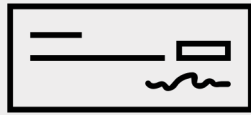




product profile

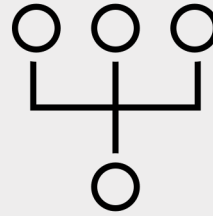
Skin Care & Cosmetics

Chen Qi

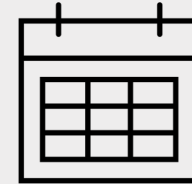


- Background
- Labeling System
- Data Collecting

Qi Chen



- Model building
- Training
- Results
- Evaluation
- Deploy & Update



- Application
- Modification

Product Profile and Sentiment analysis

Background

Pain points

- Lack of knowledge on customer preferences
- Mismatch of customer preferences and operations
- Passively answer questions, while they shall lead the conversations and lean towards converting
- Mass based recommendations with key products rather than individually customized

Qi Chen

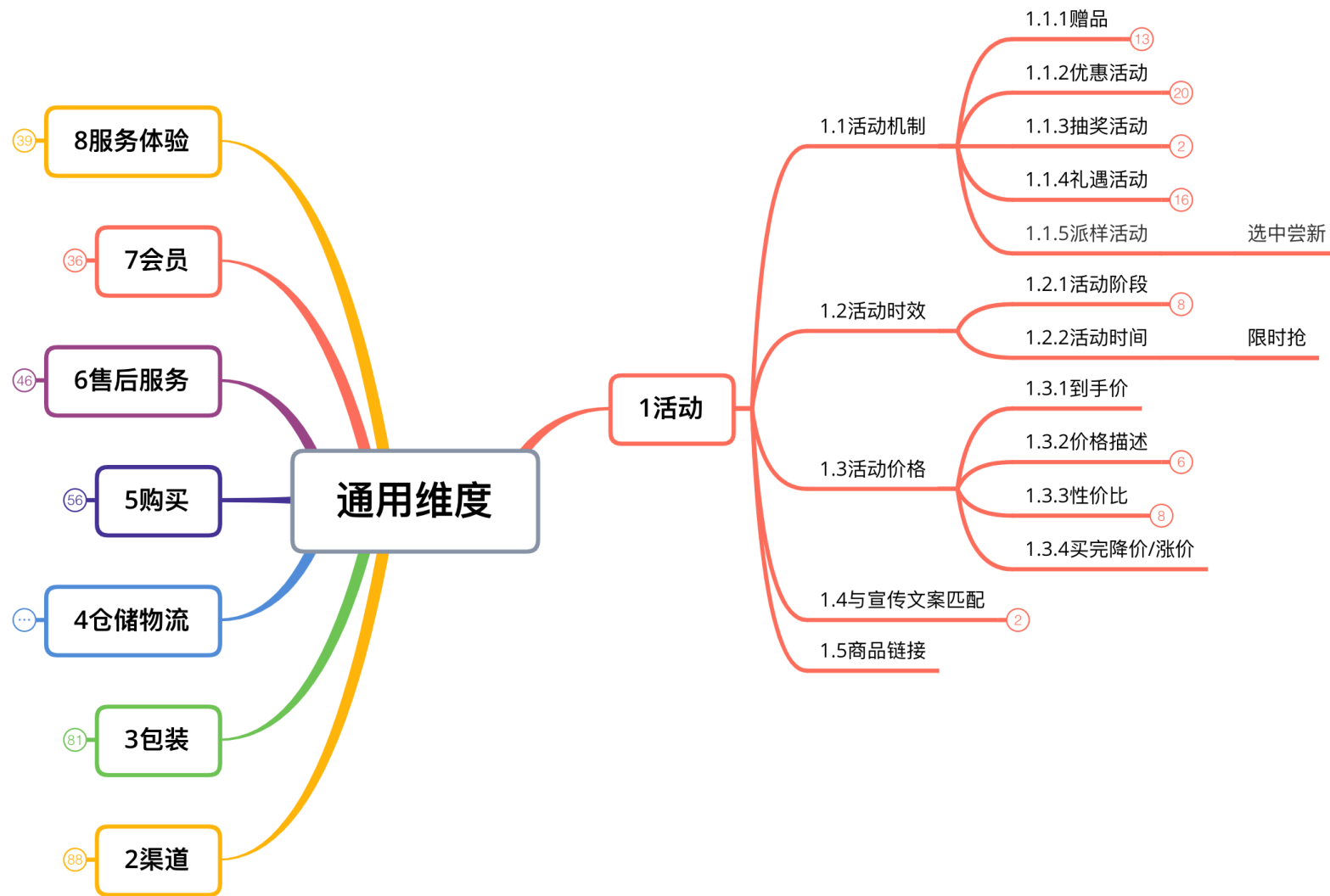
