

01VIXSM

Deep Natural Language Processing

Prof. Luca Cagliero, Giuseppe Gallipoli, Ali
Yassine

Dipartimento di Automatica e Informatica
Politecnico di Torino



Exam rules

- The exam consists of
 - a written test (max. 22 points)
 - Covering theoretical aspects introduced during the course
 - Closed and/or open-ended questions
 - A group project (max. 10 points)
 - A team project is assigned during the course
 - The project report must be submitted through the Didactic Portal (section “Elaborati”)
 - The evaluation comprises an oral discussion
 - The final score is given by the sum of the points achieved in the written part and in the evaluation of the final report

Exam rules

- The written test is scheduled during the official exam session
 - All marks will be recorded by the end of the session
 - Exam bookings are mandatory
 - The exam is closed-books

Exam rules

- The project score is valid until the end of the academic year (until September 2026)
- When you pass the written exam we automatically record the final grade by summing the achieved project score (if any)
- If you want to improve the mark of the written test you can reject it and redo the test at the next available exam
- If you want to improve the score of the project you have to redo it from scratch
 - No supplementary oral part is allowed

Exam rules

- To get the project score, the project must be delivered **1 week before** the end of the current examination session
 - Please send us a email to notify your submission
 - We evaluate your submission and set up the oral discussion **by the end of the session**
- You are allowed to deliver the project in advance
 - We may anticipate the oral discussion if we receive multiple early submissions
- All grades of the written part (plus eventually the project score) MUST be recorded within the official exam session

Project rules: example nr. 1

- I plan to give the written exam in the winter examination session (January-February 2026)

EXAM SESSION A.Y. 2025/2026

Date	Event
Monday, 12/01/2026 to Saturday, 21/02/2026	WINTER EXAMINATION SESSION – a.y. 2025/26
Monday, 08/06/2026 to Saturday, 18/07/2026	SUMMER EXAMINATION SESSION - a.y. 2025/26
Monday, 31/08/2026 to Saturday, 19/09/2026	AUTUMN EXAMINATION SESSION – a.y. 2025/26

To get the extra points my team has to deliver the project at least one week before the end of the examination session, i.e., by 14 Feb 2026

Project rules: example nr. 2

- I plan to give the written exam in the summer examination session (June-July 2026)



To get the extra points my team has to deliver the project at least one week before the end of the examination session, i.e., by 11 July 2026

Consulting on final project

- Scheduled dates
 - 26/11/2025 from 2.30pm to 5.30pm Room 3i
 - 09/12/2025 from 1pm to 2.30pm Room 3i
 - 17/12/2025 from 3pm to 5.30pm Room 3i
 - 07/01/2026 from 3pm to 5.30pm Room 3i
- Goal
 - Understanding of the assigned paper
 - Suggestions/recommendation on data retrieval and preparation, methodology, experiments, validation, method extensions
 - No code debugging

Project rules

- 10 points
 - Report: maximum **2 points**
 - Clarity of the problem statement and methodology
 - Experiments
 - Analysis of the results
 - Conclusions/takeaways
 - Reproducibility: maximum **2 points**
 - Source code
 - GitHub repository setup
 - Reproducible code
 - Live demo (if applicable)
 - Oral discussion (mandatory): maximum **1 point**
 - Extensions of the current solution: maximum **5 points**
 - More details in the next slides

Report

- Report in Latex
 - English
 - Mandatory
 - Overleaf template:
 - <https://it.overleaf.com/latex/templates/ieee-conference-template/grfzhnncsfqn>
 - Latex guide
 - https://it.overleaf.com/learn/latex/Learn_LaTeX_in_30_minutes
 - Maximum 4 pages
 - Plus eventual bibliography, appendices
 - Include an explicit link to the project repository

Report

- Abstract
 - Around 250 words
 - Highlights on
 - Context
 - Methodology
 - Proposed extensions
 - Results

Report

- Problem statement

- Theoretical formalization of the problem you have addressed
 - Expected input
 - Addressed task
 - Expected output

Report

- Methodology
 - Overview of the NLP pipeline/architecture
 - Description of each module
 - Pseudocode
 - If need be

Report

- Experiments
 - Data description
 - Experimental design
 - Hardware
 - Software/libraries
 - Validation method
 - Performance metrics
 - Execution times
 - Results
 - Tables, plots, etc.
 - Analysis of the results

