



**МАШИН ОРЧУУЛГЫН ЧАНАРЫН ҮНЭЛГЭЭНД ХИЙМЭЛ ОЮУНЫ АРГАЧЛАЛЫГ
АШИГЛАХ БОЛОМЖ БА ҮР НӨЛӨӨ**

(парафраз болон хураангуйлах хэрэглэгдэхүүн ашигласан загварчлалын жишээн дээр)

О.Мөнхчимэг³

Abstract. This research examines the potential of Artificial Intelligence (AI) techniques and algorithms to improve and assess the quality of machine translation (MT) outputs, utilizing paraphrase and summarize model tools to enhance the results. In recent years, a multitude of studies have scrutinized the defaults and limitations of machine translation when compared to human translations, particularly within the realm of translation studies. These researches emphasize the efficiency of machine translation regarding time and cost, while also underscoring various quality-related issues. The significance of this research lies in its proposal to integrate paraphrasing and summarization techniques to elevate the quality and assessment of MT. AI models and algorithms, including GPT, Google Translate, Trados, DeepL, and Microsoft translation tools, have highlighted the necessity for a comprehensive evaluation of the quality discrepancies between machine-generated and human translations. Recent findings indicate that rather than outright rejecting or endorsing machine translation, it is essential to explore strategies for enhancing its accuracy. This article aims to advance translation quality and evaluation by incorporating paraphrasing and summarization tools. The integration of AI translation algorithms with MT systems has the potential to improve translation quality, optimize workflows, expedite results, and provide economic advantages in project execution. The results under this study indicate that the application of paraphrasing and summarization models can significantly enhance the quality of MT outputs, yielding more accurate and improved translations compared to original MT methods.

Түлхүүр үг: ХОУ алгоритм, нейрон сүлжээний орчуулга (NMT), GPT, хүний оролцоотой орчуулга (HITL- Human-in-the-loop), орчуулгыг оновчлох/translation optimization, ХОУ орчуулгын интеграци

Удиртгал

Орчуулга зүйн салбарт хиймэл оюуны оролцоо, нөлөөлөл сүүлийн жилүүдэд эрчимжиж, шинэ шийдэл, өөрчлөлт, шинэчлэлт, хувьслыг авчраад байгаа билээ. Энэ ч утгаараа хиймэл оюун ухааны (ХОУ)-ийн алгоритм, загварчлал, түүний орчуулгын чанарын үнэлгээнд үзүүлдэг нөлөөг эрдэмтэд олон талаас нь судлах болжээ. (ХОУ)-ийн алгоритм орчуулгын чанарт хэрхэн нөлөөлдөг, эсхүл хэрхэн нөлөөлж болох талаар (ХОУ) суурьтай загвар нээлттэй хэлбэрээр туршигдаж, ашиглагдах болсон энэ үед орчуулгын чанарын үнэлгээний асуудал чухлаар тавигддаг. (ХОУ) технологийг ашиглан хэлний ур чадвараа сайжруулах, эсхүл орчуулгын чанарыг оновчтой

³ ХУИС, Англи хэлний тэнхим, доктор /Ph.D/, <https://orcid.org/0009-0008-5973-1736>

боловсронгуй болгох, эсхүл өөрийн үндсэн үүргийг орлоход AI ашиглах салбарууд улам бүр нэмэгдэж байгаа технологийн зуунд хэл шинжлэл, тэр дундаа орчуулга зүйн салбар ухаан нь (ХОУ) суурьтай цахим өгөгдөл боловсруулах загвар аргачлалыг ашиглан, орчуулгын чанарын үнэлгээний нарийвчлалд нэлээд нөлөөлдгийг илрүүлсэн. (ХОУ) загварчлал, модельд суурилсан орчуулгын чанар нь илүү боловсронгуй болж, орчуулгын салбарт чухал байр суурь эзлэх болсон гэдэг практик нэлээдгүй үр дүнгүүдээс харагдаж болно. Орчуулгын чанарыг үнэлэх, дүгнэх нь үргэлж орчуулгын чухал бүрэлдэхүүн хэсэг болж ирсэн бол энэхүү үйлдлийг өнөө үед (ХОУ)-аар орлуулан, орчуулгын чанарыг үнэлэх, сайжруулалтыг тухай бүртээ гүйцэтгэж, үр дүнг туршиж байна. Орчуулгын чанарыг үнэлэх, дүгнэх эцсийн зорилго нь уншигч, сонсогч, орчуулгын бүтээлийг хэрэглэгчийн хэрэгцээ, шаардлагад нийцсэн, зөв, алдаагүй байхыг чухалчилдаг билээ (Cui et al., 2021). Үүний зэрэгцээ, орчуулгын явц нь орчуулсан агуулга нь хүссэн утгыг зөв дамжуулж, зорилтот хэлний соёлын утга агуулгыг бүрэн буулгаж байгаа эсэхээр тодорхойлогдоно (Doherty et al., 2013). Үүнд: орчуулгын чанарыг үнэлэх, дүгнэхдээ хүний буюу зорилтот хэрэглэгчийн үнэлгээ, технологийн үнэлгээ, эсвэл аль алиныг нь хослуулан интеграци хийх зарчмаар явуулах боломжтой гэдгийг энэхүү парафраз болон хураангуйлах хэрэглэгдэхүүн ашигласан загварчлалын жишээгээр танилцуулж байгаа юм (Kilian et al., 2012).