Target

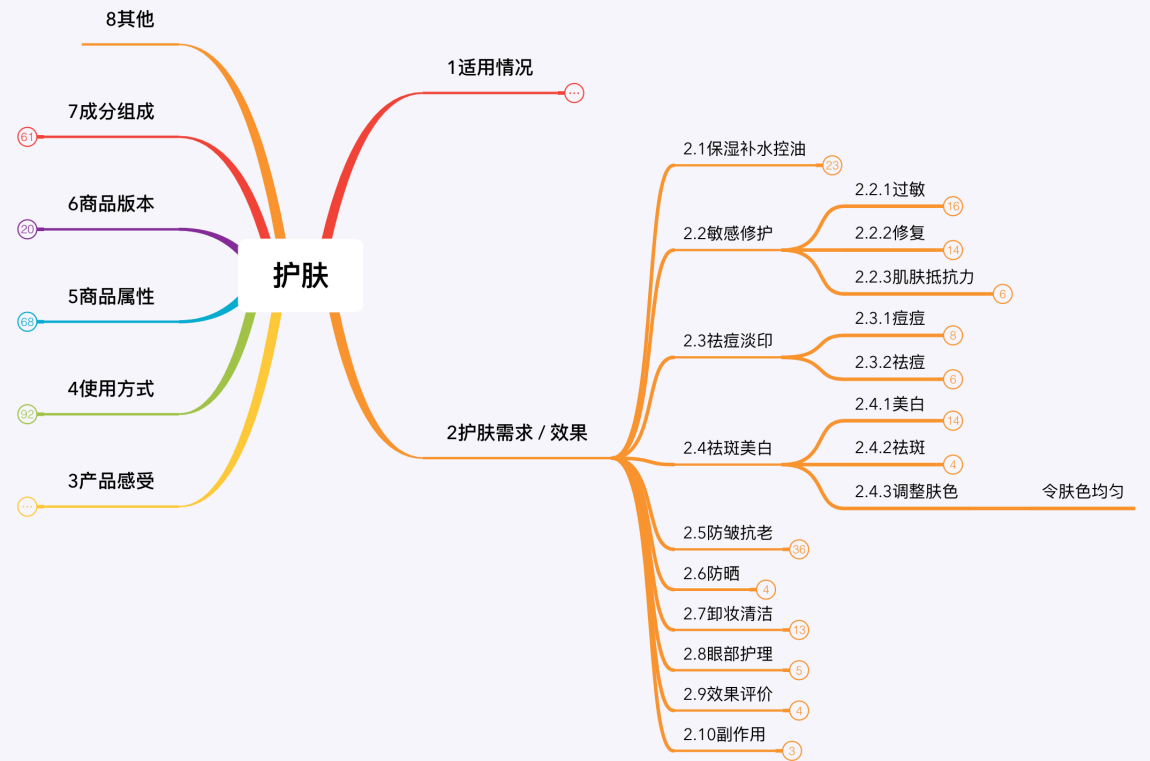
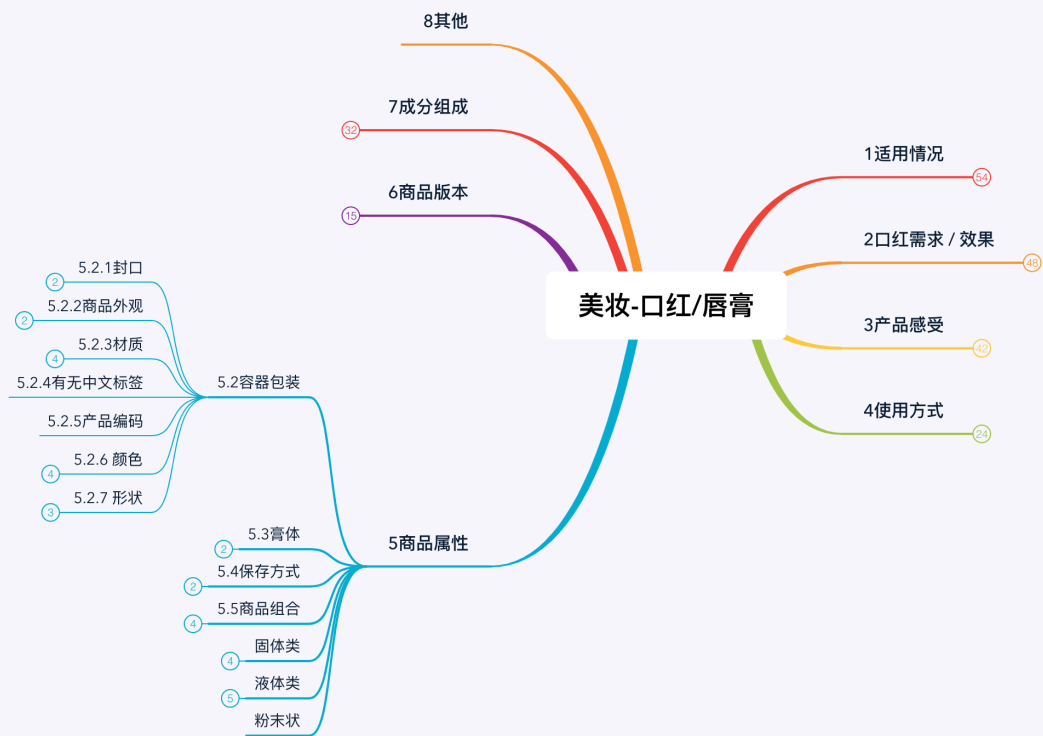
- To Provide EC customers the bricks and mortar stores-like shopping experience
- To Improve EC customer Service representatives capability towards Beauty Assistant
- Improve customer experience with prompt feedbacks on customer post-order related follow up questions

