

Planteamiento de modelos de Machine Learning para análisis de imágenes

Kevin Becerra – Laura Roncancio

Contexto

Teniendo en cuenta las ventajas que ofrece la inteligencia artificial, se busca obtener un modelo que prediga los productos que toman las personas en un supermercado, con el fin de asignar el valor de la compra sin necesidad de que pasen por la caja.

Para esto se busca un modelo de Machine Learning que aprenda a identificar la clase y subclase de producto de las imágenes que se le proporcionan. Ya teniendo un modelo con un alto nivel de precisión, se podría calcular el precio total de la compra.

Metodología

Se utilizaron 2 modelos de Machine Learning, esto para evaluar cuál es el que tiene más precisión, es decir, cuál ofrece mayor cantidad de aciertos.

- El primer modelo es uno de **Random Forest**, se realizó la carga de imágenes y se realizaron los respectivos arreglos para cada una de ellas, se tomó como variable objetivo el ID de la clase de producto.

Este modelo se encarga de crear árboles de decisiones independientes y así generar predicciones, sin embargo, es un modelo básico para el análisis de imágenes, por lo que puede generar un porcentaje de precisión bajo, generó 28,25% de accuracy.

- El segundo modelo es un **modelo de redes neuronales convolucionales (CNN)**, especialmente diseñado para el análisis de imágenes ya que son capaces de capturar patrones espaciales y visuales complejos, como bordes, texturas y colores, que son esenciales para la identificación precisa de productos. El modelo logró una precisión del 6%, lo que es considerablemente bajo. Esto indica que el modelo necesita ajustes importantes, como el uso de Transfer Learning o la recolección de más datos para mejorar la capacidad de generalización. La baja precisión puede deberse a la complejidad del problema y al desbalance de clases en el conjunto de datos.

Elección de Random Forest: La elección de Random Forest se hizo como un punto de partida rápido y económico para evaluar patrones básicos en los datos de imágenes de productos antes de invertir en modelos más complejos, mientras que la CNN se utilizó por su capacidad de extraer características complejas de imágenes, alineándose con las necesidades tecnológicas avanzadas de un supermercado inteligente; esto destaca la complejidad del problema y la necesidad de más datos y recursos.

Insights principales

- Distribución de categorías
 - Tanto para el conjunto de prueba como para el de entrenamiento, por una diferencia significativa, las categorías más predominantes en los datos son manzanas, jugo, leche, yogurt y melón, lo que sugiere que estas deberían recibir mayor atención en la optimización del sistema.
 - El desbalance de clases es evidente, lo que afecta la capacidad de los modelos para generalizar a categorías menos representadas.
- Limitaciones del modelo

- La CNN, aunque diseñada para datos visuales, no rindió bien debido a la falta de datos y a la complejidad del problema.
- La precisión del modelo Random Forest, aunque mejor que la de la CNN, sigue siendo insuficiente para un entorno de producción confiable.
- Impacto de la calidad de imágenes
 - La calidad y consistencia de las imágenes son críticas. Cambios en la apariencia de los productos, como maduración de frutas, y condiciones de iluminación pueden causar errores de predicción.

Recomendaciones para el negocio

A pesar de que el nivel de accuracy del modelo es bajo y el costo del margen de error sería bastante alto, se generan algunas recomendaciones teniendo en cuenta el entendimiento de los datos y el contexto del negocio:

- Implementar Transfer Learning
 - Usar modelos preentrenados como VGG16 o ResNet50 para mejorar el rendimiento de la CNN, aprovechando características aprendidas de conjuntos de datos masivos.
 - Se podría reducir la necesidad de una gran cantidad de datos y mejorar la precisión de manera significativa.
- Recolección y aumento de datos
 - Capturar más imágenes de productos desde múltiples ángulos y en diversas condiciones de iluminación para abordar el desbalance de clases.
 - Aplicar técnicas avanzadas de aumento de datos para generar más ejemplos y mejorar la capacidad de generalización del modelo.
- Optimización del sistema actual
 - Monitorear y ajustar la calidad de las imágenes capturadas por las cámaras. Implementar verificaciones regulares para asegurar que las imágenes sean claras y consistentes.
 - Considerar un sistema híbrido donde los productos que dependen del peso (como frutas y verduras) sean tratados de manera diferente, por ejemplo, solicitando a los clientes que pesen estos productos antes de agregarlos al carrito.
 - Con las categorías de producto con mayor relevancia se podría evaluar si están bien ubicadas en los mostradores, si se tiene suficiente stock y si se pueden hacer combos con categorías que tengan poca relevancia para así impulsar las ventas de estas últimas.
 - Teniendo en cuenta que las imágenes deben ser lo más claras posible, sería necesario mantener los productos en el mejor estado posible para evitar que al madurarse más de la cuenta cambien su color y esto cause errores en la predicción.
 - Finalmente, se sugiere evaluar el tratamiento de productos ingresados al almacén por los compradores, o estampados de bolsas (ejemplo hay bolsas ecológicas que tienen estampadas frutas).

Conclusión

El proyecto ha demostrado que un sistema automatizado es posible, pero requiere optimización para alcanzar niveles de precisión aceptables. Invertir en Transfer Learning y mejorar la calidad y cantidad de datos puede proporcionar un retorno significativo a largo plazo, mejorando la experiencia del cliente y optimizando las operaciones del supermercado.