

Algorithm FuzzyOntoExt(D , PA , Ont)**Input:** corpus D and vector of threshold values PA **Output:** a light weight fuzzy domain ontology Ont **Main Procedure:**

- 1) $Ont = \langle \rangle$
- 2) For each document $d \in D$ Do
 - a) Construct text windows $w \in d$
 - b) Remove stop words sw from w
 - c) Perform POS tagging for each term $t_i \in w$
 - d) Apply Porter stemming to each term t_i
 - e) Filter specific linguistic patterns such as NN, AN, NVN, etc.
 - f) Accumulate the frequency for $t_i \in w$ and the joint frequency for any pair $t_i, t_j \in w$
 - g) IF $lower \leq Freq(t_i) \leq upper$, THEN $A = A \cup t_i$
- 3) For each term $t_i \in A$ Do /* Concept Extraction */
 - a) compute its context vector c_i using BMI, MI, JA, CP, KL, ECH, or NGD
 - b) $C = C \cup c_i$
- 4) For each $c_i \in C$ Do /* Concept Filtering - α -cut */
 - a) IF $\exists t_i \in c_i : \mu_{c_i}(t_i) < \zeta$, THEN $C = C - c_i$
- 5) For each $c_i \in C$, $t_i \in A$ Do /* Update R_{AC} relations */
 - a) IF $\mu_{c_i}(t_i) \geq \zeta$, THEN $R_{AC} = R_{AC} \cup (t_i, c_i)$
- 6) $\forall c_i \in C$: Compute $Rel(c_i, D_j)$
- 7) IF $Rel(c_i, D_j) < \varpi$, THEN $C = C - c_i$ /* Concept Pruning */
- 8) Perform Dimensionality Reduction SVD
- 9) For each pair of concepts $(c_i, c_j) \in C$ Do
 - a) Compute the taxonomy relation (c_i, c_j) using $Spec(c_i, c_j)$
 - b) IF $\mu_{R_{CC}}(c_i, c_j) > \lambda$, THEN $R_{CC} = R_{CC} \cup (c_i, c_j)$
- 10) For each $(c_i, c_j) \in R_{CC}$ Do /* Taxonomy Pruning */
 - a) IF $\mu_{R_{CC}}(c_i, c_j) < \mu_{R_{CC}}(c_j, c_i)$, THEN $R_{CC} = R_{CC} - (c_i, c_j)$
 - b) IF $\exists P(c_i \rightarrow c_x, \dots, c_y \rightarrow c_j)$ AND $\mu_{R_{CC}}(c_i, c_j) \leq \min(\{\mu_{R_{CC}}(c_i, c_x), \mu_{R_{CC}}(c_x, c_y), \dots, \mu_{R_{CC}}(c_y, c_j)\})$, THEN $R_{CC} = R_{CC} - (c_i, c_j)$
- 11) $Ont = \langle A, C, R_{AC}, R_{CC} \rangle$
- 12) RETURN Ont

Fig. 4. The fuzzy domain ontology extraction algorithm.