Product Profile and Sentiment analysis

Labeling System

- General
- Domain category
- 3 Levels





Qi Chen

Product Profile and Sentiment analysis

Data Collecting

- Brand Classification
- Category
 - Skin Care
 - Cosmetic
 - Food
 - Baby Product

Comment	Topic Label	Sentiment Label
价格便宜	G1.3.2	1
然后这款水质地比较粘稠	S3.1.1	0
而且卖家服务态度超级好	G8.1.3	1
卖家发货很快	G4.2.1	1

Model Building

Task

1. Sentiment analysis
2. Topic classification

Approach

1. Data Pre-Processing
2. Multi-label Classification
3. Training
4. Optimization
5. Evaluation

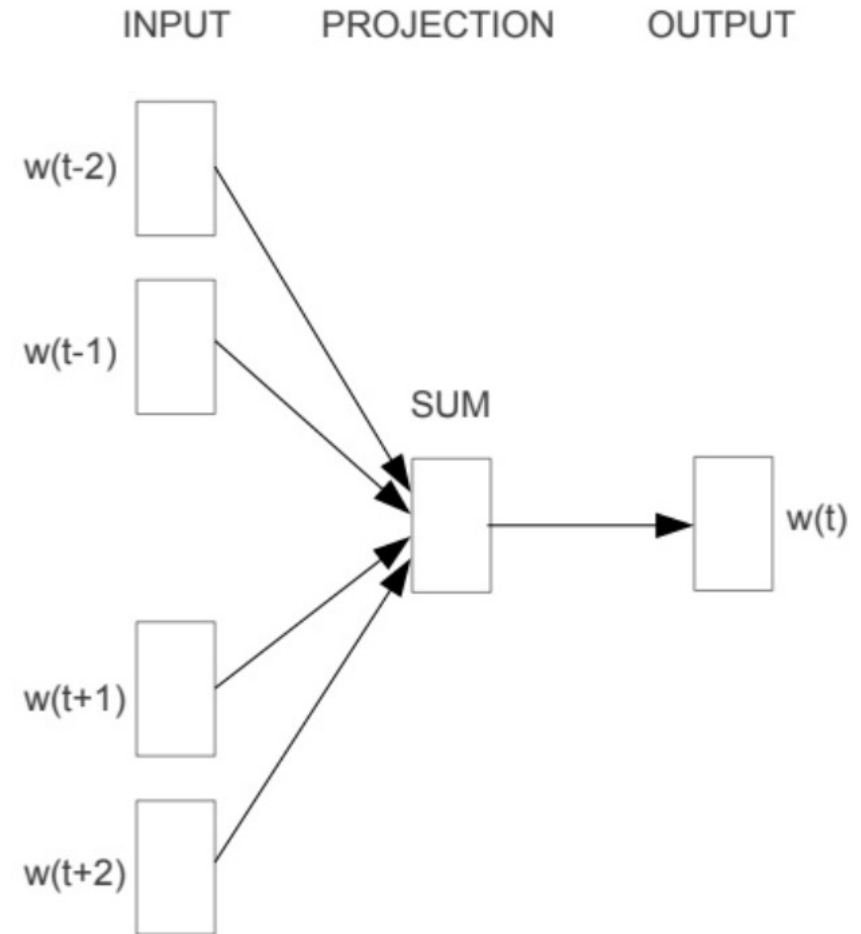
Qi Chen

Product Profile and Sentiment analysis

Data Pre-Processing

- IF-IDF
- Word2Vec
- BERT

1. convert labels into numbers
2. Enrich corpus with domain words
3. words segmentation by 'jieba' library
4. words padding to constant length
5. CBOW to words embedding



CBOW

Classification Model

SVM

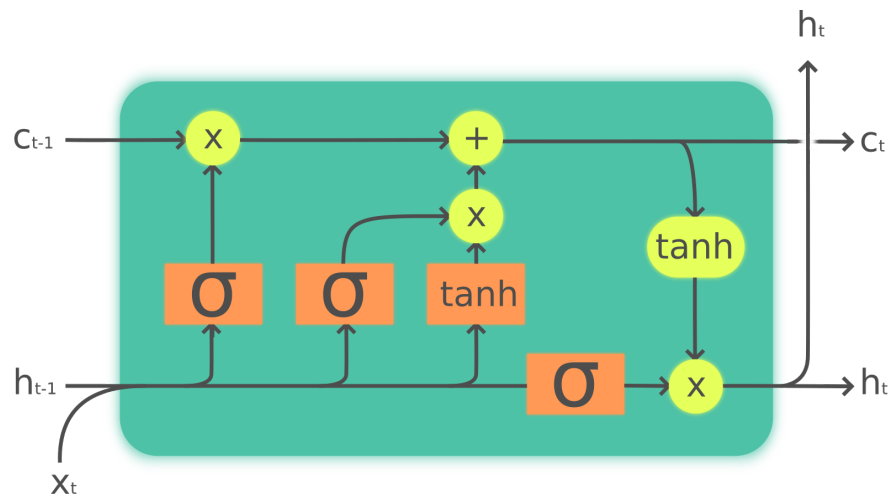
Logistic Regression

Long Short Term Memory (LSTM)




