

Blended Intensive Program

Artificial Intelligence (AI) for Humanities: from Text Simplification to Automatic Humor Analysis (practical course)



General information

Course Title	Artificial Intelligence (AI) for Humanities: from Text Simplification to Automatic Humor Analysis (practical course)
Abstract	<p>This is an introductory course on artificial intelligence and natural language processing. Students will learn:</p> <ul style="list-style-type: none"> ● how to do data analysis in Python; ● how to apply large pre-trained deep learning models, such as BLOOM, GPT-3, Google mT5, for new data and tasks, including text simplification, terminology extraction, and humor analysis; ● how to evaluate the results generated by AI; ● how to generate images based on textual prompts.
Calendar	<ul style="list-style-type: none"> ● between 06/03 and 18/03 2023: 1st Virtual part / 8h ● 20-24/03 2023: On-site Intensive Course in Brest(physical mobility) / 24h ● between 01/04 and 06/06 2023: 2nd Virtual part / 20 h of

	collaborative work
Total number of hours:	52 (24h on site + 28h online)
ECTS	5
Teacher(s) in charge	liana.ermakova@univ-brest.fr
Contact	Regarding organisational aspects: severine.allain@univ-brest.fr Regarding pedagogical aspects: liana.ermakova@univ-brest.fr
Number of participants	The minimum number of participants is 15, maximum is 24. Each SEA-EU university can propose up to 4 students. Interested teachers might contact Pr Liana Ermakova
Nominations	Nominations deadline : 15/01/2023 Students have to be nominated by their teachers / international offices through this form
Mobility costs	This mobility is eligible for Erasmus+. Please contact your university for more information.
More info about last year's course	https://www.youtube.com/watch?v=KrkYO9RWUEc

Target group

Disciplinary background	Students interested in Artificial Intelligence, especially state-of-the-art large pre-trained models, natural Language Processing and image generation from text descriptions, Scientific text simplification and Computational Humor. The course is open to Bachelor, Master and Doctoral students in Translation, Digital Humanities, Computer Science, Linguistics, and Social Science interested in AI.
Additional requirements	English B1 No specific prerequisites are required for this class but basic knowledge of programming in Python might be helpful. The course is targeted to students interested in Artificial Intelligence, especially state-of-the-art large pre-trained models, natural Language Processing and image generation from text descriptions, Scientific text simplification and Computational Humor. It is open to Bachelor, Master and Doctoral students in Translation, Digital Humanities, Computer Science, Linguistics, and Social Science interested in AI.

Any required material/software to take part to the course:	No software is required. The exercises will be available via Google Colab
Learning objectives/outcomes:	<p>The students will learn how to apply</p> <ul style="list-style-type: none"> ● textual data analysis ● transfer learning for various natural language processing tasks (machine translation, text generation, text classification,...) ● image generation from textual prompts ● various evaluation metrics

Structure of the course

<p>Introductory phase (virtual) : 8h 1h of presentation and opening of the course (synchronous) and 7h of individual practice (asynchronous)</p>	<ul style="list-style-type: none"> ● introduction into Google Colab https://colab.research.google.com/ ● various forms of data input/output, file formats (json, csv, txt, etc.) ● regular expressions https://docs.python.org/3/library/re.html <p>The objective is to obtain prerequisites for basic data processing and basic text processing with regular expressions.</p>
<p>Presential phase (presential, in UBO, Brest): 24h 20th march 2023-24th march 2023 (5 days)</p>	<p>The content of the course might be adapted according to the level of the group.</p> <ul style="list-style-type: none"> ● introduction into natural language processing ● introduction into deep learning ● few shot learning by prompt tuning for natural language processing (BLOOM, GPT-3, etc.) ● image generation from textual prompts (type DALL-E) ● data analysis library Pandas ● non-neural approaches for natural language processing ● transformer architecture ● transfer learning ● evaluation metrics <p>The objective is to learn how to solve natural language tasks and how to evaluate the obtained results.</p>
<p>Group collaborative work (virtual): 20h 18h of collaborative work (asynchronous) + 2h of restitution (synchronous)</p>	<p>Students will ask to do a collaborative projects on application of AI models to one of the following tasks (non-exhaustive list):</p> <ul style="list-style-type: none"> ● automatic summarization ● machine translation ● terminology extraction and explanation ● text simplification ● wordplay detection ● wordplay interpretation

	<ul style="list-style-type: none"> • wordplay generation • wordplay visualisation • visualisation of scientific information <p>...</p> <p>The objective is to carry out an AI project in autonomy, to analyze and to present the results.</p>
Evaluation	<p>The evaluation will be based on the quality of the written report describing the project and the quality of the oral presentation. The reports (working notes) should be submitted as working notes at CLEF. https://clef2023.clef-initiative.eu/index.php</p> <p>Accepted working notes will be published in CEUR-WS proceedings (see examples at https://ceur-ws.org/Vol-3180/)</p>

Practical information

Adress of the course	The course will take place in Brest City Center, at the Humanities Faculty, 20 rue Duquesne, 29200 Brest.
Accomodation	<ul style="list-style-type: none"> • Hotels in the city center are between 65 and 80€ / night. (Hotel Vauban, Abalys Hotel, Citotel Gare...) • You can however find a list of private options (cheaper) on the usual booking platforms (booking, airbnb..) • The Youth Hostel costs around 25€/night (dorms) and is located +/- 20 mn by bus from the city center