National College of Ireland

Configuration Manual

MSc Research Practicum
MSc Fintech Information

Pallavi Guduguntla
Student ID: x23357673

School of Computing
National College of Ireland

Supervisor: Brian Byrne

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: ……Pallavi Guduguntla……

Student ID: ……x23357673……

Programme: ……MSc Fintech Information…… **Year:** 2024-25……..

Module: MSc Research Practicum……

Lecturer: …Brian Bryne……

Submission Due Date: 11-08-2025……

Project Title: Temporal Sentiment Shift Detection in Financial Markets: Domain Expertise vs Architectural Innovation

Word Count: ……752…… **Page Count:** ……5……

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Pallavi Guduguntla

Date: 11-08-2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Pallavi Guduguntla
x23357673

Temporal Sentiment Analysis - Quick Setup Guide

This project compares how well different AI models (FinBERT vs ModernBERT) can analyze the sentiment of financial news articles, especially during market regime changes.

Key Features:

- Downloads financial news about major tech stocks (AAPL, MSFT, etc.)
- Identifies market conditions (Bull Market, Bear Market, High Volatility, etc.)
- Analyzes sentiment (Positive, Negative, Neutral) using AI models
- Detects "temporal shifts" - periods when market sentiment changes rapidly
- Compares which model handles these shifts better

Why It Matters: Financial sentiment can change quickly during market transitions. This research tests if newer models (ModernBERT) handle these changes better than older ones (FinBERT).

1 Setup Google Colab

1. Upload your notebook **pallavi_final.ipynb** to Google Colab
2. **Runtime → Change runtime type → GPU → A100 (or T4) → Save**

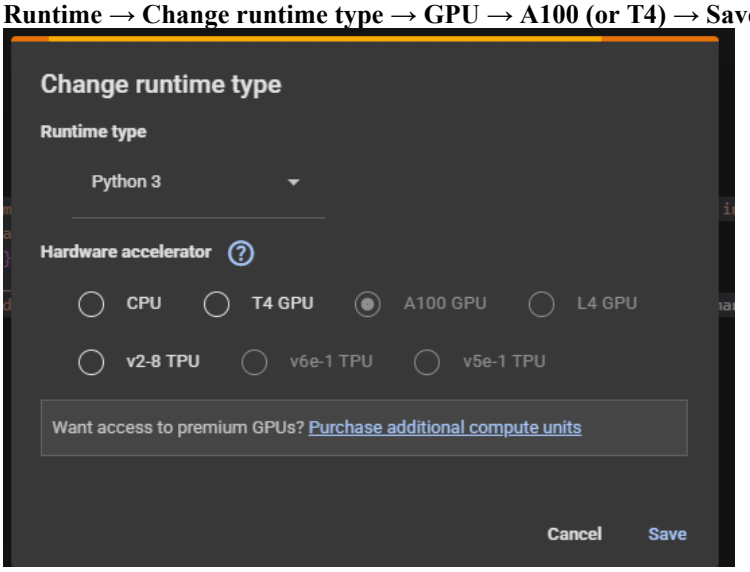


Figure 1 GPU Selection

3. Click **Connect**
4. Mount Google Drive:

```
from google.colab import drive
drive.mount('/content/drive')
```

2 Run Data Preparation (First Section)

What it does: Downloads news articles, stock prices, and market indicators. Labels each article with sentiment and market regime.

Just run all cells in order. Takes ~45 minutes.

