

Bridging Applied Experience and Research Contexts in Ukrainian NLP Education

Yurii Paniv*

Ukrainian Catholic University
str. Kozelnytska 2
Lviv, Ukraine
paniv@ucu.edu.ua

Viktoriia Makovska*

Ukrainian Catholic University
str. Kozelnytska 2
Lviv, Ukraine
makovska.pn@ucu.edu.ua

Abstract

We present an open, bachelor-level Natural Language Processing (NLP) course developed at Ukrainian Catholic University and delivered in Ukrainian. The course addresses several challenges in NLP education: adapting predominantly English-centric materials to a different linguistic and cultural context, supporting students with heterogeneous technical backgrounds, and balancing foundational theory with industry-relevant skills. All course materials, including lecture slides, notebooks, video recordings, and assignments, are publicly available. We describe our pedagogical design choices, focusing on culturally adapted tasks, integrated ethics, project-based assessment, and continuous student feedback. Our experience demonstrates that it is feasible to build a comprehensive and modern NLP curriculum from scratch in a non-English context, even when instructors come primarily from industry backgrounds.

1 Introduction

Natural Language Processing education faces increasing challenges due to emergence of Large Language models (LLMs). Educators must balance foundational concepts with cutting-edge developments. Additionally, most NLP educational resources are English-centric, limiting accessibility for students in other linguistic contexts.

While recent work has explored innovative teaching practices in NLP, including active learning, flipped classrooms (Agirrezabal, 2021), and ethics-aware curricula (Bender et al., 2020), end-to-end NLP course design delivered in non-English instructional contexts remain relatively rare.

We present an NLP course taught at Ukrainian Catholic University (UCU) during the Fall 2025 semester that explicitly addresses these challenges. Our contributions are threefold. First, we describe

the development of a full-semester NLP course taught entirely in Ukrainian, with culturally and linguistically adapted examples and assignments. Second, we demonstrate how an industry-driven teaching team can design a curriculum that bridges research foundations and real-world production concerns. Third, we release all course materials openly, contributing reusable Ukrainian-language resources to the NLP education community.¹

2 Course Design

2.1 Structure and Content

The course spans 15 weeks and combines weekly lectures with hands-on practical sessions. As illustrated in Figure 1, the curriculum is organized into five progressive layers that develop students' competencies from foundational understanding to applied and evaluative skills.

The course begins with *Foundations*, covering core NLP tasks and classical approaches such as tokenization, part-of-speech tagging, and syntactic parsing. It then introduces *Neural Methods*, including PyTorch fundamentals, recurrent neural networks, sequence-to-sequence models, and the Transformer architecture.² These methods are subsequently applied in the *Applied Tasks* module, which focuses on practical NLP problems such as Named Entity Recognition (NER), sequence labeling, machine translation, and information retrieval.

Building on this foundation, the *Modern NLP* module addresses large language models, fine-tuning techniques (e.g., LoRA Hu et al., 2022 and PEFT Lialin et al., 2024), prompt engineering, and retrieval-augmented generation (Lewis et al., 2020).

¹The complete course materials, including lecture slides, Jupyter notebooks, assignments, and video recordings in Ukrainian, are available at <https://github.com/robinhad/ucu-nlp-course> and via a dedicated YouTube playlist at <https://www.youtube.com/playlist?list=PLF5C4LaYzP2LCKeTTgXJgFoXaDDn2C0Ep>.

²<https://pytorch.org/>

*Equal contribution

Finally, the course concludes with *Production and Ethics*, which covers multi-modal models, deployment considerations, model efficiency, and ethical and alignment issues relevant to real-world NLP systems, including discussions on bias detection and bias mitigation, fairness metrics and privacy protection in LLMs.

Overall, the curriculum is intentionally designed to support increasing cognitive complexity, progressing from foundational knowledge acquisition toward application, evaluation, and creative system design, in line with Bloom’s taxonomy (Anderson and Krathwohl, 2001) with reasonable adaptations.

