

## PhD Course in Computer Sciences XXXVI cycle, a.y. 2020/2021

**n. 1 scholarship funded by the Departments, linked to the research project: “Methods and techniques for the quality of safety-critical software for railway systems”.**

### Abstract

*Although the specific details of the PhD work will be agreed with the selected candidate, the perspective topic shall be concerned with methods and techniques for the quality assurance of safety-critical system, with a particular focus on the embedded ERTMS applications that we are developing in the project that is providing the grant. For instance, the interesting topics may include testing methods for safety critical systems, automatic generation of test cases for safety critical systems, design patterns to guarantee quality and testability of safety-critical software, and so forth.*

**n.1 scholarships fundend by Fresenius Medical Care Italia S.p.A linked to the project:” *Dynamic Influence and Probabilistic Models to balance multiple, complex trade-offs in medical decision making for CKD patients*”.**

### Abstract

#### **Causal Models, Bayesian Networks and Influence Diagram Models for Decision Support Systems in Nephrology**

*There are several therapeutic options to treat patients with renal failure. The choice of the best option is a complex problem that combines many different factors (e.g., risk-benefit of the therapy, patient’s preferences, patient’s characteristics, availability of the resources, clinical practice, etc.). In this context, an intelligent system able to merge patient specific risk factors and other contextual information might help to take a more aware decision and increase the probability of better patient’s outcomes. Causal models are widely applied in medicine and in other fields for the ability to combine accuracy and explicability of the output. For this reason, the research project will be focused on the application of such paradigms to build a medical decision support system.*

## Posizioni con Percorso Executive

n.2 posizioni riservate ai dipendenti presso brumbrum S.p.a.

### Posizione n. 1

#### **Profilo**

*Computer Science, Computer Vision*

#### **Progetto**

*Computer Vision for Automotive*

*Costruzione di un algoritmo in grado di identificare i danni di carrozzeria presenti su una vettura e di classificarli su una scala di intensità a partire da fotografie*

#### **Formazione**

*Come parte dell'attività formativa, il candidato dovrà superare almeno 2 esami sui temi inerenti al progetto.*

#### **Argomento della Tesi**

*Nella seguente proposta di tesi si vuole dare applicazione a tecniche di Computer Vision all'interno del processo di ricondizionamento di una vettura che è destinata a essere rimessa in vendita come auto usata. Quando si acquista una vettura usata tramite canali B2B sia che si tratti di canali tradizionali (quindi fisici) che di canali online, è prassi del settore che ogni veicolo sia corredato da una perizia che ne rendiconti le condizioni di carrozzeria. Queste perizie possono essere strutturate, cioè redatte da società terze secondo formati standard che prevedono fotografie degli eventuali difetti e una descrizione degli stessi, sia non strutturate, composte cioè da sole fotografie senza ulteriori dettagli. Le perizie vengono valutate in fase di decisione di acquisto da un perito che in base alle immagini determina in linea di massima l'importo necessario al ricondizionamento. Una volta acquistata la vettura viene sottoposta ad una perizia dal vivo molto più accurata e viene determinato l'effettivo ammontare dei danni presenti. Per rendere scalabili entrambi i processi di perizia, pre e post acquisto, si vorrebbe costruire un algoritmo in grado di individuare difetti e/o danni di carrozzeria, di indicare quali parti (o pannelli) sono coinvolti e di dare un'indicazione dell'entità del danno in base alla classificazione standard del settore. Tale algoritmo permetterebbe di creare una pre-perizia a partire dal corredo fotografico delle vetture in vendita, sulla base della quale scremare le vetture più interessanti e di individuare con estrema precisione ogni difetto di carrozzeria dell'auto effettivamente acquistata a partire da fotografie di alta qualità scattate in un apposito set fotografico. Le difficoltà da superare sono notevoli, in particolare in virtù del fatto che tipicamente le fotografie della perizia preacquisto sono spesso di scarsa qualità. Inoltre, l'algoritmo richiesto non solo deve*

*individuare eventuali difetti, ma deve correttamente collocarli sulle parti della vettura interessate e deve altresì classificarne l' entità.*

