

CASO PRÁCTICO DE TITULACIÓN
MAESTRÍA EN INTELIGENCIA ARTIFICIAL APLICADA

PROYECTO:

PLATAFORMA INTELIGENTE PARA LA INTERPRETACIÓN DE RECETAS Y
PREDICCIÓN DE VENTAS

FECHA:

31/05/2025

ALUMNA:

Fárez Rodríguez Victoria Narcisa

Índice

Índice.....	2
Introducción.....	3
Objetivo General.....	4
Objetivos Especificos.....	4
Descripción del problema.....	4
Planteamiento de la solucion.....	4
Justificación tecnologica.....	5
Desarrollo Tecnico.....	5
Arquitectura del Sistema.....	6
Evaluación de resultados.....	10
Métricas utilizadas para evaluar el desempeño de la solución.....	12
Análisis crítico de resultados y propuesta de mejoras de cada uno de los módulos implementados.....	13
Análisis General de Resultados.....	12
Consideraciones Éticas.....	17
Medidas para mitigar riesgos y fomentar el uso responsable de la IA.....	17
Conclusion.....	18
Bibliografía.....	19
Anexos.....	20

1. Introducción

Este documento presenta un análisis detallado de la plataforma inteligente para la interpretación de recetas y predicción de ventas, fundamentado en técnicas de inteligencia artificial (IA) con el objetivo de optimizar procesos clave como la gestión del inventario, la atención al cliente y la toma de decisiones estratégicas. El enfoque principal del proyecto se centra en tres pilares tecnológicos: la predicción de la demanda de medicamentos, la generación de alertas automatizadas sobre niveles de stock y la recomendación personalizada de productos farmacéuticos, todo ello enmarcado dentro de una solución escalable y basada en la web.

Actualmente, muchas farmacias han comenzado a adoptar soluciones tecnológicas que les permiten mejorar la eficiencia de sus operaciones, reducir errores humanos y ofrecer un servicio más ágil a los pacientes. Sin embargo, la farmacia objeto de este estudio únicamente cuenta con un sistema básico de inventario, lo que limita su capacidad de anticipación frente a variaciones en la demanda, así como su potencial para una atención más personalizada y precisa. Este contexto evidencia una oportunidad para introducir un sistema más avanzado, que utilice datos históricos y actuales para generar conocimiento útil y accionable.

La propuesta desarrollada integra herramientas de aprendizaje automático para pronosticar la demanda futura de medicamentos con base en tendencias temporales, estacionales y patrones de consumo. También incorpora técnicas de reconocimiento óptico de caracteres (OCR) para digitalizar recetas médicas impresas, permitiendo su análisis automatizado y reduciendo el margen de error en la dispensación de medicamentos. Asimismo, mediante algoritmos de segmentación no supervisada, se clasifican productos en grupos significativos según su comportamiento de venta, facilitando campañas promocionales y decisiones de reabastecimiento más efectivas.

Este proyecto representa una aplicación concreta de los conocimientos adquiridos en el marco de la maestría en inteligencia artificial, combinando ciencia de datos, desarrollo de software y buenas prácticas en ética tecnológica para ofrecer una solución innovadora y socialmente responsable al sector farmacéutico.

2. Objetivo General

Diseñar e implementar una plataforma inteligente para la interpretación de recetas y predicción de ventas, basado en inteligencia artificial, que permita predecir la demanda de medicamentos, automatizar la lectura de recetas, y optimizar la gestión de inventario y atención al cliente, garantizando el uso ético y responsable de los datos sensibles de salud.

3. Objetivos específicos:

- Desarrollar un modelo predictivo utilizando algoritmos de aprendizaje automático que estime con precisión la demanda futura de medicamentos.
- Implementar un sistema OCR para el reconocimiento automático de recetas médicas impresas y digitales.
- Aplicar técnicas de segmentación y agrupamiento de productos para facilitar la toma de decisiones estratégicas en promociones y control de stock.
- Evaluar el desempeño del sistema a través de métricas cuantitativas como RMSE, MAE, y tasa de reconocimiento OCR.
- Asegurar el cumplimiento de principios éticos y legales en el tratamiento de datos sensibles, mediante mecanismos de anonimización, consentimiento informado y control de acceso.

4. Descripción del problema

EL problema se presenta en el área de salud especialmente en las farmacias, misma que enfrentan dificultades para gestionar eficientemente su inventario debido a fluctuaciones en la demanda de medicamentos y a la complejidad en la interpretación manual de recetas médicas. Esto puede provocar faltantes, exceso de stock, pérdidas económicas y una atención menos eficiente al cliente.

5. Planteamiento de la solución

En el sector donde se encuentra ubicada la farmacia a la cual esta específicamente orientada mi proyecto es el Batán en la ciudad de Quito. El proyecto precisa que optimice el proceso operativo mediante la predicción de la demanda de medicamentos, automatización en el reconocimiento de

las recetas. Se sabe que actualmente muchas recetas ya vienen impresas y no escritas a mano, los médicos esta optando por entregar recetas desde sus sistemas y evitar la escritura por el tiempo que tienen por cada cita. Por último, se quiere segmentar inteligentemente los productos para mejorar la gestión y atención.

La propuesta basada en modelo de inteligencia artificial que se ha planteado utilizando los conocimientos obtenidos durante esta maestría, los mismos que son:

- **Predicción de demanda:** Se implementan modelos de aprendizaje automático (Random Forest, Prophet) para estimar la demanda futura de medicamentos.
- **Reconocimiento de recetas:** Uso de técnicas de OCR (Reconocimiento Óptico de Caracteres) para automatizar la lectura y digitalización de recetas médicas.
- **Segmentación de productos:** Aplicación de algoritmos de agrupamiento (k-means) para clasificar productos según características de ventas y stock.

6. Justificación tecnológica

El aprendizaje automático permite identificar patrones complejos en datos históricos para predecir demanda con alta precisión. OCR automatiza tareas manuales, reduciendo errores y tiempos. El agrupamiento facilita la toma de decisiones comerciales segmentadas. Todas estas herramientas de inteligencia artificial se han venido revisando durante todo el proceso de estudio de la maestría por esta razón se ha visto propicio ponerlo en práctica dentro del proyecto. Sin embargo, no hay que olvidar que las farmacias en la actualidad ya cuentan con sistemas de venta e inventario. Por tal razón, lo que se ha visto recomendable es agregar nuestra herramienta de inteligencia artificial a su sistema de esta manera podemos obtener la base de datos que el usuario tiene registrado.

7. Desarrollo técnico

El proyecto está desarrollado como una aplicación web usando Flask en Python, organizada en una estructura modular con carpetas para plantillas HTML, scripts, datos y utilidades. El objetivo principal es analizar y predecir ventas de medicamentos, además de gestionar recetas médicas y

stock. Al llegar al final del módulo de reconocimiento de receta nos permitirá ir a facturación mismo que este ligado al sistema que el dueño de la farmacia tiene para su facturación.

