Anomaly Order Detection and Response

Subhasish Basak
Chennai Mathematical Institute

The business problem

How can Tierra Del Fuego better fulfill Pazzo's orders—even anomalous ones—and prevent cuts?

Steps to follow

We start with descriptive analysis to get the essential insights from the data. Simple EDA is done to identify the SKUs which are more prone to cuts.

➔ **Step 1**

Once the cut prone SKUs are identified restrict main focus on them only

➔ **Step 2**

Use plots/charts to visually identify patterns/trends in the features of these restricted SKUs

➔ **Step 3**

Focus more on important features which behaves distinctly in case of the selected SKUs

Modelling in a naive approach

Let us predict whether TDF is going to face a cut for a week...
A Naive approach would be Logistic Regression

Output will be boolean -> `1` : cut & `0` : no cut
Input features can be -> Ordered units, Pageviews from the previous week

For a more efficient approach, we should try :

- eliminating non significant features
- fitting the model separately for each SKU

P.S

Logistic regression is one of the simplest and widely used Classifier

predicting a cut for the upcoming weeks is not enough !!!

Afterall TDF wants to reduce cuts, hence it needs to predict future orders beforehand…

But we can use our knowledge of possibility of a cut.

Whenever there is a possibility of a cut according to our Naive Logistic regression model, we go on estimating the orders

Tip

orders can be a treated to be a Time Series.

Graphical visualization can be used to observe patterns in consumer orders for different SKUs

Feature Extraction

Pazzo has provided as set of weekly information for each SKU, corresponding to 2 years. viz.

- On Hand Inventory
- Pageview Out of Stock
- Average Price
- Change in Pageviews
- Forecasts

And our variable of interest is:

- Consumer Ordered Units

Time series analysis

We can model **Consumer Orders** as Auto-Regressive process. The intuition behind it is,

Past orders tend to impact the present and future orders as well. A typical Auto regressive model is of the following form,

$Y[t] = a^*Y[t-1] + e[t]$

Where Y[t] is the observation corresponding to time t and e[t] is the random error corresponding to the same time point.

Not only Auto-regressive !

ARIMA gives a moderate fit, we should try to incorporate the effects of other features as well..

➔ **Which features ??**

Select features from the given list

Or

Generate new features out of them

➔ **How to incorporate ??**

Using a mixture model..

A multiple regression model with Auto-regressive part

Feature generation..

- ## Weighted forecasts

Forecasts for a particular week is available with 1,4 & 8 week period. We combine them to produce a single figure associated with weights inversely proportional to their lookahead period.

- ## Change in price

Changes in prices are more significant than actual price, thus we compute the changes from the raw price figures.

P.S

The selection of the weights in computing the forecast can be estimated also.

i.e. can be incorporated as a hyperparameters in the model

The mixture model..

Our output variable is Consumer Order (Y_t)

Features:

- Weighted forecasts (W_t)
- Change in price (P_t)
- Pageview Out of Stock (Pv_t)
- Change in Pageviews (dPv_t)

the "_t"s denote the timestamp for all the variables

We added a random noise in the model.

$$Y_t = \alpha_1 + \alpha_2 W_t + \alpha_3 P_{t-1} + \alpha_4 Pv_{t-1} + \alpha_5 dPv_{t-1} + \sum_{t=1}^{m} w_t Y_{t-1} + \varepsilon_t$$

Estimation

- We have the parameters in the model that can be used by minimizing the error sum of squares, using the method of least squares.
- For the given data we compute error sum of squares from the given Consumer Orders and the predicted orders from the above model

The auto-regressive order (m) can be determined

➔ Subjectively : Using Auto-correlation plot

➔ Analytically : Using AIC optimizing criterion

—

The mixture model can be used to make predictions at one week beforehand on the Consumer orders and TDF can adjust its manufacturing process accordingly.

The model should be trained separately for different SKUs, since each of the item has different demand-supply characteristics.

Extension possibilities:
The model can be extended incorporating more features. If training data for more than 2 years is available the model can yield more accuracy.

Tip

Ideally, speak of people in very different situations, but where each could benefit from your solution.