2.2 Culturally Adapted Materials

A key design principle of the course is adaptation to the Ukrainian linguistic and cultural context. All lectures are delivered in Ukrainian, with slides and notebooks using Ukrainian-language examples where appropriate. Homework assignments feature datasets and tasks relevant to Ukrainian NLP challenges, helping students understand that NLP extends beyond English-centric problems. Through these assignments, students are encouraged to critically examine existing large language models not only as powerful tools, but also as artifacts shaped by the linguistic, cultural, and technical environments in which they are developed. This perspective highlights how model performance and biases may vary across low- and mid-resource languages.

2.3 Ethics Integration

Ethical considerations are integrated directly into the curriculum. A dedicated lecture addresses bias and fairness in NLP systems, introducing standard fairness definitions and evaluation metrics such as demographic parity (equal outcomes), equalized odds (equal true and false positive rates), and predictive parity (Hardt et al., 2016; Barocas et al., 2023). Students are also exposed to practical approaches to bias mitigation at different stages of the NLP pipeline. We encourage students to reason about responsibility, risk, and societal impact alongside technical performance. More detailed description is available in Appendix B.

2.4 Assessment Strategy

Quizzes. Interactive Kahoot sessions for engagement and formative assessment are designed to prepare students for three course tests that take part after 5th, 10th and 15th week. ³

³<https://kahoot.com/>

Homework. Kaggle-based assignments with automated evaluation and leaderboard-style feedback. ⁴ Each assignment is adapted to the local context with Ukrainian-language datasets, including sentiment analysis, NER, and instruction tuning of models to answer graduate-level examination questions.

Final Project. A comprehensive project with three evaluated components: (1) product conception and design, (2) technical implementation, and (3) evaluation methodology. Example projects include propaganda detection, multimodal RAG and sentiment analysis of entities mentions in news.

3 Pedagogical Considerations

Heterogeneous Student Backgrounds Our students enter with varying levels of programming experience and prior knowledge. Pre-course survey results (Figure 2 in Appendix C) show substantial variation in professional experience, prior experience working with neural networks, NLP tasks, and LLMs. We address this through: (1) designing assignments with flexible difficulty levels, (2) providing extensive starter code in notebooks, and (3) ensuring all practical sessions and assignments can run on free cloud resources (Google Colab) to avoid hardware barriers. ⁵

Bridging Industry and Research All instructors have industry experience, which shapes the curriculum to include practical considerations often absent from purely academic courses, such as deployment concerns, computational efficiency, data annotation workflows, and production-ready code practices, but grounding them in their research origins.

4 Conclusion

We are presenting an open, Ukrainian-language NLP course designed to balance foundational theory, modern techniques, and industry relevance. Our experience shows that high-quality NLP education can be developed outside English-dominant contexts and that industry-experienced instructors can successfully design academically grounded curricula. We hope that our openly available materials support other educators working in multilingual and underrepresented settings and contribute to a more inclusive ecosystem of NLP teaching resources.

⁴<https://www.kaggle.com/>

⁵<https://colab.google/>

Limitations

The course design process involved several practical challenges. Resource constraints by free-tier cloud platforms limited model sizes and dataset scales, but encouraged efficiency-oriented approaches. Rapid advances in NLP forced to make mid-semester updates to course materials, which we addressed through modular slide design and an emphasis on stable foundational concepts. Finally, developing the curriculum without prior institutional templates required substantial upfront effort, but enabled a clear design unconstrained by legacy structures.

Acknowledgments

We would like to express gratitude to Dmytro Chaplynskyi, Yurii Laba and Maksym Shamrai for their lectures. We thank the students of Ukrainian Catholic University for their engagement and feedback, which significantly shaped the development of the course. We are grateful to Tetiana Zakharchenko for her pedagogical guidance. Besides that, we would like to thank ELEKS for partial support of this work through a grant dedicated to the memory of Oleksiy Skrypnyk.