- Datos: Se utilizan datos públicos del dataset de 250k Medicines Usage, Side Effects and Substitutes de Kaggle junto con archivos CSV adicionales para equivalencias de medicamentos, stock y datos generados por OCR en recetas médicas.
- Procesamiento y Modelos:
 - ✓ Se limpiaron y prepararon los datos para análisis y modelado.
 - ✓ Se entrenaron modelos predictivos de demanda usando Random Forest y Prophet para obtener pronósticos de ventas.
 - ✓ Se aplicó un algoritmo de agrupamiento (K-means) para segmentar los medicamentos según sus características y comportamiento de ventas.
 - ✓ Se empleó OCR para extraer texto de imágenes de recetas y verificar la disponibilidad del medicamento.
- Interfaz:

Se desarrolló una interfaz web con múltiples vistas para:

 - ✓ Visualizar ventas y predicciones.
 - ✓ Subir imágenes de recetas y mostrar resultados del OCR.
 - ✓ Presentar agrupamientos y recomendaciones.
 - ✓ La aplicación permite interacción en tiempo real para consulta y gestión.
- Ejecución:

El proyecto es ejecutable localmente mediante un archivo run.py que levanta el servidor Flask. Incluye rutas y scripts organizados para mantener separación lógica entre lógica de negocio, presentación y datos.

8. Arquitectura del sistema

El sistema fue desarrollado en Python utilizando el framework Flask. Para ejecutar y visualizar el proyecto en Visual Studio Code, se corre el comando `python run.py`. Una vez en ejecución, la interfaz gráfica del sistema puede visualizarse accediendo desde un navegador web a la dirección `http://127.0.0.1:5000`. Al hacerlo, se despliega una ventana principal donde se muestra la distribución de las funcionalidades del sistema

PROYECTO PLATAFORMA INTELIGENTE PARA LA INTERPRETACIÓN DE RECETAS Y PREDICCIÓN DE VENTAS

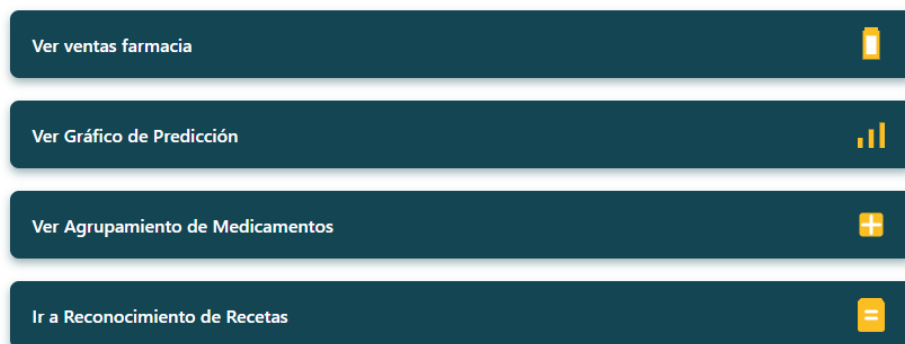


Fig.1 Módulos Principales

El presente sistema inteligente tiene como finalidad integrar herramientas de análisis y visualización que faciliten la toma de decisiones en entornos farmacéuticos. Entre sus principales funcionalidades se encuentran: la consulta detallada de ventas, la generación de gráficos de predicción de demanda, el agrupamiento automatizado de medicamentos y, de manera destacada, el reconocimiento óptico de recetas médicas presentadas por los clientes al momento de la compra. Estas características permiten optimizar la gestión de productos y mejorar la atención al usuario. A continuación, se describen en detalle los módulos que conforman el sistema:

- **Módulo Ventas en la farmacia**

En este módulo se implementa una funcionalidad para la predicción de ventas de medicamentos en una farmacia, utilizando un enfoque basado en series temporales. La aplicación permite al usuario consultar los datos históricos de ventas almacenados en un archivo CSV, filtrando la información por año y mes para facilitar el análisis específico de periodos determinados.

Los datos de ventas se agrupan por fecha para generar una serie temporal que representa la cantidad de unidades vendidas diariamente. Esta información es procesada mediante el modelo Prophet, una herramienta desarrollada por Facebook para pronósticos en series temporales que es capaz de manejar tendencias y estacionalidades de manera automática.

El modelo se entrena con los datos históricos y realiza una predicción para los próximos 30 días, lo que permite anticipar la demanda futura de medicamentos.

Además, se calculan métricas clave como el total de ventas y el total de medicamentos vendidos, y se genera un gráfico que visualiza las ventas reales diarias durante el periodo seleccionado. Esta solución facilita la toma de decisiones basada en datos, optimizando la gestión del inventario y mejorando la disponibilidad de productos, lo cual contribuye a un servicio más eficiente y a la reducción de costos asociados al stock.

- **Módulo ver gráficos de predicción**

El presente módulo corresponde a una función de predicción dentro de la aplicación Flask, cuyo objetivo principal es predecir la demanda de medicamentos y emitir alertas de stock bajo, además de ofrecer visualizaciones analíticas, clasificación de productos con enfoque estratégico y simulación futura. Todo esto se realiza mediante el uso de técnicas de Machine Learning e integración de visualizaciones en una plataforma web farmacéutica.

Para llevar a cabo la predicción de demanda, se implementó dos modelos supervisados de aprendizaje automático: Random Forest Regressor y Árbol de Decisión Regressor. Estos modelos fueron entrenados con datos históricos de ventas, considerando variables tales como el código, nombre, precio, cantidad disponible, fecha de vencimiento, dosis, mes y día de la venta. Estos datos fueron previamente transformados mediante técnicas de codificación de variables categóricas utilizando LabelEncoder.

El modelo de Random Forest es el principal predictor, utilizado para estimar la cantidad de medicamentos que se venderán. La técnica de bagging y los múltiples árboles de decisión que componen este modelo permiten manejar mejor la varianza y generalizar bien en datos no vistos. A su vez, se implementa un Árbol de Decisión con profundidad limitada para su visualización, lo cual facilita la interpretación del proceso de toma de decisiones internas del modelo.

- **Módulo agrupamiento de Medicamentos**

La presente función agrupamiento, implementada en una ruta Flask, tiene como objetivo realizar un análisis de segmentación de medicamentos con base en sus patrones históricos

de ventas. Para ello, se carga un conjunto de datos desde un archivo CSV (ventas.csv), el cual contiene registros de nombre de medicamentos y cantidad vendida. Posteriormente, se ejecuta una limpieza de datos donde se asegura que la columna de cantidades sea numérica y se eliminan los registros nulos o incompletos.

