November 15, 2021

Abstract

1

Algorithm 1: identify Row Context

Input: r_i, $Backgrd(T_i)$=$T_1, T_2, \ldots, T_n$ and similarity threshold θ_r
Output: $con(r_i)$

```
1  con(r_i) = Φ;
2  for j = 1; j ≤ n; j ≠ i do
3      float maxSim = 0;
4      r^maxSim = null;
5      while not end of T_j do
6          compute Jaro(r_i, r_m)(r_m ∈ T_j);
7          if (Jaro(r_i, r_m) ≥ θ_r) ∧ (Jaro(r_i, r_m) ≥ r^maxSim) then
8              replace r^maxSim with r_m;
9      con(r_i) = con(r_i) ∪ r^maxSim;
10 return con(r_i);
```