References

- Manex Agirrezabal. 2021. [The flipped classroom model for teaching conditional random fields in an nlp course](#). In *Proceedings of the Fifth Workshop on Teaching NLP*, pages 80–86. Association for Computational Linguistics.
- Lorin W. Anderson and David R. Krathwohl. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, New York.
- Amanda Askeff, Yuntao Bai, Anna Chen, Dawn Drain, Deep Ganguli, Tom Henighan, Andy Jones, Nicholas Joseph, Ben Mann, Nova DasSarma, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez, Jackson Kernion, Kamal Ndousse, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark, and 3 others. 2021. [A general language assistant as a laboratory for alignment](#). *Preprint*, arXiv:2112.00861.
- Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2023. *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press.
- Emily M. Bender, Dirk Hovy, and Alexandra Schofield. 2020. [Integrating ethics into the nlp curriculum](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*. Association for Computational Linguistics.
- Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. [Calibrating noise to sensitivity in private data analysis](#). In *Proceedings of the Third Conference on Theory of Cryptography, TCC'06*, page 265–284, Berlin, Heidelberg. Springer-Verlag.
- Moritz Hardt, Eric Price, and Nathan Srebro. 2016. [Equality of opportunity in supervised learning](#). *Preprint*, arXiv:1610.02413.
- Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, and Weizhu Chen. 2022. [Lora: Low-rank adaptation of large language models](#). *arXiv preprint arXiv:2106.09685*.
- Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS '20*, Red Hook, NY, USA. Curran Associates Inc.
- Vladislav Lialin, Vijeta Deshpande, Xiaowei Yao, and Anna Rumshisky. 2024. [Scaling down to scale up: A guide to parameter-efficient fine-tuning](#). *Preprint*, arXiv:2303.15647.
- Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. 2017. [Membership inference attacks against machine learning models](#). *Preprint*, arXiv:1610.05820.
- Amer Sinha, Thomas Mesnard, Ryan McKenna, Daogao Liu, Christopher A. Choquette-Choo, Yangsibo Huang, Da Yu, George Kaissis, Zachary Charles, and 1 others. 2025. [Vaultgemma: A differentially private gemma model](#). *arXiv preprint arXiv:2510.15001*. Open differential privacy LLM research.
- Laura Weidinger, John Mellor, Maribeth Rauh, Conor Griffin, Jonathan Uesato, Po-Sen Huang, Myra Cheng, Mia Glaese, Borja Balle, Atoosa Kasirzadeh, Zac Kenton, Sasha Brown, Will Hawkins, Tom Stepleton, Courtney Biles, Abeba Birhane, Julia Haas, Laura Rimell, Lisa Anne Hendricks, and 4 others. 2021. [Ethical and social risks of harm from language models](#). *Preprint*, arXiv:2112.04359.

A Course Structure

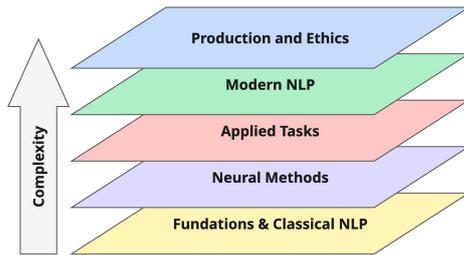


Figure 1: Structure of the NLP course, illustrating the progression from foundational concepts to applied, modern, and production-oriented topics, aligned with increasing cognitive complexity.

B Ethics Integration

Privacy risks in large language models are discussed through the lens of membership inference attacks (MIA) (Shokri et al., 2017), which demonstrate how models may unintentionally reveal information about their training data. The course covers mitigation strategies based on differential privacy (Dwork et al., 2006). As a concrete example, we discuss VaultGemma, a language model trained from scratch with differential privacy guarantees (Sinha et al., 2025).

In addition, the course suggests ways to build safer LLM-based systems, supported by real-world case studies that highlight potential harmful effects of conversational agents (Askill et al., 2021; Weidinger et al., 2021).

C Qualitative Feedback

After the course, we conducted an anonymous feedback survey. All students were informed in advance that their responses could be used in anonymized form for research and publication purposes related to this paper.

Students highlighted the practical relevance of the course and the importance of contextualized, hands-on learning activities. Representative comments include:

What did you find most valuable about this course?

“For me personally, the nuances I’ve learned during homework assignments, especially about NER and fine-tuning, were the most valuable.” [original in English]

What did you find most valuable about this course?

