|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. **Example** | **B. Pattern** | **C. Number of Related Instances** | **D. Number of Instances Identified** | **SCORE** |
| dd/mm/yy or dd/mm/yyyy | \\b\d{1,2}/\d{1,2}/\d{2,4}\\b | 1 | 1 | 1 |
| dd.mm.yyyy or dd.mm.yy | \\b\d{1,2}\.\d{1,2}\.\d{2,4}\\b | 1 | 1 | 1 |
| yyyy-mm-dd | \\b\d{4}-\d{1,2}-\d{1,2}\\b | 2 | 2 | 1 |
| Month dd, yyyy | (?:January|February|October|...|December) \d{1,2} \*\, \*\d{4}\\b | 4 | 4 | 1 |
| dd Mon,yyyy | \d{1,2} (?:Jan|Feb|Mar|Apr|...|Dec) \*\,\d{4} | 3 | 3 | 1 |
| 4.30 AM | \\b\d{1,2}\.\d{1,2} AM | 0 | 0 | - |
| 4:30 am | \\b\d{1,2}\:\d{1,2} am | 0 | 0 | - |
| 5.15 pm | \\b\d{1,2}\.\d{1,2} PM | 2 | 2 | 1 |
| 13:24 | \\b\d{2}\:\d{2} | 3 | 3 | 1 |
| 7382293451 | [^\+)\d] \d{10} | 1 | 1 | 1 |
| +917382293451 | \+\d{2}\d{10} | 1 | 1 | 1 |
| (+91) 738229345 | \(\+91\) \d{10} | 1 | 1 | 1 |
| 22469 | [^)\d] \d{5} | 1 | 1 | 1 |
| 4322469 | 043\d{5} | 1 | 1 | 1 |
| (043)22469 | \(043\)\d{5} | 1 | 1 | 1 |
| 10 dollars | \d+\.\d\* dollars | 1 | 1 | 1 |
| $ 10 | \$ \d+\.\d\* | 1 | 1 | 1 |
| 1046 | re.sub('([\d])(\d{3})','\\1,\\2',testcurcon) | Substitution | | |
| 100049 | re.sub('([\d])(\d{2})(\d{2})','\\1,\\2,\\3',testcurcon2) | Substitution | | |
| 1000496 | re.sub('([\d])(\d{3})(\d{3})','\\1,\\2,\\3',testcurcon3) | Substitution | | |
| abcd.yz [at] iiith [dot] ac [dot] in | \w+\.\w+ \[at\] iiith \[dot\] ac \[dot\] in | 0 | 0 | - |
| abcd [at] nitt.edu | \w+ \[at\] nitt.edu | 0 | 0 | - |
| abcd AT nitt DOT edu | \w+ AT nitt DOT edu | 1 | 1 | 1 |
| abcd.yz(AT) iiith (DOT) ac (DOT) in | \w+\.\w+\(AT\) iiith \(DOT\) ac \(DOT\) in | 1 | 1 | 1 |
| www.ntt.edu | [^/]www\.\w+\.\w+ | 1 | 1 | 1 |
| https://www.nitt.edu | https?\://www\.\w+\.\w+\.?\w\* | 2 | 2 | 1 |
| Rama Viswanath | [a-zA-z]{4,} [a-zA-Z]{3,} | 5 | 6 | False positives |
| Rama V. | \S{4,} \S\. | 2 | 2 | 1 |
| V.Rama | [a-zA-Z]{1}\.[a-zA-Z]{4,} | 2 | 2 | 1 |
| Subash Chandra Bose | [a-zA-z]{4,} [a-zA-Z]{3,} [a-zA-Z]{4,} | 1 | 4 | False positives |
| S. C. Bose | \\b[a-zA-Z]{1}\. [a-zA-Z]{1}\. [a-zA-Z]{4,} | 2 | 2 | 1 |
| Subash C. Bose | \\b[a-zA-Z]{3,} [a-zA-Z]{1}\. [a-zA-Z]{4,} | 1 | 1 | 1 |
| Mr. Subash Chandra Bose | Mr\. [a-zA-z]{4,} [a-zA-Z]{3,} [a-zA-Z]{4,} | 1 | 1 | 1 |
| Mr. S. C. Bose | Mr\. [a-zA-Z]{1}\. [a-zA-Z]{1}\. [a-zA-Z]{4,} | 1 | 1 | 1 |
| U.S.A | re.sub('[A-Z]\.[A-Z]\.[A-Z].','United States of America', testabr) | Substitution | | |
| I.T. | re.sub('[^A-Z\.][A-Z]\.[A-Z]\.[^A-Z]','Information Technology', testabr) | Substitution | | |