COMP7630 – Web Intelligence and its Applications

Course Outline

Valentino Santucci

(valentino.santucci@unistrapg.it)

Teacher's contacts

- Valentino Santucci

- Emails:
    - valentino.santucci@unistrapg.it
    - v-santucci-1@hkbu.edu.hk

- Personal website:
   - https://valentinosantucci.github.io/

Who am I?

- **MSc in Computer Science**
  - from University of Perugia (Italy)

- **PhD in Maths and Computer Science**
  - from University of Perugia (Italy)

- **Professor of Computer Science and Engineering**
  - at University for Foreigners of Perugia (Italy)

- **Rector's Delegate for Technological Innovation**
  - at University for Foreigners of Perugia (Italy)

My scientific interests

- Artificial Intelligence
- Evolutionary Computation
- Combinatorial Optimization
- Machine Learning
- Natural Language Processing
- Sustainability-related applications of AI

Where is Perugia?
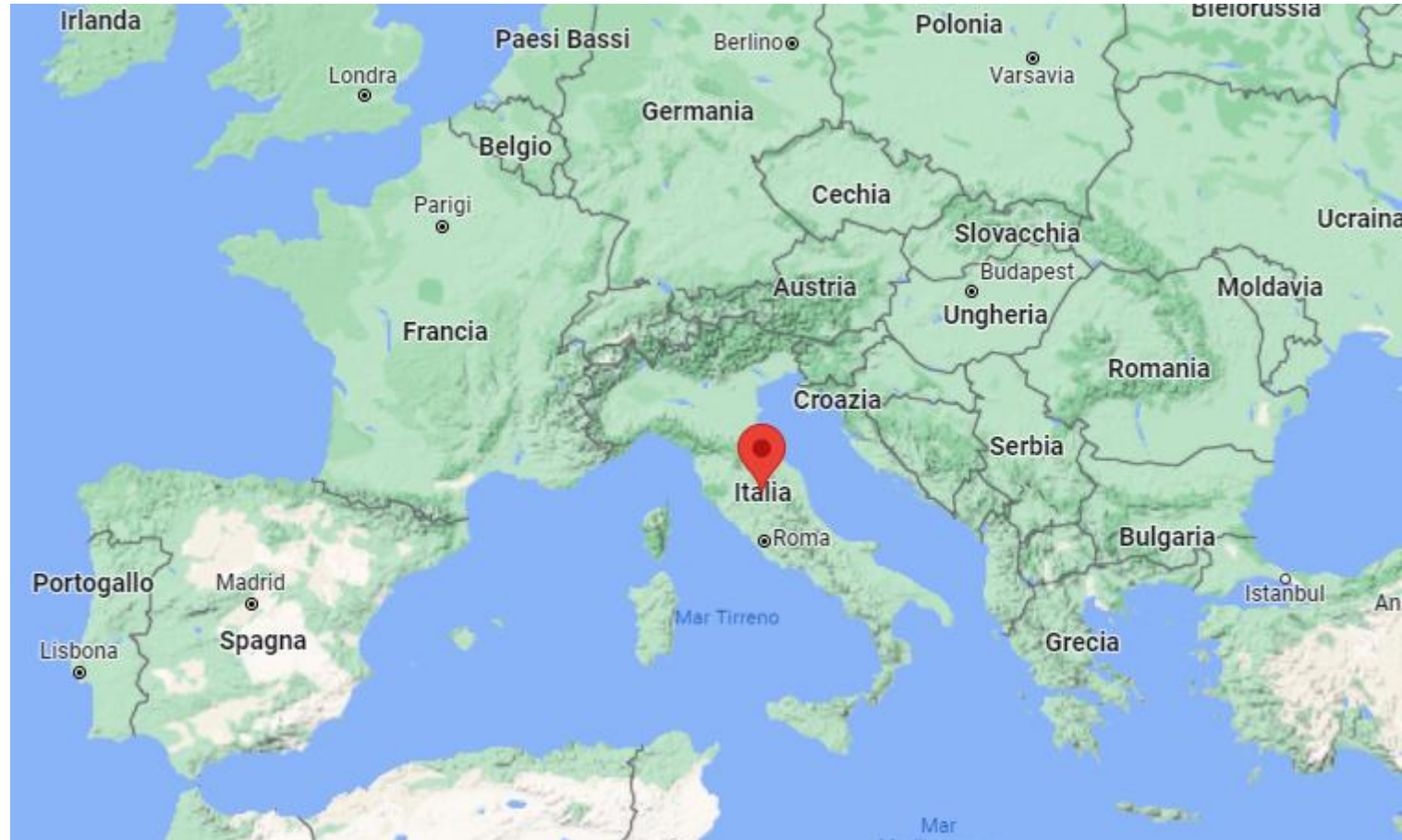
Course Description

This course introduces the fundamental concepts as well as practical applications of Web Intelligence (WI) which combines contemporary Artificial Intelligence and advanced Information Technology in the context of Web-empowered systems, environments, and activities. Also, advanced topics related to Web Intelligence (WI) as well as their possible impact to different sectors of the society. Students after taking this subject should be able to 1) identify the possible impact of Web Intelligence in the society, and 2) apply WI related techniques to advance existing Web-based systems and on-line business platforms.

Course Intended Learning Outcomes (CILO)

Knowledge (Lectures)

1. Describe the concepts and applications of contemporary Artificial Intelligence and advanced Information Technology in the context of Web empowered systems, environments, and activities

2. Explain the techniques and issues central to the development of WI computing systems

3. Explain the practical applications of Web intelligence

Professional Skill (Assignments, Labs and Group Projects)

4. Solve advanced technical problems in generic Web environment

5. Apply specific methods and techniques in a number of Intelligent Web applications

Attitude (Group Projects)

6. Develop team spirit in tackling challenging problem in Web

Course Content

0. Web Intelligence in a Nutshell

1. Linear Algebra background + Dimensionality Reduction
2. Natural Language Processing pipelines + Text Classification/Clustering
3. Tutorial on Python programming + SciKit-Learn algorithms applied to texts
