

Machine Learning Classification

Uploading an CVS file

After logging in, click the CVS file tab.

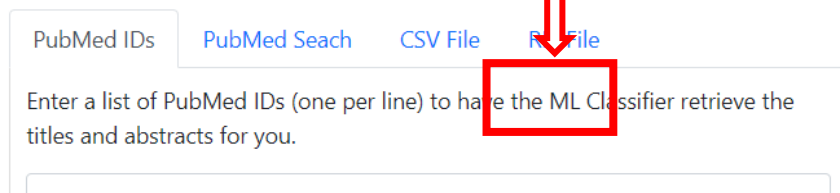
Machine Learning Classifier

Input Type

NB: A maximum of 1000 articles can be processed per request.

PubMed IDs PubMed Search CSV File **RIS File**

Enter a list of PubMed IDs (one per line) to have the ML Classifier retrieve the titles and abstracts for you.



Choose file to upload.

*Reminder that the file must be an CVS to upload and that it contains the required fields.

Input Type

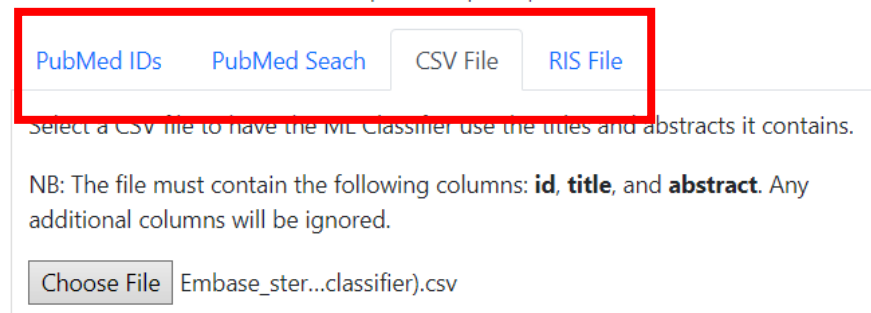
NB: A maximum of 1000 articles can be processed per request.

PubMed IDs PubMed Search CSV File **RIS File**

Select a CSV file to have the ML Classifier use the titles and abstracts it contains.

NB: The file must contain the following columns: **id**, **title**, and **abstract**. Any additional columns will be ignored.

Choose File Embase_ster...classifier).csv



Select CVS file in results format then hit classify to download.

Results Format

CSV File

Result Columns:

- **Classification** - 1 if the article is above our classification threshold, or 0 if below
- **Probability** - The probability score output by the selected model
- **Problem** - For articles where we could not calculate a Probability, this column will give the reason why

RIS File

Result Tags:

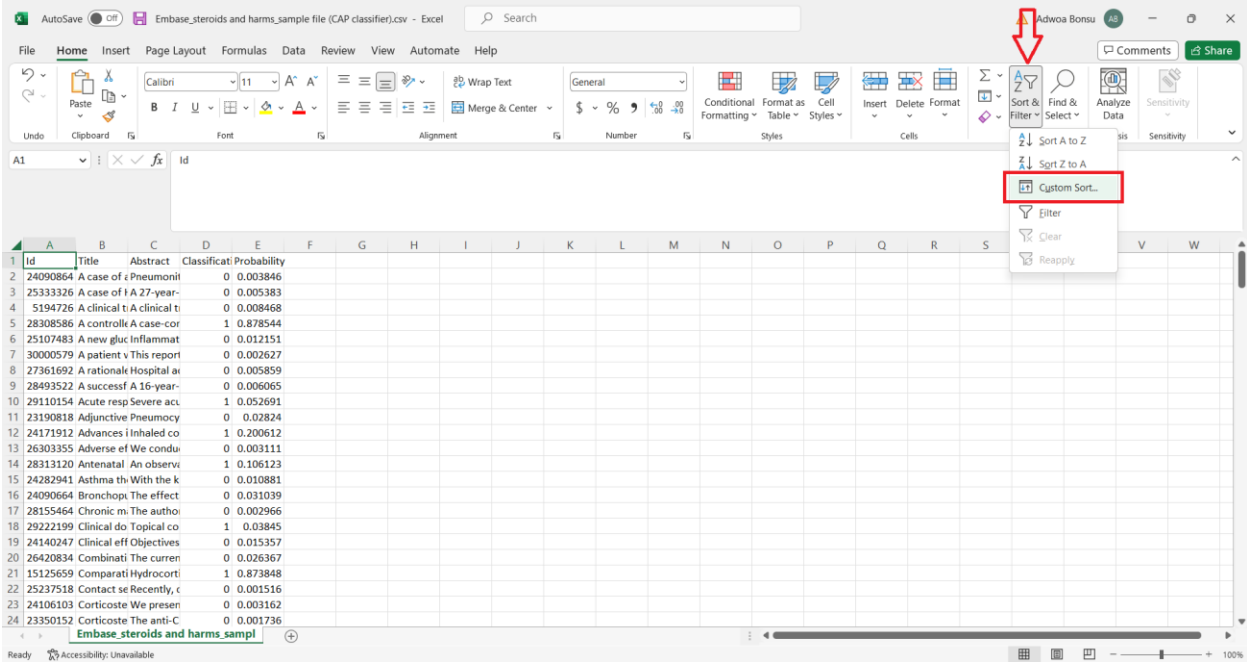
- **U1 - Classification** - 1 if the article is above our classification threshold, or 0 if below
- **U2 - Probability** - The probability score output by the selected model
- **U3 - Problem** - For articles where we could not calculate a Probability, this tag will give the reason why
- **ZT - Title** - The Title used by the ML model (only included when an RIS file is used for input, because it may contain records with multiple titles, so ZT clarifies which was used, and that it was parsed correctly)
- **ZA - Abstract** - The Abstract used by the ML model (only included when an RIS file is used for input, to ensure the Abstract was parsed correctly)

