Assignment 2

The goal of this assignment is to fine-tune a model for token classification on span-level annotated data. Through this exercise, you will become more familiar with subtask 2 (span-level detection of narrative elements) of the shared task on "Oppositional thinking analysis: Conspiracy vs critical narratives":
https://pan.webis.de/clef24/pan24-web/oppositional-thinking-analysis.html

Slides on token classification:
https://docs.google.com/presentation/d/1i5mx5bf1hyPzxUNo4uaL7dssUwq821YaL1gOCxR0HAs/edit?usp=sharing

Instructions

1. Setting up a `conda` environment is highly recommended. You can reuse the environment from the previous assignment.
2. Add the following libraries: `datasets`, `transformers` and `seqeval`.
3. We will work only on English data. Make sure you have a spaCy model for English, which you can obtain by running the following:

   ```
   python -m spacy download en_core_web_sm
   ```

4. Download the dataset (`dataset_oppositional`) and the two python scripts (`exercise_2.py` and `utils.py`) and put them in the same directory. The directory should look like this:

   ```
   ├── dataset_oppositional/
   │   ├── dataset_en_train.json
   │   └── dataset_es_train.json
   ├── exercise_2.py
   └── utils.py
   ```

5. Have a look at both scripts: `exercise_2.py` and `utils.py`. There are several missing bits of code in both files: search for the keyword `TODO` to find them. The exercises are described in the section below and in comments in the code.
6. You will only need to submit the `utils.py` file (for exercises 1, 2, 3 and 5) and a word/pdf/txt file with a small report for exercise 4. The deadline for the exercises is March 21.

Notes:
- If you prefer to work on notebooks, feel free to copy the content of `exercise_2.py` into a notebook, and run from there. You don't have to submit `exercise_2.py`.
- In `exercise_2.py`, there are some print statements that show the expected output of a function given an input.

- In this assignment, we will be working only with a subset of the dataset and of the labels. In particular, we will use only the "CONSPIRACY" instances and the following four categories: 'CAMPAIGNER', 'VICTIM', 'AGENT' and 'FACILITATOR'.
- In case of doubts, you can reach me at mcoll@prhlt.upv.es.

Exercises

- **Exercises 1, 2 and 3: Data preparation** (1.25, 1.25 and 2.5 points each)
  - In utils.py, search for "TODO Exercise 1|2|3" and follow the instructions in the comments.
- **Exercise 4: Fine-tuning and evaluating the model** (2.5 points)
  - In exercise_2.py, search for "TODO Exercise 4": you will need to fine-tune DistilBERT for token-classification, following the notebook tutorial: https://github.com/huggingface/notebooks/blob/main/examples/token_classific ation.ipynb (more instructions in the comment). You will **not** have to submit the code for fine-tuning and evaluating a model; instead, you will have to submit a small report (word/pdf/txt format) in which you include:
    - A description of the resulting fine-tuned model (hyperparameters used, etc) and of the results of applying the classifier on the test set (both the overall results and the results per category). Provide also a brief interpretation of the results (in English).
    - Add a paragraph on how the performance could be improved. In particular, suggest one idea you are quite confident about its success in improving the performance, and one idea you would like to try just out of curiosity.
- **Exercise 5: Apply the model to the test set** (2.5 points)
  - In utils.py, search for "TODO Exercise 5" and follow the instructions.

Bibliography

- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." In Proceedings of the 2019 Conference of the NAACL: HLT, Volume 1 (Long and Short Papers), pp. 4171-4186. 2019.
- D. Jurafsky, J. H. Martin. Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 3rd edition, 2023: https://web.stanford.edu/~jurafsky/slp3/. Chapter 8.
- Sanh, Victor, Lysandre Debut, Julien Chaumond, and Thomas Wolf. "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter." arXiv preprint arXiv:1910.01108 (2019).
- Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, et al.. 2020. Transformers: State-of-the-Art Natural Language Processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pages 38–45.