

## Aganitha - Python Take Home Exercise

Candidate: Aayush Patidar | ID: Naukri072025

### ✓ Problem:

Remove papers where all authors are from academic institutions.

### ✓ Tools Used:

- Python (pandas)
- LLM (ChatGPT)

### ✓ Approach:

1. Read the input CSV file.
2. Split AuthorAffiliations using semicolon.
3. Define keywords to detect academic terms (University, Institute, etc.)
4. Check each paper's affiliations:
  - If all affiliations are academic → Exclude
  - Else → Include
5. Save result in a new CSV.

### ✓ Output:

Generated `filtered\_output.csv` with correct results.

### ✓ LLM Contribution:

ChatGPT helped write clean and efficient logic, especially for string parsing and filtering.

✓ Verified:

Tested with multiple inputs. Fully academic-only papers are excluded.