

MOVIE GENRE CLASSIFICATION

TRẦN KHÔI NGUYỄN - 22520987

ĐỘNG LỰC PHÁT TRIỂN BÀI TOÁN

Trong thời đại công nghệ số, người dùng thường tìm kiếm phim dựa trên sở thích cá nhân.

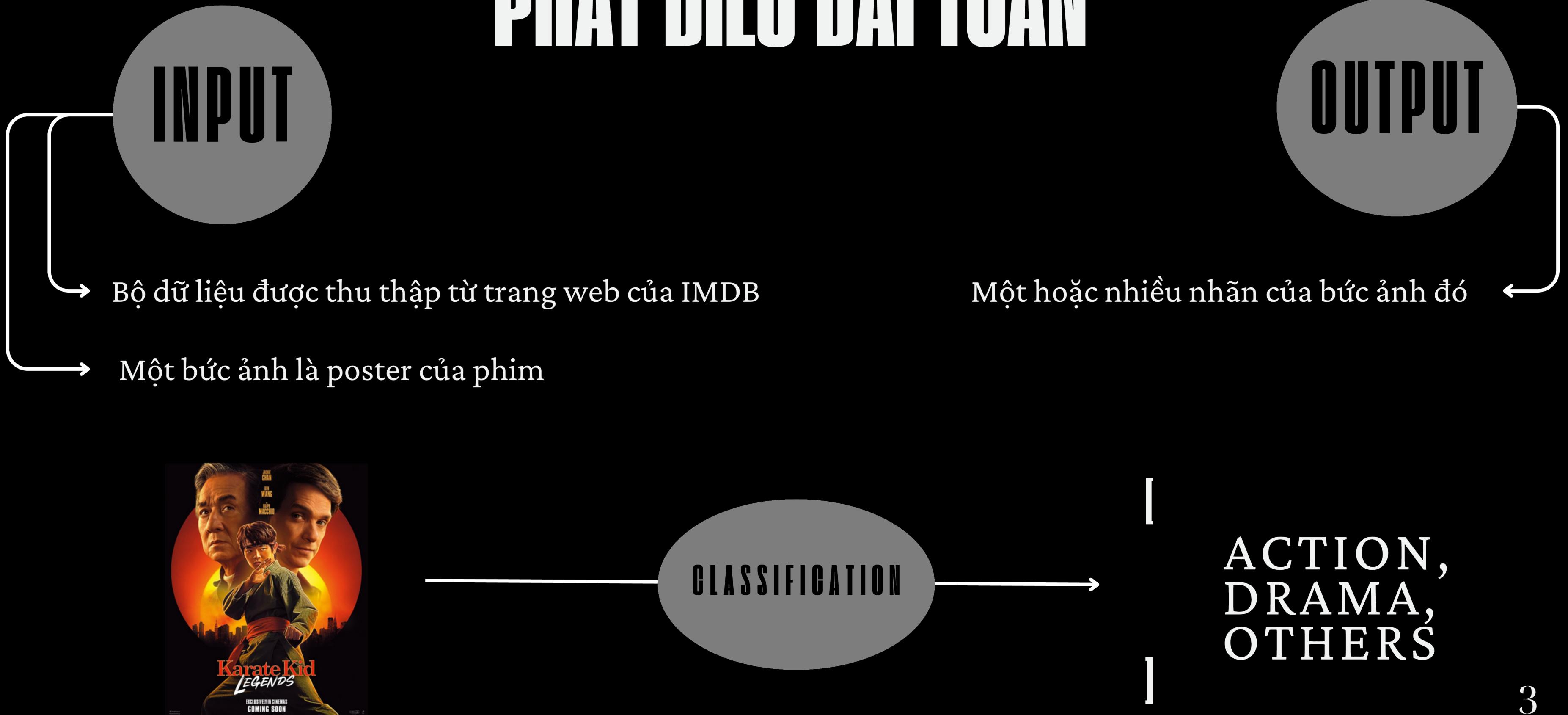
Việc tự động phân loại thể loại phim từ poster phim giúp hệ thống gợi ý phim chính xác hơn, tăng cường trải nghiệm người dùng

Một số ứng dụng thực tế có thể kể đến như là: Hệ thống gợi ý phim, chatbot,...

