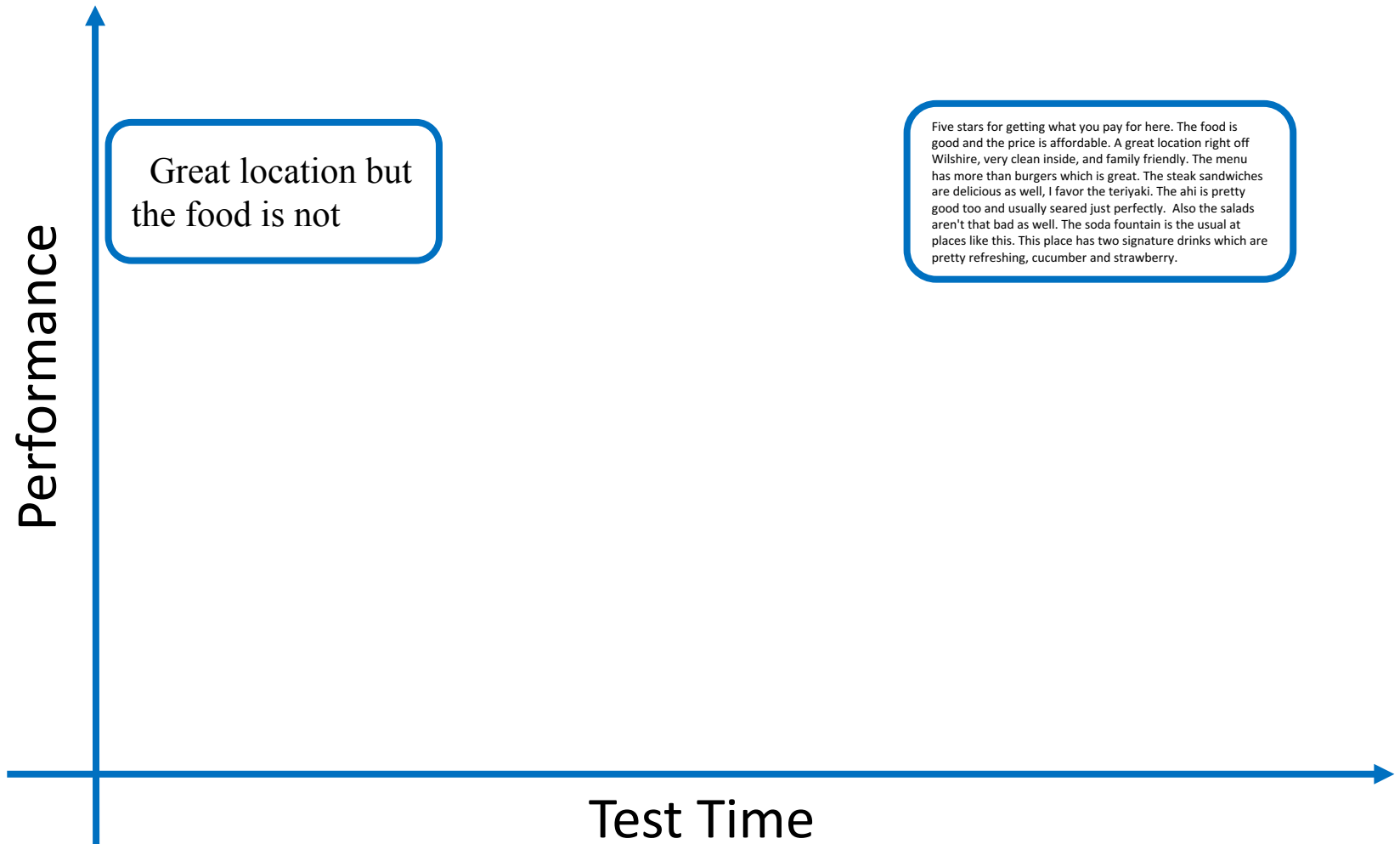# Robust Text Classifier on Test-Time Budgets

