continua en nuevas tecnologías y tendencias editoriales, y fomentar un ambiente de trabajo colaborativo. En el caso de los

editores, deberán estar capacitados con las últimas normas ortográficas, el estilo y la gramática de la Real Academia Española (RAE).

6. Implementar prácticas sostenibles

Meta: Adoptar prácticas ecológicas y sostenibles en la producción.

Estrategia: Utilizar tintas y químico certificados con bajo o nulo impacto ambiental para el proceso de impresión.

7. Responsabilidad social y cultural

Meta: Promover la lectura y el acceso a la educación de calidad.

Estrategia: Desarrollar programas de donación de libros y organizar actividades educativas en comunidades desfavorecidas.

2.5. Diagnóstico y análisis de la estrategia general

El mercado de editorial ecuatoriano en la actualidad

Según los datos de registro ISBN (International Standard Book Number) de la Cámara Ecuatoriana de Libro (CEL) del año 2023, la mayor parte de la producción nacional de libros recae en la categoría de interés general⁴, seguido por los didácticos⁵ y finalmente por los científicos, técnicos y didácticos⁶. En cuanto a la literatura ecuatoriana, juntamente con la infantil, son los rubros más importantes de la primera categoría. Es menester considerar que la mayor parte de los títulos nuevos registrados son para libros físicos (de tapa blanda o dura) y el 45 % son digitales; este promedio ha mantenido estable desde el 2020, pues, a raíz de la pandemia, en comparación con el 2019, la edición de libros digitales se duplicó (Cámara Ecuatoriana del Libro, 2023).

⁴ Obras de referencia y consulta; literatura infantil, juvenil y adulta; filosofía; psicología; generalidades, etc.

⁵ Categoría compuesta principalmente por textos escolares para Educación General Básica y Bachillerato.

⁶ Libros técnicos, universitarios y especializados.

La oferta de este mercado está compuesta por pequeñas y medianas empresas, unas pocas grandes nacionales, y principalmente por editoriales transnacionales. Además, este mercado depende de la importación de materias primas —principalmente papel, químicos y tintas— y también de la oferta internacional —libros terminados—, especialmente de Reino Unido y España (ICEX España Exportación e Inversiones, 2023).

En cuanto a la demanda, según la Encuesta de Hábitos Lectores, Prácticas y Consumos Culturales realizada por el Ministerio de Cultura y Patrimonio en 2022, el 56.7 % de la población lee libros. En promedio, en el país, se lee un libro completo y dos incompletos al año. La mayor parte de este consumo se concentra entre las personas de 5 a 18 años, es decir, estudiantes de Educación General Básica y Bachillerato General Unificado. A partir de los 18 años, el 50 % manifiestan leer libros (ICEX España Exportación e Inversiones, 2023).

En síntesis, el mercado de literatura en Ecuador es pequeño, pero ha mostrado una tendencia de crecimiento especialmente en la categoría de interés general. Publicaciones Educativas Ariel tiene una posición consolidada, pero enfrenta grandes retos debido a la competencia y la necesidad de modernización.

Tabla 1. Análisis FODA

Fortalezas	Oportunidades
<ul style="list-style-type: none"> • Historia y reputación en el mercado editorial. • Tres grandes colecciones de literatura universal y ecuatoriana. • Amplio conocimiento en ortotipografía y las normas más actuales de la Real Academia 	<ul style="list-style-type: none"> • Avances tecnológicos en OCR y NLP que pueden mejorar la eficiencia en el proceso de digitalización. • Uso de inteligencia artificial en softwares de diseño gráfico como Adobe Indesign, Photoshop e Illustrator.

<p>Española (RAE), en cuanto estilo, gramática y ortografía.</p> <ul style="list-style-type: none"> • Uso paulatino de inteligencia artificial. • Experiencia con el uso de plataformas de autopublicación. • Impresión bajo demanda para libros con baja rotación. 	<ul style="list-style-type: none"> • Uso de herramientas de inteligencia artificial, como Midjourney, que pueden generar imágenes e ilustraciones para libros, evitando así la contratación de ilustradores. • Coediciones con editoriales extranjeras para generar mayor alcance y abaratar costos. • Generación de audiolibros con inteligencia artificial.
<p>Debilidades</p> <ul style="list-style-type: none"> • Empresa pequeña con recursos limitados para adquirir más editores y diseñadores. • Falta de un departamento de promoción que trabaje en escuelas y colegios. • En imprenta, máquinas antiguas que no son eficientes en términos gráficos actuales. • Gestión de cobranza nacional. • Negocio estacional con demanda no constante. La mayor parte de ventas se dan en abril-mayo (inicio de clases 	<p>Amenazas</p> <ul style="list-style-type: none"> • Competencia de grandes editoriales nacionales y transnacionales. • Cambios en las políticas gubernamentales que pueden afectar la industria editorial. Por ejemplo, el gobierno, al buscar mayores ingresos, puede proponer a la Asamblea la eliminación del beneficio tributario que tienen los libros al no pagar IVA. • Al ser los libros un bien normal con una elasticidad alta, en una economía con crisis, los agentes, al tener menos

<p>de la región Costa y Galápagos) y en agosto-septiembre (región Sierra y Amazonía).</p>	<p>ingresos, bajarán su consumo en este bien.</p> <ul style="list-style-type: none"> • <i>Shocks</i> internos y externos que pueden afectar la oferta de insumos como papel, cartulina, tintas y químicos.
---	---

2.6. Diagnóstico y análisis de la estrategia de analítica de datos y transformación digital

Actualmente, Publicaciones Educativas Ariel y Radmandí no ha implementado una transformación digital en el sentido estricto. Sin embargo, en el área administrativa, dado que el sistema contable es antiguo y no tiene una interfaz amigable para analizar correctamente datos, sí se implementa una conexión ODBC entre su base de datos en Microsoft SQL Server y paquetes estadísticos como R y Python. De esta manera, se analizan ventas, costos, cobranzas, inventario y producción a tiempo real con estadística descriptiva e inferencial.

Con relación al área editorial, la empresa tiene amplia experiencia en la publicación de libros digitales en plataformas como Kindle Direct Publishing y iBooks. También, maneja el software de Blink Learning para la creación de libros interactivos, en donde se pueden leerlos, realizar actividades autocalificables e interactuar con video y audio. Por el lado de diseño, con la finalidad de disminuir costos en ilustradores externos, se ha utilizado inteligencia artificial para crear imágenes, especialmente para portadas de libros de literatura.

En cuanto a edición de textos escolares y guías de lectura, los editores sí utilizan a modelos de NLP para complementar y hacer más eficiente su trabajo de creación de contenido de texto. Este siempre es supervisado y verificado, pues el personal reconoce y ha sido capacitado sobre las limitaciones de esta tecnología.

En esta misma línea, cuando se trabaja en el proceso de reedición de libros de literatura de las colecciones de Publicaciones Educativas Ariel, la editorial utiliza los softwares ABBY Finereader 12 y Adobe Acrobat Pro para leer los documentos escaneados y convertirlos en texto editables; sin embargo, como se ha mencionado anteriormente, sus resultados son muy limitados e inexactos, especialmente con literatura ecuatoriana, que presenta muchos modismos y palabras vernáculas. Por lo tanto, después de tener el texto editable, es necesario de 3 a 4 revisiones por parte de un editor especializado para corregir los libros y proceder, posteriormente, con la diagramación en el área de diseño. Un proceso que es, actualmente, ineficiente.

CAPÍTULO 3. REVISIÓN DE LA LITERATURA

En esta sección, se utilizaron libros, recursos en línea y revistas indexadas reconocidas que explican el desarrollo de la tecnología de reconocimiento óptico de caracteres (OCR). De la misma manera, se expone la evolución y el funcionamiento general de las redes neuronales detrás de los grandes modelos de lenguaje natural como *transformers*. Finalmente, se abordará la literatura relacionada al uso de las interfaces de programación de aplicaciones (API, por sus siglas en inglés).

3.1. Evolución del reconocimiento óptico de caracteres (Optical Character Recognition)

Según IBM (2022), un sistema de reconocimiento óptico de caracteres —que también se lo conoce como reconocimiento de texto— «utiliza una combinación de hardware y software para convertir documentos físicos e impresos a un texto legible por una máquina». Es decir, una cámara o un escáner pueden digitalizar un documento y, posteriormente, con un software OCR, se puede conseguir un archivo editable, evitando la transcripción humana.

Los cimientos de esta tecnología se remontan a principios del siglo XX, en el contexto de la Primera Guerra Mundial y el uso del telégrafo. Emanuel Goldberg, gran físico israelí, inventó un aparato que podía leer caracteres y convertirlos en código telegráfico. Pocos años después, en 1920, este inventor desarrolló la Statistical Machine, un avance sin precedentes para gestión documental. Esta máquina capturaba imágenes de documentos y las almacenaba en forma de *microfilm*, y, rústicamente, se podía realizar búsquedas en ellos. IBM compró años después esta patente (Edmonds, 2020).

Después de hito, «las versiones anteriores de los OCR fueron entrenadas con imágenes de cada carácter y estaban limitadas a reconocer una tipografía a la vez» (Edmonds, 2020). En los años setenta, Ray Kurzweil patentó el *OCR omni-font*, que, como el prefijo *omni* indica, podía reconocer texto impreso en cualquier fuente o tipografía. Con todas las diferentes

aplicaciones de este invento, Kurzweil decidió utilizarlo para comercializar una máquina capaz de leer documentos o libros para ciegos, con la finalidad de disminuir las limitaciones propias de esta discapacidad. Es decir, se utilizó por primera vez una tecnología de aprendizaje automático de *text-to-speech*. Con este avance, Xerox, en 1980, compró la compañía de Kurzweil e implementó este invento para convertir texto de imágenes a un formato entendible para computadoras (IBM, 2022).