Check it worked:

```
import os
import pandas as pd
```

 # Should print dataset info
df = pd.read_csv("/content/temporal_sentiment_analysis/temporal_sentiment_dataset_final.csv")
print(f"✓ Dataset ready: {len(df)} rows")

```
10. Preparing final dataset...
------------------------------------------------
✓ Final dataset saved to: /content/temporal_sentiment_analysis/temporal_sentiment_dataset_final.csv
✓ Model agreement analysis saved to: /content/temporal_sentiment_analysis/model_agreement_analysis.csv

================================================================================
FINAL DATASET SUMMARY
================================================================================

📊 DATASET OVERVIEW:
------------------------------------------------
Total articles: 37,509
Date range: 2009-12-15 to 2023-10-20
Stocks: AAPL, AMZN, GOOGL, META, MSFT, NFLX, NVDA, TSLA
Temporal shift articles: 9,420 (25.1%)

📊 X AND Y SUMMARY:
------------------------------------------------
Y (Target): sentiment (positive/negative/neutral)

X (Features):
  - Text: title, content
  - Temporal: market_regime, regime transitions, sentiment persistence, rolling ratios
  - Market: VIX, S&P500 returns, drawdown
  - Price: momentum, volatility, price-sentiment alignment
  - Context: financial context, signal strength

📊 SENTIMENT DISTRIBUTION:
------------------------------------------------
  neutral: 15,553 (41.5%)
  positive: 14,823 (39.5%)
  negative: 7,133 (19.0%)

📊 MODEL AGREEMENT:
------------------------------------------------

📊 SENTIMENT DISTRIBUTION:
------------------------------------------------
  neutral: 15,553 (41.5%)
  positive: 14,823 (39.5%)
  negative: 7,133 (19.0%)

📊 MODEL AGREEMENT:
------------------------------------------------
  FinBERT-RoBERTa agreement: 62.9%
  Average confidence: 0.950

📊 SENTIMENT BY MARKET REGIME:
------------------------------------------------
+-----------------+------------+-----------+------------+
|                 |   negative |   neutral |   positive |
+=================+============+===========+============+
| Bear Market     |      0.25  |     0.368 |      0.382 |
+-----------------+------------+-----------+------------+
| Bull Market     |      0.179 |     0.431 |      0.391 |
+-----------------+------------+-----------+------------+
| High Volatility |      0.215 |     0.38  |      0.405 |
+-----------------+------------+-----------+------------+
| Normal          |      0.187 |     0.413 |      0.399 |
+-----------------+------------+-----------+------------+
| Recovery        |      0.224 |     0.349 |      0.427 |
+-----------------+------------+-----------+------------+

================================================================================
✨ DATA PREPARATION COMPLETE!
Ready for FinBERT vs ModernBERT temporal shift detection experiments
================================================================================
```

3 Run Model Training (Second Section)

What it does: Trains 3 AI models (FinBERT, ModernBERT-base, ModernBERT-large) and compares their performance on sentiment analysis, especially during market regime changes.

Just run all cells in order. Takes ~6 hours.
Results will be saved to:
/content/drive/MyDrive/temporal_shift_detection/

4 Parameters You Can Change

4.1 Quick Changes (Before Running)

4.1.1 Fewer Stocks (Faster Processing)

In Data Preparation section:
'selected_stocks': ['AAPL', 'MSFT', 'GOOGL'] *# Just 3 instead of 8*

```
# ============================================
# CONFIGURATION
# ============================================
CONFIG = {
    'batch_size': 32,
    'max_length': 512,
    'output_dir': '/content/temporal_sentiment_analysis',
    'checkpoint_every': 1000,
    'selected_stocks': ['AAPL', 'GOOGL', 'AMZN', 'MSFT', 'TSLA', 'NFLX', 'META', 'NVDA'],
    'temporal_shift_window': 7,  # Days around regime change to mark as temporal shift
    'test_split_ratio': 0.2,
    'random_state': 42
}
```

4.1.2 Fewer Training Epochs (Faster Training)

In Model Training section, change all models:
'epochs': 3 *# Instead of 10*

```
# ============================================
# CONFIGURATION
# ============================================
CONFIG = {
    'batch_size': 32,
    'max_length': 512,
    'output_dir': '/content/temporal_sentiment_analysis',
    'checkpoint_every': 1000,
    'selected_stocks': ['AAPL', 'GOOGL', 'AMZN', 'MSFT', 'TSLA', 'NFLX', 'META', 'NVDA'],
    'temporal_shift_window': 7,  # Days around regime change to mark as temporal shift
    'test_split_ratio': 0.2,
    'random_state': 42
}
```

