Topic Modeling for Research Articles

Researchers have access to large online archives of scientific articles. As a consequence, finding relevant articles has become more difficult. Tagging or topic modelling provides a way to give token of identification to research articles which facilitates recommendation and search process.

Given the abstract and title for a set of research articles, predict the topics for each article included in the test set.

Note that a research article can possibly have more than 1 topic. The research article abstracts and titles are sourced from the following 6 topics:

1. Computer Science

2. Physics

3. Mathematics

4. Statistics

5. Quantitative Biology

6. Quantitative Finance

**Data Dictionary**

**train.zip**

**train.csv**

|  |  |
| --- | --- |
| ID | Unique ID for each article |
| TITLE | Title of the research article |
| ABSTRACT | Abstract of the research article |
| Computer Science | Whether article belongs to topic computer science (1/0) |
| Physics | Whether article belongs to topic physics (1/0) |
| Mathematics | Whether article belongs to topic Mathematics (1/0) |
| Statistics | Whether article belongs to topic Statistics (1/0) |
| Quantitative Biology | Whether article belongs to topic Quantitative Biology (1/0) |
| Quantitative Finance | Whether article belongs to topic Quantitative Finance (1/0) |

**test.csv**

|  |  |
| --- | --- |
| ID | Unique ID for each article |
| TITLE | Title of the research article |
| ABSTRACT | Abstract of the research article |

**sample\_submission.csv**

|  |  |
| --- | --- |
| ID | Unique ID for each article |
| TITLE | Title of the research article |
| ABSTRACT | Abstract of the research article |
| Computer Science | Whether article belongs to topic computer science (1/0) |
| Physics | Whether article belongs to topic physics (1/0) |
| Mathematics | Whether article belongs to topic Mathematics (1/0) |
| Statistics | Whether article belongs to topic Statistics (1/0) |
| Quantitative Biology | Whether article belongs to topic Quantitative Biology (1/0) |
| Quantitative Finance | Whether article belongs to topic Quantitative Finance (1/0) |