4. Evolutionary Computation for Web Intelligence
5. Singular Value Decomposition + Probability Theory background

6. Association Rules Mining
7. Web Information Retrieval
8. Collaborative Filtering
9. Social Network Analysis
10. Web Usage Mining
11. Sentiment Analysis

What will you know after the course?

- Which are the main applications and approaches of Web Intelligence

- Undestand the maths behind Web Intelligence techniques

- Implement prototypical Web Intelligence techniques in Python

Assessment

- **Continuous Assessment**   **40%**
  - 1 Assignment   15%
  - 1 Group Project   25%

- **Examination**   **60%**
  - Analysis based questions will be used to assess the understanding of WI computing systems.
  - Problem solving questions will be used to assess the students' ability in tackling Web Intelligence applications.

Assignment

- **What?**
  Programming questions + Written questions + Exercises

- **When?**
  After approximately 1/2 of the course
  (exact dates will be announced after the add/drop period)

- **How to submit?**
  Through Moodle

Group Project

- Two options: either a development project or a term paper

- You are expected to form a group and submit a project proposal (to be approved by teacher) after half of the course period

- You are expected to submit project reports + a video presentation where group members talk over slides

- Number of students per group and exact deadlines will be announced later on the basis of the number of students following the course

Development Project / Term Paper

- **Development project**:
  - You are expected to submit a working software prototype by going through the steps of web data acquisition, pre-processing, mining, and presentation.
  - <u>Example</u> projects (you can either choose one of them or propose any other reasonable subject):
    - Automatic suggestion of hashtags using association rules mining (for Instagram, Twitter, Moodle, etc.);
    - Text classification and/or clustering and/or ranking for web-related job listings from JOBSDB, Linkedin, etc.;
    - Characters importance through the chapters of narrative books (by using Named Entity Recognition + Social Network Analysis tools);
    - Emotional analysis of tendencies in web-published user reviews;
    - Evolutionary algorithms for optimizing the design of a web advertisement banner (by simulated users feedback);
    - Evolutionary algorithms for learning association rules of e-commerce items.