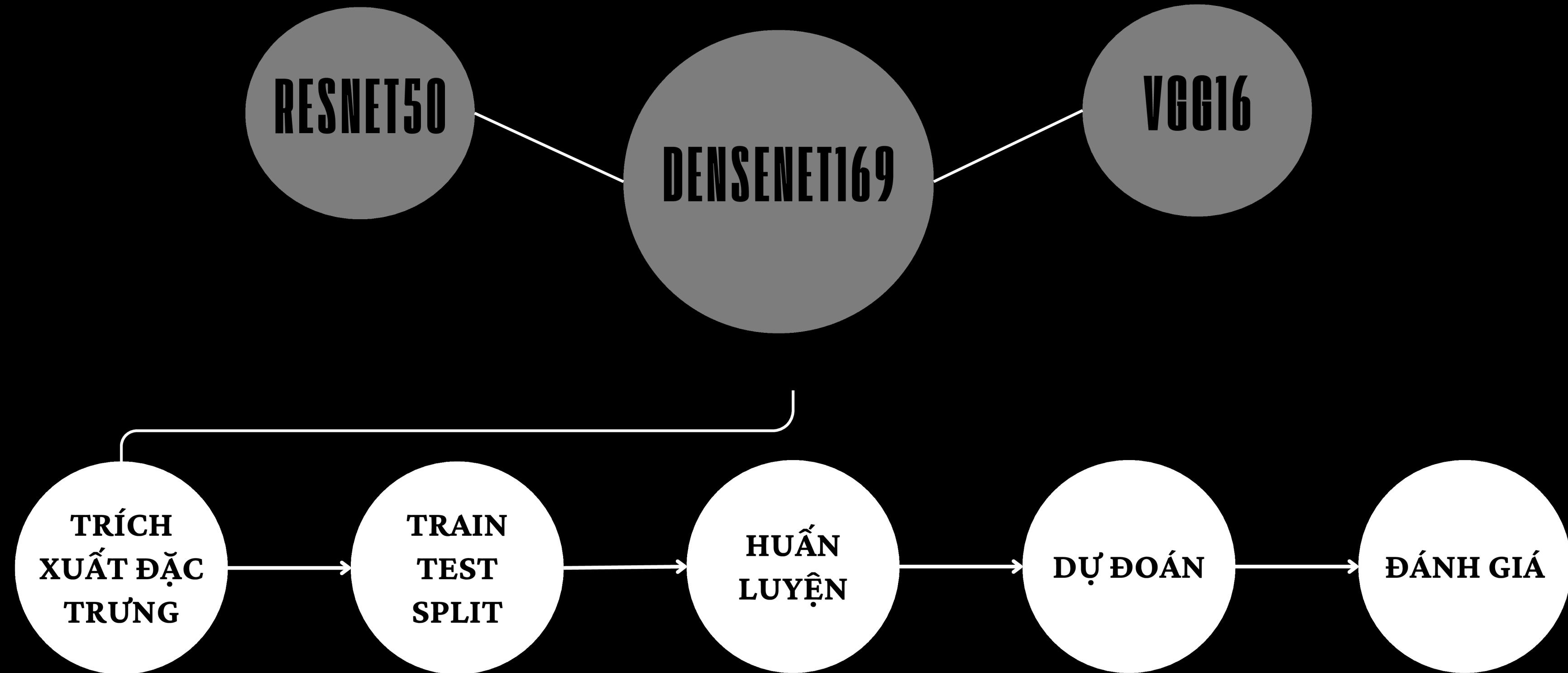
Nguồn dữ liệu phong phú: IMDB, TMDb



