POLITECNICO DI TORINO

UNIVERSITÀ DEGLI STUDI DI TORINO

Dottorato/PhD in Matematica Pura e Applicata / Pure and Applied Mathematics

(in convenzione con Università degli Studi di Torino e il Politecnico di Torino / jointly activated by Università degli Studi di Torino and Politecnico di Torino)

XXXV ciclo

(aggiornamento del 14 maggio 2019 / update on 14th May 2019)

Elenco delle tematiche per specifiche borse di Dottorato / List of research topics bound to PhD scholarships

1. Modeling, Simulation, Prediction, Control *(funded by Politecnico di Torino in the framework of the iniziative the incentive scheme for doctoral studies and of MIUR Dipartimenti di eccellenza for the Department of Mathematical Sciences "G. L. Lagrange" – n. 2 scholarships)*
2. Artificial Intelligence and Simulation for tackling complexity in engineering applications: deriving data driven models and reduced order models for fast evaluation. *(funded by Politecnico di Torino, Centro Interdipartimentale SmartaData in the framework of the iniziative the incentive scheme for doctoral studies and of MIUR Dipartimenti di eccellenza 2018-2022)*
3. Scientific impact and technological consequences of the ICH E9(R1) addendum to Statistical Principles for Clinical Trials on Choosing Appropriate Estimands and Defining Sensitivity Analyses in Clinical Trials. *(scholarship of Politecnico di Torino, DISMA, funded by GSK Vaccines s.r.l.)*
4. Learning and Optimization in Large-Scale Networks. *(funded by Politecnico di Torino in the framework of the iniziative the incentive scheme for doctoral studies and of MIUR Dipartimenti di eccellenza 2018-2022 - Machine Learning Reply)*

PhD in Pure and Applied Mathematics

(jointly activated by Università degli Studi di Torino and Politecnico di Torino)

Research Title: Artificial Intelligence and Simulation for tackling complexity in engineering applications: deriving data driven models and reduced order models for fast evaluation

Funded by	Centro Interdipartimentale SmartaData@Polito MIUR grant Dipartimenti di Eccellenza 2018-2022

Supervisor	Berrone Stefano stefano.berrone@polito.it Canuto Claudio claudio.canuto@polito.it Pieraccini Sandra sandra.pieraccini@polito.it

Contact	https://www.polito.it/geoscore http://smartdata.polito.it

Context of the research activity	In many engineering applications uncertainty on data and models is becoming a relevant issue, and many techniques are under development in order to provide fast and reliable tools to perform uncertainty quantification analyses. In several applications, the process under investigation requires performing simulations on a very complex model, or on a model affected by a large amount of uncertainty. Some situations also call for the sequential solution of well assessed models, but at different unknown scales or with unknown interactions. When a large number of simulations for such models is required, deep learning can be a useful tool to investigate a wide range of configurations or parameter values. Furthermore, in recent works deep learning has been applied, in some contexts, to derive data driven models able to capture properties at different scales, or an unknown interaction between different phenomena, or to derive equivalent parameters via data driven homogeization techniques. Several issues concerning the use of deep learning for these applications are still not sufficiently understood, and deserve a deep and wide investigation, as for example the choice of optimal

	learning sets, neural network performance testing and optimization. The project is therefore multidisciplinary, as it involves several fields of investigation, in the framework of a specific engineering application: deep learning, numerical simulation, HPC techniques, optimization, statistics.

Objectives	The research that the Ph.D. student will carry on will contribute to understand the role of Deep Learning in several engineering fields where simulations, complex models, and multi-scale phenomena play a fundamental role. A topic which deserves investigation concerns the optimal construction of learning sets in order to conform to the statistical properties of the quantities of interest; their construction strategies and stabilization methods are also worth being analyzed. Quite often a trained neural network displays a very reliable behavior in predicting quantities of interest, but in some situations a large discrepancy between the actual quantities and those predicted by the neural network appears. It is worth investigating the possibility to perform some “a posteriori” rigorous or statistical (but also heuristic) analysis to detect and fix these discrepancies. The main field of interest for the research will be geophysics; in particular the study of fractured porous media with respect to the hydro-mechanical properties. For instance, the PhD student will apply deep learning techniques to analyze the possibility to predict underground flows depending on field probabilistic properties of fractures and porosity of the surrounding rock matrix, possibly considering several distributions of different pore scales. Results obtained could be applied to the same context, but to different phenomena (mechanical behaviour, fault slips, fault reactivation) or to different contexts with similar multi-scale heterogeneous nature. This research will be relevant in all applications related to underground resources exploitation (Oil&Gas, geothermal applications) or protection (water), or to geological storage (CO2, nuclear waste). In particular, the long term research will aim at contributing in providing reliable tools for risk evaluation.

