

Importing Libraries

```
In [220]: import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score, auc, f1_score, roc_curve
from tensorflow.keras.models import load_model
import tensorflow as tf
from sklearn.metrics import average_precision_score
from sklearn.metrics import precision_score
import matplotlib.pyplot as plt
import pickle
import torch
import cv
import sys
import glob
import json
```

Data Reading and Model Training

```
In [278]: # read and train the model
probabilities = dict()
def read_files():
    file_names = glob.glob('tt*')
    for i in file_names:
        print(i)
        for i in file_names:
            print(i)
            x_train = pd.read_csv(i+'x_train.csv', index_col=0)
            y_train = pd.read_csv(i+'y_train.csv', index_col=0)
            x_test = pd.read_csv(i+'x_test.csv', index_col=0)
            y_test = pd.read_csv(i+'y_test.csv', index_col=0)
            clf = RandomForestClassifier(random_state=42)
            clf = clf.fit(x_train, y_train.values.ravel())
            prob = clf.predict_proba(x_test)
            y_score = clf.decision_function(x_test)
            probabilities[i] = prob
```

```
In [279]: # Here the list of files that are read by the read_files()
read_files()

-->For tt0052357
-->For tt0063442
-->For tt0068646
-->For tt0073195
-->For tt0075314
-->For tt0078788
-->For tt0082089
-->For tt0086180
-->For tt0088247
-->For tt0088944
-->For tt0092099
-->For tt0096320
-->For tt0099423
-->For tt0100405
-->For tt0103776
-->For tt0103855
-->For tt0107290
-->For tt0108399
-->For tt0110604
-->For tt0112573
-->For tt0113272
-->For tt0114746
-->For tt0115759
-->For tt0116282
-->For tt0116522
-->For tt0117060
-->For tt0118715
-->For tt0119488
-->For tt0120382
-->For tt0120689
-->For tt0120890
-->For tt0123755
-->For tt0124315
-->For tt0127523
-->For tt0145025
-->For tt0172495
-->For tt0178868
-->For tt0190332
-->For tt0217505
-->For tt0253474
-->For tt0281358
-->For tt0319061
-->For tt0361748
-->For tt0379786
-->For tt0398201
-->For tt0409459
-->For tt0440963
-->For tt0443272
-->For tt0479884
-->For tt0780571
-->For tt0822832
-->For tt0845513
-->For tt0976051
-->For tt1001508
-->For tt1038919
-->For tt1099212
-->For tt1119646
-->For tt1205489
-->For tt1375666
-->For tt1412386
-->For tt1707386
-->For tt2024544
-->For tt2488496
-->For tt2582846
```



```

def scene_transition_ground_truths:
    pr_dict: Scene transition predictions.
    shot_to_end_frame_dict: End frame index for each shot.
    threshold: A threshold to filter the predictions.
Returns:
    Mean MIoU, and a dict of MIoU for each movie.

def iou(xn, y):
    eo, eo_x = x
    el, el_y = y
    amin, smax = (s0, s1) if el > s0 else (s1, s0)
    emin, emax = (e0, e1) if el > e0 else (el, e0)
    return (emin - smax + 1) / (emax - amin + 1)

def scene_frame_ranges(scene_transitions, shot_to_end_frame):
    end_shots = np.where(scene_transitions[:, 0])
    scenes = np.zeros((len(end_shots) + 1, 2), dtype=end_shots.dtype)
    scenes[:, -1, 1] = shot_to_end_frame[end_shots]
    scenes[:, 1, 1] = shot_to_end_frame[len(scene_transitions)]
    scenes[1:, 0] = scenes[:, -1, 1] + 1
    return scenes

def miou(gt_array, pr_array, shot_to_end_frame):
    gt_scenes = scene_frame_ranges(gt_array, shot_to_end_frame)
    pr_scenes = scene_frame_ranges(pr_array, shot_to_end_frame)
    assert gt_scenes[-1, -1] == pr_scenes[-1, -1]

    m = gt_scenes.shape[0]
    n = pr_scenes.shape[0]

    # iou for (gt scene, pr scene) pairs

```

```

    smax, smax = (s0, s1) if s1 > s0 else (s1, s0)