Md Rizwan Parvez,
Kai-Wei Chang,
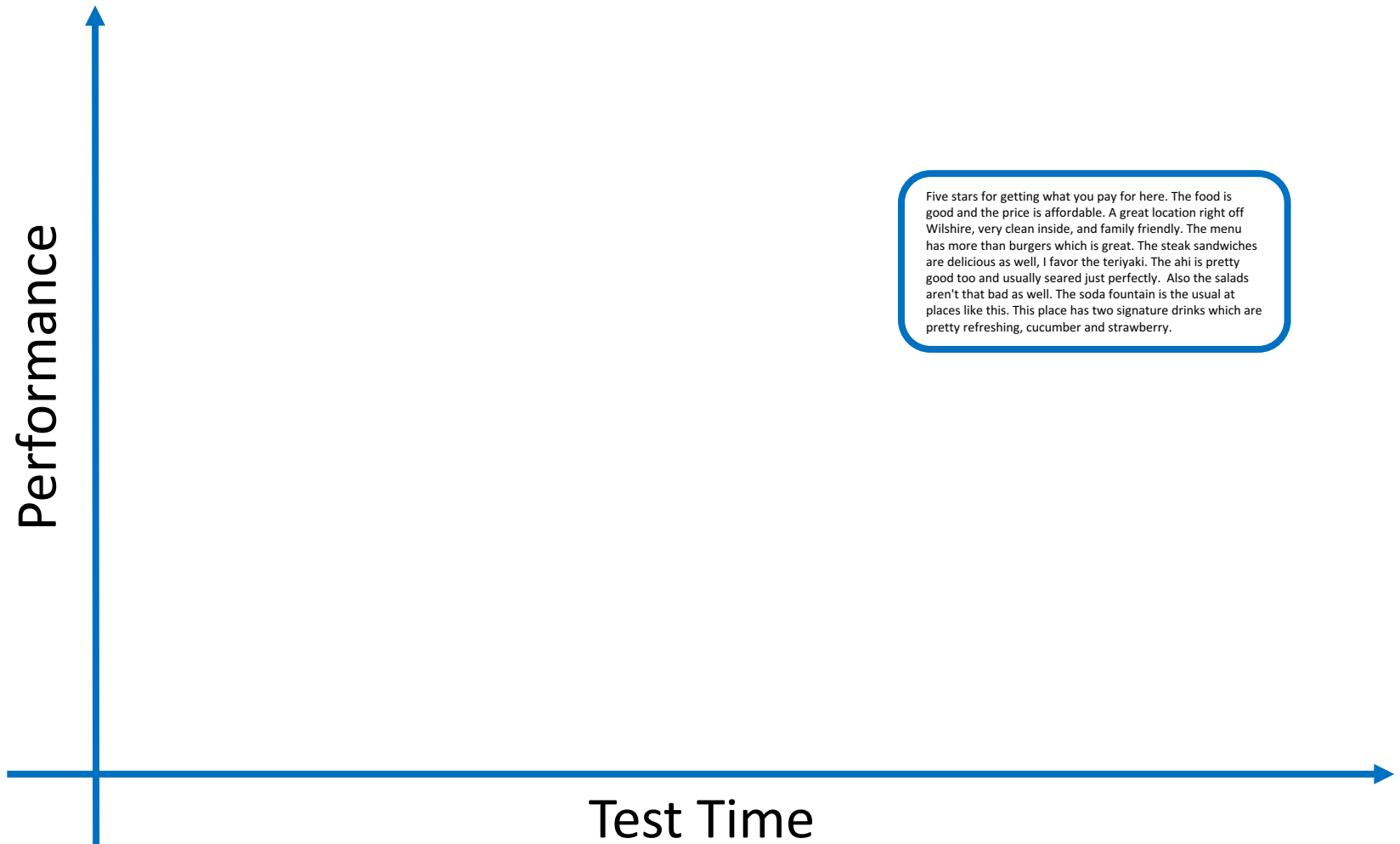
Tolga Bolukbasi,
Venkatesh Saligrama

Motivation

Performance (y-axis) vs Test Time (x-axis)

Great location but the food is not

Five stars for getting what you pay for here. The food is good and the price is affordable. A great location right off Wilshire, very clean inside, and family friendly. The menu has more than burgers which is great. The steak sandwiches are delicious as well, I favor the teriyaki. The ahi is pretty good too and usually seared just perfectly. Also the salads aren't that bad as well. The soda fountain is the usual at places like this. This place has two signature drinks which are pretty refreshing, cucumber and strawberry.

Motivation

Performance (y-axis)

Test Time (x-axis)

