# ABLATION STUDY

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 1.0, Sarcasm weight: 1.0, Fine-grained weight: 1.0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 1654.0674

✅ New best overall model at epoch 1 with main F1 = 0.5638

📊 Validation Metrics:

main: Acc = 0.7326, F1 = 0.5638, Precision = 0.6652, Recall = 0.5359

stereo: Acc = 0.7062, F1 = 0.6967, Precision = 0.7031, Recall = 0.7062

sarcasm: Acc = 0.7676, F1 = 0.5498, Precision = 0.7200, Recall = 0.5599

implicit\_fine: Acc = 0.5994, F1 = 0.5213, Precision = 0.5357, Recall = 0.5408

📈 Test Metrics:

main: Acc = 0.7311, F1 = 0.5556, Precision = 0.6331, Recall = 0.5310

stereo: Acc = 0.6954, F1 = 0.6803, Precision = 0.6950, Recall = 0.6952

sarcasm: Acc = 0.7550, F1 = 0.5268, Precision = 0.6597, Recall = 0.5440

implicit\_fine: Acc = 0.5606, F1 = 0.4820, Precision = 0.5053, Recall = 0.5016

🔁 Epoch 2/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 1080.9415

✅ New best overall model at epoch 2 with main F1 = 0.6114

📊 Validation Metrics:

main: Acc = 0.7358, F1 = 0.6114, Precision = 0.6205, Recall = 0.6083

stereo: Acc = 0.7300, F1 = 0.7353, Precision = 0.7456, Recall = 0.7300

sarcasm: Acc = 0.7514, F1 = 0.6191, Precision = 0.6520, Recall = 0.6090

implicit\_fine: Acc = 0.6201, F1 = 0.5357, Precision = 0.5383, Recall = 0.5400

📈 Test Metrics:

main: Acc = 0.7286, F1 = 0.5942, Precision = 0.6075, Recall = 0.5900

stereo: Acc = 0.7421, F1 = 0.7461, Precision = 0.7527, Recall = 0.7420

sarcasm: Acc = 0.7277, F1 = 0.5932, Precision = 0.6134, Recall = 0.5869

implicit\_fine: Acc = 0.6071, F1 = 0.5248, Precision = 0.5353, Recall = 0.5277

🔁 Epoch 3/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 649.7761

✅ New best overall model at epoch 3 with main F1 = 0.6156

📊 Validation Metrics:

main: Acc = 0.7521, F1 = 0.6156, Precision = 0.6541, Recall = 0.5956

stereo: Acc = 0.7349, F1 = 0.7406, Precision = 0.7535, Recall = 0.7349

sarcasm: Acc = 0.7153, F1 = 0.6134, Precision = 0.6159, Recall = 0.6113

implicit\_fine: Acc = 0.5935, F1 = 0.5443, Precision = 0.5925, Recall = 0.5320

📈 Test Metrics:

main: Acc = 0.7484, F1 = 0.5990, Precision = 0.6327, Recall = 0.5814

stereo: Acc = 0.7389, F1 = 0.7444, Precision = 0.7553, Recall = 0.7389

sarcasm: Acc = 0.7017, F1 = 0.6068, Precision = 0.6059, Recall = 0.6078

implicit\_fine: Acc = 0.5921, F1 = 0.5131, Precision = 0.5348, Recall = 0.5126

🔁 Epoch 4/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 423.0418

📊 Validation Metrics:

main: Acc = 0.7326, F1 = 0.6145, Precision = 0.6529, Recall = 0.6052

stereo: Acc = 0.7072, F1 = 0.7052, Precision = 0.7292, Recall = 0.7072

sarcasm: Acc = 0.7153, F1 = 0.5722, Precision = 0.5902, Recall = 0.5682

implicit\_fine: Acc = 0.6171, F1 = 0.5901, Precision = 0.6272, Recall = 0.5787

📈 Test Metrics:

main: Acc = 0.7260, F1 = 0.5950, Precision = 0.6297, Recall = 0.5891

stereo: Acc = 0.7286, F1 = 0.7296, Precision = 0.7429, Recall = 0.7285

sarcasm: Acc = 0.7061, F1 = 0.5783, Precision = 0.5877, Recall = 0.5744

implicit\_fine: Acc = 0.5937, F1 = 0.5539, Precision = 0.6195, Recall = 0.5431

🔁 Epoch 5/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 280.0227

✅ New best overall model at epoch 5 with main F1 = 0.6258

📊 Validation Metrics:

main: Acc = 0.7486, F1 = 0.6258, Precision = 0.6653, Recall = 0.6099

stereo: Acc = 0.6993, F1 = 0.7075, Precision = 0.7297, Recall = 0.6993

sarcasm: Acc = 0.7081, F1 = 0.5818, Precision = 0.5912, Recall = 0.5777

implicit\_fine: Acc = 0.6289, F1 = 0.5866, Precision = 0.6243, Recall = 0.5822

📈 Test Metrics:

main: Acc = 0.7500, F1 = 0.6103, Precision = 0.6400, Recall = 0.5983

stereo: Acc = 0.7484, F1 = 0.7519, Precision = 0.7689, Recall = 0.7483

sarcasm: Acc = 0.6945, F1 = 0.5617, Precision = 0.5696, Recall = 0.5591

implicit\_fine: Acc = 0.6252, F1 = 0.5693, Precision = 0.6757, Recall = 0.5658

⏱ Total time for config: 1248.61 seconds

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 0, Sarcasm weight: 0, Fine-grained weight: 1.0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=0, sarcasm=0, fine=1.0

Loss: 1073.1558

📊 Validation Metrics:

main: Acc = 0.7277, F1 = 0.5036, Precision = 0.7148, Recall = 0.4972

implicit\_fine: Acc = 0.6024, F1 = 0.5188, Precision = 0.5175, Recall = 0.5323

📈 Test Metrics:

main: Acc = 0.7367, F1 = 0.5099, Precision = 0.7339, Recall = 0.5036

implicit\_fine: Acc = 0.6157, F1 = 0.5285, Precision = 0.5293, Recall = 0.5425

🔁 Epoch 2/5 | config: stereo=0, sarcasm=0, fine=1.0

Loss: 722.7487

📊 Validation Metrics:

main: Acc = 0.7332, F1 = 0.5804, Precision = 0.6168, Recall = 0.5681

implicit\_fine: Acc = 0.6299, F1 = 0.5416, Precision = 0.5442, Recall = 0.5420

📈 Test Metrics:

main: Acc = 0.7370, F1 = 0.5852, Precision = 0.6299, Recall = 0.5698

implicit\_fine: Acc = 0.6283, F1 = 0.5514, Precision = 0.6104, Recall = 0.5445

🔁 Epoch 3/5 | config: stereo=0, sarcasm=0, fine=1.0

Loss: 465.7626

📊 Validation Metrics:

main: Acc = 0.7469, F1 = 0.6091, Precision = 0.6524, Recall = 0.5912

implicit\_fine: Acc = 0.6093, F1 = 0.5734, Precision = 0.5845, Recall = 0.5756

📈 Test Metrics:

main: Acc = 0.7505, F1 = 0.6018, Precision = 0.6488, Recall = 0.5841

implicit\_fine: Acc = 0.6094, F1 = 0.5353, Precision = 0.5597, Recall = 0.5407

🔁 Epoch 4/5 | config: stereo=0, sarcasm=0, fine=1.0

Loss: 286.9538

📊 Validation Metrics:

main: Acc = 0.7550, F1 = 0.6208, Precision = 0.6470, Recall = 0.6045

implicit\_fine: Acc = 0.6152, F1 = 0.5500, Precision = 0.5570, Recall = 0.5552

📈 Test Metrics:

main: Acc = 0.7528, F1 = 0.6101, Precision = 0.6382, Recall = 0.5939

implicit\_fine: Acc = 0.6165, F1 = 0.5435, Precision = 0.5519, Recall = 0.5483

🔁 Epoch 5/5 | config: stereo=0, sarcasm=0, fine=1.0

Loss: 176.8253

✅ New best overall model at epoch 5 with main F1 = 0.6273

📊 Validation Metrics:

main: Acc = 0.7530, F1 = 0.6273, Precision = 0.6303, Recall = 0.6245

implicit\_fine: Acc = 0.6122, F1 = 0.5594, Precision = 0.5660, Recall = 0.5640

📈 Test Metrics:

main: Acc = 0.7565, F1 = 0.6308, Precision = 0.6344, Recall = 0.6276

implicit\_fine: Acc = 0.6079, F1 = 0.5367, Precision = 0.5452, Recall = 0.5411

⏱ Total time for config: 855.07 seconds

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 1.0, Sarcasm weight: 1.0, Fine-grained weight: 0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=1.0, sarcasm=1.0, fine=0

