

Overall comparison 2

Model	Dataset	LoRA / Adapter	Domain Prefix	Class Weights	Accuracy	Macro F1	Weighted F1
SFT1	MIX (BIO+FIN)	No	No	Yes	0.7884	0.6545	0.6545
SFT2	BIO (drug)	No	No	No	0.5900	0.5893	0.5893
SFT3	FIN (news)	No	No	Yes	0.7260	0.6857	0.7325
SFT4	MIX (BIO+FIN)	No	Yes	Yes	0.6450	0.6374	0.6467
SFT5	BIO (drug)	LoRA	No	No	0.5693	0.5694	0.5694
SFT6	FIN (news)	LoRA	No	Yes	0.7364	0.6974	0.7387
SFT7	MIX (BIO+FIN)	LoRA	Yes	Yes	0.6375	0.6302	0.64114

Model	LoRA	Training Setup	Eval Set	FIN Acc	FIN Macro F1	BIO Acc	BIO Macro F1
SFT2	No	BIO	BIO	–	–	0.5513	0.5521
SFT3	No	FIN	FIN	0.7260	0.6857	–	–
SFT8	No	FIN → BIO	MIX	0.3993	0.3801	0.5593	0.5599
SFT9	Yes	FIN → BIO	MIX	0.4613	0.4182	0.5400	0.5419

Zero Shot testing on Game review dataset

Model	Zero-shot Dataset	N	Accuracy	Macro F1	Weighted F1
SFT4 (No Adapter)	GAME	5000	0.5852	0.5186	0.6062
SFT7 (LoRA)	GAME	5000	0.5486	0.5013	0.5788

Per Domain

Domain	SFT4 (No Adapter) Acc / Macro F1	SFT7 (LoRA) Acc / Macro F1
FIN	0.735048 / 0.695242	0.7413 / 0.6974
BIO	0.568 / 0.56520	0.538 / 0.5321

ZEROSHOT

SFT4

=== Zero-shot on GAME (N=5000) ===

Overall accuracy: 0.5852

Overall macro F1: 0.5186

Overall weighted F1: 0.6062

SFT7

=== Zero-shot on GAME (N=5000) ===

Overall accuracy: 0.5486

Overall macro F1: 0.5013

Overall weighted F1: 0.5788

Takeaway

Strongest financial model:

- **SFT6 (FIN + LoRA)** slightly beats **SFT3 (FIN, no LoRA)**
 - Acc: 0.7364 vs 0.7260
 - Macro F1: 0.6974 vs 0.6857

BIO models:

- **SFT2 (no adapter) > SFT5 (LoRA)**
 - Acc drops from 0.5513 → 0.5340 when adding LoRA.

Mixed-domain models (MIX BIO+FIN):

- **SFT4 (no adapter) > SFT7 (LoRA)**
 - Overall acc: 0.6300 → 0.6013
 - FIN domain acc: 0.7360 → 0.7073
 - BIO domain acc: 0.5240 → 0.4953

Raw Stats

SFT2_bio_without_adapter:

Config:

```
"learning_rate": 3e-5,  
"batch_size": 16,  
"num_epochs": 3,  
"weight_decay": 0.01,
```

Dataset:

drug_review

Weighted_class:

No. Since the data has already balanced

Domain_prefix:

No

Results:

Accuracy: : 0.5513

Macro F1: : 0.5521

Weighted F1: : 0.5521

SFT3_financial_without_adapter:

Config:

"learning_rate": 3e-5,

"batch_size": 16,

"num_epochs": 3,

"weight_decay": 0.01,

Dataset:

financial_review

Weighted_class:

Yes

Domain_prefix:

No

Results:

Accuracy: 0.7260

Macro F1: 0.6857

Weighted F1: 0.7325

SFT4_mix_without_adapter:

Config:

"learning_rate": 3e-4,

"batch_size": 16,

"epochs": 3,

"use_lora": False,

"use_class_weights": True,

Dataset:

Mixed with drug_review and financial_review

Weighted_class:

Yes

Domain_prefix:

Yes

Result:

Accuracy: 0.6300

Macro F1: 0.6221

Weighted F1: 0.6345

Domain: FIN

Accuracy: 0.7360

Macro F1: 0.6897

Domain: BIO

Accuracy: 0.5240

Macro F1: 0.5231

SFT5_Bio_with_adapter:

config:

"learning_rate": 1e-4,

"batch_size": 16,

"epochs": 10,

"use_lora": True,

"use_class_weights": True,

LoRA config (you can tweak these)

"lora_r": 64,

"lora_alpha": 128,

"lora_dropout": 0.1,

Dataset:

drug_review

Weighted_class:

Yes

Result:

Accuracy: 0.5340
Macro F1: 0.5339
Weighted F1: 0.5339

SFT6_financial_with_adapter:

config:

```
"learning_rate": 1e-4,  
  "batch_size": 16,  
  "epochs": 10,
```

```
"use_lora": True,  
"use_class_weights": True,
```

LoRA config (you can tweak these)

```
"lora_r": 64,  
"lora_alpha": 128,  
"lora_dropout": 0.1,
```

Dataset:

Financial_news

Weighted_class:

Yes

Result:

Accuracy: 0.7364
Macro F1: 0.6974
Weighted F1: 0.7387

SFT7_mix_with_adapter

Config:

```
"learning_rate": 1e-4,  
"batch_size": 16,  
"epochs": 10,
```

```
"use_lora": True,  
"use_class_weights": True,
```

LoRA config

"lora_r": 64,
"lora_alpha": 128,
"lora_dropout": 0.1,

Dataset:

Mixed with drug_review and financial_review

Weighted_class:

Yes

Domain_prefix:

Yes

Accuracy: 0.6013

Macro F1: 0.5873

Weighted F1: 0.6039

Domain: FIN

Accuracy: 0.7073

Macro F1: 0.6370

Domain: BIO

Accuracy: 0.4953

Macro F1: 0.4956