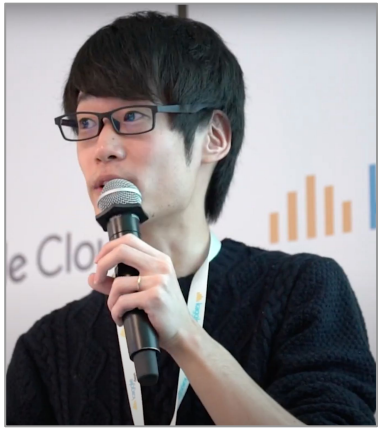
学生生活紹介

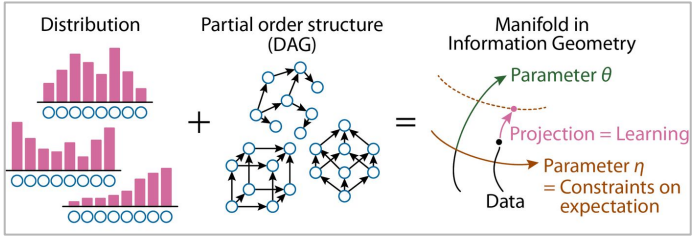
情報学コース　杉山麿人研究室
加納龍一

自己紹介

- 加納龍一 (かのうりゅういち)
  - IT企業に勤務しながら、社会人学生として研究
  - 5年一貫コースの4年目



- 所属：情報学専攻 杉山研究室
  - 機械学習
    - 離散数学
    - 統計学 (情報幾何学など)
    - ...



入学の動機

- Kaggleなどで使用されている様々なテクニックの「根拠」が気になった
  - 「何故か分からないけれどうまくいく」というテクニックが大量に存在
  - 経験的な成功が先行する中で根拠が理解できれば、世界に対するインパクトも大きい！

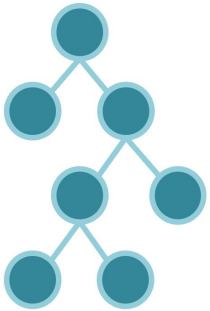
- 指導教員と興味が似ていた
  - 離散数学的な側面と統計学的な側面の両方からのアプローチ

- いつかは博士の学位が欲しいと思っていた
  - 自分の人生の中での成し遂げたいことのひとつ
  - 社会人も3年目に入り、少しずつ仕事に慣れてきた

研究紹介

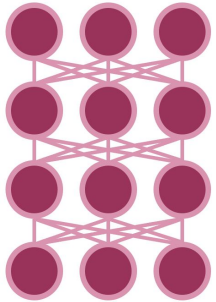
- 決定木とニューラルネットワークとを融合させた手法(Soft-Tree)の理論研究
  - 二大巨塔の強みを合わせた現代的なモデルに対するテクニックの裏付けを与えたい