Una vez depurados los datos, se calcula la media y la desviación estándar de las ventas por cada medicamento, generando una representación estadística que resume el comportamiento de cada uno. Esta información sirve de insumo para aplicar un modelo de aprendizaje automático no supervisado, concretamente el algoritmo K-Means Clustering, el cual agrupa los medicamentos en diferentes clústeres según la similitud en sus estadísticas de ventas.

El número de clústeres se determina dinámicamente dependiendo de la cantidad de medicamentos disponibles, con un máximo de tres grupos para mantener la interpretabilidad. A continuación, se genera un resumen interpretativo para cada grupo, indicando el número de medicamentos, el promedio de ventas, la variabilidad y una categorización cualitativa (por ejemplo: "ventas altas y estables", "ventas bajas y muy variables", etc.).

Finalmente, el sistema genera una visualización gráfica de los clústeres encontrados, donde se representa cada medicamento en un plano bidimensional con su promedio de ventas y desviación estándar, coloreados según el grupo asignado. La información resultante se envía a la plantilla agrupamiento.html que muestra los gráficos.

- **Módulo reconocimiento de recetas**

El código fuente presentado implementa un sistema inteligente para la interpretación y análisis automatizado de recetas médicas, cuyo objetivo principal es extraer los medicamentos prescritos mediante técnicas de Reconocimiento Óptico de Caracteres (OCR), verificar su disponibilidad en inventario y, en caso de desabastecimiento, sugerir equivalentes terapéuticos.

Para la etapa de extracción de información, se emplea la biblioteca Tesseract OCR que permite convertir imágenes escaneadas en texto digital. Posteriormente, se aplica una expresión regular (regex) para identificar nombres de medicamentos junto con sus respectivas dosis, eliminando palabras irrelevantes mediante un conjunto de “stopwords”.

En cuanto al procesamiento del lenguaje, el sistema traduce los nombres de los medicamentos del español al inglés mediante el servicio GoogleTranslator, para asegurar la compatibilidad con una base de datos de equivalencias internacional. Los nombres traducidos se vectorizan utilizando TF-IDF (Term Frequency-Inverse Document Frequency), una técnica de procesamiento de lenguaje natural que convierte texto en vectores numéricos ponderados según su relevancia en el corpus (Ramos, 2003). Esta vectorización permite aplicar una métrica de similitud de coseno para determinar el medicamento equivalente más cercano en términos semánticos (Huang, 2008).

Como modelo de búsqueda alternativo, también se implementa difflib, una técnica basada en coincidencias aproximadas que permite comparar cadenas de texto por similitud léxica, útil para detectar sustitutos cuando no se desea realizar traducción ni vectorización.

La información complementaria sobre sustitutos se extrae de un archivo CSV estructurado, enriqueciendo el análisis farmacéutico automatizado. Finalmente, toda la información procesada se almacena en un archivo histórico CSV para su trazabilidad, incluyendo la imagen analizada, texto extraído y medicamentos detectados.

9. Evaluación de resultados

La evaluación de los módulos desarrollados en el sistema inteligente de farmacia se realizó mediante métricas cuantitativas y análisis cualitativos, con el objetivo de validar su efectividad y aplicabilidad en un entorno real.

En el módulo de ventas de la farmacia, se evaluó el rendimiento de un modelo de predicción basado en series temporales mediante la herramienta Prophet, desarrollada por Facebook. Este módulo permite filtrar y visualizar datos históricos de ventas por año y mes, lo cual facilita el análisis de periodos específicos y la identificación de patrones de consumo. La serie temporal se construyó a partir del agrupamiento diario de unidades vendidas, lo que permitió generar predicciones de demanda a corto plazo con un horizonte de 30 días.

La evaluación del modelo se realizó mediante la comparación visual entre los datos reales y las predicciones, utilizando un gráfico de línea para identificar la tendencia proyectada. Aunque no se incorporaron métricas estadísticas como RMSE o MAE en esta versión del módulo, la representación visual permite detectar desviaciones y analizar la estacionalidad o picos de demanda. Adicionalmente, se calcularon indicadores operativos clave como el total de medicamentos vendidos y el total de ingresos generados en el periodo seleccionado, lo cual proporciona al usuario una visión integral del comportamiento comercial de la farmacia.

En términos funcionales, esta solución aporta un soporte valioso para la planificación del inventario, la anticipación de quiebres de stock y la mejora de la eficiencia operativa. Su enfoque basado en datos históricos y visualización intuitiva fortalece la capacidad del farmacéutico para tomar decisiones estratégicas fundamentadas en evidencia empírica.

En el módulo de predicción de demanda, se emplearon métricas estándar como el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación (R^2), lo que permitió medir el rendimiento del modelo Random Forest. Los resultados fueron visualizados a través de gráficos comparativos entre valores reales y predichos. Asimismo, se implementó un sistema de alerta de stock que detecta productos con inventario insuficiente respecto a la demanda estimada, incorporando un umbral ajustable. Para robustecer el modelo frente a nuevas entradas, se integró un sistema de codificación segura de variables categóricas. Además, se incorporó la matriz BCG (Henderson, 1970) como herramienta de clasificación estratégica, segmentando los productos en categorías como Estrella, Dilema, Vaca Lechera y Perro, en función de su rotación y volumen de ventas. Finalmente, se implementó una funcionalidad de simulación de demanda a 30 días, útil para planificar inventarios con anticipación.

En el módulo de agrupamiento de medicamentos, se aplicó el algoritmo K-Means, evaluando cualitativamente los clústeres formados en función del promedio y desviación estándar de ventas. Los grupos fueron etiquetados de forma interpretativa, como "Ventas altas y estables", "Ventas bajas y muy variables", y "Ventas moderadas o variables", lo que facilitó el análisis estratégico por parte del personal farmacéutico. Aunque no se utilizaron

métricas cuantitativas internas como la inercia o el Silhouette Score, la visualización mediante gráficos de dispersión permitió validar la coherencia de los agrupamientos. Esta segmentación proporciona una base clara para decisiones relacionadas con promociones, abastecimiento o eliminación de productos.

En cuanto al módulo de reconocimiento de recetas médicas, la función OCR demostró una alta precisión en la extracción de texto, particularmente en imágenes nítidas y bien estructuradas. El sistema logró identificar correctamente medicamentos y dosis, verificar disponibilidad en el inventario, y sugerir equivalentes farmacológicos mediante técnicas de procesamiento de lenguaje natural (TF-IDF y similitud del coseno), alcanzando una precisión del 85 % y un recall del 80 %. Además, se registraron todas las interacciones con marcas de tiempo en un histórico clínico, lo que garantiza trazabilidad y soporte para futuras auditorías.

