



UNIVERSITÀ DEGLI STUDI DI TORINO

Titoli dei progetti di ricerca

Dottorato di Ricerca in MODELING AND DATA SCIENCE

Research Topics

PhD Programme in MODELING AND DATA SCIENCE

TEMATICHE GREEN

Referente scientifico/tutor	Titolo del progetto Requisiti	Descrizione sintetica
Rosa Meo	ITA: Monitoraggio delle portate fognarie come strumento di supporto nello sviluppo di strategie gestionali atte alla mitigazione dell'impatto dei cambiamenti climatici	ITA: Inquadramento del contesto/problema Il cambiamento climatico provoca grandi esondazioni e la caduta improvvisa di enormi quantità d'acqua che provocano danni ingenti quali allagamenti superficiali se non ben

ENG:

Monitoring of sewerage flow rates as a tool in the development of strategies to mitigate the impact of climate change.

indirizzate. Le reti fognarie di molte zone italiane sono vetuste e con condotti insufficienti a garantire il deflusso di grandi quantità d'acqua. Per affrontare questa criticità si propone la soluzione di studiare un modello di piano fognario e i possibili interventi per migliorarlo con l'azienda (BrianzAcque) che affianca il Gestore del servizio idrico.

Risultati attesi

Con l'espletamento delle attività legate agli step di progetto, che prevedono la messa a punto di modelli predittivi, e i possibili interventi di correzione alla rete, si otterrà un "piano fognario integrato" che permette una visione globale, individua gli interventi di correzione, le priorità e il costo, utili ai fini della pianificazione degli interventi.

Competenze richieste

Il dottorand* deve avere competenze di analisi dati, Data Mining o Machine Learning, per la creazione

di modelli predittivi delle serie storiche, preferibilmente con conoscenza e background di processi stocastici.

La conoscenza della lingua Inglese (lettura e scrittura) è indispensabile.

ENG:

Framework of the doctoral project:

The climate change provokes big emergencies and the fall of enormous amount of water in a little time that is hard to manage if not well addressed.

		<p>The sewer network of many areas in Italy are old and insufficient to manage big amounts of water. To afford this emergency we propose as a solution to study a predictive model of the sewer network response and the possible interventions to improve it with the help of the company that has knowledge of the specific domain and collaborates with the managers of the network itself (BriazAcque).</p> <p>Expected results</p> <p>With the deployment of the research project that includes the preparation of predictive models of the emergencies and the study of the possible interventions, we will obtain a sewer integrated plan that will permit a global view of the factors that determine the emergencies, allows the interventions on the network with a priority and their costs, useful for the planning and the management of the emergencies.</p> <p>Required skills</p> <p>The candidate should have competences of data analysis, Data Mining. or Machine Learning, for the creation of predictive models of time series, with preference knowledge on stochastic processes.</p> <p>Knowledge of English (reading and writing) is mandatory.</p>
Stefano Ferraris	<p>ITA: Analisi di big data relativi alle risorse idriche sotterranee</p> <p>ENG: Groundwater big data analysis</p>	<p>ITA</p> <p>Inquadramento del contesto/problema</p> <p>La prevenzione di eventi critici come la siccità agricola e potabile richiede un uso estensivo di dati di monitoraggio ambientale. Oggi è disponibile una quantità impressionante di dati</p>

e una varietà di strumenti statistici e informatici che possono essere applicati per lo studio di tali big data.

Risultati attesi

Per il Piemonte disponiamo di informazioni nello spazio e nel tempo. Tali dati comprendono più di vent'anni di livello di acque sotterranee, portate fluviali, dati meteorologici di ARPA Piemonte e dati satellitari.

Unendo modelli matematici specifici e un approccio di machine learning si prevede di valutare quantitativamente le risorse idriche disponibili in un contesto di cambiamento climatico. Inoltre, l'utilizzo di tali risultati suggerirà azioni di prevenzione al fine di limitare i rischi idrogeologici legati ai cambiamenti climatici.

Competenze richieste

Background di fisica e matematica (master in matematica, fisica o ingegneria). Lingua inglese.

ENG

Framework of the doctoral project

Prevention of critical events such as agricultural and drinking water droughts require an extensive use of environmental monitoring data. Today, an impressive amount of data are available and a variety of statistical and computer science tools can be applied for the study of these big data.

Expected results

For the Piemonte region, we possess information in space and time. These data include more than twenty years of groundwater level, river discharge, meteorology data of ARPA Piemonte and satellite data.