A partir de los 2000, el uso del OCR se hizo común en el mercado (Edmonds, 2020). En la actualidad, entre los más utilizados se encuentran Adobe Acrobat Pro, Abby Finereader y Tesseract OCR. De la misma manera, al comprar un escáner o una impresora, este software ya viene integrado con ellos; también, al digitalizar un documento con un *smartphone*, existen herramientas como Google Docs, que permiten convertir imágenes en archivos editables. Por otro lado, Google Cloud Platform ofrece dentro de sus herramientas a Document AI y la API de Cloud Vision —especialmente para desarrolladores—, en el que se integran modelos avanzados de aprendizaje profundo para mejorar los resultados del reconocimiento de caracteres con la comprensión del contexto.

3.2. Inteligencia artificial

Lo que actualmente se conoce como inteligencia artificial ha sido el producto de siglos de desarrollo del discernimiento humano. Muchas ciencias y campos de conocimiento, como la filosofía, la filología, matemáticas, física, economía, psicología, la neurociencia y, por supuesto, la computación, han contribuido para construirla.

Aunque la historia de la inteligencia artificial se remonta a principios del siglo XIX, con la invención del telar de Jacquard y con la máquina analítica de Charles Babbage —que no la construyó pero sí diseñó—, no es sino durante y después de la Segunda Guerra Mundial, con la creación de la máquina enigma y los avances en la ciencia computacional de Alan Turing,

que el verdadero desarrollo de esta ciencia comienza (Grzybowski et al., 2024). La contribución de este último científico es tal que, incluso, para identificar si un programa o computadora está dentro del espectro de la inteligencia artificial, comúnmente, se utilizan la prueba de Turing y su prueba total (Russell & Norvig, 2022).

En el primer caso, se necesita que la computadora pueda comunicarse en lenguaje natural —lenguaje humano—; que pueda representar el conocimiento para almacenar lo que sabe u oye; que tenga la capacidad de razonamiento, y que pueda aprender y adaptarse a circunstancias nuevas: detectar y extrapolar patrones. Para la prueba total, que es un alcance propuesto por otros investigadores, además de los elementos analizados, se necesita que la computadora pueda reconocer el mundo a través de visión (*computer vision*) y reconocimiento de habla (*speech recognition*); finalmente, con la robótica, tener la capacidad de manipular objetos y moverse (Russell & Norvig, 2022).

Por aquellos mismos años, con la influencia de Turing y la neurofisiología, en 1943, Warren McCulloch y Walter Pitts propusieron un modelo de neuronas artificiales. Estas, con una lógica booleana, podían «prenderse» o «apagarse». La activación ocurría cuando recibían un estímulo de un número suficiente de neuronas vecinas. Finalmente, concluyeron que «cualquier función computable podía realizarse por una red de neuronas conectadas y que los conectores lógicos (*AND*, *OR*, *NOT*, etc.) podrían implementarse por estructuras neuronales simples» (Russell & Norvig, 2022).

Con estas bases, un sinnúmero de aportes se realizaron a lo largo de la segunda mitad del siglo xx y la actualidad. Es menester mencionar que, en 1956, un programa comprobó, racionalmente, 38 teoremas matemáticos del libro *Principia Mathematica* de Bertrad Russell y Whitehead, y pudo desarrollar otras pruebas «más elegantes» de algunas de ellas (Grzybowski et al., 2024).

Russell & Norvig (2022) mencionan que los mayores avances en esta ciencia pueden resumirse en los ganadores del Premio Turing. Incluso, desde la literatura de ciencia ficción, con Isaac Asimov, que formuló las tres leyes de la robótica (Grzybowski et al., 2024). También es relevante traer a colación que, en 1956, el término *inteligencia artificial* fue acuñado por primera vez en la Conferencia de Dartmouth por un selecto grupo de científicos (Moor, 2006).

Actualmente, el concepto de inteligencia artificial es muy amplio y abarca un sinnúmero de técnicas que permiten a las computadoras o programas imitar el comportamiento humano; sin embargo, es importante considerar dos áreas principales que están embebidas dentro de ella: *machine learning* y *deep learning* (Goodfellow et al., 2016).

3.3. Aprendizaje automático y profundo

El aprendizaje automático es una rama de la inteligencia artificial que utiliza métodos estadísticos y algorítmicos para que las máquinas puedan aprender y hacer predicciones basadas en datos. Es decir, al disponer de mucha información, pueden tomar decisiones considerando el contexto. Por ejemplo, calcular la probabilidad promedio de fallo de pago de un crédito de una persona que cumple con ciertas características socioeconómicas o categorizar a usuarios de redes sociales según lo que publican.

Es fundamental distinguir la finalidad predictiva de la explicativa de este tipo de aprendizaje con otros campos, como la econometría, que emplean los mismos métodos estadísticos, pero con el propósito de explicar las relaciones entre variables, encontrar causalidad y también el pronóstico. En el caso del *machine learning*, la predicción correcta es lo más valorado (Hair et al., 2019).

Aunque estos métodos estadísticos eran ya conocidos en el siglo XX, con los grandes avances tecnológicos de las últimas décadas, como tarjetas gráficas más potentes, las máquinas

pueden ahora «resolver problemas con conocimiento del mundo real y tomar decisiones que son, al parecer, subjetivas» (Goodfellow et al., 2016).

Según Goodfellow et al. (2016), en el *machine learning* clásico, es necesario que una persona seleccione manualmente las características más relevantes de un conjunto de datos. Sin embargo, es posible que los modelos puedan aprender automáticamente sin la necesidad de que alguien guíe y elija las variables a considerar. Cuando esto sucede, un nuevo concepto es necesario dentro del *machine learning*: el *representation learning*. En él, los modelos pueden descubrir patrones más complejos por sí mismos. Los autores ilustran el caso de un modelo que pueda identificar carros en imágenes, en que la *simple* tarea de distinguir una llanta, que puede tener diferentes tonos, sombras, etc, puede ser tan compleja que resultaría muy demandante que una persona le indique sobre qué características considerar para lograr llegar al resultado. Este tipo de modelos pueden aprender no solo «el mapeo al resultado sino la representación en sí misma».

Finalmente, dentro del *machine* y el *representation learning*, se encuentra a un nivel más profundo, como su nombre indica, el *deep learning*. Este aprendizaje tiene las mismas particularidades del anteriormente revisado, pero con la diferencia de que introduce «representaciones que son expresadas en términos de otras más simples» y con múltiples capas (Goodfellow et al., 2016). Estas últimas son denominadas *redes neuronales*, que son la base del *deep learning*.

3.4. Redes neuronales recurrentes (RNN) y convolucionales (CNN)

Una red neuronal tiene como insumo a distintas variables que interactúan entre sí en distintas capas para proveer un resultado. Bajo este principio, una red neuronal recurrente utiliza los resultados obtenidos como nuevos insumos para poder tener nuevos resultados. Es decir, este método resulta bastante útil en datos secuenciales, en donde los resultados actuales dependen

en gran parte de los sus estados anteriores, de manera análoga cuando se trabaja con series de tiempo. Una gran ventaja de las RNN es que pueden capturar dependencias altamente complejas, lo que resulta ampliamente eficaz para predecir, por ejemplo, precios de *stocks* financieros, secuencia de letras, palabras y oraciones, y reconocimiento de voz (Serrano, 2021). En este sentido, hay algunas variaciones de estas redes como el *Long Short-Term Memory* y el *Gated Recurrent Unit* (GRU) que tienen muy buenos resultados especialmente en estas tareas (Goldberg, 2021).

Para analizar insumos más complejos como imágenes, audio o videos, se utilizan las *redes neuronales convolucionales*. Este último término proviene de una operación matemática denominada *convolución*, que es un tipo de operación lineal (Goodfellow, 2016).

Según IBM, estas redes neuronales aplican tres tipos de capas: convolucionales, de *pooling* y capas completamente conectadas (FC). «Por cada capa adicional, las CNN incrementan su complejidad, identificando mayores porciones de una imagen». Generalmente, las capas iniciales identifican elementos como colores o bordes, y, conforme, avanza, características más complejas como formas de un objeto, hasta que finalmente se lo pueda reconocer.

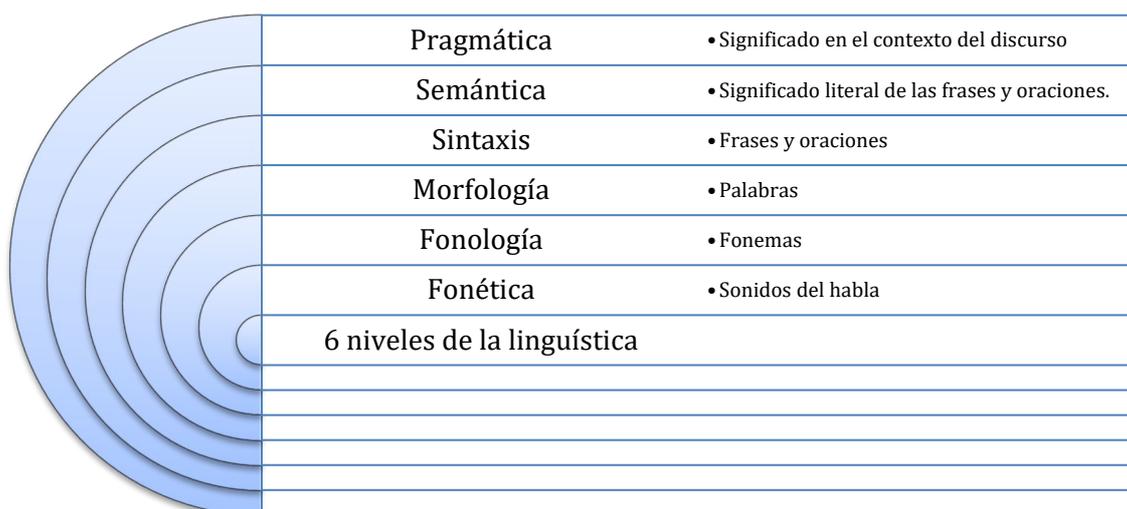
3.5. Procesamiento de lenguaje natural (NLP)

Según Lee (2024), el procesamiento del lenguaje natural (NLP) es un campo de la inteligencia artificial (IA) que permite a las máquinas entender y analizar el lenguaje humano para mejorar la interacción humano-computadora. Entre las principales aplicaciones en el que se utiliza, están la recuperación de información, los resúmenes automáticos, el análisis de sentimientos, el reconocimiento del habla, la generación de texto, traducción, etc. Esta área de la inteligencia artificial utiliza gran parte de los métodos anteriormente descritos, además de

otros campos de conocimiento como la filosofía (lógica), la psicología, la lingüística y, por supuesto, la computación.

