

DMV302 Data Analysis and Pre-Processing

Model Report + Data Analysis

Introduction

Data mining is an important and key part of the retail industry. Supermarkets around the world benefit from data mining in so many ways every day. This report details some ways that data mining can be used at Woolworths to great effect. The data mining uses are Customer Segmentation, Market Basket Analysis, Sales Forecasting, Price Optimization, Production Recommendations Systems and Fraud Detection.

Customer Segmentation

Customers can be grouped and organized into categories, demographics and other relevant things that can be useful for Woolworths. By segmenting customers Woolworths can target specific groups of people with tailored marketing campaigns, offer personalized promotions and improve overall customer satisfaction. For instance, energy drinks can be targeted at younger adults specifically. This increases the likelihood of companies wanting to buy ads and promotions with Woolworths. Another way this can be used is by giving newer members more introductory deals and older members different kinds of deals depending on the needs of Woolworths. The main way this can be done is through cluster algorithms like k-means or hierarchical clustering. The data required for this can be wide in variety depending on the kind of segmentation Woolworths wants to do, some basic ones would include customer data like age, join date, average shopping spend and more. This data can be collected mainly through the membership program at Woolworths.

Market Basket Analysis

Analyzing what products are bought together can provide a great insight into your store and how to optimize it. By collecting data on what people buy together in one shop we can increase impulse buying and store layout optimization. Through data mining we can find that it's common that someone who buys eggs and milk is also likely to buy something essential such as bread. With this knowledge these 3 items can be placed in proximity to each other to thus increase the chances of them buying the bread and to increase the customer throughput of the store. The data required for this can be found in sales history; if the basket data is not stored, a system to store each customer's shop/basket must be implemented. This form of data mining works by associated rule systems and finding common patterns.

Sales Forecasting

Being able to predict what sales Woolworths will get in the next x amount of time is a valuable thing to know. Things like stock required, staff required, and revenue gained can all be predicted leading to the smoother operation of Woolworths. There are many cases where stock management and staff rosters can become imbalanced due to busier months, promotional periods and even holidays. With this data mining system implemented Woolworths will know how much stock to order and how many staff to roster all year round including the holidays, which is where things can be much harder to predict manually. The data needed can easily be obtained by looking at the past sales history of Woolworths not just in the store but also in the local area and national Woolworths stores. This technique is done through regression analysis or time series forecasting algorithms.

Price Optimization

Pricing your products at Woolworths could be revamped to include data mining technology. Using data mining we can determine the optimal pricing of the products based on things like demand, supply, competition prices and customer behavior. Instead of manually prices items, automatic systems can be set up to determine the price of the products, reducing the workload on your staff. Using sales data the best price can be found that boosting your profits. This can be done by looking at sale history and using predictive algorithms.

Product Recommendation Systems

Recommending products to shoppers online can be a great way to increase sales and profits. This can be done by suggesting products based on the user's purchase history and preferences. Additionally, we can suggest products to user based on what they added to their online basket like if a user adds bread, we could suggest butter or jam as well as products they maybe didn't know existed that they can be interested in. This technique can be combined with customer segmentation to give tailored suggestions based on their age group, gender, tendencies and more. All these can lead to better customer experience and higher profit margins.

Fraud Detection

Fraud is an ongoing battle at Woolworths both online and in-person. Data mining can help alleviate this problem and reduce fraud. We can detect suspicious transaction patterns

and flag them as being dangerous. This will help protect Woolworths from fraud and help maintain customer trust. This is done with anomaly detection algorithms.

Challenges

Data mining does come with its challenges. The first and most obvious challenge is collecting of the required data, luckily at Woolworths there are already a lot of good systems that collect all kinds of data, but this may need to be expanded and then centralized to a main server/database. Keeping the sensitive data safe is a priority and a great challenge. We must make sure that the best levels of cyber security are used and only very trusted staff members like managers and IT staff have access to this data. Ensuring that data collection is ethical is needed too. Another challenge is the preparation of the data as the data may be stored inconsistently and in different formats.

Conclusion

In conclusion, Woolworths would benefit greatly from data mining in a variety of ways that I have outlined. Greater customer engagement and satisfaction will make customers enjoy shopping at Woolworths and will give customers an overall better experience. Higher profits and better management of stock and staff are some other benefits of introducing data mining into your company. Data mining does come with its challenges that can be overcome.