Орчуулгын салбар нь хиймэл оюун (ХОУ)-ны үүсэхтэй холбоотойгоор томоохон өөрчлөлтөд орсон. Орчуулгын чанарыг үнэлэх, дүгнэхэд (ХОУ) алгоритмын ашиглах нь улам бүр нэмэгдэж байгаа. Энэ нь мэдээж эх хэмжээний текст бичвэр, өгөгдлийг хурдан шуурхай, зөв, чанартай гүйцэтгэхтэй холбоотой асуудал.

Хиймэл оюун (ХОУ)-ны оролцоотой загвар, хэрэглэгдэхүүн нь бичвэр дэх буруу ташаа орчуулгын алдааг илрүүлэх, орчуулгын чанарыг илүү сайжруулах, мөн хиймэл оюун (ХОУ)-ын өөрийн алгоритмууд нь орчуулгын чанарыг үнэлэх, дүгнэхэд зарцуулж буй хугацаа болон зардал хүний гүйцэтгэлтэй харьцуулшгүй бага байгаа нь бид (ХОУ)-д тийм* гэж хэлэх хэрэгцээ, шаардлагыг бий болгоод байна (Gladkoff et al., 2021).

Энэхүү судалгаа нь орчуулгын чанарыг сайжруулах, үнэлэх, дүгнэхэд хиймэл оюун загваруудын нөлөөг судлахдаа парафраз болон хураангуйлах хэрэглэгдэхүүн ашигласан загварчлалын жишээгээр танилцуулж нотлох харуулах явдал юм. Судалгаандаа хиймэл оюун загвар ашигласан парафраз болон хураангуйлах түүлүүдийг ашиглан судалгааны жишээ бичиг баримт дахь алдааг, буруу ойлголтыг залруулах, машин орчуулгын үр дүнг (ХОУ)-нд суурилсан парафраз болон хураангуйлах хэрэглэгдэхүүнтэй интеграцлах зарчмыг тодорхойлоход оршино. Ингэснээр орчуулгын чанарыг сайжруулах, үнэлэх, дүгнэхэд машин орчуулгыг, хиймэл оюунаар бүтээгдсэн бусад загваруудтай хослуулах замаар илүү нарийвчилсан орчуулгыг бэлтгэн боломжийг эрэлхийлж байгаа юм Lee et al. (2021). Судалгааны үр дүн нь орчуулгын чанарыг үнэлэх, дүгнэхэд хиймэл оюун загвар, алгоритмын үр нөлөөг судлах болон орчуулгын чанарыг сайжруулахад парафраз болон хураангуйлах түүлүүдийг хэрхэн ашиглах боломжтой талаарх ойлголтыг өгөхөд оршино.

Судлагдсан байдлын тойм.

Сүүлийн 5-10 жилд орчуулгын салбар нь хиймэл оюун (ХОУ) технологийн стратегийн шинэчлэлтийн нөлөөгөөр чухал эрс өөрчлөгдөж байгаа. Хиймэл оюун нь орчуулгын стандарт, нарийвчлалыг сайжруулах, орчуулагч нарт ажлыг хөнгөвчлөхөд зориулагдан хэрэглэгдэж байгаа бөгөөд энэ нь орчуулгын олон шат дамжлага, явцыг автоматжуулах болж иржээ. Хиймэл оюун нь орчуулгын чанар, тэр дундаа орчуулгын үнэлгээнд шууд нөлөөлж байгаа. Машины орчуулгын шинэ эрин үед интеграцилах зарчмаар, тухайлбал, парафраз болон хураангуйлах хэрэглэгдэхүүнийг ашиглах замаар шинэ боломжийг орчуулгын салбарт олгож чадсан. Жишээ дурдахад хиймэл оюун

(ХОУ) технологийн тусламжтайгаар АНУ-ын Агаарын цэргийн хүчин өдөрт 100,000 орчим орос хэлний техникийн үг, хэллэгийг англи хэл рүү орчуулах боломжтой болсон. Иймд машины орчуулгын түүхийн хөгжил болон компьютерт суурилсан орчуулгын шинэ модель, загваруудын интеграцийн талаар судалгааны тойм, практикт гарч буй шинэ туршлагаас дор өгүүлэх болно.