Para entender la generación de texto y la comprensión del lenguaje por parte de los humanos y la complejidad de trasladar ese conocimiento a las máquinas desde una perspectiva lingüística, es útil considerar los seis niveles de análisis (figura 1) y las posibles ambigüedades presentes. Por ejemplo, puede haber ambigüedades léxicas⁷, sintácticas⁸, semánticas⁹ y pragmáticas¹⁰ (Lee, 2024). Entenderlas es de suma importancia, pues las máquinas pueden hacer inferencias correctas si analizan por contexto.

Figura 1. Los seis niveles de la lingüística



Fuente: Lee (2024)

⁷ Por ejemplo, la palabra *juego* puede ser un verbo y un sustantivo.

⁸ En «Vi a la persona con binoculares», puede haber dos interpretaciones: el sujeto vio con unos binoculares a esa persona o el sujeto vio a una persona que tenía binoculares.

⁹ En «Pedro está buscando la batería», las interpretaciones posibles son las siguientes: Pedro está buscando el instrumento musical o una batería (pila).

¹⁰ Por ejemplo, en «Voy a salir a correr», no se conoce el contexto, si la persona va a salir a correr para una competencia deportiva o solo para ejercitarse normalmente.

3.6. Redes neuronales de atención y transformers

Las redes neuronales recurrentes y de convolución dominaron el campo del procesamiento de lenguaje natural (NLP) por muchas décadas. A pesar de tener muy buenos resultados, alcanzaban rápidamente su límite cuando se enfrentaban a secuencias grandes y muchos parámetros (Rothman, 2024).

Sin embargo, en 2017, con la publicación de *Attention is all you need*, un nuevo método revolucionó el NLP y los avances futuros de la inteligencia artificial: la aplicación de *transformers* con el uso de mecanismos de atención. Esta arquitectura es la base actual de los principales modelos grandes de lenguaje natural como Chat-GPT, de OpenAI, y Gemini, de Google.

La principal diferencia de este nuevo enfoque con las redes neuronales recurrentes es que «*transformers* no analiza tokens en secuencias, pero sí cada token con otros en una secuencia» (Rothman, 2024). Un token es la unidad mínima dentro del contexto de NLP, puede que sea una palabra o incluso una parte de ella.

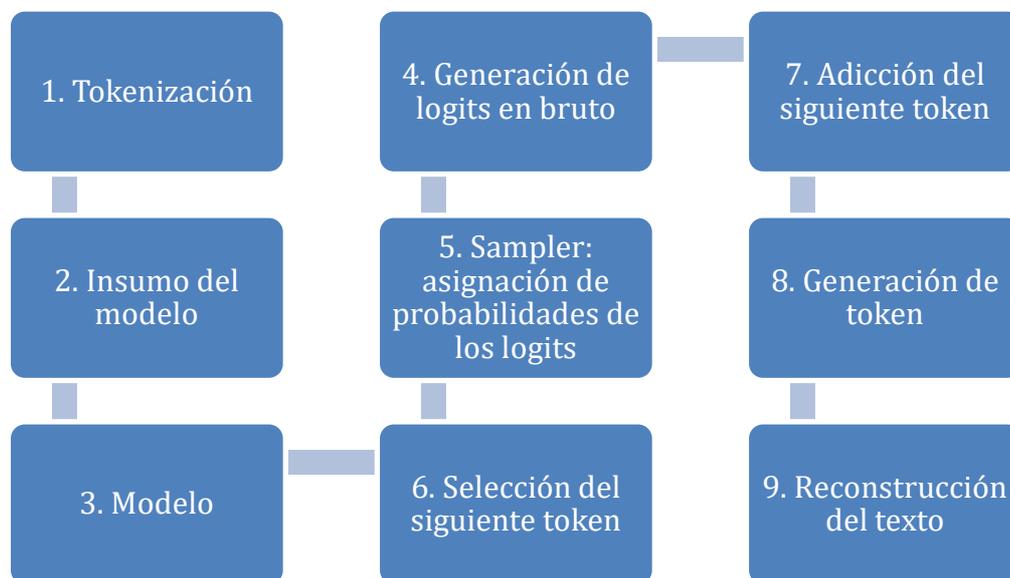
Desde 2017, con el modelo original de *transformers*, que tiene alrededor de 65 millones de parámetros, con los datos más recientes, se conoce que el modelo GPT-3.5 tiene aproximadamente 175 billones de ellos. Esto refleja que, conforme se van complejizando y desarrollando estos modelos, van adquiriendo mayor capacidad de generación de contenido al estar más contextualizados y descubrir relaciones más complejas en los datos de los que se alimentan. Actualmente, hay escasa información sobre cuántos parámetros tiene el nuevo modelo GPT-4o (Rothman, 2024).

3.7. Funcionamiento de *transformers*

Rothman (2024) explica el proceso utilizado de los modelos generativos basados en *transformers* en nueve pasos. En el primero, se recibe una secuencia de palabras —una oración,

por ejemplo— y se separara en tokens, las unidades mínimas. A este proceso se lo denomina *tokenización*. Después, se recibe la secuencia tokenizada en el modelo *Generative Pretrained Transformer* (GPT). En el tercer paso, el modelo procesa los insumos a través de distintas capas de *transformer* hasta la de salida. Posteriormente, se genera *logits*¹¹ en bruto. En el quinto paso, cada *logit* es asignado con una probabilidad, cuya selección va a depender de los hiperparámetros del modelo. Inmediatamente, el siguiente token es seleccionado basado en su probabilidad. En el séptimo paso, se adiciona otro token a la secuencia de entrada y se repite el proceso desde el paso tres hasta que se alcance el límite máximo de tokens. En el penúltimo paso, la generación del texto terminará cuando se haya alcanzado el límite máximo de tokens o uno de fin de secuencia se haya detectado. Finalmente, hay una reconstrucción del texto, en el que el tokenizador convierte la secuencia final de tokens en una cadena de texto.

Figura 2. Inferencia de un token de un modelo de GPT de OpenAI



Fuente: Rothman (2024)

¹¹ Por ejemplo, en la oración «El perro come...», los *logits* son posibles palabras que pueden completar la oración para que tenga sentido sintáctico. Cada palabra tiene asignada una probabilidad, según con qué ha sido alimentado anteriormente el modelo GPT.

3.8. Uso de una interfaz de programación de aplicaciones (API)

La información que se visualiza en páginas web, por lo general, está hecha para que sea visualmente atractiva y fácil de acceder para una persona promedio. Por ejemplo, si alguien necesita comprar un pasaje de avión, a través de la página web, se accede a información como precios y disponibilidades de asientos. Por otro lado, cuando se requiere acceder a ella a través de una computadora para realizar operaciones y usos más complejos, la tarea es diferente. Las API permiten hacer una conexión entre una aplicación con otra para poder extraer información o acceder, en el caso de este proyecto de negocio, al servicio que ofrece una plataforma (Cooksey, 2014). Es decir, las API actúan como mensajeros.

De esta manera, se pretende utilizar la API de OpenAI para acceder al modelo GPT 4- o a través de un cuaderno de Python. Entre las ventajas de utilizar estos protocolos, está en que estos modelos se pueden aplicar a las necesidades específicas de un negocio o trabajo, facilita la automatización, se pueden ajustar los parámetros y, lo más importante, la escalabilidad.

El usuario común de Chat GPT o Gemini tiene muchas limitaciones, pues no puede alimentar al modelo con una gran cantidad de datos, por ejemplo, con un libro de cien páginas. Y, en el caso de hacerlo, se enfrenta a problemas como respuestas parciales e incluso alucinaciones. Esto se debe a que estos modelos tienen un límite de tokens por solicitud y pierden memoria y contexto de lo que se está procesando. La solución óptima es programar directamente con las API y automatizar el proceso.

CAPÍTULO 4. IMPLEMENTACIÓN DEL MODELO

En esta sección, se explicará la implementación del modelo utilizado para realizar la transcripción inteligente del libro propuesto de literatura ecuatoriana. Para ello, se implementó, a través de su respectiva API, el modelo basado en *transformers* GPT-4o, que, actualmente, es el más avanzado en el campo del procesamiento de lenguaje natural y *computer vision*.

4.1. Datos

Se utilizó el archivo escaneado de la edición de 1970 de *Cosas de mi tierra* de José Antonio Campos, publicado por la Biblioteca de Autores Ecuatorianos de Clásicos Ariel. El archivo está en escala de grises, tiene 119 páginas y pesa 46.3 MB.

Se eliminaron las páginas en blanco, además de las que tenían información no relevante para la transcripción, como la relación de títulos de la editorial, la portadilla y el índice. En este sentido, el total de páginas para alimentar el modelo son 114.