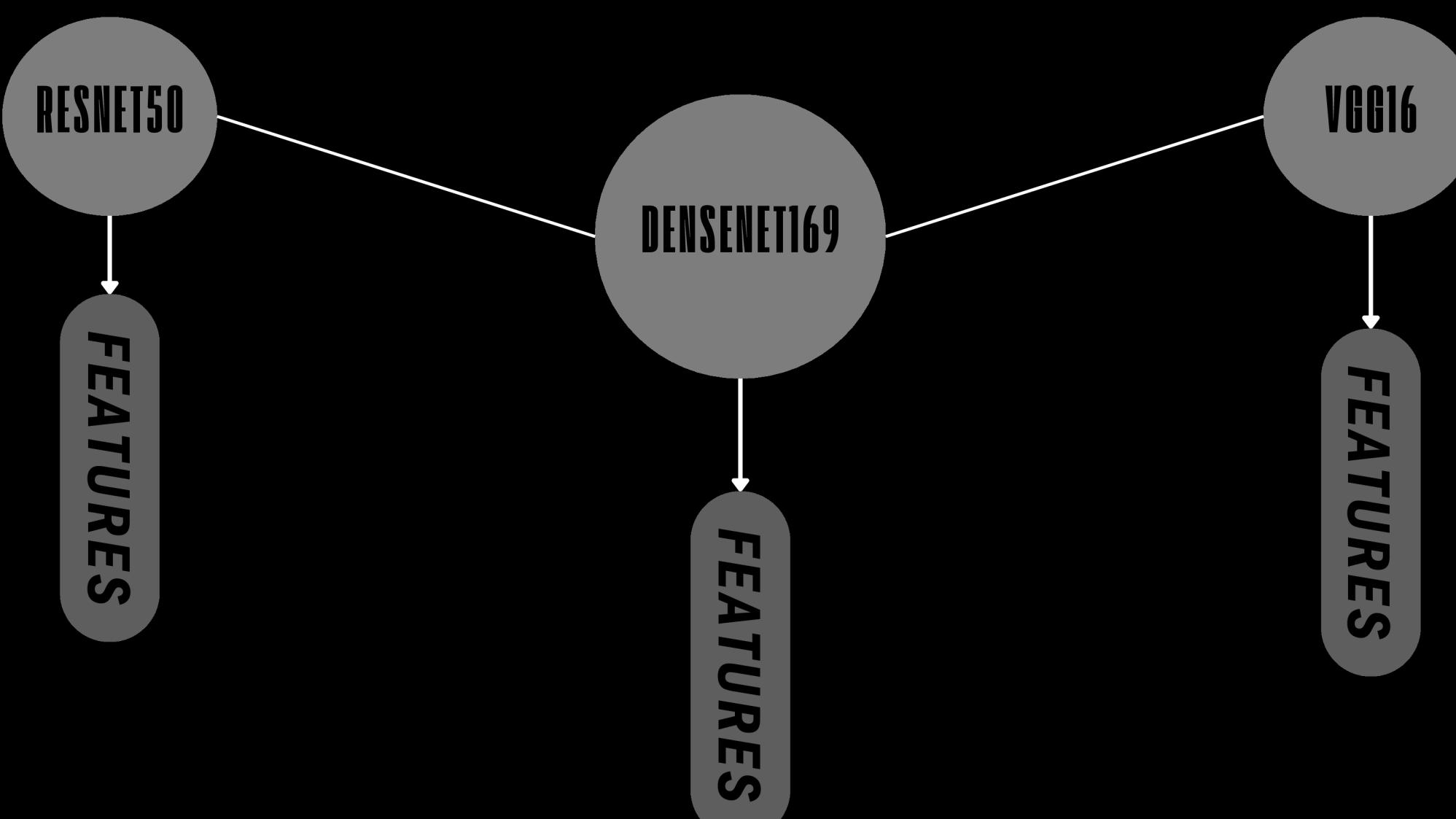
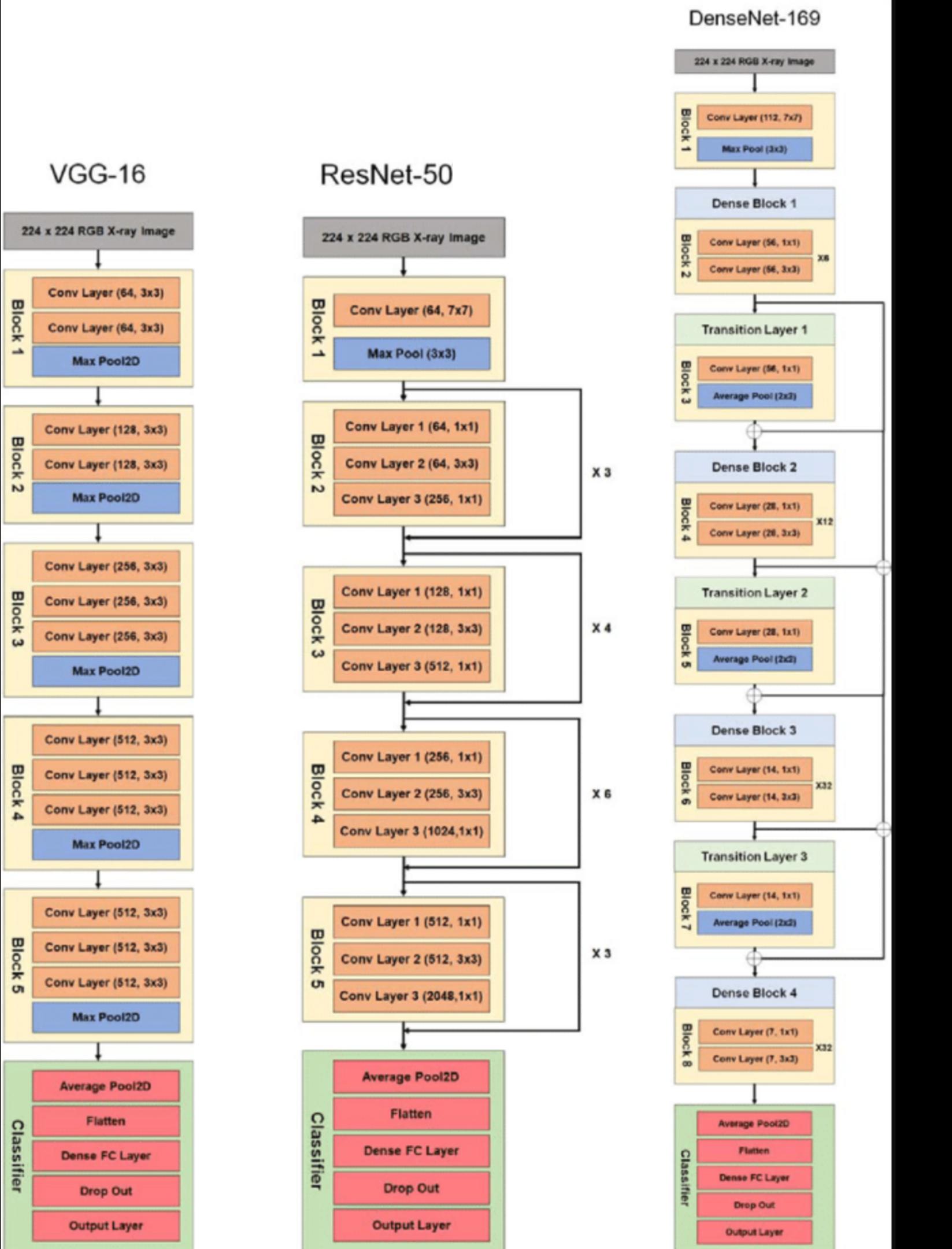
PHÁT BIỂU BÀI TOÁN



