**Đánh giá quá trình gán nhãn-Tính độ đồng thuận**

Code: <https://colab.research.google.com/drive/1iSq8Ipk8-MHScypWdhBwtJv0Bhe7oKYh?usp=sharing>

Bộ dữ liệu được gán nhãn bởi hai tình nguyện viên độc lập nhau. Việc gán nhãn được thực hiện sau khi các tình nguyện viên đọc kĩ và hiểu rõ guideline cũng như được giải thích rõ ràng các thắc mắc.

Xác suất đồng thuận hai người: Pr(a) = 0.5953177257525084

Xác xuất đồng thuận hai người đối với từng nhãn:

Autos & Vehicles : 0.0008053601190143287

Comedy : 0.0008053601190143287

Education : 0.0008053601190143287

Entertainment : 0.0008053601190143287

Film and Animation : 0.0008053601190143287

Gaming : 0.0008053601190143287

Howto & Style : 0.0008053601190143287

Music : 0.0008053601190143287

News & Politics : 0.0008053601190143287

Nonprofits & Activism : 0.0008053601190143287

People & Blogs : 0.0008053601190143287

Pets & Animals : 0.0008053601190143287

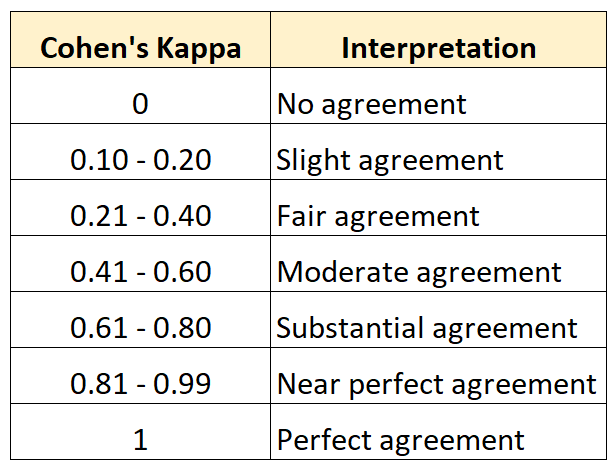
Science & Technology : 0.0008053601190143287

Sports : 0.0008053601190143287

Travel & Events : 0.0008053601190143287

Xác suất giả định của 2 annotator trên 15 nhãn: Pr(e) = 0.012080401785214934

Độ đồng thuận của 2 người: K = 0.5903692213629828



Dựa theo các khoảng giá trị của thang Kappa, thì 0.59 nằm ở mức Moderate agreement(0.41-0.6). Do đó, bản Guideline hiện tại và bộ dữ liệu hiện tại sẽ tiếp tục được sử dụng cho quá trình gán nhãn diện rộng.