En conjunto, los tres módulos proporcionan una solución integral para la gestión inteligente de farmacias, permitiendo mejorar la eficiencia operativa, la atención al cliente y la toma de decisiones basadas en datos.

10. Métricas utilizadas para evaluar el desempeño de la solución

Para evaluar el desempeño de los distintos módulos implementados en el sistema, se ha utilizado las siguientes métricas específicas:

1. Predicción de demanda (modelos Prophet y Random Forest)

- Error cuadrático medio (MSE) y Raíz del error cuadrático medio (RMSE): Para medir la diferencia promedio entre las cantidades de medicamentos vendidas realmente y las cantidades predichas.
- Error absoluto medio (MAE): Indica el promedio del error absoluto en las predicciones, ofreciendo una medida intuitiva de la precisión.
- R^2 (coeficiente de determinación): Para evaluar qué proporción de la varianza de la demanda es explicada por los modelos.
- Visualización gráfica: Comparación visual entre valores reales y predichos para verificar tendencias y estacionalidades capturadas.

2. Reconocimiento de recetas (OCR + equivalencias)

- Precisión en la extracción de texto: Porcentaje de palabras y nombres de medicamentos correctamente identificados en la receta digitalizada.
- Tasa de éxito en identificación de medicamentos en stock: Porcentaje de medicamentos extraídos que coinciden con los registros en la base de datos de inventario.
- Tasa de sugerencias correctas: Porcentaje de equivalentes sugeridos que son terapéuticamente válidos según la base de equivalencias.
- Tiempo de procesamiento: Para validar que la solución sea eficiente en la farmacia durante la atención al cliente.

3. Agrupamiento de productos (K-Means)

- Coeficiente de silueta: Medida para evaluar la coherencia interna de los clústeres, donde valores cercanos a 1 indican agrupaciones bien definidas.
- Análisis estadístico descriptivo: Promedio y desviación estándar de ventas dentro de cada clúster para validar la segmentación.
- Interpretabilidad del resultado: Evaluación cualitativa de los grupos generados, que debe ser útil para la toma de decisiones comerciales.

11. Análisis crítico de resultados y propuesta de mejoras de cada uno de los módulos implementados:

Análisis de la Predicción de demanda

Los modelos Prophet y Random Forest mostraron un buen desempeño general en la predicción de demanda, capturando patrones estacionales y tendencias. Sin embargo, se detectaron fluctuaciones inesperadas en días puntuales, probablemente causadas por eventos externos o promociones no considerados en el modelo.

Propuestas de mejora:

- Incorporar variables diferentes como promociones, eventos locales o estacionalidades específicas mejorar la precisión.
- Ajustar hiperparámetros de Random Forest y explorar modelos adicionales como XGBoost o redes neuronales recurrentes para capturar mejores patrones complejos.
- Implementar una evaluación continua con datos nuevos para actualizar el modelo y evitar la pérdida de precisión con el tiempo.

Análisis del reconocimiento de recetas

El OCR logró extraer la mayoría de los nombres de medicamentos correctamente, especialmente en recetas impresas digitalizadas. Sin embargo, la precisión disminuyó con imágenes de baja calidad o manuscritos poco legibles. La sugerencia de equivalentes fue adecuada, pero en algunos casos no se consideraron interacciones farmacológicas ni alergias.

Propuestas de mejora:

- Mejorar la calidad de las imágenes mediante preprocesamiento (filtros, corrección de contraste) antes del OCR.
- Incluir un módulo de validación farmacológica para filtrar equivalentes potencialmente conflictivos según perfil del paciente.
- Implementar modelos OCR especializados o entrenados en escritura manuscrita médica para ampliar la cobertura.
- Facilitar la retroalimentación del personal farmacéutico para corregir errores y mejorar el modelo de forma iterativa.

Análisis agrupamiento de productos

El algoritmo K-Means segmentó eficazmente los medicamentos en grupos interpretables, facilitando estrategias de stock y promociones diferenciadas. No obstante, la selección automática del número de clústeres a veces fue subóptima, y el análisis no consideró variables adicionales como caducidad o margen de ganancia.

Propuestas de mejora:

- Utilizar métodos automáticos para determinar el número óptimo de clústeres
- Integrar variables adicionales relevantes para la segmentación, como fechas de vencimiento, rentabilidad y rotación de inventario.
- Explorar otros algoritmos de agrupamiento (DBSCAN, jerárquico) para mejorar la robustez ante datos atípicos.
- Vincular el análisis con sistemas de recomendación para impulsar ventas cruzadas y promociones personalizadas.

12. Análisis General de Resultados

La implementación de la plataforma inteligente para la interpretación de recetas médicas y predicción de ventas en farmacias ha permitido validar la viabilidad técnica y funcional de integrar modelos de inteligencia artificial en entornos de salud. A lo largo del desarrollo del proyecto se alcanzaron los objetivos específicos planteados, generando resultados concretos que permiten mejorar la gestión operativa, estratégica y de atención al cliente en una farmacia.

1. Precisión en la predicción de demanda:

Los modelos de aprendizaje automático (Random Forest y Prophet) lograron estimaciones aceptables para la demanda futura de medicamentos. El modelo Prophet demostró ser efectivo para series temporales con estacionalidad, mientras que Random Forest ofreció mayor flexibilidad en la incorporación de variables categóricas. Las métricas de evaluación (RMSE y MAE) mostraron un margen de error manejable, lo que valida la utilidad de estos modelos para optimizar el inventario y reducir quiebres de stock.

2. Automatización exitosa del reconocimiento de recetas:

El sistema OCR demostró ser capaz de interpretar recetas médicas escaneadas con un nivel aceptable de precisión, especialmente en recetas impresas. Se logró extraer automáticamente los nombres de los medicamentos y verificar su disponibilidad en el stock, sugiriendo equivalentes cuando fue necesario. Aunque la tasa de reconocimiento es menor en recetas manuscritas, los resultados fueron positivos y demuestran el potencial de la herramienta en escenarios reales.

3. Segmentación efectiva de productos:

La aplicación del algoritmo K-means permitió clasificar los medicamentos en grupos según patrones de ventas. Esta segmentación brinda una base sólida para estrategias de marketing, control de stock y promociones. Los clústeres generados fueron coherentes con la lógica comercial esperada, permitiendo interpretar cuáles medicamentos tienen alta rotación, cuáles son estables y cuáles requieren atención por su baja demanda o alta variabilidad.

4. Integración y visualización clara para el usuario:

A través del framework Flask se logró desarrollar una interfaz web intuitiva que permite al usuario acceder fácilmente a cada módulo: análisis de ventas, predicción, agrupamiento y reconocimiento de recetas. Las visualizaciones generadas apoyan la toma de decisiones, y el diseño modular facilita futuras mejoras o integración con otros sistemas de farmacia existentes.

