

Poster Session Program

Tuesday, July 5 5:30-6:30 PM

Poster Nb	Name	Ecole/Company	Poster title
1	François AMAT	Hi! PARIS PhD at IP PARIS - Telecom Paris	Rule mining in tabular databases
2	Jean-Rémy CONTI	IP PARIS - Telecom Paris	Mitigating Gender Bias of Pre-Trained Face Recognition Models with an Ethical Module
3	Simon DELARUE	Hi! PARIS PhD at IP PARIS - Telecom Paris	Sparse Graph Neural Networks
4	Andréa EPIVENT	Hi! PARIS PhD at IP PARIS - ENSAE	Algorithms on Digital Platforms: Empirical & Experimental Evidence
5	Qi GAN	IP PARIS - Telecom Paris	Evaluation of sprinter gestures with RGB videos
6	Luca HEISING	Tilburg University, Netherlands	(X)AI for treatment error detection in radiotherapy: on the challenge of explaining multidimensional AI decisions with high precision.
7	Jing NIU	HEC Paris	Beyond follower size--investigating influencer's endorsing content creation strategies
8	Dilia Carolina OLIVO	IP PARIS - Telecom Paris	Graphs for XAI : Leveraging Graph Structure for Explainable Automation of Fraudulent Wire Transfer Detection
9	David PERERA	Hi! PARIS PhD at IP PARIS - Telecom Paris	Detection and Classification of Sound Events
10	Marie REINBIGLER	Hi! PARIS PhD at IP PARIS - Telecom SudParis	Frugal deep learning for neuromuscular tissue analysis for tomorrow's gene therapies
11	Yufei SHEN	Hi! PARIS PhD at HEC Paris	The Value of Privacy in Online Advertising

Poster Abstracts

For more details on the posters: <https://www.summerschool.hi-paris.fr/program/poster-session/>

Poster 1 - Rule mining in tabular databases

Speaker: François AMAT, Hi! PARIS PhD at IP PARIS - Telecom Paris, francois.amat@telecom-paris.fr

Abstract:

Poster 2 - Mitigating Gender Bias of Pre-Trained Face Recognition Models with an Ethical Module

Speaker: Jean-Rémy CONTI, IP PARIS - Telecom Paris, jean-remy.conti@telecom-paris.fr

Abstract: In spite of the high performance and reliability of deep learning algorithms in a wide range of everyday applications, many investigations tend to show that a lot of models exhibit biases, discriminating against specific subgroups of the population (*e.g.* gender, ethnicity). This urges the practitioner to develop fair systems with a uniform/comparable performance across sensitive groups. In this work, we investigate the gender bias of deep Face Recognition networks. In order to measure this bias, we introduce two new metrics, BFAR and BFRR, that better reflect the inherent deployment needs of Face Recognition systems. Motivated by geometric considerations, we mitigate gender bias through a new post-processing methodology which transforms the deep embeddings of a pre-trained model to give more representation power to discriminated subgroups. It consists in training a shallow neural network by minimizing a Fair von Mises-Fisher loss whose hyperparameters account for the intra-class variance of each gender. Interestingly, we empirically observe that these hyperparameters are correlated with our fairness metrics. In fact, a careful selection significantly reduces gender bias. This paper has been accepted at ICML 2022.

Poster 3 - Sparse Graph Neural Networks

Speaker: Simon DELARUE, Hi! PARIS PhD at IP PARIS - Telecom Paris, simon.delarue@telecom-paris.fr

Abstract: It has been showed that embedding techniques can lead to performance increases on downstream prediction tasks in several domains, while overcoming high computational costs related to analytics. Applied to graphs, these embedding methods allow to build vectorized feature spaces that encode nodes structural properties. In the last few years, deep learning techniques that showed great performance in numerous research areas including computer vision and language analysis, were adapted to form Graph Neural Networks (GNN), which extend neural networks models specifically for graphs. We propose an efficient implementation of GNNs in Scikit-network, a Python package inspired by scikit-learn for the analysis of large graphs. By relying on the sparse nature of real-world networks, Scikit-network provides state-of-the-art graph algorithms with limited number of dependencies (Numpy and Scipy).

Poster 4 - Algorithms on Digital Platforms: Empirical & Experimental Evidence

Speaker: Andr ea EPIVENT, Hi! PARIS PhD at IP PARIS – ENSAE, andrea.epivent@ensae.fr

Abstract: Internet platforms rely on artificial intelligence based tools to shape the content users see online. In particular, platforms may use algorithms to automate their price-setting decisions, recommend content to their users or improve their ad targeting.

While algorithms offer many benefits to platforms and users, there are concerns about potential negative consequences they may also bring. For instance, competition authorities and scholars are currently investigating whether pricing algorithms could learn to tacitly collude, users also question how their recommendations are produced.

The aim of my research is to investigate the impact algorithms have on internet users and how we could give them more power. In doing so, my first project aims at studying the link between the use of pricing algorithms by third-party sellers on Amazon and price dispersion on the platform. In my second project, I plan to conduct an experiment to study whether giving internet users the opportunity to “shop” for their preferred recommendation algorithms on digital platforms can improve their welfare.

Poster 5 - Evaluation of sprinter gestures with RGB videos

Speaker: Qi GAN, IP PARIS - Telecom Paris, qi.gan@telecom-paris.fr

Abstract: Gesture analysis has played an important role in improving the performance of athletes and monitoring the rehabilitation of patients with injury. However, to acquire precise movement information of a subject normally requires wearing sensors to help measurement. The convenience, efficiency and cost become major concerns of this method. Thanks to the invention of convolutional deep neural networks, the research in computer vision has bloomed in recent years, which makes it possible to capture skeleton models of human motion with an accuracy of around 20mm on public datasets. On the other hand, modern artificial intelligence technology has proved its ability in exploring key features in sports performance prediction. This project is a natural intersection of gesture analysis and artificial intelligence with application to sports, especially sprint.