Se decidió trabajar con imágenes PNG, pues este formato captura más información en comparación con los JPG o JPEG. Además, cada imagen pesa en promedio 2 MB, tiene una dimensión de 1510 x 2358 pixeles, con una resolución de 300 dpi. Adicionalmente, el libro tiene alrededor de 55 017 palabras.

Este libro resulta desafiante para un OCR tradicional e incluso para cualquier modelo de inteligencia artificial, pues está plagado de modismos y errores ortográficos intencionales en palabras que tratan de representar de manera escrita la fonología de la Costa del Ecuador. Además, hay que sumar los errores originales de la edición, el uso incompleto de signos de exclamación e interrogación, la mezcla entre los diálogos de los personajes y los del narrador (en los que se pedirá al modelo que inserte rayas de diálogos para diferenciarlos), y las manchas y errores de impresión del libro.

4.2. Modelo GPT-4o de OpenAI

La plataforma de OpenAI ofrece tres modelos: GPT-3.5 Turbo, GPT-4 Turbo y, el más avanzado, GPT-4o, lanzado en mayo de 2024 e incluso más económico que su predecesor.

Por evidentes razones, se escogió el modelo GPT-4o, que puede trabajar con datos estructurados y no estructurados. Es decir, para la finalidad de este proyecto, imágenes escaneadas y texto, haciendo uso de *computer vision* y NLP.

Una vez configurada la clave de la API¹² y con los insumos del libro digital escaneado¹³, se procedió a realizar el modelo. Primero, se transforman las imágenes a un tamaño adecuado para que no sobrepase la cantidad máxima de tokens por solicitud. Una vez subidas las imágenes en el entorno, para que el modelo pueda analizarlas, se las deben codificar a *base64*, pues este formato proporciona una forma estándar antes de que se conviertan en un formato numérico.

Con estos insumos ya listos, se deben elaborar los *prompts* principales: el de sistema y del usuario. En el primero, se debe establecer el rol que va a tomar el modelo, y, en el segundo, las especificaciones de la transcripción y la corrección ortográfica y gramatical para el libro a analizar. Cabe mencionar que la elaboración de este último *prompt* es de vital importancia, pues el resultado depende en gran parte de cómo se lo elabore. Además, el modelo solamente puede procesar un cierto número de tokens por solicitud —tanto en el *completion* y en los *prompts*— y, en este sentido, se tiene que ser directo y eficaz al elaborar las indicaciones.

Posteriormente, dado que se va a realizar una transcripción inteligente de un libro, se necesita que exista una gran fidelidad y que infiera únicamente en las palabras que le resultan

¹² Esta clave de uso personal y no puede compartirse por ningún motivo, pues genera costos y almacena información sensible del usuario.

¹³ Para optimizar los costos de almacenamiento y procesamiento, para libros, lo recomendable es trabajar con imágenes a escala de grises y con una resolución mínima de 300 dpi.

ilegibles, además de hacer una corrección profunda. De esta manera, se necesita ser conservador y que el modelo no sea muy creativo. Para ajustar este comportamiento, es posible experimentar con los parámetros *temperature* y *top_p*. Valores bajos en el primero permiten que el modelo sea más determinista, y el segundo parámetro es un límite de probabilidad acumulada para considerar el rango de palabras para utilizar: entre más bajo, menos opciones de palabras.¹⁴

Con la experimentación de los parámetros del modelo en una página compleja, que contenga muchos términos vernáculos y errores de edición, y una vez escogidos los valores óptimos, se procedió a automatizar el proceso en Python para que lo replique a todas las páginas del libro. Finalmente, el resultado fue exportado a formato *.json* con el objetivo de también guardar los metadatos de la transcripción, como tiempo, número de páginas, tokens procesados y palabras en las que el modelo pueda tener dudas.

4.3. Tiempo de procesamiento

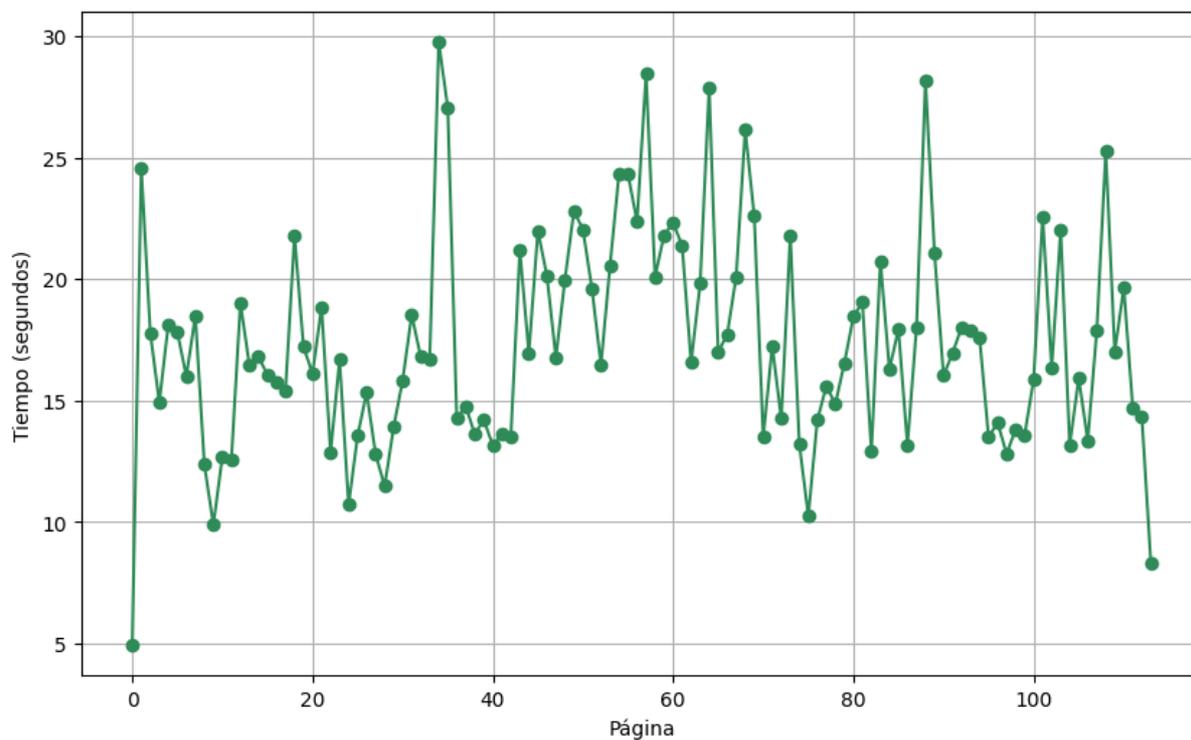
Las 114 páginas le tomaron procesar al modelo 1981.27 segundos, es decir 33 minutos, con un acelerador de hardware CPU en Google Colab. Este tiempo puede ser mucho menor al contratar unidades de procesamiento gráfico más avanzados. En este caso, se pagó la suscripción de esta plataforma, pero, al momento de correr el modelo, no hubo disponibilidad para utilizar tarjetas gráficas más potentes como una TPU v2 o una T4 GPU.

En promedio, el modelo procesó cada página en 17.38 segundos, con una desviación estándar de 4.37. En la figura 3, se puede observar el tiempo de procesamiento individual. Como

¹⁴ Por ejemplo, si se tiene un valor de 0.9, el modelo va a tener un mayor rango de palabras a considerar. En el caso de la oración: «El perro come...» la palabra más probable puede ser *croquetas*, seguida de *carne* y después *pan*. La probabilidad de la primera palabra puede ser de 0.6, la de la segunda, 0.29. y de la otra, 0.05. Con el valor de 0.9, únicamente va a considerar las primeras dos palabras.

es fácil inferir, en un libro, con un similar número de palabras a lo largo de todas sus páginas, se tiene un comportamiento estacionario en la variable.

Figura 3. Fluctuación del tiempo de procesamiento por página¹⁵



4.4. Resultados y limitaciones del modelo

En el anexo B, se pueden visualizar los resultados de la transcripción inteligente de la primera página del cuento «El torneo de los estribillos», que es el primer cuento del libro, después del prólogo de Hernán Rodríguez Castelo. Se eligió esta página debido a que el estudio introductorio tiene un español neutro y no es tan desafiante para el modelo como lo es el comienzo del libro.

Los resultados son sobresalientes: debido al *prompt*, el modelo introdujo rayas de diálogo faltantes en la edición original; completó signos de interrogación y exclamación que no

¹⁵ La variable *página* no representa una cantidad, sino el índice de cada página.

había por errores de edición o del mismo autor; corrigió faltas ortográficas como tildes en mayúsculas¹⁶; insertó puntos y comas; entregó un resultado sin contaminaciones comunes como encabezados, números de página o caracteres que el OCR reemplaza por manchas en las imágenes, y, principalmente, respetó, en su mayoría, los errores ortográficos intencionales que representan la fonología de la Costa del Ecuador. Es importante mencionar que, para libros con estilos de carácter como cursiva o negrita, se puede adaptar el código para que lo exporte en formato *markdown*, respetando estos elementos claves de la ortotipografía.

No obstante, también presentó limitaciones. Debido a que el modelo carece de información de modismos americanos y ecuatorianos; por ejemplo, reemplazó la palabra *peñiscar* por *peinica*, que no existe. También, hubo dos casos que no respetó la fonología escrita: corrigió *er* por *el* y *fuerte* en vez de *juerte*. Adicionalmente, reformuló un diálogo que era una oración interrogativa por una afirmativa. Asimismo, en un caso de la página, hizo la transcripción literal de la imagen sin insertar rayas de diálogo. En otras páginas, sí se observaron casos alucinación, especialmente cuando el texto no es tan legible.