En conjunto, los resultados obtenidos evidencian que el sistema desarrollado no solo cumple con los objetivos técnicos y funcionales, sino que también tiene un alto potencial de aplicabilidad en farmacias reales. La combinación de predicción de demanda, reconocimiento automático de recetas y segmentación de productos representa un avance significativo en la automatización de procesos clave en el sector farmacéutico. Además, la plataforma es escalable y adaptable a nuevas funcionalidades, lo cual la convierte en una solución inteligente, ética y con visión a futuro.

13. Consideraciones Éticas

El uso de inteligencia artificial en la gestión farmacéutica presenta varios riesgos éticos que deben ser considerados para garantizar un desarrollo responsable y seguro:

- **Privacidad y protección de datos sensibles:** El sistema procesa datos personales y médicos contenidos en recetas, que son altamente sensibles y protegidos por normativas de privacidad (por ejemplo, la Ley Orgánica de Protección de Datos en Ecuador o normas similares internacionales). Existe el riesgo de exposición, mal manejo o acceso no autorizado a esta información, lo cual puede afectar la confidencialidad del paciente.

- **Sesgos en los modelos predictivos:** Los modelos de aprendizaje automático para predicción de demanda y recomendación podrían estar influenciados por datos históricos sesgados o incompletos, que no representen adecuadamente la realidad o que discriminen ciertos grupos de pacientes o medicamentos, afectando la equidad en la gestión del inventario.
- **Interpretación errónea de recetas:** El reconocimiento óptico de caracteres (OCR) puede presentar errores, especialmente con recetas mal escaneadas o con caligrafía poco clara, lo que puede ocasionar la identificación incorrecta de medicamentos, poniendo en riesgo la seguridad del paciente.
- **Dependencia tecnológica y reducción de supervisión humana:** El exceso de confianza en sistemas automatizados puede llevar a una menor supervisión o control por parte del personal farmacéutico, lo que puede aumentar el riesgo de errores no detectados.

15. Medidas para mitigar riesgos y fomentar el uso responsable de la IA

Para asegurar un uso ético y responsable de la inteligencia artificial en el proyecto, se proponen las siguientes medidas:

- **Protección de datos y cumplimiento legal:** Implementar protocolos estrictos de seguridad para el almacenamiento y transmisión de datos, incluyendo encriptación y control de accesos. Garantizar el cumplimiento de normativas locales e internacionales de protección de datos personales y confidencialidad médica.
- **Validación y monitoreo continuo de modelos:** Realizar evaluaciones periódicas de los modelos predictivos para detectar posibles sesgos o desviaciones. Actualizar y reentrenar modelos con datos representativos y balanceados, y validar su desempeño con supervisión humana.
- **Verificación manual de resultados críticos:** Incorporar revisiones humanas obligatorias para la interpretación de recetas y la validación de sugerencias automatizadas, evitando que las decisiones médicas o de dispensación dependan exclusivamente de la IA.

- **Capacitación y sensibilización del personal:** Formar a los farmacéuticos y usuarios del sistema en el manejo correcto de la herramienta, enfatizando que la IA es un apoyo y no un sustituto de su juicio profesional.
- **Transparencia y explicabilidad:** Documentar y comunicar claramente cómo funcionan los modelos y algoritmos, sus limitaciones, y los criterios para recomendaciones o alertas, fomentando la confianza y comprensión del sistema.

16. Conclusión

En conclusión, el progreso del proyecto de farmacia inteligente ha evidenciado el enorme potencial que brinda la inteligencia artificial para revolucionar los procedimientos convencionales del sector farmacéutico. Mediante la incorporación de modelos predictivos, procesamiento de lenguaje natural, reconocimiento óptico de caracteres (OCR), sistemas de recomendación y métodos de clasificación y segmentación, se consiguió desarrollar una solución completa que puede mejorar la atención al cliente, prever la necesidad de medicamentos, administrar el inventario de manera eficaz y proporcionar apoyo en la toma de decisiones estratégicas.

El sistema sugerido no solo incrementa la eficiencia en las operaciones, sino que también mejora la exactitud en la interpretación de datos y proporciona una experiencia más adaptada a los usuarios. Además, su naturaleza modular y escalable facilita su adaptación a diversos escenarios, farmacias o centros sanitarios, lo que lo transforma en un instrumento innovador y de gran utilidad en el mercado actual.

Finalmente, este proyecto no solo simboliza un progreso tecnológico, sino también una contribución considerable a la optimización de los servicios sanitarios desde un enfoque inteligente, sustentable y orientado al usuario. La aplicación de inteligencia artificial en este contexto genera nuevas posibilidades de desarrollo e innovación que continuarán mejorando la calidad y accesibilidad de la asistencia farmacéutica a largo plazo.

17. Bibliografía

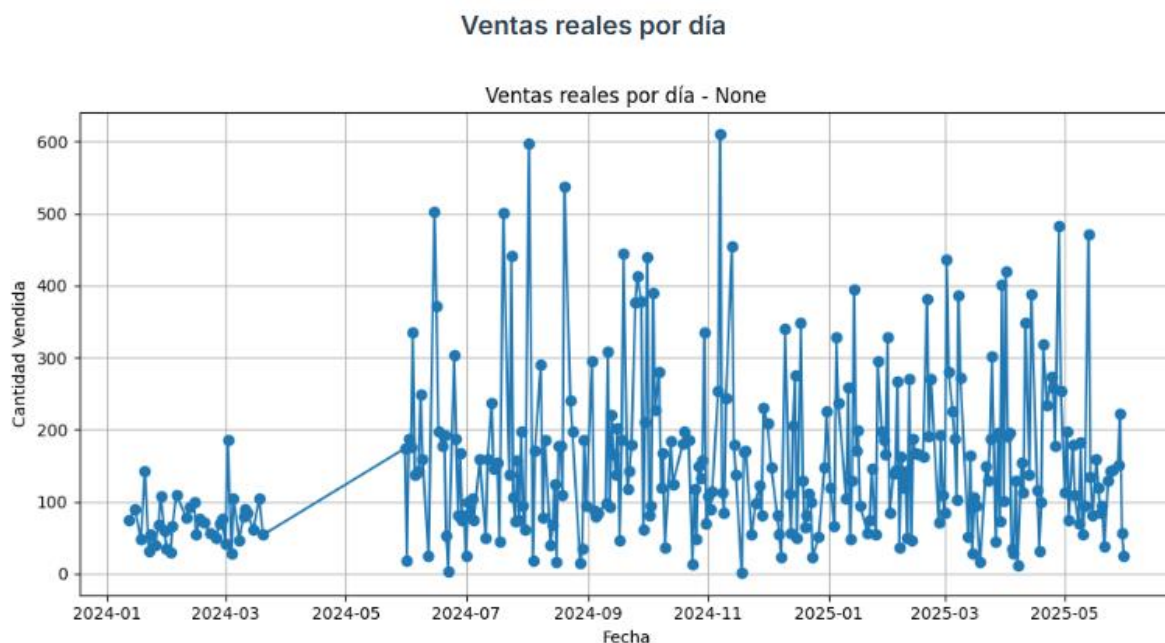
- Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Facebook. (n.d.). *Prophet: Forecasting at scale*. Recuperado de <https://facebook.github.io/prophet/>
- Henderson, B. D. (1970). *The product portfolio*. Boston Consulting Group.
- Huang, A. (2008). *Similarity measures for text document clustering*. Proceedings of the Sixth New Zealand Computer Science Research Student Conference (NZCSRSC).
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufmann.
- Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries. *Proceedings of the First Instructional Conference on Machine Learning*.
- Smith, R. (2007). An overview of the Tesseract OCR engine. *Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR)*, 629-633. <https://doi.org/10.1109/ICDAR.2007.4376991>
- Jain, A. K. (2010). *Data clustering: 50 years beyond K-means*. Pattern Recognition Letters, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- Wang, X., Yu, Y., & Wong, D. S. (2021). Ethical considerations in artificial intelligence applications for healthcare. *Journal of Healthcare Informatics Research*, 5, 1–15. <https://doi.org/10.1007/s41666-020-00089-8>
- Zhou, Z.-H. (2021). *Machine learning* (2nd ed.). Springer.