The goal of the PhD thesis is to explore movement parameters to help evaluate the gestures of athletes during sprint using monocular RGB video of their training. To achieve this goal, the research is conducted in three steps: 1) Developing an algorithm to better estimate sprinter’s pose during running, 2) Extracting parameters of the sprinter in one run based on estimated pose, 3) Exploring the relationship between the parameters and the performance of sprinters, so that to help coach improve techniques of the athletes.

Poster 6 - (X)AI for treatment error detection in radiotherapy: on the challenge of explaining multidimensional AI decisions with high precision

Speaker: Luca HEISING, Tilburg University, Netherlands, l.m.heising@tilburguniversity.edu

Abstract: AI has become a ubiquitous topic with many different applications. In recent years, healthcare has been the main focus for many AI researchers with success as a myriad of AI applications have been introduced to aid clinicians in their work. One major drawback is, however, the lack of transparency and explainability of AI models. For this reason, Explainable Artificial Intelligence (XAI) has been emerging in the field of AI, also in healthcare. This poster presents some of the challenges faced when exploring XAI for healthcare, and more specifically, for radiotherapy due to the need for high precision and complexity of the data.

Poster 7 - Beyond follower size--investigating influencer’s endorsing content creation strategies

Speaker: Jing NIU, HEC Paris, jing.niu@hec.edu

Abstract: Social media influencers are playing essential roles in brands' communication with target consumers. Optimizing social media campaigns with influencers remains a hot topic in recent years. While most research has focused on influencer selection based on their network degrees, expertise, or other characteristics, this paper investigates influencers' content creation strategy(ies) once they have been selected by the company. We propose that influencers utilize more than one strategy to create content for a campaign based on their multiple roles (a brand agent, a peer to the followers, and a self-branding agent). Our empirical approach employs quantified content data of more than 2400 sponsored Instagram posts to conduct a latent cluster analysis. We document distinct communication strategies influencers used in practice and provide some evidence that an influencers' strategy choices impact the return on investment of one campaign regardless of their follower sizes and experience level. These findings suggest that marketing managers should consider influencers' communication strategies in paid campaigns when developing assessment criteria.

In collaboration with with Kristine De Valck and Peter Ebbes

Poster 8 - Graphs for XAI : Leveraging Graph Structure for Explainable Automation of Fraudulent Wire Transfer Detection

Speaker: Dilia Carolina OLIVO, IP PARIS - Telecom Paris, dilia.olivo@telecom-paris.fr

Abstract:

Poster 9 - Detection and Classification of Sound Events

Speaker: David PERERA, Hi! PARIS PhD at IP PARIS - Telecom Paris, david.perera@telecom-paris.fr

Abstract: Invariant-based learning is a promising approach in deep learning. Among other benefits, it can mitigate the lack of diversity of available datasets and increase the interpretability of trained models. Practitioners often use a consistency cost penalizing the sensitivity of a model to a set of carefully selected data augmentations. However, there is no consensus about how these augmentations should be selected. We have studied the impact of several types of augmentations for training a given model. We consider the task of sound event detection and classification for our experiments. In particular, we show that transformations operating on the internal layers of a deep neural network are beneficial for this task.

Poster 10 - Frugal deep learning for neuromuscular tissue analysis for tomorrow's gene therapies

Speaker: Marie REINBIGLER, Hi! PARIS PhD at IP PARIS - Telecom SudParis, marie.reinbigler@telecom-sudparis.eu

Abstract: Généthon, a non-profit research institute, is a pioneer in gene therapy vectors development that can benefit to several hundred neuromuscular diseases of genetic origin. Currently pre-clinical studies are conducted to gain precise knowledge of muscle physiology and disease based on the analysis of histological sections, i.e. slices of organs observed by microscopy. This work aims at processing these large histological slices by exploiting deep learning to analyze and classify the neighborhoods of pathological areas of interest using a 'frugal' computing platform, i.e., inexpensive local hardware resources to enable access democratization as well as total control on data privacy. The core of our approach is to build a scaling processing architecture composed of a pair of differentiable blocks. First, an image-based block will extract a segmentation of the slices into related components. In a second step, a graph-based approach, built on the set of previously detected components, will be deployed to perform the classification. At each step, our goal will be to maximize the computational performance of the execution medium, in order to make our approach compatible with modest computational resources - an open problem in the context of deep learning, which is particularly computationally and energy intensive.

Poster 11 - The Value of Privacy in Online Advertising

Speaker: Yufei SHEN, Hi! PARIS PhD at HEC Paris, yufei.shen@hec.edu

Abstract: In the industry of online advertising, publishers rely on the collection of consumer personal data for ad targeting, which has raised widespread concern about privacy intrusion. While recent data regulations in China, Europe, and the U.S. require giant tech firms to give users more privacy control over their personal data, it is still not clear how much is the value of consumer privacy to various stakeholders in the online advertising market. Using a unique user-level panel dataset collected from a leading social media platform and publisher in China, we find that after opting out of ad targeting, users would be delivered ad impressions that generate 22.64% less revenue for the publisher, compared to their opt-in peers. More importantly, we show that ads from small and medium-sized advertisers, rather than larger advertisers, would be significantly less likely to display to opt-out users. These results indicate that privacy protection has a price of market inefficiency and using personal data in online advertising plays an important role in facilitating product information diffusion for small and medium-sized enterprises.