AI for Life Scientists: from Basics to Applications

Report of Contributions

Introduction to AI in Life Sciences

Contribution ID: 9

Type: not specified

Introduction to AI in Life Sciences

Thursday 11 April 2024 14:00 (1h 30m)

The talk will provide an introduction to Artificial Intelligence (AI) applications in life sciences. Beginning with fundamental concepts such as regression and classification tasks, it will explain the key metrics used to assess model performance and showcase specific applications in life sciences. The lecture will also address the challenges and limitations of AI models, like data scarcity, potential biases, and AI's 'black-box'nature. Additionally, it explores the integration of AI in life sciences, discussing its potential in advancing research and clinical solutions.

Presenter: PIRAUD, Marie (Helmholtz AI, Helmholtz Munich) **Session Classification:** Lecture Series

Supervised and Unsupervised Lear ...

Contribution ID: 10

Type: not specified

Supervised and Unsupervised Learning

Thursday 18 April 2024 14:00 (1h 30m)

This talk provides an overview of key concepts in machine learning, focusing on the distinctions and applications of Supervised and Unsupervised learning. It dives into various supervised model types, including linear, tree-based, and ensemble approaches and neural networks, but also touches on unsupervised methods like clustering and dimensionality reduction. The discussion encompasses the fundamentals of these models, their applications, and key differences, providing a comprehensive overview suitable for understanding various machine learning approaches.

Presenter: LOEWE, Axel (KIT, IBT) **Session Classification:** Lecture Series

Data Preprocessing

Contribution ID: 11

Type: not specified

Data Preprocessing

Thursday 2 May 2024 14:00 (1h 30m)

This talk will give a general overview of data analysis steps commonly preceding the application of actual AI modeling approaches. Possible challenges and solutions in handling real-world data will be illustrated with examples of omics datasets. After highlighting the importance of data representation, normalization, quality control and batch effect correction, the lecture will address how to deal with missing values, outliers, class imbalance and different data types. Furthermore, a main focus will be on dimensionality reduction techniques, including methods for feature transformation, feature aggregation and feature selection.

Presenter: GEORGII, Elisabeth (Helmholtz AI, Helmholtz Munich) **Session Classification:** Lecture Series

Diffusion Models

Contribution ID: 12

Type: not specified

Diffusion Models

Thursday 23 May 2024 14:00 (1h 30m)

Unfortunately, the lecture by Bastian Riek planned for 23 May has to be rescheduled. We will inform all participants registered for the lecture series by email about a new date.

Diffusion models are a novel class of generative machine learning models that have recently gained significant traction in different fields, including biomedical applications. Such models operate by *gradually* injecting additional noise into data, yielding a distribution of random noise. To generate new unseen instances, the random noise distribution is sampled and subjected to a denoising step. This process results in a new instance from the target distribution. In the context of biomedical applications, diffusion models hold promise for a wide range of different tasks, from drug discovery to medical imaging analysis. This talk provides a brief overview of their potential and their challenges for large-scale biomedical applications.

Presenter: RIECK, Bastian (Helmholtz Munich, Technical University of Munich)

Ethical Considerations for AI Models

Contribution ID: 13

Type: not specified

Ethical Considerations for AI Models

Thursday 27 June 2024 14:00 (1h 30m)

This lecture provides a comprehensive overview of the ethical landscape surrounding AI models. We will briefly cover the history of AI ethics to understand concepts particularly relevant in the life sciences, such as non-maleficence, bias, fairness, transparency, and responsibility. From there, we will look at life-science-relevant AI cases to discuss the potential real-world implications of these concepts. We will round up the session by discussing state-of-the-art tools to mitigate ethical challenges when creating AI in the life sciences.

Presenter: WILLEM, Theresa (Helmholtz Munich, Institute of History and Ethics in Medicine, Department of Clinical Medicine, TUM School of Medicine and Health)

Transformer Models

Contribution ID: 14

Type: not specified

Transformer Models

Thursday 13 June 2024 14:00 (1h 30m)

The rapid evolution of transformer-based large language models (LLMs) has revolutionized numerous fields, with their impact on genomics promising to unlock new frontiers in biology and medicine. In this lecture, we will cover basics of language models, explaining how they predict and interpret sequences of symbols, and how we process DNA sequences for computational processing. The core of the lecture will delve into transformer models, emphasizing their self-attention mechanisms and their ability to handle long-range dependencies in data, making them ideal for analyzing the complex sequences found in DNA and RNA. By the end of the lecture, attendees will understand the inner workings of the transformer architecture and the application of such language models in genomics.