Bidirectional LSTM

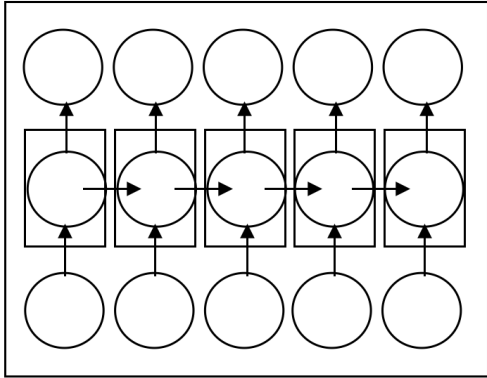
Gated Recurrent Unit (GRU)

CNN

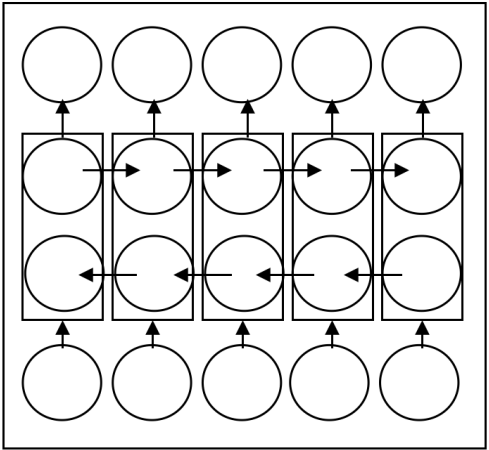


Legend:

	Layer	Pointwise op	Copy
			



(a)



(b)

Structure overview
 (a) unidirectional RNN
 (b) bidirectional RNN

Training & Evaluation

- Output Probability: $Softmax = \frac{e^{\theta_i^T x}}{\sum_{j=1}^k e^{\theta_j^T x}}$
- Corss-Entropy Loss: $H(p, q) = - \sum_x p(x) \log q(x) \rightarrow = q(x) - p(x)$
- Back propagation Through Time
- Gradient Descent for hyper-parameters update