## **Posizione n.2**

### **Profilo**

*Computer Science, Bayesian Methods*

### **Porgetto**

*Bayesian Methods for Marketing Analytics*

*Esplorazione delle tecniche Bayesiane per la soluzione di problemi di ottimizzazione e attribuzione su tematiche di marketing digitale.*

### **Formazione**

*Bayesian Methods for Marketing Analytics*

*Come parte dell ' attività formativa, il candidato dovrà superare almeno 2 esami sui temi inerenti al progetto*

### **Argomento della tesi**

*Nella seguente proposta di tesi si vogliono investigare diversi problemi di ottimizzazione e attribuzione all' interno della digital marketing analytics con tecniche Bayesiane.*

*Tecniche e algoritmi fondati su concetti e risultati propri della Statistica Bayesiana stanno vivendo un momento di grande popolarità a traino della grande diffusione di metodi di Machine e Deep Learning, in cui trovano vasta applicazione. Anche settori della marketing analytics, come la conversion rate optimization e il paid marketing , sono toccati da questa ondata di popolarità e cominciano a nascere idee di applicazioni di metodi Bayesiani , anche se poi trovano difficile applicazione soprattutto per via di gap culturale degli operatori del settore.*

*Ad esempio , alcuni dei campi in cui tecniche Bayesiane potrebbero trovare fertile applicazione sono:*

*1. A/B test ing: la procedura standard per confronto statistico tra le performance di due o più versioni del sito per identificare quella che garantisce maggiori coefficienti di conversione, fa riferimento a tecniche di inferenza frequentista. Tali tecniche richiedono un' elevata numerosità del campione per innescare i criteri di applicabilità del Teorema del Limite Centrale e per garantire la giusta significatività e potenza. Tecniche Bayesiane possono invece ridurre la numerosità del campione, riducendo il cosiddetto costo di test, e si prestano ad ulteriori sofisticazioni ad esempio attraverso algoritmi Multi Armed Bandit.*

*2. Ottimizzazione di campagne paid: i display pubblicitari sui social e/o su Google sono tipicamente gestiti con aste real time e prevedono quindi l' individuazione di un target, di un bid e di un budget giornaliero. La soluzione al problema dell'ottimizzazione simultanea di più campagne (target diversi) sotto il vincolo di un unico budget giornaliero, in modo da garantire il miglior ritorno dell' investimento, pone numerose sfide e ben si presta ad essere individuata attraverso tecniche di Reinforcement Learning.*

*3. Modello di attribuzione: è assai raro che un cliente arrivi al sito attraverso un solo canale di marketing , diventa quindi importante per valutare l'efficacia di ogni singolo canale riuscire a valutare nel modo più accurato possibile il ruolo e il ritorno sull' investimento. Modelli Generativi e Reti Causali sono solo alcune delle tecniche Bayesiane che possono essere impiegate per affrontare questo problema.*

*L' obiettivo della tesi è quello di individuare una soluzione in ambiente Bayesiano per ciascuna delle problematiche evidenziate e di portare la soluzione di almeno un problema allo stato dell' arte.*

**n.2 posizioni riservate ai dipendenti presso Social Things Srl, vincolate al progetto: "Optimized Social Artificial Intelligence"**

**Profilo**

*Ricercatore in ambito Intelligenza Artificiale*

**Esami obbligatori** *Il candidato seguirà dei corsi come previsto dal regolamento della scuola di dottorato in informatica.*

**Abstract**

*The project is aimed at defining the conceptual, systemical and logical design of an Explainable and Neural Network Recommender System solution able to provide users with empowered intelligent suggestions so as to improve their actions, interactions and decision-making processes in complex socio-technical systems thanks to social friend information.*