Машин орчуулгын эхэн үе. Машины орчуулга хөгжих хоёр үндсэн шалтгаан байгаа. Нэгдүгээрт, мэдээ мэдээллийн урсгалыг хязгааргүй эрин үе, хоёрдугаарт, компьютерын технологийн хурдтай хөгжиж буй үе – эдгээр хүчин зүйлс нь хүний үйл ажиллагааг улам их механикжуулж, автоматжуулах шаардлага үүссэнтэй холбоотой. Тухайлбал: ОХУ-ын шинжлэх ухааны салбарын сэтгүүлчид өдөрт олон зуун өгүүлийн үр дүнг англи хэл рүү орчуулах хэрэгцээ шаардлагатай тулгардаг. Орчуулагчдын хомсдол болон хөдөлмөрийн зардал нэмэгдэх нь шинжлэх ухааны салбарт орчуулгын хэрэгцээг хязгаарлаж, ямар нэгэн оруулах шийдлийг эрэлхийлж байсан үеийн машин орчуулгын хөгжлийн эхэн үе гэлтэй. 1950-аад оны сүүлчээр машины орчуулгын загварыг боловсруулах, хөгжүүлэх боломжийг судалж эхэлсэн бөгөөд анхны туршилтаар орчуулгын хөтөлбөрийг Георгтаун их сургууль болон IBM компани хамтран хийж эхэлсэн түүхтэй. Энэхүү туршилт нь 1953 оны сүүлчээр дууссан бөгөөд тухайн туршилтын үр дүнд 100-300 үгийг орчуулж чадсан. Машины орчуулгын асуудал нь анхандаа хэлний асуудал байсан бол цаашдаа зөвхөн орос-англи хэлний толь бичиг болон хэлзүйн дүрмийн асуудлаар тогтоогүй гэдэг нь өнөөдрийн бидний хөгжил шууд харуулж буй. Машины орчуулгын хөгжүүлэхийн тулд хоёр үндсэн зарчмаар ажиллаж эхэлсэн ба үүний нэг нь шууд хэрэглээний орчуулга хийх хүсэлтэй эмпирик аргыг ашиглах, нөгөө нь хэлний анализ болон синтаксик шинжилгээг хийх илүү нарийн төвөгтэй арга замыг судлах шинэ хандлага руу чиглэсэн юм. IBM-ийн туршилт нь шууд орчуулгын системийг бүтээсэн бөгөөд орос хэлнээс англи хэл рүү 10,000 үг орчуулж чаддаг машин орчуулгын программ үүсгэж байжээ (Martins et al., 2017).

Цаашлаад 1959 онд Европын Шинжлэх Ухааны Өгөгдлийн боловсруулалтын төвийг байгуулсны дараа машины орчуулгыг судалгааны хүрээнд Euratom байгууллагаас эхэлсэн. 1961 онд Ispra дахь төвдөө том хэмжээний компьютер суурилуулж, энэ үеэс Георгтаун их сургуулийн орчуулгын алгоритмын системийг ашиглах болжээ. Тус систем нь Евратомд маш хурдан хугацаанд орчуулга хийхэд ашиглагдаж байсан. Евратом нь орчуулах үйл явцын хурд болон өртгийг хамгийн чухал зүйлс гэж үзэж, орчуулгын чанар болон үр дүнгийн тухай хөндлөнгийн үнэлгээг төдийлөн тооцоолоогүй юм (Beck et al., 2013). Гэсэн хэдий ч тухайн үедээ тус машин орчуулгын систем нь их хэмжээний орчуулгын хэрэгцээг хурдан хангах боломжийг олгож байсан гэж үздэг. Машины орчуулга нь одоогийн байдлаар эдийн засгийн болон хугацаанд хүчин зүйлсийг нэн тэргүүнд анхаарах болжээ.

Хоёрдох үеийн машины орчуулгын загварууд. (GPT-4, Google Translate, DeepL, Microsoft орчуулгын жишээгээр) Машин орчуулгын (МО) хөгжлийн хоёр дахь үе шатыг авч үзвэл бидний өнөөдөр хэрэглэгдэж буй МО алгоритмууд нь 10 гаруй жилийн өмнө бүтээгдсэн бөгөөд эдгээр системчилсэн загварууд нь компьютерын хөгжил, автоматжуулах хурд, санах ой, хандалтын төхөөрөмж зэргийг дахин программчилж шинэ систем боловсруулах хэрэгцээ үүссэн. Энэхүү шинэ системийн хөгжүүлэлт нь орчуулгын чанарыг шууд сайжруулахыг зорьж, мөн хэрэглэгчдийн дата өгөгдөлд суурилан тасралтгүй хөгжүүлэх боломжтой байхаар алгоритмтай бүтээгдэхээр зорьж байв (Lee et al., 2021). Хэл шинжлэлийн ухагдахуунд гол анхаарал хандуулж, өгүүлбэр зүйн бүтцийг сайжруулах, холбоо үгсийн хоорондын уялдаа холбоог тодорхойлох зэргээр төрөл бүрийн орчуулгын хувилбаруудыг туршиж, ялгаатай байдалд шинжилгээ хийж сайжруулалтыг хийж байв. Тухайлбал, тухайн үгийн утга, орчуулгын ялгааг тодорхойлох шаардлагатай нөхцөлийг тодорхойлж,

үгсийн утга, үгсийн дарааллыг хянаж, орчуулгын баримт бичгийн бүтэц болон өөрчлөлтийг боловсронгуй болгох тал дээр ажиллаж байв(Kong, 2022). Түүнчлэн, орчуулгын хурдыг нэмэгдүүлэн шинэчилсэн ба энэ нь орчуулгын хурдыг 300,000 үг/цаг хүртэл нэмэгдүүлж чадсан.

Шинэ үеийн машин орчуулгын систем нь өгүүлбэр зүй, утга зүйн шинжилгээг боловсронгуй болгож, орчуулгын хурд, чанарыг сайжруулахад чиглэж байгаа бөгөөд цаашид тасралтгүй хөгжүүлэлт хийх боломжтой шинэ арга технологийг ашиглах зорилготой юм.