Five stars for getting what you pay for here. The food is good and the price is affordable. A great location right off Wilshire, very clean inside, and family friendly. The menu has more than burgers which is great. The steak sandwiches are delicious as well, I favor the teriyaki. The ahi is pretty good too and usually seared just perfectly. Also the salads aren't that bad as well. The soda fountain is the usual at places like this. This place has two signature drinks which are pretty refreshing, cucumber and strawberry.
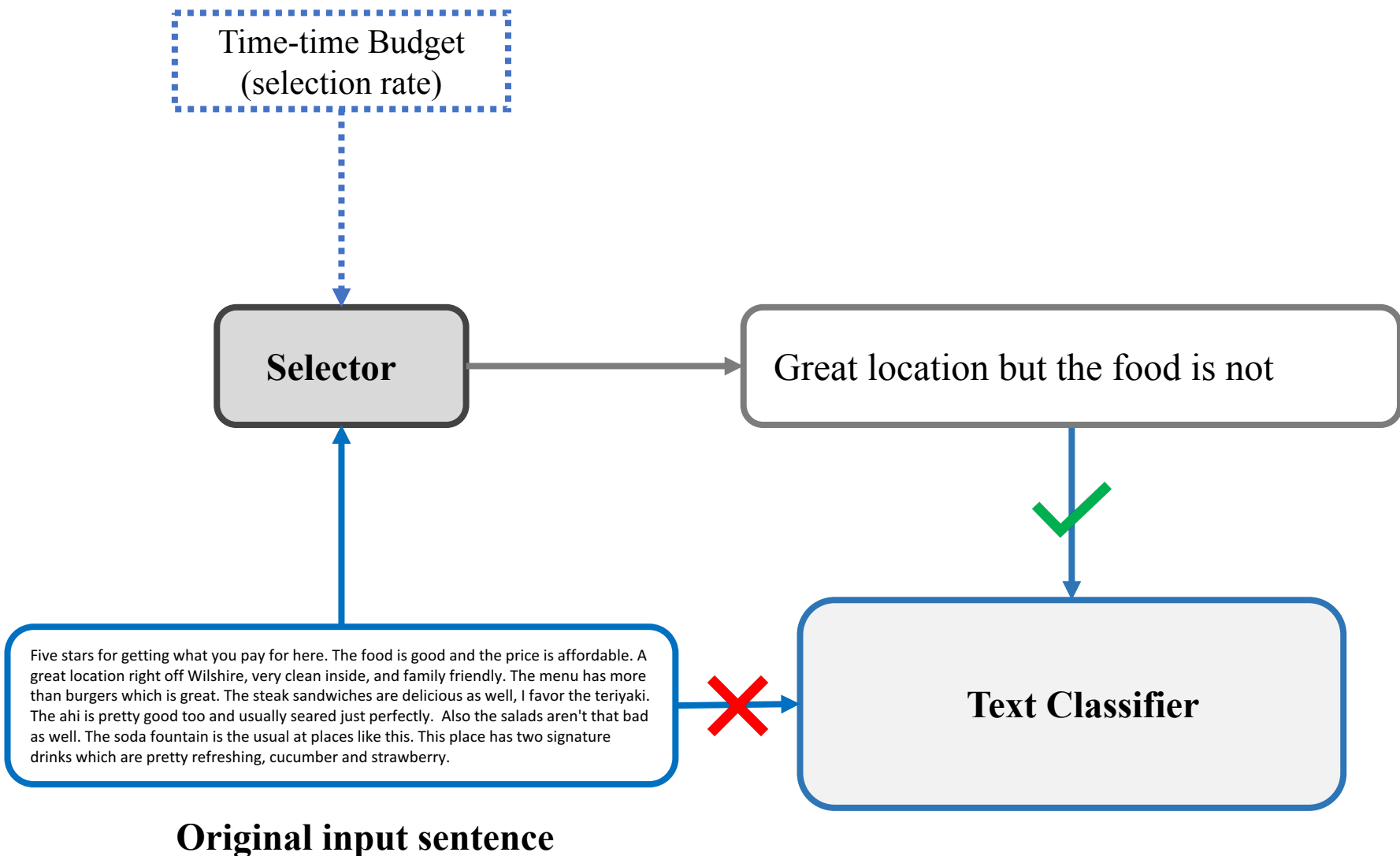
intuition