“I liked the homework assignments, because after completing them I understood the material better, especially data preparation, which later became very useful.” [translated from Ukrainian]

What would you suggest to improve?

“It might be beneficial to add some small practical tasks. [...] It is important that such micro-tasks are not too difficult and that they require applying what was discussed in the lecture or studied independently.” [original in English]

Several students also suggested extending the time devoted to foundational machine learning concepts and introducing additional, more granular practical exercises. These suggestions will inform future iterations of the course.

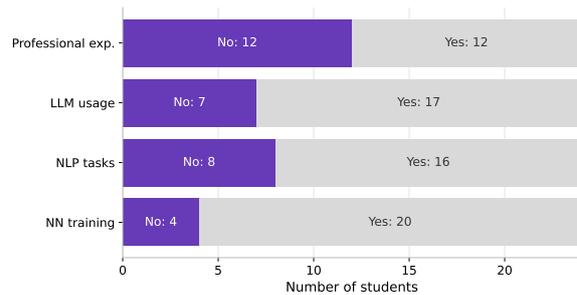


Figure 2: Pre-course self-reported experience of students (N=24). **NN training** indicates prior implementation or training of neural networks; **NLP tasks** refers to prior work on NLP-related tasks or tools; **LLM usage** denotes prior use of large language models; and **Professional exp.** indicates any prior professional experience. Bars show the proportion of students with (Yes) and without (No) prior experience.

Week	Topic	Module	Key Concepts / Skills
1	Course overview; NLP tasks and use cases	Foundations	NLP problem types, pipeline overview, real-world applications
2	Classical NLP methods (lecture + practice)	Foundations	Tokenization, n-grams, POS tagging, parsing
3	Transition to machine learning; PyTorch basics	Neural Methods	Classification vs. regression, loss functions, gradient descent
4	Sequence models and Transformers	Neural Methods	RNNs, seq2seq, attention, Transformer architecture
5	Applied NLP tasks (Part I)	Applied Tasks	Text classification, sequence modeling, practical datasets
6	Applied NLP tasks (Part II)	Applied Tasks	NER, sequence labeling, machine translation
7	Evaluation in NLP	Applied Tasks	Precision/recall, F1, BLEU, task-specific metrics
8	Data processing and annotation	Applied Tasks	Dataset creation, annotation workflows, data quality
9	Large Language Models	Modern NLP	Pretraining objectives, scaling, tokenization, LLM usage
10	Prompt engineering	Modern NLP	Prompt design, zero-/few-shot prompting, evaluation pitfalls
11	Parameter-efficient fine-tuning and IR	Modern NLP	LoRA, PEFT, sentence transformers, information retrieval
12	Retrieval-augmented generation	Modern NLP	RAG pipelines, dense retrieval, grounding
13	Multimodal models and deployment	Production	CLIP, multimodal retrieval, efficiency, quantization
14	Ethics, safety, and alignment	Production & Ethics	Bias, fairness metrics, privacy, responsible deployment
15	Final project presentations	Synthesis	End-to-end system design and evaluation

Table 1: Week-by-week overview of course topics and their progression from foundational NLP concepts to modern large language models and production considerations.

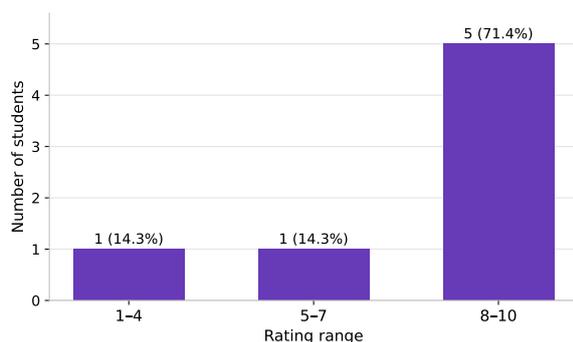


Figure 3: Post-course ratings of the overall course structure (N=7), aggregated into low (1–4), medium (5–7), and high (8–10) score ranges.

