船井情報科学振興財団　第 2 回留学報告書

齋藤 優太 *

2021 年 12 月

2021 年 8 月より、Cornell University の Computer Science 専攻 Ph.D. 課程に在籍している齋藤優太と申します。秋学期がほぼ終了し、年明けに控えるいくつかの重要な国際会議への論文投稿に向けて研究に集中しているところです。今回の報告書では、大学院での研究活動や講義、生活状況等についてご報告します。

1　研究

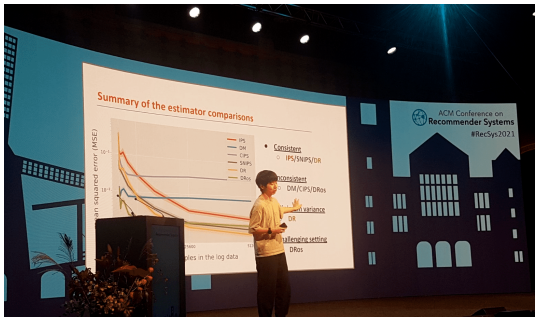
学期が始まってからすぐに研究関連の大きな行事として、ACM Conference on Recommender Systems (RecSys) という推薦システムに関するトップ国際会議[*1]における Tutorial 開催がありました。機械学習関連の国際会議では、毎回いくつかの Tutorial が開催されます。これは一言でいうと、あるトピックについて数時間で構成される集中講義といった位置付けの企画で、その分野に明るい研究者が自身の Ph.D. 学生などと共に企画・開催します。機械学習関連の主要国際会議で日本人の方が Tutorial を開催しているのはあまり見たことがなかったのですが、自分や自分の研究プロジェクトを宣伝するのにとても良い場だと思っており、Ph.D. の期間中に何度か開催するつもりでいます。今回はその最初の挑戦として、RecSys における 3 時間の Tutorial 枠を 6 月の段階で確保していました。よって 8 月中旬に渡米してからしばらくは、この Tutorial で納得のいく発表をすべく、資料や発表練習などの準備に勤しんでいました。

図 1　（左）アムステルダム中央駅。観光客も含め、ほとんどマスクをしていない。（右）学会会場。欧州近隣国からの参加者が大多数だったが、会場はほぼ満員でかなりの盛り上がりだった。

* ys552@cornell.edu

*1 https://recsys.acm.org/recsys21/

今年の RecSys は、オランダの首都アムステルダムで開催されました。多くの学会がまだ完全にオンラインで開催されている状況でしたが、RecSys は現地参加とオンライン参加の両方が可能なハイブリッドの形態で開催されました。去年は完全にオンラインだったので、東京から遠隔で参加するしかなかったのですが、今回は 2 年ぶりに現地参加でき、9 月最後の 1 週間はアムステルダムに滞在していました。基調講演や一般発表等が行われる学会のメイン会場は常時かなりの割合の席が埋まっており、久しぶりに物理学会の雰囲気を楽しむことができました。また当たり前ではありますが、現地参加した方が圧倒的に色んな人と交流しやすく、多くの企業研究者と知り合いになれたことは現地参加したからこその収穫でした。なお、アムステルダム市街ではマスクを着用している人を見つけるのが大変なほどパンデミック以前の生活に戻っている印象でした。



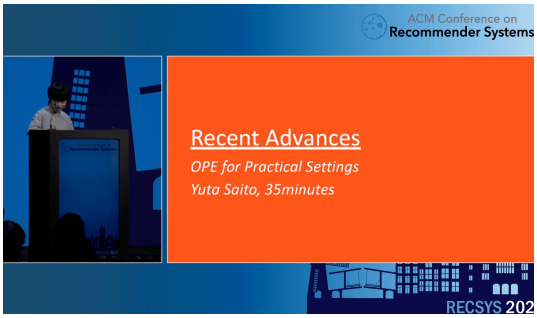


図 2　チュートリアル発表の様子。英語で 2 時間ほどの発表はこれまでで最長だったが、必要な練習は事前に積んでいたので、特に緊張せずやり遂げることができた。発表は右の写真のようにオンラインでの参加者向けに生中継されていた。