A single text classifier for various test-time budgets

Performance (y-axis), Test Time (x-axis)
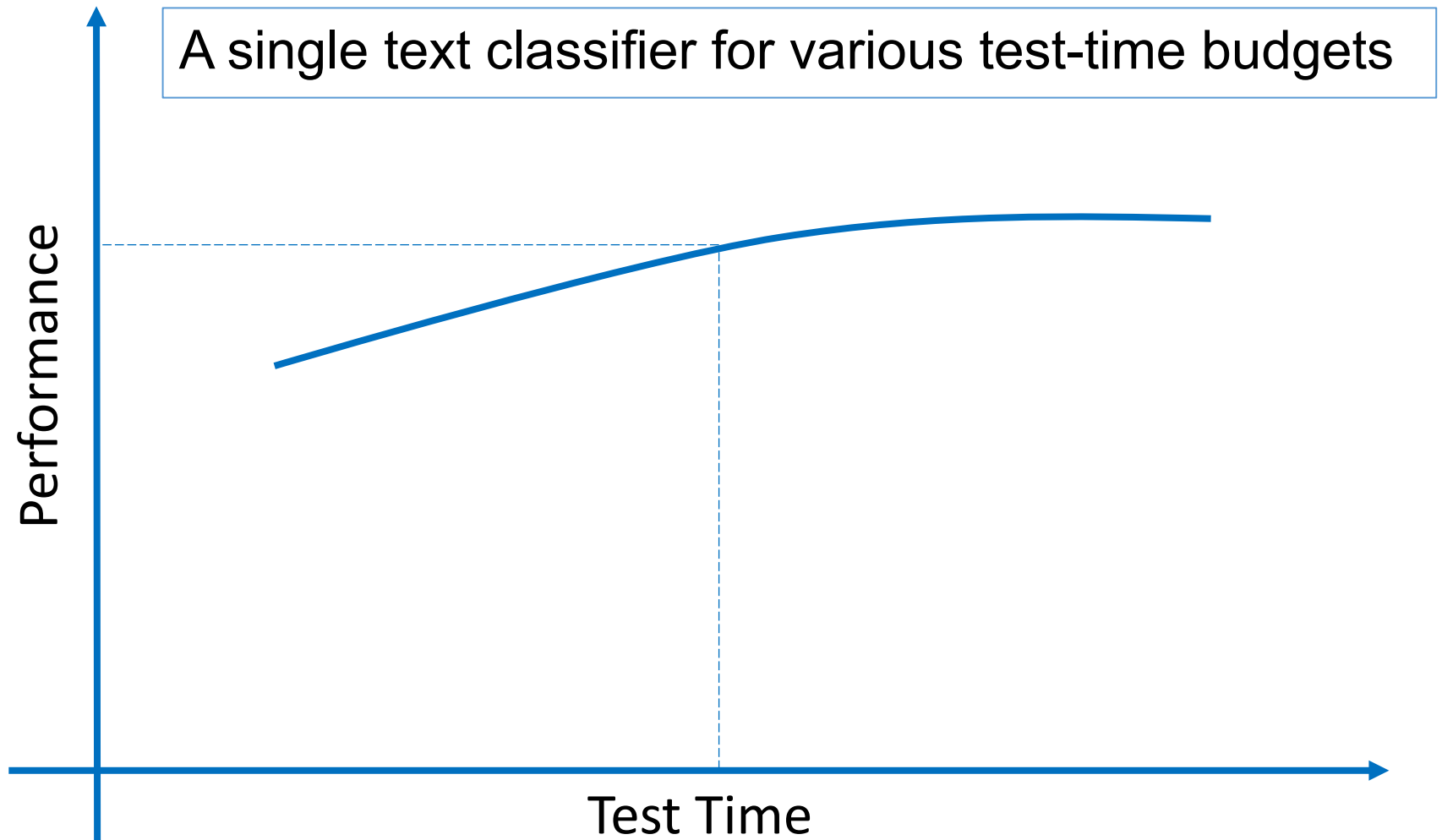
Intuition

Business	Telecom Austria taps the Bulgarian **market**. **Telecom** Austria, Austria's **largest** telecoms operator, obtained ..
Sci/Tech	.. Reuters - **Software security** companies **and** handset makers, including Finland's **Nokia** (NOK1V.HE), **are** ..

Document Classification

A Meta Selector-Classifier Framework

Time-time Budget
(selection rate)

Selector → Great location but the food is not

Five stars for getting what you pay for here. The food is good and the price is affordable. A great location right off Wilshire, very clean inside, and family friendly. The menu has more than burgers which is great. The steak sandwiches are delicious as well, I favor the teriyaki. The ahi is pretty good too and usually seared just perfectly. Also the salads aren't that bad as well. The soda fountain is the usual at places like this. This place has two signature drinks which are pretty refreshing, cucumber and strawberry.

Text Classifier

Original input sentence

A Meta Selector-Classifier Framework

Budget
(selection rate)

