

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

🔄 Epoch 1/2 | config: stereo=1.0, sarcasm=1.0, fine=1.0

Loss: 1613.0035

✓ New best overall model at epoch 1 with main F1 = 0.7354

📊 Validation Metrics:

main: Acc = 0.7760, F1 = 0.7354, Precision = 0.7638, Recall = 0.7238

stereo: Acc = 0.6924, F1 = 0.6790, Precision = 0.7373, Recall = 0.6924

sarcasm: Acc = 0.7640, F1 = 0.6274, Precision = 0.6750, Recall = 0.6150

implicit_fine: Acc = 0.6161, F1 = 0.5289, Precision = 0.5345, Recall = 0.5366

📈 Test Metrics:

main: Acc = 0.7791, F1 = 0.7370, Precision = 0.7703, Recall = 0.7244

stereo: Acc = 0.7041, F1 = 0.6877, Precision = 0.7519, Recall = 0.7040

sarcasm: Acc = 0.7363, F1 = 0.5851, Precision = 0.6198, Recall = 0.5793

implicit_fine: Acc = 0.6039, F1 = 0.5212, Precision = 0.5257, Recall = 0.5277

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

Loss: 1039.3074

✓ New best overall model at epoch 2 with main F1 = 0.7628

📊 Validation Metrics:

main: Acc = 0.7876, F1 = 0.7628, Precision = 0.7670, Recall = 0.7594

stereo: Acc = 0.7339, F1 = 0.7378, Precision = 0.7474, Recall = 0.7339

sarcasm: Acc = 0.7532, F1 = 0.6103, Precision = 0.6531, Recall = 0.6006

implicit_fine: Acc = 0.6299, F1 = 0.5449, Precision = 0.5450, Recall = 0.5500

📈 Test Metrics:

main: Acc = 0.7764, F1 = 0.7512, Precision = 0.7540, Recall = 0.7489

stereo: Acc = 0.7516, F1 = 0.7535, Precision = 0.7570, Recall = 0.7515

sarcasm: Acc = 0.7147, F1 = 0.5869, Precision = 0.5990, Recall = 0.5821

implicit_fine: Acc = 0.6024, F1 = 0.5220, Precision = 0.5267, Recall = 0.5251

⌚ Total time for config: 500.38 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

🔄 Epoch 1/2 | config: stereo=0, sarcasm=0, fine=1.0

Loss: 948.3626

📊 Validation Metrics:

main: Acc = 0.7702, F1 = 0.7527, Precision = 0.7485, Recall = 0.7595

implicit_fine: Acc = 0.6181, F1 = 0.5290, Precision = 0.5328, Recall = 0.5391

📈 Test Metrics:

main: Acc = 0.7680, F1 = 0.7503, Precision = 0.7462, Recall = 0.7569

implicit_fine: Acc = 0.6063, F1 = 0.5185, Precision = 0.5215, Recall = 0.5284

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

Loss: 636.4763

✓ New best overall model at epoch 2 with main F1 = 0.7636

📊 Validation Metrics:

main: Acc = 0.7922, F1 = 0.7636, Precision = 0.7749, Recall = 0.7563

implicit_fine: Acc = 0.6339, F1 = 0.5388, Precision = 0.5579, Recall = 0.5500

📈 Test Metrics:

main: Acc = 0.7943, F1 = 0.7668, Precision = 0.7767, Recall = 0.7602

implicit_fine: Acc = 0.6150, F1 = 0.5296, Precision = 0.5392, Recall = 0.5345

⌚ Total time for config: 342.04 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: 1.0, Sarcasm weight: 1.0, Fine-grained weight: 0

🔄 Epoch 1/2 | config: stereo=1.0, sarcasm=1.0, fine=0

Loss: 957.8608

📊 Validation Metrics:

main: Acc = 0.7640, F1 = 0.7260, Precision = 0.7446, Recall = 0.7171

stereo: Acc = 0.7102, F1 = 0.7168, Precision = 0.7280, Recall = 0.7102

sarcasm: Acc = 0.7495, F1 = 0.4284, Precision = 0.3748, Recall = 0.5000

📈 Test Metrics:

main: Acc = 0.7744, F1 = 0.7392, Precision = 0.7567, Recall = 0.7302

stereo: Acc = 0.7453, F1 = 0.7485, Precision = 0.7548, Recall = 0.7452

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

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

Loss: 646.1372

📊 Validation Metrics:

main: Acc = 0.7637, F1 = 0.7451, Precision = 0.7413, Recall = 0.7509

stereo: Acc = 0.7369, F1 = 0.7414, Precision = 0.7488, Recall = 0.7369

sarcasm: Acc = 0.7171, F1 = 0.6047, Precision = 0.6114, Recall = 0.6005

📈 Test Metrics:

main: Acc = 0.7643, F1 = 0.7451, Precision = 0.7416, Recall = 0.7500

stereo: Acc = 0.7555, F1 = 0.7582, Precision = 0.7627, Recall = 0.7554

sarcasm: Acc = 0.7219, F1 = 0.5983, Precision = 0.6111, Recall = 0.5926

🕒 Total time for config: 405.35 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

🔄 Epoch 1/2 | config: stereo=0, sarcasm=0, fine=0

Loss: 223.4448

 Validation Metrics:

main: Acc = 0.7726, F1 = 0.7130, Precision = 0.7859, Recall = 0.6986

 Test Metrics:

main: Acc = 0.7791, F1 = 0.7231, Precision = 0.7923, Recall = 0.7078

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


Loss: 160.6681


 Validation Metrics:

main: Acc = 0.7818, F1 = 0.7490, Precision = 0.7646, Recall = 0.7403

 Test Metrics:


main: Acc = 0.7930, F1 = 0.7634, Precision = 0.7770, Recall = 0.7551

 Total time for config: 211.52 seconds

 Saved overall best model to best_model_overall_lr2e-05_do0.1_mb32_ab8_mw1.0_st0_sa0_fi1.0.pt with F1 = 0.7636


✓ All experiments completed. Results saved to grid_search_epochwise_results.csv

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