



## Batch 10 – Project – Data Mining -1 (Unsupervised Learning)

The project has **SIX** deliverable components. This project is a team assignment (honor code – 3N-b) except for the second and the fifth component listed in this document; the second and the fifth components are individual components (honor code – 4N).

- 
- Team size: 4-6, preferably 5.
  - Teams across sections are discouraged.

### Contents

Project Title:.....	1
The objective of the project: .....	1
Phase I (intermediate) deliverables .....	2
1. Intermediate project presentation.....	2
2. Evaluating other teams.....	2
3. Intermediate project report.....	2
Phase II (final) deliverables .....	3
4. Final project presentation.....	3
5. Evaluating other teams.....	3
6. Final project report.....	3

### Project Title:

Analysis of *WTO-UNCTAD-ITC annual trade in services* using *unsupervised* learning methods

### The objective of the project:

Using the techniques discussed in the class discover valuable insights through mining WTO-UNCTAD-ITC annual trade in services dataset publicly available under the following URL.

[https://www.wto.org/english/res\\_e/statistics\\_e/trade\\_datasets\\_e.htm](https://www.wto.org/english/res_e/statistics_e/trade_datasets_e.htm)

Your analysis should be based primarily on the data available on the above website. However, you are more than encouraged to refer to any other publicly available resources to familiarize yourself with the contextual knowledge on trades. If necessary, you may also use any secondary data or knowledge to substantiate your findings or interpretations of what you discover in the primary dataset. Please note that you are not necessarily required to use all of the data provided on this website. Furthermore, you may not be able to use the data as is without further processing the data. You may use only a subset of the data, and thereafter, prepare the final dataset based on the scope of the problem you decide to go with. The

scope of the problem that you pursue should be meaningful and well-motivated, and it should create social and/or economic values for the identified stakeholder; you must precisely identify the stakeholder for this exercise.

## Phase I (intermediate) deliverables

### 1. Intermediate project presentation

**Teamwork \* presentation during Session 06 \* upload slides—in PDF format only—on the LMS, at least, by the midnight before the day of Session 06**

---

- Duration of the presentation: up to 7 minutes
- Slide 1 (title, team members)
- In no more than 5 slides discuss your current findings, the methodology used, social & economic values of the findings along with actionable recommendations.
- Be very precise and focus on educating your peers about your team's approach and justifying the value of the insights that you are presenting to them.
- You are discouraged to use your own machine to present in the class. Please bring a copy of the PDF version of the slide deck on a pen drive and use the machine in the classroom for the presentation.
- Anyone in the team or more than one member of the team can present. It is absolutely up to the team to decide.

### 2. Evaluating other teams

**Individual work \* during Session 6**

---

You are required to evaluate others' presentations, as well. You'll be given a specific evaluation format during the presentation session for this purpose. Please bring your own laptop. You'll have to submit your evaluation online as the presentations by different teams proceed. You should NOT evaluate your own team.

### 3. Intermediate project report

**Teamwork \* upload the report—in PDF format only—on the LMS, at least, by the midnight before the day of Session 06**

---

**Page 1 (cover page)**

- Meaningful title of the project based on what you decided to be the scope of the project
- List of team members
- Executive Summary (in no more than 500 words)

**In up to 5 pages, excluding the cover page, present the insights that you find worth sharing.**

- Describe each of the key insights in not more than one sentence each, followed by ...
- Steps (data processing, tools used, parameters chosen, etc.) to reproduce the insight
- Social and/or economic values of the insight(s) to the identified stakeholder
- How the value of those insights can be harnessed—implementation/deployment issues, by the stakeholder in this exercise. Additional recommendations, if any.
- Please do not clutter your project report with insignificant insights as it will only dilute the value of other significant findings from your analysis.

- The bullets-points-only-like writing is strongly discouraged. The report is expected to be of the professional quality that conforms to business writing standard.

**Appendix (if necessary, you may add some tables, charts, etc.)**

## **Phase II (final) deliverables**

### **4. Final project presentation**

**Teamwork \* presentation over WebEx on the scheduled date (after the second residency and toward the end of the term; exact date to be announced later; stay tuned) \* upload slides—in PDF format only—on the LMS, at least, by the midnight before the day of the presentation**

- Duration of the presentation: up to 7 minutes
- Slide 1 (title, team members)
- In no more than 5 slides discuss your current findings, the methodology used, social & economic values of the findings along with actionable recommendations.
- Be very precise and focus on educating your peers about your team's approach and justifying the insights that you are presenting to them.
- Anyone in the team or more than one member of the team can present. It is absolutely up to the team to decide.

### **5. Evaluating other teams**

**Individual work \* during the final presentations**

You are required to evaluate others' presentations, as well. You'll be given a specific evaluation format before the presentations start. You'll have to submit your evaluation online as the presentations by different teams proceed. You should not evaluate your own team.

### **6. Final project report**

**Teamwork \* upload the report—in PDF format only—on the LMS, at least, by the midnight before the day of the presentation.**

- You should follow same guidelines as was expected in the intermediate project report, except that for the final report you may use up to 10 pages.
- The final report should be a significantly improved version of the intermediate report; with all comments and feedback (that applies to you), received from peers and instructor during the presentation, incorporated; new tools, which were introduced in the second half of the class, applied.
- The bullets-points-only-like writing style is strongly discouraged.
- The writing quality must be of a professional quality that conforms to business writing standard.

*After submission of the project deliverables, team members may be asked to fill a peer evaluation report, if multiple team members indicate insufficient contribution of a member. An appropriate penalty will apply to concerned members.*