		<p>Joining specific mathematical models and a machine learning approach it is expected to quantitatively evaluate the water resources in a context of climate change. Furthermore, the use of such results will suggest prevention actions in order to limit the hydrogeological risks related with climate changes.</p> <p>Required skills</p> <p>Physics and maths background (master in mathematics, physics, or engineering). English language.</p>
Stefano Ferraris	<p>ITA: Modellazione del bilancio idrico di bacini idrografici piemontesi</p> <p>ENG: Catchment water balance modeling for river flood prevention</p>	<p>ITA</p> <p>Inquadramento del contesto/problema</p> <p>Le inondazioni sono importanti conseguenze del cambiamento climatico e la prevenzione di tali eventi catastrofici è diventata un compito primario. L'attuale disponibilità di big data è un aiuto importante per lo sviluppo di modelli realistici o per l'utilizzo di metodi di machine learning. I modelli matematici possono quindi consentire di prevedere possibili rischi e suggerire interventi sul territorio quando ci si concentra su singoli bacini idrografici.</p> <p>Risultati attesi</p> <p>Il progetto di dottorato mira a utilizzare i dati di monitoraggio ambientale disponibili per la prevenzione di eventi critici come alluvioni in aree specifiche del Piemonte. Oltre all'utilizzo dei dati disponibili nello spazio e nel tempo, la ricerca comprenderà lo sviluppo di una metodologia a scala di bacino (area di ordine 100 km²). Tale scala è importante per una corretta gestione del territorio e lo sviluppo di modalità adeguate di organizzazione dei dati acquisiti saranno uno degli obiettivi del progetto, con particolare riferimento al bacino della Regione Piemonte. L'utilizzo delle informazioni estratte dai dati consentirà di sviluppare la seconda parte della ricerca, che prevede lo sviluppo di modelli.</p>

Il funzionamento idrologico e idraulico di un bacino idrografico in Piemonte sarà studiato attraverso metodi statistici nel dominio del tempo o della frequenza, nonché con metodi di apprendimento automatico e approccio fisico. Particolare enfasi sarà dedicata all'analisi delle piene a scala di bacino con la simulazione di eventi piovosi potenzialmente critici, con e senza l'adozione di tecniche di mitigazione/prevenzione.

I risultati, dopo essere stati implementati in WebGIS, consentiranno un'applicazione diretta da parte di decisori, amministratori locali e personale tecnico anche in altri bacini dell'arco alpino.

Competenze richieste

Background di fisica e matematica (master in matematica, fisica o ingegneria). Lingua inglese.

ENG

Framework of the doctoral project

Floods are important consequences of climate change and prevention of such catastrophic event has become a primary task. The current availability of big data is an important help for the development of realistic models or for the use of machine learning methods. Mathematical models may therefore allow to forecast possible risks and suggest interventions in the territory when one focus on a single catchment.

Expected results

This doctoral project aims to use of available environmental monitoring data for prevention of critical events such as floods in specific areas of Piemonte. Besides the use of available data in space and time, the research will include the development of a methodology at the catchment scale (order 100 km² area). Such scale is important for a correct land management and suitable methods for organizing the acquired data will be one of the goals

		<p>of the project, with special reference to a Piemonte catchment. The use of information extracted from the data will allow to develop the second part of the programmed research with the development of models.</p> <p>The hydrological and hydraulic functioning of a catchment in Piemonte will be studied through statistical methods in time or frequency domain as well as with machine learning methods together with physical approach. Particular emphasis will be devoted to the analysis of flooding at basin scale with the simulation of potentially critical rain events with and without adoption of mitigation / prevention techniques.</p> <p>The results, after being implemented with WebGIS, will allow direct application by decision makers, local administrators and technical staff also in other catchments in the Alps.</p> <p>Required skills</p> <p>Physics and maths background (master in mathematics, physics, or engineering). English language</p>
Cristina Zucca	<p>ITA: Sviluppo di un sistema di monitoraggio energetico aziendale</p> <p>ENG: Development of a company energy monitoring system</p>	<p>ITA</p> <p>Inquadramento del contesto/problema</p> <p>Il risparmio energetico e l'utilizzo di energie rinnovabili giocano un ruolo fondamentale nella ricerca di uno sviluppo aziendale sostenibile. Questo richiede che, da una parte le aziende si organizzino per l'utilizzo di energia rinnovabile e dall'altro individuino metodi di produzione volti al risparmio energetico.</p> <p>La multinazionale Agilent nella sede torinese si sta muovendo su entrambi questi fronti. Questo progetto di dottorato è in parte finanziato da Agilent ed è rivolto a un ripensamento dei processi produttivi introducendo il vincolo di minimizzare il consumo energetico. Occorre individuare i processi a maggior consumo energetico e studiare come ridurre i consumi o come ottimizzare i tempi di richiesta di energia, evitando momenti di picco energetico.</p>

Il progetto partirà dall'analisi statistica di dati esistenti a cui verranno affiancati i dati provenienti da nuovi sensori opportunamente collocati.

