

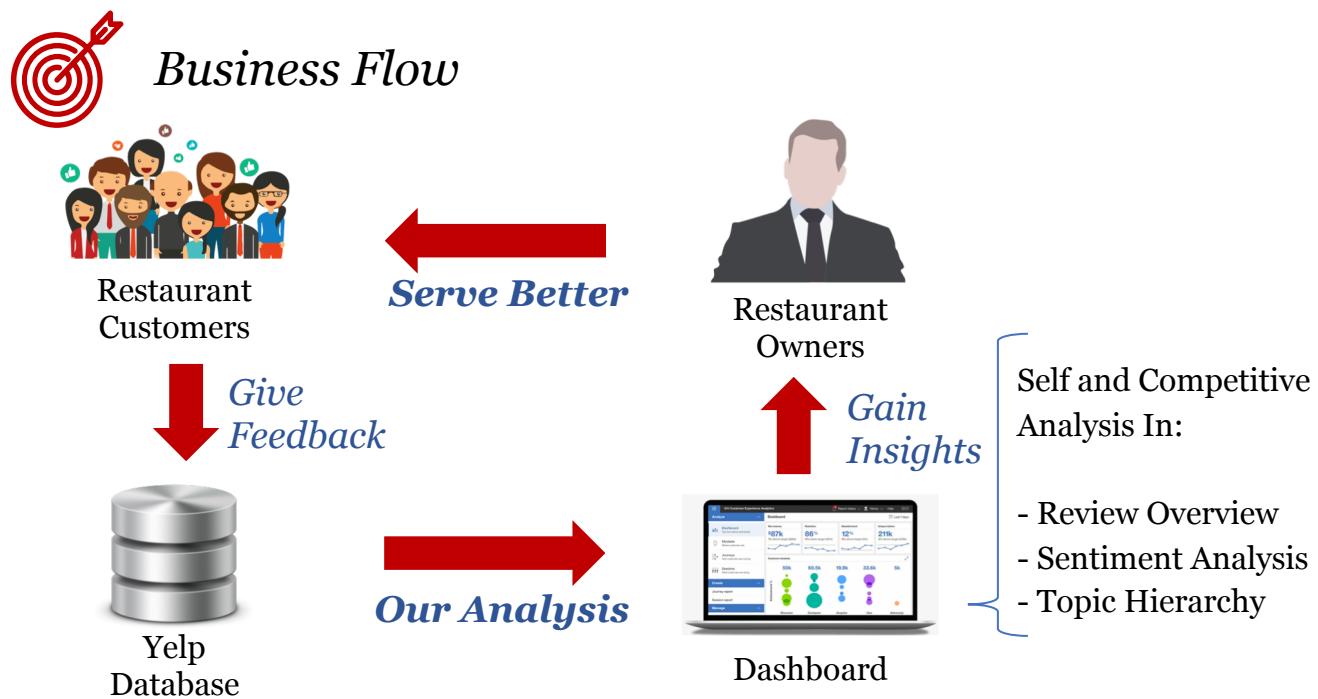
# Do You Really Know What Your Customers Are Talking About?

Elizabeth Zhu, Lin Xu, Yusha Wang, Boyang Wei, Xiangke Chen, Yi-Li Yu



## Pain Points of Restaurant Owners

- Restaurant reviews are highly volatile and usually in large quantity
- Received ratings do not represent actual comments
- Received textual comments can not be systematically analyzed
- Received textual comments are not being used for benchmark analysis



For more information, please contact us:

zhuoo392@umn.edu, weioo157@umn.edu  
xuooo532@umn.edu, yuooo297@umn.edu  
wang6064@umn.edu, chen6789@umn.edu

Dashboard Demo



Find us on GitHub





# Big Data Platforms & Tools



## NLTK, LDA, Textblob, TF-IDF



Amazon cloud computing platforms and APIs which allows users to have at their disposal a virtual clusters of computers

### Storage



### PySpark



Databricks develops a web-based platform for Spark and provides cluster management and IPython-style notebooks

Tableau software query data in varied formats to generate graph-type data visualizations.

### Visualization



## Technical Architecture

### Data input

Textual Reviews

### Pre-processing

Tokenization

Stemming

Removing Stop Words

Lemmatization

### Transformation

Bag of Words

TF-IDF

### Analysis

Sentiment Analysis

Topic Modeling

### Visualization

Dashboard

Our Analysis Can Compatible with Many Big Data Platforms

ORACLE

CLOUDERA

mongoDB

Microsoft Azure

SISENSE

collibra

