Intro to Analytics Modeling HW 9

Shelf Space Optimization

2024-07-16

Question 18.1
Describe analytics models and data that could be used to make good recommendations to the retailer. <u>How much shelf space should the company have, to maximize their sales or their profit?</u>

Of course, there are some restrictions – for each product type, the retailer imposed a minimum amount of shelf space required, and a maximum amount that can be devoted; and of course, the physical size of each store means there's a total amount of shelf space that has to be used. But the key is the division of that shelf space among the product types.

For the purposes of this case, I want you to ignore other factors – for example, don't worry about promotions for certain products, and don't consider the fact that some companies pay stores to get more shelf space. Just think about the basic question asked by the retailer, and how you could use analytics to address it.

As part of your answer, I'd like you to think about how to measure the effects. How will you estimate the extra sales the company might get with different amounts of shelf space – and, for that matter, how will you determine whether the effect really exists at all? Maybe the retailer's hypotheses are not all true – can you use analytics to check?

Think about the problem and your approach. Then talk about it with other learners, and share and combine your ideas. And then, put your approaches up on the discussion forum, and give feedback and suggestions to each other.

You can use the {given, use, to} format to guide the discussions: Given {data}, use {model} to {result}.

One of the key issues in this case will be data – in this case, thinking about the data might be harder than thinking about the models.

Breakdown of the Problem Statement

- Hypotheses
    1. More shelf space -> more sales
    2. More sales -> more complementary product sales
    3. Larger effect if complementary products are adjacent
- Problem Statement
    - Maximize sales
        - Constraints
            - Min / Max amounts for each product
            - Total space in store
    - Measuring how changes in the amount of shelf space dedicated to different product types affect the overall sales and profit of the store
        - I might need to know how much extra sales the company might get with different amounts of shelf space

Proposed Solution

Steps

1. Using the hypothesis 1, get the amount of the shelf space allocated for each product that maximize the sales. We need to set up a linear programming optimization model (Model1)
2. For the hypothesis 2, set up a new linear programming optimization model (Model2)
3. For the hypothesis 3, set up a new linear programming optimization model (Model3)
4. Compare the Model1 with Model2
    1. Check which model shows higher sales. If Model 1's sales are greater than Model 2's sales, we disprove that more sales mean more complementary product sales, and vice versa.
5. Compare the Model1 with Model3
    1. Check which model shows higher sales. If Model 1's sales are greater than Model 3's sales, we disprove that larger effect if complementary products are adjacent, and vice versa.

For step 1:

Required Data (for each customer)

- Min/Max amount of shelf space allocated to each product *i*
- Sales or profit per unit of shelf space for each product *i*
- Total available shelf space in the store, *T*

Given the set of required data, we can use a linear programming optimization to maximize the total sales. Once we obtain the optimal result (Model1), the output will include the amount of shelf space allocated for each product to maximize the sales. This can be used to prove/disprove the hypotheses 2 & 3.

Here's a set up for the linear programming optimization.

Decision Variables:

- x_i: Amount of shelf space allocated to product *I*

Objective Function:

- Maximize total sales. Let's denote S_i as the sales per unit of shelf space for product i.

$$Maximize \sum_i S_i * x_i$$

Constraints:

- Minimum and Maximum shelf space for each product:

$$L_i \leq x_i \leq U_i \forall i$$

- Total shelf space

$$\sum_i x_i \leq T$$

For step 2:

Required Data (for each customer)

- Min/Max amount of shelf space allocated to each product *i*
- Sales or profit per unit of shelf space for each product *i*
- Total available shelf space in the store, *T*
- Also, assume that we can estimate the additional sales gained by the complementary products *i* and *j*. It's denoted by C_{ij}
  - For the additional sales gained by the complementary products sales, please refer to the Appendix section at the end.

