

Which of the following is not represented by a word sequence?

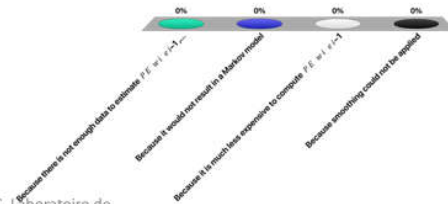
- A. Named Entity
- B. Part-of-Speech tag
- C. Concept
- D. Key Phrase
- E. Word n-gram



©2018, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

$P(W|E)$ is approximated by $P_E(w_i|e_{i-1})$
and not using $P_E(w_i|e_{i-1}, \dots, e_1)$

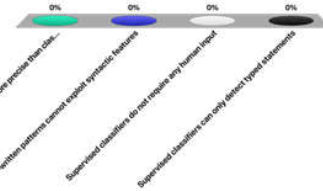
- A. Because there is not enough data to estimate $P_E(w_i|e_{i-1}, \dots, e_1)$
- B. Because it would not result in a Markov model
- C. Because it is much less expensive to compute $P_E(w_i|e_{i-1})$
- D. Because smoothing could not be applied



©2018, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Which is true?

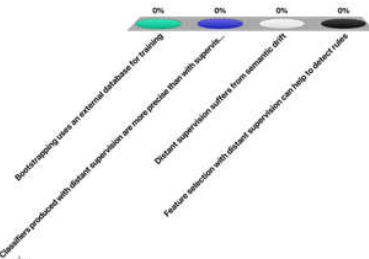
- A. Hand-written patterns are in general more precise than classifiers
- B. Hand-written patterns cannot exploit syntactic features
- C. Supervised classifiers do not require any human input
- D. Supervised classifiers can only detect typed statements



©2018, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis

Which is true?

- A. Bootstrapping uses an external database for training
- B. Classifiers produced with distant supervision are more precise than with supervised learning
- C. Distant supervision suffers from semantic drift
- D. Feature selection with distant supervision can help to detect rules



©2018, Karl Aberer, EPFL-IC, Laboratoire de systèmes d'informations répartis