

2024年3月4日(月)

チュートリアル・特別講演 | チュートリアル

10:00 ~ 11:30 | 大会議室：409 アクリエひめじ 4F

**[TU-A-1] クラウド環境で駆動する生成系AIの最先端**

10:00 ~ 11:30

[TU-A-1] クラウド環境で駆動する生成系AIの最先端

杉村 勇馬(株式会社G-gen CTO), 山本 紘暉(クラスメソッド株式会社 新規事業部 生成AIチーム), 遠藤 雅樹(職業能力開発総合大学校/WorldSkills クラウドコンピューティング職種エキスパート), 横山 昌平(東京都立大学/クラウド技術者養成協会)

チュートリアル・特別講演 | チュートリアル

10:00 ~ 11:30 | 大会議室：408 アクリエひめじ 4F

**[TU-B-1] LLMと音声理解・生成の最新動向**

10:00 ~ 11:30

[TU-B-1] LLMと音声理解・生成の最新動向

西田京介, 安藤厚志 (NTT 人間情報研究所)

チュートリアル・特別講演 | チュートリアル

10:00 ~ 11:30 | 大会議室：407 アクリエひめじ 4F

**[TU-C-1] 高次元空間の最近傍探索 (Approximate Nearest Neighbor Search in High-dimensional Space)**

10:00 ~ 11:30

[TU-C-1] 高次元空間の最近傍探索 (Approximate Nearest Neighbor Search in High-dimensional Space)

Kejing Lu (名古屋大学)

チュートリアル・特別講演 | チュートリアル

10:00 ~ 11:30 | 中会議室：401~403 アクリエひめじ 4F

**[TU-D-1] 連合学習のすゝめ。(前半)**

10:00 ~ 11:30

[TU-D-1a] 連合学習のすゝめ。(前半)

佐々木 勇和 (大阪大学), 西尾 理志 (東京工業大学), Seng Pei Liew (LINEヤフー), Yang Cao (北海道大学), 鈴木 雄太 (トヨタ自動車)

チュートリアル・特別講演 | チュートリアル

13:10 ~ 13:40 | 中会議室：401~403 アクリエひめじ 4F

**[TU-D-1b] 連合学習のすゝめ。(後半)**

13:10 ~ 13:40

[TU-D-1b] 連合学習のすゝめ。(後半)

佐々木 勇和 (大阪大学), 西尾 理志 (東京工業大学), Seng Pei Liew (LINEヤフー), Yang Cao (北海道大学), 鈴木 雄太 (トヨタ自動車)

チュートリアル・特別講演 | チュートリアル

13:10 ~ 14:40 | 大会議室：409 アクリエひめじ 4F

**[TU-A-2] 統計的形狀分析による経済データ分析**

13:10 ~ 14:40

[TU-A-2] 統計的形狀分析による経済データ分析

白田由香利 (学習院大学)

チュートリアル・特別講演 | チュートリアル

13:10 ~ 14:40 | 大会議室：408 アクリエひめじ 4F

**[TU-B-2] OSSデータベースの開発コミュニティに参加しよう！**

13:10 ~ 14:40

[6C2-1310-1add] [TU-B-2] OSSデータベースの開発コミュニティに参加しよう！

藤井 雅雄, 加藤 慎也 (株式会社NTTデータグループ)

チュートリアル・特別講演 | チュートリアル

13:10 ~ 14:40 | 大会議室：407 アクリエひめじ 4F

**[TU-C-2] 大規模言語モデルに基づく検索モデル**

13:10 ~ 14:40

[TU-C-2] 大規模言語モデルに基づく検索モデル

加藤 誠, 阿部 健也, 薄羽 阜太 (筑波大学)

チュートリアル・特別講演 | チュートリアル

13:50 ~ 14:50 | 中会議室：401~403 アクリエひめじ 4F

**[TU-D-2] LLMの嘘：ハルシネーション解説**

13:50 ~ 14:50

[6C4-1350-1add] チュートリアル

三島 惇也, 北村 拓斗, 桑原 悠希, 高橋 巧実, 鈴木 優 (岐阜大学)

チュートリアル・特別講演 | チュートリアル

📅 2024年3月4日(月) 10:00 ~ 11:30 | 🏢 大会議室：409 アクリエひめじ 4F

**[TU-A-1] クラウド環境で駆動する生成系AIの最先端**

10:00 ~ 11:30

[TU-A-1] クラウド環境で駆動する生成系AIの最先端

杉村 勇馬(株式会社G-gen CTO), 山本 紘暉(クラスメソッド株式会社 新規事業部 生成AIチーム), 遠藤 雅樹 (職業能力開発総合大学校 / WorldSkills クラウドコンピューティング職種エキスパート), 横山 昌平(東京都立大学 / クラウド技術者養成協会)

