

Machine Learning in Grocery Stores

Background

By: Aidan Housenbold

Before machine learning, people had to wait in long lines at grocery stores and their shopping was less efficient. Also, staffing shortages made some stores struggle to operate. Adding technology to make autonomous stores allows for quicker shopping, a more stable economy, and data to refine supply for a stores customers. I like autonomous shopping for convenience; I have visited one of the stores and I loved being able to walk right out instead of waiting in long grocery lines. I tried out autonomous shopping in an airport and it allowed me to get a snack quickly without missing my flight.

Recent Developments



The Halo Cart: To combat Japan's aging population and subsequent labor shortage, Halo cart implements a camera ring and artificial intelligence to create contactless checkout with the basket



Amazon Go: Amazon go makes a just walk out shopping experience using cameras and machine learning to track and identify each person and item as they move through the store.

Bias

The main type of bias is overfitting. There are so many inputs for the model due to the unique body shapes, movements, item shapes, and item labels that the model considers too many of the small data and fails to track if an item was taken or not. I think this can be a big issue in the entangled phase of recognition. Couples, friend, and strangers will all interact differently when close together, and it could be hard to detect who takes what and if they are keeping the item shared. I think more training data that is refined through giving simulated training data would help eliminate the over fitting

Technical Aspects

Amazon has been a little vague on how they are achieve each task, but, they have said they use computer vision based machine learning models and neural networks to create a “just walk out” shopping experience. Amazon uses ML to solve for 4 main issues for Go stores:

Person Detection

The go store uses RGB video and distance calculation to identify each person and track them. However, they run into issues when and entangled space happens, meaning two people are overlapping. To solve for this, Amazon uses an ML model to guess who is who, then mark each person in the entangled state to later confirm who they are.

Object Recognition

The goal of these programs is to detect when an item is picked up off a shelf and in someone's hand. The two main issues is detecting similar products and unideal lighting deformities. To identify the right product amazon uses residual neural networks that do refined product recognition after the convolution neural network classifies the item. The lighting issues is solve though using of training data from similar situations.

Pose Estimation

Amazon used top down cameras initially, but they did not work well enough to track customers, so Amazon created a model that makes a stick figure of each person. They used convolution neural network with a cross entropy to identify each joints, then a linear regression to connect the figure and pairwise function to group the connections together to create a human stick figure.

Activity Analysis

There is almost an infinite amount of poses that a human can have when picking up an object, so the issue is that there is not enough training data to collected from actual people to train a model in activity analysis. Instead, Amazon created a simulator of people and fed it to the machine to speed up the train process. About a months worth of traditional training data could be processed in a day with the simulated models.