Given the set of required data, we can use a linear programming optimization to maximize the total sales. Once we obtain the optimal result (Model2), the output will be compared with the output from Model1 and the comparison result will be used to prove/disprove the hypotheses 2. If Model 1's sales are greater than Model 2's sales, we disprove that more sales mean more complementary product sales, and vice versa.

Here's a set up for the linear programming optimization.

Decision Variables:

- x_i: Amount of shelf space allocated to product I

Objective Function:

- Maximize total sales. Let's denote S_i as the sales per unit of shelf space for product i.

$$Maximize \sum_i S_i * x_i + \sum_{i<j} C_{ij}$$

Constraints:

- Minimum and Maximum shelf space for each product:

$$L_i \leq x_i \leq U_i \forall i$$

- Total shelf space

$$\sum_i x_i \leq T$$

For step 3:

Required Data (for each customer)

- Min/Max amount of shelf space allocated to each product i
- Sales or profit per unit of shelf space for each product i
- Total available shelf space in the store, T
- Additional sales gained by the complementary products i and j. It's denoted by C_{ij}
- Binary variable indicating if product i and j are placed adjacent to each other

Given the set of required data, we can use a linear programming optimization to maximize the total sales. Once we obtain the optimal result (Model3), the output will be compared with the output from Model1 and the comparison result will be used to prove/disprove the hypotheses 2. If Model 1's sales are greater than Model 2's sales, we disprove that more sales mean more complementary product sales, and vice versa.

Here's a set up for the linear programming optimization.

Decision Variables:

- x_i: Amount of shelf space allocated to product I
- y_{ij}: Binary variable indicating if product i and j are placed adjacent to each other ($y_{ij} = 1$ if adjacent, $y_{ij} = 0$ otherwise)

Objective Function:

- Maximize total sales. Let's denote S_i as the sales per unit of shelf space for product i.

$$Maximize \sum_i S_i * x_i + \sum_{i<j} C_{ij} * y_{ij}$$

Constraints:

- Minimum and Maximum shelf space for each product:

$$L_i \leq x_i \leq U_i \forall i$$

- Total shelf space

$$\sum_i x_i \leq T$$

For step 4:

Comparison between Model1 & Model2

Comparison of the maximum sales between Model 1 and Model 3. The outcomes will be used to prove or disprove Hypothesis 2. If Model 1's sales are greater than or equal to Model 2's sales, we disprove that more sales mean more complementary product sales, and vice versa. (In the case where Model 1 performs better, we want to consider equal sales too because Model 2 has potential additional sales from complementary product pairs.)

For step 5:

Comparison between Model1 & Model3

If we prove that more sales mean more complementary product sales, we can now test whether there is a larger effect if complementary products are adjacent. We compare the maximum sales between Model 2 and Model 3. The outcomes will be used to prove or disprove Hypothesis 3. If Model 2's sales are greater than Model 3's sales, we disprove that there is a larger effect when complementary products are adjacent, and vice versa.

Appendix

Derivation of the additional sales gained by the complementary products sales.

Required Data

- Historical sales data for all products
- Identify pairs of complementary products through domain knowledge. (This can be done through market basket analysis but we assume that there exits a domain knowledge for this and it's more accurate.)

Given the set of required data, we use linear regression modeling to estimate the additional sales of product. Note that we assume that there's no other factors such as promotions, seasons, and holidays for simplicity. Here's a detailed set up for the linear regression modeling for the additional sales estimation.

Step 1: Fit a linear regression model where the dependent variable is the sales of product *j* and the independent variable is the sales of product *i*:

$$Sales_j = a + b * Sales_i$$

where

- $Sales_j$ is the sales of the complementary *j*
- $Sales_i$ is the sales of the primary product *i*
- a is the intercept and *b* represents the impact of the sales of the product *i*

Step 2: Estimate the additional sales by using the fitted model from Step 1:

$$\Delta Sales_j = b * \Delta Sales_i$$

Where $\Delta Sales_i$ represents the change in the sales of product i.