チュートリアル・特別講演 | チュートリアル

## [TU-A-1] クラウド環境で駆動する生成系 AIの最先端

2024年3月4日(月) 10:00 ~ 11:30 大会議室：409 (アクリエひめじ 4F)

10:00 ~ 11:30

### [TU-A-1]クラウド環境で駆動する生成系 AIの最先端

杉村 勇馬(株式会社G-gen CTO), 山本 紘暉(クラスメソッド株式会社 新規事業部 生成AIチーム), 遠藤 雅樹 (職業能力開発総合大学校/WorldSkills クラウドコンピューティング職種エキスパート), 横山 昌平(東京都立大学/クラウド技術者養成協会)

対象：研究に生成系AIを取り入れてみたい学部生・院生そして教員

ChatGPTを始めとした生成AIの話題で持ちきりの一年でした。データ工学や情報マネジメントの研究コミュニティにおいても、生成AIを使う研究が多数出てきています。膨大な計算リソースとデータを準備し独自のAIモデルを構築する事もできますが、クラウド環境を使えば、容易に生成AIを活用する事ができます。皆さんの研究にすぐに活用できるように、本チュートリアルではAWSとGoogle Cloudの二つの環境において、生成AIを使う方法を伝授します。対象とする技術レベルは情報系の大学生・大学院生です。また、チュートリアル終盤では、なるべく定額でクラウドを利用するTips等、大学の研究においてクラウドを利用する際にしばしば最大の障害となる『契約手続き』について、パネル形式でお伝えします。

#### 主なトピック

- ・ 山本紘暉 (クラスメソッド株式会社 新規事業部 生成AIチーム)  
「AWSで生成AIのRAGを使ったチャットボットを作ってみよう」 - 35分
- ・ 杉村勇馬 (株式会社G-gen CTO)  
「日本企業における生成 AI のリアル」 - 35分
- ・ パネル  
「研究におけるAWS/Google Cloudの活用：定額利用や契約のあれこれ」 - 20分

司会：横山昌平(東京都立大学/NPO法人クラウド技術者養成協会理事長)

チュートリアル・特別講演 | チュートリアル

🎵 2024年3月4日(月) 10:00 ~ 11:30 | 🏢 大会議室：408 アクリエひめじ 4F

**[TU-B-1] LLMと音声理解・生成の最新動向**

10:00 ~ 11:30

[TU-B-1] LLMと音声理解・生成の最新動向

西田京介, 安藤厚志 (NTT 人間情報研究所)

チュートリアル・特別講演 | チュートリアル

## [TU-B-1] LLMと音声理解・生成の最新動向

2024年3月4日(月) 10:00 ~ 11:30 大会議室：408 (アクリエひめじ 4F)

10:00 ~ 11:30

## [TU-B-1] LLMと音声理解・生成の最新動向

西田京介, 安藤厚志 (NTT 人間情報研究所)

対象：学部4年生レベル～

ChatGPTの登場以降、LLMをベースとした汎用人工知能（AGI）の実現が現実味を帯びつつある。OpenAIのGPT-4は、言語理解・生成の高度な能力に加えて視覚入力の処理も可能となり、一部のタスクでは平均的なヒトを超える能力を既に実現している。さらに、GoogleのGeminiは、言語と視覚に加えて音声に対する学習も組み込むことで、ヒトとAIの共生社会に革命的な進化をもたらす可能性を秘めている。このチュートリアルでは、大規模言語モデル（Large Language Model; LLM）および、LLMのマルチモーダル拡張における音声の理解と生成の分野に着目する。前半では、LLMの基本から最新の進展について詳しく解説し、後半ではLLM関連技術が音声理解・生成に与えた影響やLLMを活用したマルチモーダルなアプローチについて紹介する。

チュートリアル・特別講演 | チュートリアル

2024年3月4日(月) 10:00 ~ 11:30 | 大会議室：407 アクリエひめじ 4F

**[TU-C-1] 高次元空間の最近傍探索 (Approximate Nearest Neighbor Search in High-dimensional Space)**

10:00 ~ 11:30

[TU-C-1] 高次元空間の最近傍探索 (Approximate Nearest Neighbor Search in High-dimensional Space)

Kejing Lu (名古屋大学)

