

# GROUP OUTLYING ASPECTS MINING PREDICT FUTURE SALE

Shaoni Wang<sup>1</sup>, Gang Li<sup>2</sup>MingZhu Kang

<sup>1</sup> Xi'an Shiyou University, China

<sup>2</sup> Deakin University, Australia

## Introduction

In deep learning projects, data visualization operations often need to be carried out, including visualization of original image data, visualization of loss and accuracy, etc.

In this study, the data in the data set are collated and analyzed to analyze the impact of commodity sales volume, product category, operating income and the economic scale of the city where the stores are located on the income.

The model is used to train the data set, and the historical sales data is used as the model feature to realize the commodity sales forecast. Effectively guide the store to conduct reasonable inventory management.

## Data Sets

- Preprocessing of project data sets  
The preprocessing of project data set includes training set, commodity set, commodity data set, commodity category set and test set.  
through the data of the training set we determined the parameters of the fitting curve to filter the obvious outliers and checked the abnormal conditions of prices.
- Training set data cleaning  
through the data of the training set, we determined the parameters of the fitting curve.

## Structured Data And Analysis

Sale Analysis



The 11<sup>th</sup> International Conference on Knowledge Science, Engineering and Management (KSEM 2018), 17-19/08/2018, Changchun, China

Revenue in the 23rd month was up sharply from the 11th month, while sales were down year-on-year. One item's gross revenue was unusually high.

Product Category Influence  
The influence of product category on sales volume The first three categories that contribute the most to total sales are the 40th, 30th and 55th categories.

## Predict Future Trend

By processing closed stores and goods that are not for sale, Keep only stores open for the last six months and items that sell. Combining the features of the store data set and the merchandise data set.

The goal is to use historical sales data to predict future sales. Using the historical sales data as the characteristics of the model, this month's sales results as labels to build a model for regression analysis.

## Model Training

## Conclusion

~~Formalize the problem of Group Outlying Aspects Mining by extending outlying aspects mining.~~  
~~Propose GOAM algorithm to solve the Group~~Through this project, I have learned a lot, including the effective aspects of problem cutting, code implementation of analysis algorithm, design of analysis process, etc.

~~Outlying Aspects Mining problem.~~  
~~Utilize the pruning strategies to~~From data analysis methods to feature engineering and prediction model construction, a lot of time has been spent to study and comb.

~~reduce time complexity.~~In the process of predictive analysis, the theoretical and data support for feature analysis and model construction is not concise and powerful enough, which needs to be strengthened.