Хоёрдох үеийн машины орчуулгын загварууд: Машины орчуулгын хоёрдох үеийн загварууд нь орчин үеийн хиймэл оюун ухааны хамгийн сүүлийн үеийн ололтод тулгуурлан бүтээгдсэн. Эдгээр системчлэгдсэн загварууд нь орчуулгын чанарыг эрс сайжруулсан бөгөөд олон хэлний орчуулгыг богино хугацаанд үр дүнтэй гүйцэтгэх боломжтой нөхцөлийг бүрэлдүүлж чадсан. GPT-4, Google Translate, DeepL, Microsoft орчуулга зэрэг орчуулгын системүүд нь цахим орчинд идэвхтэй ашиглагдаж, байнгын өгөгдлийн сайжруулалтыг AI тусламжтай гүйцэтгэж байна.

GPT-4 нь чанарын өндөр түвшинд бичвэрийн орчуулгыг үйлдвэрлэх, олон хэлний орчуулга хийх боломжтой томоохон хэлний модель юм. Энэ нь зөвхөн орчуулга хийхээс гадна, өгөгдөл, даалгавар боловсруулах сантай төдийгүй, гол онцлог нь бичвэрийн цааш утга санаа, контекстийг сайн ойлгож, нийгэм, соёлын холбогдолтой мэдлэгийг агуулсан орчуулга хийх үед орчуулгын чанар харьцангуй өндөр түвшинд хийгдэж байна. Google Translate нь олон төрлийн хэлний орчуулгыг хоромхон хугацаанд гүйцэтгэдэг монголд ашиглагдаж буй хамгийн түгээмэл машин орчуулгын систем юм. DeepL нь орчуулгын чанарт ихээхэн анхаарал хандуулан, төрөл бүрийн хэлзүйн бүтэц, зохион байгуулалтыг илүү нарийвчлалтайгаар орчуулах тал дээр давуу талтай. Харин Microsoft орчуулгын систем нь эдгээр загваруудаас хурд, хялбар гормоороо ялгардаг гэж үздэг. Эдгээр машин орчуулгын системүүдийн хөгжүүлэлт нь орчуулгын салбарт шинэ боломжийг нээж, хэрэглэгчдийн цаг, санхүүг хэмнэхээс гадна үнэн зөв орчуулга хийх боломжийг олгож, ажлыг хөнгөвчилж өгч байна.

Орчуулгын чанарын үнэлгээ. Тэгвэл машин орчуулгын чанарыг үнэлэхдээ орчуулсан материал нь эх бичвэрийн санаа, зорилго, утгыг зөв дамжуулж байгаа эсэхийн гол хэмжүүр болгож үздэг. Орчуулгын чанарыг үнэлэх нь орчуулгын салбарын хөгжилд чухал төдийгүй орчуулсан материал нь чанартай байж, зорилтот уншигчдын шаардлагыг хангана. Нэн ялангуяа, орчуулгын чанарыг үнэлэх нь хууль, эмнэлэг, техникийн болон уран зохиолын орчуулгын салбарт чухал ач холбогдолтой. Жишээ нь, хууль зүйн орчуулгад орчуулгын алдаа үүсэх нь хуулийн дүрмийг буруу ойлгуулах, эрх, үүрэг, хуулийн шийдвэр зэргээс үүдэлтэй үр дагаварт хүргэж болзошгүй. Эмнэлгийн орчуулгад буруу мэдээлэл нь буруу оношилгоо эсвэл эмчилгээний алдаа гаргах, бүр үхэлд хүргэж болно. Уран зохиолын орчуулгад алдаа гарвал тухайн бүтээлийн утга болон утга зохиолын үнэ цэнэ буурах эрсдэлтэй.

Энэ ч утгаараа тухайлбал, хэт нарийн агуулгатай, эсхүл хууль бус үйлдэл, хүчирхийлэл, бэлгийн мөлжлөг, хүний эрх, терроризм, хорлон сүйтгэх ажиллагаа зэрэг мэдээллийн төөрөгдөл үүсгэхүйц, бусдад хохирол учруулах агуулгыг орчуулахгүй байх автомат тохиргоог чат GPT алгоритмд оруулсан байдаг(Martins et al., 2017).

Хүний гүйцэтгэх орчуулгын сул тал: Орчуулгын чанарыг үнэлэхэд хүний гүйцэтгэсэн орчуулгын чанарын үнэлгээг хийхэд шинэ хэрэгсэл, аргачлалуудыг бий болгох хэрэгцээ, шаардлагатай. Машин орчуулга (МТ) нь хиймэл оюун ухаан (ХОУ)-ыг ашиглан нэг хэлнээс нөгөө хэл рүү агуулгыг илүүтэй үгчилсэн хэлбэрээр орчуулдаг. Иймд бид энд (ХОУ)-д суурилсан үнэлгээний аргачлалууд нь хүний бүтээсэн орчуулгын чанарыг сайжруулахад туслах боломжтой гэж үзээд парафраз болон хураангуйлах хэрэглэгдэхүүн ашигласан загварчлалын жишээгээр кэйс

судалгааны үр дүнг боловсруулсан юм. Хүний бүтээсэн орчуулгын үзэмжийн шинжтэй субъектив хандлага нь орчуулгын чанарыг үнэлэхэд хязгаарлалт болдог тул хүмүүсийн ойлголтын зөрөө, ялгаатай байдал нь орчуулгын нарийвчлал, оновчлол хийхэд харилцан ялгаатай байж болно. Мөн соёл, хэлний онцлог ялгаатай байдал нь орчуулгын чанар болон утга зөрүү үүсгэж болзошгүй.

