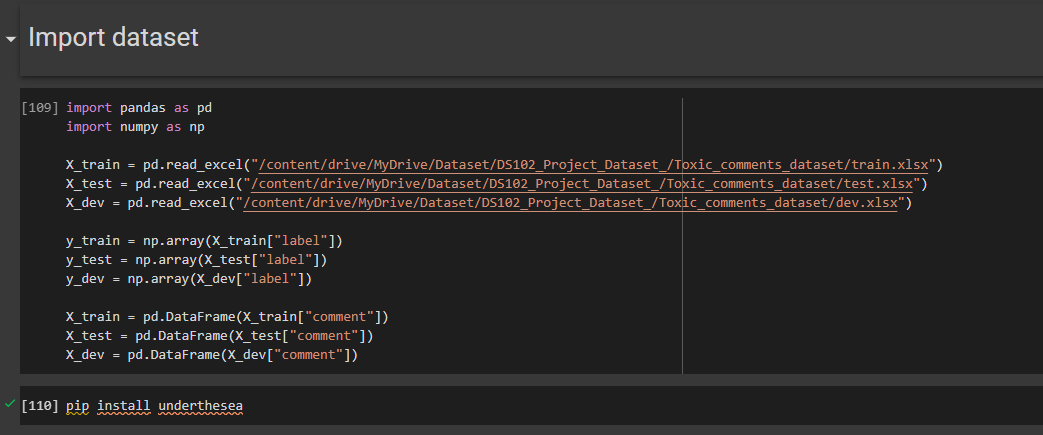
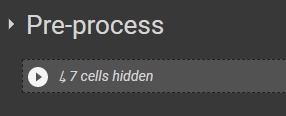
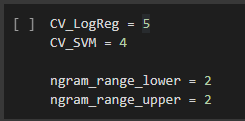
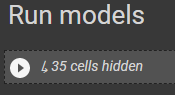
1. Import Dataset ở Cell Import dataset (thay đổi đường dẫn)
2. Ở cell Pre-Process: có thể chạy hoặc không, tùy thuộc vào phân công. 

* Ở cell này bao gồm xóa stop-word, xóa kí tự thừa, segment từ, xóa khoảng trắng,...

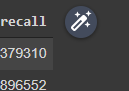
1. Cell Run models: Mở rộng cell Run models ra, cell đầu tiên sẽ có 4 thông số

Tùy thuộc phân công sẽ thay đổi thông số cụ thể. Ở file DS102\_model\_no\_gridsearch.ipynb sẽ không có 2 thông số CV\_LogReg và CV\_SVM

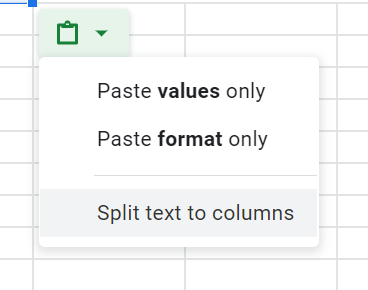
1. Sau khi chỉnh các thông số, thu gọn cell Run models lại, và chạy toàn bộ, không cần mở rộng các cell nhỏ hơn.
2. Khi chạy xong, mở rộng cell Run models, tìm cell Result, lưu và ghi lại bảng thông số vào file sheet Results trong thư mục chính.

Cách copy vào sheet:

* Chọn Convert this dataframe to an interactive table.



* Chọn Copy table
* Chọn định dạng CSV và chọn Copy
* Paste vào Sheet, chọn Split text to columns



Chú ý: không cần quan tâm các cell còn lại.

Các công việc cần làm:

* Ở file DS102\_model\_gridsearch.ipynb:
* Không chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Không chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Không chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Không chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Ở file DS102\_model\_no\_gridsearch.ipynb:
* Không chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 1, ngram\_range\_upper = 1
* Không chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Không chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Không chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Không chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 2, CV\_SVM = 2, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 3, CV\_SVM = 3, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 4, CV\_SVM = 4, ngram\_range\_lower = 2, ngram\_range\_upper = 2
* Chạy Pre-process, CV\_LogReg = 5, CV\_SVM = 5, ngram\_range\_lower = 2, ngram\_range\_upper = 2