チュートリアル・特別講演 | チュートリアル

## [TU-C-1] 高次元空間の最近傍探索 (Approximate Nearest Neighbor Search in High-dimensional Space)

2024年3月4日(月) 10:00 ~ 11:30 大会議室：407 (アクリエひめじ 4F)

10:00 ~ 11:30

### [TU-C-1]高次元空間の最近傍探索 (Approximate Nearest Neighbor Search in High-dimensional Space)

Kejing Lu (名古屋大学)

対象：大学院1年レベル

ビッグデータの時代では多くのアプリケーションで大規模なデータセットを扱う必要があるが、その際に必ず直面する問題が最近傍探索問題 (NNS) である。この問題は非常に基本的であり、データマイニング、画像検索、推薦システムなど、ほとんど全てのアプリケーションに現れる。特に近年、人工知能・機械学習の分野で注目されている「ベクトル検索」もNNSに深く関連する。この問題を解くことは、特に高次元空間においては非常に困難である。そのため、近似最近傍探索問題 (ANNS) が注目されている。このチュートリアルでは、ANNS問題の包括的な紹介を行う。チュートリアルは以下の3つの部分からなる。

第1部では、ANNSの背景と、高次元空間におけるその難しさについて述べる。この問題の目的は極めて理解しやすいが、以下の2つの理由により、この問題を実際に解くことは容易ではない。(1)大規模なデータセットに対する線形スキャン (総当り) は、その複雑さゆえに常に受け入れがたい。(2)次元の呪いと呼ばれる現象が存在する。実際のところ、高次元空間では、低次元空間ではうまく機能していた木構造が、線形スキャンに劣る可能性がある。すなわち、高次元でこの難問を解くには、本質的に異なるアイデアが必要になる。

第2部では、高次元ユークリッド空間において有効な3種類のANNS手法である、木構造、Locality Sensitive Hashing (LSH)、類似性グラフについて紹介する。これらは一般に、高次元空間では線形スキャンよりもはるかに良く機能する。これらの基本的な考え方を紹介した後、これらに関連するANNS手法について紹介する。

第3部では、ANNSの今後の動向について紹介する。例えば、一般的にグラフベース検索が最も有効であることは知られているが、その問合せ品質の理論的保証など、その動作メカニズムについてはまだほとんど分かっていない。グラフベースの検索をより理論的に説明できる研究者や、グラフベースの性能を継続的に向上させる産業界の研究者が必要である。将来アカデミアに進む者も、また、企業に就職する者も、このテーマには面白い仕事が見つかるかもしれない。学生や若手研究者がANNSのコミュニティに参加する動機付けになれば幸いである。

チュートリアル・特別講演 | チュートリアル

📅 2024年3月4日(月) 10:00 ~ 11:30 | 🏢 中会議室：401~403 アクリエひめじ 4F

**[TU-D-1] 連合学習のすすめ。(前半)**

10:00 ~ 11:30

[TU-D-1a] 連合学習のすすめ。(前半)

佐々木 勇和 (大阪大学), 西尾 理志 (東京工業大学), Seng Pei Liew (LINEヤフー), Yang Cao (北海道大学), 鈴木 雄太 (トヨタ自動車)

チュートリアル・特別講演 | チュートリアル

## [TU-D-1] 連合学習のすゝめ。(前半)

2024年3月4日(月) 10:00 ~ 11:30 中会議室: 401~403 (アクリエひめじ 4F)

10:00 ~ 11:30

### [TU-D-1a] 連合学習のすゝめ。(前半)

佐々木 勇和 (大阪大学), 西尾 理志 (東京工業大学), Seng Pei Liew (LINEヤフー), Yang Cao (北海道大学), 鈴木 雄太 (トヨタ自動車)

対象: 大学院1年レベルから機械学習初学者

連合学習は各クライアントが保持するデータをサーバと共有せずに, 学習済みモデルを共有することで, 高精度なモデルを構築する分散型の学習アプローチである. プライバシの保護および通信量の削減が可能であるため, 研究開発だけではなく実際のサービスに既に応用されている. 本チュートリアルでは, 連合学習の重要性や研究の面白さを伝え, 多くの学生や研究者に連合学習を浸透させることを目的とする. 背景や概要を説明した後, データ不均一性, プライバシ, および通信効率といった様々な観点から行われている手法を紹介する. さらに, 各個人がもつようなモバイルデバイスへの応用と各企業や組織がもつようなデータサイロへの応用を紹介する. 最後に, 我々が考えている今後の展望について述べる. 本チュートリアルを見れば, 膨大な数の論文が増え続けている連合学習の研究動向が把握できます!

