

Full ML/NLP Deployment Demo

Description:

Deployment of the NLP baseline as a FastAPI REST API service, demonstrating system integration and applied engineering skills.

Methodology:

FastAPI REST API serving pre-trained model via joblib

Endpoints: /predict and /health

Minimal unit tests included for reproducibility

Supports JSON-based predictions for integration with other applications

Outcomes:

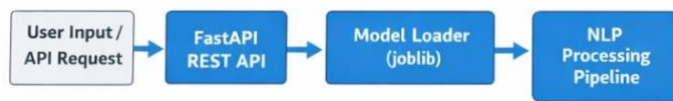
Highlights deployment, integration, and applied engineering

Complements NLP baseline for MSc-level research + application demonstration

Architecture Diagram:

User Text Input → FastAPI → Model Loader → ML Pipeline → API Response

ML/NLP Deployment Architecture



Literature / Research Context:

Demonstrates practical implementation of reproducible ML pipelines

Shows applied engineering skills aligned with intelligent systems research

Future Work:

Extend to interactive dashboards

Integrate with neural or transformer-based models

Automated testing of endpoints and performance metrics