*The approach will be primarily validated and tested in an educational learning management system: "WhoTeach", so as to improve its intelligence in guiding users throughout the creation or the exploitation of dynamic and customized learning and training paths. Nevertheless, the Explainable Recommender System to propose could also be applied in other socio-technical systems: then, given the primary test of the system feasibility in the educational environment, we will also try to identify other possible applications to domains where this paradigm could fit and help. Artificial Intelligence (AI) is facing a fast evolution, as regards both methodology and techniques. Nevertheless, most AI and Machine Learning (ML) models appear to be as a black-box. Explainable AI is a quite recent approach in the AI field, aimed at providing intelligent systems with the ability to explain the decisions, recommendations, predictions or actions made by it and the process through which they are made. In particular, Recommender Systems have become extremely appealing in all fields, and especially for all Technology Enhanced Learning (TEL) researches aimed to design, develop and test technical innovations which support and enhance learning and teaching practices of both individuals and organizations. In this scenario the goal of the project is to exploit the new emerging paradigm of Explainable and Neural Network Recommender Systems, leveraging social information to provide (social) explanations in order to supply users with his/her friends' public interests as explained recommendation. In this research we primarily apply this paradigm to our educational platform called "WhoTeach", an innovative and original system to integrate knowledge discovery, social networks analysis, and educational services. In particular, we provide "WhoTeach" environment with optimized Social Explainable Recommendations oriented to design new teachers' programme and courses.*

## **n.2 posizioni riservate ai dipendenti presso Intesa San Paolo**

### **Posizione 1**

#### **Title**

*ML-based pipelines and frameworks in the financial sector*

*Ricerca su tematiche di Explainable AI (XAI) e Fairness all'interno dei modelli di Machine Learning. Oggi giorno e nel prossimo futuro sarà sempre più richiesto che i modelli sviluppati, spesso tramite black box, soddisfino criteri di trasparenza e di non discriminazione.*

*Durante il progetto di ricerca si vuole sviluppare in questo senso algoritmi e tool che supportino:*  
*1) l'investigazione all'interno dei modelli sviluppati, per permettere una maggior comprensione delle logiche utilizzate dall'algoritmo, in modo da permettere un miglior fine tuning dell'algoritmo e una migliore interazione con chi deve utilizzare il modello.*

*2) l'identificazione di problemi di discriminazione all'interno delle previsioni del modello e la rimozione di queste.*

*Questo tipo di ricerca ha sia impatti sociali che economici per le aziende che sviluppano modelli di classificazione. Riuscire a produrre un modello che sia trasparente, affidabile e non discriminatorio, protegge le aziende da problemi di natura regolatoria, aiuta ad aumentare la reputazione e la fiducia da parte dei propri clienti e permette l'ottimizzazione dei processi aziendali. Dall'altra parte, anche a livello sociale, ottenere una spiegazione trasparente e condivisibile della motivazione per cui si è visti ottenere o no una determinata concessione, per mezzo di un algoritmo, aiuta ad aumentare l'accettazione di questo.*

### **Posizione 2**

#### **Title**

*Explainable AI & Fairness*

#### **Introduction**

*Artificial intelligence (AI) and Machine Learning (ML) are being extensively adopted for a number of purpose and diverse range of applications across the financial services industry. Given high volume, accurate historical records and quantitative nature of the field, data-driven self-learning methods offer unprecedented opportunities for applied research in the field.*

#### **Aims and Objectives**

*The project will concentrate on the design and realisation of ML-based pipelines and frameworks in the financial sector. The first objective consists in identifying and evaluating the appropriateness of different methodologies that can be applied to approach a specific use case. The central idea is to explore different sub-fields of AI and ML in order to give a contribution in the field of AI and ML applied to financial data, empowering their ability in supporting human decision making. Among the possible specific scenarios, supervised learning algorithms can be applied to classification and prioritization use cases where it is crucial not to just achieve high performance metrics but to maximise the business impacts and benefits. Unsupervised techniques such as identification of anomalies and pattern recognition can be applied in anti-money laundering or fraud detection processes, where traditional rule-based systems are not promising given the dynamical nature of such phenomena and the unbalanced structure of data. The same problems could be approached using Network Analysis, where graphs bearing from interaction data can be analysed. Another research field that is recently growing concerns structured and unstructured text mining, including sentiment analysis on external data. Lastly, considering that the algorithm result and accuracy are*



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

*not less important than the design and development process itself, the aim is to focus on the statistical and theoretical principles beyond algorithms to better understand potentialities and limitations of specific methodologies.*