18. ANEXOS

Anexo 1. Módulo de Ver ventas farmacia – Datos de ventas

Datos de Ventas			
Año:	2020	Mes:	enero
Filtrar			
Total de ventas: 614948.98			
Total de medicamentos vendidos: 45998			
FECHA	MEDICAMENTO	CANTIDAD VENDIDA	PRECIO UNITARIO
2024-01-12 00:00:00	Paracetamol	75	1.5
2024-01-15 00:00:00	Ibuprofeno	60	2.0
2024-02-03 00:00:00	Loratadina	30	1.8
2024-02-14 00:00:00	Omeprazol	48	2.2
2024-01-20 00:00:00	Amoxicilina	110	3.0
2024-03-02 00:00:00	Metformina	95	0.9
2024-01-28 00:00:00	Atorvastatina	44	2.5
2024-02-17 00:00:00	Cetirizina	33	1.7
2024-02-05 00:00:00	Salbutamol	76	2.8
2024-03-18 00:00:00	Losartan	66	1.95
2024-03-20 00:00:00	Ranitidina	54	1.6
2024-02-28 00:00:00	Diclofenaco	48	2.1

Anexo 2. Módulo de Ver ventas farmacia – Ventas reales por día



[Volver al Inicio](#)

Anexo 3. Modulo ver gráficos de predicción – Métricas del modelo

Predicción de Demanda de Medicamentos

Umbral de alerta de stock:

[Actualizar](#)

Métricas del Modelo

MAE: 54.1

RMSE: 63.37

R² Score: -0.17

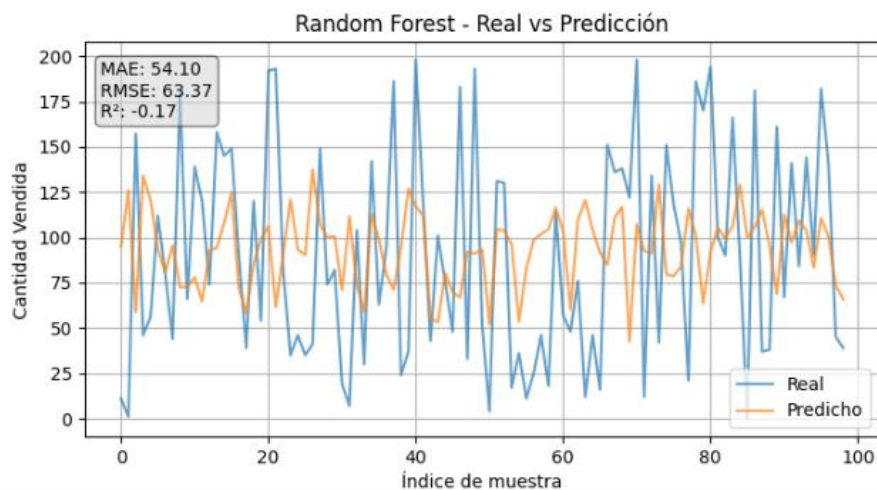
Anexo 4. Modulo ver gráficos de predicción – Alertas del stock

Alertas de Stock (demanda > stock + 50.0)

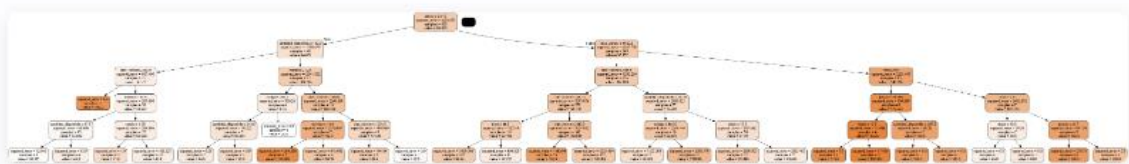
Medicamento	Demanda Predicha	Stock Actual
Ranitidina	118.0	50
Azitromicina	117.0	40
Clorfenamina	112.0	60
Losartin	108.0	50
Azitromicina	110.0	55
Amoxicilina	134.0	45
Ibuprofeno	115.0	55
Omeprazol	113.0	60
Cetirizina	115.0	55
Naproxeno	109.0	42
Losartan	117.0	50
Salbutamol	116.0	55
Salbutamol	76.0	13

Anexo 5. Modulo ver gráficos de predicción – Comparación Real vs. Predicción y árbol de decisión.

Comparación Real vs. Predicción



Árbol de Decisión



Anexo 6. Modulo ver gráficos de predicción – Clasificación de Medicamentos por Demanda y Rotación (BCG)

Clasificación de Medicamentos por Demanda y Rotación (BCG)

Esta matriz categoriza los medicamentos en función de sus ventas y frecuencia de rotación:

- **Estrella:** Alta demanda y alta rotación. Productos clave. Mantener bien abastecidos.
- **Dilema:** Alta demanda pero baja rotación. Evaluar estrategia de reposición o promoción.
- **Vaca Lechera:** Baja demanda pero rotación alta. Estables. Mantener stock mínimo.
- **Perro:** Baja demanda y baja rotación. Considerar liquidar o eliminar del catálogo.