Орчуулгын чанар, үнэлгээг сайжруулахад санал оруулж буй (ХОУ) загварууд: Орчуулгын чанар, бүтээмжийг сайжруулах зорилгоор олон төрлийн (ХОУ) загварууд ашиглагдаж туршигдаж байна. Эдгээр загварууд нь орчуулгын чанарыг үнэлэхдээ нарийвчлал, тохиромжтой байдал, үнэн зөв байдал гэх мэт олон төрлийн шалгуурыг голчлон ашигладаг.

- Машин сургалтын (ML) загварууд (NMT, SVM, RF, RNN)-ын алгоритмын тусламжтай орчуулгын чанарыг сайжруулах замаар, үнэлгээг хийх боломжтой. Тухайлбал: нейрон сүлжээний машин орчуулга (NMT): Энэ нь хүний мэдрэлийн тогтолцоонд суурилан загварчилсан хүний тархины моделийг машин сургалтын аргын тусламжтайгаар нэг хэлнээс нөгөө рүү орчуулга хийх машин орчуулгын загвар юм. NMT загвар нь нейрон сүлжээ ашиглан орчуулгын чанарыг сайжруулж, хурдан хугацаанд илүү нарийвчлалтай орчуулга хийх боломжийг олгодог. Нейрон сүлжээ нь бичвэрийн контекстийг ойлгох тал дээр хөгжүүлэлт хийгдсэн.
- Хэлзүйн дүрэмд суурилсан (ХОУ) загварууд/Syntax-Based (HOU) нь хэлзүйн дүрэмд суурилсан загварууд нь орчуулгын чанарыг үнэлэхэд орчуулгын бүтээлд хүний ойлголт, мэдлэгийг тусгах чухал хэрэгсэл болдог. Эдгээр загварууд нь орчуулгын чанарыг хэмжих болон засварлахад хэрэглэгддэг. Орчуулсан агуулга болон эх бичвэрийн хоорондын зөрүүг хэмжих хэмжүүр нь орчуулгын чанарыг шууд тогтоох, орчуулгын алдааны түвшнийг илрүүлэхэд ашиглагддаг бөгөөд орчуулгын чанарыг үнэлэх, мэргэжлийн орчуулагчдын түвшинд хүргэх үндсэн зорилготой. Үүнд: BLEU, TER зэрэг алгоритмууд багтана. Эдгээр алгоритмын тусламжтай хэлзүйн дүрэм, үгийн сангийн алдаа, өгүүлбэрийн дэд гишүүд, үгийн дарааллыг зөв тогтооход оршино. Орчуулгын чанарыг сайжруулах зорилгоор ашиглагддаг ХОУ загварууд нь орчуулгын нарийвчлал, оновчтой, үнэн зөв байдал зэрэг олон шалгуурыг тусгаж өгдөг. Машин сургалтын загварууд болон хэлзүйн дүрэмд суурилсан загварууд нь орчуулгын чанарыг үнэлэх, сайжруулахад чухал үүрэг гүйцэтгэдэг бөгөөд орчуулгын тэргүүлэх шийдэл гэж үздэг.

Судалгааны хэсэг

Иймд энэхүү судалгаа нь хиймэл оюун загварууд орчуулгын чанарын үнэлгээг боловсруулахдаа дээр шалгуур үзүүлэлтээр үнэлгээ тогтоож, хиймэл оюун загваруудад суурилсан парафраз болон хураангуйлах хэрэглэгдэхүүнтэй интеграцичлах замаар орчуулгын чанарт нөлөөлөх оролдлого хийсэн туршилт болно. Ингэснээр хиймэл оюун ухаанаар ажиллаж буй орчуулгын чанар, оновчлол, илэрцийг авч үзэж, чанарын үнэлгээг сайжруулахад чухал ач холбогдолтой гэж үзсэн.

Машины орчуулгын чанарыг сайжруулах, үнэлэх асуудал нь орчуулга хийх явцад чухал хэсэг төдийгүй эцсийн орчуулга үр дүн хэр сайн, үнэн зөв болохыг тодорхойлох шууд нөлөөлнө. Хиймэл оюун ухаан(AI)-д суурилсан парафраз болон хураангуйлах хэрэглэгдэхүүн нь орчуулгын чанарыг сайжруулахад чухал нөлөө үзүүлж байна. Энэ хэсэгт бид хиймэл оюуны аргачлалуудыг ашиглах боломж, үр нөлөөг судалж, парафраз болон хураангуйлах загварчлалын жишээн дээр үндэслэн орчуулгын чанарын сайжруулах талаар шинжилгээ боловсруулав.

