ABSTRACT
In this project we explored LSTM neural network for the fine-grained sentiment analysis of restaurant customer reviews in Chinese language. For this aspect-level multi-class classification task, we trained one model separately for each of the elements under each model architecture.

DATA SET
We used the datasets provided by AI challenger official. The training and validation datasets are manually labelled. The four classes are: positive (1), neutral (0), negative (-1), not mentioned (-2). Training (105,000), validation (7500), and test (7498) sets have same class distributions.
Here we just list the 6 categories for the all 20 elements:
Within “Others”, the 2 elements are “overall experience” and “willing to consume again”.

LSTM FOR CLASSIFICATION
We constructed a two-layer LSTM (Long Short-Term Memory) neural network. This model takes the 105,000 outputs of size (1, 400) from feature extraction step as the input. The embedding matrix takes weights from the feature extraction step, and is not trainable. Apart from hyperparameters shown in the graph, we also modified class weights according to class distributions in each element. We used categorical cross entropy as the loss function:

\[ L(\theta) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{K} y_{ij} \log(p_{ij}) \]

For 7 elements with relatively low performance from the structure above, we update the following hyperparameters:
- Number of hidden units in LSTM layer: 256
- Dropout: 0.2
- Batch size: 32 or 64
The training process is slow, and is still on going...but the validation loss and accuracies are clearly improved.

RESULTS AND DISCUSSIONS
Here is an example of prediction:
“一直经过这条路 第一次进去挨饿…之前说说环境还是很不错的 感觉很适合小情侣来 比较安静的环境 精致下午茶感觉特别好 晚上也不怎么吵很舒服 快 但是周末 中午人太多 人很多很安静 非常喜欢这里的气氛 再说说美食点了一个新推出的冰激凌披萨 薄薄的披萨真的很美味呢 其他尝试只适合 餐点的披萨 真的是特别 去的那天起 多数是点的 比较少 但是有点偏咸了…玫瑰巧克力 和香草味都很好吃的挺好 可是汤汤水水 目前大众点评开 虽然能享受9折 但真的是划算 以后还会经常光顾的”

Keywords: good environment, good service, quiet, tasty, large portion, next time must try again, discount, will come often

FUTURE WORKS
1. Data augmentation;
2. Improve the quality of language representation models; Implement contextual representation, e.g. BERT;
3. Apply Attention mechanism for extracting key information.

REFERENCES
https://github.com/AlChallenger/Al.Challenger_2018
Yan Song, Shuming Shi, Jing Li, and Haitong Zhang. Directional Skip-Gram: Explicitly Distinguishing Left and Right Context for Word Embeddings. NAACL 2018 (Short Paper)