LADy (🕺):
A Benchmark Toolkit for Latent Aspect Detection Enriched with Backtranslation Augmentation

Farinam Hemmatizadeh[1], Christine Wong[1], Alice Yu[2] and Hossein Fani[1]
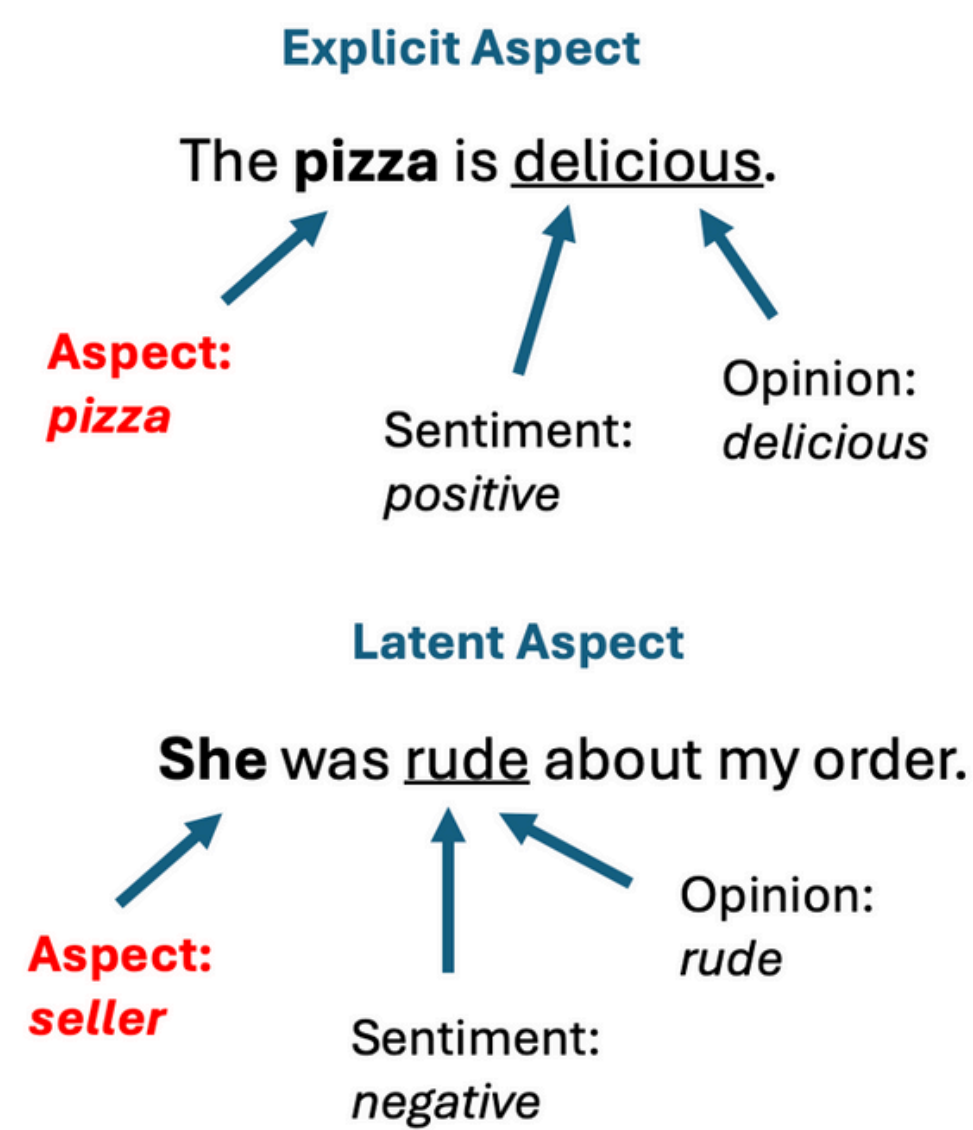[1] University of Windsor, Canada
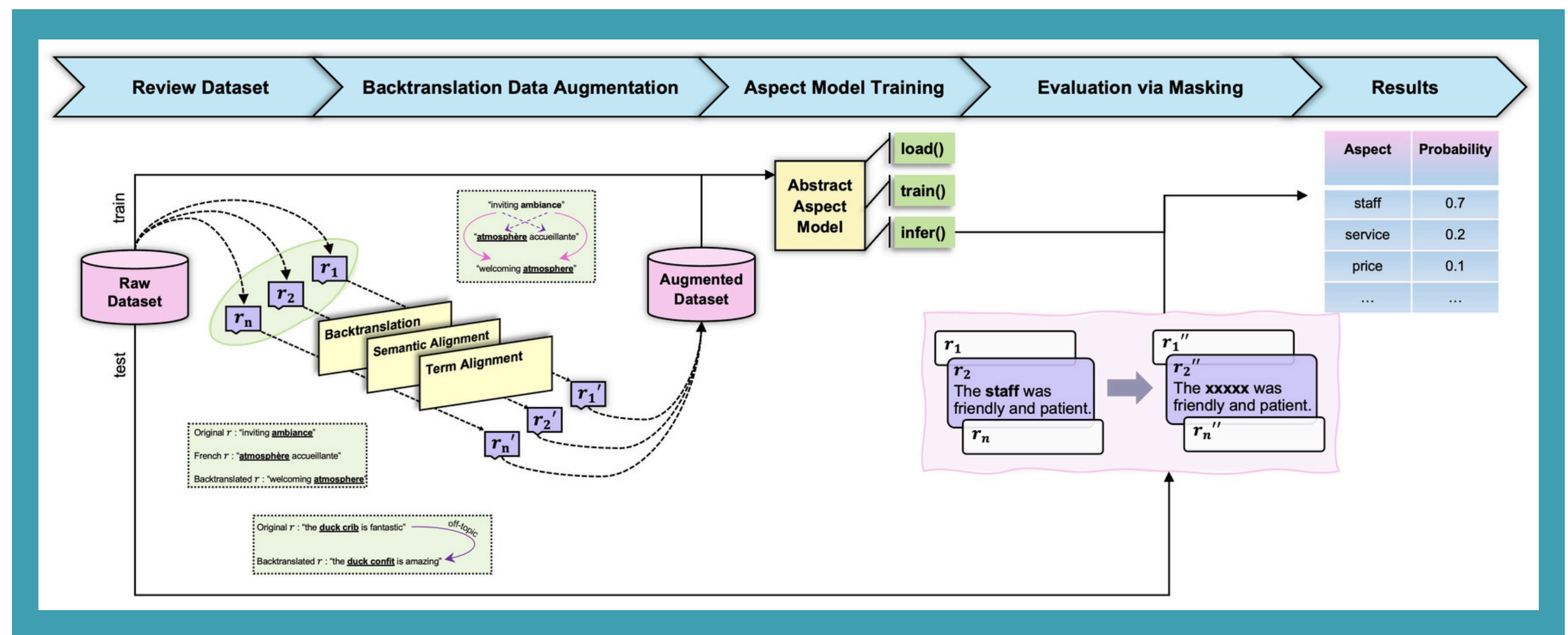[2] Vincent Massey Secondary School, Canada