1. How to build an effective/efficient selector

Selector

Great location but the food is not

Five stars for getting what you pay for here. The food is good and the price is affordable. A great location right off Wilshire, very clean inside, and family friendly. The menu has more than burgers which is great. The steak sandwiches are delicious as well, I favor the teriyaki. The ahi is pretty good too and usually seared just perfectly. Also the salads aren't that bad as well. The soda fountain is the usual at places like this. This place has two signature drinks which are pretty refreshing, cucumber and strawberry.
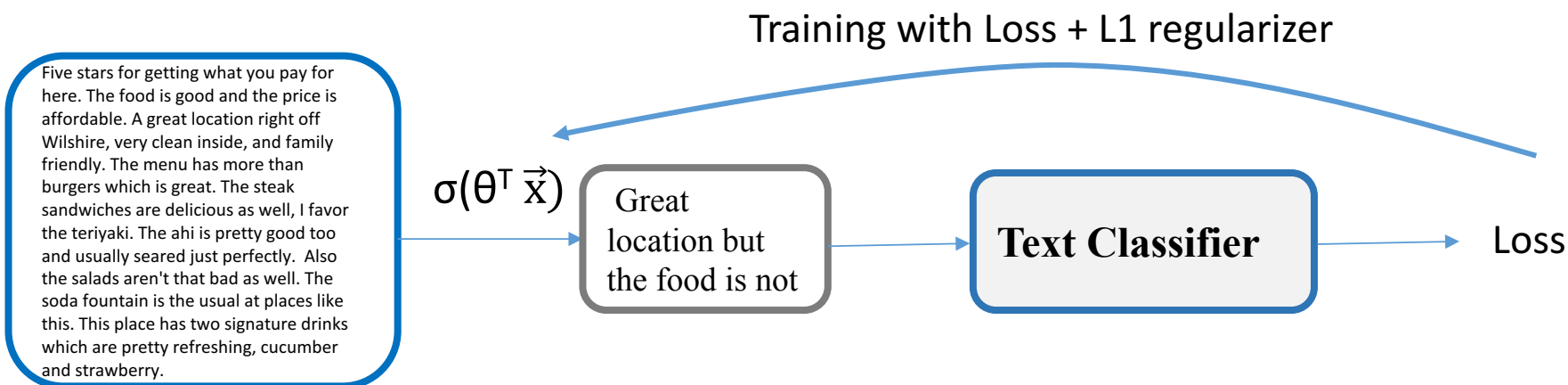
Text Classifier

Original input sentence

Kai-Wei Chang (http://kwchang.net)