PHƯƠNG PHÁP



GIỚI THIỆU PHƯƠNG PHÁP TRÍCH SUẤT ĐẶC TRƯNG



VGG16

Input là một ảnh có kích thước 224 x 224 x 3

Có tổng cộng 21 lớp, 5 lớp Max Pooling, 16 lớp có tham số có thể huấn luyện

Là kiến trúc với số lớp lớn, số filter của từng block convolution tăng dần từ 64 - 128 - 256 - 512

Đạt vị trí 1st và 2nd cho tác vụ object detection và classification

RESNET50

Input là một ảnh có kích thước 224 x 224 x 3

Sử dụng Residual Block để giải quyết vấn đề Vanishing Gradient

Với 50 lớp, kiến trúc có thể học được nhiều các đặc trưng sâu hơn của ảnh

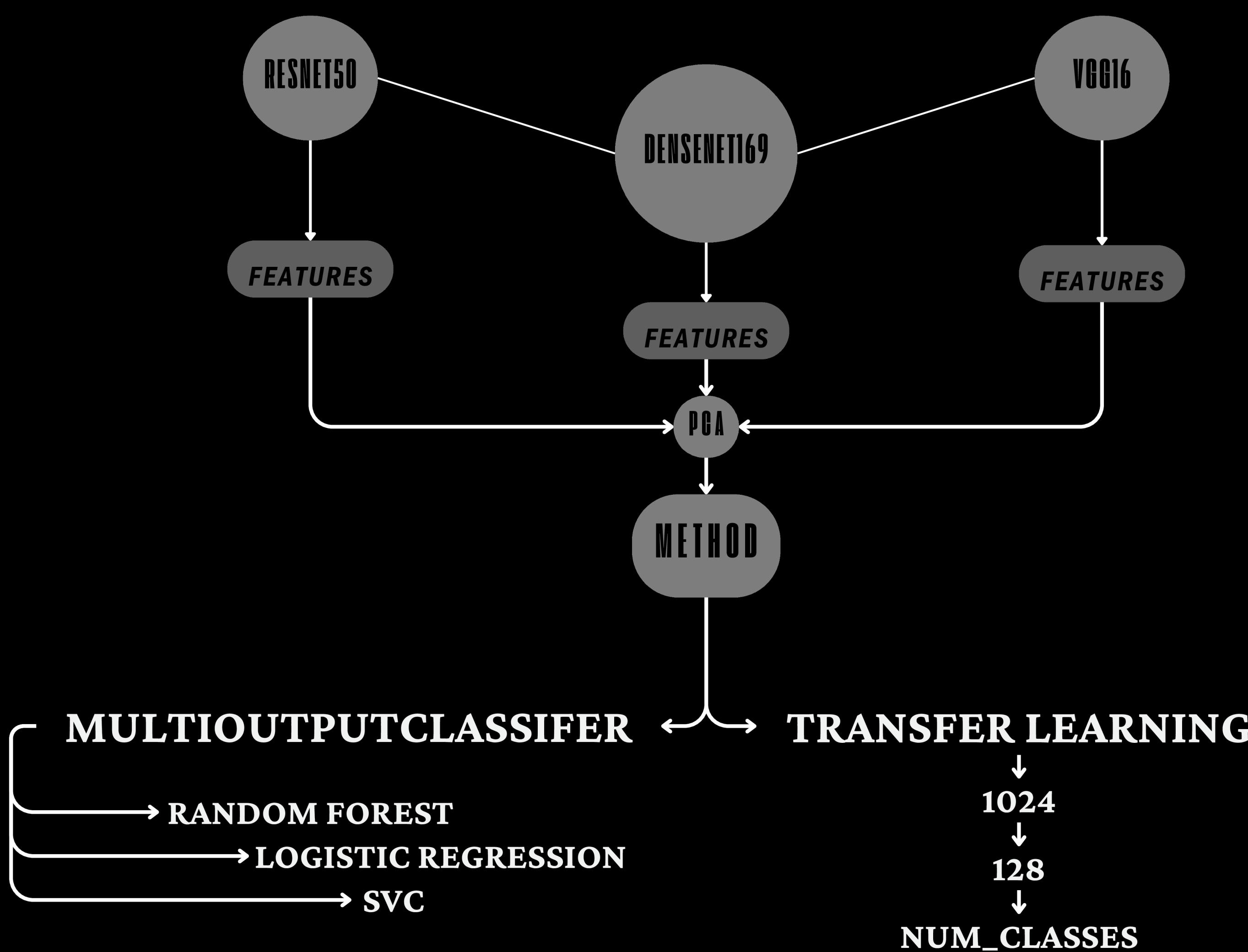
DENSENET169

Input là một ảnh có kích thước 224 x 224 x 3

Các phiên bản của Densenet từ 121 đến 264 lớp

Nổi bật với lớp Dense Block và Transition Layer giúp cải thiện khả năng học của mô hình

HUẤN LUYỆN



LOSS FUNCTION

SOFT F1 MACRO

$$\text{Soft-F1} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN} + \epsilon}$$

Where:

- **True Positives (TP)** = $\sum y_{\text{hat}} \cdot y$ (overlap between predictions and ground truth)
- **False Positives (FP)** = $\sum y_{\text{hat}} \cdot (1 - y)$ (predicted positive but not actually positive)
- **False Negatives (FN)** = $\sum (1 - y_{\text{hat}}) \cdot y$ (actual positive but not predicted as such)

sigmoidF1: A Smooth F1 Score Surrogate Loss for Multilabel Classification

DATASET: BỘ DỮ LIỆU

6 NHÃN
CHÍNH

[DOCUMENTARY, DRAMA, HORROR, MYSTERY, ROMANCE, SCIFI, THRILLER]

TỔNG
CỘNG 104
NHÃN

Lọc →

13 NHÃN

[
ACTION, ADVENTURE,
COMEDY, CRIME,
DOCUMENTARY, DRAMA,
FANTASY, HORROR,
MYSTERY, THRILLER,
ROMANCE, SCIFI, OTHERS
]

