

留学報告書

安永迪弘

2022年6月

2019年9月より Stanford大学にてコンピュータサイエンス(CS)の博士学生をしております安永迪弘と申します。今回は3年目冬～春学期の生活を振り返りたいと思います。

1. 研究

Googleなどの検索エンジン, Siriやカスタマーサービスなどの質問応答システムを始め, 多くの自然言語処理タスクは解答を出すために関連する背景知識を集めそれに基づいて正しく推論をする必要があります。そのためPhD課程の研究テーマとして, より高度に知識と推論を駆使する自然言語処理の手法を開発することを目指しています。この半年間, 主に次の2つの研究課題に取り組みました。

(1) 言語モデルが持つ知識を増やすには？

BERTやGPT-3などの大規模言語モデルは, Wikipediaなど大量の文書で事前学習することで様々な知識を学習¹でき, 自然言語処理システムの根幹を成しています。しかし, 既存の言語モデルは各文書を単体で学習し文書同士の関連性を考慮に入れないため, 各文書内で言及された知識を学習できても関連する複数文書にまたがる知識は効果的に学習できない²という課題があります。そのため, 関連する複数文書を同時にモデリングする手法を確立し言語モデルが学習できる知識を増やすという研究を行いました。具体的には, ウェブ文書のハイパーリンク・学術文書の引用リンクなど, 我々は日々文書リンクを使って有用な関連知識を学んでいることに注目し, リンク付けられた文書を同時に学習する言語モデル「LinkBERT」を実装しました。LinkBERTは実際に質問応答タスクの精度を大きく向上でき, 面白い研究になったと感じています。

論文: M. Yasunaga, J. Leskovec, P. Liang. "LinkBERT: Pretraining Language Models with Document Links". ACL 2022. (<https://arxiv.org/abs/2203.15827>)

¹ 例えば「ワシントンDC」の文書から「ワシントンDCはアメリカの首都である」といった知識を学習できます

² 例えば「ワシントンDCのタイダルベイスンでは日本の桜が見られる」という知識は「ワシントンDC」の文書単体では得られず, 関連する「タイダルベイスン」の文書と組み合わせて初めて得られます。

同時に、大規模言語モデルを実際に一から学習し実験するプロジェクトに取り組んだのは今回が初めてで、大変だったことも多く、良い勉強になりました。パラメーター数の多い大規模言語モデルの学習は莫大な計算資源と時間がかかり³、仮説検証のサイクルに時間がかかってしまいます。そのため、手当たり次第実験するのではなく小さいモデルで注意深く仮説検証をして見通しを掴んでから徐々にモデルを拡大していく、システムティックな研究をすることを訓練させられたプロジェクトになりました。また関連して、言語モデルの挙動がモデルサイズ拡大によってどう変化するのか見積もる「Scaling law」研究⁴がOpenAIを始め最近人気ですが、自分自身言語モデルを一から学習する経験をしたことで、Scaling law 研究の重要性や面白さを改めて実感するようになりました。

(2) 複数の知識源を組み合わせて推論するには？

知識と推論を駆使する自然言語処理システムを開発する上で更なる課題の1つは、背景知識となる情報源が様々な形式のデータ (テキストに加え、画像、知識グラフ、データベースなど) に散らばっており、それらを統合して推論する方法が必要であることです。前回発表した論文「QAGNN」 (<https://arxiv.org/abs/2104.06378>) ではテキストと知識グラフを統合する手法を提案しましたが、今回はそれを改善した手法「GreaseLM」、そしてテキストと知識グラフに加えて更に画像情報を統合する手法「VQA-GNN」をラボの仲間と共同研究しました。

X. Zhang, A. Bosselut, M. Yasunaga, H. Ren, P. Liang, C.D. Manning, J. Leskovec. "GreaseLM: Graph Reasoning Enhanced Language Models for Question Answering". ICLR 2022. (<https://arxiv.org/abs/2201.08860>)

Y. Wang, M. Yasunaga, H. Ren, S. Wada, J. Leskovec. "VQA-GNN: Reasoning with Multimodal Semantic Graph for Visual Question Answering" (<https://arxiv.org/abs/2205.11501>)

この共同研究のおかげでかなり自分の経験幅や視野を広げることができたと感じており、ラボ仲間に感謝しています。

³ 例えばBERT-largeサイズの言語モデルを十分学習するには最新のA100 GPUを16個並列して1ヶ月程かかります。

⁴ Scaling Laws for Neural Language Models. <https://arxiv.org/abs/2001.0836>

2. TA経験

冬学期にはStanfordで初めてTeaching Assistant (TA)の経験をしました。Chris Manning先生が教えている自然言語処理のクラス ([CS224n: Natural Language Processing with Deep Learning](#)) です。学部生から社会人修士生まで500人以上の学生が履修する巨大なクラスで、緊張感もありましたがやりがいのある経験になりました。TAとして毎週オフィス・アワーを行うのですが、大変優秀で鋭い質問をする学生さんが多かったです。そのため、自分も毎週念入りに準備することで自然言語処理各論の理解を深められ、学生との対話で新たな知見を得ることもできました。

またChris先生から授業準備のコツを伺うことができました。例えば、学生を退屈にさせない・そして授業後に何か印象に残って帰ってもらうことが大事であり、そのためには一回の授業で多くのコンセプトを伝えようとするのではなく本当に重要なコンセプトだけ強調すること、各コンセプトに対して面白い例を使い具体的に掘り下げることが大事だと仰っていました。

3. 最後に

PhD 3年目は自分の研究を積み上げていくこと・共同研究を通じて研究の幅とインパクトを広げることに奔走した1年でした。先日PhD課程のQualification Examを無事通過することができ、一つの良い区切りになりました。この先の約2年間をどう有意義に使うかじっくり考えようと思っています。船井財団にはいつもサポートしていただき、本当に感謝しております。

また、今年の夏はシアトルにある Meta AI (Facebook AI研究所) でインターンをする予定で楽しみにしています。Meta AIは "[Open Pre-trained Transformer](#)" や "[Retrieval-augmented Language Model](#)" プロジェクトなど言語モデルの研究開発を牽引している研究所の一つで、メンターや周りの研究者から沢山学び吸収しようと思っています。