さて学会では、ウェブ企業等における実用性の高さから注目が集まっているオフ方策評価（Off-Policy Evaluation）に関するチュートリアルを開催しました。[*2]　具体的には、過去にとあるアルゴリズム（例えば、YouTube や Amazon などのサービスで使われている推薦アルゴリズム）によって蓄積されたデータを用いて、新たなアルゴリズムの性能を評価したりするための統計技術に関する基礎から最近の研究、実装方法などを広く扱いました。もちろん新たに開発したアルゴリズムを一定期間サービスに実装すればその性能を正確に知れるわけですが、そのアルゴリズムの性能が悪かったとしたら、実験中に大きな損害を被ってしまうかもしれません。したがって、新たに開発したアルゴリズムの性能を、それを実装することなく既に手元にある蓄積データのみから安全かつ正確に評価したい、というモチベーションが生まれます。これは学会に参加していた多くの企業研究者が抱える推薦システムのオフライン評価に対する課題、及び蓄積データを有効に活用したいというモチベーションと関連が深いテーマだったからか、現地会場では大きな注目を集めることに成功しました。実際、学会後に Criteo という機械学習の研究で有名なフランス企業の研究所が公開した学会報告[*3]において（以下リンク先からの引用）、

- **One of the highlights of this year's RecSys was the counterfactual tutorial delivered by Yuta Saito and Thorsten Joachims** which was especially clear on the range of inverse

*2 チュートリアルのウェブサイト: https://sites.google.com/cornell.edu/recsys2021tutorial
*3 https://medium.com/criteo-engineering/recsys-2021-an-in-person-conference-at-long-last-in-the-wonderful-city-of-amsterdam-166ba805f021

propensity score style estimators and their trade-offs. The fact that there was no easy solution was emphasized and the presentation gave excellent intuition. They also announced their simulation platform Open Bandit Pipeline which appears to be an excellent resource.

と好意的に言及してもらうなどしました。本チュートリアルを通じて多くの参加者とコミュニケーションを取ることができ、新たな人脈につながるなどしています。こうして良い発表を地道に繰り返すことが自身の分野におけるプレゼンスの向上に繋がるとこれまでの経験から実感しているので、来年以降の RecSys や別の主要国際会議でも積極的にチュートリアルを開催しようと思っているところです。

さてオランダから帰国してからは、本格的に Cornell での研究に着手しています。元々は、前述の Tutorial のテーマでもあったオフ方策評価に関する研究をメインで行う予定だったのですが、それとは少し離れた推薦システムや情報検索における公平性の分野でとても良いテーマを偶然見つけ、最近はどちらかというとそちらに力を注いでいます。推薦システムは、Amazon や Netflix, YouTube, Spotify などの Web サービスにおいてユーザが好む商品や映画、動画、楽曲をユーザの過去の行動パターンに基づいてうまく推薦するための技術を扱う分野です。推薦システムの分野では長らく、「いかにユーザを満足させる推薦を行うか」というユーザ志向のモチベーションに基づきほとんどの研究が行われてきたのですが、ここ数年、「アイテムやコンテンツを提供する側」への影響を考慮に入れた研究も増え始めています。例えば、Spotify のトップ画面において人気歌手の楽曲ばかりを推薦する推薦アルゴリズムを実装したとします。するとすでに世間によく知られたの歌手の楽曲にすぐにありつけるため、ユーザを（少なくとも短期的に）一定程度満足されることが可能かもしれません。しかし、このアルゴリズムは一定の地位を確立した人気歌手の楽曲しか推薦しないわけですから、これからスターになろうとしているその他多くの歌手は、自身の楽曲がなかなか推薦されず不満を持つでしょう。そうなると、別のプラットフォーム（Amazon Music など）に主戦場を移してしまうかもしれません。すなわち、プラットフォームの長期的な繁栄を考えた場合にはユーザを満足させるだけでは十分ではなく、コンテンツを提供する側（例えば歌手）にとって公平な推薦を行うアルゴリズムを考える必要がでてきます。Cornell での advisor である Thorsten Joachims がこの推薦システムにおける公平性の問題にかなり初期に取り組み始めた人だということは以前から知っていたのですが、こちらに来るまではあまりキャッチアップできていませんでした。しかし、後述のセミナー形式の講義をきっかけにいくつか論文を読んでみたところ色々研究できそうな雰囲気を感じ、実際筋が良さそうな研究アイデアに辿り着くことができました。Thorsten も現在のアイデアにかなり満足しているようで、また研究グループ内で進捗を共有したときも評判が良さそうだったので、講義がない冬休みの間はこのアイデアを論文にまとめることに集中するつもりでいます（次の報告書のときにもう少し詳細を書けるように進捗を出せたら、と思っています）。