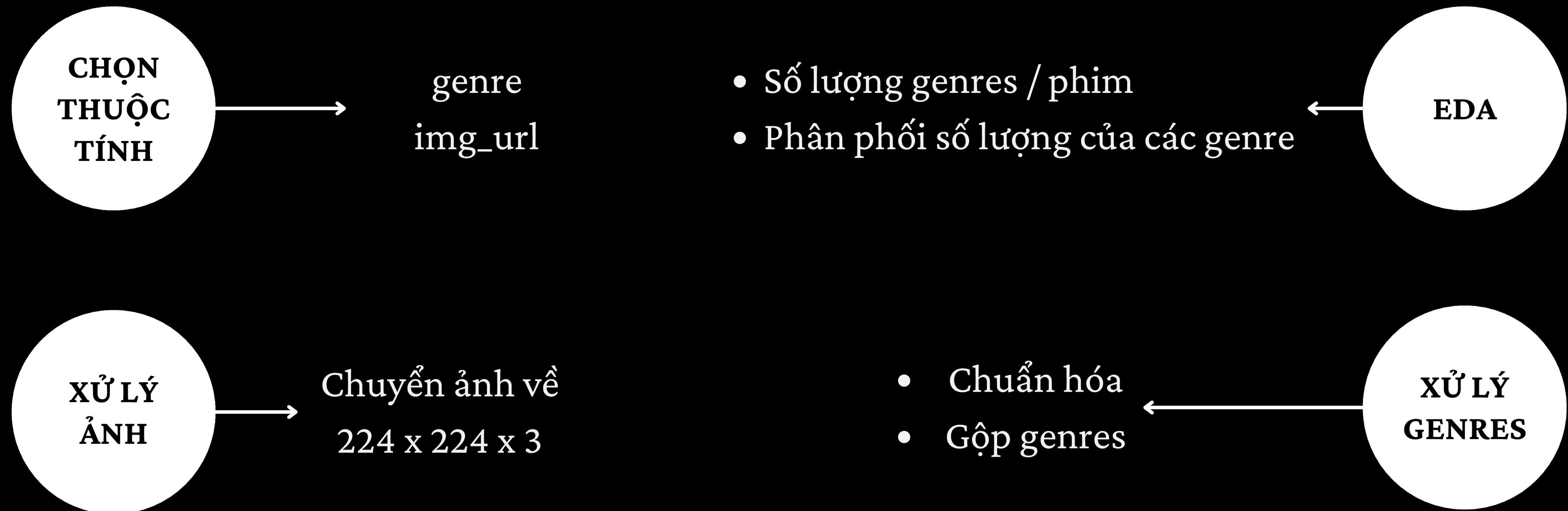
19525
PHIM

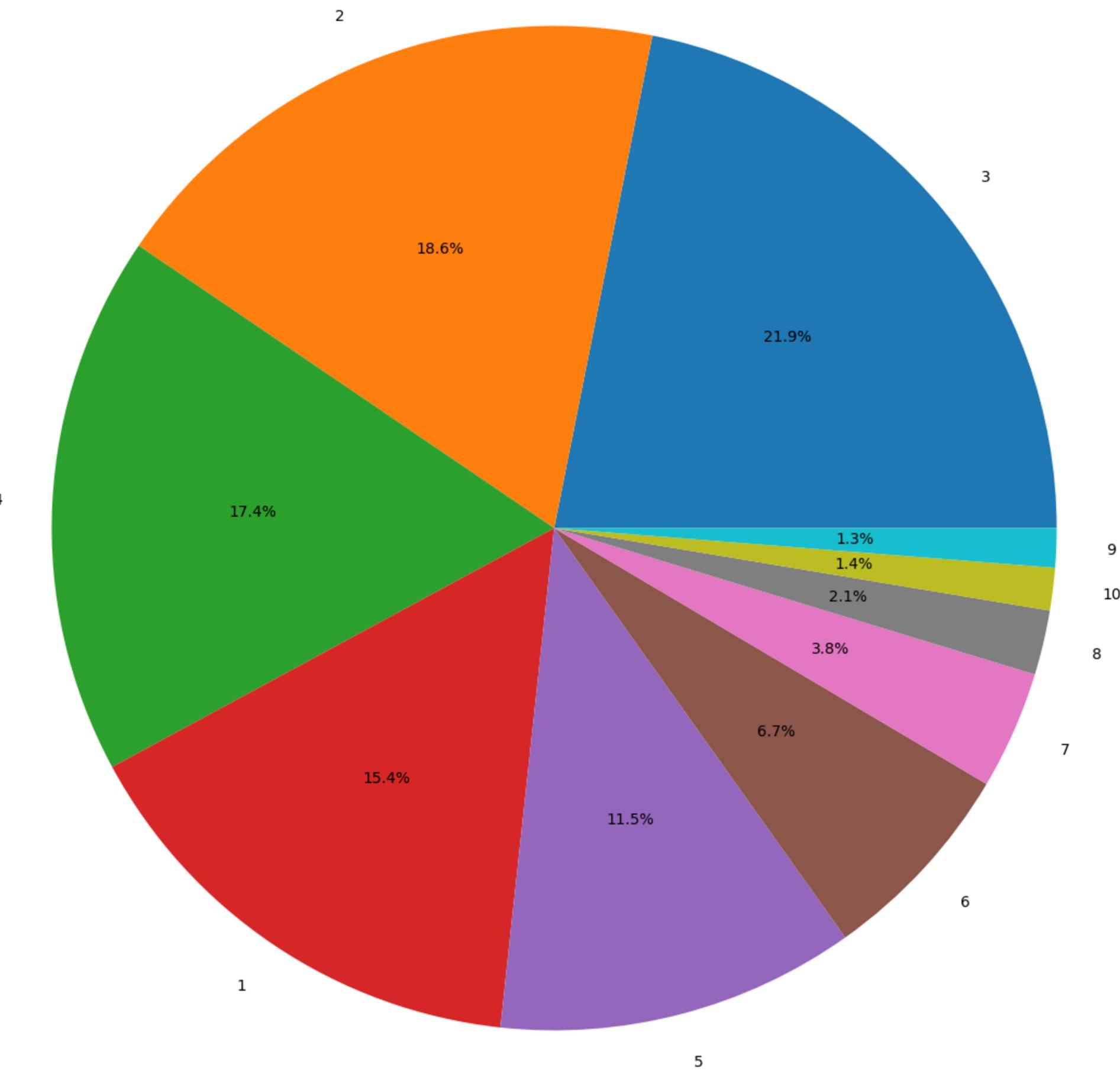
← BeautifulSoup →

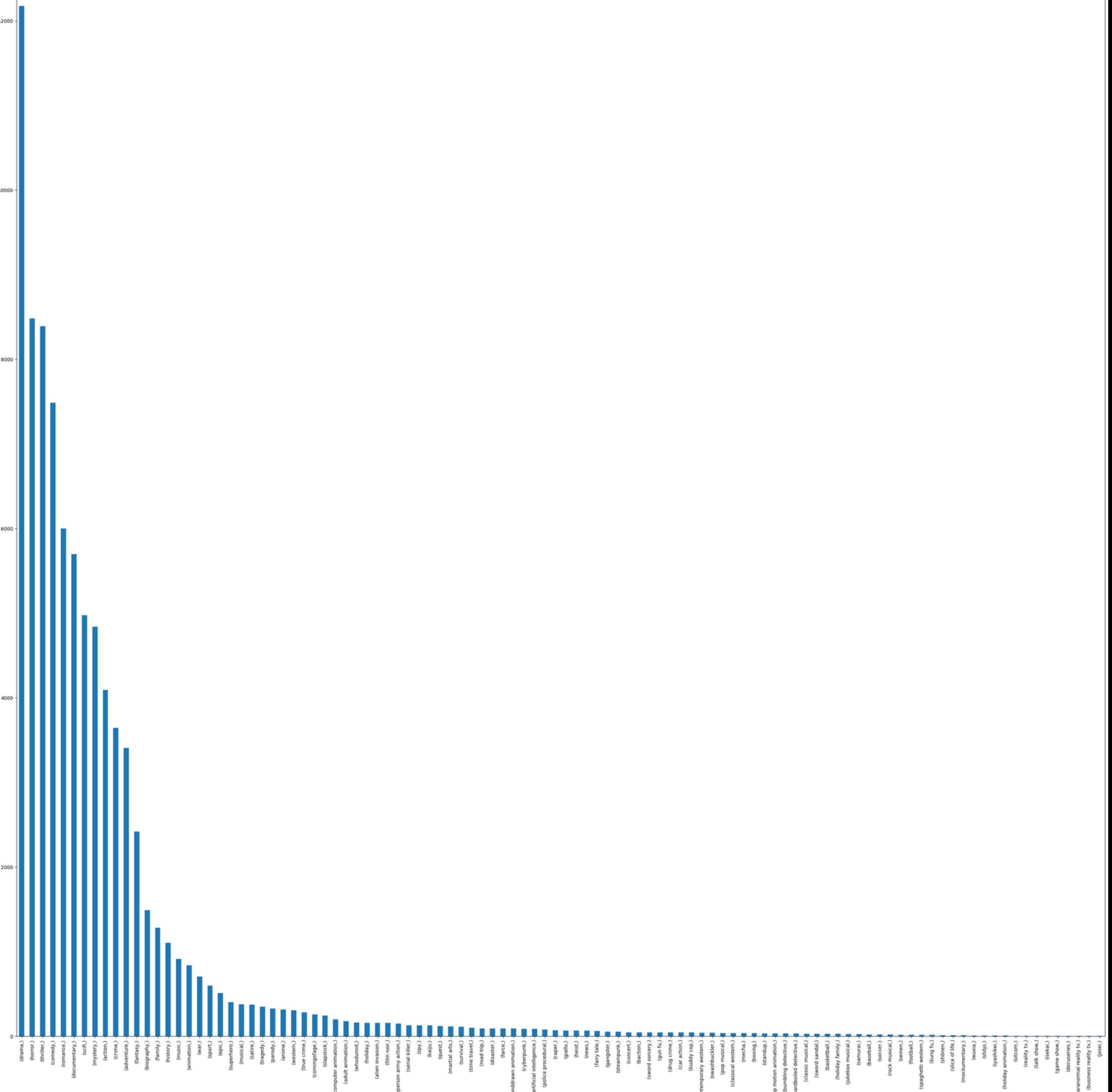