University of Windsor

Fani's Lab!

SIGIR 2024 Washington, D.C.

SIGIR Student Travel Grant

Background

Explicit Aspect

The **pizza** is <u>delicious</u>.

Aspect: *pizza*
Sentiment: *positive*
Opinion: *delicious*

Latent Aspect

She was <u>rude</u> about my order.

Aspect: *seller*
Sentiment: *negative*
Opinion: *rude*

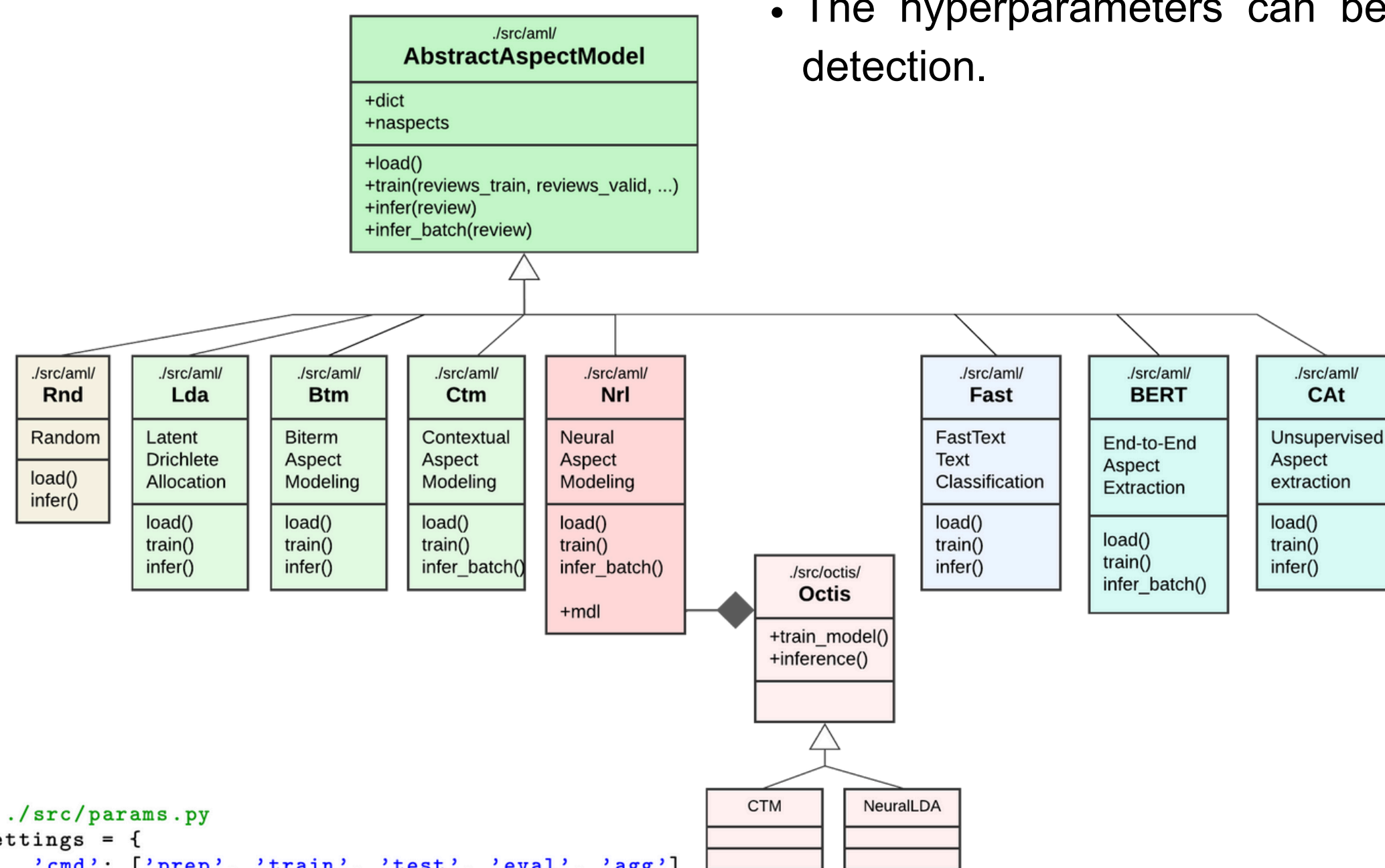## Workflow
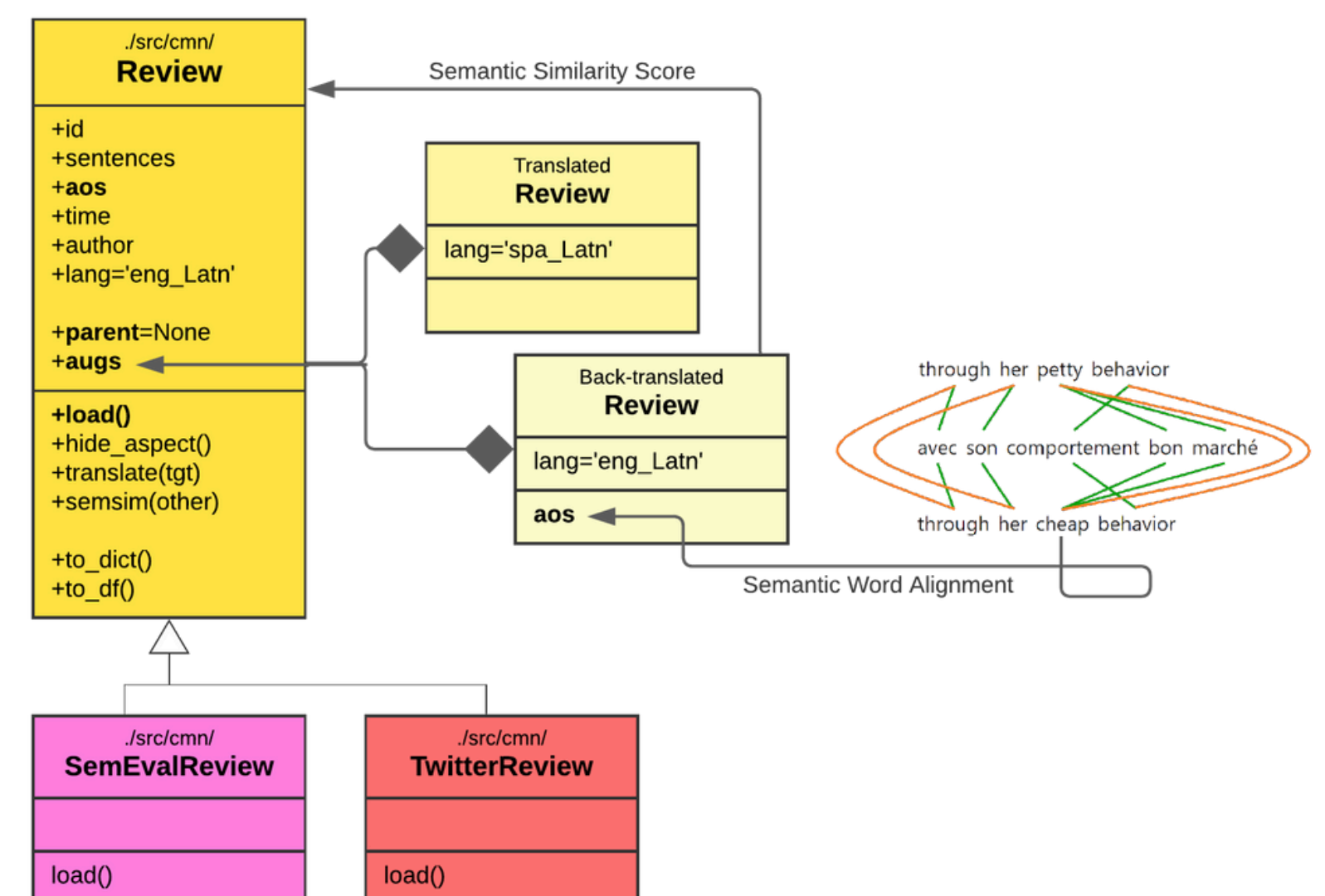