A pesar de estas restricciones, el resultado del modelo es incomparable con el proceso de edición OCR común para este tipo de literatura. Existe una mejora en los tiempos de edición sin precedentes. Es asertivo afirmar que, con los nuevos modelos grandes de lenguaje natural, el mercado editorial al nivel mundial va a experimentar una gran expansión en su curva de posibilidades de producción.

4.4. Comparación de tiempos de edición

¹⁶ Antes de la Ortografía de la Real Academia Española del año 2010, no se solían poner tildes en mayúsculas. Por ejemplo, si el pronombre personal *él* iniciaba una oración y debía comenzar con mayúscula, la tilde se omitía, como es el caso del libro.

Una vez realizado el modelo y conocer el tiempo promedio de transcripción inteligente de un libro de literatura ecuatoriana de 114 páginas, es menester preguntarse cuánto tiempo ahorra este método en el proceso de edición normal.

Normalmente, cuando se va a reeditar un libro antiguo, el flujo de trabajo es el siguiente: 1) reconocimiento del libro escaneado por un software OCR; 2) selección de los elementos gráficos que no se quiere considerar para el reconocimiento;¹⁷ 3) limpieza general del texto en el software¹⁸; 4) exportación a un editor de texto como Word; 5) limpieza de impurezas y elementos incensarios de formato;¹⁹ 6) uniformización de elementos ortotipográficos²⁰; 7) exportar el texto editable a un Word con estilos; 8) implementar estilos de párrafo y de carácter en el documento; 9) tres revisiones completas de texto, y 10) envío del documento al departamento de diseño para diagramar el libro. Una vez realizada la diagramación, se realiza dos correcciones de estilo adicionales.

En promedio, un libro de las características presentadas, para llegar al paso siete y comenzar con la verdadera edición, se necesitan de 4 días de trabajo de edición. Del paso siete al diez, alrededor de siete días. Es decir, toma en promedio 11 días en elaborar el documento completo para que pase a la etapa de diagramación, y tres días adicionales para las dos

¹⁷ El software reconoce elementos gráficos como encabezados, número de páginas, imágenes y manchas que deben eliminarse manualmente para no considerarlos en el reconocimiento.

¹⁸ Se compara la página original con el resultado y, cuando el software cree que hay algún error de transcripción, lo resalta automáticamente en cian. A pesar de ser bastante útil esta característica, comete bastantes errores, y, en un libro de baja calidad de imagen, puede resultar muy abrumador trabajar en una página con cientos de errores.

¹⁹ Muchos elementos innecesarios de formato son exportados como *tabs*, saltos de sección, de página, guiones de separación de palabras, separaciones de párrafos, espacios, saltos de línea, entre otros.

²⁰ Reemplazo de guiones por rayas de diálogo, separación de decimales en puntos, puntos suspensivos correctos, completación de signos de interrogación y exclamación, entre otros.

correcciones ortotipográficas en el posdiseño. En total, 14 días de trabajo de edición para un libro de estas características.

Se estima que, con el nuevo uso de la transcripción inteligente, se necesita verificar su contenido en dos días para que pase a la etapa de diagramación. Después de esta etapa, el modelo resulta muy útil para hacer la siguiente corrección ortotipográfica, evitando errores humanos y dando resultados en tiempos récord y también ahorrando tiempo en el proceso de posdiagramación. En este sentido, se puede hacer una revisión general en apenas dos días.

En síntesis, se estima que el trabajo de edición actual, que involucra a dos editores, para un libro como el analizado, puede tomar once días. Con este modelo, apenas tres.

CAPÍTULO 5. ANÁLISIS FINANCIERO²¹

Una vez analizados los resultados del modelo, para desarrollar el segundo objetivo de este proyecto, se realizó un análisis financiero para la implementación de la digitalización de 100 títulos de las colecciones de Ariel Clásicos Ecuatorianos y Ariel Clásicos Universales para un periodo de dos años.

5.1. Estimaciones de egresos e ingresos

Costos operacionales

Los costos fijos de este proyecto se distribuyen a lo largo de 8 meses, y están compuestos por los costos de digitalización, edición, diagramación y de subir los libros a plataformas digitales. El principal costo variable es el de impresión bajo demanda para los libros físicos del mercado interno.

La mayor parte de libros de estas dos colecciones tienen una similar cantidad de páginas. Por esta razón, se consideró como referencia al libro objeto de este proyecto, que tiene 114. Para los 100 libros a digitalizar, se estimó un costo de uso del modelo de USD 22 800, que considera los costos de procesamiento y del *know-how*.

En el caso de la edición y corrección de estilo, se estima que los dos editores de planta puedan corregir 1.6 libros a la semana, cumpliendo con los 100 en 8 meses de trabajo. Con un salario mensual de USD 1100 para cada uno, se estima un costo de USD 17 600.

²¹ Para un análisis más profundo, dirigirse a la hoja de cálculo que se utilizó para elaborar los flujos de caja proyectados con el VAN y el TIR.

Con respecto al diseño y diagramación, se calculó un costo de USD 137.50 por libro, es decir, USD 15 750 en total.²² Además, existe otro costo para subir los archivos del libro a plataformas digitales como Kindle Direct Publishing y Apple Books de USD 20 por libro; es decir, USD 2000 en total.

Por el lado de la producción de los libros físicos para el mercado interno, dado que no tienen tanta rotación como los libros actualmente reeditados por no estar en la malla curricular presente, se los realizará a través del mecanismo de impresión bajo demanda. Aunque este proceso de producción no tiene economías de escala como la impresión offset, en donde se consigue costos medios mínimos, es óptimo para libros de baja y media rotación, ya que tiene baja inversión inicial y menos costos de inventarios. El costo promedio de impresión se consideró en la hoja de cálculo y no se puede presentar en este estudio.

En la tabla 2, se resumen estas estimaciones de los costos totales para los 12 meses del proyecto.

Tabla 2. Costos fijos para la reedición 100 títulos (8 meses)²³

Costo digitalización	\$ 22 800
Costo corrección de estilo y edición	\$ 17 600
Costo diagramación y portada total	\$ 15 750
Costo subir libros en plataformas	\$ 2000
Costo total	\$ 58 150

²² Hace poco tiempo, la editorial implementó un mecanismo de diagramación automática basada en estilos en el programa InDesign, que minimiza los tiempos de diagramación para los libros de literatura. La mayor parte del costo recae en la elaboración de la portada y la creación del libro en formato *epub* para subirlos a plataformas digitales.

²³ Estos costos son contables. Es decir, son los acumulados en los 4 meses de reedición. Para visualizar los costos desagregados por cada mes, dirigirse al anexo digital del flujo de caja.

Costos no operacionales

Para financiar este proyecto, se consideró en el flujo de caja proyectado un préstamo de USD 15 000 para 23 meses plazo con una tasa de interés activa del 12 %. Con una amortización francesa, en total, los egresos de intereses serán de USD 1866.²⁴

Ingresos operacionales

Hay tres tipos de ingresos: venta de libros electrónicos en plataformas digitales, venta de libros físicos en Amazon.com (impresión bajo demanda)²⁵, y, finalmente, la demanda interna de libros físicos, que es el rubro más representativo en este proyecto.

Se estima que cada mes se reediten 13 libros, completando todos en el noveno. A partir de allí, se calcula una venta mensual de 165 libros electrónicos, de 300 libros bajo el mecanismo de impresión bajo demanda de Amazon, que vende al nivel mundial, y 1200 libros físicos en el mercado interno.²⁶ Para los ingresos de los primeros meses, se calculó la proporción respectiva de estos datos.

En los 8 primeros meses de implementación, dado que hay una cantidad variable de libros, se estiman USD 27 570 de ingresos bajo los tres rubros mencionados. A partir de septiembre, los ingresos mensuales serán de USD 7877.

En la tabla 3, se puede visualizar un resumen de los ingresos mensuales cuando ya se encuentren reeditados los 100 libros.

²⁴ Para mayor detalle, dirigirse a la tabla de amortización de la hoja de cálculo del anexo digital.

²⁵ En estos dos primeros rubros, solo hay costos fijos, que se consideraron en el cuarto mes. Después, únicamente generan regalías.

²⁶ Para que se llegue a la meta de venta de 1200 libros mensuales, se deberían vender 12 ejemplares de cada título.

Tabla 3. Ingresos mensuales estimados a partir del noveno mes

Venta de libros electrónicos en plataformas digitales	\$ 288
Venta de libros físicos en Amazon.com	\$ 989
Venta de libros físicos en el mercado interno	\$ 6,600
Total	\$ 7877

Ingresos no operacionales

Para este rubro, se consideró una inversión por parte de la empresa de USD 20 000, además de un préstamo de USD 15 000.

5.2. Métodos de evaluación: VAN (valor actual neto) y TIR (tasa interna de retorno)

Como se mencionó al principio de este capítulo, se propuso un flujo de caja proyectado para dos años. En la tabla 4, se puede encontrar el resumen de los rubros principales de este análisis para cada año, que se lo hizo de manera mensual.