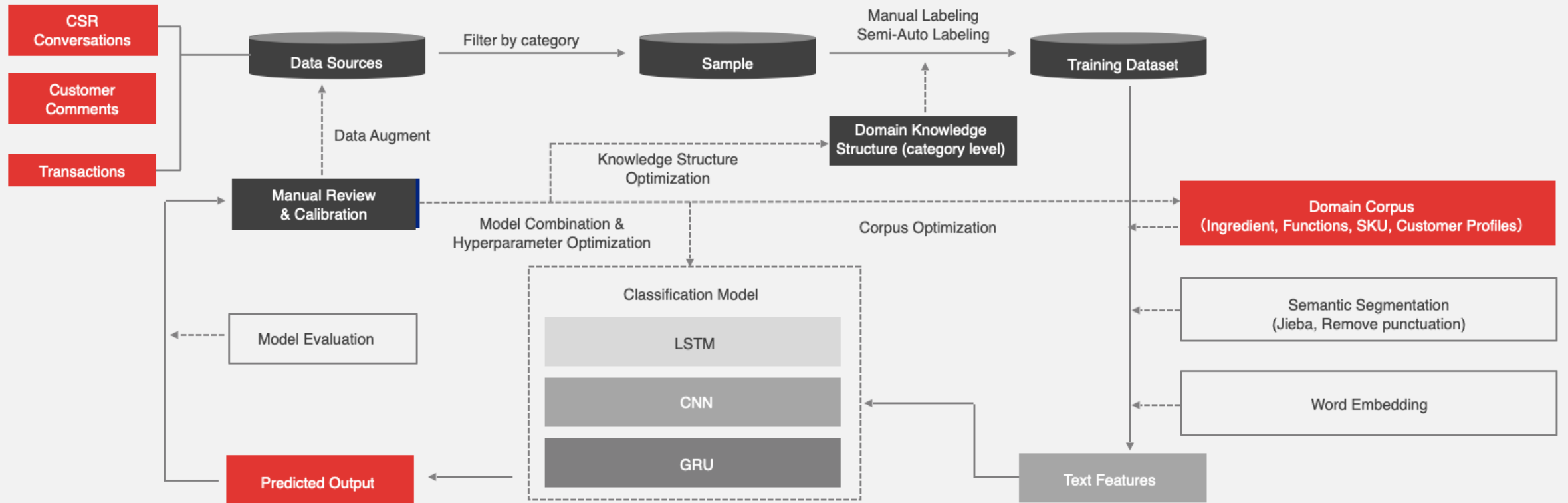
Online Deploy



- Word Embedding Vector
- Model Structure
- Model Weights

Layer (type)	Output Shape	Param
embedding (Embedding)	(None, 87, 300)	87
bidirectional (Bidirectional)	(None, 87, 512)	1140736
lstm ₁ (<i>LSTM</i>)	(None, 87, 256)	787456
lstm ₂ (<i>LSTM</i>)	(None, 128)	197120
dropout (Dropout)	(None, 128)	0
dense (Dense)	(None, 64)	8256
dense ₁ (<i>Dense</i>)	(None, 1)	87
Total params: 17,133,633		
Trainable params: 2,133,633		

End to End Solutions



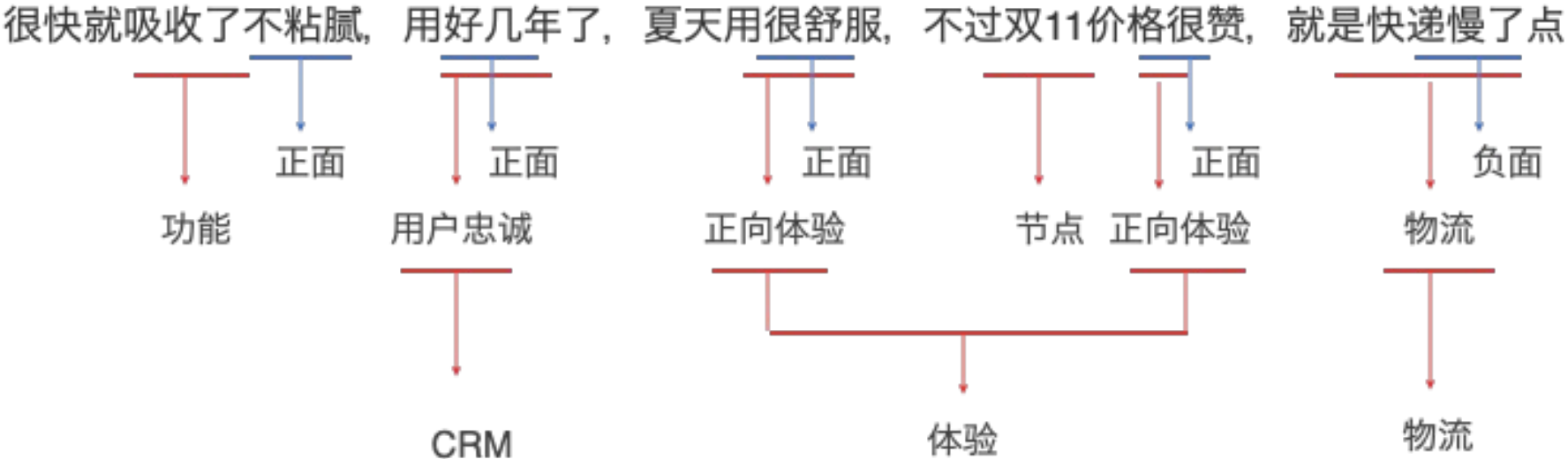
Product Profile and Sentiment analysis

Comment

Sentiment Analysis

Topic Label

Semantic classification



Results

Modification

Data & Model

- Knowledge Graph
- Node Embedding
- Relational Reasoning - Pairwise relations

between entities

Applications

- Fully Optimize the material on the consumer content contact
- Content optimization model and analysis dimension
- Page structure analysis
- Good/ Bad comments rate
- Promotion points
- Repetitive
- Competing goods content analysis
- Selling point to refine
- Information density
- User Comments from other source
- Channel analysis
- Functional analysis



Customer service

- Recommendation
 - Based on historical data
- Conversion rate
 - Old Customers - Repurchase
 - New Customers - Purchase

at analysis