Conclusions and takeaways

- Remarkable outcomes
- Main challenges
- Lesson learnt

Extensions

- An extension is an additional exploration/analysis/study that does not appear in the original paper
- Examples of extensions
 - Ablation study
 - Model exploration
 - Data enrichment (with experimental results)
 - Domain adaptation
 - Apply the same model to other domains
 - Multilingual extension
- Each extension: max 2.5 points
 - We will consider the level of complexity and completeness of the analysis
- It would be possible to discuss about other types or ideas for extensions with the teachers

Oral discussion

- Group discussion
 - 15 minutes
- Send an email to schedule an appointment
 - After the project report submission
- 10-minute presentation
 - Each team member has to present part of the work
- Q&A
 - To **all** team members

Project topics' preferences

Topic ID	Topic name	Preferred Count	Alternative Count
T1	Question Answering & Machine Translation	14	14
T2	Text Summarization	6	9
T3	Transformers for Time Series Forecasting	10	5
T4	Figurative Language Understanding	2	5
T5	Temporal Reasoning	2	1

Project assignments

- Number of teams per topics
 - T1 Question Answering and Machine Translation: 8
 - T2 Text Summarization: 7
 - T3 Transformers for Time Series Forecasting: 7
 - T4 Figurative Language Understanding: 7
 - T5 Temporal Reasoning: 5

Project assignments

- Link to the official assignments
 - DeepNLP 2025-2026 - Team Projects Definitive.pdf on the Didactic Portal (folder “Project”)
- Teams must be composed of 4 or 5 persons
- Some teams are nameless
 - If you would like to specify a team name, please send an email to Giuseppe Gallipoli.
- Missing teams
 - If a team or individual student(s) did not sign up in time, send an email to Luca Cagliero **with Giuseppe Gallipoli and Ali Yassine in CC.**
 - You will be assigned a topic at the discretion of the course staff to ensure a balanced distribution of projects across teams.

Project assignments

ASKQE: Question Answering as Automatic Evaluation for Machine Translation

 : Question Answering & Machine Translation

 : <https://aclanthology.org/2025.findings-acl.899.pdf>

 : <https://github.com/dayeonki/askqe>

 : Ali Yassine (ali_yassine@polito.it)

Project assignments

Salient Information Prompting to Steer Content in Prompt-based Abstractive Summarization

 : Text Summarization

 : <https://aclanthology.org/2024.emnlp-industry.4.pdf>

 : <https://github.com/amazon-science/SigExt>

 : Giuseppe Gallipoli (giuseppe.gallipoli@polito.it)

Project assignments

Chronos-2: From Univariate to Universal Forecasting

 : Transformers for Time Series Forecasting

 : <https://arxiv.org/abs/2510.15821v1>

 : <https://huggingface.co/amazon/chronos-2>

 : Ali Yassine (ali_yassine@polito.it)

Project assignments

BESSTIE: A Benchmark for Sentiment and Sarcasm Classification for Varieties of English

 : Figurative Language Understanding

 : <https://aclanthology.org/2025.findings-acl.441.pdf>

 : <https://huggingface.co/datasets/unswnlporg/BESSTIE>

 : Giuseppe Gallipoli (giuseppe.gallipoli@polito.it)

Project assignments

Learning to Reason Over Time: Timeline Self-Reflection for Improved Temporal Reasoning in Language Models

 : Temporal Reasoning

 : <https://aclanthology.org/2025.acl-long.1358.pdf>

 : <https://github.com/amazon-science/TISER>

 : Ali Yassine (ali_yassine@polito.it) and Giuseppe Gallipoli (giuseppe.gallipoli@polito.it)

Acknowledgements and copyright license

- Copyright licence
 - Attribution + Noncommercial + NoDerivatives
- Acknowledgements
 - I would like to thank Dr. Giuseppe Gallipoli and Dr. Ali Yassine, who collaborated to the writing and revision of the teaching content.
- Affiliation
 - The author and his staff are currently members of the Database and Data Mining Group at Dipartimento di Automatica e Informatica (Politecnico di Torino) and of the SmartData interdepartmental centre
 - <https://dbdmg.polito.it>
 - <https://smartdata.polito.it>



Thank you!