

# INL, projekt 2, podobieństwo znaczeniowe zdań

Zadanie polega na zbudowaniu aplikacji oceniającej w skali 0-5 podobieństwo znaczeniowe pary zdań.

Ręcznie anotowane dane: na stronie <http://zil.ipipan.waw.pl/Scwad/CDSCorpus> dostępna jest lista 1000 par zdań wraz z oceną związku między nimi – kolumna 'round(RE)' – w skali 0-5.

Na przykład:

Młody mężczyzna stoi w rozkroku, trzymając małą piłkę w prawej ręce.	Mężczyzna z ochraniaczami na nogach stoi na trawie w rozkroku.	3
--	--	---

Metoda rozwiązania zadania jest dowolna – reguły, metody statystyczne, sieci neuronowe. Jako dane treningowe czy rozwojowe należy wybrać 90% zdań, a pozostałe 10% (najlepiej wzięte nie po kolei ale wylosowane) użyć jako dane ewaluacyjne.

Termin nadesłania wyników: 26 styczeń 2018.