

Final Projects of the NLU Course

Mahed Mousavi, Gabriel Roccabruna, Dr. Evgeny Stepanov
Prof. Giuseppe Riccardi

Signals and Interactive Systems Lab

Department of Information and Communication Technology

University of Trento

What is this about and Objectives

- Course project & report concur 80% of the final exam grade.
- We provide default projects in this presentation
- They are interesting and relevant topics
- We assign to all of you a topic randomly
- ~~● You can still exchange topics amongst yourselves till May 25.~~

Plan for the next weeks

- We will have Project review and Q&A in the last TWO classes (May 31 and June 6) or post your project clarification question on piazza.com so everybody benefits
- **IMPORTANT**: Use the "proj" label and the project ID when you ask a question about that (eg. [proj DP4]) in the subject of the post
- ⇒ This way you can browse previous Q&A and
- We can give you to-the-point answers/suggestions

Instructions

- Projects **must be done individually** (only 1 student)
- The delivery due to at least 7 days prior the exam date.
- The delivery must include:
 - The code, the data used and the structure
 - A report of max 4-pages (+1 page for references)
 - We will provide you the report template

Main Index

- Projects
- Report guidelines

Three Topics for the Exam Project

- **Language Modeling** [LM]
 - LM with RNN
 - Fine-Tune a pre-trained LM
- **NLU task : “Intent & Slots”** [NLU]
 - Join Intent Classification and Slot Filling
- **Sentiment Analysis** [SA]
 - Subjectivity and Polarity
 - Aspect-based Sentiment Analysis

Note:

Given a Topic, which ids are [LM], [NLU] and [SA], you can pick any proposed project in that Topic.

Language Modeling [LM]

Computing $P(W)$ and Predicting $P(w_i | h(w_i))$

$P(\text{lizard} | \text{the, other, day, I, was, walking, along, and, I, saw, a})$

Evaluating the language models

Language Modeling with RNN

- **Description:** Implementation of a Language Model using one of the RNN architectures (eg. Vanilla, LSTM, GRU)
- **Dataset:** Penn Tree bank ([download](#))
- **Baseline:** Vanilla: 140 PPL, LSTM: 90.7 PPL
- **References:**
 - MIKOLOV, Tomáš, et al. Extensions of recurrent neural network language model. In: 2011 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2011. p. 5528-5531.
 - Merity, S., Keskar, N. S., & Socher, R. (2017). Regularizing and optimizing LSTM language models. arXiv preprint arXiv:1708.02182.
- **Hints:**
 - Try some regularization techniques (eg. dropout or [Merity et al.](#))
 - Make sure you compute the PPL in the right way.
 - Batch size, pad tokens etc.

Fine-tune a pre-trained LM

- **Description:** Evaluate and fine-tune the GPT2-(small|medium) model, on datasets from at least three different domains (e.g. financial, e-commerce, news)
- **Dataset:** To collect and to present (origin, content, statistics and split)
- **Baseline:** It depends on the dataset (fine-tuning PPL < 40)
- **References:**
 - **GPT2:** https://huggingface.co/transformers/model_doc/gpt2.html
 - *RADFORD, Alec, et al. Language models are unsupervised multitask learners. OpenAI blog, 2019, 1.8: 9.*
- **Hints:**
 - Compare your results, and understand and explain why your model works well or not.
 - PPL in GPT2 <https://huggingface.co/docs/transformers/perplexity>

NLU - Sentence Level [NLU]

— Intent classification

Intent classification is a text classification task in which the objective is to assign an intent for a given sentence or utterance.

Utterance: Can you help me find out about flights ?

Intent: InfoRequest

What is the intention of the user ?

Concept tagging (or slot filling)

Slot filling is a sequence labelling task where the objective is to map a given sentence or utterance to a sequence of domain-slot labels.

What are the properties of the request ?

Utterance: I want to travel from nashville to tacoma

Concepts: O O O O O B-fromloc.city_name O B-to loc.city_name

Intent Classification and Slot Filling

- **Description:** Implement a neural network that predicts intents and slots in a multitask learning setting.
 - *Since the datasets are tiny, you have to train and test your model from scratch at least 5 times. Report average and standard deviation.*
- **Dataset:** ATIS and SNIPS
- **Baseline:** (you have to improve this of at least 2-3 %)
 - ATIS -> Slot F1: 92.0%, Intent Acc.: 94.0%;
 - SNIPS -> Slot F1: 80.0%, Intent Acc.: 96.0%
- **References:**
 - *Multitask learning:*
 - Interesting reading: <https://runder.io/multi-task/>
 - Han, S.C., Long, S., Li, H., Weld, H., & Poon, J. (2021). Bi-directional Joint Neural Networks for Intent Classification and Slot Filling. Interspeech.
 - Chen, Q., Zhuo, Z., & Wang, W. (2019). Bert for joint intent classification and slot filling. arXiv preprint arXiv:1902.10909.
- **Hints:**
 - Experiment with different architectures (Seq2Seq, Bi-LSTM + CRF, etc.)
 - Use conll.py script to evaluate your results
 - Try to use pre-trained models (eg. BERT, GPT2, T5)

Sentiment Analysis [SA]

Sentiment Analysis is a natural language processing task that automatically extracts writers' orientation/attitude off given text.

The objective of the project is to automatically assign a polarity, which could be either **positive**, **negative** or **neutral**, to a given text.

Examples:

[Movie Review]

The first Star Wars movie is terrific! - **Positive**

[Financial Doc]

Tokyo stock exchange closes in deep red - **Negative**

[Product Review]

This keyboard isn't the best, but it gets the job done - **Neutral**

Subjectivity & Polarity

- **Description:** Implement Sentiment Analysis that consists of (1) Subjectivity Detection and (2) Polarity Classification.
- **Dataset:**
 - MovieReviews: NLTK (movie_reviews),
 - Subjectivity Dataset: NLTK (subjectivity)
- **Baseline:** NB: 84 ACC
- **References:**
 - *A sentimental education: Sentiment analysis using subjectivity, Bo Pang and Lillian Lee, Proceedings of ACL, pp. 271--278, 2004*
- **Hints:**
 - Make sure you make correct use of training data for polarity classification

Aspect-based Sentiment Analysis

- **Description:** Implement target/aspect identification and polarity classification on SemEval 2014 (laptop).
- **Dataset:** [SemEval 2014](#)
- **Baseline:** RNN: 57.9 F1
- **References:**
 - Hu, M., Peng, Y., Huang, Z., Li, D., & Lv, Y. (2019, July). Open-Domain Targeted Sentiment Analysis via Span-Based Extraction and Classification. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 537-546).

Report guidelines

- **Description:**
 - The project report must follow the template that is linked in this slide
 - The maximum length of the report is **4 pages** (+1 for references)
 - As a reference, we propose you two examples of excellent reports
- **Template:** Final NLU project template (zip file here)
- **Example of excellent reports:**
 - Sample 1
 - Sample 2

Still questions?

- Get on **piazza.com** platform.
- Use the "proj" label and the project ID when you ask a question about that (eg. [proj DP4]) in the subject of the post