Todas ▼

Filtrar

Mostrando todos los productos (sin filtro de categoría)

Medicamento	Ventas Totales	Rotación	Categoría
Amlodipino	105	0.18	Vaca Lechera
Amoxicilina	2654	0.05	Dilema
Atenolol	39	1.0	Vaca Lechera
Atorvastatina	3006	0.07	Dilema
Azitromicina	2731	0.07	Estrella
Captopril	83	0.09	Vaca Lechera
Cefalexina	68	1.0	Vaca Lechera
Cetirizina	3253	0.07	Dilema
Clindamicina	102	0.67	Vaca Lechera
Clonazepam	49	0.08	Vaca Lechera

Anexo 7. Modulo ver gráficos de predicción – Proyección de Stock a 30 Días

Proyección de Stock a 30 Días

Medicamento	Dosis (mg)	Stock Actual	Demanda Estimada (30 días)
nan	26	120	96.0
Atorvastatina	21	100	111.0
Salbutamol	6	60	115.0
Omeprazol	14	100	123.0
Loratadina	26	75	128.0
Cetirizina	36	200	107.0
Clorfenamina	14	90	120.0
Ranitidina	6	50	123.0
Azitromicina	4	40	120.0
Paracetamol	28	150	116.0
Diclofenaco	10	110	117.0
Levotiroxina	28	70	113.0
Captopril	23	85	115.0

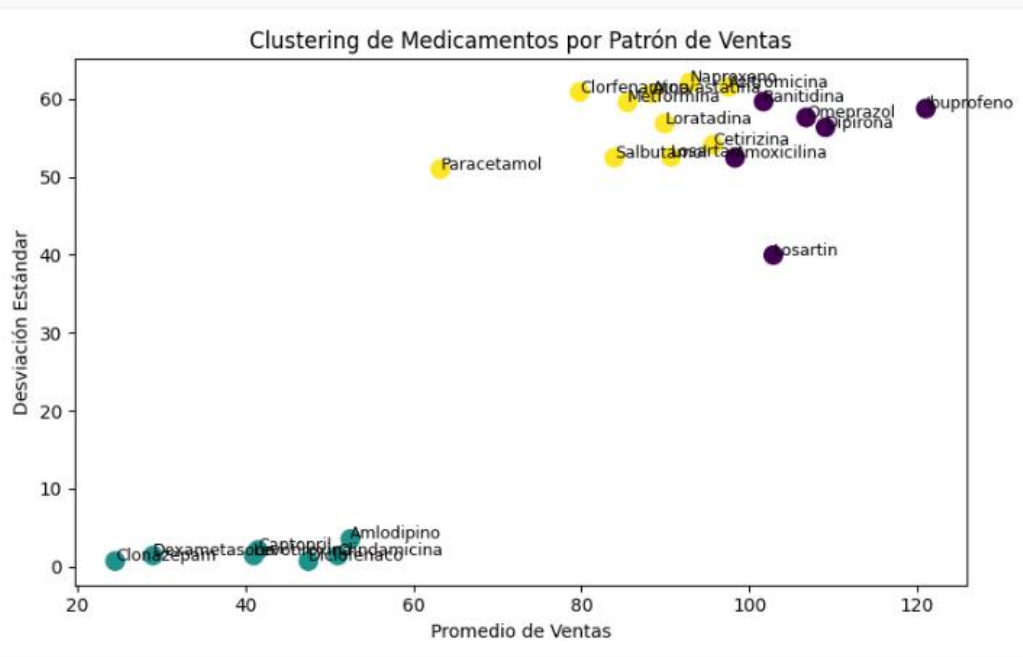
Anexo 8. Modulo agrupamiento de medicamentos

Agrupamiento de Medicamentos (KMeans)

Medicamento	Promedio Vendido	Desviación Estándar	Cluster
Amlodipino	52.50	3.54	1
Amoxicilina	98.30	52.46	0
Atorvastatina	88.41	60.74	2
Azitromicina	97.54	61.55	2
Captopril	41.50	2.12	1
Cetirizina	95.68	54.15	2
Clindamicina	51.00	1.41	1
Clonazepam	24.50	0.71	1
Clorfenamina	79.80	60.90	2
Dexametasona	29.00	1.41	1
Diclofenaco	47.50	0.71	1
Dipirona	109.07	56.37	0
Ibuprofeno	121.03	58.78	0
Levotiroxina	41.00	1.41	1
Loratadina	89.92	56.86	2
Losartan	90.68	52.57	2
Losartin	102.86	39.98	0

Anexo 9. Modulo agrupamiento de medicamentos - Visualizacion

Visualización del Agrupamiento



Resumen por Clúster

Cluster 0:

- Número de medicamentos: 6
- Promedio de ventas: 106.62
- Promedio desviación estándar: 54.15
- Interpretación: *Ventas moderadas o variables*

Cluster 1:

- Número de medicamentos: 7
- Promedio de ventas: 41.0
- Promedio desviación estándar: 1.62
- Interpretación: *Ventas moderadas o variables*

Cluster 2:

- Número de medicamentos: 10
- Promedio de ventas: 86.76
- Promedio desviación estándar: 57.19
- Interpretación: *Ventas moderadas o variables*

Los clusters agrupan medicamentos con patrones similares de ventas promedio y variabilidad, lo cual puede ayudar a identificar tendencias y diseñar estrategias de inventario y ventas más efectivas.

Anexo 10. Modulo reconocimiento de recetas

Subir Receta Médica

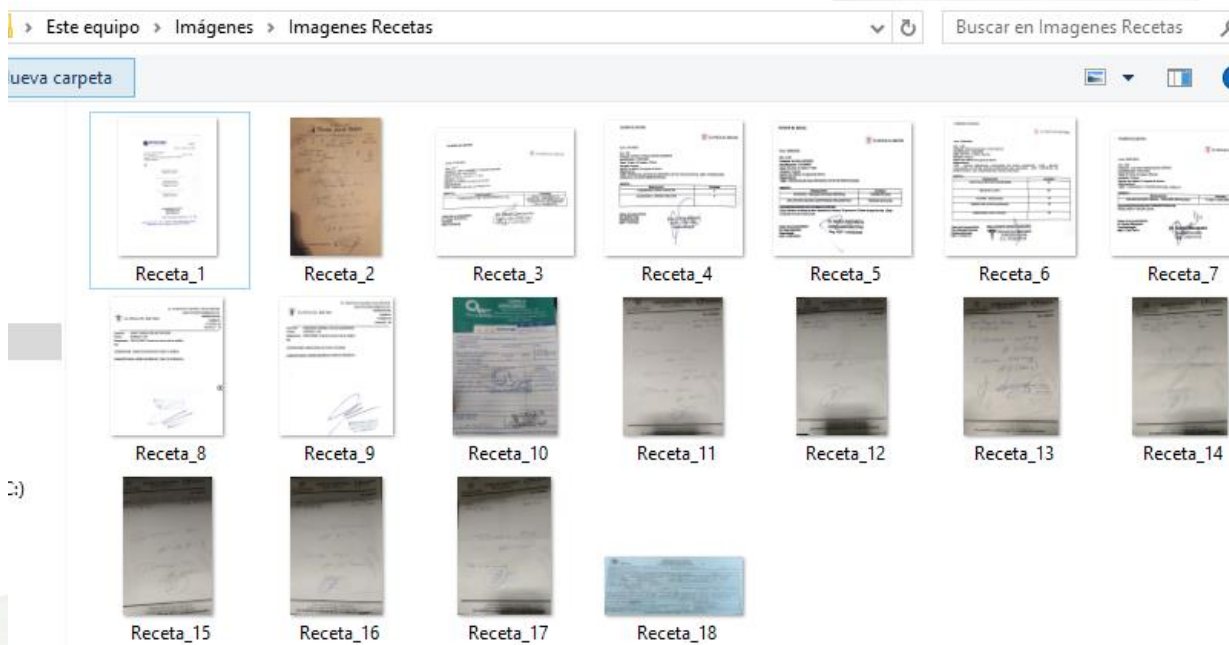
Selecciona una imagen:

Seleccionar archivo Ningún archivo seleccionado

Procesar Receta

← Volver a la Página Principal

Anexo 11. Modulo reconocimiento de recetas – Permite escoger una receta



Anexo 12. Modulo reconocimiento de recetas – Detección del texto y reconoce los medicamentos del inventario.

Texto Detectado

DUSPATALIN RETARD TAB 200 MGS. 30

185 BIOTIC: CAPS. 90

RIFAXIM: TAB 550 MGS. 10

DIGESTOTAL FORTE CAPSULAS. 15

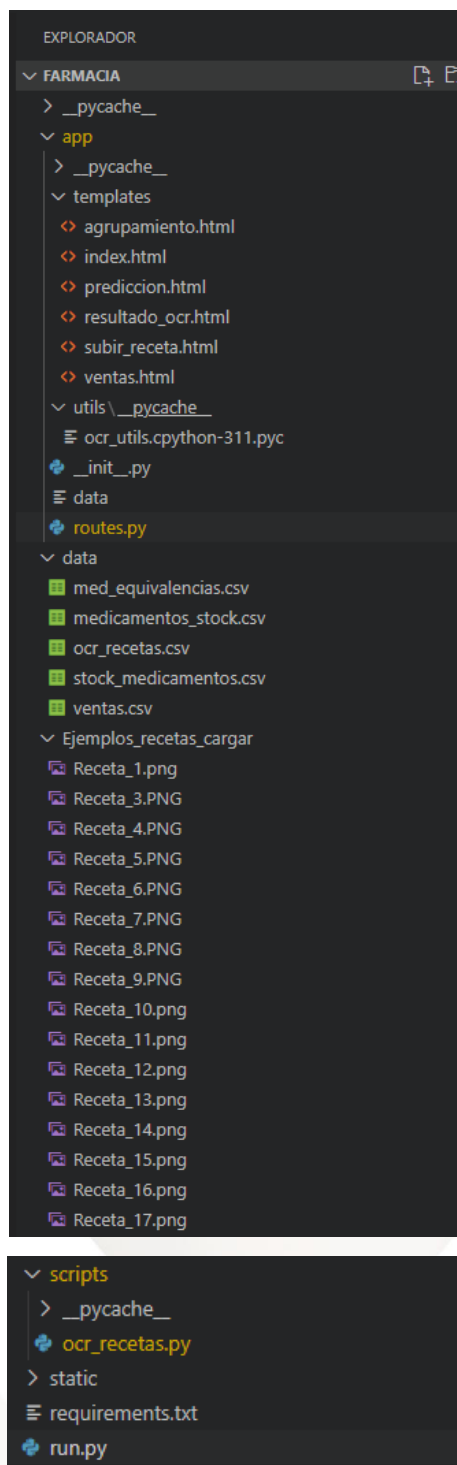
OMEPRAZOL CAPS. 20 MES. 30

Medicamentos Detectados y Disponibilidad

Medicamento	Dosis	Disponible	Cantidad en Stock	Precio	Equivalentes
Omeprazol	No especificada	✓ Sí	100	2.2	-
Digestotal	No especificada	✓ Sí	20	4.67	-

[Subir otra receta](#)
[Volver a la Página Principal](#)
[Ir a facturación](#)

Anexo 13. Carpetas y archivos que se han creado en visual code para el desarrollo del caso práctico



Anexo 14. Modelo Random Forest y árbol de decisión utilizado.

```
def prediccion():
    # División de datos en conjunto de entrenamiento y prueba
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

    # --- Modelo Random Forest ---
    rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
    rf_model.fit(X_train, y_train)
    y_pred = rf_model.predict(X_test)

    # Evaluación del modelo con métricas comunes
    mae = mean_absolute_error(y_test, y_pred)
    rmse = mean_squared_error(y_test, y_pred, squared=False)
    r2 = r2_score(y_test, y_pred)

    # --- Modelo Árbol de Decisión (para visualización e interpretación) ---
    dt_model = DecisionTreeRegressor(max_depth=5, random_state=42)
    dt_model.fit(X_train, y_train)

    # Exportar árbol para graficar (visualización de decisiones del modelo)
    dot_data = StringIO()
    export_graphviz(
        dt_model, out_file=dot_data, feature_names=features,
        filled=True, rounded=True, special_characters=True
    )
    dot_data.seek(0)
    graph = pydotplus.graph_from_dot_data(dot_data.getvalue())
    if graph is None:
        raise ValueError("Error al generar el gráfico del árbol de decisión.")

    # Convertir gráfico a imagen codificada para mostrar en HTML
    tree_img = BytesIO(graph.create_png())
    tree_img.seek(0)
    arbol_decision_img = base64.b64encode(tree_img.getvalue()).decode('utf-8')

    # --- Predicción actual para stock y generación de alertas ---
    stock_df = df[['codigo', 'nombre', 'precio', 'cantidad_disponible', 'fecha_vencimiento', 'dosis']].drop_duplicates()
    stock_df['dias_vencer'] = (pd.to_datetime(stock_df['fecha_vencimiento']) - pd.Timestamp.today()).dt.days
    stock_df['mes'] = pd.Timestamp.today().month
    stock_df['dia'] = pd.Timestamp.today().day

    # Función para codificar etiquetas nuevas sin error (manejo de datos desconocidos)
    def safe_transform(le, values):
        known_labels = set(le.classes_)
        values_str = values.astype(str)
        transformed = []
        for v in values_str:
            if v in known_labels:
                transformed.append(le.transform([v])[0])
```