2　講義など

秋学期はいわゆる講義を 1 つ、セミナー形式の講義を 1 つ、その他セミナーに 2 つ参加しています。

まず、CS6787 Advanced Machine Learning Systems という機械学習システムに関する講義を受講しました。これまでは機械学習といっても統計的な側面の勉強しかしてこなかったのですが、この講義では最適化や並列化、ハードウェアなどこれまであまり触れてこなかった研究アプローチを知ることができ、教養として楽しめました。また（この講義に限らずですが）講義の最終プロジェクトも基本自分の研究と関連付けてテーマ設定できるため、講義を出発点にして研究を進めることもでき、とても良かったです。これに加えて、CS7792 Fairness and Dynamics of Learning Systems というセミナー形式の講義を受講しました。これは毎週学生が

持ち回りで関連の論文を発表して、みんなで議論するという形式のものです。名前の通り推薦システムやより広範な機械学習の公平性（Fairness）に関する内容で、気になっていたもののあまり論文を読めていなかった公平性の分野のサーベイを自分に強制するために受講しました。これまでは自分の研究に関係しそうな論文しか読んでこなかったのですが、たまに新しい分野の勉強をしてみると面白いもので、最近はむしろこちらの公平性の研究に力を入れるようになっていることは先に述べた通りです。なお上記 2 つの講義・セミナーでは、まあまあの頻度で学生同士のディスカッションの時間があったのですが、そのたびに自分の英語力の無さを痛感しました。ここで前までは真面目に、「英語の勉強を頑張らないと」と思っていた気がしますが、最近は英語で失敗しても気にならないくらい他に強みを持とう、という開き直った考えになってきています。

その他、CS のあらゆる分野における研究の雰囲気を感じることができる CS Colloquium というセミナーと、より自分の研究分野と関連が深い Artificial Intelligence Seminar に、それぞれ週 1 回参加しています。基本は外部から人を呼んで内部の Ph.D. 学生や先生方が発表を聞くというスタイルなのですが、AI Seminar の方は自分の advisor が仕切っていることもあり「発表したければしてもいいよ」と言われていて、春学期はこの辺のセミナーで発表して知り合いを増やせたらと画策しているところだったりもします。

図 3　（左）キャンパスは坂を登った高い場所にあり、夜はとても良い景色を眺めることができる。（右）CS 専攻が入っている Gates Hall。名前の通り、ビル・ゲイツからの寄付で 10 年ほど前に立った新しい建物。最近また大口の寄附があり、数年以内に CS 専攻用の新たな建物が立つとのこと。