Selenium

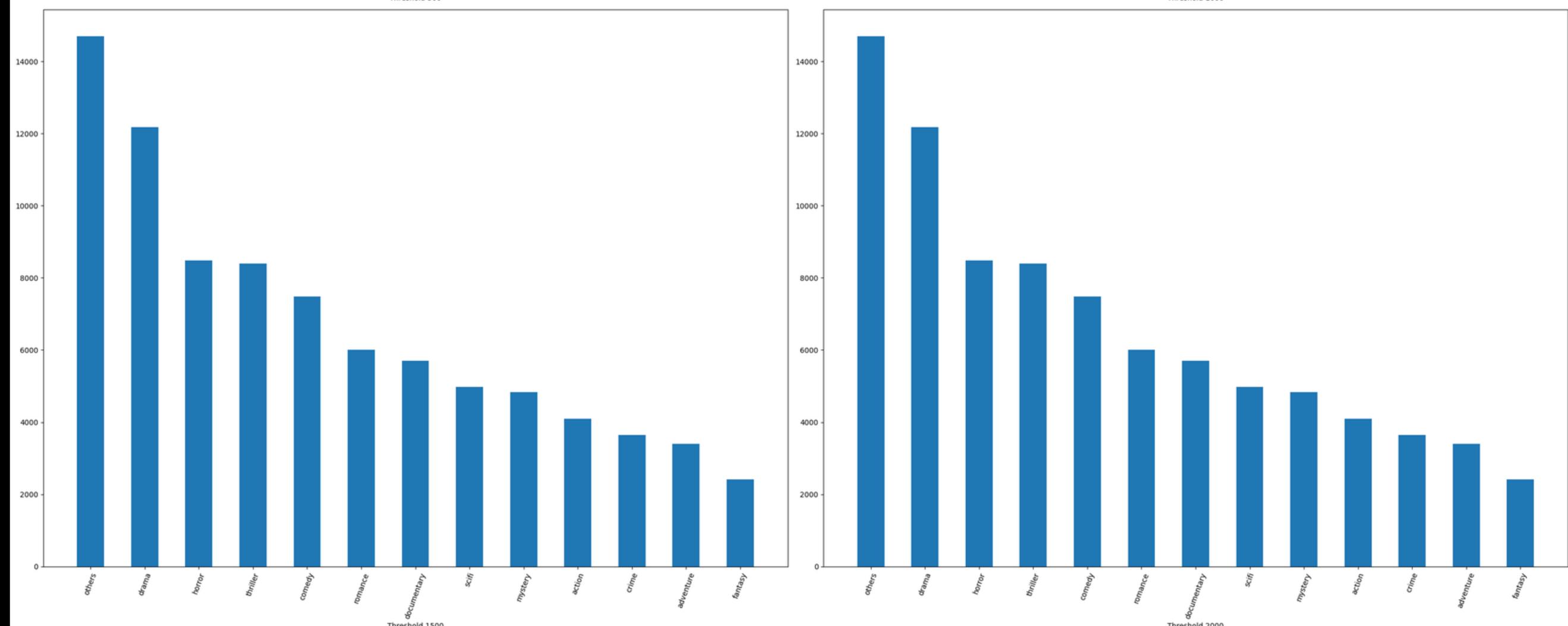
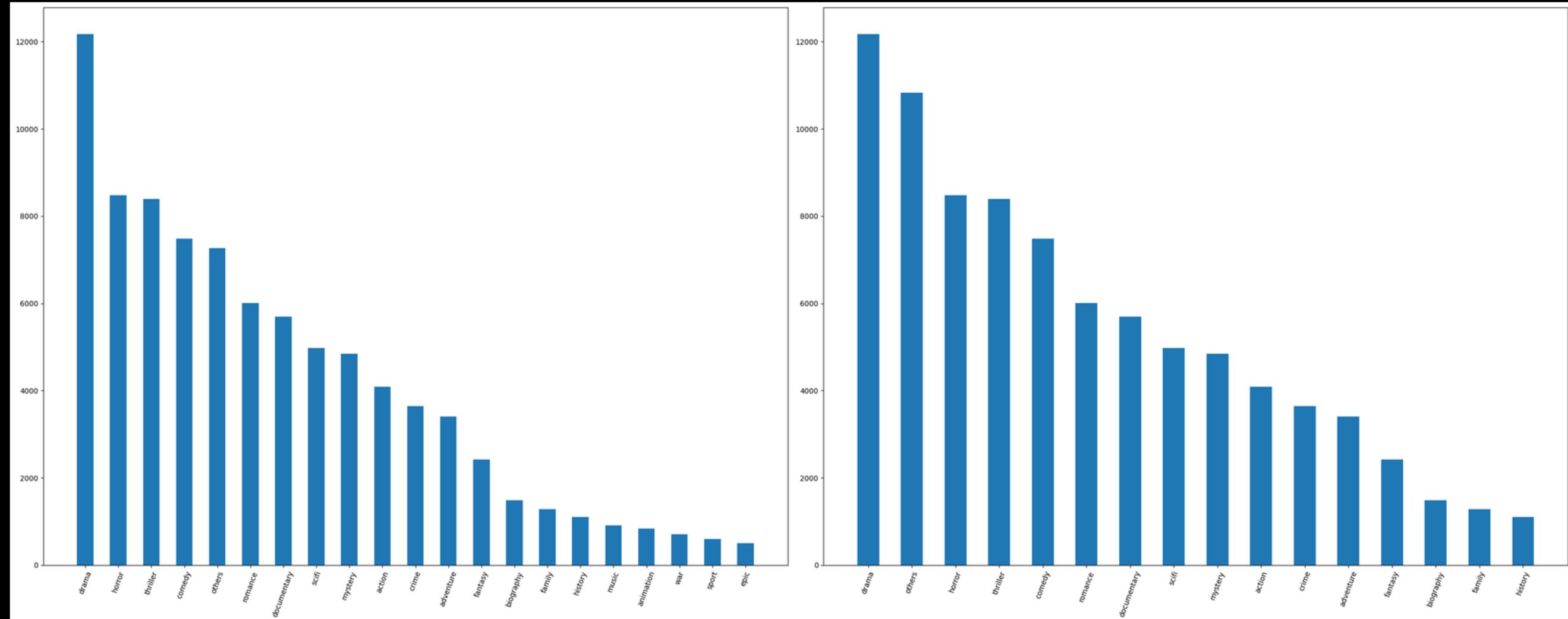
name, year, age, time, genre, storyline, img_url