Presenter: MARTINEK, Vasikmil (CEITEC MU)

Explainable AI

Contribution ID: 15

Type: not specified

Explainable AI

Thursday 4 July 2024 14:00 (1h 30m)

This talk dives into Explainable AI (XAI), exploring its significance in demystifying complex Machine Learning (ML) and Deep Learning (DL) models. It addresses the necessity of understanding model predictions, especially in biology, to reveal mechanisms behind biological systems. The talk covers model-agnostic and model-specific methods, highlighting the challenge of using XAI methods when features are not independent. It emphasises how XAI contributes to scientific understanding, safety, and ethical considerations, ensuring models are accurate but also fair and trustworthy

Presenter: THIEBES, Scott (KIT) Session Classification: Lecture Series

Biological Data

Contribution ID: 16

Type: not specified

Biological Data

Thursday 25 April 2024 14:00 (1h 30m)

This talk provides an introduction to omics data, highlighting the different omics data types, such as genomics and proteomics, and explaining experimental methods like the collection of transcriptomics data via RNA sequencing. In addition, it addresses the challenges of collecting omics data and their inherent biases. The aim is to provide a fundamental understanding of omics data characteristics, which is essential for applying machine learning in omics research and applications.

Presenter: USADEL, Björn (Forschungszentrum Jülich, IBG) **Session Classification:** Lecture Series

AI for Drug Discovery

Contribution ID: 17

Type: not specified

Al for Drug Discovery

Thursday 11 July 2024 14:00 (1h 30m)

This lecture introduces drug discovery, emphasising the role of Machine Learning (ML) and Deep Learning (DL) in modern pharmaceutical research. It covers the application of these technologies in predicting drug-target interactions, highlighting their importance in developing new therapies. The session aims to offer a comprehensive overview of the intersection between AI and drug discovery, focusing on its practical implications in healthcare.Lecture 10: AI for Genomics

Presenter: LOTFOLLAHI, Mohammed (Wellcome Sanger Institute, CCAIM, University of Cambridge)

AI for Omics

Contribution ID: 18

Type: not specified

Al for Omics

Thursday 18 July 2024 14:00 (1h 30m)

This lecture introduces applications of AI models to omics data in the context of precision medicine. The key goals are 1) to understand the molecular processes underlying complex human phenotypes to be able to rationally device new therapeutic strategies; 2) to identify individuals at high risk of disease and 3) to predict the response to treatment. The lecture will introduce recent work from the Heinig lab addressing these aims for cardiovascular diseases.

Presenter: HEINIG, Matthias (Helmholtz Munich, TUM School of Computation, Information and Technology)

ChatGPT in Action

Contribution ID: 19

Type: not specified

ChatGPT in Action

Thursday 25 July 2024 14:00 (1h 30m)

This lecture provides a high-level overview of the capabilities and future perspectives of ChatGPT and GPT-4, focusing on their applications in the scientific world. It provides insights into the workings of large language models (LLMs) and offers practical guidance on how to use ChatGPT as a scientist effectively. Attendees will learn about enhancing their research and writing through advanced prompt engineering techniques. The session also covers best practices in coding with ChatGPT with practical examples.

Presenters: CEA, Donatella (Helmholtz AI, Helmholtz Munich); MEKKI, Ilhem Isra (Helmholtz AI, Helmholtz Munich)

Diffusion Models

Contribution ID: 20

Type: not specified

Diffusion Models

Monday 15 July 2024 14:00 (1h 30m)

Diffusion models are a novel class of generative machine learning models that have recently gained significant traction in different fields, including biomedical applications. Such models operate by gradually injecting additional noise into data, yielding a distribution of random noise. To generate new unseen instances, the random noise distribution is sampled and subjected to a denoising step. This process results in a new instance from the target distribution. In the context of biomedical applications, diffusion models hold promise for a wide range of different tasks, from drug discovery to medical imaging analysis. This talk provides a brief overview of their potential and their challenges for large-scale biomedical applications.

Presenter: RIECK, Bastian (Helmholtz Munich, Technical University of Munich)