**PhD in INGEGNERIA DELL'INFORMAZIONE / INFORMATION TECHNOLOGY - 36th cycle**

Research Area n. 1 - Computer Science and Engineering

Research Field: DEEP LEARNING AND NATURAL LANGUAGE PROCESSING FOR ADVANCED INDUSTRIAL PROCESS MINING

<table>
<thead>
<tr>
<th>Monthly net income of PhD scholarship (max 36 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>€ 1400.0</td>
</tr>
</tbody>
</table>

In case of a change of the welfare rates during the three-year period, the amount could be modified.

<table>
<thead>
<tr>
<th>Context of the research activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>The research proposal addresses the investigation and application of Deep Learning models usually employed for Natural Language Processing, in the field of Process Mining at the industrial plant level. Natural Language Processing is the branch of artificial intelligence that deals with the computer aided automatic analysis and processing of information expressed in a natural language (human language). In the last years, there has been a great growth in this discipline, thanks to the creation of powerful deep learning algorithms, having applications in, e.g., intelligent keyboards, sophisticated chat-bots, and document understanding. Process Mining deals with the study of series of discrete data from a process and the extraction of knowledge from them, through the use of specific data mining algorithms applied to event logs. Capturing the similarities between discrete process data and language (meant as series of words), the theme of the doctorate is the study of a theoretical framework, and a general approach, for industrial process mining taking advantage from the latest improvement in Deep Learning models for Natural Language Processing. A possible application of derived methods and techniques will be sought to events coming from an industrial plant (e.g., from the SCADA system) in order to forecast...</td>
</tr>
</tbody>
</table>

Motivation and objectives of the research in this field
### Methods and techniques that will be developed and used to carry out the research

Starting from the study of the state of the art in the fields of Deep Learning (for Natural Language Processing) and Process Mining, the research will develop methods derived from Natural Language Processing to be applied to the mining of an industrial process (i.e., a foundry plant). Classical NLP models based on Recurrent Neural Networks, such as LSTM and GRU, will be used after learning proper embedding for industrial plant data. Subsequently models based on the Transformer will be studied and adapted to the specific domain in order to learn proper attention mechanisms.

### Educational objectives

The PhD program offers the highest level of formation, leading to a high expertise in the chosen field, along with a training in the latest topics research currently explored by the scientific community. The candidate will be proficient in Deep Learning for discrete events sequences and industrial process mining.

### Job opportunities

The doctoral graduates have opportunities both in academy and in industry. The collaboration with industrial partners allowed past PhD graduates to find satisfactory job position in private company. Opportunities in academy are also made possible.

### Composition of the research group

- 1 Full Professors
- 2 Associated Professors
- 0 Assistant Professors
- 3 PhD Students

### Name of the research directors

Matteo Matteucci

### Contacts

Ph.: 0223993470
Email: matteo.matteucci@polimi.it
https://www.deib.polimi.it/eng/people/details/267262

The research group belong to the research line on Artificial Intelligence and Robotics
https://www.deib.polimi.it/eng/research-lines/details/118
### Additional support - Financial aid per PhD student per year (gross amount)

<table>
<thead>
<tr>
<th>Housing - Foreign Students</th>
<th>1st year</th>
<th>2nd year</th>
<th>3rd year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1500.0 € per student</td>
<td>1000.0 € per student</td>
<td>1000.0 € per student</td>
</tr>
</tbody>
</table>

Max number of financial aid available: 4, given in order of merit.

| Housing - Out-of-town residents (more than 80Km out of Milano) | -- |

### Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information

**LIST OF UNIVERSITIES, COMPANIES, AGENCIES AND/OR NATIONAL OR INTERNATIONAL INSTITUTIONS THAT ARE COOPERATING IN THE RESEARCH:** Politecnico di Milano - DEIB (deib.polimi.it); STIIMA-CNR (www.stiima.cnr.it)

The PhD will be co-supervised by the research team in Machine and Manufacturing Control Systems (2MaCS) at the Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing of the National Research Council of Italy. The objective of the 2MaCS group research activities is to develop factories as integrated systems capable to respond to market needs in a fast, efficient, safe and sustainable way, by means of the development and the integration of intelligent control, machine learning and automation systems. The PhD candidate will be asked to contribute to the MSs students' supervision, and to do regular visits the 2MaCS research Lab.

**INCREASE IN THE SCHOLARSHIP FOR STAYS ABROAD:** Euro 566.36 per month, for up to 6 months

**EDUCATIONAL ACTIVITIES** (purchase of study books and material, including computers, funding for participation in courses, summer schools, workshops and conferences): financial aid per PhD student per year
- 2nd year: euros per student (1534)
- 3rd year: euros per student (1534)

**TEACHING ASSISTANSHIP:** (availability of funding in recognition of supporting teaching activities by the PhD student)

There are various forms of financial aid for activities of support to the teaching practice. The PhD student is encouraged to take part in these activities, within the limits allowed by the regulations.

**COMPUTER AVAILABILITY:**
- 1st year: individual use
- 2nd year: individual use
- 3rd year: individual use

**DESK AVAILABILITY:**
1st year: individual use
2nd year: individual use
3rd year: individual use