決定木



・階層的なデータ分割に基づく予測
・探索ベースでhardな決定境界を構築

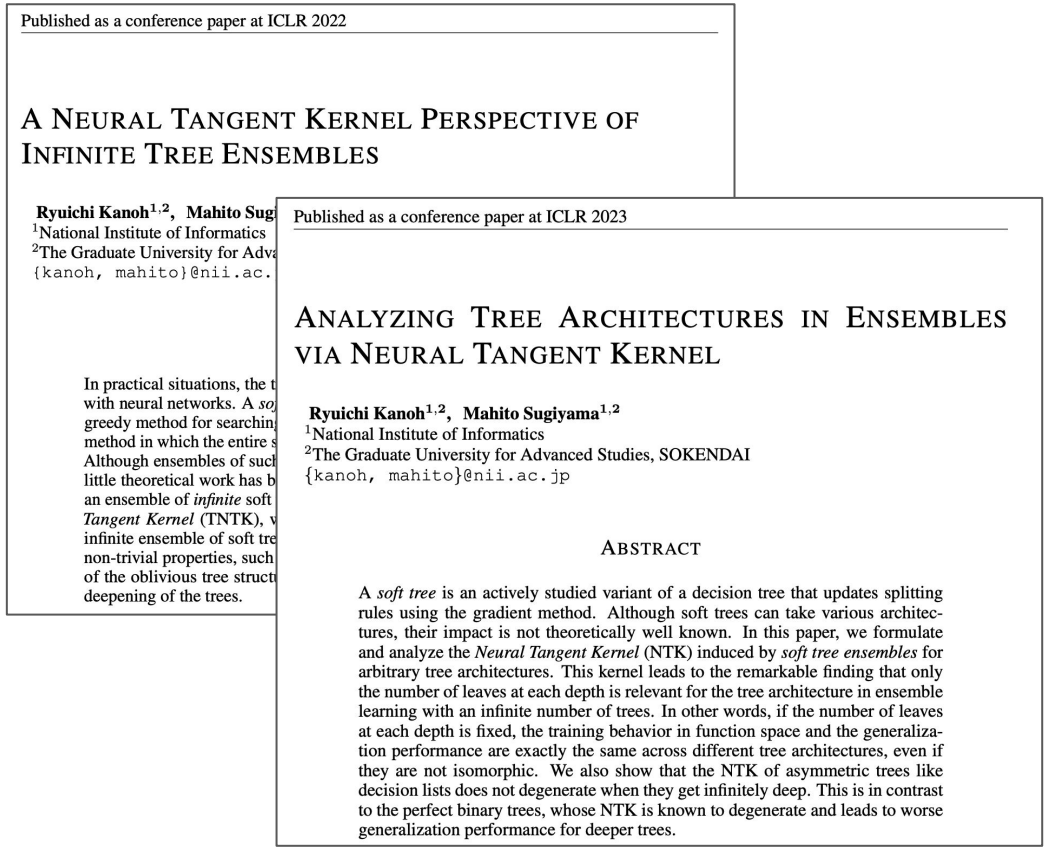
ニューラルネット



・階層的な特徴量抽出に基づく予測
・全処理を微分可能な演算とし、勾配法で学習

実績

- トップカンファレンス採択
  - ICLR2022, ICLR2023
  - 各種ワークショップ等
- 人工知能学会で各種受賞
  - 研究会優秀賞 (2021)
  - 全国大会優秀賞 (2023)
- などなど



Published as a conference paper at ICLR 2022

A NEURAL TANGENT KERNEL PERSPECTIVE OF INFINITE TREE ENSEMBLES

Ryuichi Kanoh[1,2], **Mahito Sug**
[1]National Institute of Informatics
[2]The Graduate University for Adva
{kanoh, mahito}@nii.ac.

In practical situations, the t
with neural networks. A *so*
greedy method for searchin
method in which the entire s
Although ensembles of such
little theoretical work has b
an ensemble of *infinite* soft
Tangent Kernel (TNTK), w
infinite ensemble of soft tre
non-trivial properties, such
of the oblivious tree structu
deepening of the trees.

Published as a conference paper at ICLR 2023

ANALYZING TREE ARCHITECTURES IN ENSEMBLES VIA NEURAL TANGENT KERNEL

Ryuichi Kanoh[1,2], **Mahito Sugiyama**[1,2]
[1]National Institute of Informatics
[2]The Graduate University for Advanced Studies, SOKENDAI
{kanoh, mahito}@nii.ac.jp

ABSTRACT

A *soft tree* is an actively studied variant of a decision tree that updates splitting rules using the gradient method. Although soft trees can take various architectures, their impact is not theoretically well known. In this paper, we formulate and analyze the *Neural Tangent Kernel* (NTK) induced by *soft tree ensembles* for arbitrary tree architectures. This kernel leads to the remarkable finding that only the number of leaves at each depth is relevant for the tree architecture in ensemble learning with an infinite number of trees. In other words, if the number of leaves at each depth is fixed, the training behavior in function space and the generalization performance are exactly the same across different tree architectures, even if they are not isomorphic. We also show that the NTK of asymmetric trees like decision lists does not degenerate when they get infinitely deep. This is in contrast to the perfect binary trees, whose NTK is known to degenerate and leads to worse generalization performance for deeper trees.