チュートリアル・特別講演 | チュートリアル

📅 2024年3月4日(月) 13:10 ~ 13:40 | 🏢 中会議室：401~403 アクリエひめじ 4F

**[TU-D-1b] 連合学習のすゝめ。(後半)**

13:10 ~ 13:40

[TU-D-1b] 連合学習のすゝめ。(後半)

佐々木 勇和 (大阪大学), 西尾 理志 (東京工業大学), Seng Pei Liew (LINEヤフー), Yang Cao (北海道大学), 鈴木 雄太 (トヨタ自動車)

チュートリアル・特別講演 | チュートリアル

## [TU-D-1b] 連合学習のすゝめ。(後半)

2024年3月4日(月) 13:10 ~ 13:40 中会議室：401~403 (アクリエひめじ 4F)

13:10 ~ 13:40

## [TU-D-1b]連合学習のすゝめ。(後半)

佐々木 勇和 (大阪大学), 西尾 理志 (東京工業大学), Seng Pei Liew (LINEヤフー), Yang Cao (北海道大学), 鈴木 雄太 (トヨタ自動車)

対象：大学院1年レベルから機械学習初学者

連合学習は各クライアントが保持するデータをサーバと共有せずに、学習済みモデルを共有することで、高精度なモデルを構築する分散型の学習アプローチである。プライバシーの保護および通信量の削減が可能であるため、研究開発だけではなく実際のサービスに既に応用されている。本チュートリアルでは、連合学習の重要性や研究の面白さを伝え、多くの学生や研究者に連合学習を浸透させることを目的とする。背景や概要を説明した後、データ不均一性、プライバシー、および通信効率といった様々な観点から行われている手法を紹介する。さらに、各個人がもつようなモバイルデバイスへの応用と各企業や組織がもつようなデータサイロへの応用を紹介する。最後に、我々が考えている今後の展望について述べる。本チュートリアルを見れば、膨大な数の論文が増え続けている連合学習の研究動向が把握できます！

チュートリアル・特別講演 | チュートリアル

🎵 2024年3月4日(月) 13:10 ~ 14:40 | 🏢 大会議室：409 アクリエひめじ 4F

**[TU-A-2] 統計的形狀分析による経済データ分析**

13:10 ~ 14:40

[TU-A-2] 統計的形狀分析による経済データ分析

白田由香利 (学習院大学)

チュートリアル・特別講演 | チュートリアル

## [TU-A-2] 統計的形狀分析による経済データ分析

2024年3月4日(月) 13:10 ~ 14:40 大会議室：409 (アクリエひめじ 4F)

13:10 ~ 14:40

### [TU-A-2] 統計的形狀分析による経済データ分析

白田由香利 (学習院大学)

対象：データサイエンティスト、リーマン距離の応用に興味のあるかた

データ散布図をひとつの形状(shape, formation)とみなすことで、データ集合の時系列変化は形状のフォーメーション変化(deformation)として表現可能となる。この変化は非線形であり、従来定量的計測が困難であったが、統計的形狀分析により、主成分ごとに変化を測定可能となった。統計的形狀分析は、生物学的進化(骨格の変化など)の計測手段としてMardia達がセオリーを開発した。しかし我々のチームでは経済経営データに対して適応してきた。例えば、国のGDP成長率と貧困レベルの散布図の時系列変化分析などである。本講義では、経済データによる応用を説明した後、数学的セオリーを説明する。形状は、センタリングおよびスケールングを行い、preshapeに変換する。Preshapeの形状空間はリーマン多様体としてみなすことができ、形状間の距離はリーマン距離として定義できる。セオリーの2つ目のポイントは、薄型平板スプラインによる補間である。2次元形状の点集合間に、仮想的に透明な板があるとみなし、その板を3次元空間でどのように曲げると与えられた2次元形状点集合ができるかを、ベンディング・エネルギー最小化問題として解く。データの点の数だけ、基底関数を並べ、その関数の和として形状変換モデルを構築する。時系列データ分析手法として非常に役立つ統計的形狀分析の応用と数学理論を、楽しいグラフィクスで分かり易く説明する。殆どのかたが、見たことのない新しい理論と手法だと思うので、是非、この機会に覚えて使ってください。

教材サイト：Yukari SHIROTA, Riri Fitri Sari, Tetri Widiyani, and Takako Hashimoto: "" Visually Do Statistical Shape Analysis!,"" Visual Math Lecture Note Series, the tutorial lecture in Data Science and Advanced Analytics (DSAA) Tokyo, Japan: IEEE, 21st Oct. 2017. <https://www-cc.gakushuin.ac.jp/~20010570/mathABC/SELECTED/ShapeAnalysis/>