МО-ын гүйцэтгэл, харьцуулалт: Машин орчуулгын хэрэглэгдэхүүн нь хоорондоо алгоритмын хувьд харилцан адилгүй тул үгсийн утга, хэв маягийг өөрчлөн сайжруулах боломжтой. Машины орчуулга хийхэд хэд хэдэн хүндрэлүүд тулгардаг. Эдгээр нь үгийн утга, соёлын ялгаа,

болон тусгай хэрэглээний орчуулга зэрэг асуудлуудыг багтаадаг. Машины орчуулга нь зарим үед хэлц утга бүхий соёлын ялгааг зөв орчуулах тал дээр учир дутагдалтай, хэлц хэллэгийг орчуулахад автомат орчуулга нь шууд утга, буруу ташаа утга илэрц гаргах магадлал харьцангуй өндөр гэдэг нь практикт ажиглагдсан.

Chat GPT болон DeepL зэрэг хиймэл оюуны ухаантай интеграци хийгдсэн орчуулгын моделиуд нь парафраз болон хураангуйлах аргуудыг орчуулгын чанарыг сайжруулахад хэрхэн ашигладаг талаар олон судалгаанд тусгагдсан. Эдгээр технологиуд нь орчуулгын хурдыг нэмэгдүүлэх, үгсийн зохисгүй байдлыг арилгах, хэрэглэгчдэд илүү олон төрлийн орчуулгын хувилбар санал болгодог.

GPT хувилбарын парафраз ба хураангуйлах аргачлал: GPT нь парафраз болон хураангуйлахыг маш өндөр түвшинд гүйцэтгэдэг буюу их хэмжээний бичвэр, мэдээллийг боловсруулж, хурдан хугацаанд ойлгомжтой, зөв орчуулгыг бэлтгэхэд түлхэц болно. GPT хураангуйлах аргачлал нь нийт эх бичвэрийн утгыг хадгалан богино цөөн өгүүлбэрээр товч орчуулгыг бэлтгэх боломжтой.

DeepL- парафраз интеграцичлал: Орчуулгыг илүү оновчтой, үндсэн санаа, утгыг хадгалж орчуулах боломжтой тул орчуулгын чанарыг илүү чанаржуулж өгч байна. Орчуулгын алгоритмд парафразын техникийн интеграци хийснээр орчуулгыг илүү уялдаа сайтай, хэрэглэгчдэд ойлгомжтой орчуулах боломжтой.

Парафраз болон хураангуйлах хэрэгслийг ашиглалт: Машины орчуулгын чанарыг үнэлэхэд парафраз болон хураангуйлах аргууд нь орчуулгын үгсийн утгыг хадгалж, бүтэц, агуулгаас хамааран хамгийн зөв хэлбэрийг олж чадна. Эдгээр аргууд нь хэлний бүтэц болон утгыг нарийвчлан ойлгох, хураангуйлах үед алдааг багасгах, мөн орчуулгын хурдыг нэмэгдүүлэхэд чухал үүрэг гүйцэтгэдэг.

Парафраз хэрэгсэл: Энэ нь машин орчуулгаар бэлтгэгдсэн өгүүлбэрийн буюу эх бичвэрийг агуулгыг алдагдуулахгүйгээр, олон төрлийн жанраар буюу өөр үгс, өгүүлбэрийн бүтэц, дараалал, хэлзүйн дүрэм ашиглан санааг ижил түвшинд илэрхийлэх аргачлал юм. Парафраз нь зөвхөн үгсийг солих бус, үгийн утга, түүний нийлэмжийг хадгалж, орчуулгыг илүү оновчтой, ойлгомжтой болгодог учир машины орчуулгын үр дүнгээс харьцангуй чанартай бүтээлийг боловсруулна. Машин орчуулгын систем нь парафразыг ашиглан, хослуулах зарчмаар хэрэглэгчдэд илүү тодорхой, орчуулгын хэлэнд ууссан үг хэллэгийг ашиглан орчуулах боломжтой.

Жишээ: A woman has a lot of work to do гэдэг өгүүлбэрийг парафразын аргаар орчуулбал дараах байдлаар илэрц гарах боломжтой, үүгээр үл хязгаарлагдана.

1. She has numerous tasks to complete. 2. She is faced with a significant amount of work. 3. She has an extensive workload ahead of her. 4. She is required to accomplish a great deal of work. 5. She has a substantial number of responsibilities to fulfill.

Хураангуйлах хэрэгсэл: уг аргачлал нь орчуулгын агуулгыг аль болох бага үгс ашиглан, хамгийн гол мэдээллийг хадгалаад цогц байдлаар утгыг нэг хэлээс нөгөө хэлэнд орчуулах боломжийг олгож байна. Машины орчуулгын хувьд хураангуйлах нь урт өгүүлбэр, утга санааг эсвэл олон цогцолбор дэд сэдвийн хувьд гүйцэд ойлгомжтой, товч байдлаар орчуулахад ашиглагддаг.

Жишээ: Climate change is increasingly recognized as a major threat to global ecosystems, leading to shifts in species distributions, altered community structures, and increased rates of extinction. Rising global temperatures, coupled with changes in precipitation patterns and increased frequency of extreme weather events, are creating conditions that are increasingly inhospitable for many species and habitats. For example, coral reefs, which are highly sensitive to temperature fluctuations, are experiencing widespread

bleaching and mortality due to ocean acidification and warming waters. Furthermore, shifts in vegetation zones and changes in the timing of seasonal events, such as migration and flowering, are disrupting ecological interactions and potentially leading to cascading effects throughout entire ecosystem – гэсэн цогцолборыг: Climate change is a significant threat to ecosystems, causing shifts in species locations, changes in community structures, and higher extinction rates. Rising temperatures and extreme weather make conditions harder for many species and habitats, like coral reefs. Additionally, changes in vegetation and seasonal events disrupt ecological interactions, creating further impacts on ecosystems гэж хураангуйлах оруулах боломжтой.