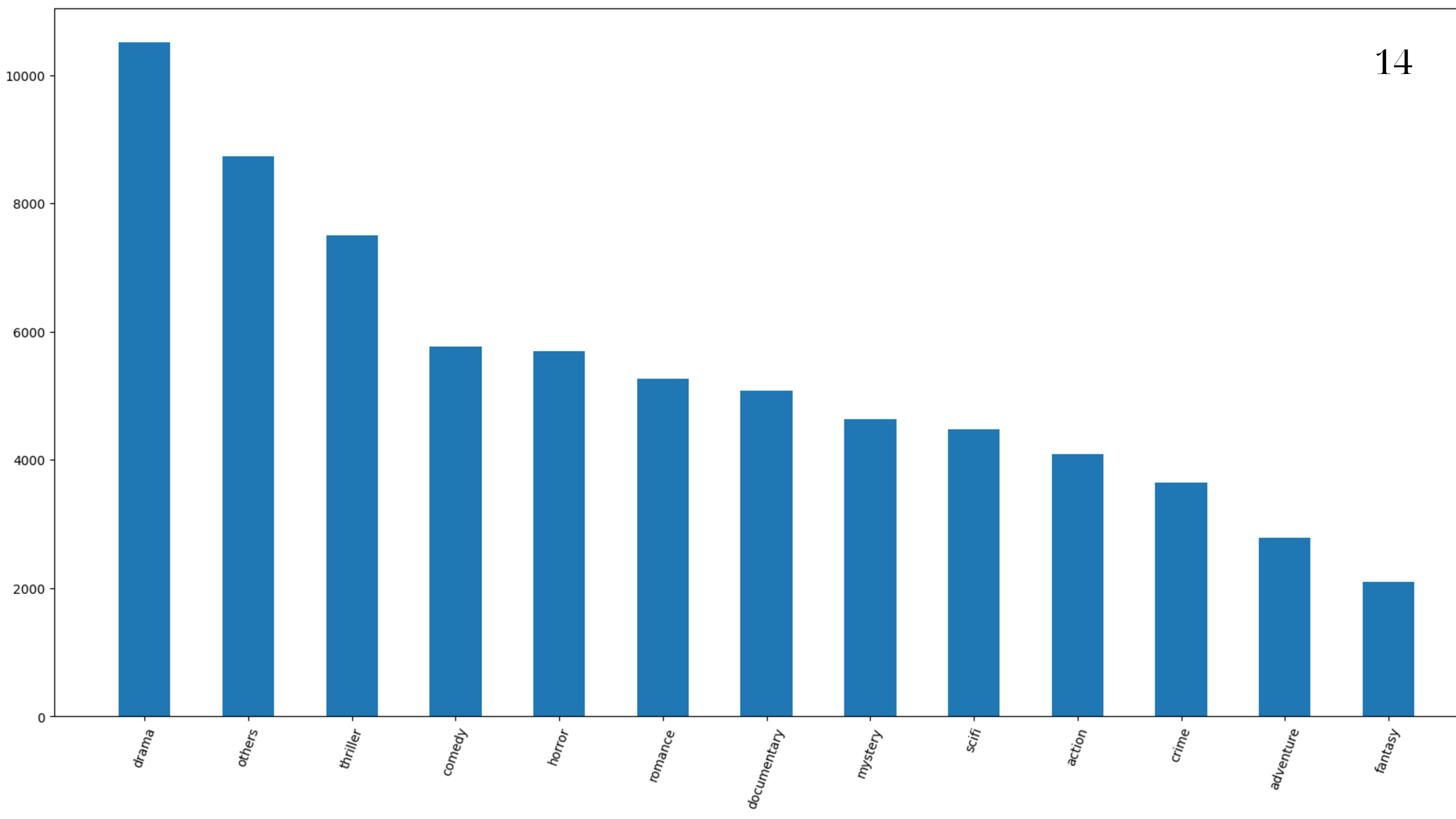
DATASET: TIỀN XỬ LÝ



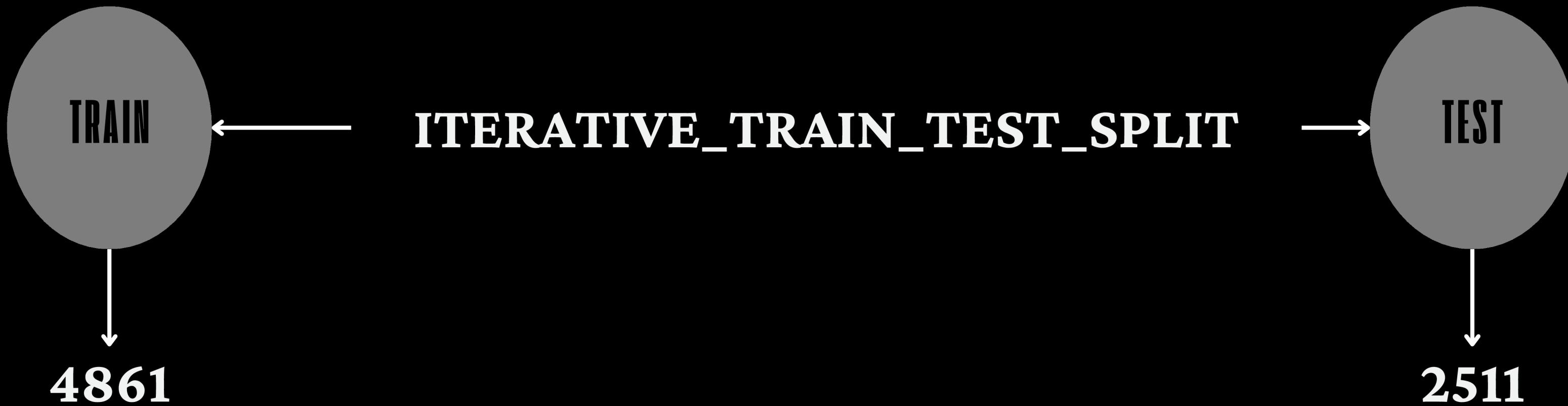








TRAIN TEST SPLIT



METRICS

F1
MACRO

$$F_1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

HAMMING
LOSS

$$\text{Hamming Loss} = \frac{1}{nL} \sum_{i=1}^n \sum_{j=1}^L I(y_i^j \neq \hat{y}_i^j)$$