```

100% 100%


```

"t00052357": tensor([ 65, 1149, 1537, ..., 15948, 18955, 18617]),
dtype=torch.int32),
"t006642": tensor([ 31, 5922, 6198, ..., 24890, 24892, 25448]),
dtype=torch.int32),
"t0073195": tensor([ 1380, 1527, 1548, ..., 179465, 179522, 178463]),
dtype=torch.int32),
"t007514": tensor([ 1602, 1758, 1816, 2588, 2986, 3289, 3460, 3627, 4195,
4252, 4303, 4341, 4388, 4543, 4629, 4653, 4725, 4977,
4891, 4934, 5143, 5191, 5229, 5270, 5441, 5780, 5934,
5950, 5958, 5960, 5962, 5964, 5966, 5968, 5970, 5972,
5974, 5976, 5978, 5980, 5982, 5984, 5986, 5988, 5990,
5992, 5994, 5996, 5998, 6000, 6002, 6004, 6006, 6008,
6010, 6012, 6014, 6016, 6018, 6020, 6022, 6024, 6026,
6028, 6030, 6032, 6034, 6036, 6038, 6040, 6042, 6044,
6046, 6048, 6050, 6052, 6054, 6056, 6058, 6060, 6062,
6064, 6066, 6068, 6070, 6072, 6074, 6076, 6078, 6080,
6082, 6084, 6086, 6088, 6090, 6092, 6094, 6096, 6098,
6100, 6102, 6104, 6106, 6108, 6110, 6112, 6114, 6116,
6118, 6120, 6122, 6124, 6126, 6128, 6130, 6132, 6134,
6136, 6138, 6140, 6142, 6144, 6146, 6148, 6150, 6152,
6154, 6156, 6158, 6160, 6162, 6164, 6166, 6168, 6170,
6172, 6174, 6176, 6178, 6180, 6182, 6184, 6186, 6188,
6190, 6192, 6194, 6196, 6198, 6200, 6202, 6204, 6206,
6208, 6210, 6212, 6214, 6216, 6218, 6220, 6222, 6224,
6226, 6228, 6230, 6232, 6234, 6236, 6238, 6240, 6242,
6244, 6246, 6248, 6250, 6252, 6254, 6256, 6258, 6260,
6262, 6264, 6266, 6268, 6270, 6272, 6274, 6276, 6278,
6280, 6282, 6284, 6286, 6288, 6290, 6292, 6294, 6296,
6298, 6300, 6302, 6304, 6306, 6308, 6310, 6312, 6314,
6316, 6318, 6320, 6322, 6324, 6326, 6328, 6330, 6332,
6334, 6336, 6338, 6340, 6342, 6344, 6346, 6348, 6350,
6352, 6354, 6356, 6358, 6360, 6362, 6364, 6366, 6368,
6370, 6372, 6374, 6376, 6378, 6380, 6382, 6384, 6386,
6388, 6390, 6392, 6394, 6396, 6398, 6400, 6402, 6404,
6406, 6408, 6410, 6412, 6414, 6416, 6418, 6420, 6422,
6424, 6426, 6428, 6430, 6432, 6434, 6436, 6438, 6440,
6442, 6444, 6446, 6448, 6450, 6452, 6454, 6456, 6458,
6460, 6462, 6464, 6466, 6468, 6470, 6472, 6474, 6476,
6478, 6480, 6482, 6484, 6486, 6488, 6490, 6492, 6494,
6496, 6498, 6500, 6502, 6504, 6506, 6508, 6510, 6512,
6514, 6516, 6518, 6520, 6522, 6524, 6526, 6528, 6530,
6532, 6534, 6536, 6538, 6540, 6542, 6544, 6546, 6548,
6550, 6552, 6554, 6556, 6558, 6560, 6562, 6564, 6566,
6568, 6570, 6572, 6574, 6576, 6578, 6580, 6582, 6584,
6586, 6588, 6590, 6592, 6594, 6596, 6598, 6600, 6602,
6604, 6606, 6608, 6610, 6612, 6614, 6616, 6618, 6620,
6622, 6624, 6626, 6628, 6630, 6632, 6634, 6636, 6638,
6640, 6642, 6644, 6646, 6648, 6650, 6652, 6654, 6656,
6658, 6660, 6662, 6664, 6666, 6668, 6670, 6672, 6674,
6676, 6678, 6680, 6682, 6684, 6686, 6688, 6690, 6692,
6694, 6696, 6698, 6700, 6702, 6704, 6706, 6708, 6710,
6712, 6714, 6716, 6718, 6720, 6722, 6724, 6726, 6728,
6730, 6732, 6734, 6736, 6738, 6740, 6742, 6744, 6746,
6748, 6750, 6752, 6754, 6756, 6758, 6760, 6762, 6764,
6766, 6768, 6770, 6772, 6774, 6776, 6778, 6780, 6782,
6784, 6786, 6788, 6790, 6792, 6794, 6796, 6798, 6800,
6802, 6804, 6806, 6808, 6810, 6812, 6814, 6816, 6818,
6820, 6822, 6824, 6826, 6828, 6830, 6832, 6834, 6836,
6838, 6840, 6842, 6844, 6846, 6848, 6850, 6852, 6854,
6856, 6858, 6860, 6862, 6864, 6866, 6868, 6870, 6872,
6874, 6876, 6878, 6880, 6882, 6884, 6886, 6888, 6890,
6892, 6894, 6896, 6898, 6900, 6902, 6904, 6906, 6908,
6910, 6912, 6914, 6916, 6918, 6920, 6922, 6924, 6926,
6928, 6930, 6932, 6934, 6936, 6938, 6940, 6942, 6944,
6946, 6948, 6950, 6952, 6954, 6956, 6958, 6960, 6962,
6964, 6966, 6968, 6970, 6972, 6974, 6976, 6978, 6980,
6982, 6984, 6986, 6988, 6990, 6992, 6994, 6996, 6998,
7000, 7002, 7004, 7006, 7008, 7010, 7012, 7014, 7016,
7018, 7020, 7022, 7024, 7026, 7028, 7030, 7032, 7034,
7036, 7038, 7040, 7042, 7044, 7046, 7048, 7050, 7052,
7054, 7056, 7058, 7060, 7062, 7064, 7066, 7068, 7070,
7072, 7074, 7076, 7078, 7080, 7082, 7084, 7086, 7088,
7090, 7092, 7094, 7096, 7098, 7100, 7102, 7104, 7106,
7108, 7110, 7112, 7114, 7116, 7118, 7120, 7122, 7124,
7126, 7128, 7130, 7132, 7134, 7136, 7138, 7140, 7142,
7144, 7146, 7148, 7150, 7152, 7154, 7156, 7158, 7160,
7162, 7164, 7166, 7168, 7170, 7172, 7174, 7176, 7178,
7180, 7182, 7184, 7186, 7188, 7190, 7192, 7194, 7196,
7198, 7200, 72
```