Optimizing SIG’s sales and operations planning by forecasting customer demand for packages in Thailand market

Professor: Galit Shmueli
Team members: Pei Pei Chen, Lynn Pan, Saskia Thus, Amber Lin
Introduction

- SIG is a global company that produces packages and filling machines for food and beverage.
- SIG has multiple different clients in Europe and Asia.
- SIG customizes the package with different sizes and print to meet client’s demand.

Asia Pacific Market

Thailand, Indonesia, Vietnam, Taiwan, Korea, Bangladesh, Malaysia, India, Philippines.

Our focus

Thailand
**Forecast Goal**

**Forecast objective**

Product demand (1,000 unit)

One year forecast per month

Train | Valid | Forecast

Time

- 2009/1
- 2017/12
- 2018/12
- 2019/12

**Purpose**

Give SIG accurate forecasts on which they can base and optimize their sales and operations planning.
Workflow

Series more than 2 years

Regular customer

Last 2 years all zero

Past customer

Forecast value all="0"

New customer

Active Customer

Seasonal Naive

Moving Average

Auto Arima

Linear regression

Use best Model to Predict the 2019 data
Data - Thailand

2009 to 2018 (monthly)

120 records in each series
85 series in total

Thousands of package unit

No trend & seasonality | trend | seasonality
---|---|---
52 (61%) | 29 (34%) | 4 (5%)

Using ets model to find the XYZ
Methods

Best Forecast:
- Moving average
- Exponential smoothing (ets)
- Linear regression
- Auto arima
- Seasonal Naive forecast

Apply different model to different pattern of product categories.

Quick Forecast:
- Moving average

Benchmark:
Seasonal Naive forecast

Evaluation:
- V RMSE
- V Forecast Error plot

Train Period: 2009-2017
Validation Period: 2018
Performance & Result

Benchmark

- Seasonal Naïve Error Distribution
- RMSE: 168572.19

Best Model

- Best Model Error Distribution
- RMSE: 127786.6

Quick Model

- Moving Average Error Distribution
- RMSE: 216227.1696

Reduce overfitting
Future work: External linear regression

External data: sales person’s manually forecast

By adding external data to the model, the prediction is closer to the actual result.

Nov, Dec
Operational Requirements/Constraints

Data collection
1. Ongoing analysis requiring collecting new data.
2. External data (SOP data) requires collecting new data.

Company Policy
1. The limit amount of sales data is close to the forecasting goal (k=12) months, the more data is added to the time series, the forecast performance would become more stable.
2. Policy changed every two years including discount and product generation. This influences the orders and the forecasts.

Interface
Forecast the daily traffic from Facebook fan page to content website for TC Incubator

Team 3 Jay Lee, Sam Kuo, Astro Yan, Serina Hung
Base on result from this week to plan next week posting.

Marketing team has no serious A/B testing method to plan the posting strategy.

See the posting result.
**Business Goal**

**Goal**

Provide a tool that TC will be more convenient to compare the traffic from the different categories of posting.

**Stakeholders**

Social Media Marketing department (Social Manager)

**Challenge/Opportunity**

1. More Efficiency for the planning of posting.
2. Increase website traffic.
3. Apply it to other business client.
4. Without data analyst maintain and implement forecasting.
"Predict traffic from the Facebook page into the TC official website in the next 7 days, with different categories of posts."
Forecasting Process

- Define business & forecasting goal
- Get data: Both GA and performance of Facebook posts
- Explore & Visualize series: By using Tableau
- Pre-process data:
  - Aggregate data (from hourly to daily)
  - Label & theme the theme of the Facebook post (include Jinrih toolbox, Jinrih planning, Jinrih growth, Jinrih check-in, Jinrih attitude, Jinrih brand)
  - Exclude outlier data
  - Check each lag of value & 6 theme of post
- Implement: Provide TC a tool which evaluate in what time post which kind of theme could get best traffic to their content website.
- Evaluate & Compare performance:
  - Chart
  - MAPE
- Apply forecasting method:
  - Linear Regression
  - Neural Network
  - Ensemble
- Partition series:
  - The last two week validation period
Data Description

Source: Google Analytics set up by TC
Measure: hourly traffic
Time period: 2018.08.01~2018.11.31
Type: Hourly

Pre-processing

Aggerate data (hourly - daily) → Remove the outlier → Label & Map the theme of post → Check each lag of value & six themes of post

<table>
<thead>
<tr>
<th>Time</th>
<th>Traffic (Total in day)</th>
<th>Jinrih toolbox</th>
<th>Jinrih planning</th>
<th>Jinrih growth</th>
<th>Jinrih check-in</th>
<th>Jinrih attitude</th>
<th>Jinrih brand</th>
<th>External information</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018/9/1</td>
<td></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2018/9/2</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Method

Benchmark
- Naive

Model
- Regression, Neural Network, Ensemble
  (with external data including 6 different posting theme and Lag)

Performance
- 1. Chart
- 2. MAPE
Results & Evaluation

**Series**
- Ensemble
- NNet
- Linear
- Actual
<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>347.6232</td>
<td>52.999353</td>
</tr>
<tr>
<td>NNet</td>
<td>421.7101</td>
<td>57.05525</td>
</tr>
<tr>
<td>Ensemble</td>
<td>287.8759</td>
<td>40.70686</td>
</tr>
<tr>
<td>Naive</td>
<td>322.7974</td>
<td>45.5156</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>504.17</td>
<td>100.417</td>
</tr>
<tr>
<td>NNet</td>
<td>417.9051</td>
<td>68.5472</td>
</tr>
<tr>
<td>Ensemble</td>
<td>373.5245</td>
<td>68.40184</td>
</tr>
<tr>
<td>Naive</td>
<td>322.7974</td>
<td>45.5156</td>
</tr>
</tbody>
</table>
Implement

Update the traffic data (From GA)

Fill the past data in the excel file

Run R code and get the suitable post strategy next week.

Use the forecast outcome.

Social media manager can try new post strategy for next week of post theme

Control Group

Test Group
Recommendation

We suggest them to use forecast as a tool to do A/B testing more serious.

We suggest them to record the adjustment and promotion they have done in A/B testing, cause variables effect a lots.

If shiny interface is needed by client, we can assist them to conduct it.

In the future, we can combine outsource API to automate data collection to offer complete solution.