**Features Used**

In our research we have selected some common features and some novel ones. As a

summary the features that we have used are listed in Table 2.

Table 2 List of features used in the research.

|  |  |  |
| --- | --- | --- |
| ID | Feature Name | Tool used to extract |
| 1 | Word | Programming Language |
| 2 | Next word | Programming Language |
| 3 | Previous word | Programming Language |
| 4 | POS tag of Word | Stanford NLP |
| 5 | POS tag of next word | Stanford NLP |
| 6 | POS tag of previous word | Stanford NLP |
| 7 | Next two words | Programming Language |
| 8 | Previous two words | Programming Language |
| 9 | Length of the word | Programming Language |
| 10 | Semantic group of the word | UMLS |
| 11 | Semantic group of next word | UMLS |
| 12 | Semantic group of previous word | UMLS |
| 13 | Exact match of bigram | UMLS |
| 14 | Exact match of trigram | UMLS |
| 15 | Exact match of reverse bigram | UMLS |
| 16 | CUI of the word | UMLS |
| 17 | MetaMap match of the word | MetaMap |
| 18 | MetaMap match of next word | MetaMap24 |
| 19 | MetaMap match of previous word | MetaMap |
| 20 | Lemmatized version of the word | Stanford NLP |
| 21 | Parent of the word in dependency tree | Stanford NLP |
| 22 | Abbreviation full name | List of Abbreviations |
| 23 | Abbreviation full name exact match into UMLS | UMLS |
| 24 | Abbreviation full name semantic group | UMLS |