4.1.3 Skip ModernBERT-Large (Save 3 Hours)

In Model Training section, comment out:
python
MODELS = {
'finbert': {...},
'modernbert_base': {...},
 # 'modernbert_large': {...} # Comment this out
}

```
# Model configurations optimized for A100
MODELS = {
    'finbert': {
        'model_name': 'yiyanghkust/finbert-tone',
        'max_length': 512,
        'batch_size': 64,      # Increased for A100
        'eval_batch_size': 128, # Larger eval batch
        'learning_rate': 2e-5,
        'epochs': 10,
        'gradient_accumulation_steps': 1  # No accumulation needed with larger batch
    },
    'modernbert_base': {
        'model_name': 'answerdotai/ModernBERT-base',
        'max_length': 2048,
        'batch_size': 16,       # Increased for A100
        'eval_batch_size': 32,  # Larger eval batch
        'learning_rate': 1e-5,
        'epochs': 10,
        'gradient_accumulation_steps': 2
    },
    'modernbert_large': {
        'model_name': 'answerdotai/ModernBERT-large',
        'max_length': 2048,
        'batch_size': 8,        # Conservative for large model
        'eval_batch_size': 16,
        'learning_rate': 1e-5,
        'epochs': 10,
        'gradient_accumulation_steps': 4
    }
```

4.1.4 Check Results

What you'll see:

- Performance comparison between models
- Which model handles temporal shifts better
- Accuracy across different market conditions

After training completes:

 # View final report
with open("/content/drive/MyDrive/temporal_shift_detection/results/final_report.txt", 'r') as f:
  print(f.read())

 # View comparison plot
from IPython.display import Image
Image("/content/drive/MyDrive/temporal_shift_detection/results/model_comparison.png")

5 Common Issues

5.1 Out of Memory Error:

- Restart runtime: **Runtime → Restart runtime**
- Clear GPU:

python
import torch
torch.cuda.empty_cache()

5.2 Can't Find Dataset:

- Make sure you ran Data Preparation first
- Check the file exists:

python
!ls /content/temporal_sentiment_analysis/

5.3 Training Too Slow:

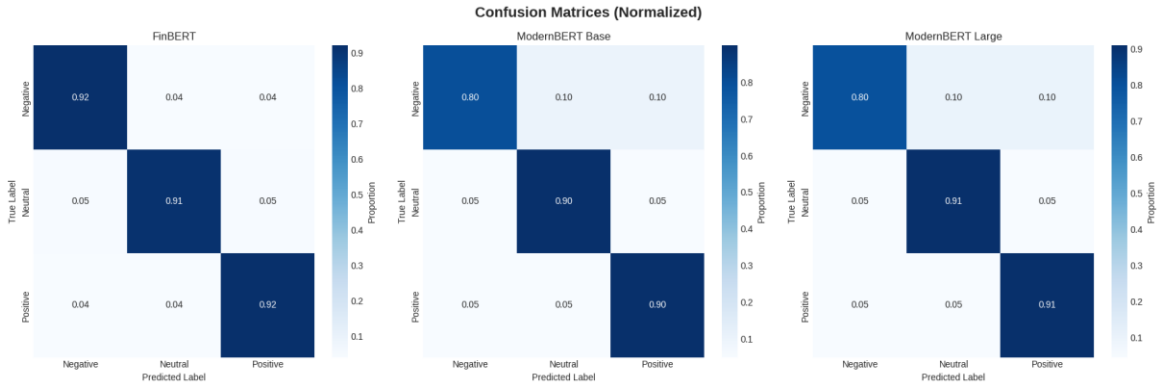
- Use fewer epochs (3 instead of 10)
- Train only FinBERT (skip ModernBERT)
- Make sure you have GPU (not CPU)

5.4 That's It!

1. Run Data Prep → 45 minutes
2. Run Model Training → 6 hours
3. Check results in Google Drive

All outputs are automatically saved to your Google Drive.
Expected Outcome: You'll have a complete comparison showing which model (FinBERT or ModernBERT) performs better at financial sentiment analysis, especially during volatile market periods.

6 Results



REFERENCES
Google Colab. (n.d.). *Colab.google*. colab.google. https://colab.google/