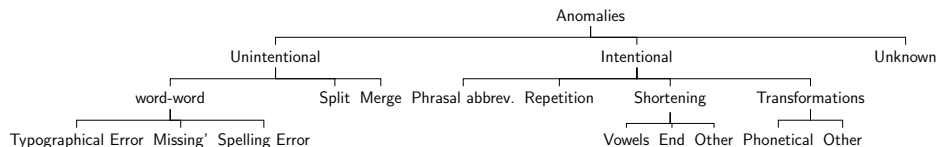


- 1 Automatic Classification of Normalization Replacements
- 2 Universal Dependencies for English Tweets

Automatic Classification of Normalization Replacements

What?



Automatic Classification of Normalization Replacements

What?

| | | | | | | | |
|-------------|---------|--------|-----------|----|-------|---|--------------|
| drinkingggg | spirite | during | chrstms | is | gr8 | . | LOL |
| drinking | spirits | during | christmas | is | great | . | Laughing out |

Automatic Classification of Normalization Replacements

What?

| | | | | | | | |
|-------------|---------|--------|-----------|----|-------|---|--------------|
| drinkingggg | spirite | during | xmas | is | gr8 | . | LOL |
| drinking | spirits | during | christmas | is | great | . | Laughing out |
| 7 | 1 | 0 | 8 | 0 | 12 | 0 | 6 |

Automatic Classification of Normalization Replacements

Why?

- Automatic in-depth evaluation of normalization systems
- Can it be mapped to other languages?
- We can exclude categories during training

Automatic Classification of Normalization Replacements

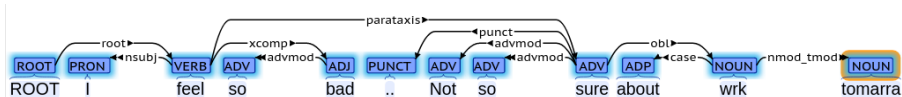
How?

- Take ideas from work from: grammatical error correction/machine translation
- Use training corpus (<https://bitbucket.org/robvander/normcat>)
- More information in paper

Universal Dependencies for English Tweets

- 329 dev and 323 test trees
- (almost only) noisy tweets
- No in-domain training (EWT)
- tokenization, POS (*2), normalization and dependency trees

Universal Dependencies for English Tweets



Universal Dependencies for English Tweets

What?

- Generate VerbForms or lemmas

Universal Dependencies for English Tweets

What?

- Generate VerbForms or lemmas
- Parsing:

Universal Dependencies for English Tweets

What?

- Generate VerbForms or lemmas
- Parsing:
 - Data selection/weighing from EWT (5 domains)

Universal Dependencies for English Tweets

What?

- Generate VerbForms or lemmas
- Parsing:
 - Data selection/weighing from EWT (5 domains)
 - Can we use Italian training data?

Universal Dependencies for English Tweets

What?

- Generate VerbForms or lemmas
- Parsing:
 - Data selection/weighing from EWT (5 domains)
 - Can we use Italian training data?
 - ...

Universal Dependencies for English Tweets

Why?

- Extremely interesting topic

Universal Dependencies for English Tweets

Why?

- Extremely interesting topic
- Improve domain adaptation

Universal Dependencies for English Tweets

Why?

- Extremely interesting topic
- Improve domain adaptation
- ...

Universal Dependencies for English Tweets

How?

- ...