WEEK 3 PLAN (MON → SUN)

🎯 Week 3 Goal

Lock a **research-grade baseline** (TF-IDF + Logistic Regression) with proper evaluation outputs saved in `/results/`, plus small error analysis and 1–2 controlled baseline variants. Start the deep learning direction lightly (setup/planning or notebook skeleton).

🟢 MONDAY — Baseline audit + standardise (DONE)

Time: 1–1.5h
✅ Baseline runs cleanly
✅ Saved artefacts moved to `/results/`
✅ Commit done

Tiny checklist to patch later

- Add `random_state=42` to Logistic Regression
- Add a short markdown cell at top (“What this notebook does”)

🟢 TUESDAY — Evaluation completeness (assessor-grade) DONE

Time: 2–3h
Coding

- Ensure baseline evaluation is complete and saved:
    - Accuracy, macro-F1, weighted-F1
    - confusion matrix (already saved CSV)
- (Optional but good) Save confusion matrix as **PNG** for the report later.

Deliverables

- `/results/baseline_metrics.json`
- `/results/confusion_matrix.csv`
- `/results/valid_predictions.csv`
- (Optional) `/results/confusion_matrix.png`

✅ End-of-day feeling: “I have real results I can cite.”

🟢 WEDNESDAY — Error analysis (small but impressive) DONE

Time: 1.5–2.5h
Coding / Analysis

- Identify most common confusion pairs (e.g., half-true ↔ mostly-true)
- Extract 10 misclassified examples with short notes (why hard)
- Save a table for later Chapter 4 discussion

Deliverables

- `/results/error_analysis_examples.csv`
- 5 bullet points summarising failure patterns (for assessor meeting)

✅ End-of-day feeling: “I can explain what’s going wrong and why.”

🟢 THURSDAY — Controlled baseline variants (still classical) DONE

Time: 2–3h
Coding
Run *controlled* variations (keep it simple):

- TF-IDF unigrams vs TF-IDF bigrams (1,2)
- (Optional) stopwords toggle

Record the same metrics for each.

Deliverables

- `/results/baseline_variants.csv`
- One short conclusion: “Variant X improved macro-F1 by … / didn’t help”

✅ End-of-day feeling: “I’m improving things systematically, not randomly.”

🟢 FRIDAY — Write Week 3 into Chapter 4 (short + factual) + commit DONE

Time: 1.5–2h
Writing
Update Chapter 4 “Preliminary Baseline Results”:

- Insert your actual metrics (accuracy + macro-F1)
- Mention 1 confusion matrix observation
- Mention 1–2 error analysis insights
- One sentence: “This motivates deep learning”

Repo

- Restart+Run all on notebooks
- Commit + push

✅ End-of-week feeling: “My baseline story is complete and defensible.”

🟡 SATURDAY — Deep learning direction setup (light but real) DONE

Time: 2–3h
Goal: satisfy Ammar: you've started deep learning direction

Tasks

- Create a new notebook:
    - `03_deep_learning_baseline.ipynb`
- Install/check dependencies if needed (transformers, torch)
- Load LIAR text-only pipeline (statement + label)
- Tokenise with a pretrained model tokenizer (DistilBERT recommended)
- Run a tiny "smoke test" training setup (even 1 epoch / small subset is fine)

Deliverables

- Notebook exists and runs at least to tokenisation + dataloaders
- (Best case) you produce preliminary validation macro-F1 (even if low)

✅ End-of-day feeling: "Deep learning direction is real, not just talk."

🟡 SUNDAY — Plan Week 4 + prepare talking points DONE

Time: 45–90 min
Tasks
Write 8 "assessor meeting" bullets:

- what baseline is + why
- what metrics you used + why macro-F1
- your baseline result (numbers)
- confusion pairs
- what errors look like

- what controlled variant did
- what deep learning approach you're starting (DistilBERT)
- what you'll compare and how

✅ End-of-day feeling: "I can explain my work confidently."
