

# **diiP Summer School 2024 - dSDS**

## **Rapport sur les contributions**

ID de Contribution: 1

Type: **Non spécifié**

# AI for Research

*lundi 10 juin 2024 14:30 (1h 30m)*

**Auteur principal:** BECHERINI, Yvonne (Université Paris Cité)

**Orateurs:** PALPANAS, Themis (LIPADE - Paris Descartes University); BECHERINI, Yvonne (Université Paris Cité)

ID de Contribution: 2

Type: **Non spécifié**

## Large Language Models (LLMs)

*lundi 10 juin 2024 16:30 (1h 30m)*

Large language models have become ubiquitous. This lecture introduces to language modelling and large language models. We will discover what they are, where they come from and the primary motivations behind their design. We then provide an overview of the properties of these models when trained at the current scale of very large language models.

If time remains, we introduce the problematic of explaining their behaviour.

**Orateur:** Prof. CRABBÉ, Benoît (Université Paris Cité)

ID de Contribution: 3

Type: **Non spécifié**

## Supervised Learning

*mardi 11 juin 2024 09:00 (1h 30m)*

**Orateur:** Prof. IOSIFIDIS, Alexandros (Aarhus University)

ID de Contribution: 4

Type: **Non spécifié**

## **Supervised Learning, Hands-on session**

*mardi 11 juin 2024 11:00 (1 heure)*

**Orateur:** Prof. IOSIFIDIS, Alexandros (Aarhus University)

ID de Contribution: 5

Type: **Non spécifié**

## **AI in Medicine/Biology**

*mardi 11 juin 2024 12:00 (1 heure)*

**Orateur:** Prof. ASSIÉ, Guillaume (Institut Cochin, Inserm CNRS Université de Paris Cité)

ID de Contribution: 6

Type: **Non spécifié**

## Representation Learning

*mardi 11 juin 2024 14:30 (1h 30m)*

**Orateur:** Dr RAITOHARJU, Jenni (University of Jyväskylä)

ID de Contribution: 7

Type: **Non spécifié**

## **Representation Learning - Hands-on session**

*mardi 11 juin 2024 16:30 (1 heure)*

**Orateur:** Dr RAITOHARJU, Jenni (University of Jyväskylä)



ID de Contribution: **8**

Type: **Non spécifié**

## Knowledge-guided Data Science

*mercredi 12 juin 2024 09:00 (1h 30m)*

This lecture presents an overview of knowledge-guided data science, a rising methodology in machine learning which fuses data with domain knowledge. We will present numerous case studies on this methodology to showcase how to unleash its potential in real-world data science applications.

**Orateur:** LIANG, Shen

ID de Contribution: 9

Type: **Non spécifié**

## **Knowledge-guided Data Science - Hands-on session (Basic)**

*mercredi 12 juin 2024 11:00 (2 heures)*

**Orateur:** LIANG, Shen

ID de Contribution: **10**

Type: **Non spécifié**

## **Language Models**

ID de Contribution: 11

Type: **Non spécifié**

## High-Dimensional Vector Similarity Search

*mercredi 12 juin 2024 14:30 (1h 30m)*

Very large amounts of high-dimensional data are now omnipresent (ranging from traditional multidimensional data to time series and deep embeddings), and the performance requirements (i.e., response-time and accuracy) of a variety of applications that need to process and analyze these data have become very stringent and demanding. In the past years, high-dimensional similarity search has been studied in its many flavors. Similarity search algorithms for exact and approximate, one-off and progressive query answering. Approximate algorithms with and without (deterministic or probabilistic) quality guarantees. Solutions for on-disk and in-memory data, static and streaming data. Approaches based on multidimensional space-partitioning and metric trees, random projections and locality-sensitive hashing (LSH), product quantization (PQ) and inverted files, k-nearest neighbor graphs and optimized linear scans. Surprisingly, the work on data-series (or time-series) similarity search has recently been shown to achieve the state-of-the-art performance for several variations of the problem, on both time-series and general high-dimensional vector data. In this talk, we will touch upon the different aspects of this interesting story, and present some of the state-of-the-art solutions.

**Orateur:** PALPANAS, Themis (LIPADE - Paris Descartes University)

ID de Contribution: 12

Type: **Non spécifié**

## An Overview of Anomaly Detection for Time Series

*jeudi 13 juin 2024 09:00 (1h 30m)*

Anomaly detection is an important problem in data analytics with applications in many domains. In recent years, there has been an increasing interest in anomaly detection tasks applied to time series. In this talk, we take a holistic view of anomaly detection in time series, starting from the core definitions and taxonomies related to time series and anomaly types, to an extensive description of the anomaly detection methods proposed by different communities in the literature. We will then present new benchmarks capturing diverse domains and applications for the purpose of evaluating anomaly detection methods. We will then conclude on Ensembling and Model Selection for time series anomaly detection, discussing different strategies applicable to automatically selecting the appropriate methods for a specific time series.

**Orateur:** Dr BONIOL, Paul (INRIA & ENS)

ID de Contribution: **13**

Type: **Non spécifié**

## AI in Industry

*jeudi 13 juin 2024 12:00 (1 heure)*

**Orateur:** Dr PONCET, Paul (Engie)

ID de Contribution: 14

Type: **Non spécifié**

## **Time Series Anomaly Detection in Practice - Hands-on session**

*jeudi 13 juin 2024 11:00 (1 heure)*

**Orateur:** Dr BONIOL, Paul

ID de Contribution: 15

Type: **Non spécifié**

## Graph Neural Networks

*jeudi 13 juin 2024 14:30 (1h 30m)*

**Orateur:** Prof. VAZIRGIANNIS, Michalis



ID de Contribution: 16

Type: **Non spécifié**

## **Graph Generative AI + Applications**

*jeudi 13 juin 2024 16:30 (1h 30m)*

**Orateur:** Prof. VAZIRGIANNIS, Michalis (École polytechnique, Palaiseau)

ID de Contribution: 17

Type: **Non spécifié**

# Generative adversarial networks and Active Learning

*vendredi 14 juin 2024 09:00 (1h 30m)*

**Orateur:** Prof. SAADALLAH, Amal (TU Dortmund University)

ID de Contribution: **18**

Type: **Non spécifié**

## **Generative adversarial networks and Active Learning - Hands-on session**

*vendredi 14 juin 2024 11:00 (1 heure)*

**Orateur:** Prof. SAADALLAH, Amal (TU Dortmund University)

ID de Contribution: **19**Type: **Non spécifié**

## AI in Particle Physics

*vendredi 14 juin 2024 12:00 (1 heure)*

Particle physics deals with gigantic machines, large quantities of experimental data and computer simulations, complex and lengthy theoretical calculations. It is the perfect playground to take advantage of machine learning algorithms. After a short introduction to high energy physics, this lecture will show how one can speed up steps like event generation or detector simulation, better measure parameters, improve classification of events as signal or background, or discover anomalies in data taking, through various applications of machine learning in this field.

**Orateur:** Dr COADOU, Yann (CPPM, Aix-Marseille Université, CNRS/IN2P3)