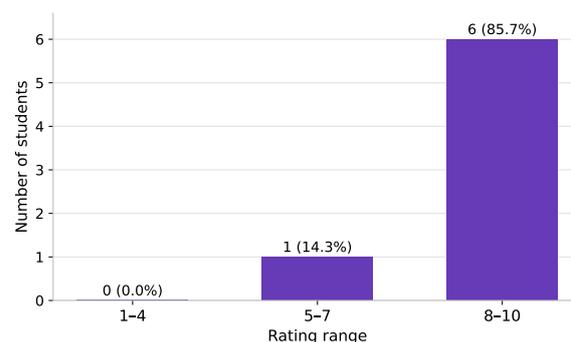


Figure 4: Post-course ratings of lecture slides and notebooks quality (N=7), aggregated into low (1–4), medium (5–7), and high (8–10) score ranges.

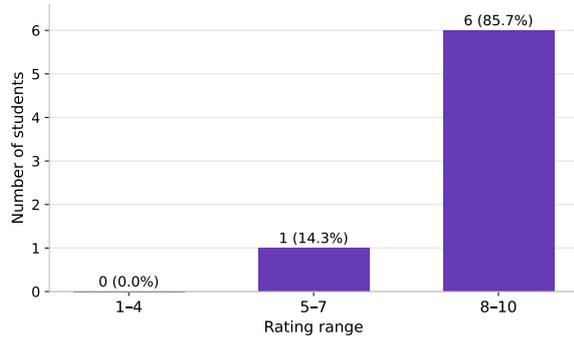


Figure 5: Post-course ratings of assignment quality (N=7), aggregated into low (1–4), medium (5–7), and high (8–10) score ranges.

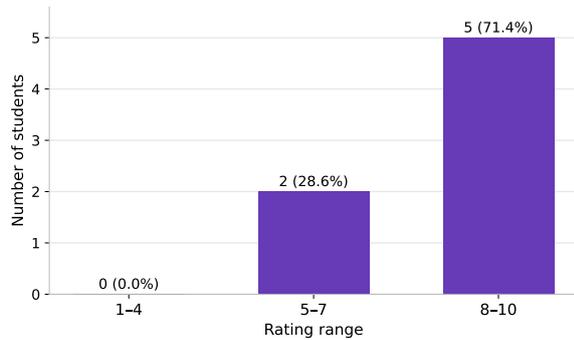


Figure 6: Post-course self-reported confidence in applying basic NLP techniques (N=7), aggregated into low (1–4), medium (5–7), and high (8–10) score ranges.

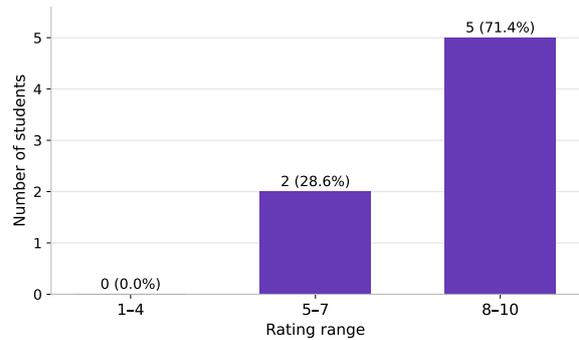


Figure 8: Post-course self-reported preparedness for NLP-related interviews (N=7), aggregated into low (1–4), medium (5–7), and high (8–10) score ranges.

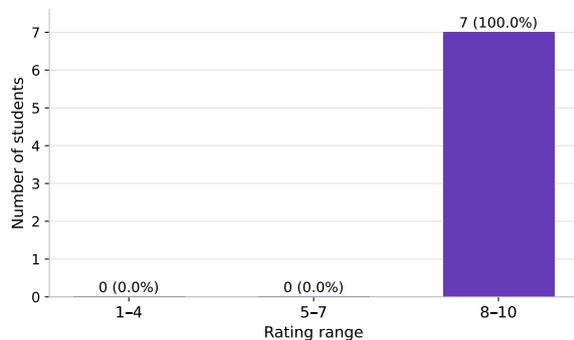


Figure 7: Post-course self-reported understanding of NLP evaluation metrics (N=7), aggregated into low (1–4), medium (5–7), and high (8–10) score ranges.