[Reset](#)

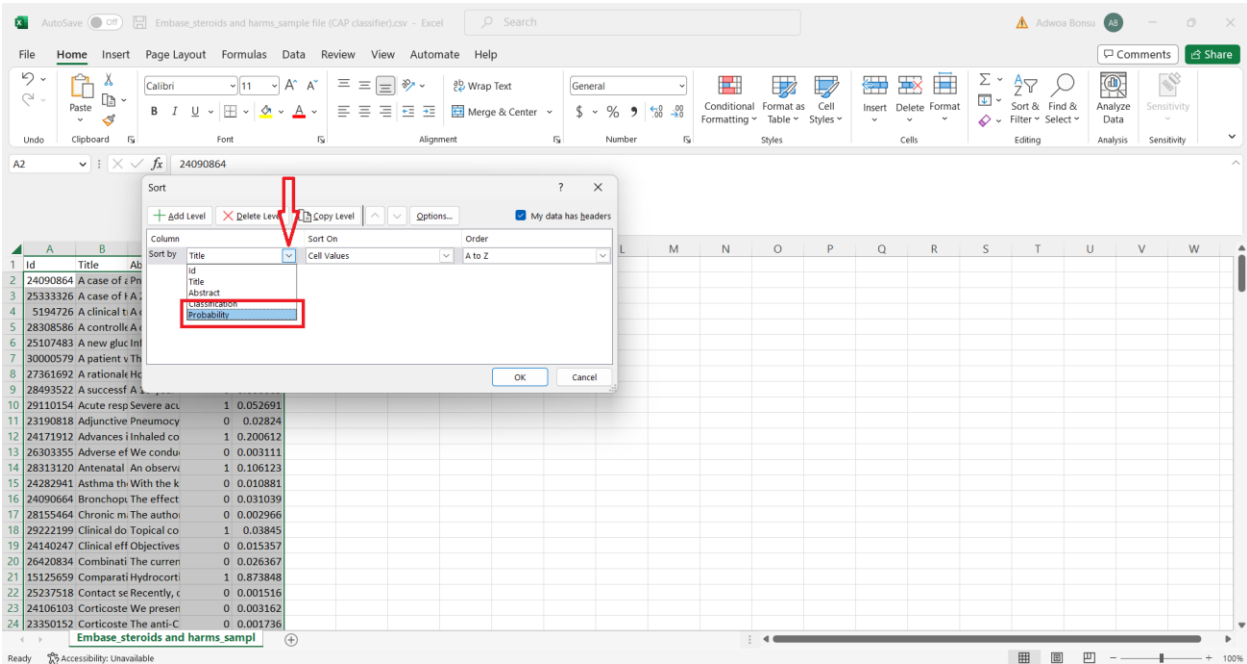
*The Classification threshold is the point where the studies reach 99% sensitivity for rigour. At this point, they may be considered for inclusion in HiRU Critical Appraisal Process; however, they are not guaranteed to pass for inclusion.

Open file in Microsoft Excel to view and sort results.

To sort click on Sort & Filter to Customize Sort



Sort by Probability



Organize by Highest to Lowest to view highest probability

The screenshot shows the Microsoft Excel interface with a 'Sort' dialog box open. The dialog box has the following settings:

- Column: Probability
- Sort On: Cell Values
- Order: Highest to Lowest (highlighted with a red box and a red arrow)
- My data has headers: checked

The background spreadsheet contains a table with the following data:

Id	Title	Ab
24351877	Ranitidine O	
28210520	Granisetron A	
24352717	Dexameth A	
24090472	The effect A	
29511128	Prospectiv In	
20217511	Prenatal ri To	
26246391	The treatm A	
27041694	The role o M	
6.14E+08	Cost effec Inhaled co	1.0935289
18184645	Treatment A randomi	1.0920248
22296393	Mometasc The safety	1.0913974
24096097	Rhinitis in Allergic rhi	1.0896398
29143305	Role of de Thirty pati	1.0896236
28308586	A controll A case-cor	1.0878544
15125659	Comparati Hydrocorti	1.0873848
29110946	Methylypre In five dou	1.0854161
24096094	Treating se Nonsedati	1.0779957
18240994	Report on Between s	1.0758002
18155176	Dexameth Nine childr	1.0618229
19247983	The use of In a multie	1.0458015
15136639	Hydrocorti In a left-ri	1.0298234
29454708	Tocolysis s OBJECTIVE	1.0277004
24171912	Advances i Inhaled co	1.0200612