3　生活

在籍している Cornell University のメインキャンパスは、ニューヨーク州のイサカ（Ithaca）という人口 3,4 万人の小都市に位置しています。ニューヨーク州ではありますが、多くの人が思い浮かべるニューヨーク “市” までは車で 4,5 時間かかり、都会と呼べる感じではありません。日本からイサカへの移動は少し大変で、渡米にはまあまあの時間がかかりました。まず羽田から 12 時間ほどのフライトでデトロイトに行き、デトロイトで 1 泊した後に、1 時間半ほどのフライトでイサカ空港へ、というルートです。これを 8 月に経験していたので、先のアムステルダムへの旅はかなり短く感じました。

大学選択の際、Cornell 以外は全てカリフォルニアの住みやすそうな街にある大学から合格をもらっており、生活環境の面で進路選択をかなり迷った、といったことを第 1 回報告書に書いていました。しかし、数ヶ月住んでみたところ、生活環境についてあまり悩む必要はなかったなと感じるくらい、生活面で不満な点は出てきていません。そもそも学期が始まってからは基本研究に集中しているので、自分の部屋と大学の行き来しかし

ておらず、あまり街の大きさや不便さが気になりません（正直、街中をほとんど探索していないので、イサカのことはよくわかっていません...）。確かに他の有名な街と比べるとエンターテインメントやレストラン、レジャーなどは少ないと思いますが、自分の場合はそもそもあまり外に出るタイプではないので、蓋を開けたら関係なかったというところです。日常の食事や生活用品などは、Instacart[*4]や Amazon で完全に事足りているので、その辺も特に不便なく生活できています。[*5] ちなみに、Instacart はスーパーで売っている食材・惣菜・冷食などを、部屋の前まで配達してくれるサービスです。スーパーに出向いて買い物するよりも配達料等で高くつきますが、その分買い物にかかる時間を圧倒的に節約でき、現在はほぼ Instacart 経由でしか食品を調達していません。ちなみに、アメリカは食事のクオリティが低く大変だという話をよく聞きますが、この点についても今のところ何も気になっていません。今思えば東京に住んでいた時も、最寄りのコンビニエンスストアで毎回同じようなものを買い、ほぼ全ての食事を済ませていました。こちらに来てからも最初の数週間で食べれそうなものを特定し、以後、ほぼ同じ注文を数日おきに Instacart に投げ続けています。このように東京にいたときと変わらない満足度で毎日食事が取れてしまっているなど、生活の多くの面に関して無頓着な点は留学に向いていたのかもしれません。

気候については予想通りといったところで、10 月下旬あたりから少し寒い日が増えてきました。最近は 0 度からちょっと低い気温をふらついており、雪もたまに降ります。しかし、街中・キャンパス内の除雪はしっかりしており、寒いといっても耐えられないほどではありません。街の大きさについても、気候についても、かなり北海道の地元に似ている印象で、来る前は生活環境や気候が少し不安要素でしたが、来てみると「こんなものか」と思えてしまっています。結局は研究・勉強が大事でそれ以外のことはあまり気に留めなくなるので、大学選択のときは研究室や大学のプログラムを優先して考えるのが良いのかもしれません。兎にも角にも、これまでより多くの面において恵まれた環境で研究させてもらえているなと日々実感しており、引き続き、研究に力を入れて頑張りたいと思っているところです。

4　さいごに

渡米する前は初めての海外生活で研究に集中できるか心配していましたが、来てみるとあまり日本と変わらない環境で、程よいプレッシャーを感じつつこれまで以上に研究に集中できています。また分野をリードする研究者にかなり気軽にアクセスでき、世界中から集まった優秀な学生たちから刺激を受けられる最高の環境に身を置けて、とても満足した研究生活を送れています。最後になりますが、留学に際して様々な面から手厚いサポートをしてくださっている船井科学振興財団の皆さまに深く御礼申し上げます。

---

*4 https://instacart.com/

*5 Amazon は注文してから配達まで 3,4 日かかり、強いて言えばこれは不便な点なのかもしれません。しかし、地元に住んでいたときも配達類は少し時間がかかっていたので、同じような感覚です。