

## RELAZIONE ESERCITAZIONE 1

*Roger Ferrod, Simone Cullino, Davide Giosa*

### SOMMARIO

La seguente esercitazione richiede di calcolare, tramite *WordNet*, la similarità massima tra coppie di termini e confrontarla, tramite indici di correlazione, con un gold value annotato manualmente. Per il calcolo della similarità sono state utilizzate tre metriche: *Wu & Palmer*, *Shortest Path* e *Leacock & Chodorow*. Mentre gli indici di correlazione presi in considerazione sono *Pearson* e *Spearman*.

### IMPLEMENTAZIONE

Per lo svolgimento di questa esercitazione abbiamo diviso il codice in 4 moduli:

- `WordNetDriver.py` contiene i metodi che si interfacciano a *WordNet*
- `Metrics.py` contiene l'implementazione delle tre metriche: *Wu & Palmer*, *Shortest Path* e *Leacock & Chodorow*
- `Indeces.py` contiene l'implementazione degli indici di correlazione: *Pearson* e *Spearman*
- `Es1.py` punto di avvio dell'esercitazione, chiama iterativamente i vari metodi utilizzando tutte le metriche e gli indici fortini dagli appositi moduli.

### RISULTATI

L'output dell'esercitazione consiste in due file: il primo file contiene, per ogni coppia di termini, le tre concept-similarity calcolate e il golden-value; il secondo file contiene invece i due indici di correlazione calcolati per ogni metrica.

Abbiamo deciso inoltre di fornire un grafico per ogni metrica contenente sull'asse delle ascisse il golden-value e sull'asse delle ordinate la similarità calcolata. In tal modo è possibile verificare visivamente l'allineamento dei dati e confrontarli con i coefficienti di correlazione calcolati. I risultati sono riportati di seguito:

- *Wu & Palmer*:
  - *Pearson*: 0.299
  - *Spearman*: 0.060
- *Shortest Path*:
  - *Pearson*: 0.207
  - *Spearman*: 0.016
- *Leacock & Chodorow*:
  - *Pearson*: 0.331
  - *Spearman*: 0.016