チュートリアル・特別講演 | チュートリアル

🎵 2024年3月4日(月) 13:10 ~ 14:40 | 🏢 大会議室：408 アクリエひめじ 4F

📌 [TU-B-2] OSSデータベースの開発コミュニティに参加しよう！

13:10 ~ 14:40

[6C2-1310-1add] [TU-B-2] OSSデータベースの開発コミュニティに参加しよう！

藤井 雅雄, 加藤 慎也 (株式会社NTTデータグループ)

チュートリアル・特別講演 | チュートリアル

## [TU-B-2] OSSデータベースの開発コミュニティに参加しよう！

2024年3月4日(月) 13:10 ~ 14:40 大会議室：408 (アクリエひめじ 4F)

13:10 ~ 14:40

### [6C2-1310-1add][TU-B-2] OSSデータベースの開発コミュニティに参加しよう！

藤井 雅雄, 加藤 慎也 (株式会社NTTデータグループ)

対象：学部4年生レベル、OSS開発に興味のある方

オープンソースソフトウェア (OSS) のリレーショナルデータベース管理システムであるPostgreSQLは、世界中の開発者が協力して開発され、多くの企業や組織で利用されています。また、開発においては学術研究の成果が様々に活用されています。本チュートリアルでは、OSSやPostgreSQLの基本を解説し、PostgreSQL開発コミュニティへの参加方法に焦点を当てます。

具体的には、PostgreSQLのバグの調査からコミュニティへの報告、修正パッチの投稿方法について解説します。また、ライブコーディングを通じて、簡単な機能の実装、コンパイル、実行を実演します。さらに、今までにNTTデータグループが開発したPostgreSQLの機能を説明し、最後にアカデミックなデータベースの研究とPostgreSQLの関連についても紹介します。

本チュートリアルに参加することで、PostgreSQLの基本から、PostgreSQLコミュニティへの参加までの具体的なステップを学ぶことができます。

チュートリアル・特別講演 | チュートリアル

🎵 2024年3月4日(月) 13:10 ~ 14:40 | 🏢 大会議室：407 アクリエひめじ 4F

**[TU-C-2] 大規模言語モデルに基づく検索モデル**

13:10 ~ 14:40

[TU-C-2] 大規模言語モデルに基づく検索モデル

加藤 誠, 阿部 健也, 薄羽 皐太 (筑波大学)

---

チュートリアル・特別講演 | チュートリアル

## [TU-C-2] 大規模言語モデルに基づく検索モデル

2024年3月4日(月) 13:10 ~ 14:40 大会議室：407 (アクリエひめじ 4F)

---

13:10 ~ 14:40

## [TU-C-2]大規模言語モデルに基づく検索モデル

加藤 誠, 阿部 健也, 薄羽 阜太 (筑波大学)

対象：情報検索に関心のある学部4年生レベル以上の方

チュートリアル・特別講演 | チュートリアル

📅 2024年3月4日(月) 13:50 ~ 14:50 | 🏢 中会議室：401~403 アクリエひめじ 4F

📌 [TU-D-2] LLMの嘘：ハルシネーション解説

13:50 ~ 14:50

[6C4-1350-1add] チュートリアル

三島惇也, 北村拓斗, 桑原悠希, 高橋巧実, 鈴木優 (岐阜大学)

チュートリアル・特別講演 | チュートリアル

## [TU-D-2] LLMの嘘：ハルシネーション解説

2024年3月4日(月) 13:50 ~ 14:50 中会議室：401~403 (アクリエひめじ 4F)

13:50 ~ 14:50

### [6C4-1350-1add]チュートリアル

三島惇也, 北村拓斗, 桑原悠希, 高橋巧実, 鈴木優 (岐阜大学)

対象：LLMに興味を持っており、これから勉強したいと考えている学部生、大学院生

Llama2やGPT4をはじめとするLLM(Large Language Model)が登場し、パラダイムシフトが起きた。ChatGPTやGitHub copiloteをはじめとするLLMを用いたサービスを利用する人が増えてきた。LLMは様々なタスクで高い性能を示すが、LLMは嘘（ハルシネーション）を生成することもある。本チュートリアルでは、LLMのハルシネーションに焦点を当てる。LLMの構築方法からハルシネーション解消に向けたアプローチまでを解説する。少しでもLLMをわかって（わかった気になって）くれると嬉しい。また、本チュートリアルは学生が主導して行う。今回の試みが将来的に、学生によるチュートリアル発表が行われる契機となり、本研究会の発展や学生間の交流につながることを我々は期待している。