Skills and competencies for the development of the activity	The candidate should have a good background in computer programming, mathematical modeling, numerical methods and statistics.

PhD in Pure and Applied Mathematics

(jointly activated by Università degli Studi di Torino and Politecnico di Torino)

Research Title: Scientific impact and technological consequences of the ICH E9(R1) addendum to Statistical Principles for Clinical Trials on Choosing Appropriate Estimands and Defining Sensitivity Analyses in Clinical Trials

Funded by	GSK Vaccines s.r.l. DISMA

Supervisor	Referente aziendale: Marco Costantini marco.x.costantini@gsk.com Referente PoliTo: Mauro Gasparini mauro.gasparini@polito.it

Contact	http://www.disma.polito.it/la_ricerca/gruppi/statistica_biologia_e_industria

Context of the research activity	GSK studies must follow international guidance for clinical trials conduct and analysis. In such a dynamic and continuously changing environment, it is key for the company to be up to date on the most recent methodological developments and to be prepared to implement top-notch techniques. The statistics group led by professor Gasparini in Politecnico has the appropriate competence and experience to help companies and to ducate students on state-of-the-art biostatistics for clinical trials. The R1 addendum to the ICH E9 guideline on “Statistical Principles for Clinical Trials” is having a significant impact on the design, conduct, analysis and interpretation of clinical studies, with the introduction of the concept of *Estimand* which translates the trial objective into an unambiguous and well-defined definition of the treatment effect to be estimated. The role of the intercurrent events, such as treatment discontinuation, use of rescue medication, treatment switch and death, in the interpretation of the treatment effect will have to be carefully evaluated in the study protocol; the statistical analysis will need to be aligned with the

	estimand and a series of sensitivity analyses, targeting the same estimand under different assumptions, should be foreseen in order to assess the robustness of the results of the study, including for example the main estimates of the treatment effects and possible deviations from any underlying modeling assumptions.

Objectives	The PhD candidate will develop a framework to serve as a general guidance on estimands for all clinical teams that operate in the context of vaccine clinical trials. This will enhance trial planning and will introduce more appropriate and accurate statistical methods in line with the new ICH E9 R1 requirements. The outcome of this research will be immediately utilized within the ongoing COPD development program that will enter Phase III in 2021. The appropriate treatment of estimands, with special focus on data collected via e-Diaries, in COPD is expected to be a requirement from the Health Authorities and, in fact, GSK Pharma has already started to develop some methodologies in that disease area. Some academic publications are an expected outcome of the research. The journal aimed at are the usual outlet of biostatistical research such as Biometrics, Biometrical Journal, Statistical Methods in Medical Research, Pharmaceutical Statistics and the like.

Skills and competencies for the development of the activity	The student must have received a well-rounded education on mathematics and its applications. In addition, the candidate will have acquired around 30 credits specifically in probability, statistics and applied statistics.

PhD in Pure and Applied Mathematics

(jointly activated by Università degli Studi di Torino and Politecnico di Torino)

Research Title: Learning and Optimization in Large-Scale Networks

Funded by	Machine Learning Reply MIUR grant Dipartimenti di Eccellenza 2018-2022

Supervisor	Giacomo Como giacomo.como@polito.it Fabio Fagnani fabio.fagnani@polito.it Giorgia Fortuna g.fortuna@reply.it

Contact	http://staff.polito.it/giacomo.como/ http://calvino.polito.it/~fagnani/indexeng.html https://www.reply.com/en/

Context of the research activity	This PhD project will be a collaboration between the Department of Mathematical Sciences of Politecnico di Torino and Machine Learning Reply, a consulting company of the Reply group specialized in the development of solutions related to natural language, hot topic detection, optimization, forecasting, and anomaly detection. The research activity will be focused on the development of supervised and unsupervised machine learning methods for anomaly detection in multi-agent systems, of generative models for synthetic data generation from known but hardly modellable probability distributions, and of distributed optimization algorithms for the estimation of model parameters.

Objectives	The two main objectives of this project are: (i) to develop theory for generative models for both textual and graph-structured data; (ii) to apply the models above to problems of current interest such as fake news detection in social networks or text-generation. The goal is to develop distributed probabilistic optimization models that would help estimating probability densities able to generate certain predetermined data. The estimated models will be applied to generate new synthetic data and to determine whether new data is consistent with such distribution. Particular focus will be on graphical-structured data and evolutionary.

Skills and competencies for the development of the activity	The candidates are expected to have a solid mathematical background with a strong interest both in applications to large-scale optimization, learning, and networks.