

abstract

label

Samples: 300,000

Avg Length: 900 chars

Unique Words: 981,219

Samples: 300,000

Unique Classes: 5

Distribution: **Balanced**

Sample Labels: math, cs, cond-mat, astro-ph, physics

Preprocessing Options:

- ☒

☒ Text Cleaning (remove special chars) ?
- ☒

☒ Data Validation (remove nulls) ?
- ☒

☒ Category Mapping (convert to numeric) ?
- ☒

☒ Memory Optimization ?

Advanced Preprocessing Options:

- ☐

☐ Rare Words Removal ?
- ☐

☐ Context-aware Stopwords ?
- ☐

☐ Lemmatization ?
- ☐

☐ Phrase Detection ?

Column Preview:

	abstract	label
0	we describe a new algorithm the kellpebble game with colors and use it obtain a characterization of the family of kellsparse gra	math
1	the evolution of earthmoon system is described by the dark matter field fluid model proposed in the meeting of division of part	physics
2	we show that a determinant of stirling cycle numbers counts unlabeled acyclic singlesource automata the proof involves a bijec	math
3	in this paper we show how to compute the lambda_alpha norm alphage using the dyadic grid this result is a consequence of th	math
4	we study the twoparticle wave function of paired atoms in a fermi gas with tunable interaction strengths controlled by feshbac	cond-mat

Show Original Data

Save Column Configuration

Previous

Next