Парафраз, Хураангуйлах хэрэгслүүдийг интеграцчилсан зарчим: Парафраз хэрэгсэл: Машины орчуулгын бичвэрийн утга санааг алдагдуулахгүй, машины орчуулгын чанарыг улам илүү сайжруулж, утгын уялдаа холбоотой байх гаргалгааг санал болгож байна. Энэ нь орчуулгын үр дүн, чанарыг сайжруулж, хэлний бүтэц, утгыг илүү нарийн, тодорхой гаргаж өгч байна.

Сүүлийн үеийн судалгаанд парафраз болон хураангуйлах аргууд нь орчуулгын чанарыг сайжруулахад бодит үр дүнг үзүүлж байна. Машины орчуулгын системүүдийг нэгтгэн ашиглах нь орчуулгын чанарыг илүү сайн сайжруулах бөгөөд хүний оролцоог багасгах боломжтой гэж үзэж байна. Мөн орчуулгын асуудлуудыг шийдвэрлэхэд AI болон хүний оролцоог хосолсон загварууд ихээхэн үр дүнтэй байна.

Судалгааны үр дүн:

- **WMT 2020-ийн судалгаа:** Машины орчуулга хийхэд парафразын болон хураангуйлах аргууд нь орчуулгын нарийвчлалыг сайжруулсан бөгөөд орчуулгын чанарт эерэг нөлөө үзүүлсэн.
- **BLEU болон TER:** Хүний орчуулгын үнэлгээний үзүүлэлтүүдийн (BLEU, TER) хувьд парафраз болон хураангуйлах аргууд нь машины орчуулгын үнэлгээнд илүү өндөр оноогоор үнэлэгдсэн.

Хиймэл оюуны алгоритм, загвар: хүний оролцоотой хиймэл оюуны ухааны загварыг хэрэглэх нь хиймэл оюун болон мэргэжлийн орчуулагчийн оролцоог хослуулан орчуулгын чанарыг сайжруулах аргачлал юм. Энэ арга нь орчуулга хийх хурдыг сайжруулж, ажлын үр дүнг сайжруулахаас гадна, орчуулагчийн мэргэжлийн ур, чадварыг хөгжүүлэхэд түлхэц үзүүлнэ. AI загварууд болон хэрэгслүүд нь машин орчуулгын чанарыг дээшлүүлэх тал дээр шинэ үеийн хөгжилд ирээд байна, тухайлбал, DeepL, Google Translate, GPT зэрэг хэрэгслүүд нь хэлний бүтэц болон утга санааг илүү сайн нарийн ойлгож байгаа нь туршилтын үр дүнд харагдсан.

Парафраз болон хураангуйлах хэрэгслүүдийн интеграци, түүний үр нөлөө: Хиймэл оюун ухаан (AI) болон хүний оролцоотой (HITL) Парафраз болон хураангуйлах аргачлалыг ашиглан бэлтгэсэн орчуулга нь машины орчуулгын чанарыг эрс сайжруулж байна. GPT-4, DeepL, Google Translate зэрэг хэрэгслүүд нь орчуулгын утга, хэлний бүтэцэд илүү гүнзгий ойлголттой, илүү бодит үг хэллэгийг хэрэглэдэг бол, хүний оролцоотой бэлтгэсэн Парафраз болон хураангуйлах системүүд нь нарийвчилсан, мэргэжлийн орчуулгыг бэлтгэхэд чухал үүрэг гүйцэтгэнэ. Гэвч зарим сул талууд тулгарч болно.

Практик хэрэглээнд орчуулгын чанар, үр дүнг сайжруулах боломж: AI загваруудыг парафраз болон хураангуйлах аргуудтай хослуулан ашиглах нь орчуулгын зардлыг бууруулах, хурдыг нэмэгдүүлэхэд чухал үүрэг гүйцэтгэдэг. Эдгээр технологиуд нь орчуулгын төсөл, ажлыг илүү хурдан хугацаанд чанартай гүйцэтгэх магадлал, хурдыг нэмэгдүүлж, ажлын ачааллыг 80 хувь хүртэл бууруулдаг гэсэн судалгаа бий.

Судалгааны хязгаарлалт. Машины орчуулга (МО) нь сүүлийн жилүүдэд сайжран хөгжсөн ч орчуулгын чанарыг сайжруулах, үнэлэх нь асуудал шийдэгдээгүй хэвээр байна. Ерөнхийдөө хүний бүтээсэн орчуулгын хувьд хэмжээ, цар хүрээнээс хамаарч автомат орчуулгыг чухалчлах шаардлага үүсэн гарч байна. Харин HITL болон AI-ийн тусламжтайгаар Парафраз болон хураангуйлах аргачлалыг нь орчуулгын салбарт ашиглах нь орчуулгын чанарыг улам бүр сайжруулах хандлагатай байна.

Судалгааны гол дүгнэлт.