Tabla 4. Flujo de caja proyectado para dos años

	2025	2026
INGRESOS OPERACIONALES		
Venta de libros electrónicos en plataformas digitales	\$ 2161	\$ 3458
Venta de libros físicos en Amazon.com	\$ 7418	\$ 11 869
Venta de libros físicos en el mercado interno	\$ 49 500	\$ 79 200
TOTAL INGRESO OPERACIONAL	\$ 59 079	\$ 94 527

EGRESOS OPERACIONALES		
Costos digitalización	\$ 22 800	
Costo corrección de estilo	\$ 17 600	
Costo diagramación y portada	\$ 15 750	
Costo subir libros en plataformas	\$ 2 000	
Costo impresión	\$ 16 200	\$ 25,920
TOTAL EGRESOS OPERACIONALES	\$ 74 350	\$ 25,920
FLUJO OPERACIONAL		
	\$ -15 270.46	\$ 68 607.27
INGRESOS NO OPERACIONALES		
Desembolsos del préstamo bancario	\$ 15 000	
Otros ingresos no operacionales		
TOTAL INGRESOS NO OPERACIONALES	\$ 15 000	
EGRESOS NO OPERACIONALES		
Pagos de capital	\$ 6 746	\$ 8 253
Pago de intereses	\$ 1 319	\$ 546
Otros egresos no operacionales		
TOTAL EGRESOS NO OPERACIONALES	\$ 8 066	\$ 8 799
FLUJO NO OPERACIONAL		
	\$ 6 933	\$ -8 799
FLUJO DE CAJA NETO		
	\$ -8 336	\$ 59 807
DESCUENTO 15% TRABAJADORES		\$ 8 971

UTILIDADES DESPUÉS DE DESCUENTO		\$ 50 836
TRABAJADORES		
IMPUESTO A LA RENTA (25%)		\$ 12 709
UTILIDAD NETA	\$ -8336	\$ 38 127

Para el cálculo de valor actual neto (VAN), se consideró una tasa de descuento del 15%, que la componen la tasa de interés pasiva —que se encuentra alrededor de 8%—, una inflación esperada del 2% y una prima de riesgo del 5%. Con los valores estimados de la utilidad neta, se obtuvo un VAN de USD 1580 y una tasa interna de retorno de 18.79%.

Los valores de estos indicadores muestran que el modelo es financieramente viable. Es decir, con los supuestos anteriormente mencionados, se puede recuperar la inversión de USD 20 000. Además, es importante mencionar que, a partir del segundo año, los costos totales disminuyen considerablemente y solo se tendría el costo variable de impresión bajo demanda para el mercado interno. A largo plazo, se seguirán recibiendo regalías por los libros tanto digitales como físicos de las plataformas digitales.

5.3. Punto de equilibrio y análisis de sensibilidad

Las variables más importantes en este proyecto son los costos de procesamiento por página y la cantidad de libros físicos vendidos en el mercado interno por mes, que, como se analizó en un apartado anterior, tienen más demanda que los digitales.

De esta manera, para considerar posibles cambios en estos supuestos y sus efectos en la rentabilidad del proyecto, se realizó un análisis de sensibilidad (tabla 5). Para la primera variable, se consideraron costos desde USD 1 a USD 4, y, para la segunda, cantidades vendidas por mes desde 500 a 3500 libros.

Tabla 5. Análisis de sensibilidad

		Cantidad libros físicos vendidos mensualmente en el mercado interno						
VAN	\$ 1,580.60	900	1000	1100	1200	1300	1400	1500
Costo de procesamiento por página	\$ 1.00	\$ -2166	\$ 2387	\$ 6940	\$11 494	\$16 047	\$20 600	\$25 154
	\$ 1.25	\$ -4645	\$-91	\$ 4462	\$ 9015	\$13 569	\$18 122	\$22 675
	\$ 1.50	\$ -7123	\$ -2569	\$ 1984	\$ 6537	\$11 090	\$15 644	\$20 197
	\$ 1.75	\$ -9601	\$ -5048	\$ -494	\$ 4059	\$ 8612	\$13 165	\$17 719
	\$ 2.00	\$ -12 079	\$ -7526	\$ -2973	\$ 1581	\$ 6134	\$10 687	\$15 241
	\$ 2.25	\$ -14 558	\$ -10 004	\$ -5451	\$ -898	\$ 3656	\$ 8209	\$12 762
	\$ 2.50	\$ -17 036	\$ -12 483	\$ -7929	\$ -3376	\$ 1177	\$ 5731	\$10 284

En este sentido, también se identificó el punto de equilibrio de este proyecto, considerando la variable de ventas en el mercado interno. Las ventas mínimas para que haya rentabilidad económica en estos dos años no deben ser menores a 1166 libros mensuales.

CAPÍTULO 6. CONCLUSIONES Y RECOMENDACIONES

6.1. Conclusiones

El uso de los grandes modelos de lenguaje natural basados en *transformers* en el proceso de reedición de literatura clásica ha demostrado ser viable tanto técnica como financieramente. La implementación de la API de OpenAI para usar GPT-4o ha permitido una transcripción inteligente que aprende según el contexto, superando ampliamente a los softwares OCR tradicionales.

Además de la calidad del resultado, los tiempos de edición disminuyen drásticamente. Se estimó que, normalmente, un libro con las características de *Cosas de mi tierra*, con alrededor de 50 000 palabras, le puede tomar 11 días a un editor experimentado para corregirlo. Con la implementación de inteligencia artificial en este proceso, puede tomar alrededor de 2 a tres días.

Para tener resultados óptimos es necesario experimentar con los principales parámetros del modelo, y ajustarlos según la necesidad de la tarea. En el caso de una transcripción de un libro antiguo con errores de impresión, con manchas y con errores de edición, se necesita ser conservador, pero también se debe permitir que el modelo pueda ser en cierta medida creativo para inferir palabras que no son legibles, corregir la gramática y ortografía, y respetar los errores ortográficos intencionales de los autores que tratan de representar de manera escrita la fonología de regiones particulares. Es menester mencionar que se escogió un libro de literatura ecuatoriana por los desafíos que presenta. Es fácil inferir que, para las obras de literatura universal, con un español neutro, los resultados van a ser mucho más óptimos.

A pesar de los sobresalientes resultados del modelo, la inteligencia artificial no es perfecta. Es necesario que quienes trabajan con ella siempre supervisen su trabajo. A pesar de lo conservador que fue el modelo, hubo casos de alucinaciones, transcripciones inexactas y la inconsistencia de correcciones, esta última no tan frecuente. También, a pesar de que este

modelo tiene billones de parámetros, se puede inferir que no está lo suficientemente alimentado con términos específicos del español americano y, peor aún, de ecuatorianismos.

El uso de las API es muy importante para agilizar el flujo de trabajo. Para el usuario normal de inteligencia artificial, que utiliza las plataformas disponibles, es imposible llegar a estos resultados en tiempos tan cortos. Incluso si pagan la suscripción de Chat-GPT, tienen limitaciones de procesamiento, además, el modelo alucina fácilmente cuando es alimentado con mucha información y se le pide tareas tan complejas como la realizada en este proyecto. También, si sube un *pdf* a la plataforma, cuando alcanza el número máximo de tokens generados, el modelo para.

Actualmente, no se conoce casos en que la industria editorial ecuatoriana haya implementado este tipo de tecnología para agilizar su proceso de reedición de libros antiguos, pues el modelo más avanzado de OpenAI apenas salió hace cuatro meses. De la misma manera, con una reunión realizada el 26 de junio con, posiblemente, la editorial de literatura más grande del país, por su trayectoria y cantidad de publicaciones de literatura ecuatoriana contemporánea, se pudo verificar que esta tecnología no se está aplicando. Por otro lado, sí se pudo confirmar que, en la actualidad, las editoriales sí generan contenido gráfico con inteligencia artificial; generalmente, para la elaboración de portadas de libros.

En el caso de uso de inteligencia artificial para edición a gran escala, utilizando *natural language processing* y *computer vision*, Publicaciones Educativas Ariel podría considerarse innovadora en la industria editorial ecuatoriana, como alguna vez lo fue en los años setenta y ochenta. Sin embargo, es importante direccionar correctamente el proceso de cambio de una pequeña empresa para que utilice estas tecnologías y la implemente de manera integral en todos sus departamentos. Esta iniciativa es un gran primer paso.

Adicionalmente, es importante reconocer que, en el corto y mediano plazo, como lo notó NVIDIA, la creación y la ingeniería de *prompts* va a resultar de vital importancia en esta

nueva revolución industrial. La programación pasa a un segundo nivel. Se tiene que aprender a construir indicaciones directas y precisas para comunicarse con los modelos de inteligencia artificial. De esta manera, se puede inferir que, en el mercado general, no solamente en el editorial, la búsqueda de profesionales con esta nueva profesión estará en auge. Un buen comienzo para introducirse a este mundo es capacitarse de cómo funciona el lenguaje humano y de la ciencia que está detrás de la gramática, y de la importancia de saber comunicarse y escribir bien.

Por otro lado, en el mediano plazo, las editoriales, al nivel nacional e internacional, tendrán la necesidad de contar con científicos de datos que puedan agilizar su proceso de edición; sin embargo, no solo basta el conocimiento en programación y estadística, sino que también estén capacitados en temas de edición, corrección ortotipográfica y gramática. Es decir, se enfatiza en la formación holística de los profesionales presentes y futuros en este campo — y en general—, para que desarrollen sus habilidades blandas y no solo se enfoquen en su área de especialidad. Estas capacidades serán determinantes al momento de insertarse en el mercado laboral y les permitirán adquirir una mayor ventaja competitiva con respecto a otros profesionales.

Finalmente, con respecto a la literatura ecuatoriana clásica, es importante rescatarla y, a través de las facilidades que ofrecen las plataformas digitales, difundirla internacionalmente, como siempre quiso Benjamín Carrión.

6.2. Recomendaciones

Se recomienda —y de hecho se va a implementar en un corto plazo en la editorial— correr el modelo y posteriormente realizar un *fine-tuning* para alimentarlo con otros libros que tengan el mismo estilo literario, como *Los que se van* y demás libros de la literatura ecuatoriana de la primera mitad del siglo XX.

Finalmente, se recomienda utilizar otros modelos de procesamiento de lenguaje natural basados en *transformers* para poder comparar los resultados con el utilizado en este proyecto.