Loss: 1003.5245

📊 Validation Metrics:

main: Acc = 0.7288, F1 = 0.4993, Precision = 0.6778, Recall = 0.4970

stereo: Acc = 0.7181, F1 = 0.7227, Precision = 0.7336, Recall = 0.7181

sarcasm: Acc = 0.7153, F1 = 0.6190, Precision = 0.6196, Recall = 0.6185

📈 Test Metrics:

main: Acc = 0.7328, F1 = 0.5030, Precision = 0.6198, Recall = 0.5008

stereo: Acc = 0.7255, F1 = 0.7287, Precision = 0.7334, Recall = 0.7254

sarcasm: Acc = 0.6787, F1 = 0.5870, Precision = 0.5847, Recall = 0.5906

🔁 Epoch 2/5 | config: stereo=1.0, sarcasm=1.0, fine=0

Loss: 669.9247

📊 Validation Metrics:

main: Acc = 0.7387, F1 = 0.5664, Precision = 0.7011, Recall = 0.5354

stereo: Acc = 0.7319, F1 = 0.7353, Precision = 0.7487, Recall = 0.7319

sarcasm: Acc = 0.6991, F1 = 0.5964, Precision = 0.5971, Recall = 0.5957

📈 Test Metrics:

main: Acc = 0.7412, F1 = 0.5685, Precision = 0.6968, Recall = 0.5388

stereo: Acc = 0.7429, F1 = 0.7466, Precision = 0.7536, Recall = 0.7428

sarcasm: Acc = 0.6801, F1 = 0.5837, Precision = 0.5821, Recall = 0.5858

🔁 Epoch 3/5 | config: stereo=1.0, sarcasm=1.0, fine=0

Loss: 409.6738

📊 Validation Metrics:

main: Acc = 0.7422, F1 = 0.5999, Precision = 0.6629, Recall = 0.5740

stereo: Acc = 0.7616, F1 = 0.7657, Precision = 0.7763, Recall = 0.7616

sarcasm: Acc = 0.6649, F1 = 0.5770, Precision = 0.5746, Recall = 0.5824

📈 Test Metrics:

main: Acc = 0.7432, F1 = 0.5935, Precision = 0.6441, Recall = 0.5714

stereo: Acc = 0.7484, F1 = 0.7493, Precision = 0.7600, Recall = 0.7483

sarcasm: Acc = 0.6614, F1 = 0.5777, Precision = 0.5752, Recall = 0.5847

🔁 Epoch 4/5 | config: stereo=1.0, sarcasm=1.0, fine=0

Loss: 248.5972

📊 Validation Metrics:

main: Acc = 0.7349, F1 = 0.5743, Precision = 0.6516, Recall = 0.5461

stereo: Acc = 0.7567, F1 = 0.7579, Precision = 0.7619, Recall = 0.7567

sarcasm: Acc = 0.7477, F1 = 0.5517, Precision = 0.6313, Recall = 0.5563

📈 Test Metrics:

main: Acc = 0.7398, F1 = 0.5632, Precision = 0.6504, Recall = 0.5383

stereo: Acc = 0.7555, F1 = 0.7560, Precision = 0.7566, Recall = 0.7554

sarcasm: Acc = 0.7522, F1 = 0.5856, Precision = 0.6479, Recall = 0.5803

🔁 Epoch 5/5 | config: stereo=1.0, sarcasm=1.0, fine=0

Loss: 169.9010

📊 Validation Metrics:

main: Acc = 0.7291, F1 = 0.5842, Precision = 0.6122, Recall = 0.5671

stereo: Acc = 0.7329, F1 = 0.7384, Precision = 0.7518, Recall = 0.7329

sarcasm: Acc = 0.6919, F1 = 0.5867, Precision = 0.5874, Recall = 0.5861

📈 Test Metrics:

main: Acc = 0.7325, F1 = 0.5888, Precision = 0.6184, Recall = 0.5707

stereo: Acc = 0.7358, F1 = 0.7409, Precision = 0.7520, Recall = 0.7357

sarcasm: Acc = 0.6974, F1 = 0.5988, Precision = 0.5985, Recall = 0.5992

⏱ Total time for config: 1048.24 seconds

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 0, Sarcasm weight: 0, Fine-grained weight: 0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 293.1129

📊 Validation Metrics:

main: Acc = 0.7416, F1 = 0.5969, Precision = 0.6317, Recall = 0.5795

📈 Test Metrics:

main: Acc = 0.7402, F1 = 0.5774, Precision = 0.6162, Recall = 0.5619

🔁 Epoch 2/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 220.8149

📊 Validation Metrics:

main: Acc = 0.7556, F1 = 0.6107, Precision = 0.6627, Recall = 0.5845

📈 Test Metrics:

main: Acc = 0.7565, F1 = 0.5954, Precision = 0.6514, Recall = 0.5705

🔁 Epoch 3/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 160.4168

📊 Validation Metrics:

main: Acc = 0.7527, F1 = 0.6027, Precision = 0.6722, Recall = 0.5729

📈 Test Metrics:

main: Acc = 0.7526, F1 = 0.6038, Precision = 0.6678, Recall = 0.5753

🔁 Epoch 4/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 96.8237

📊 Validation Metrics:

main: Acc = 0.7486, F1 = 0.6063, Precision = 0.6459, Recall = 0.5864

📈 Test Metrics:

main: Acc = 0.7428, F1 = 0.6092, Precision = 0.6497, Recall = 0.5883

🔁 Epoch 5/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 60.0177

📊 Validation Metrics:

main: Acc = 0.7253, F1 = 0.6073, Precision = 0.6161, Recall = 0.6046

📈 Test Metrics:

main: Acc = 0.7160, F1 = 0.5922, Precision = 0.6020, Recall = 0.5886

⏱ Total time for config: 559.20 seconds

💾 Saved overall best model to best\_model\_overall\_lr2e-05\_do0.1\_mb32\_ab8\_mw1.0\_st0\_sa0\_fi1.0.pt with F1 = 0.6273

# BASELINE (2 EXPERIMENTS)

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 0, Sarcasm weight: 0, Fine-grained weight: 0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 293.8113

✅ New best overall model at epoch 1 with main F1 = 0.5303

📊 Validation Metrics:

main: Acc = 0.7460, F1 = 0.5303, Precision = 0.7613, Recall = 0.5247

📈 Test Metrics:

main: Acc = 0.7360, F1 = 0.4936, Precision = 0.6724, Recall = 0.5021

🔁 Epoch 2/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 219.3531

✅ New best overall model at epoch 2 with main F1 = 0.5770

📊 Validation Metrics:

main: Acc = 0.7588, F1 = 0.5770, Precision = 0.7256, Recall = 0.5472

📈 Test Metrics:

main: Acc = 0.7593, F1 = 0.5700, Precision = 0.7097, Recall = 0.5436

🔁 Epoch 3/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 154.2848

✅ New best overall model at epoch 3 with main F1 = 0.6046

📊 Validation Metrics:

main: Acc = 0.7285, F1 = 0.6046, Precision = 0.6461, Recall = 0.5944

📈 Test Metrics:

main: Acc = 0.7260, F1 = 0.6144, Precision = 0.6452, Recall = 0.6063

🔁 Epoch 4/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 93.8435

✅ New best overall model at epoch 4 with main F1 = 0.6147

📊 Validation Metrics:

main: Acc = 0.7326, F1 = 0.6147, Precision = 0.6179, Recall = 0.6126

📈 Test Metrics:

main: Acc = 0.7321, F1 = 0.6054, Precision = 0.6036, Recall = 0.6074

🔁 Epoch 5/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 55.1781

📊 Validation Metrics:

main: Acc = 0.7358, F1 = 0.6088, Precision = 0.6499, Recall = 0.5911

📈 Test Metrics:

main: Acc = 0.7353, F1 = 0.6094, Precision = 0.6426, Recall = 0.5938

💾 Saved model for current config to model\_lr2e-05\_do0.1\_mb32\_ab8\_mw1.0\_st0\_sa0\_fi0\_baseline.pt

⏱️ Total time for config: 561.06 seconds

✅ All experiments completed. Results saved to grid\_search\_epochwise\_results\_baseline.csv

I have no name!@job-71dfdf9f9695-0-0:~/codes-v2$ python scripts/yorgunum-baseline.py

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 0, Sarcasm weight: 0, Fine-grained weight: 0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 296.6973

✅ New best overall model at epoch 1 with main F1 = 0.5259

📊 Validation Metrics:

main: Acc = 0.7428, F1 = 0.5259, Precision = 0.7148, Recall = 0.5093

📈 Test Metrics:

main: Acc = 0.7500, F1 = 0.5309, Precision = 0.7137, Recall = 0.5151

🔁 Epoch 2/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 221.3607

✅ New best overall model at epoch 2 with main F1 = 0.6181

📊 Validation Metrics:

main: Acc = 0.7571, F1 = 0.6181, Precision = 0.6765, Recall = 0.5909

📈 Test Metrics:

main: Acc = 0.7584, F1 = 0.5980, Precision = 0.6449, Recall = 0.5764

🔁 Epoch 3/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 160.4021

📊 Validation Metrics:

main: Acc = 0.7507, F1 = 0.6102, Precision = 0.6896, Recall = 0.5833

📈 Test Metrics:

main: Acc = 0.7416, F1 = 0.5821, Precision = 0.6431, Recall = 0.5627

🔁 Epoch 4/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 97.5464

📊 Validation Metrics:

main: Acc = 0.7477, F1 = 0.6133, Precision = 0.6815, Recall = 0.5903

📈 Test Metrics:

main: Acc = 0.7363, F1 = 0.5940, Precision = 0.6705, Recall = 0.5724

🔁 Epoch 5/5 | config: stereo=0, sarcasm=0, fine=0

Loss: 56.5221

✅ New best overall model at epoch 5 with main F1 = 0.6206

📊 Validation Metrics:

main: Acc = 0.7285, F1 = 0.6206, Precision = 0.6179, Recall = 0.6253

📈 Test Metrics:

main: Acc = 0.7123, F1 = 0.5882, Precision = 0.5816, Recall = 0.5970

^[[A💾 Saved model for current config to model\_lr2e-05\_do0.1\_mb32\_ab8\_mw1.0\_st0\_sa0\_fi0\_baseline.pt

⏱️ Total time for config: 561.34 seconds

# AUXILARY TASK WEIGHT

I have no name!@job-71dfdf9f9695-0-0:~/codes-v2$ python scripts/yorgunum.py

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 0.5, Sarcasm weight: 0.5, Fine-grained weight: 0.5

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=0.5, sarcasm=0.5, fine=0.5

Loss: 1703.5977

✅ New best overall model at epoch 1 with main F1 = 0.5948

📊 Validation Metrics:

main: Acc = 0.7483, F1 = 0.5948, Precision = 0.6754, Recall = 0.5632

stereo: Acc = 0.7082, F1 = 0.7081, Precision = 0.7261, Recall = 0.7082

sarcasm: Acc = 0.7495, F1 = 0.4354, Precision = 0.6252, Recall = 0.5024

implicit\_fine: Acc = 0.6112, F1 = 0.5276, Precision = 0.5240, Recall = 0.5450

📈 Test Metrics:

main: Acc = 0.7453, F1 = 0.5648, Precision = 0.6438, Recall = 0.5382

stereo: Acc = 0.7191, F1 = 0.7151, Precision = 0.7321, Recall = 0.7190

sarcasm: Acc = 0.7493, F1 = 0.4283, Precision = 0.3746, Recall = 0.5000

implicit\_fine: Acc = 0.6079, F1 = 0.5255, Precision = 0.5260, Recall = 0.5391

🔁 Epoch 2/5 | config: stereo=0.5, sarcasm=0.5, fine=0.5

Loss: 1136.9300

✅ New best overall model at epoch 2 with main F1 = 0.6185

📊 Validation Metrics:

main: Acc = 0.7489, F1 = 0.6185, Precision = 0.6457, Recall = 0.6013

stereo: Acc = 0.7399, F1 = 0.7423, Precision = 0.7525, Recall = 0.7399

sarcasm: Acc = 0.7586, F1 = 0.6036, Precision = 0.6632, Recall = 0.5946

implicit\_fine: Acc = 0.6378, F1 = 0.5988, Precision = 0.6056, Recall = 0.5986

📈 Test Metrics:

main: Acc = 0.7488, F1 = 0.6087, Precision = 0.6462, Recall = 0.5893

stereo: Acc = 0.7429, F1 = 0.7457, Precision = 0.7510, Recall = 0.7428

sarcasm: Acc = 0.7522, F1 = 0.6046, Precision = 0.6507, Recall = 0.5956

implicit\_fine: Acc = 0.6220, F1 = 0.5661, Precision = 0.6059, Recall = 0.5639

🔁 Epoch 3/5 | config: stereo=0.5, sarcasm=0.5, fine=0.5

Loss: 652.4374

📊 Validation Metrics:

main: Acc = 0.7370, F1 = 0.6087, Precision = 0.6065, Recall = 0.6138

stereo: Acc = 0.7369, F1 = 0.7405, Precision = 0.7489, Recall = 0.7369

sarcasm: Acc = 0.7153, F1 = 0.6227, Precision = 0.6221, Recall = 0.6233

implicit\_fine: Acc = 0.6309, F1 = 0.5569, Precision = 0.5774, Recall = 0.5615

📈 Test Metrics:

main: Acc = 0.7407, F1 = 0.6186, Precision = 0.6161, Recall = 0.6229

stereo: Acc = 0.7500, F1 = 0.7529, Precision = 0.7585, Recall = 0.7499

sarcasm: Acc = 0.6873, F1 = 0.6137, Precision = 0.6093, Recall = 0.6250

implicit\_fine: Acc = 0.6268, F1 = 0.5544, Precision = 0.5843, Recall = 0.5599

🔁 Epoch 4/5 | config: stereo=0.5, sarcasm=0.5, fine=0.5

Loss: 368.6786

📊 Validation Metrics:

main: Acc = 0.7527, F1 = 0.6154, Precision = 0.6498, Recall = 0.6000

stereo: Acc = 0.7191, F1 = 0.7216, Precision = 0.7322, Recall = 0.7191

sarcasm: Acc = 0.7550, F1 = 0.5916, Precision = 0.6544, Recall = 0.5850

implicit\_fine: Acc = 0.6093, F1 = 0.5602, Precision = 0.5678, Recall = 0.5677

📈 Test Metrics:

main: Acc = 0.7595, F1 = 0.6302, Precision = 0.6921, Recall = 0.6060

stereo: Acc = 0.7373, F1 = 0.7400, Precision = 0.7460, Recall = 0.7372

sarcasm: Acc = 0.7378, F1 = 0.5929, Precision = 0.6248, Recall = 0.5860

implicit\_fine: Acc = 0.6157, F1 = 0.5426, Precision = 0.5519, Recall = 0.5517

🔁 Epoch 5/5 | config: stereo=0.5, sarcasm=0.5, fine=0.5

Loss: 234.9710

📊 Validation Metrics:

main: Acc = 0.7507, F1 = 0.6144, Precision = 0.6615, Recall = 0.6006

stereo: Acc = 0.7369, F1 = 0.7401, Precision = 0.7478, Recall = 0.7369

sarcasm: Acc = 0.7514, F1 = 0.5690, Precision = 0.6435, Recall = 0.5683

implicit\_fine: Acc = 0.5945, F1 = 0.5345, Precision = 0.5408, Recall = 0.5350

📈 Test Metrics:

main: Acc = 0.7498, F1 = 0.6082, Precision = 0.6732, Recall = 0.5910

stereo: Acc = 0.7540, F1 = 0.7554, Precision = 0.7578, Recall = 0.7538

sarcasm: Acc = 0.7536, F1 = 0.5942, Precision = 0.6521, Recall = 0.5870

implicit\_fine: Acc = 0.6063, F1 = 0.5335, Precision = 0.5359, Recall = 0.5347

💾 Saved model for current config to model\_lr2e-05\_do0.1\_mb32\_ab8\_mw1.0\_st0.5\_sa0.5\_fi0.5.pt

⏱️ Total time for config: 1244.32 seconds

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 1.0, Sarcasm weight: 1.0, Fine-grained weight: 1.0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 1695.9587

📊 Validation Metrics:

main: Acc = 0.7341, F1 = 0.5383, Precision = 0.6400, Recall = 0.5218

stereo: Acc = 0.6904, F1 = 0.6871, Precision = 0.7332, Recall = 0.6904

sarcasm: Acc = 0.7586, F1 = 0.5003, Precision = 0.7046, Recall = 0.5324

implicit\_fine: Acc = 0.6024, F1 = 0.5116, Precision = 0.5443, Recall = 0.5284

📈 Test Metrics:

main: Acc = 0.7421, F1 = 0.5396, Precision = 0.6394, Recall = 0.5239

stereo: Acc = 0.7025, F1 = 0.6967, Precision = 0.7353, Recall = 0.7024

sarcasm: Acc = 0.7608, F1 = 0.5111, Precision = 0.7154, Recall = 0.5383

implicit\_fine: Acc = 0.6063, F1 = 0.5129, Precision = 0.5431, Recall = 0.5284

🔁 Epoch 2/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 1108.1485

📊 Validation Metrics:

main: Acc = 0.7454, F1 = 0.6179, Precision = 0.6326, Recall = 0.6064

stereo: Acc = 0.7409, F1 = 0.7446, Precision = 0.7558, Recall = 0.7409

sarcasm: Acc = 0.7243, F1 = 0.6247, Precision = 0.6278, Recall = 0.6221

implicit\_fine: Acc = 0.5994, F1 = 0.5524, Precision = 0.5793, Recall = 0.5588

📈 Test Metrics:

main: Acc = 0.7426, F1 = 0.5977, Precision = 0.6140, Recall = 0.5859

stereo: Acc = 0.7294, F1 = 0.7311, Precision = 0.7391, Recall = 0.7293

sarcasm: Acc = 0.6931, F1 = 0.5809, Precision = 0.5834, Recall = 0.5791

implicit\_fine: Acc = 0.6236, F1 = 0.5499, Precision = 0.6158, Recall = 0.5573

🔁 Epoch 3/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 668.3683

📊 Validation Metrics:

main: Acc = 0.7588, F1 = 0.6037, Precision = 0.6706, Recall = 0.5815

stereo: Acc = 0.7240, F1 = 0.7283, Precision = 0.7399, Recall = 0.7240

sarcasm: Acc = 0.7189, F1 = 0.6103, Precision = 0.6159, Recall = 0.6065

implicit\_fine: Acc = 0.6122, F1 = 0.5536, Precision = 0.5647, Recall = 0.5633

📈 Test Metrics:

main: Acc = 0.7579, F1 = 0.6084, Precision = 0.6710, Recall = 0.5857

stereo: Acc = 0.7373, F1 = 0.7418, Precision = 0.7495, Recall = 0.7373

sarcasm: Acc = 0.6974, F1 = 0.5973, Precision = 0.5973, Recall = 0.5973

implicit\_fine: Acc = 0.6386, F1 = 0.5625, Precision = 0.5802, Recall = 0.5736

🔁 Epoch 4/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 410.1065

📊 Validation Metrics:

main: Acc = 0.7617, F1 = 0.6065, Precision = 0.6875, Recall = 0.5792

stereo: Acc = 0.7161, F1 = 0.7201, Precision = 0.7328, Recall = 0.7161

sarcasm: Acc = 0.6919, F1 = 0.6053, Precision = 0.6022, Recall = 0.6100

implicit\_fine: Acc = 0.6014, F1 = 0.5456, Precision = 0.5675, Recall = 0.5331

📈 Test Metrics:

main: Acc = 0.7623, F1 = 0.6080, Precision = 0.7079, Recall = 0.5776

stereo: Acc = 0.7168, F1 = 0.7182, Precision = 0.7360, Recall = 0.7167

sarcasm: Acc = 0.6888, F1 = 0.6112, Precision = 0.6070, Recall = 0.6202

implicit\_fine: Acc = 0.6102, F1 = 0.5264, Precision = 0.5357, Recall = 0.5235

🔁 Epoch 5/5 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 274.8039

✅ New best overall model at epoch 5 with main F1 = 0.6310

📊 Validation Metrics:

main: Acc = 0.7544, F1 = 0.6310, Precision = 0.6436, Recall = 0.6234

stereo: Acc = 0.7339, F1 = 0.7401, Precision = 0.7559, Recall = 0.7339

sarcasm: Acc = 0.7387, F1 = 0.6157, Precision = 0.6349, Recall = 0.6078

implicit\_fine: Acc = 0.5935, F1 = 0.5330, Precision = 0.5655, Recall = 0.5263

📈 Test Metrics:

main: Acc = 0.7419, F1 = 0.6129, Precision = 0.6206, Recall = 0.6093

stereo: Acc = 0.7318, F1 = 0.7354, Precision = 0.7511, Recall = 0.7317

sarcasm: Acc = 0.7104, F1 = 0.5909, Precision = 0.5987, Recall = 0.5868

implicit\_fine: Acc = 0.6150, F1 = 0.5336, Precision = 0.5368, Recall = 0.5348

💾 Saved model for current config to model\_lr2e-05\_do0.1\_mb32\_ab8\_mw1.0\_st1.0\_sa1.0\_fi1.0.pt

⏱️ Total time for config: 1252.04 seconds

🚀 Starting new config:

Learning rate: 2e-05, Dropout: 0.1, Main batch size: 32, Aux batch size: 8

Main weight: 1.0, Stereo weight: 2.0, Sarcasm weight: 2.0, Fine-grained weight: 2.0

📊 Dataset sizes for MAIN task:

Train: 13747 samples

Val: 3437 samples

Test: 4296 samples

🔁 Epoch 1/5 | config: stereo=2.0, sarcasm=2.0, fine=2.0

Loss: 1724.4993

📊 Validation Metrics:

main: Acc = 0.7184, F1 = 0.4544, Precision = 0.4643, Recall = 0.4602

stereo: Acc = 0.6726, F1 = 0.6617, Precision = 0.7324, Recall = 0.6726

sarcasm: Acc = 0.7495, F1 = 0.5676, Precision = 0.6388, Recall = 0.5671

implicit\_fine: Acc = 0.6053, F1 = 0.5172, Precision = 0.5366, Recall = 0.5148

📈 Test Metrics:

main: Acc = 0.7104, F1 = 0.4451, Precision = 0.4607, Recall = 0.4510

stereo: Acc = 0.7057, F1 = 0.6921, Precision = 0.7657, Recall = 0.7056

sarcasm: Acc = 0.7478, F1 = 0.5325, Precision = 0.6286, Recall = 0.5449

implicit\_fine: Acc = 0.5921, F1 = 0.5027, Precision = 0.5327, Recall = 0.5023

🔁 Epoch 2/5 | config: stereo=2.0, sarcasm=2.0, fine=2.0

Loss: 1136.7569

📊 Validation Metrics:

main: Acc = 0.7265, F1 = 0.4829, Precision = 0.7292, Recall = 0.4761

stereo: Acc = 0.7349, F1 = 0.7384, Precision = 0.7451, Recall = 0.7349

sarcasm: Acc = 0.7532, F1 = 0.6020, Precision = 0.6518, Recall = 0.5934

implicit\_fine: Acc = 0.6033, F1 = 0.5191, Precision = 0.5167, Recall = 0.5277

📈 Test Metrics:

main: Acc = 0.7297, F1 = 0.4752, Precision = 0.6836, Recall = 0.4730

stereo: Acc = 0.7595, F1 = 0.7602, Precision = 0.7645, Recall = 0.7594

sarcasm: Acc = 0.7378, F1 = 0.5816, Precision = 0.6205, Recall = 0.5764

implicit\_fine: Acc = 0.6173, F1 = 0.5358, Precision = 0.5344, Recall = 0.5451

🔁 Epoch 3/5 | config: stereo=2.0, sarcasm=2.0, fine=2.0

Loss: 716.3513

📊 Validation Metrics:

main: Acc = 0.7454, F1 = 0.5366, Precision = 0.7487, Recall = 0.5178

stereo: Acc = 0.7161, F1 = 0.7217, Precision = 0.7339, Recall = 0.7161

sarcasm: Acc = 0.7514, F1 = 0.6166, Precision = 0.6514, Recall = 0.6066

implicit\_fine: Acc = 0.6270, F1 = 0.5589, Precision = 0.6106, Recall = 0.5584

📈 Test Metrics:

main: Acc = 0.7421, F1 = 0.5328, Precision = 0.6949, Recall = 0.5130

stereo: Acc = 0.7563, F1 = 0.7591, Precision = 0.7680, Recall = 0.7562

sarcasm: Acc = 0.7291, F1 = 0.5839, Precision = 0.6103, Recall = 0.5783

implicit\_fine: Acc = 0.6252, F1 = 0.5405, Precision = 0.5352, Recall = 0.5492

🔁 Epoch 4/5 | config: stereo=2.0, sarcasm=2.0, fine=2.0

Loss: 499.4968

📊 Validation Metrics:

main: Acc = 0.7591, F1 = 0.5993, Precision = 0.6794, Recall = 0.5743

stereo: Acc = 0.7399, F1 = 0.7429, Precision = 0.7493, Recall = 0.7399

sarcasm: Acc = 0.7459, F1 = 0.6070, Precision = 0.6408, Recall = 0.5982

implicit\_fine: Acc = 0.6063, F1 = 0.5551, Precision = 0.5963, Recall = 0.5488

📈 Test Metrics:

main: Acc = 0.7437, F1 = 0.5793, Precision = 0.6431, Recall = 0.5588

stereo: Acc = 0.7595, F1 = 0.7625, Precision = 0.7672, Recall = 0.7594

sarcasm: Acc = 0.7118, F1 = 0.5573, Precision = 0.5779, Recall = 0.5553

implicit\_fine: Acc = 0.6244, F1 = 0.5553, Precision = 0.5863, Recall = 0.5566

🔁 Epoch 5/5 | config: stereo=2.0, sarcasm=2.0, fine=2.0

Loss: 363.6996

📊 Validation Metrics:

main: Acc = 0.7626, F1 = 0.6013, Precision = 0.6836, Recall = 0.5799

stereo: Acc = 0.7260, F1 = 0.7309, Precision = 0.7395, Recall = 0.7260

sarcasm: Acc = 0.7351, F1 = 0.5903, Precision = 0.6201, Recall = 0.5838

implicit\_fine: Acc = 0.6171, F1 = 0.5298, Precision = 0.5313, Recall = 0.5319

📈 Test Metrics:

main: Acc = 0.7542, F1 = 0.6055, Precision = 0.6652, Recall = 0.5835

stereo: Acc = 0.7437, F1 = 0.7472, Precision = 0.7534, Recall = 0.7436

sarcasm: Acc = 0.7277, F1 = 0.5912, Precision = 0.6124, Recall = 0.5850

implicit\_fine: Acc = 0.6031, F1 = 0.5300, Precision = 0.5416, Recall = 0.5287

💾 Saved model for current config to model\_lr2e-05\_do0.1\_mb32\_ab8\_mw1.0\_st2.0\_sa2.0\_fi2.0.pt

⏱️ Total time for config: 1256.52 seconds

✅ All experiments completed. Results saved to grid\_search\_epochwise\_results.csv