- **Term paper**:
  - You are expected to submit a literature review paper presenting the technological development and applications of a particular area in Web Intelligence. In order to get a high mark, each group is suggested to pick at least 3 x [group size] papers which are good and recent enough, read them in detail, compare and contrast them, and discuss possible future work and new opportunities of application.
  - <u>Example</u> topics (you can either choose on them of propose any other reasonable subject):
    - Evolutionary algorithms for web intelligence applications;
    - Word embeddings such as Word2Vec, FastText, Glove, etc.;
    - Large language models and transfer learning;
    - Zero-Shots and Few-Shots learning techniques.

- **FURTHER DETAILS ABOUT GROUP PROJECTS WILL BE ANNOUNCED LATER.**

Important notes

- Assignment and project reports, if not otherwise specified, have to be written in your own words instead of copy-and-paste from the Web or generate with AI tools, and submitted via HKBU Moodle (upload options will be available at the right time).

- Plagiarism and AI-generated cases once caught will receive heavy penalty.

Assessment rubric

	Excellent (A)	Good (B)	Satisfactory (C)	Fail (F)
Describe concepts and applications of AI and advanced IT in Web empowered systems,	Thorough description of almost all concepts and applications	Description of most of the concepts and applications	Description of some of the concepts and applications	Description of a limited number of concepts and applications
Explain the techniques and issues central to the development of WI computing systems	Thorough explanation of almost all techniques and issues	Explanation of most of techniques and issues	Explanation of some of techniques and issues	Explanation of a limited number of techniques and issues
Explain the practical applications of Web intelligence	Thorough explanation of almost all applications	Explanation of most of the applications	Explanation of some of the applications	Explanation of only a few applications
Solve advanced technical problems in generic Web environment	Solving almost all technical problems	Solving most of the technical problems	Solving some of the technical problems	Solving of only a very small number of technical problems
Apply specific methods and techniques in a number of Intelligent Web applications	Application of almost all relevant methods and techniques to applications, including those involving novel solutions	Application of most of the correct methods and techniques to applications	Application of some of the correct methods and techniques to applications	Application of only a very small number of correct methods and techniques to applications

Course Reference Material

[1] Teaching Material provided by the teacher on the Moodle course page

   URL: https://buelearning.hkbu.edu.hk/course/view.php?id=115052

[2] Liu Bing. Web Data Mining – Exploring Hyperlinks, Contents and Usage Data, Springer, 2011

   URL: https://link.springer.com/book/10.1007/978-3-642-19460-3

   (Note: you may try using a search engine to locate a downloadable soft copy)

[3] Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu. Social Media Mining: An Introduction, Cambridge University Press, 2014

   URL: http://www.socialmediamining.info/

   (Free download of a soft copy)

Further shorter references may be provided in the Moodle page, together with the indications of the related chapters of books [2] and [3].

Provisional calendar of the classes

- #01 – 15 January
- #02 – 22 January
- No class on 29 January
- #03 – 05 February
- #04 – 12 February
- #05 – 19 February
- #06 – 26 February
- #07 – 05 March
- #08 – 12 March

- #09 – 19 March
- #10 – 26 March
- #11 – 2 April
- #12 – 9 April
- #13 – 16 April

- 2–17 May: Examination period

According to https://calendar.ar.hkbu.edu.hk/2024-2025/academic-year-and-academic-calendar/academic-calendar-2024-2025

Teaching Assistant

- Hongtao Wang

- cshtwang@comp.hkbu.edu.hk (please always add me in cc)

- Available for face-to-face meetings on Tuesday afternoons (office no. DLB625F), just in case of some request I cannot handle via email, Moodle forum or Zoom (which remain the priority communication channels)