Hi everyone, I’m Yuming from SME iTeam. This innovation idea is E-shop Comment Analysis System. This is about how we serve for E-commerce companies.

Shopping online is more and more popular today. People buy products and leave comments on big E-commerce platforms such as Amazon. Most E-commerce companies have E-shops on these platforms. Their products received many customer comments there and it was inefficient for them to get feedback about their products by manually reading the comments.

So what they need is an assistant that could provide multi-dimensional comment analysis and extract customers’ opinions.

Our solution is to build an analysis server and integrate it with ERP system like B1. Our customers could finally see the comment analysis result in a dashboard in their ERP systems.

Let’s see a demo about this. This is the product information page in B1 and there is a comment analysis dashboard here. If our customers want to know how buyers evaluate their product, they just need a click now. This product is a down coat selling on a Chinese E-commerce platform. There are 171 comments collected. And this dashboard displays our comment analysis result on this down coat.

First let’s look through this grid with a donut. This is the year to date average sentiment index. We do sentiment analysis towards all comments and got an average sentiment index. 0 means all negative and 100 means all positive. This down coat receives a score of 77, while the industry average is 79, which implies that this down coat may receive less compliments than other down coats. This donut below tells us the distribution of buyers’ comments. How many positive, how many neural, and how many negative. And this grid with a line chart shows the trend of the average sentiment index: current month compared to last month, and the change over last several months.

Next, let’s look at the first grid. This shows what we call multi-dimensional analysis or opinion mining. These Chinese characters are topics: quality, appearance, service. We integrate buyers’ comments into several topics and do sentiment analysis under each topic. So that we can know buyers’ general opinion on each topic. For example, this grid tells us there are 52 comments talking about quality and the sentiment index of quality is 82, 4% higher than industry average, which means most buyers think the down coat’s quality is good. While the topic service receive a score much lower than industry average, which means that this E-commerce company probably should improve their service.

And if you want to know more specific buyer opinions, you can see through the keywords grid, the stack of latest 10 positive comments, the stack of latest 10 negative comments, and the stacks of latest comments on topics.

I believe this dashboard could give the E-commerce companies important product feedback and help them adjust their strategy. In this way, our idea could perfect our solutions to E-commerce companies.

The following flowchart will show how we analysis the comments. The analysis core mainly consists of four units, keyword extractor, similarity calculation and clustering, sentiment analysis and data preparation and training. These units together enable us to provide multi-dimensional comment analysis and extract customers’ opinions.

The following flowchart will show how we analysis the comments. First our analysis server get the product’s item ID on E-commerce platforms from ERP systems (B1). Then our comments extractor will extract the product’s comments from the E-commerce platforms. Next, we use technology of TF-IDF, TextRank and POStag to mine keywords in the comments. These keywords will then be put into a similarity calculation unit. In this unit, we use a trained word2vec model to embed these keywords in a vector space. So that we can cluster the keywords with similar meanings as opinions by spectral clustering. Next, we do sentiment analysis toward each opinion. We use technology of word segmentation and Naive Bayes Classifier to quantify customers’ sentiment according to their expressions on each opinion. To ensure the accuracy of this Naive Bayes Classifier as well as the word2vec model mentioned above, we collected corpus from Wiki and E-commerce platforms and labelled some of them. The labelled corpus are used to train the Naive Bayes Classifier and the labelled corpus are used to train Word2vec model. Finally, all the analysis result are returned to the ERP systems and displayed in one dashboard.