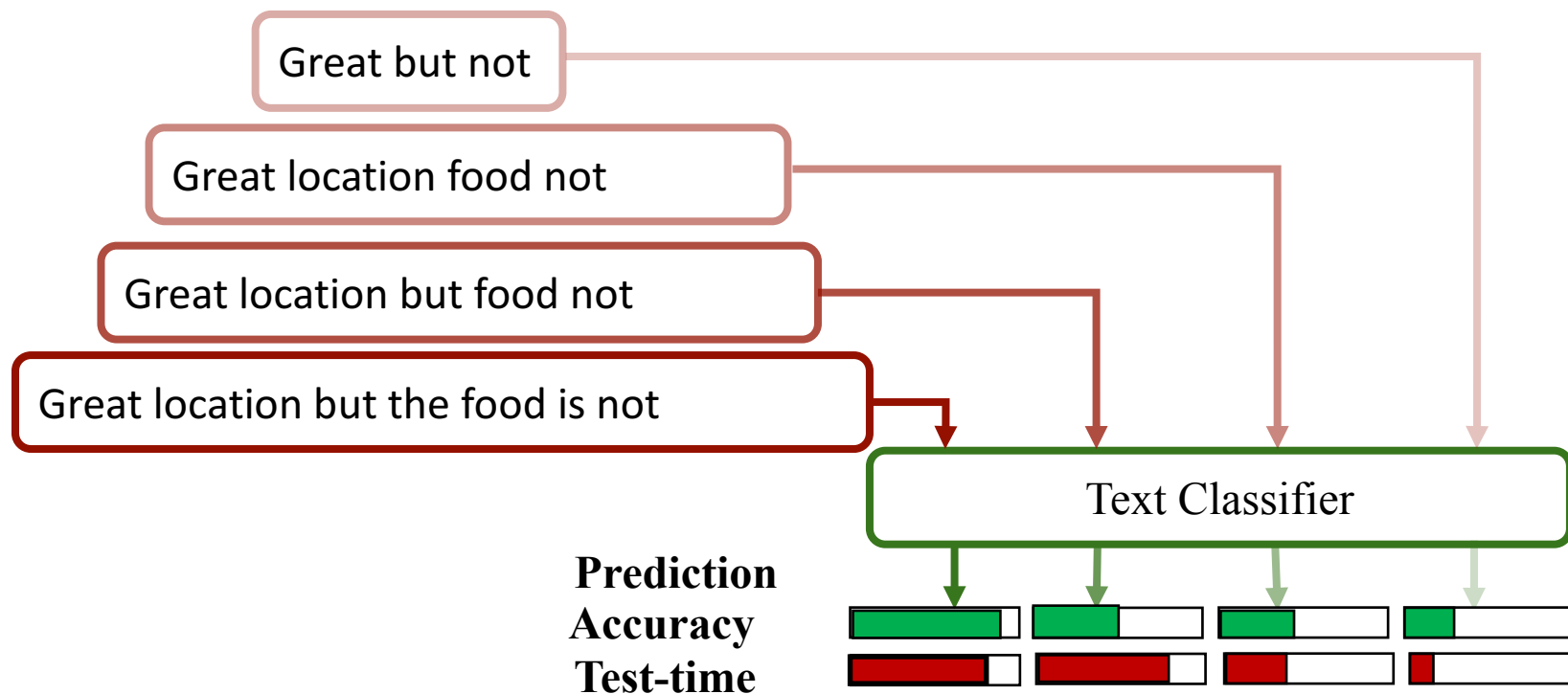**2. How to train a robust classifier?**

End-to-End Training for WE Selector

- Selector has to be light-weight

- We consider a selector based on word-embedding
  - P(word x is selected) = $\sigma(\theta^T \vec{x})$
  - **End-to-end** training [Lei et al. (2016)]
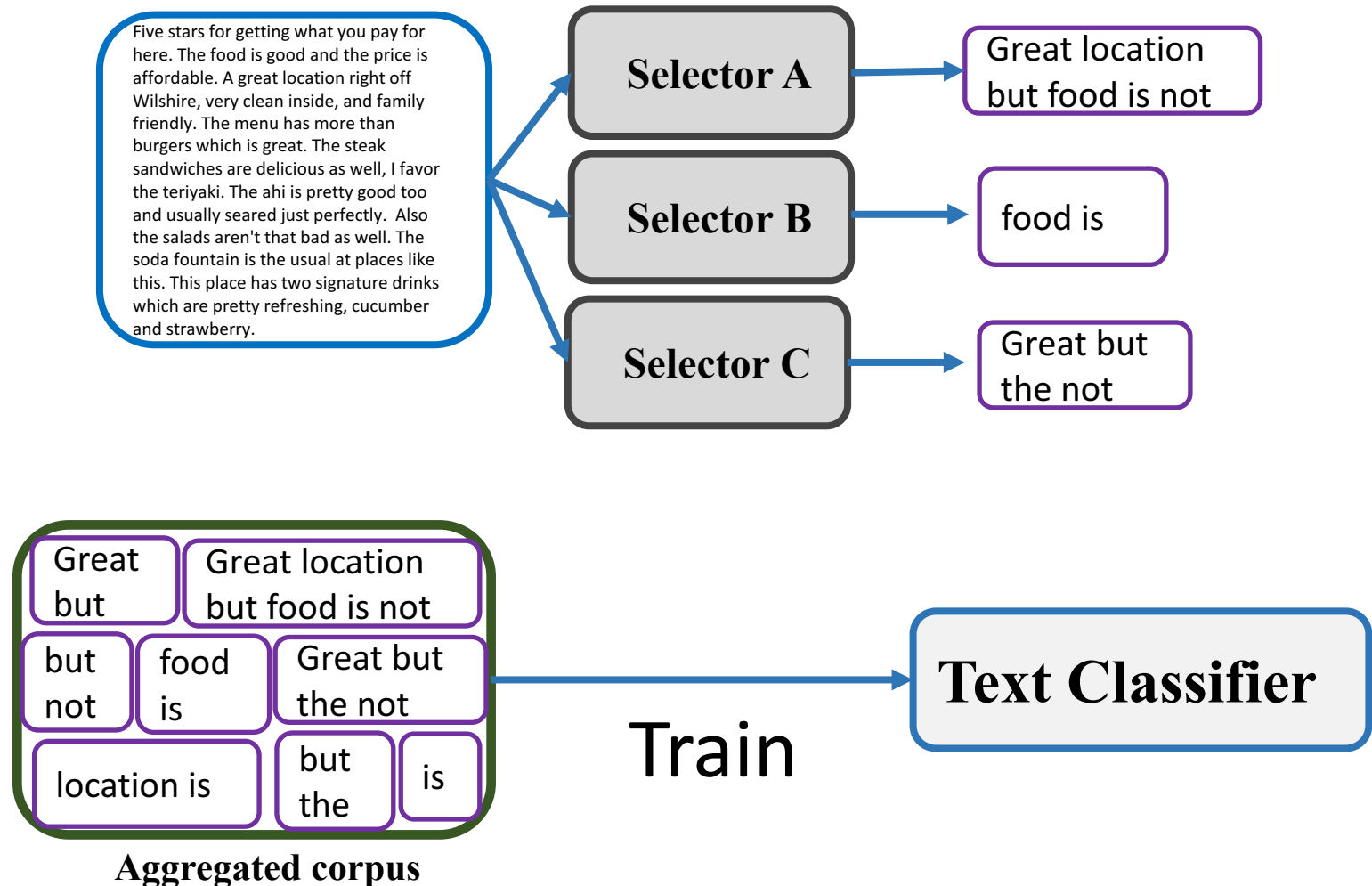 - L1-regularizer to encourage sparsity

Training with Loss + L1 regularizer

Five stars for getting what you pay for here. The food is good and the price is affordable. A great location right off Wilshire, very clean inside, and family friendly. The menu has more than burgers which is great. The steak sandwiches are delicious as well, I favor the teriyaki. The ahi is pretty good too and usually seared just perfectly. Also the salads aren't that bad as well. The soda fountain is the usual at places like this. This place has two signature drinks which are pretty refreshing, cucumber and strawberry.

