

深層学習による日中機械翻訳エンジンをリリース

【ポイント】

- 世界初のカスタマイズ可能な日中のニューラル機械翻訳
- 従来型の機械翻訳に比べ翻訳精度が大幅に向上
- 書き言葉系の日中ニューラル機械翻訳エンジンのデモを、みらい翻訳ホームページにて提供開始

株式会社みらい翻訳(みらい翻訳、本社: 東京都渋谷区、代表取締役社長: 栄藤 稔)は、国立研究開発法人情報通信研究機構(NICT、本部: 東京都小金井市、理事長: 徳田 英幸)との共同研究のもと、深層学習技術を用いたニューラル機械翻訳エンジンの、日本語・中国語(簡体字)(以下、「中国語」もしくは「中」といいます。)を開発しました。中国語解析の開発には、都科摩[北京]通信技術研究中心有限公司の協力も得ました。

日・中双方のビジネスシーンを想定した書き言葉系機械翻訳エンジン開発を行い、従来の統計的機械翻訳に比べ翻訳精度が大幅に向上しました。当エンジンの訳質はみらい翻訳のホームページにてデモとして体験いただけます(<https://miraitranslate.com/trial/>)。

また、ニューラル機械翻訳エンジンを搭載した企業向け機械翻訳サービスである Mirai Translator™ にも本年夏頃を目処に、この度開発した日・中機械翻訳エンジンを追加し、サービスを拡張しての提供を開始します。

【背景】

日本経済における対中取引は、輸出入額および対中直接投資額も、製造業を中心に 2017 年においても堅調な推移をみせております。中国が日本経済にとって重要なパートナーである状況はこれからも継続されていき、ビジネスにおける中国との協業やコミュニケーションは今後も増加傾向が続くでしょう。一方で、深層学習による機械翻訳技術の進歩は目覚ましく、日英言語間のサービスはより日常的にビジネスシーンで利用されるようになってきています。みらい翻訳では、市場の要望に応えるためにも、この深層学習技術を活用した日本語・中国語間のニューラル機械翻訳エンジンの開発にいち早く取り組んで参りました。

【今回の成果】

みらい翻訳は、NICT との共同研究のもと、深層学習技術を導入することにより、日・中、双方のビジネス向け書き言葉系の精度を大幅に向上した機械翻訳エンジンを開発しました。特に、ビジネスコミュニケーションや製造系等のビジネス文章などに注力した精度向上を実現しております。ニューラル機械翻訳エンジンは、従来型の機械翻訳エンジン(統計的機械翻訳エンジン)と比較し、翻訳精度が大幅に向上しました。

【今後の展望】

今回開発した日中の書き言葉系機械翻訳エンジンは、みらい翻訳の企業向け機械翻訳サービスである Mirai Translator™ や、みらい翻訳のパートナーが提供するパートナーサービスに組み込まれ、実用性の高い機械翻訳ソリューションとして提供していく予定です。当サービスでは、お客様の固有名詞に対応するユーザ辞書を登録してご利用いただけるようになります。また、みらい翻訳では、日本語・中国語間の翻訳エンジンでも、過去に人手翻訳したお客様の対訳データをお預かりし、翻訳エンジンを機械学習し直すことにより、お客様固有表現を翻訳処理できるように、日中エンジンもカスタマイズできるサービスも提供して参ります。

¹ ニューラル機械翻訳エンジンとは、脳の神経回路を模したニューラルネットワークを用いた機械翻訳技術です。膨大な対訳データによりニューラルネットワークを機械学習することで、高精度な翻訳を実現します。

【今回の成果の詳細】

従来型の統計的機械翻訳エンジンの翻訳処理結果と、今回開発したニューラル機械翻訳エンジンを、ビジネスコミュニケーションおよび製造系等のビジネス文章を評価文として、人手による精度評価を行いました。

- 比較対象
1. 統計的機械翻訳エンジン(従来エンジン)
 2. ニューラル機械翻訳エンジン

従来エンジンに対するニューラル機械翻訳エンジンの粗読率²を確認したところ、中国語から日本語への翻訳処理結果の粗読率は 60% (対従来エンジンからの改善率は 80%)、日本語から中国語への翻訳処理結果の粗読率は 80% (対従来エンジンからの改善率は 30%) を超えました。

今回の結果では、特に日本語から中国語への翻訳精度が高く、日本企業内にある日本語文書を、中国の現地法人に情報共有する場合に役立つと想定できます。製造業の現場では、製造マニュアル等の日本語書類は充実しているが、中国等の現地法人では十分な資料が揃っていないと聞きます。そのような場合、機械翻訳で大量の日本語書類を処理し、中国語で共有することが可能になります。

<翻訳例文>

中国語→日本語

例文 1:

原文	据称, 由于就业人口的减少及老龄化进程快速发展的影响, 日本的农业在今后数年内将迎来大量离农时代。除了熟练农户的技能传承及后继人才的培育之外, 农田的大规模化及农作业的省力化等生产效率提高问题已成为了紧要课题。
人手翻訳	日本の農業は、就業人口の減少と高齢化の急速な進展により今後数年で大量離農時代を迎えると言われており、熟練農家のノウハウの継承・後継者育成に加え、農地の大規模化や農作業の省力化などの生産性向上が喫緊の課題となっています。
統計的機械翻訳	就業者の減少と高齢化が急速に進むの影響で、日本の農業は今後数年のうちに大量の農時代を迎えた熟練の農家の技能の伝承および後継者の育成以外にも、農地の大規模化や農作業の省力化などの生産効率を高める問題はすでに緊急課題となっている。
ニューラル機械翻訳	就業人口の減少や高齢化の進展により、日本の農業は今後数年で大量の離農時代を迎えるといわれている。熟練農家の技能伝承や後継人材の育成に加えて、農地の大規模化や農作業の省力化などの生産性向上が急務となっている。

例文 2:

原文	公司公布了2016财年(2016年4月1日-2017年3月31日)至2018财年(2018年4月1日-2019年3月31日)的新中期经营计划《结构改革第三阶段》。
人手翻訳	2017年3月期から2019年3月期までの新しい中期経営計画となる「構造改革ステージ3」を発表しました。
統計的機械翻訳	同社の2016年度(2016年4月1日-2017年3月31日)2018年度(2018年4月1日～2019年3月31日)の新中期経営計画「構造改革の第3段階」を発表した。
ニューラル機械翻訳	2016年度(2016年4月1日の-2017年3月31日)から2018年度(2018年4月1日の-2019年3月31日)までの新中期経営計画《構造改革第三段階》を発表した。

²機械翻訳の実用性を確認する評価軸。翻訳結果において、原文の情報かどの程度翻訳処理結果で読み取れるかを5段階で評価し「5」を満点とする。3から5点の文章の総数を文章全体数で割った割合を粗読率として明示し、機械翻訳結果が理解可能な割合を表す。

例文 3:

原文	焊接皮革保護手套及安全靴應分別使用符合JIS T 8113(焊接皮革保護手套)及JIS T 8101(安全靴)規定的產品。
人手翻譯	溶接用かわ製保護手袋及び安全靴は、それぞれJIS T 8113(溶接用かわ製保護手袋)及び JIS T 8101(安全靴)に規定されたものを使用する。
統計的機械翻訳	溶接レザーを保護する手袋及び安全靴それぞれJIS T8113(溶接皮革保護手袋)およびJIS T8101(安全靴)の規定に適合した製品を使用すべきである。
ニューラル機械翻訳	溶接皮革の保護手袋と安全靴はそれぞれJIS T8113(溶接革保護手袋)及びJIS T8101(安全ブーツ)規定に適合する製品を使用すべきである。

日本語→中国語

例文 1:

原文	所得税半減の優遇政策を受けられる小規模企業・零細企業の範囲を大幅に拡大します。
人手翻訳	大幅扩展享受减半征收所得税优惠政策的小微企业范围。
統計的機械翻訳	会大幅扩大所得税减半的优惠政策获得小规模企业和零散企业的范围。
ニューラル機械翻訳	将大幅度扩大可享受所得税减半优惠政策的小规模企业、小型企业的范围。

例文 2:

原文	それだけでなく、このプラットフォームは、ハイスループット衛星の容量を現在の20Gから300Gまでさらに増やすことができます。
人手翻訳	不仅如此, 这个平台还可以进一步将高通量卫星的容量由目前的20G提高到300G。
統計的機械翻訳	不仅如此, 该平台、高通量的卫星的容量目前为20到300G能增加。
ニューラル機械翻訳	不仅如此, 该平台还可将高吞吐量卫星的容量从目前的20G增加到300G。

例文 3:

原文	ますます多くの通信事業者が、5Gの商業利用のために準備を整えており、2020年までに5Gの全面的な商業利用を実現するという目標を達成する力を養っています。
人手翻訳	越来越多的运营商正在为5G商用做好准备, 构建2020年实现5G全面商用的目标的能力。
統計的機械翻訳	越来越多的通信运营商, 5G的商业利用的, 所以准备了这个数据到2020年实现5G的全面的商业利用达成目标的能力。
ニューラル機械翻訳	越来越多的通信运营商为5G商业利用做准备, 在2020年之前培养实现5G全面商业利用目标的能力。

【リンク】

お試し翻訳	https://miraitranslate.com/trial/
機械翻訳サービス Mirai Translator™	https://miraitranslate.com/service/miraitranslator/
みらい翻訳のパートナーサービス	https://miraitranslate.com/service/#partner

< サービスに関するお問い合わせ >

株式会社みらい翻訳
 セールス&マーケティング部
 古谷 利昭
 Tel: 03-6434-1340
 E-mail: info@miraitranslate.com
<https://miraitranslate.com>

< 広報・報道関係のお問い合わせ先 >

みらい翻訳PR事務局 (株式会社アンティル内)
 担当: 横山、田辺、南口
 Tel: 03-5572-6063
 E-mail: miraitranslate@vectorinc.co.jp