Intelligent Systems

Assignment 2: Detecting Contradictions in Law Documents

Deadline: January 11, 2026

1 Introduction

In this seminar assignment, your goal is to develop an NLP system capable of identifying **contradictions in legal text**. You will be provided with two datasets containing law sections and annotated contradictions:

- A **Slovene dataset**, containing law sections and a list of contradicting sections.
- An **English dataset**, formatted as pairs of text segments (premise, hypothesis) labeled as `Entailment`, `Contradiction`, or `NotMentioned`.

You may choose to work with either dataset or combine both to explore multilingual approaches. The ultimate goal is to create a system that could be used by legal professionals: when drafting a new section of a law or contract, the system should automatically identify potentially **contradicting existing sections**.

2 Task 1 - Data preparation and exploration (10%)

Load the provided JSONL datasets, transform them into a format suitable for use with your models. For the Slovene dataset, this may involve creating pairs of sections with binary labels (contradicting / not contradicting). For the English dataset, use the provided labels directly or merge classes into binary categories if appropriate.

Perform basic exploratory analysis. Some examples of questions you might want to answer are:

- Are there missing values or duplicate entries?
- How many examples are in the chosen dataset?
- How long are typical sections?
- Are contradictions balanced across the datasets?

Split the data into `train`, `validation`, and `test` sets. Prepare tokenized text representations suitable for traditional machine learning models.

Additionally, include a short **data presentation and visualization section**. Visualize the data distribution, text lengths, and example contradictions using tables and graphs.

3 Task 2 - Basic Machine Learning (20%)

Train and evaluate several **traditional machine learning classifiers** on your formatted data. You may use models such as:

- Logistic Regression
- Decision Trees or Random Forests
- Support Vector Machines (SVM)

Use appropriate text representations (e.g., TF-IDF, Bag-of-Words) and apply **hyperparameter tuning and cross-validation** to improve model performance. Evaluate your models using appropriate metrics (accuracy, F1, precision, recall, etc.), and use the test set only for final evaluation.

4 Task 3 - Transformer-based Classifier (30%)

In this part, move from traditional ML to **deep learning approaches**. Your goal is to fine-tune a **transformer-based model** (e.g., BERT, SloBERTa, or multilingual models such as CroSloEngual BERT) to classify pairs of legal text as contradicting or not.

Steps to include:

- Select a suitable pre-trained transformer model.
- Prepare sentence-pair inputs for contradiction classification.
- Fine-tune the model using your training data.
- Evaluate the model on the test set and compare with traditional approaches.

Experiment with model choice, fine-tuning parameters, and text preprocessing strategies. Discuss how well the transformer captures legal contradictions compared to simpler models.

5 Task 4 - Sentence Embedding Model for Contradiction Search (30%)

In the final task, you will train a **custom sentence transformer model** to generate embeddings for legal text sections. The goal is to obtain embeddings such that sections that **contradict each other** have **high similarity scores**.

- Train or fine-tune a sentence transformer (e.g., using `sentence-transformers` library).
- Evaluate whether embeddings correctly capture contradiction relationships.
- Implement a retrieval approach: given a new law section, retrieve the top k most similar (potentially contradicting) sections from the dataset.
- Optionally, apply the classifier from Task 3 as a post-filter to select true contradictions from candidates.

Hint: You can learn more about training sentence transformers in this tutorial: `https://huggingface.co/blog/how-to-train-sentence-transformers`

6 Task 5 - Results and Report (10%)

Prepare a comprehensive report summarizing your approach, results, and conclusions. Include:

- Tables and graphs comparing model performance across all tasks.
- Visualizations showing the effect of different models or embedding strategies.
- Qualitative examples of correctly and incorrectly classified contradictions.
- Discussion of which approaches worked best and why.

Ensure that graphs are readable, clearly labeled, and accompanied by short interpretive text.

7 Submission

- **Deadline:** January 11, 2026
- **Format:** Jupyter Notebook
- **Group Work:** Maximum of two students per group
- **Use of Generative AI:** You may use ChatGPT or similar tools to help initiate your solution. Please attach the relevant conversation as a single `.txt` file alongside your submission.