$\sigma(\theta^T \vec{x})$

Great location but the food is not → **Text Classifier** → Loss

[Lei et al., (2016)] **References**: Tao Lei, Regina Barzilay, and Tommi S. Jaakkola. 2016. Rationalizing neural predictions. In EMNLP

Key Challenges - Fractured Texts

❖ Text classifier is not trained to deal with fractured texts

Great but not

Great location food not

Great location but food not

Great location but the food is not

Text Classifier

Prediction
Accuracy
Test-time

Training a Robust Classifier with Data Aggregation

Five stars for getting what you pay for here. The food is good and the price is affordable. A great location right off Wilshire, very clean inside, and family friendly. The menu has more than burgers which is great. The steak sandwiches are delicious as well, I favor the teriyaki. The ahi is pretty good too and usually seared just perfectly. Also the salads aren't that bad as well. The soda fountain is the usual at places like this. This place has two signature drinks which are pretty refreshing, cucumber and strawberry.

Selector A → Great location but food is not

Selector B → food is

Selector C → Great but the not

Great but | Great location but food is not

but not | food is | Great but the not

location is | but the | is

Aggregated corpus

Train → **Text Classifier**

An Alternative Selector

- L1-regularized Bag-of-Words Model
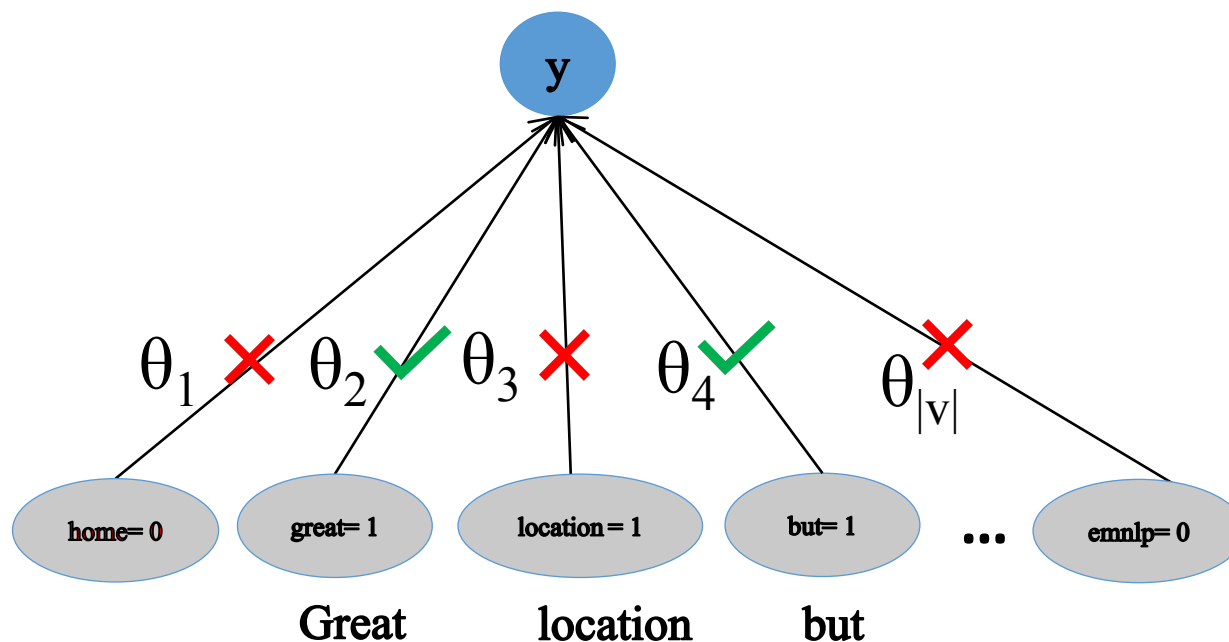
$$\theta^* = \text{argmin}_\theta \ \sum_{(x,y)} \log \left(1 + \exp(-y \ \theta^T \cdot \vec{x})\right) + \frac{1}{b} \ ||\theta||_1$$