研究室メンバーの研究内容

- 応用から理論まで、様々な内容に取り組んでいる
  - 理論
    - 情報幾何の観点から高速な低ランク近似の手法を提案
    - 非負値行列分解における過剰パラメータ化の解析
  - 応用
    - 時系列データの外れ値検出
    - 信号分離

NeurIPS, ICLR, AISTATSなど、幅広い機械学習関係の学会で発表

研究の日常

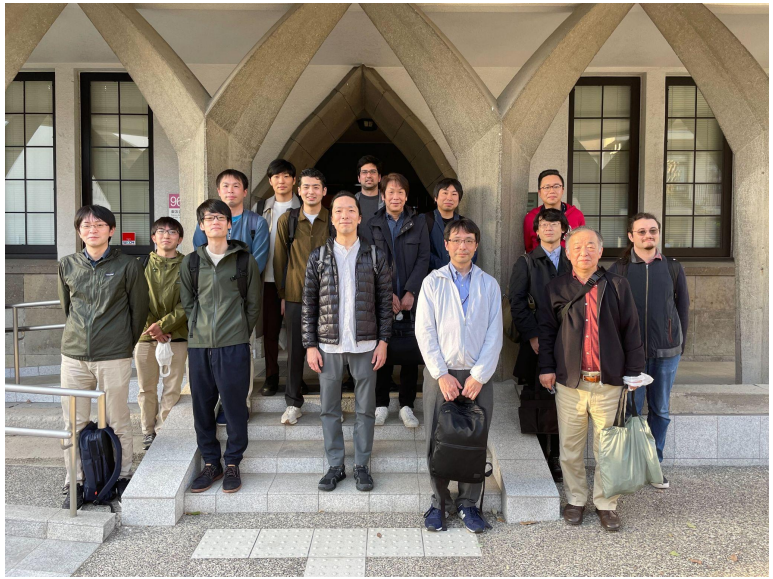
- オンライン/オフラインのハイブリッドでコミュニケーション
  - 研究室によって雰囲気は違うかも？

- メンバーとの交流機会も色々
  - 毎週指導教員とディスカッション
  - ランチセミナー (論文紹介、雑談etc)
  - 研究室全体セミナー (研究成果共有)
  - 各種研究会、学会

研究の日常

- 学会や研究会もオフライン開催されるようになってきて楽しい
  - (少し前はオンラインのみのことが多く、寂しかった…)

研究室合宿@軽井沢



JST-CREST研究会@京都大学

ICML2023@ハワイ

仕事と研究の両立

- 仕事も研究も在宅でできているので、移動などがネックになることはない
  - コロナ禍を超えたからこそ、柔軟にやりやすくもなってきていると思う
  - 自分は業務後にそのまま家で研究をすることが多い
  - (研究会などオフラインのイベントは、業務と都合をつけたりするのは大変かも)

- 研究の時間は基本的に業務後or土日
  - 仕事が忙しかったりすると、あまり研究ができない時期も
  - ひとつひとつのマイルストーン(論文締め切りなど)を見据えた計画性は重要
  - ディスカッションの時間などは、うまく時間を見つけながら調整

研究室訪問などでの確認をおすすめします

おわりに

- 成果をあげるには、総研大の情報学専攻は良い選択肢のひとつ
  - 教員に対して学生の数も少なく、指導教員と議論などもじっくりとできる
  - 同じ志を持つ仲間と共に研究に取り組む日々は刺激的
    - 留学生や社会人、他大学の方々など、多様なメンバーとの新しい繋がりも魅力
  - 主体性を持って行動すれば、どんどんチャンスを広げることができる

- 一言に「総研大」といっても、研究室によって文化はバラバラ(だと思う)
  - 興味のある研究室にいくつか話を聞いてみることをお勧めします
  - そのうえで、自分にマッチした環境がどこかにはあるのではないかと思います