COMPUTER AND CONTROL ENGINEERING

Urban data science

Funded By	Dipartimento DAUIN Centro Interdipartimentale SmartData@PoliTO
Supervisor	CHIUSANO SILVIA ANNA - silvia.chiusano@polito.it
Contact	CERQUITELLI TANIA - tania.cerquitelli@polito.it
Context of the research activity	In the urban ecosystem a multitude of intertwined systems coexists, varying from people sociality to transport systems. While each of these urban facets already represents a complex system, their interconnection is definitively a challenging scenario. Urban data science represents a new inter-disciplinary field of data science research, that addresses the development of competences and technical infrastructures required to study and address urban challenges from a data-driven perspective.
	The PhD student will work on the study, design and development of proper data models and novel solutions and for the acquisition, integration, storage, management and analysis of big volumes of heterogeneous urban data. The research activity involves multidisciplinary knowledge and skills including database, machine learning techniques, and advanced programming. Different case studies in urban scenarios such as urban mobility, citizen-centric contexts, and healthy city will be considered to conduct the research activity. The objectives of the research activity consist in identifying the peculiar characteristics and challenges of each considered application domain and devise novel solutions for the management and analysis of urban data for each domain. More urban scenarios will be considered with the aim of exploring the different facets of urban data and evaluating how the proposed solutions perform on different data collections. More in detail, the following challenges will be addressed during the PhD: - Suitable data fusion techniques and data representation paradigms should be devised to integrate the heterogeneous collected data into a unified representation describing all facets of the targeted domain. For example, since urban data are often collected with different spatial and temporal granularities, suitable data fusion techniques should be devised to support a spatio-temporal alignment of collected data. - Adoption of proper data models. The storage of heterogeneous urban data collections requires the use of alternative data representations to the relational model such as NoSQL databases (e.g., MongoDB), also able to manage geo-referenced data. - Design and development of algorithms for big data analytics. Huge volume of data demands the definition of novel data analytics strategies also exploiting recent analysis paradigms and cloud based platforms. Moreover,

Objectives	urban data is usually charaterized by spatio-temporal coordinates describing when and where data has been acquired, which entails the design of suitable data analytics methods. Outline of work plan 1st Year. The PhD student will review the recent literature on urban computing to identify the up-to-date research directions and the most relevant open issues in the urban scenario. Based on the outcome of this preliminary explorative analysis, an application domain, such as urban mobility, will be selected as a first reference case study. The selected domain will be investigated to (i) identify the open research issues, (ii) identify the most relevant data analysis perspectives for gaining useful insights, and (iii) assess of main data analysis issues. The student will perform an exploratory evaluation of state-of-the-art technologies and methods on the considered domain, and will present a preliminary proposal for the optimization techniques of these approaches. 2nd and 3rd Year. Based on the results of the 1st year activity, the PhD student will design and develop a suitable data analysis framework including innovative analytics solutions to efficiently extract useful knowledge in the considered domain, aimed at overcoming weaknesses of state-of-the-art methods. Moreover, during the 2nd and 3rd year, the student will progressively consider a larger spectrum of application domains in the urban scenario. The student will evalute if and how his/her proposed solutions can be applied to the new considered domains as well as he/she will propose novel analytics solutions. During the PhD, the student will have the opportunity to cooperate in the development of solutions applied to the research projects on smart cities. The student will also complete his/her background by attending relevant courses. The student will participate to conferences presenting the results of his/her research activity.

Skills and competencies for the development of the activity	knowldege of machine learning techniques, good programming skills