Risultati attesi

Il target finale sarà la costruzione di modelli matematici predittivi del consumo energetico dei singoli macchinari e delle linee di produzione. Tali modelli serviranno a ridurre i consumi e ottimizzare la richiesta energetica.

Competenze richieste

Background di fisica e matematica (master in matematica, fisica o ingegneria). Conoscenza del calcolo delle probabilità e della statistica (processi stocastici e serie storiche). Lingua inglese.

ENG

Framework of the doctoral project

Energy saving and the use of renewable energies play a fundamental role in the search for sustainable business development. This requires companies, on the one hand, to organize themselves for the use of renewable energy and, on the other, to identify production methods aimed at saving energy.

The multinational Agilent (Turin) is moving on both these fronts. This PhD project is partly funded by Agilent and is aimed at rethinking production processes by introducing some constraints to minimize energy consumption. It is necessary to identify the most energy intensive industrial processes and to study how to reduce energy consumption or how to optimize energy demand times, avoiding moments of peak energy.

The project will start from the statistical analysis of existing data which will be flanked by data from new sensors conveniently located.

		<p>Expected results</p> <p>The final target will be the construction of predictive mathematical models for the energetic consumption of individual machines and production lines. These models will be used to reduce consumption and to optimize energy demand.</p> <p>Required skills</p> <p>Physics and maths background (master in mathematics, physics, or engineering). English language. Expertize in probability and statistics (stochastic processes and time series) are appreciated. English language</p>
Alessandra De Rossi	<p>ITA</p> <p>Strumenti avanzati di Intelligenza Artificiale, Machine Learning e Data Science per la modellizzazione e la previsione del traffico</p> <p>ENG</p> <p>Advanced Artificial Intelligence, Machine Learning and Data Science Tools for Modeling and Prediction of Traffic Scenarios</p>	<p>ITA</p> <p>Inquadramento del contesto/problema</p> <p>L'obiettivo del Progetto è quello di analizzare e costruire nuovi strumenti dell'Intelligenza Artificiale, del Machine Learning e della Data Science che saranno implementati nell'ambiente di calcolo MATLAB. L'applicazione di questi strumenti computazionali servirà a progettare un modello del traffico atto anche alla previsione degli spostamenti nel territorio urbano e suburbano dando un contributo allo studio e alle ricerche nel campo della mobilità sostenibile.</p> <p>Le ricerche verteranno principalmente su tre filoni: Support Vector Machine, kernel da usare nel Machine Learning, Graph Machine Learning. Tali strumenti saranno poi applicati al problema della modellazione e previsione di scenari del traffico. Il Progetto è in collaborazione con MathWorks Srl.</p> <p>Risultati attesi</p>

Per quanto riguarda i metodi e gli algoritmi numerici riguardanti Machine Learning e Data Science sono attesi i seguenti risultati:

- ottimizzazione della scelta dei parametri in kernel usati nelle Support Vector Machine e negli algoritmi di trattamento di dati e segnali su grafi per la classificazione;
- aumento dell'efficienza di algoritmi di Support Vector Machine locali;
- miglioramento delle tecniche computazionali usate per l'elaborazione di grandi quantità di informazioni e segnali sui grafi;
- parallelizzazione dei programmi MATLAB di Machine Learning implementati.

Questi metodi e algoritmi efficienti saranno utilizzati nello sviluppo del Progetto sulla modellazione e previsione del traffico. Questo prevederà come risultati attesi:

- analisi di data set, processamento e visualizzazione dei dati;
- selezione delle soluzioni algoritmiche più utili per il trattamento dei dati;
- formulazione matematica e algoritmica del problema;
- uso di MATLAB per implementare la soluzione algoritmica;
- test dei metodi e codici sui dati, studio della loro efficienza, test del modello su dati reali.

Competenze richieste

Ottima conoscenza della lingua inglese; ottima conoscenza dell'ambiente Matlab; buone basi di Analisi Numerica e Statistica.

ENG

Framework of the doctoral project

The aim of the PhD Project is to build and analyze new tools of Artificial Intelligence, Machine Learning and Data Science that will be implemented in the MATLAB software. The application of these computational tools will serve to design a traffic model also suitable for the prediction of movements in the urban and suburban territories, giving a contribution to the study and research in the field of sustainable mobility. The research will focus mainly on three topics: Support Vector Machines, kernels to be used in Machine Learning applications, Graph Machine Learning. These tools will then be applied to the problem of modeling and design of traffic scenarios. The PhD Project is in collaboration with MathWorks Srl.

Expected results

With regard to the numerical methods and algorithms concerning Machine Learning and Data Science, the following results are expected:

- optimization of the choice of kernel parameters used in Support Vector Machines and in algorithms for processing data and signals on graphs for classification;
- increase the efficiency of local Support Vector Machine algorithms;
- improvement of computational techniques used for the processing of large amounts of data and signals on graphs;
- parallelization of the implemented MATLAB codes.