Энэхүү судалгаа нь парафраз болон хураангуйлах хэрэгсэл ашиглан орчуулгын чанарыг сайжруулах, ХОУ-ны алгоритмыг хэрхэн ашиглах талаар дэлгэрэнгүй гаргаж, шинэ боломжуудыг судалж байна. ХОУ-ийн тусламжтайгаар орчуулгын чанарыг эрс сайжруулж, хөдөлмөр, зардал хэмнэх боломжтой болохыг нотолжээ.

Машин орчуулгын чанарыг үнэлэхэд хиймэл оюун ухааны парафраз болон хураангуйлах аргуудыг ашиглах нь чухал ач холбогдолтой юм. Эдгээр аргачлалууд нь орчуулгын утга, бүтэц болон агуулгын нарийвчлалыг хадгалж, илүү тодорхой, ойлгомжтой орчуулга гаргах боломжийг бүрдүүлдэг. Парафраз нь зөвхөн үгсийн хувиргалт бус, утгын зөрүүг багасган, орчуулга илүү байгалийн хэллэгтэй болгох бол хураангуйлах нь агуулгыг хамгийн гол, товч мэдээллээр хүргэхэд чухал үүрэг гүйцэтгэдэг. GPT болон DeepL зэрэг хиймэл оюун ухаан ашигласан моделиуд эдгээр аргуудыг амжилттай ашиглаж, орчуулгын чанарыг сайжруулахад шууд нөлөөлдөг. Судалгаагаар, парафраз ба хураангуйлах технологиуд нь орчуулгын хурдыг нэмэгдүүлж, зардлыг бууруулдаг бөгөөд орчуулгын чанарыг хэвийн түвшинд хадгалж байна гэж үзлээ. Ийнхүү, хиймэл оюуны системтэй хослуулан ашиглах уг санал нь орчуулгын салбарт олон давуу талуудыг авчирч, орчуулгын нарийвчлал, оновчлол, үр дүнг сайжруулах, хэрэглэгчдэд илүү өндөр чанартай үйлчилгээ хүргэх боломжтой гэж дүгнэв.

Цаашдын чиг хандлага

Машины орчуулгын чанарыг үнэлэхэд хиймэл оюун ухааныг ашигласан парафраз ба хураангуйлах аргуудыг ашиглах нь үр дүнтэй бөгөөд орчуулгын утга санаа, хэл зүй, өгүүлбэрийн бүтцийн чанарыг сайжруулахад чухал үүрэг гүйцэтгэдэг. GPT, DeepL, Google зэрэг хиймэл оюун ухаанд суурилсан загварууд нь орчуулгын чанарт бодитой үр дүнг үзүүлж байгаа бөгөөд парафраз болон хураангуйлах аргачлалуудыг ашигласнаар илүү нарийвчилсан, оновчтой орчуулга хийх боломжтой гэж үзэв. Цаашид энэ талаарх бодит практик дээр нэвтрүүлэх нэлээдгүй судалгааны чиглэл бий.

Ашигласан бүтээлийн жагсаалт

Cui, Q., Huang, S., Li, J., Geng, X., Zheng, Z., Huang, G., & Chen, J. (2021). DirectQE: Direct pretraining for machine translation quality estimation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(14), 12719–12727.

Doherty, S., Gaspari, F., Groves, D., & Van Genabith, J. (2013). Mapping the industry I: Findings on translation technologies and quality assessment. [Online] Available from: <https://core.ac.uk/download/pdf/16754865.pdf> [Accessed 17/06/23].

Gladkoff, S., Sorokina, I., Han, L., & Alekseeva, A. (2021). Measuring uncertainty in translation quality evaluation (TQE). Available from: <https://arxiv.org/pdf/2111.07699.pdf>.

Kilian, N., Krause, M., Runge, N., & Smeddinck, J. (2012). Predicting crowd-based translation quality with language-independent feature vectors. In *Proceedings of the 4th Human Computation Workshop* (pp. 114–115). IEEE Computer Society.

- Lee, D., Ahn, J., Park, H., & Jo, J. (2021). IntelliCAT: Intelligent machine translation post-editing with quality estimation and translation suggestion. Available from: <https://arxiv.org/pdf/2105.12172.pdf>.
- Lommel, A. (2018). Metrics for translation quality assessment: A case for standardising error typologies. In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment. Machine translation: Technologies and applications* (Vol. 1, pp. 109–127). Cham: Springer.
- Li, X. (2022). An English translation quality evaluation model integrating knowledge transfer and wireless network. *Mobile Information Systems*, 2022, 2086486.
- Martins, A. F., Junczys-Dowmunt, M., Kepler, F. N., Astudillo, R., Hokamp, C., & Grundkiewicz, R. (2017). Pushing the limits of translation quality estimation. *Transactions of the Association for Computational Linguistics*, 5, 205–218.
- Beck, D., Shah, K., Cohn, T., & Specia, L. (2013). SHEF-Lite: When less is more for translation quality estimation. In *Proceedings of the Eighth Workshop on Statistical Machine Translation* (pp. 337–342). Association for Computational Linguistics. Available from: <https://statmt.org/wmt13/pdf/WMT41.pdf>.
- Chen, Z., Tan, Y., Zhang, C., Xiang, Q., Zhang, L., Li, M., & Wang, M. (2017). Improving machine translation quality estimation with neural network features. In *Proceedings of the Second Conference on Machine Translation* (pp. 551–555). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W17-4761>.