

Sentiment Analysis Repository Refactor Report

Summary

- Reorganized source code into a standard src/ layout under src/sentiment_analysis.
- Created a top-level scripts/ directory for runnable entry points and utilities.
- Updated imports across the codebase to use the new package paths.
- Added `__init__.py` files to make packages explicit and reliable.

Detailed Changes and Rationale

1) Introduced a src/ package layout

- Moved core modules (`pretrained_embeddings.py`, `feedforward_neural_network.py`) into `src/sentiment_analysis/embeddings` and `src/sentiment_analysis/models`.
- Moved the existing functions/ tree into `src/sentiment_analysis/functions`.
- Rationale: The src/ layout keeps import paths clean and prevents accidental shadowing of packages when running code from the repository root.

2) Standardized entry points

- Relocated execution scripts (`main.py`, `main_pretrained.py`, `landing.py`, `check_vectorized_comments.py`, `test.py`) into a top-level `scripts/` directory.
- Rationale: Separating runnable scripts from importable modules clarifies responsibilities and keeps the package focused on reusable code.

3) Updated imports for the new structure

- Repointed script and module imports to `sentiment_analysis.*` so the refactored paths resolve consistently.
- Updated intra-package references (e.g., `glove_vector_training.py` and `keras_feedforward_neural_network.py`) to use the new package paths.
- Rationale: Keeps internal references aligned with the new layout and avoids ambiguous relative imports.

4) Package initialization

- Added missing `__init__.py` files throughout the moved package tree.
- Rationale: Makes package boundaries explicit, improving IDE support and import reliability.

Notes

- Data files and figures remain in their original locations; only code organization and imports were updated.