b: selection rate
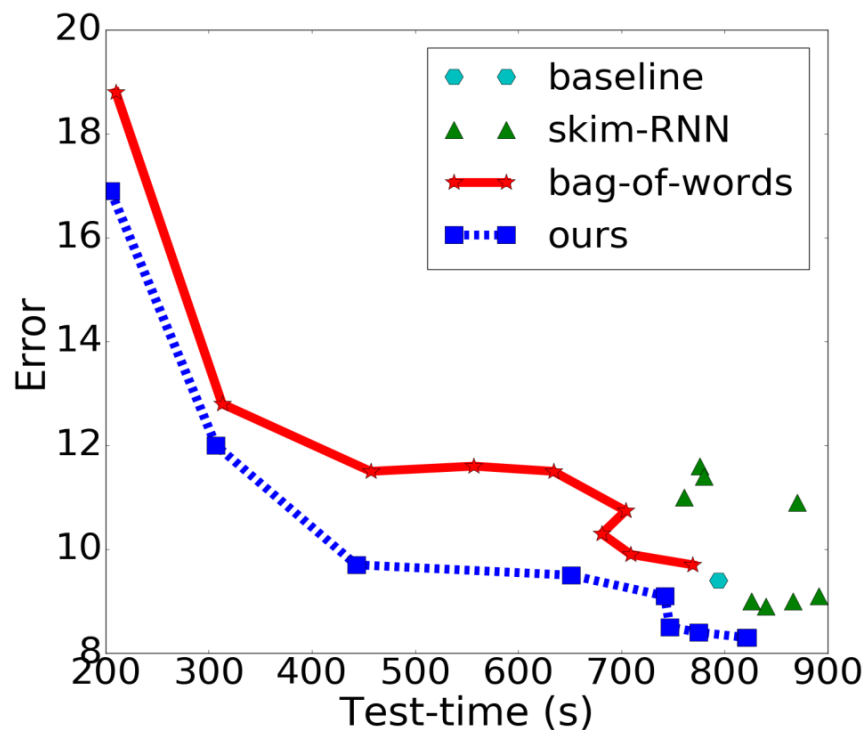
Related Work

❖ Skip/Skim unimportant Text

   ❖ Skip-RNN [Campos et al., 2017]

   ❖ Skim-RNN [Seo et al., 2018]

   ❖ Our approach is agonistic to text classifier
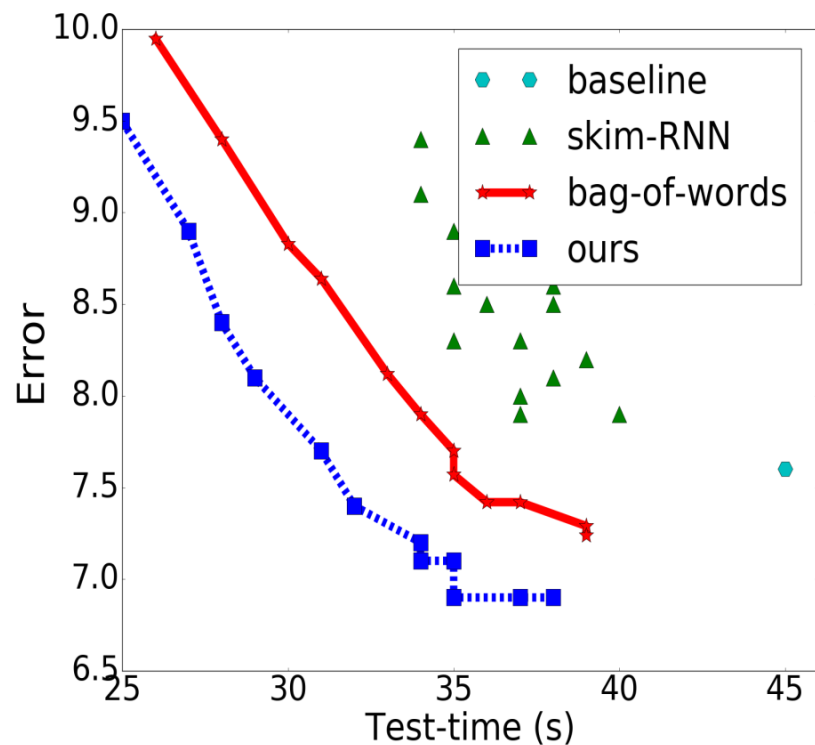
❖ Model Compression [Bucila et al., 2006]

Experiment Setting

- Datasets: SST-2, IMDB, AGNews, and Yelp

- Classifiers: LSTM, BCN [McCann et al., 2017]

- Aggregate corpus from both selectors:
  - Selection rate = {50%, 60%, 70%, 80%, 90%, 100%}

Error Rate v.s. Test-Time

IMDB

AGNews

See our paper for more results

Conclusion

- We present a selector-classifier meta framework for building efficient text classifier

- We propose efficient selectors and data aggregation scheme to train classifiers that are compatible with fractured text selected at any rate.

Source codes are available at:

https://github.com/uclanlp/Fast-and-Robust-Text-Classification

Thanks!!!!