- Backtranslation is used to enhance the extraction of latent aspects, while semantic alignment and term alignment are used to ensure the augmented reviews remain semantically similar and contain accurate aspect labels.
- The hyperparameters can be adjusted for explicit aspect detection instead of latent aspect detection.

Class Structures


```
# ./src/params.py
settings = {
    'cmd': ['prep', 'train', 'test', 'eval', 'agg']
    'prep': {'langaug':'pes_Arab', 'spa_Latn', ...],
            'translator': 'nllb'},
    'train': {'ratio': 0.85, 'nfolds': 5, 'nwords': 20,
            'langaug_semsim': 0.5,
            'qualities': ['coherence', 'perplexity']},
    'test': {'h_ratio': 0.5},
    'eval': {'metrics': ['precision', 'recall', ...],
            'topkstr': [1, 5, 10]}
```
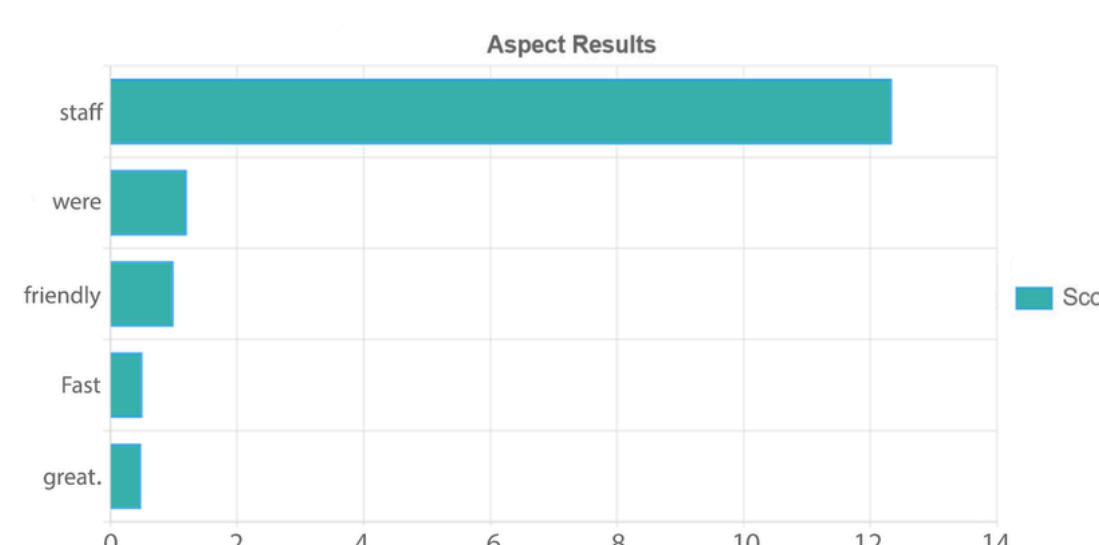

Results and Web Application

🕺 LADy: Latent Aspect Detection from Online Unsolicited Reviews

	bert-tfm [15]			cat [24]			loclda [2]			btm [13, 30]			ctm [1]			random		
	pr1	rec5	ndcg5	pr1	rec5	ndcg5	pr1	rec5	ndcg5	pr1	rec5	ndcg5	pr1	rec5	ndcg5	pr1	rec5	ndcg5
none	**0.7000**	0.6897	0.6757	0.3327	0.5248	0.4343	0.2320	0.3549	0.2925	0.1872	**0.3133**	0.2500	0.0560	0.0493	0.0485	0.0000	0.0005	0.0005
+Chinese	0.6661	0.6928	0.6699	0.3723	0.5287	0.4596	0.1968	0.3408	0.2647	0.1760	0.2783	0.2261	0.0624	0.0717	0.0637	**0.0016**	0.0028	0.0022
+Farsi	0.6742	0.6707	0.6608	0.3703	0.5386	0.4592	0.1840	0.3494	0.2689	0.1776	0.2834	0.2303	0.0560	0.0823	0.0722	0.0000	0.0002	0.0002
+Arabic	0.6661	0.6898	0.6671	0.4139	0.5683	0.4939	0.2000	0.3654	0.2887	0.1568	0.2956	0.2269	0.0592	0.0649	0.0577	0.0000	0.0000	0.0000
+French	0.6565	0.7030	0.6734	0.4040	0.5584	0.4883	0.2512	0.3577	0.3032	0.1968	0.3048	0.2481	0.0720	0.0837	0.0733	0.0000	0.0008	0.0006
+German	0.6710	0.6927	0.6721	0.3980	0.5505	0.4787	0.2416	0.3648	0.2976	0.1808	0.2975	0.2242	0.0560	0.0717	0.0603	0.0000	**0.0061**	**0.0036**
+Spanish	0.6645	0.7099	0.6769	0.3921	0.5663	0.4842	0.2224	0.3737	0.3035	0.2000	0.2975	0.2466	0.0464	0.0531	0.0458	0.0000	0.0000	0.0000
+all	0.6613	**0.7182**	**0.6823**	0.5980	**0.7861**	**0.7096**	0.2592	**0.3744**	**0.3104**	0.2128	0.2986	0.2515	**0.2192**	**0.2470**	**0.2263**	0.0016	0.0008	0.0010

Table caption: semeval-15-restaurant

Codebase

Paper