6.3. Impacto del proyecto en la empresa

Este proyecto y la implementación de este modelo está siendo útil para agilizar los procesos de edición y evitar los errores humanos. La capacidad de digitalizar y corregir textos permitirá a la empresa expandir su catálogo de libros físicos y digitales, tanto a nivel nacional como internacional. En este sentido, la reducción en los tiempos y costos asociados permitirá desarrollar una ventaja competitiva diferenciadora, incrementando así los ingresos de la empresa y su sostenibilidad a mediano y largo plazo.

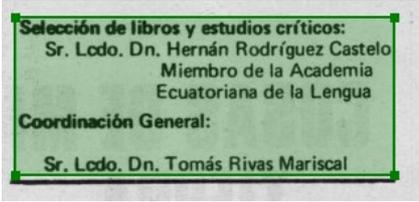
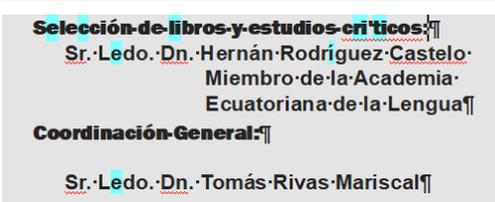
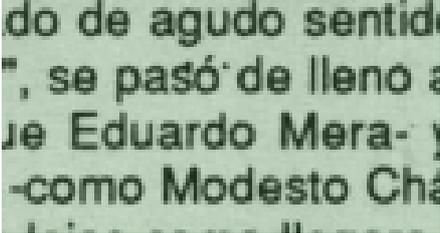
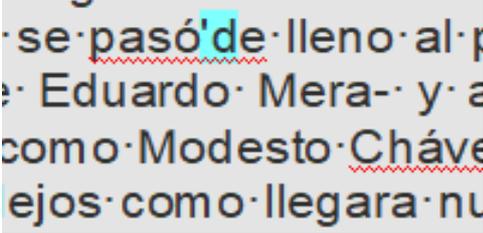
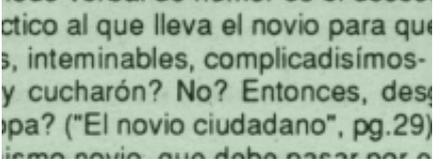
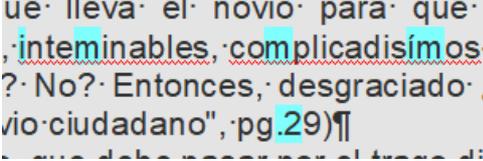
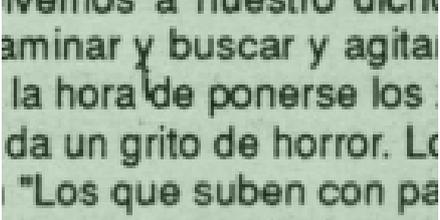
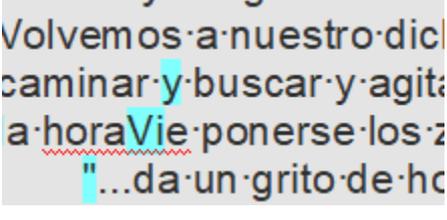
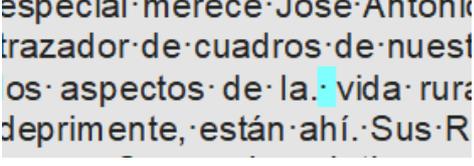
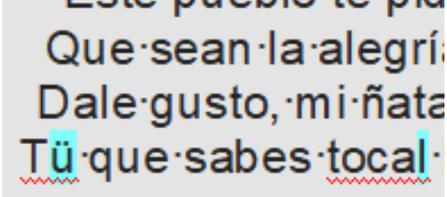
REFERENCIAS

- Cámara Ecuatoriana del Libro. (2024). *Gráficos Estadísticas CEL*.
<https://www.celibro.org.ec/pagina/wp-content/uploads/2024/03/Graficos-Estadisticas-CEL.pdf>
- Campos, J. A. (1979). *Cosas de mi tierra* (1st ed.). Publicaciones Educativas Ariel.
- Cooksey, B. (2014). *An Introduction to APIs* (B. Landers & D. Schreiber, Eds.). Zapier, Inc.
- Edmonds, M. (2020). *A Brief History of Optical Character Recognition (OCR)*. Pitney Bowes. https://www.pitneybowes.com/content/dam/pitneybowes/uk/en/shipping-and-mailing/e-invoicing/Blog_E-invoicing-The_Brief_History_of_OCR.pdf?itid=lk_inline_enhanced-template
- El Ministerio de Cultura y Patrimonio presentó los resultados de la Encuesta de Hábitos Lectores, Prácticas y Consumos Culturales*. (2022, June 14). Ministerio de Cultura y Patrimonio. <https://www.culturaypatrimonio.gob.ec/el-ministerio-de-cultura-y-patrimonio-presento-los-resultados-de-la-encuesta-de-habitos-lectores-practicas-y-consumos-culturales/>
- Goldberg, Y. (2017). *Neural Network Methods for Natural Language Processing*. Morgan & Claypool Publishers.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. The MIT Press.
- Grzybowski, A., Pawlikowska-Łagód, K., & Lambert, W. C. (2024). A History of Artificial Intelligence. *Clinics in Dermatology*, 42(3), 221–229.
<https://doi.org/https://doi.org/10.1016/j.clindermatol.2023.12.016>
- Hair, J., Black, W., Babin, B., & Anderson, R. (2019). *Multivariate Data Analysis* (8th ed.). Cengage.
- Hernán Rodríguez Castelo: El interés por los libros vino de familia. (2012, July 27). *El Universo*. <https://www.eluniverso.com/2012/07/27/1/1380/hernan-rodriguez-castelo-el-interes-libros-vino-familia.html>
- Lee, R. (2024). *Natural Language Processing: A Textbook with Python Implementation*. Springer.
- Los álbumes de cromos divirtieron y dieron conocimientos a niños y jóvenes. (2008, April 27). *El Universo*.
<https://www.eluniverso.com/2008/04/27/0001/18/82904E375424433997C6629266A0FDAD.html>
- Moor, J. (2006). The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazine*, 27.
<https://aaai.org/ojs/index.php/aimagazine/article/view/1911/1809>

- Pazmiño, J. (2010, April). *La economía en el gobierno de José María Velasco Ibarra*.
https://the.pazmiño.com/VELASCO_IBARRA-1968-72.pdf
- Pérez Pimentel, R. (1987). *Diccionario biográfico del Ecuador: Vol. V*. Universidad de Guayaquil.
- Rodríguez Castelo, H. (n.d.-a). *Historia de la literatura*.
https://www.hernanrodriguezcastelo.com/?page_id=67
- Rodríguez Castelo, H. (n.d.-b). *Vida y obra*.
<https://www.hernanrodriguezcastelo.com/biografia.htm>
- Rothman, D. (2024). *Transformers for natural language processing and computer vision: explore generative AI and large language models with Hugging Face, ChatGPT, GPT-4V, and DALL-E3* (3rd ed.). Packt.
- Russell, S., & Norving, P. (2022). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- Serrano, L. (2021). *Grokking Machine Learning*. Manning Publications.
- Tapia, E. (2024, April 16). FMI: La economía de Ecuador, la que menos crecerá en 2024 en América Latina. *Primicias*. <https://www.primicias.ec/noticias/economia/fmi-pib-proyeccion-crecimiento-ecuador-venezuela/#:~:text=El%20FMI%20prev%C3%A9%20que%20la,ser%C3%A1%20la%20que%20m%C3%A1s%20crecer%C3%A1>
- Vision*. (n.d.). OpenAI Platform. <https://platform.openai.com/docs/guides/vision>
- What are Convolutional Neural Networks?* (n.d.). IBM.
<https://www.ibm.com/topics/convolutional-neural-networks>
- What Is Optical Character Recognition (OCR)?* (2022, January 5). IBM Cloud Education.
<https://www.ibm.com/blog/optical-character-recognition/>