PRECISION
MACRO

$$\text{Precision} = \frac{\text{Precision}_{\text{Class } A} + \text{Precision}_{\text{Class } B} + \dots + \text{Precision}_{\text{Class } N}}{N}$$

$$\text{Recall} = \frac{\text{Recall}_{\text{Class } A} + \text{Recall}_{\text{Class } B} + \dots + \text{Recall}_{\text{Class } N}}{N}$$

$$\text{Precision}_{\text{Class } A} = \frac{\text{TP}_{\text{Class } A}}{\text{TP}_{\text{Class } A} + \text{FP}_{\text{Class } A}}$$

$$\text{Recall}_{\text{Class } A} = \frac{\text{TP}_{\text{Class } A}}{\text{TP}_{\text{Class } A} + \text{FN}_{\text{Class } A}}$$

	VGG					DENSENET169				RESNET		
Metrics	f1_score (avg)	precision (avg)	recall (avg)	hamming loss	f1_score (avg)	precision (avg)	recall (avg)	hamming loss	f1_score (avg)	precision (avg)	recall (avg)	hamming loss
RANDOM FOREST	<u>0.35</u>	0.31	0.44	0.32	0.27	0.23	0.37	0.38	0.29	0.23	0.42	0.4
LOGISTIC REGRESSION	0.18	0.36	0.13	0.23	0.1	0.34	0.07	0.23	0.08	0.35	0.05	0.23
SVC	0.1	0.36	0.07	0.22	0.07	0.32	0.05	0.22	0.07	0.45	0.05	0.22

		precision	recall	f1-score	support
A C T I O N	0	0.33	0.27	0.30	475
A D V E N T U R E	1	0.25	0.12	0.17	318
C O M E D Y	2	0.43	0.55	0.48	640
C R I M E	3	0.23	0.13	0.17	375
D O C U M E N T A R Y	4	0.33	0.49	0.39	508
D R A M A	5	0.43	0.95	0.59	1051
F A N T A S Y	6	0.07	0.02	0.03	226
H O R R O R	7	0.33	0.58	0.42	569
M Y S T E R Y	8	0.20	0.14	0.17	463
T H R I L L E R	9	0.35	0.86	0.50	850
R O M A N C E	10	0.42	0.56	0.48	600
S C I F I	11	0.38	0.35	0.36	525
O T H E R S	12	0.35	0.73	0.47	749
micro avg		0.37	0.55	0.44	7349
macro avg		0.31	0.44	0.35	7349
weighted avg		0.34	0.55	0.40	7349
samples avg		0.38	0.56	0.42	7349

FANTASY, ADVENTURE, CRIME, MYSTERY

Hiệu suất rất thấp (F1-Score < 0.2), dữ liệu huấn luyện cho 4 nhãn này ít khiến cho mô hình không học được đặc trưng của các thể loại này

COMEDY, DRAMA, THRILLER, ROMANCE, OTHERS

- Hiệu suất cao nhất trong tất cả các thể loại, dữ liệu huấn luyện cho 5 nhãn này nhiều, khiến cho mô hình học được nhiều đặc trưng của các thể loại này (Recall cao)
- Tuy nhiên, do phân bố lệch về các thể loại này, hiện tượng bias có thể xảy ra

	VGG					DENSENET169					RESNET		
Metrics	f1_score (avg)	precision (avg)	recall (avg)	hamming loss	f1_score (avg)	precision (avg)	recall (avg)	hamming loss	f1_score (avg)	precision (avg)	recall (avg)	hamming loss	
thres = 0.25	0.4	0.36	0.53	0.32	0.38	0.27	0.76	0.54	0.37	0.26	0.86	0.63	
thres = 0.5	0.37	0.4	0.43	0.29	0.37	0.29	0.71	0.5	0.37	0.26	0.84	0.63	
thres = 0.75	0.32	0.43	0.34	0.27	0.36	0.31	0.64	0.47	0.37	0.27	0.83	0.62	

		precision	recall	f1-score	support
A C T I O N	0	0.40	0.44	0.42	475
A D V E N T U R E	1	0.34	0.21	0.26	318
C O M E D Y	2	0.47	0.51	0.49	641
C R I M E	3	0.25	0.27	0.26	375
D O C U M E N T A R Y	4	0.37	0.41	0.39	508
D R A M A	5	0.42	1.00	0.59	1051
F A N T A S Y	6	0.23	0.06	0.10	226
H O R R O R	7	0.43	0.55	0.48	569
M Y S T E R Y	8	0.22	0.64	0.33	463
T H R I L L E R	9	0.34	1.00	0.51	850
R O M A N C E	10	0.46	0.56	0.51	600
S C I F I	11	0.43	0.43	0.43	525
O T H E R S	12	0.36	0.78	0.50	750
	micro avg	0.37	0.62	0.46	7351
	macro avg	0.36	0.53	0.40	7351
	weighted avg	0.38	0.62	0.45	7351
	samples avg	0.37	0.62	0.44	7351

FANTASY, ADVENTURE, CRIME

Hiệu suất rất thấp nhất ($F1\text{-Score} < 0.3$), dữ liệu huấn luyện cho 3 nhãn này ít khiến cho mô hình không học được đặc trưng của các thể loại này

COMEDY, DRAMA, THRILLER, ROMANCE, HORRORS OTHERS

- Hiệu suất cao nhất trong tất cả các thể loại, dữ liệu huấn luyện cho 6 nhãn này nhiều, khiến cho mô hình học được nhiều đặc trưng của các thể loại này (Recall cao)

SO SÁNH

- Tốt hơn khi cho hiệu suất tốt hơn các mô hình ML truyền thống

THANK YOU