

Progress Tracker

Current Step

Step 2/6: Column Selection & Preprocessing

Status: In Progress

Step Status

Step 1: Dataset Selection (Completed)

Step 2: Column Selection & Preprocessing (Completed)

Step 3: Model Configuration (Pending)

Step 4: Training Execution (Pending)

Step 5: Results Analysis (Pending)

Step 6: Text Classification (Pending)

Reset Step

Save Progress

Topic Modeling - Auto Classifier

Intelligent Text Classification with Machine Learning

STEP 2/6: Column Selection & Preprocessing

Column Selection:

Chọn cột văn bản:

abstract

Chọn cột nhãn:

categories

Samples: 1,000

Unique Classes: 309

Avg Length: 201 chars

Distribution: Imbalanced

Unique Words: 7,789

Sample Labels: astro-ph, quant-ph, hep-th, hep-ph, cond-mat.mes-hall, ... (+304 more)

All validations passed! Columns are ready for processing.

Preprocessing Options:

- ☒

☒ Text Cleaning (remove special chars)
- ☒

☒ Data Validation (remove nulls)
- ☒

☒ Category Mapping (convert to numeric)
- ☒

☒ Memory Optimization

Column Preview:

	abstract	categories
0	A fully differential calculation in perturbative quantum chromodynamics is presented for the production of massive phot	hep-ph
1	We describe a new algorithm, the (k, ℓ) -pebble game with colors, and use it obtain a characterization of the family of \mathcal{H}_k	math.CO cs.CG
2	The evolution of Earth-Moon system is described by the dark matter field fluid model proposed in the Meeting of Division	physics.gen-ph
3	We show that a determinant of Stirling cycle numbers counts unlabeled acyclic single-source automata. The proof involve	math.CO
4	In this paper we show how to compute the $\ \Lambda_\alpha\ $ norm, $\ \alpha\ \geq 0$, using the dyadic grid. This result is :math:CA math.FA	math.CA math.FA
5	We study the two-particle wave function of paired atoms in a Fermi gas with tunable interaction strengths controlled by F	cond-mat.mes-hall
6	A rather non-standard quantum representation of the canonical commutation relations of quantum mechanics systems, b	gr-qc
7	A general formulation was developed to represent material models for applications in dynamic loading. Numerical metho	cond-mat.mtrl-sci
8	We discuss the results from the combined IRAC and MIPS c2d Spitzer Legacy observations of the Serpens star-forming reg	astro-ph
9	Partial cubes are isometric subgraphs of hypercubes. Structures on a graph defined by means of semicubes, and Djokovi\	math.CO

Save Column Configuration

◀ Previous

Next ▶

Skip to End