Crawl dữ liệu từ trang web dạng infinity scroll:

[Hướng Dẫn Lấy Dữ Liệu Từ Web Bằng Python Dễ Hiểu](https://www.mcivietnam.com/blog-detail/huong-dan-lay-du-lieu-tu-web-bang-python-de-hieu/)

Chuyển đổi Unicode:

[Chuyển đổi Unicode dựng sẵn & tổ hợp với Python](https://quan.hoabinh.vn/post/2020/07/chuyen-doi-unicode-dung-san-to-hop-voi-python)

Giới thiệu tiền xử lý trong tiếng Việt

[trannguyenhan/preprocessing-data: Tiền xử lý dữ liệu tiếng Việt với 4 bước](https://github.com/trannguyenhan/preprocessing-data?tab=readme-ov-file)

[duyvuleo/VNTC: A Large-scale Vietnamese News Text Classification Corpus](https://github.com/duyvuleo/VNTC)

Phân loại tin tức tiếng Việt:

[Phân loại văn bản tiếng Việt sử dụng machine learning – Lập Trình Không Khó](https://nguyenvanhieu.vn/phan-loai-van-ban-tieng-viet/)

Phân loại tin tức và trực quan hóa dữ liệu với word cloud và spare matrix:

[Phân tích dữ liệu khám phá để xử lý ngôn ngữ tự nhiên: Hướng dẫn đầy đủ về các công cụ Python](https://neptune.ai/blog/exploratory-data-analysis-natural-language-processing-tools)

[Text Classification with NLP: Tf-Idf vs Word2Vec vs BERT | by Mauro Di Pietro | Towards Data Science](https://towardsdatascience.com/text-classification-with-nlp-tf-idf-vs-word2vec-vs-bert-41ff868d1794)

Sử dụng kỹ thuật bagging:

[Bagging và Boosting là gì? So sánh hai kỹ thuật Machine Learning, cách hoạt động, và ứng dụng thực tế](https://statio.vn/blog/bagging-va-boosting-la-gi-so-sanh-hai-ky-thuat-machine-learning-cach-hoat-ong-va-ung-dung-thuc-te)

**Document của scikit-learn, pandas, tqdm, pandarallel.**