These efficient methods and algorithms will be used in the development of the Project on traffic modelling and scenarios. The following results are expected:

- data set analysis, data processing and visualization;
- selection of the most useful algorithmic solutions for data processing;
- mathematical and algorithmic formulation of the problem;
- use of MATLAB to implement the algorithmic solution;
- testing of methods and codes on data, study of their efficiency, testing of the model on real data.

		<p>Required skills</p> <p>Excellent knowledge of English language; excellent knowledge of the Matlab environment; good abilities in Numerical Analysis and Statistics.</p>
GIUSEPPE COSTA	<p>ITA</p> <p>L'importanza delle disuguaglianze sociali in relazione ai pericoli derivanti dal cambiamento climatico. Un caso di comparazione fra Piemonte e Campania</p> <p>ENG</p> <p>The importance of social inequalities with respect to the climate change dangers. A case of comparison between Piedmont and Campania</p>	<p>ITA</p> <p>Inquadramento del contesto/problema</p> <p>Si prevede che gli eventi estremi di alte temperature siano i principali fattori responsabili di ulteriori morti a causa dei cambiamenti climatici. Gli aumenti più gravi delle temperature sono previsti in Europa, con il maggior riscaldamento nell'area del Mediterraneo. Le condizioni climatiche future, combinate con la persistente urbanizzazione di queste regioni e la distribuzione spaziale delle città, influenzano la capacità di resilienza dei territori, esponendo la popolazione ad un maggior rischio di mortalità. Un modello accurato che consideri la relazione complessa fra temperature estreme e salute è infatti cruciale per comprendere gli impatti sulla salute della popolazione sotto le variazioni del cambiamento climatico. Nella presente ricerca, si esploreranno - insieme all'analisi del clima delle aree interessate - i fattori di vulnerabilità che insieme all'esposizione al pericolo climatico determinano il potenziale rischio climatico a livello di zona geografica oggetto di studio (e.g. città, provincia, regione).</p> <p>Risultati attesi</p> <p>In principio, l'analisi del contesto porterà ad una valutazione mirata che metterà in paragone il contesto della regione Campania in relazione al cambiamento climatico rispetto alla situazione Piemontese. Questo permetterà di dare informazioni di dettaglio su due fra le Regioni più grandi d'Italia, caratterizzate da climi diversi, seppur entrambi localizzate nell'area mediterranea. Da qui emergeranno, seppur qualitativamente, le principali criticità dei territori legate agli impatti del cambiamento climatico in relazione alle temperature estreme. A seguire, verranno fornite delle stime di rischio relativo (RR) e mortalità</p>

attribuibile (AF) sia per il caldo sia per il freddo, al fine di fornire una overview generale della relazione temperatura-mortalità all'interno delle due diverse regioni. Nel caso in cui i dati relativi ai fattori socioeconomici fossero disponibili all'interno due contesti di riferimento, il paragone fornito sarà ulteriormente di dettaglio, definendo anche le differenze di rischio in relazione alle disuguaglianze sociali della popolazione.

Competenze richieste

Costituiscono titolo preferenziale le lauree in matematica, fisica, ingegneria, informatica, statistica o pregressa attività di ricerca sulle tematiche del progetto

ENG

Framework of the doctoral project

Extreme high-temperature events are projected to be the primary drivers of additional deaths due to climate change. The most severe increases in temperatures are projected in Europe, with the greatest warming in the Mediterranean area. Future climate conditions, combined with the persistent urbanization of these regions and the spatial distribution of cities, affect the resilience of territories, exposing populations to a greater risk of mortality. Indeed, an accurate model that considers the complex relationship between temperature extremes and health is crucial to understanding population health impacts under variations in climate change. In the present research, we will explore - along with the analysis of the climate of the affected areas - the vulnerability factors that together with climate hazard exposure determine the potential climate risk at the level of the geographic area under study (e.g., city, province, region).

Expected results

In principle, the context analysis will lead to a targeted assessment comparing the context of the Campania region in relation to climate change with the situation in Piedmont. This

will give detailed information on two of the largest regions of Italy, characterized by different climates and population, although both located in the Mediterranean area. From here will emerge, albeit qualitatively, the main criticalities of the territories related to the impacts of climate change in relation to extreme temperatures. Following this, estimates of relative risk (RR) and attributable mortality (AF) will be provided for both hot and cold weather, in order to provide a general overview of the temperature-mortality relationship within the two different regions. In case data related to socio-economic factors were available within two reference contexts, the comparison provided will be further detailed, also defining the differences in risk in relation to social inequalities of the population.

Required skills

Degrees in mathematics, physics, engineering, computer science, statistics or previous research activities on the themes of the project are preferential qualifications.