ANEXOS

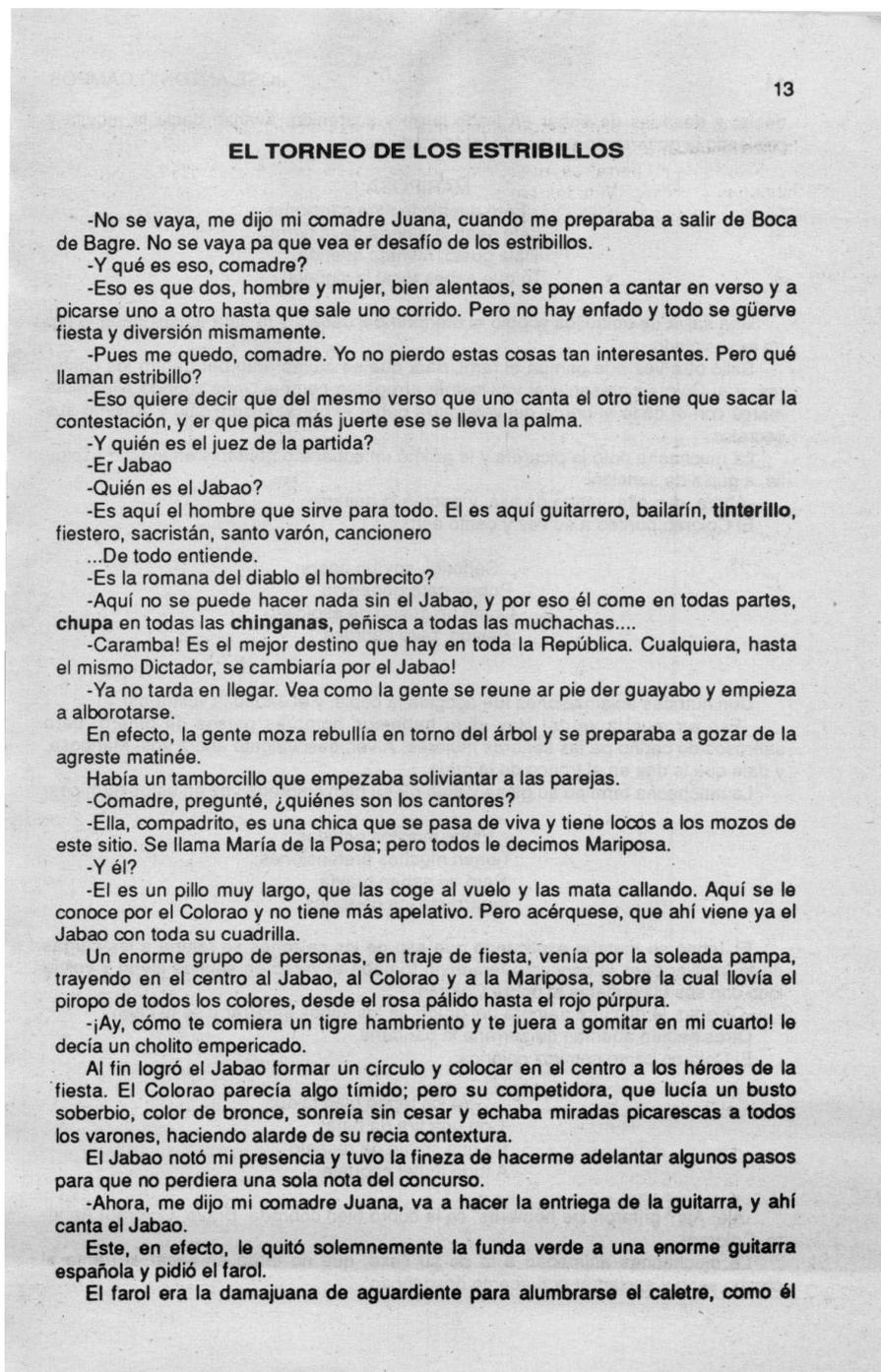
Anexo A: Errores de transcripción comunes de un OCR²⁷

Imagen escaneada	Transcripción OCR
 <p>Selección de libros y estudios críticos: Sr. Lcdo. Dn. Hernán Rodríguez Castelo Miembro de la Academia Ecuatoriana de la Lengua Coordinación General: Sr. Lcdo. Dn. Tomás Rivas Mariscal</p>	 <p>Selección de libros y estudios críticos¶ Sr. Lcdo. Dn. Hernán Rodríguez Castelo· Miembro de la Academia· Ecuatoriana de la Lengua¶ Coordinación General:¶ Sr. Lcdo. Dn. Tomás Rivas Mariscal¶</p>
 <p>do de agudo sentido ", se pasó de lleno a de Eduardo Mera- y -como Modesto Chá</p>	 <p>· se pasó de lleno al p · Eduardo Mera- y a como Modesto Cháve ejos como llegara nu</p>
 <p>ctico al que lleva el novio para que s, inteminables, complicadísimo- y cucharón? No? Entonces, des opa? ("El novio ciudadano", pg.29) ... que debe pasar por el t...</p>	 <p>ue lleva el novio para que· · inteminables, complicadísimo- ?· No?· Entonces,· desgraciado· · novio ciudadano",· pg.29)¶ ... que debe pasar por el t...</p>
 <p>vemos a nuestro dic aminar y buscar y agita la hora de ponerse los da un grito de horror. Le "Los que suben con pa</p>	 <p>Volvemos a nuestro dic caminar y buscar y agita a hora Vie ponerse los z "...da un grito de ho</p>
<p>trazador de cuadros de nue los aspectos de la vida ru deprimente, están ahí. Sus y sus Cosas de mi tierra co</p>	 <p>especial merece Jose Anton trazador de cuadros de nuest os aspectos de la vida rura deprimente, están ahí. Sus R</p>
<p>Que sean la alegría Dale gusto, mi ñata Tú que sabes total l</p>	 <p>Este pueblo te pla Que sean la alegrí Dale gusto, mi ñata Tü que sabes total·</p>

²⁷ En este caso, se utilizó el software OCR de ABBY FineReader.

<p>-El que me llevó a la pila! -Miente usted! -Lo juro por la cruz! -Vuelves a llorar? Oye, Ca por llorar a mi también</p>	<p>-El que me llevó a la pila! -Miente usted! -Lo juro por</p>
--	--

Anexo B: Comparación entre una imagen y el texto producido por GPT 4o²⁸



²⁸ De todo el libro de *Cosas de mi tierra*, se decidió analizar los resultados de la primera página del primer cuento, evitando el prólogo, pues la tarea de transcripción es mucho más compleja para este tipo de literatura que utiliza muchos términos vernáculos y que representa —especialmente en los diálogos— la fonología de la Costa de Ecuador.

EL TORNEO DE LOS ESTRIBILLOS²⁹

—¡No se vaya! —me dijo mi comadre Juana, cuando me preparaba a salir de Boca de Bagre—. ¡No se vaya pa que vea er desafío de los estribillos!³⁰

—¿Y qué es eso, comadre?³¹

—Eso es que dos, hombre y mujer, bien **alentaos**, se ponen a cantar en verso y a picarse uno a otro hasta que sale uno corrido. Pero no hay enfado y todo se güerve fiesta y diversión mismamente.

—Pues me quedo, comadre. Yo no pierdo estas cosas tan interesantes. ¿Pero qué llaman estribillo?³²

—Eso quiere decir que del **mesmo** verso que uno canta el otro tiene que sacar la contestación, y **el** que pica más **fuerte** ese se lleva la palma.

—¿Y quién es el juez de la partida?

—**Er** Jabao.

—¿Quién es el Jabao?

—Es aquí el hombre que sirve para todo. **Él³³** es aquí guitarrero, bailarín, tinterillo, fiestero, sacristán, santo varón, **cancionero**.

—¡De todo entiende!

²⁹ Las palabras o caracteres resaltados en amarillo son correcciones acertadas; los que están en cian son cambios intermedios, que no son originales del autor, pero que, sin duda, mejoran la lectura, y los que están en rojo son equivocaciones del modelo.

³⁰ Debido al *prompt* y a los parámetros escogidos del modelo, en este párrafo, se puede observar que hace las correcciones de rayas en los diálogos. Como se puede observar en la imagen del libro en su primera edición, no se diferencian los diálogos de los personajes con las intervenciones del narrador. De la misma manera, respeta la escritura de fonología costeña: *pa* y *er*.

³¹ El modelo añade el signo de interrogación inicial.

³² Aunque lo correcto es que el signo de interrogación inicial esté después de *pero*, es mejor añadirlo a que no esté, como en el libro original.

³³ Antes de la Ortografía del 2010, era común que las mayúsculas no se tilden, como en este caso. El modelo es asertado.

—Es la romana del diablo el hombrecito³⁴

—Aquí no se puede hacer nada sin el Jabao, y por eso él come en todas partes, chupa en todas las chinganas, **peinica**³⁵ a todas las muchachas...

—¡Caramba! Es el mejor destino que hay en toda la República... **C**ualquiera, hasta el mismo **D**³⁶ictador, se cambiaría por el Jabao!³⁷

—Ya no tarda en llegar. Vea **cómo** la gente se **reúne ar pie der** guayabo y empieza a alborotarse.

En efecto, la gente moza rebullía en torno³⁸ del árbol y se preparaba a gozar de la agreste matinée.

Había un tamborcillo que empezaba soliviantar a las parejas.

—Comadre **—pregunté—**, ¿quiénes son los cantores?

—Ella, compadrito, es una chica que se pasa de viva y tiene locos a los mozos de este sitio. Se llama María de la Posa; pero todos le decimos Mariposa.

—**¿**Y él?

—**Él** es un pillo muy largo, que las coge al vuelo y las mata callando. Aquí se le conoce por el Colorao y no tiene más apelativo. Pero acérquese, que ahí viene ya el Jabao con toda su cuadrilla.

³⁴ Este diálogo originalmente es una pregunta.

³⁵ Originalmente, es *peñiscar*, un americanismo de *pellizcar*.

³⁶ A pesar de que se indicó que haga una corrección ortotipográfica con las normas actuales de la RAE, no utilizó minúscula en este caso.

³⁷ Hay un buen OCR, porque el modelo transcribe textualmente, pero no corrige añadiendo un signo de exclamación inicial.

³⁸ En el caso de ABBY FineReader en este tipo de palabras confunde a la *r* y *n* como una *m*.

Un enorme grupo de personas, en traje de fiesta, venía por la soleada pampa, trayendo en el centro al Jabao, al Colorao y a la Mariposa, sobre la cual llovía el piropo de todos los colores, desde el rosa pálido hasta el rojo púrpura.

—¡Ay, cómo te comiera un tigre hambriento y te **juro**³⁹ a gomitar en mi cuarto! **—**le decía un cholo empecirado.

Al fin logró el Jabao formar un círculo y colocar en el centro a los héroes de la fiesta. El Colorao parecía algo tímido; pero su competidora, que lucía un busto soberbio, color de bronce, sonreía sin cesar y echaba miradas picarescas a todos los varones, haciendo alarde de su recia contextura.

El Jabao notó mi presencia y tuvo la fineza de hacerme adelantar algunos pasos para que no perdiera una sola nota del concurso.

—¡Ahora **me dijo mi comadre Juana,**⁴⁰ va a hacer la **entrega** de la guitarra, y ahí canta el Jabao.

Este, en efecto, le quitó solemnemente la funda verde a una enorme guitarra española y pidió el farol.

El farol era la damajuana de aguardiente para alumbrarse el caletre, como él [...]

³⁹ Te juera.

⁴⁰ Transcripción exacta, pero no implementó las rayas de intervención del narrador.